ESAIM: COCV 30 (2024) 60
https://doi.org/10.1051/cocv/2024048

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

FINITE ELEMENT ERROR ANALYSIS OF AFFINE OPTIMAL CONTROL PROBLEMS

Nicolai Jork*

Abstract. This paper is concerned with error estimates for the numerical approximation for affine optimal control problems subject to semilinear elliptic PDEs. To investigate the error estimates, we focus on local minimizers that satisfy certain local growth conditions. The local growth conditions we consider in this paper appeared recently in the context of solution stability and contain the joint growth of the first and second variation of the objective functional. These growth conditions are especially meaningful for affine control constrained optimal control problems because the first variation can satisfy a local growth, which is not the case for unconstrained problems. The main results of this paper are the achievement of error estimates for the numerical approximations generated by a finite element scheme with piecewise constant controls or a variational discretization scheme. Even though the growth conditions considered are weaker than those appearing in the recent literature on finite element error estimates for affine problems, this paper substantially improves the existing error estimates for both the optimal controls and the states when a Hölder-type growth is assumed.

Mathematics Subject Classification. 35J61, 49K20, 49M25.

Received October 30, 2024. Accepted June 3, 2024.

1. Introduction

Affine optimal control problems, by which we mean problems where the controls appear at most in an affine way in the objective functional and the constraining equation, are a relatively recent subject of study, especially when PDE constraints are considered. For the analysis of affine optimal control problems subject to ODE constraints, we refer to the papers [1–10] which contain results related to sufficient second-order conditions and the metric regularity and stability of the optimal control problems, especially for those with bang-bang structure of the optimal controls. An application of the regularity and stability investigations is for error estimates for the discretized problem, for instance, the Euler discretization, which can be found in [7]. For PDE-constrained problems, the earliest work related to affine problems is, to the best knowledge of the author, the work [11], which was extended to problems with different constraining PDEs or objective functionals, for instance, in [12–17]. Recently, the study of affine PDE-constrained optimal control problems was done in the works [18–20] under assumptions resembling partially the assumptions that appeared in the context of ODE-optimal control in [7]. A typical example of an affine problem is the tracking-type objective functional, which is a common type of objective functional used in many applied situations, including engineering, finance, and, more

Keywords and phrases: Semilinear elliptic equations, optimality conditions, finite element error estimation.

Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria.

* Corresponding author: nicolai.jork@tuwien.ac.at

recently, machine learning. Often, a so-called Tikhonov regularization term, a quadratic term with respect to the controls, is added to the objective functional. This is mainly for two reasons. First, in some situations, it is of interest to penalize the control cost. Second, incorporating the Tikhonov regularization term has some significant implications for analyzing the control problem. One of which is that by adding such a term, under mild additional assumptions, a quadratic growth of the second variation of the objective functional can be guaranteed, making the problem coercive. This then has many implications in the analysis of the problem, such as the study of error estimates for the numerical approximation. On the other hand, this comes with the price that adding such a term represents a distortion of the original problem; the optimal controls and states of the regularized problem can have substantially different structures. For instance, the bang-bang property of optimal controls can be expected in affine problems but not in regularized ones. To compensate for the missing coercivity due to the absence of a Tikhonov regularization term, the analysis of affine (unregularized) optimal control problems, in general, builds on certain assumptions related to the growth of the objective functional at local minimizers. In this paper, to study error estimates for the numerical approximation, we rely on assumptions that were recently studied in the context of strong metric subregularity and solution stability. These growth conditions encompass the local joint growth of the first and second variations and are weaker than the growth of the objective functional satisfied by Tikhonov regularized problems and the usual assumptions made in the study of affine problems. Using these assumptions, which we will specify below, we consider error estimates for the numerical approximation generated by a finite element scheme with piecewise constant controls and a variational discretization scheme. The analysis is motivated by the works [13, 21–23] on which the results of this paper build. Still, in comparison to the results therein, the error estimates for the optimal controls and states are under the assumptions introduced in [18, 19], which are weaker than the ones assumed in [13, 21–23]. For a detailed comparison in case of a parabolic constraining PDE, we refer to [20, 24], for the convenience of the reader, we provide a short discussion at the end of Section 3. Among others, we utilize the following assumption: given a reference optimal control $\bar{u}$ and a number $\gamma \in (0, 1]$, there exist positive constants α and c such that

$$J(u) - J(\bar{u}) \geq c\|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} \text{ for all feasible controls } u \text{ with } \|u - \bar{u}\|_{L^1(\Omega)} < \alpha. \tag{1.1}$$

Conditions of the type (1.1) arise naturally in characterizing strict bang-bang optimal controls. They also appear due to sufficient second-order optimality conditions and the structural assumption on the adjoint state, see [22]. A slightly stronger assumption that implies (1.1) was first considered in [7] for affine ODE optimal control problems and [19] for PDE optimal control problems. Recently, (1.1) appeared in [25, 26] in the context of eigenvalue optimization problems. There, it was shown that for a certain type of eigenvalue optimization problem, condition (1.1) is implied by a growth of the second-order shape derivatives.

In this paper, we do not explicitly consider a sparsity-promoting term in the objective functional as it is done, for instance, in [21]. Still, the proofs can be easily adapted to include such a term following the arguments in [21]. Also, we expect the approach presented in the paper to apply to the situation of having a semilinear elliptic non-monotone and non-coercive state equation as in [27] using the results of [28].

To the author's best knowledge, the assumptions considered in this paper are the weakest so far that still allow error estimates for the numerical approximation for problems where the control appears at most in an affine way in the objective functional, and we expect the approach discussed in this paper to be easily adaptable to optimal control problems constrained by various other PDEs. It may also be feasible to achieve error estimates for the numerical approximation for a 2-dimensional Neumann boundary control problem. But we postpone the analysis to future work.

Let us list the novelties in the paper. In Proposition 3.7, we answer a question raised in [18] on the structure of optimal controls satisfying one of the assumptions introduced in [18]. This result allows us to apply the assumption for the investigation of error estimates for the numerical approximation later on. Under conditions similar to the one introduced in [18] in the context of solution stability and conditions (1.1), we derive error estimates for a finite element discretization scheme with piecewise constant controls in Theorem 4.8. In the first part of the proof of the main theorem, Theorem 4.8, we argue similarly as in the first steps of the proof of [21], Theorem 7. In contrast to the proof in [21], Theorem 7, we employ arguments centered around the linearized state. This allows us to improve the error estimates for the optimal controls for $\gamma \in (0,1)$, from $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le ch^{\gamma^2}$ to $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le ch^{\gamma}$ (and similarly for the states). Under additional assumptions on the structure of the level set of the switching function, which is given by the adjoint state for particular tracking type problems, we can further improve the error estimate from $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le ch^{\gamma}$ to $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le ch^{\frac{2\gamma}{\gamma+1}}$ in Theorem 4.10. Using the assumptions of [18, 19], we prove error estimates for a variational discretization scheme in Theorem 4.11 and discuss a relationship with solution stability afterward.

The paper is structured as follows: In the remainder of this section, we state the main assumptions that hold throughout the paper and state some additional remarks on the notation. In Section 2, we collect results on the involved PDEs, and in Section 3, the optimal control problem is discussed. In Section 4, we define the discretization schemes and prove error estimates.

Let $\Omega \subset \mathbb{R}^n$, $n \in \{2,3\}$, be a bounded domain with $C^{1,1}$-boundary. Given constants $u_a, u_b \in \mathbb{R}$ such that $u_a < u_b$, define the set of feasible controls by

$$\mathcal{U} := \{u \in L^\infty(\Omega) |\ u_a \le u(x) \le u_b \ \text{ for a.a. } x \in \Omega\} \tag{1.2}$$

and consider the optimal control problem

$$\min_{u \in \mathcal{U}} \Big\{ J(u) := \int_\Omega L(x, y(x), u(x))\, \mathrm{d}x \Big\}, \tag{1.3}$$

subject to

$$\begin{cases} Ay + f(\cdot, y) = u & \text{in } \ \Omega, \\ y = 0 & \text{on } \Gamma. \end{cases} \tag{1.4}$$

Denote by y_u the unique solution of the state equation that corresponds to the control u. The objective integrand L appearing in (1.3) satisfies additional smoothness conditions, given below in Assumption 1.2.

1.1. Main assumptions and notation

The following assumptions, close to those in [18, 21, 22, 24], are standing in all of the paper.

Assumption 1.1. The following statements are fulfilled.

(i) The operator $A : H_0^1(\Omega) \to H^{-1}(\Omega)$, is given by

$$Ay := -\sum_{i,j=1}^n \partial_{x_j}(a_{i,j}(x)\partial_{x_i} y),$$

where $a_{i,j} \in C^{0,1}(\bar{\Omega})$. Further, the $a_{i,j}$ satisfy the uniform ellipticity condition

$$\exists \lambda_A > 0 : \ \lambda_A |\xi|^2 \le \sum_{i,j=1}^n a_{i,j}(x)\xi_i\xi_j \quad \text{for all } \xi \in \mathbb{R}^n \ \text{ and a.a. } x \in \Omega.$$

(ii) We assume that $f : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying:

$$\begin{cases} f(\cdot,0) \in L^\infty(\Omega) \text{ and } \frac{\partial f}{\partial y}(x,y) \geq 0 \ \forall y \in \mathbb{R}, \\ \forall M > 0 \ \exists C_{f,M} > 0 \text{ s. t. } \left|\dfrac{\partial f}{\partial y}(x,y)\right| + \left|\dfrac{\partial^2 f}{\partial y^2}(x,y)\right| \leq C_{f,M} \ \forall |y| \leq M, \\ \forall \rho > 0 \text{ and } \forall M > 0 \ \exists \ \varepsilon > 0 \text{ such that} \\ \left|\dfrac{\partial^2 f}{\partial y^2}(x,y_2) - \dfrac{\partial^2 f}{\partial y^2}(x,y_1)\right| < \rho \ \forall |y_1|, |y_2| \leq M \text{ with } |y_2 - y_1| \leq \varepsilon, \end{cases}$$

for almost every $x \in \Omega$.

Assumption 1.2. The function $L : \Omega \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ is Carathéodory and of class C^2 with respect to the second variable. In addition, we assume that

$$\begin{cases} L(x,y,u) = L_a(x,y) + L_b(x,y)u \ \text{ with } \ L_a(\cdot,0), L_b(\cdot,0) \in L^1(\Omega), \\ \forall M > 0 \ \exists C_{L,M} > 0 \text{ such that} \\ \left|\dfrac{\partial L}{\partial y}(x,y,u)\right| + \left|\dfrac{\partial^2 L}{\partial y^2}(x,y,u)\right| \leq C_{L,M} \ \forall |y|, |u| \leq M, \\ \forall \rho > 0 \text{ and } M > 0 \ \exists \varepsilon > 0 \text{ such that} \\ \left|\dfrac{\partial^2 L}{\partial y^2}(x,y_2,u) - \dfrac{\partial^2 L}{\partial y^2}(x,y_1,u)\right| < \rho \ |y_1|, |y_2| \leq M \text{ with } |y_2 - y_1| \leq \varepsilon, \end{cases}$$

for almost every $x \in \Omega$.

In the paper, we denote by c a positive constant that may change its value from line to line.

2. Auxiliary results for the state equation

We collect properties of solutions to linear and semilinear elliptic PDEs. The results in this section are standard by now; we refer to [18, 21]. In [18], the results are obtained for a non-monotone and non-coercive semilinear elliptic PDE. The PDE considered in this paper can be seen as a special case, and the results apply. Let $a_0 \in L^\infty(\Omega)$ be a nonnegative function. We consider the properties of solutions to the linear equation

$$Az + a_0 z = v \text{ in } \Omega, \ z = 0 \text{ on } \Gamma. \tag{2.1}$$

Theorem 2.1. *[18], Lemma 2.2 Let $v \in L^r(\Omega)$ with $r > n/2$. Then the linear equation (2.1) has a unique solution $z_v \in H_0^1(\Omega) \cap C(\bar{\Omega})$. Further there exists a positive constant C_r independent of a_0 and v such that*

$$\|z_v\|_{H_0^1(\Omega)} + \|z_v\|_{C(\bar{\Omega})} \leq C_r \|v\|_{L^r(\Omega)}. \tag{2.2}$$

Lemma 2.2. *[18], Lemma 2.3 Assume that $s \in [1, \frac{n}{n-2})$, s' is its conjugate, and let $a_0 \in L^\infty(\Omega)$ be a nonnegative function. Then, there exists a constant $C_{s'}$ independent of a_0 such that*

$$\|z_v\|_{L^s(\Omega)} \leq C_{s'} \|v\|_{L^1(\Omega)}, \quad \forall v \in H^{-1}(\Omega) \cap L^1(\Omega), \tag{2.3}$$

where z_v satisfies the equation (2.1), and $C_{s'}$ is given by (2.2) with $r = s'$.

For the semilinear state equation, we cite the following regularity result.

Theorem 2.3. *[21], Theorem 1 For every $u \in L^r(\Omega)$ with $r > n/2$ there exists a unique $y_u \in Y := H_0^1(\Omega) \cap C(\bar{\Omega})$ solution of (1.4). Moreover, there exists a constant $T_r > 0$ independent of u such that*

$$\|y_u\|_{H_0^1(\Omega)} + \|y_u\|_{C(\bar{\Omega})} \leq T_r(\|u\|_{L^r(\Omega)} + \|f(\cdot, 0)\|_{L^\infty(\Omega)}).$$

If $u_k \rightharpoonup u$ weakly in $L^r(\Omega)$, then we have the strong convergence

$$\|y_{u_k} - y_u\|_{H_0^1(\Omega)} + \|y_{u_k} - y_u\|_{C(\bar{\Omega})} \to 0.$$

Further if $u \in L^\infty(\Omega)$, we have $y_u \in W^{2,r}(\Omega)$ for all $r < \infty$ and

$$\|y_u\|_{W^{2,r}(\Omega)} \leq M_0 r\Big(\|u\|_{L^\infty(\Omega)} + \|f(\cdot, 0)\|_{L^\infty(\Omega)}\Big)$$

for a positive constant M_0 independent of u and r.

For each $r > n/2$, we define the map $G_r : L^r(\Omega) \to H_0^1(\Omega) \cap C(\bar{\Omega})$ by $G_r(u) = y_u$.

Theorem 2.4. *[18], Theorem 2.6 Let Assumption 1.1 hold. For every $r > \frac{n}{2}$ the map G_r is of class C^2, and the first and second derivatives at $u \in L^r(\Omega)$ in the directions $v, v_1, v_2 \in L^r(\Omega)$, denoted by $z_{u,v} = G_r'(u)v$ and $z_{u,v_1,v_2} = G_r''(u)(v_1, v_2)$, are the solutions of the equations*

$$Az + \frac{\partial f}{\partial y}(x, y_u)z = v \text{ in } \Omega, \; z = 0 \text{ on } \Gamma, \tag{2.4}$$

$$Az + \frac{\partial f}{\partial y}(x, y_u)z = -\frac{\partial^2 f}{\partial y^2}(x, y_u)z_{u,v_1}z_{u,v_2} \text{ in } \Omega, \; z = 0 \text{ on } \Gamma, \tag{2.5}$$

respectively.

Lemma 2.5. *[18], Lemma 2.7 The following statements are fulfilled. Suppose that $s \in [1, \frac{n}{n-2})$. Then, there exist a constant M_s depending on s such that for every $u, \bar{u} \in \mathcal{U}$*

$$\|y_u - y_{\bar{u}} - z_{\bar{u},u-\bar{u}}\|_{L^s(\Omega)} \leq M_s\|y_u - y_{\bar{u}}\|^2_{L^2(\Omega)}. \tag{2.6}$$

There exists $\varepsilon > 0$ such that for all $u, \bar{u} \in \mathcal{U}$ with $\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} < \varepsilon$ the following inequality is satisfied

$$1/2\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} \leq \|z_{\bar{u},u-\bar{u}}\|_{C(\bar{\Omega})} \leq 3/2\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})}. \tag{2.7}$$

$$1/2\|y_u - y_{\bar{u}}\|_{L^2(\Omega)} \leq \|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)} \leq 3/2\|y_u - y_{\bar{u}}\|_{L^2(\Omega)}. \tag{2.8}$$

We need the following lemma, which is standard, proofs (2.10)can be found in [29], Theorem 8.30 or [30], Theorem 4.2 respectively. The estimate (2.9) is standard and also appears in [21], Lemma 1.

Lemma 2.6. *Let $u, \bar{u} \in \mathcal{U}$. Then for $n < p$ and some positive constant $\check{C}_p$ independent of u and $\bar{u}$ it holds*

$$\|y_u - y_{\bar{u}}\|_{H_0^1(\Omega)} \leq 1/\lambda_A\|u - \bar{u}\|_{W^{-1,2}(\Omega)}. \tag{2.9}$$

$$\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} \leq \check{C}_p\|u - \bar{u}\|_{W^{-1,p}(\Omega)}. \tag{2.10}$$

3. The optimal control problem

The optimal control problem (1.2)–(1.3) is well posed under Assumptions 1.1 and 1.2. By the direct method of calculus of variations, we obtain the existence of at least one global minimizer, see [31], Theorem 5.7. In this section, we calculate the first and second variation of the objective functional, state the first-order necessary optimality conditions, and introduce the sufficient conditions for optimality.

Definition 3.1. We say that $\bar{u} \in \mathcal{U}$ is an $L^r(\Omega)$-weak local minimum of problem (1.2)–(1.4), if there exists some positive ε such that

$$J(\bar{u}) \leq J(u) \quad \forall u \in \mathcal{U} \text{ with } \|u - \bar{u}\|_{L^r(\Omega)} \leq \varepsilon.$$

We say that $\bar{u} \in \mathcal{U}$ is a strong local minimum of (1.2)–(1.4), if there exists $\varepsilon > 0$ such that

$$J(\bar{u}) \leq J(u) \quad \forall u \in \mathcal{U} \text{ with } \|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} \leq \varepsilon.$$

We say that $\bar{u} \in \mathcal{U}$ is a strict weak (strong) local minimum if the above inequalities are strict for $u \neq \bar{u}$.

Strong local minimizers were first considered in [32], Definition 1.6. For a discussion of these notions of optimality, we refer to [24], Lemma 2.8.

Theorem 3.2. *For every $r > \frac{n}{2}$, the functional $J : L^r(\Omega) \longrightarrow \mathbb{R}$ is of class C^2. Moreover, given $u, v, v_1, v_2 \in L^r(\Omega)$ we have*

$$J'(u)v = \int_\Omega \Big[\frac{\partial L}{\partial y}(x, y_u, u)\Big] z_{u,v} + \Big[\frac{\partial L}{\partial u}(x, y_u, u)\Big] v \,\mathrm{d}x = \int_\Omega \Big[\varphi_u + L_b(x, y_u)\Big] v \,\mathrm{d}x,$$

$$J''(u)(v_1, v_2) = \int_\Omega \Big[\frac{\partial^2 L}{\partial y^2}(x, y_u, u) - \varphi_u \frac{\partial^2 f}{\partial y^2}(x, y_u)\Big] z_{u,v_1} z_{u,v_2} \,\mathrm{d}x + \int_\Omega \Big[\frac{\partial L_b}{\partial y}(x, y_u)\Big](z_{u,v_1} v_2 + z_{u,v_2} v_1) \,\mathrm{d}x.$$

Here, $\varphi_u \in H_0^1(\Omega) \cap C(\bar{\Omega})$ is the unique solution of the adjoint equation

$$\begin{cases} A\varphi + \dfrac{\partial f}{\partial y}(x, y_u)\varphi = \dfrac{\partial L}{\partial y}(x, y_u, u) \text{ in } \Omega, \\ \varphi = 0 \text{ on } \partial\Omega. \end{cases} \tag{3.1}$$

Due to the standing assumptions of the paper, we can even infer that $\varphi_u \in W^{2,p}$, $p < \infty$. To obtain this regularity for a variational solution to (3.1) the boundary must have regularity $C^{1,1}$, see [33], Section 2. We define the Hamiltonian $\Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ni (x, y, \varphi, u) \mapsto H(x, y, \varphi, u) \in \mathbb{R}$ by

$$H(x, y, \varphi, u) := L(x, y, u) + \varphi(u - f(x, y)). \tag{3.2}$$

The following local form of the Pontryagin type necessary optimality conditions for problem (1.2)–(1.3) stated below, is well known (see *e.g.* [24, 31, 34] and [21], Thm. 4).

Theorem 3.3. *If $\bar{u}$ is a weak or strong local minimizer for problem (1.3)-(1.2), then there exist unique elements $\bar{y}, \bar{\varphi} \in H_0^1(\Omega) \cap L^\infty(\Omega)$ such that*

$$\begin{cases} A\bar{y} + f(x, \bar{y}) = \bar{u} \text{ in } \Omega, \\ \bar{y} = 0 \text{ on } \partial\Omega. \end{cases} \tag{3.3}$$

$$\begin{cases} A\bar\varphi = \dfrac{\partial H}{\partial y}(x,\bar y,\bar\varphi,\bar u) \ \textit{in}\ \Omega, \\ \bar\varphi = 0 \ \textit{on}\ \partial\Omega. \end{cases} \tag{3.4}$$

$$\int_\Omega \frac{\partial H}{\partial u}(x,\bar y,\bar\varphi,\bar u)(u-\bar u)\,\mathrm{d}x \ge 0 \quad \forall u \in \mathcal{U}. \tag{3.5}$$

3.1. Sufficient assumption for local optimality

In this subsection, we discuss three assumptions of different strengths that all imply strict local optimality and appeared recently in the context of affine PDE-constrained optimal control problems in [18, 19]. In what follows, $(\bar u, \bar y, \bar\varphi)$ denotes a fixed triplet satisfying the first-order necessary optimality condition. To shorten the notation, we denote $\bar H_u(x) := \frac{\partial H}{\partial u}(x, \bar y(x), \bar\varphi(x), \bar u(x))$.

Assumption 3.4. Let $\gamma \in (2/(2+n), 1]$ and $\beta \in \{1/2, 1\}$ be given. There exist positive constants κ and α such that

$$J'(\bar u)(u-\bar u) + \beta J''(\bar u)(u-\bar u)^2 \ge \kappa \|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} \tag{3.6}$$

for all $u \in \mathcal{U}$ with $\|u-\bar u\|_{L^1(\Omega)} < \alpha$.

Assumption 3.4($\beta = 1$) was first considered in the context of elliptic PDE-constrained optimization in [19]. If $\bar u$ satisfies the first-order optimality condition (3.5), Assumption 3.4($\beta = 1$) implies Assumption 3.4($\beta = 1/2$). If the second variation of the objective funcional at the control $\bar u$ is nonnegative, the cases $\beta \in \{1/2, 1\}$ are equivalent. Indeed, the second variation can be negative at $\bar u$ for certain directions for box-constrained optimal control problems; see, for instance, [20], Example 2. Further Assumption 3.4 implies the bang-bang structure of the control $\bar u$, see [19], Proposition 4.1. At this point, let us also remark that if the control $\bar u$ is bang-bang, then the convergence $u_k \rightharpoonup^\star \bar u$ in $L^\infty(\Omega)$ implies the strong convergence of $\{u_k\}_{k=1}^\infty$ to $\bar u$ in $L^1(\Omega)$, see [19], Lemma 4.2.

Let us consider the two assumptions on the optimal control problem introduced in [18]. Due to (2.3), they present a weakening of Assumption 3.4.

Assumption 3.5. Let $\beta \in \{1/2, 1\}$ be given. There exist positive constants κ and α with

$$J'(\bar u)(u-\bar u) + \beta J''(\bar u)(u-\bar u)^2 \ge \kappa \|z_{\bar u, u-\bar u}\|_{L^2(\Omega)} \|u-\bar u\|_{L^1(\Omega)} \tag{3.7}$$

for all $u \in \mathcal{U}$ with $\|y_u - y_{\bar u}\|_{C(\bar\Omega)} < \alpha$.

Assumption 3.6. Let $\beta \in \{1/2, 1\}$ be given. There exist positive constants κ and α with

$$J'(\bar u)(u-\bar u) + \beta J''(\bar u)(u-\bar u)^2 \ge \kappa \|z_{\bar u, u-\bar u}\|^2_{L^2(\Omega)} \tag{3.8}$$

for all $u \in \mathcal{U}$ with $\|y_u - y_{\bar u}\|_{C(\bar\Omega)} < \alpha$.

Assumption 3.6, the weakest of the three assumptions and does not imply the bang-bang property of the optimal controls. Assumption 3.6 is especially interesting as it is the weakest assumption so far that allows for solution stability estimates of the optimal states, see [18]. The interest of Assumption 3.5 stems from the fact that it is the weakest assumption so far that still allows for solution stability for the optimal controls, which is also discussed in [18]. Further, in [18], it was conjectured that Assumption 3.5 may also be satisfied by optimal controls that are not bang-bang. If $\frac{\partial L_b}{\partial y} = 0$, we can answer this negatively in the following proposition.

Proposition 3.7. *Let Assumption 3.5 be satisfied and $\frac{\partial L_b}{\partial y} = 0$. Then, $\bar u$ is bang-bang.*

Proof. Assume that $\bar{u}$ is not bang-bang and let it satisfy Assumption 3.5. Since $\bar{u}$ is not bang-bang, there exists a set of positive measure $E \subset \Omega$, such that $\bar{H}_u = 0$ on E. Let v_E denote a control with $v_E = \bar{u}$ on $\Omega \setminus E$ and $\|v_E - \bar{u}\|_{L^1(\Omega)} < \alpha$. Then the first variation in direction $v_E - \bar{u}$ is zero and by (3.7), we find

$$\beta J''(\bar{u})(v_E - \bar{u}, v_E - \bar{u}) \geq \kappa \|z_{\bar{u}, v_E - \bar{u}}\|_{L^2(\Omega)} \|v_E - \bar{u}\|_{L^1(\Omega)}. \tag{3.9}$$

By the affine structure of the optimal control problem, Assumption 1.1(ii), Assumption 1.2, the fact that $\bar{y}, \bar{\varphi} \in C(\bar{\Omega})$, the calculations in Theorem 3.2 and the assumption that $\frac{\partial L_b}{\partial y} = 0$, we infer the existence of a positive constant c independent of the control v_E such that

$$\beta J''(\bar{u})(v_E - \bar{u}, v_E - \bar{u}) \leq \beta c \|z_{\bar{u}, v_E - \bar{u}}\|^2_{L^2(\Omega)}. \tag{3.10}$$

Thus, using (3.9) and (3.10), we conclude for all controls v_E with $v_E = \bar{u}$ on $\Omega \setminus E$

$$\|\bar{u} - v_E\|_{L^1(\Omega)} \leq \beta c/\kappa \|z_{\bar{u}, v_E - \bar{u}}\|_{L^2(\Omega)}. \tag{3.11}$$

Since $\bar{u}$ is not bang-bang, we can select an $\varepsilon > 0$ and a set of positive measure $\hat{E}$ such that $\bar{u}(x) \in [u_a(x) + \varepsilon, u_b(x) - \varepsilon]$ for a.e. $x \in \hat{E}$. Now consider a sequence $\{v_\varepsilon^k\}_{k=1}^\infty$ with $v_\varepsilon^k \in \{-\varepsilon, \varepsilon\}$ a.e. on Ω and $v_\varepsilon^k \rightharpoonup^* 0$ in $L^\infty(\Omega)$ for $k \to \infty$. Finally, define a sequence $\{\bar{v}_\varepsilon^k\}_{k=1}^\infty$ by $\bar{v}_\varepsilon^k := \bar{u}$ on $\Omega \setminus \hat{E}$ and $\bar{v}_\varepsilon^k := \bar{u} + v_\varepsilon^k$ on $\hat{E}$. It is clear that $\bar{v}_\varepsilon^k \rightharpoonup^* \bar{u}$ in $L^\infty(\Omega)$ and $\|\bar{v}_\varepsilon^k - \bar{u}\|_{L^1(\Omega)} = \varepsilon|\hat{E}|$ for all $k \in \mathbb{N}$. On the other hand, by Theorem 2.3, $\bar{v}_\varepsilon^k \rightharpoonup^* \bar{u}$ in $L^\infty(\Omega)$ implies $\|z_{\bar{u}, \bar{v}_\varepsilon^k - \bar{u}}\|_{L^2(\Omega)} \to 0$ as $k \to \infty$. This contradicts (3.11). □

Consequently, the notion of strong or weak local minimizer is equivalent under Assumption 3.4 and Assumption 3.5.

Lemma 3.8. *Let $\gamma \in (0,1)$ and $\beta \in \{1/2, 1\}$ be given. It is equivalent:*

1. *There exist positive constants κ and α such that*

$$J'(\bar{u})(u - \bar{u}) + \beta J''(\bar{u})(u - \bar{u})^2 \geq \kappa \|u - \bar{u}\|^{1 + \frac{1}{\gamma}}_{L^1(\Omega)} \tag{3.12}$$

   for all $u \in \mathcal{U}$ with $\|u - \bar{u}\|_{L^1(\Omega)} < \alpha$.
2. *There exist positive constants c and α such that (3.12) holds for all $u \in \mathcal{U}$ with $\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} < \alpha$.*

Further, if the objective integrand satisfies $\frac{\partial L_b}{\partial y} = 0$ it is equivalent

1. *There exist positive constants κ and α such that*

$$J'(\bar{u})(u - \bar{u}) + \beta J''(\bar{u})(u - \bar{u})^2 \geq \kappa \|u - \bar{u}\|_{L^1(\Omega)} \|z_{\bar{u}, u - \bar{u}}\|_{L^2(\Omega)} \tag{3.13}$$

   for all $u \in \mathcal{U}$ with $\|u - \bar{u}\|_{L^1(\Omega)} < \alpha$.
2. *There exist positive constants κ and α such that (3.13) holds for all $u \in \mathcal{U}$ with $\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} < \alpha$.*

Proof. The statement of Lemma 3.8 for Assumption 3.4 with $\gamma = 1$ was proven in [18], Proposition 5.2. The proof relies on the fact that (3.12) implies the bang-bang structure of $\bar{u}$. But if $\gamma \in (0,1)$, (3.12) still implies the bang-bang structure and the arguments in [18], Poposition 5.2 hold true for $\gamma \in (0,1)$. By Poposition 3.7, the Assumption 3.5 implies the control $\bar{u}$ to be bang-bang, thus the results can be obtained by the arguments as in [18], Proposition 5.2. □

The next lemmas are needed for the estimations later on. Their well-known statement was proven for objective functionals with varying generality for the case $\gamma = 1$, [19], Lemma 11. The proof for $\gamma \in (2/(2+n), 1)$ follows by the same arguments. For Lemma 3.10 below, see for instance [11], Lemma 2.7.

Lemma 3.9. *Given $\gamma \in (2/(2+n), 1]$ and $\bar{u}, u \in \mathcal{U}$. Define $u_\theta := \bar{u} + \theta(u - \bar{u})$ for some measurable function θ with $0 \leq \theta(x) \leq 1$. For all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\left| J''(\bar{u})(u-\bar{u})^2 - J''(u_\theta)(u-\bar{u})^2 \right| \leq \epsilon \|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}$$

for all $\|u - \bar{u}\|_{L^1(\Omega)} < \delta$.

Lemma 3.10. *Given $\bar{u}, u \in \mathcal{U}$. Let $\frac{\partial^2 L}{\partial u y} = 0$ and define $u_\theta := \bar{u} + \theta(u - \bar{u})$ for some measurable function θ with $0 \leq \theta(x) \leq 1$. For all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\left| J''(\bar{u})(u-\bar{u})^2 - J''(u_\theta)(u-\bar{u})^2 \right| \leq \epsilon \|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)}^2$$

for all $\|y_u - y_{\bar{u}}\|_{C(\bar{\Omega})} < \delta$.

As a consequence of the Lemmas 3.9 and 3.10, we obtain strict local optimality under Assumptions 3.4, 3.5 and 3.6.

Theorem 3.11. *Let $\bar{u} \in \mathcal{U}$ be given and $\beta \in \{1/2, 1\}$ in the assumptions reference below.*

1. *Let Assumption 3.4 hold. Then there exists positive constants κ and α such that*
$$J(u) - J(\bar{u}) \geq \kappa \|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}, \tag{3.14}$$
for all $u \in \mathcal{U}$ with $\|u - \bar{u}\|_{L^1(\Omega)} < \alpha$.
2. *Let Assumption 3.6 hold for $\bar{u} \in \mathcal{U}$. There exist positive constants κ and α such that*
$$J(u) - J(\bar{u}) \geq \kappa \|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)}^2 \tag{3.15}$$
for all $u \in \mathcal{U}$ with $\|y_{\bar{u}} - y_u\|_{C(\bar{\Omega})} < \alpha$.
3. *Let Assumption 3.5 hold for $\bar{u} \in \mathcal{U}$. Then there exist positive constants κ and α such that*
$$J(u) - J(\bar{u}) \geq \kappa \|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)} \|\bar{u} - u\|_{L^1(\Omega)} \tag{3.16}$$
for all $u \in \mathcal{U}$ with $\|y_{\bar{u}} - y_u\|_{C(\bar{\Omega})} < \alpha$.

Proof. The statements with the growths (3.15) and (3.16) were proved in [18]. The statement for (3.14) follows by the same arguments using Lemma 3.9. ☐

Remark 3.12. Assumption 3.4 with $\gamma \in (2/(n+2), 1]$, together with Lemma 3.9 is used to guarantee the existence of positive constants κ and α such that

$$J(u) - J(\bar{u}) \geq \kappa \|u - \bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}, \text{ for all } u \in \mathcal{U} \text{ with } \|u - \bar{u}\|_{L^1(\Omega)} < \alpha. \tag{3.17}$$

If $\frac{\partial^2 L}{\partial y \partial u} = \frac{\partial L_b}{\partial y} = 0$ holds for the objective integrand, as a consequence of Lemma 3.10, (3.17) can be obtained by considering Assumption 3.4($\beta = 1/2,\ \gamma \in (0, 1]$) together with Assumption 3.6($\beta = 1/2$).

To see this, let κ and α be positive constants for that Assumptions 3.4 and 3.6 are satisfied simultaneously. Then, if $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega)}$ is sufficiently small, applying Taylor's theorem, Assumption 1.2, and Lemma 3.10 yields

$$\begin{aligned} J(u) &= J(\bar{u}) + J'(\bar{u})(u-\bar{u}) + 1/2J''(\bar{u}_\theta)(u-\bar{u})^2 \geq J(\bar{u}) + 1/2J'(\bar{u})(u-\bar{u}) + 1/4J''(\bar{u})(u-\bar{u})^2 \\ &\quad + 1/2J'(\bar{u})(u-\bar{u}) + 1/4J''(\bar{u})(u-\bar{u})^2 - 1/2\Big|J''(\bar{u}_\theta)(u-\bar{u})^2 - J''(\bar{u})(u-\bar{u})^2\Big| \\ &\geq \kappa/2\|u-\bar{u}\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} + \kappa/4\|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)}^2. \end{aligned}$$

Thus, the constraint $\gamma \in (2/(2+n), 1]$ can be weakened to $\gamma \in (0,1]$ for the cost of making both, Assumptions 3.4 and 3.6 at the same time.

3.2. A short comparison with growth-related conditions in the literature

We provide a short discussion of the relationship of Assumptions 3.4, 3.5 and 3.6 and the by now classical assumptions used for the analysis of affine PDE-constrained optimal control problems in the literature. By classical assumptions, we understand the ones considered, for instance, in [11, 21, 22, 28]. For this, let us define cones appearing in affine PDE-constrained optimal control.

Definition 3.13. We consider the set

$$\Big\{v \in L^2(\Omega)\Big|\ v \geq 0 \text{ a.e. on } [\bar{u} = u_a] \text{ and } v \leq 0 \text{ a.e. on } [\bar{u} = u_b]\Big\}. \tag{3.18}$$

Given $\tau > 0$, we define the sets

$$\begin{aligned} D_{\bar{u}}^\tau &:= \Big\{v \in L^2(\Omega)\Big|\ v \text{ satisfies (3.18) and } v(x) = 0 \text{ if } \Big|\frac{\partial \bar{H}}{\partial u}(x)\Big| > \tau\Big\}, \\ G_{\bar{u}}^\tau &:= \Big\{v \in L^2(\Omega)\Big|\ v \text{ satisfies (3.18) and } J'(\bar{u})(v) \leq \tau\|z_{\bar{u},v}\|_{L^1(\Omega)}\Big\}, \\ C_{\bar{u}}^\tau &:= D_{\bar{u}}^\tau \cap G_{\bar{u}}^\tau. \end{aligned}$$

Here, $\bar{H}$ denotes the Hamiltonian (3.2) corresponding to the control $\bar{u}$, that is $\bar{H}(x) := H(x, \bar{y}(x), \bar{\varphi}(x), \bar{u}(x))$.

Usually, the following two assumptions are made for the analysis of affine problems. The first is the structural assumption on the switching function: There exist positive constants c and $\gamma \in (0,1]$ such that

$$|\{x \in \Omega|\ |\bar{H}_u| \leq \varepsilon\}| \leq c\varepsilon^\gamma. \tag{3.19}$$

It is well known that this assumption implies for a possible different constant c that

$$J'(\bar{u})(u-\bar{u}) \geq c\|u-\bar{u}\|_{L^1(\Omega)}^{1+1/\gamma} \quad \text{for all } u \in \mathcal{U}. \tag{3.20}$$

The second assumption is the so-called second-order sufficient condition

$$J''(\bar{u})(u-\bar{u})^2 \geq c\|z_{\bar{u},u-\bar{u}}\|_{L^2(\Omega)}^2 \quad \text{for all } u \in \mathcal{U} \text{ with } (u-\bar{u}) \in C_{\bar{u}}^\tau. \tag{3.21}$$

Utilizing (3.19) and (3.21), error estimates for the numerical approximation are provided in [21].

To compare these assumptions with the one used in this paper, we first notice that it is equivalent to consider Assumption 3.4 only for $u \in \mathcal{U}$ with $(u-\bar{u}) \in D_{\bar{u}}^\tau$, see [19], Proposition 6.2 for elliptic problems and for parabolic problems see [18], Corollary 14. Further, we have the following theorem that relates Assumptions 3.4 to (3.19), (3.20) and (3.21).

Theorem 3.14. *Let $\frac{\partial L_b}{\partial y}=0$ and let there exist positive constants c,k and α with $k<c$ such that*

$$J'(\bar u)(u-\bar u)\geq c\|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} \text{ for all } u\in\mathcal{U},$$

and

$$J''(\bar u)(u-\bar u)\geq -k\|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}} \text{ for all } (u-\bar u)\in C_{\bar u}^{\tau} \text{ with } \|u-\bar u\|_{L^1(\Omega)}<\alpha.$$

Then Assumption 3.4, $\beta\in\{1/2,1\}$, holds for some appropriate constants. Further, let $\frac{\partial L_b}{\partial y}=0$. Then condition (3.21) implies Assumption 3.6.

Proof. It is sufficient to prove the statement for the Assumption 3.4 on the cone $D_{\bar u}^{\tau}$. Thus, we only need to consider the case $(u-\bar u)\notin G_{\bar u}^{\tau}$. But by definition of $(u-\bar u)\notin G_{\bar u}^{\tau}$, $J'(\bar u)(u-\bar u)>\tau\|z_{\bar u,u-\bar u}\|_{L^1(\Omega)}$. Using Theorem 3.2, it is straight forward to estimate for some constant d independent of u

$$\left|J''(\bar u)(u-\bar u)^2\right|\leq d\|z_{\bar u,u-\bar u}\|_{L^\infty(\Omega)}\|z_{\bar u,u-\bar u}\|_{L^1(\Omega)}.$$

By the assumption of this theorem, it also holds

$$J'(\bar u)(u-\bar u)\geq c\|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}.$$

Thus combining the estimates we obtain for $\|u-\bar u\|_{L^1(\Omega)}$ sufficiently small

$$\begin{aligned}J'(\bar u)(u-\bar u)+J''(\bar u)(u-\bar u)^2&\geq\frac{1}{2}J'(\bar u)(u-\bar u)+(\frac{1}{2}\tau-d\|z_{\bar u,u-\bar u}\|_{L^\infty(\Omega)})\|z_{\bar u,u-\bar u}\|_{L^1(\Omega)}\\&\geq c/2\|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}+(\frac{1}{2}\tau-d\|z_{\bar u,u-\bar u}\|_{L^\infty(\Omega)})\|z_{\bar u,u-\bar u}\|_{L^1(\Omega)}\geq c/2\|u-\bar u\|_{L^1(\Omega)}^{1+\frac{1}{\gamma}}.\end{aligned}$$

The claim regarding Assumption 3.6 is straightforwardly obtained by similar arguments. □

4. Discrete model and error estimates

We come to the main part of this manuscript. The goal is to prove error estimates for the numerical approximation under Assumption 3.4 for $\gamma\in(2/(n+2),1]$ and Assumptions 3.5 and 3.6.

4.1. The finite element scheme

The finite element scheme we consider is close to the one in [21]; we also refer to [35] for an overview of the finite element method. In this section, we assume Ω to be convex, see [36], Section 5.2. Let $\{\tau_h\}_{h>0}$ be a quasi-uniform family of triangulations of $\bar\Omega$. That is, for each $T\in\tau_h$, $\rho(T)$ denotes the diameter of T, and $\sigma(T)$ denotes the diameter of the largest ball inscribed in T. The mesh size is defined by $h:=\max_{T\in\tau_h}\rho(T)$. We assume that there exist two positive constants $\tilde\sigma$ and $\tilde\rho$ such that

$$\frac{\rho(T)}{\sigma(T)}\leq\tilde\sigma \text{ and } \frac{h}{\rho(T)}\leq\tilde\rho, \tag{4.1}$$

for all $T\in\tau_h$ and all $h>0$. Denote $\bar\Omega_h=\cup_{T\in\tau_h}T$ and define $\Omega_h:=\text{int}\bar\Omega_h$, and assume that every boundary node of Ω_h is a point of Γ. Suppose that there exists a constant $C_\Gamma>0$ independent of h such that the distance

d_Γ satisfies $d_\Gamma(x) < C_\Gamma h^2$ for every $x \in \Gamma_h = \partial\Omega_h$. As a consequence, we infer the existence of a constant $C_\Omega > 0$ independent of h such that

$$|\Omega \setminus \Omega_h| \le C_\Omega h^2, \tag{4.2}$$

where $|\cdot|$ denotes the Lebesgue measure, see [36], (5.2.19). We define the finite-dimensional space

$$Y_h = \{z_h \in C(\bar{\Omega}) : z_{h|T} \in P_1(T)\ \forall T \in \tau_h \ \text{ and } z_h \equiv 0 \text{ on } \Omega \setminus \Omega_h\},$$

where $P_1(T)$ denotes the polynomials in T of degree at most 1.
For $u \in L^2(\Omega)$, the associated discrete state is the unique element $y_h(u) \in Y_h$ that solves

$$a(y_h, z_h) + \int_{\Omega_h} f(x, y_h) z_h \,\mathrm{d}x = \int_{\Omega_h} u z_h \,\mathrm{d}x \ \ \forall z_h \in Y_h, \tag{4.3}$$

where

$$a(y, z) = \sum_{i,j=1}^{n} \int_\Omega a_{ij} \partial_{x_i} y \partial_{x_j} z \,\mathrm{d}x \ \ \forall y, z \in H^1(\Omega).$$

The proof of the existence and uniqueness of a solution for (4.3) is standard; see, for instance, [37].

Lemma 4.1. *[21], Lemma 3. There exists a constant $c > 0$, depending on the data of the problem but independent of the discretization parameter h, s. t. for every $u \in \mathcal{U}$*

$$\|y_h(u) - y_u\|_{L^2(\Omega)} \le ch^2, \tag{4.4}$$

$$\|y_h(u) - y_u\|_{L^\infty(\Omega)} \le ch^2 |\log h|^2. \tag{4.5}$$

The set of feasible controls for the discrete problem is given by

$$U_h := \{u_h \in L^\infty(\Omega_h) : u_{h|T} \in P_0(T)\ \forall T \in \tau_h\}.$$

By Π_h we denote the linear projection onto U_h in the $L^2(\Omega_h)$ given by

$$(\Pi_h u)_{|T} = \frac{1}{|T|} \int_T u \,\mathrm{d}x, \quad \forall T \in \tau_h.$$

By $u_h \rightharpoonup u$ weak* in $L^\infty(\Omega)$ we mean, as in [21], that

$$\int_{\Omega_h} u_h v \,\mathrm{d}x \to \int_\Omega uv \,\mathrm{d}x \ \ \forall\ v \in L^1(\Omega).$$

Lemma 4.2. *[21], Lemma 4 Given $1 < p < \infty$ there exists a positive constant $\hat{C}_p$ that depends on p and Ω but is independent of h such that*

$$\|u - \Pi_h u\|_{W^{-1,p}(\Omega_h)} \le \hat{C}_p h \|u\|_{L^p(\Omega)} \ \ \forall\ u \in L^p(\Omega).$$

We define $J_h(u) := \int_{\Omega_h} L(x, y_h(u), u)\,\mathrm{d}x$ and $\mathcal{U}_h := U_h \cap \mathcal{U}$. Then the discrete problem is given by

$$\min_{u_h \in \mathcal{U}_h} J_h(u_h). \tag{4.6}$$

The set $\mathcal{U}_h$ is compact and nonempty, and the existence of a global solution of (4.6) follows from standard arguments. For $u \in L^2(\Omega)$, the discrete adjoint state $\varphi_h(u) \in Y_h$ is the unique solution of

$$a(z_h, \varphi_h) + \int_{\Omega_h} \frac{\partial f}{\partial y}(x, y_h(u))\varphi_h z_h \,\mathrm{d}x = \int_{\Omega_h} \frac{\partial L}{\partial y}(x, y_h(u), u) z_h \,\mathrm{d}x \quad \forall z_h \in Y_h. \tag{4.7}$$

Again the proof of the existence and uniqueness of a solution for (4.3) is standard, see [37]. One can calculate that $J_h'(u)(v) = \int_{\Omega_h} (L_b(x, y_h(u)) + \varphi_h(u))v\,\mathrm{d}x$. A local solution of (4.6) satisfies the variational inequality

$$J_h'(\bar{u}_h)(u_h - \bar{u}_h) \geq 0 \quad \forall u_h \in \mathcal{U}_h.$$

In the following, similar as in [21], we identify $u_h = \bar{u}$ on $\Omega \setminus \Omega_h$. The existence of a sequence of solutions to the discrete problem that converges to an optimal solution of (1.3) is provided in the next theorem.

Theorem 4.3. *[21], Theorem 6 Let $\bar{u}$ be a strict strong local minimizer of (1.3). Then, there exists a sequence $\{\bar{u}_h\}_h$ of local minimizes of (4.6) such that $\bar{u}_h \rightharpoonup \bar{u}$ weak* in $L^\infty(\Omega)$. Moreover, there exists $h_0 > 0$ such that*

$$J_h(\bar{u}_h) \leq J_h(u_h) \quad \textit{for all } u_h \in \mathcal{U}_h \textit{ with } \|y_h(u_h) - y_h(\bar{u}_h)\|_{L^\infty(\Omega_h)} \leq \rho, \textit{ for all } h \leq h_0. \tag{4.8}$$

Conversely, let $\{\bar{u}_h\}_h$ be a sequence of local minimizers of (4.6) satisfying (4.8) for some given $\rho > 0$ and such that $\bar{u}_h \rightharpoonup^* \bar{u}$ in $L^\infty(\Omega)$. Then $\bar{u}$ is a strong local solution of (1.3) satisfying

$$J(\bar{u}) \leq J(u) \textit{ for all } u \in \mathcal{U} \textit{ with } \|y_u - \bar{y}\|_{L^\infty(\Omega)} < \rho. \tag{4.9}$$

Remark 4.4. Let $\bar{u} \in L^r(\Omega)$ and $u_h \rightharpoonup^* \bar{u}$ in $L^\infty(\Omega)$. Then $\|y_h(\bar{u}_h) - y_{\bar{u}}\|_{L^\infty(\Omega)} \to 0$ as $h \to 0$. This follows trivially since the right-hand side of

$$\|y_h(\bar{u}_h) - y_{\bar{u}}\|_{L^\infty(\Omega)} \leq \|y_h(\bar{u}_h) - y_{\bar{u}_h}\|_{L^\infty(\Omega)} + \|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^\infty(\Omega)} \tag{4.10}$$

tends to zero for $h \to 0$ due to Theorem 2.3 and Lemma 4.1, see also the related statement in the proof of [21], Theorem 6.

For the estimations for the variational discretization, we need the following theorem. The proof of Theorem 4.5 is done along the proof of [21], Theorem 9 using the arguments from the proof of [21], Lemma 3.

Theorem 4.5. *Let $\bar{u}_h$ denote a solution to (4.6). We denote by $y_{\bar{u}_h}$ and $\varphi_{\bar{u}_h}$ the solution to the continuous state equation and to the corresponding adjoint equation with respect to $\bar{u}_h$. By $\varphi_h(\bar{u}_h)$ we denote the discrete adjoint equation corresponding to $\bar{u}_h$ and $\varphi^h_{\bar{u}_h}$ denotes the solution to the following equation*

$$\begin{cases} \mathcal{A}^*\varphi + \frac{\partial f}{\partial y}(\cdot, y_h(\bar{u}_h))\varphi & = \frac{\partial L}{\partial y}(\cdot, y_h(\bar{u}_h)) \quad \textit{in } \Omega, \\ \varphi & = 0 \quad \textit{on } \Gamma. \end{cases}$$

There exists a positive constant c, which depends on the data of the problem but is independent of the parameter h and a positive number h_0 such that for all $h < h_0$ it holds

$$\|\varphi_{\bar{u}_h} - \varphi^h_{\bar{u}_h}\|_{L^\infty(\Omega)} \leq ch^2, \tag{4.11}$$

$$\|\varphi_h(\bar u_h) - \varphi^h_{\bar u_h}\|_{L^\infty(\Omega)} \le ch^2|\log h|^2. \tag{4.12}$$

4.2. Discretization with piece-wise constant controls

The two main goals of this section are to prove that Assumptions 3.4, 3.5 and 3.6 allow finite element error estimates and to improve the error estimates in the literature for $\gamma \in (0,1)$. Before stating the main theorems, let us consider two preliminary lemmas. Let us recall that $\bar H_u(x) := H_u(x, \bar y(x), \bar\varphi(x), \bar u(x)) = \bar\varphi(x) + L_b(x, \bar y(x))$. The assumption that $\bar H_u$ is Lipschitz is not a significant constraint for tracking-type objective functionals where $\frac{\partial L_b}{\partial y} = 0$. This is because of the assumptions on the control problem in this section; the adjoint state, $\varphi_{\bar u}$, is already Lipschitz continuous. If $L_b \neq 0$, the significance of the constraints depends on the regularity of $L_b(\cdot, y)$, which comes down to the regularity of the state y. Let us also recall that $\bar\varphi \in W^{2,p}(\Omega)$ as a consequence of the regularity of the domain, the regularity of the coefficients of the elliptic operator, and the boundedness of the right-hand side of the adjoint equation due to the boundedness of the solution to the state equation. Furthermore, the solution to the state equation, $\bar y$ is due to Theorem 2.3 in $W^{2,p}(\Omega)$ and thus Lipschitz as well. Finally, due to Assumption 1.2, which gives us the needed regularity of L_b, we can infer the Lipschitz continuity of $\bar H_u$.

Lemma 4.6. *Consider a bang-bang control $\bar u \in \mathcal U$ satisfying the first order optimality condition. Then $J'(\bar u)(\Pi_h \bar u - \bar u) \le \mathrm{Lip}_{\bar H_u} h \|\bar u - \Pi_h \bar u\|_{L^1(\Omega_h)}$.*

Proof. For the reader's convenience, we present a proof which follows the arguments in [21], Lemma 7. Let T be a triangle such that $\bar H_u$ changes its sign in T. Since $\bar H_u$ is Lipschitz continuous, there exists a point $x_0 \in T$ with $\bar H_u(x_0) = 0$. For $x \in T$ we obtain

$$|\bar H_u(x)| = |\bar H_u(x) - \bar H_u(x_0)| \le Lip_{\bar H_u}|x - x_0| \le Lip_{\bar H_u} h. \tag{4.13}$$

Let us denote by S_h the union of elements T such that $\bar H_u$ changes the sign. Then on the set S_h, we have the estimate $\|\bar H_u\|_{L^\infty(S_h)} \le Lip_{\bar H_u} h$. If $\bar H_u$ does not change the sign on an element T, the bang-bang structure implies $\Pi_h \bar u = \bar u$ on $\Omega \setminus S_h$. As a consequence, we obtain the estimate

$$J'(\bar u)(\Pi_h \bar u - \bar u) = \int_{\Omega_h} \bar H_u (\Pi_h \bar u - \bar u) \;\mathrm{d}x \le Lip_{\bar H_u} h \|\Pi_h \bar u - \bar u\|_{L^1(S_h)}.$$

□

The next lemma estimates the L^1-distance of a bang-bang reference solution and its projection. It is needed to obtain error estimates for the numerical approximation later on.

Lemma 4.7. *Let $\bar u \in \mathcal U$ satisfy Assumption 3.4($\beta = 1/2$), here we allow $\gamma \in (0,1]$. There exists positive constants c (independent of h) and a h_0, such that for all $h < h_0$:*

$$\|\bar u - \Pi_h \bar u\|_{L^1(\Omega_h)} \le ch^\gamma. \tag{4.14}$$

Proof. Since $\bar u \in \mathcal U$ satisfies Assumption 3.4, there exist positive constants κ and α such that

$$J'(\bar u)(\Pi_h \bar u - \bar u) + 1/2 J''(\bar u)(\Pi_h \bar u - \bar u)^2 \ge \kappa \|\Pi_h \bar u - \bar u\|^{1+1/\gamma}_{L^1(\Omega_h)} \quad \forall h \text{ s.t. } \|\Pi_h \bar u - \bar u\|_{L^1(\Omega_h)} < \alpha.$$

By Lemma 4.6, we obtain

$$J'(\bar u)(\Pi_h \bar u - \bar u) = \int_\Omega (\varphi_{\bar u} + L_b(\cdot, y_{\bar u}))(\Pi_h \bar u - \bar u)\,\mathrm{d}x \le \mathrm{Lip}_{\bar H_u} h \|\Pi_h \bar u - \bar u\|_{L^1(\Omega_h)}.$$

We recall that according to Theorem 3.2, it holds

$$\begin{aligned} J''(\bar u)(\Pi_h\bar u-\bar u)^2 &= \int_\Omega \Big[\frac{\partial^2 L}{\partial y^2}(\cdot,y_{\bar u},\bar u)-\varphi_{\bar u}\frac{\partial^2 f}{\partial y^2}(\cdot,y_{\bar u})\Big] z^2_{\bar u,\Pi_h\bar u-\bar u}\,\mathrm{d}x \\ &\quad + 2\int_\Omega \Big[\frac{\partial L_b}{\partial y}(\cdot,y_{\bar u})\Big] z_{\bar u,\Pi_h\bar u-\bar u}(\Pi_h\bar u-\bar u)\,\mathrm{d}x = I_1+I_2. \end{aligned} \tag{4.15}$$

Given $n<p$, $n/2<r$, using Assumption 1.2, Lemma 2.2, Lemma 2.5, Lemma 2.6 and Lemma 4.2, the first term of the second variation is estimated by

$$\begin{aligned} |I_1| &\le \Big\|\frac{\partial^2 L}{\partial^2 y}(\cdot,y_{\bar u})-p_{\bar u}\frac{\partial^2 f}{\partial^2 y}(\cdot,y_{\bar u})\Big\|_{L^\infty(\Omega)}\|z_{\bar u,\Pi_h\bar u-\bar u}\|^2_{L^2(\Omega)} \\ &\le 2(C_{L,M}+C_{L,M}C_{f,M}|\Omega|^{1/r})|\Omega|^{1/2}\|y_{\Pi\bar u}-y_{\bar u}\|_{L^\infty(\Omega)}\|z_{\bar u,\Pi_h\bar u-\bar u}\|_{L^2(\Omega)} \\ &\le 2C_2\check C_p(C_{L,M}+C_{L,M}C_{f,M}|\Omega|^{1/r})|\Omega|^{1/2}\|\Pi_h\bar u-\bar u\|_{W^{-1,p}(\Omega_h)}\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)} \\ &\le 2C_2\check C_p\hat C_p(C_{L,M}+C_{L,M}C_{f,M}|\Omega|^{1/r})|\Omega|^{(2+r)/(2r)}\max\{|u_a|,\ |u_b|\}h\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)}. \end{aligned}$$

It is left to estimate the term in the second line. Again using Assumption 1.2, Lemma 2.5 and Lemma Lemma 4.2, we obtain

$$\begin{aligned} |I_2| &\le 2C_{L,M}\|z_{\bar u,\Pi_h\bar u-\bar u}\|_{L^\infty(\Omega)}\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)} \\ &\le 4C_{L,M}\|y_{\Pi\bar u}-y_{\bar u}\|_{L^\infty(\Omega)}\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)} \\ &\le 4C_{L,M}\hat C_p|\Omega|^{1/r}\max\{|u_a|,\ |u_b|\}h\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)}. \end{aligned} \tag{4.16}$$

Thus, we infer the existence of a positive constant c such that

$$\kappa\|\Pi_h\bar u-\bar u\|^{1+1/\gamma}_{L^1(\Omega_h)} \le ch\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)}.$$

Dividing both sides by $\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)}$, completes the proof. ☐

Now, we are ready to state the main theorems of this section.

Theorem 4.8. *Let $\bar u$ be a local solution of (1.2)–(1.4). Consider the constant α corresponding to the Assumptions 3.4, 3.5 or 3.6. Consider a sequence of discrete optimal controls $\bar u_h \in \mathcal{U}_h$ of (4.6) that satisfy $\|\bar u_h-\bar u\|_{L^1(\Omega_h)}<\alpha$.*

1. *Let $L_b=0$ in the objective functional and let $\bar u$ satisfy Assumption 3.6($\beta=1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|y_h(\bar u_h)-\bar y\|_{L^2(\Omega)}+\|\varphi_h(\bar u_h)-\bar\varphi\|_{L^2(\Omega)} \le c\sqrt{h}\quad \text{for all } h\le h_0. \tag{4.17}$$

2. *Let $\frac{\partial L_b}{\partial y}=0$ in the objective functional and let $\bar u$ satisfy Assumption 3.5($\beta=1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|y_h(\bar u_h)-\bar y\|_{L^2(\Omega)}+\|\varphi_h(\bar u_h)-\bar\varphi\|_{L^2(\Omega)} \le c\sqrt{h^2+h\|\Pi_h\bar u-\bar u\|_{L^1(\Omega_h)}}\quad \text{for all } h\le h_0. \tag{4.18}$$

3. *Let $\bar u$ satisfy Assumption 3.4($\beta=1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|\bar u_h-\bar u\|_{L^1(\Omega_h)}+\|y_h(\bar u_h)-\bar y\|_{L^2(\Omega)}+\|\varphi_h(\bar u_h)-\bar\varphi\|_{L^2(\Omega)} \le ch^\gamma\quad \text{for all } h\le h_0. \tag{4.19}$$

If additionally Assumption 3.6($\beta = 1/2$) holds then (4.19) holds for $\gamma \in (0, 1]$.

Proof. Let us consider a discrete control $\bar{u}_h$ that satisfies the theorem's assumptions. We first prove the existence of a positive constant c such that

$$J(\bar{u}_h) - J(\bar{u}) \le \big|J(\bar{u}_h) - J_h(\bar{u}_h)\big| + J_h(\bar{u}_h) - J_h(\Pi_h \bar{u}) + \big|J_h(\Pi_h \bar{u}) - J(\Pi_h \bar{u})\big| \\ + \big|J(\Pi_h \bar{u}) - J(\bar{u})\big| = |I_1| + I_2 + |I_3| + |I_4| \le ch^{1+\gamma}.$$

To estimate the first term, we use Assumption 1.2, the estimates in Lemma 4.1, (4.2) and the mean value theorem to obtain for intermediate functions y_θ and y_ϑ that

$$\begin{aligned} |I_1| &= \left| \int_{\Omega\setminus\Omega_h} L(x, y_{\bar{u}_h}, \bar{u}_h)\,\mathrm{d}x + \int_{\Omega_h} L(x, y(\bar{u}_h), \bar{u}_h) - L(x, y_{\bar{u}_h}, \bar{u}_h)\,\mathrm{d}x \right| \\ &\le \Big(\Big\| \frac{\partial L_a}{\partial y}(\cdot, y_\theta) \Big\|_{L^2(\Omega)} + \Big\| \frac{\partial L_b}{\partial y}(\cdot, y_\vartheta)\bar{u}_h \Big\|_{L^2(\Omega_h)} + C_\Omega \|L(\cdot, y_{\bar{u}_h}, \bar{u}_h)\|_{L^\infty(\Omega)} \Big) h^2 \\ &\le (C_{L,M} + C_{L,M} \max\{|u_a|,\ |u_b|\} + C_\Omega K_{u_a,u_b,T_r}) h^2, \end{aligned}$$

where the constant K_{u_a,u_b,T_r} with $\|L(\cdot, y_u, u)\|_{L^\infty(\Omega)} \le K_{u_a,u_b,T_r}$ for all $u \in \mathcal{U}$, as indicated by the subscripts, depends on the control model though u_a,u_b and T_r. We have $I_2 \le 0$ for the second term since $\bar{u}_h$ is a minimizer of (4.6). The term I_3 can be estimated similarly as the first term,

$$\begin{aligned} |I_3| &= \left| \int_\Omega L(x, y_{\Pi_h \bar{u}}, \Pi_h \bar{u})\,\mathrm{d}x - \int_{\Omega_h} L(x, y(\Pi_h \bar{u}), \Pi_h \bar{u})\,\mathrm{d}x \right| \\ &= \left| \int_{\Omega\setminus\Omega_h} L(x, y_{\Pi_h \bar{u}}, \Pi_h \bar{u})\,\mathrm{d}x + \int_{\Omega_h} L(x, y_{\Pi_h \bar{u}}, \Pi_h \bar{u}) - L(x, y(\Pi_h \bar{u}), \Pi_h \bar{u})\,\mathrm{d}x \right| \\ &\le \Big(\Big\| \frac{\partial L_a}{\partial y}(x, y_\theta) \Big\|_{L^2(\Omega)} + \Big\| \frac{\partial L_b}{\partial y}(x, y_\vartheta)\Pi_h \bar{u} \Big\|_{L^2(\Omega_h)} + C_\Omega \|L(x, y_\vartheta, \Pi_h \bar{u})\|_{L^\infty(\Omega)} \Big) h^2 \\ &\le (C_{L,M} + C_{L,M} \max\{|u_a|,\ |u_b|\} + C_\Omega K_{u_a,u_b,T_r}) h^2. \end{aligned}$$

We come to the crucial part of the proof, the estimate of the last term I_4. The estimation of the term I_4 determines the overall convergence rate since the other terms already satisfy the good rate $|I_1|, |I_2| \le ch^2$. To shorten the notation, let us denote by $L_{a,y}$ and $L_{b,y}$ the derivatives of L_a and L_b by y. By the mean value theorem, for some intermediate function y_θ we have

$$\begin{aligned} I_4 &= \int_\Omega L(x, y_{\Pi_h(\bar{u})}, \Pi_h(\bar{u})) - L(x, y_{\bar{u}}, \bar{u})\,\mathrm{d}x = \int_\Omega L_{a,y}(x, y_\theta)(y_{\Pi_h \bar{u}} - y_{\bar{u}})\,\mathrm{d}x \\ &+ \int_\Omega L_{b,y}(x, y_\theta)(y_{\Pi_h \bar{u}} - y_{\bar{u}})\Pi_h \bar{u}\,\mathrm{d}x + \int_\Omega L_b(x, y_{\bar{u}})(\Pi_h \bar{u} - \bar{u})\,\mathrm{d}x \\ &= \int_\Omega L_{a,y}(x, y_{\bar{u}})(y_{\Pi_h \bar{u}} - y_{\bar{u}})\,\mathrm{d}x + \int_\Omega L_{b,y}(x, y_{\bar{u}})(y_{\Pi_h \bar{u}} - y_{\bar{u}})\bar{u}\,\mathrm{d}x \\ &+ \int_\Omega L_b(x, y_{\bar{u}})(\Pi_h \bar{u} - \bar{u})\,\mathrm{d}x + \int_\Omega (L_{a,y}(x, y_\theta) - L_{a,y}(x, y_{\bar{u}}))(y_{\Pi_h \bar{u}} - y_{\bar{u}})\,\mathrm{d}x \\ &+ \int_\Omega L_{b,y}(x, y_{\bar{u}})(y_{\Pi_h \bar{u}} - y_{\bar{u}})(\Pi_h \bar{u} - \bar{u})\,\mathrm{d}x + \int_\Omega (L_{b,y}(x, y_\theta) - L_{b,y}(x, y_{\bar{u}}))(y_{\Pi_h \bar{u}} - y_{\bar{u}})\Pi_h \bar{u}\,\mathrm{d}x. \end{aligned}$$

Thus,

$$
\begin{aligned}
I_4 &= \Big[\int_\Omega (L_{a,y}(x,y_{\bar u}) + L_{b,y}(x,y_{\bar u})\bar u) z_{\bar u,\Pi_h \bar u - \bar u}\,\mathrm{d}x \Big] + \Big[\int_\Omega L_b(x,y_{\bar u})(\Pi_h \bar u - \bar u)\,\mathrm{d}x \Big] \\
&+ \Big[\int_\Omega (L_{a,y}(x,y_{\bar u}) + L_{b,y}(x,y_{\bar u})\Pi_h \bar u)(y_{\Pi_h \bar u} - y_{\bar u} - z_{\bar u,\Pi_h \bar u - \bar u})\,\mathrm{d}x \Big] \\
&+ \Big[\int_\Omega (L_{a,y}(x,y_\theta) - L_{a,y}(x,y_{\bar u}))(y_{\Pi_h \bar u} - y_{\bar u})\,\mathrm{d}x \Big] + \Big[\int_\Omega (L_{b,y}(x,y_\theta) - L_{b,y}(x,y_{\bar u}))(y_{\Pi_h \bar u} - y_{\bar u})\Pi_h \bar u\,\mathrm{d}x \Big] \\
&+ \Big[\int_\Omega L_{b,y}(x,y_{\bar u})(y_{\Pi_h \bar u} - y_{\bar u})(\Pi_h \bar u - \bar u)\,\mathrm{d}x \Big] = \sum_{i=1}^{6} K_i.
\end{aligned}
$$

Let us first consider the arguments for the estimation (4.19). We remind that since $\bar u$ satisfies Assumption 3.4, it is bang-bang. We estimate the terms K_1 and K_2 together, that is, integrating by parts, using Lemma 4.6 and Lemma 4.7 guarantee the existence of a positive constant c such that

$$
\begin{aligned}
|K_1 + K_2| &= \Big| \int_\Omega (p_{\bar u} + L_b(x,y_{\bar u}))(\Pi_h \bar u - \bar u)\,\mathrm{d}x \Big| \\
&\le \mathrm{Lip}_{\bar H_u} \|\bar H_u\|_{L^\infty(\Omega)} \|\Pi_h \bar u - \bar u\|_{L^1(\Omega_h)} \le c\mathrm{Lip}_{\bar H_u} h^{1+\gamma}.
\end{aligned}
$$

The term K_3 is estimated, using Assumption 1.2, Lemma 2.5, Lemma 2.6 and Lemma 4.2,

$$
\begin{aligned}
|K_3| &\le C_{L,M}(1 + \max\{|u_a|,|u_b|\}) \|y_{\Pi_h \bar u} - y_{\bar u} - z_{\bar u,\Pi_h \bar u - \bar u}\|^2_{L^1(\Omega)} \\
&\le C_{L,M} M_1 (1 + \max\{|u_a|,|u_b|\}) \|y_{\Pi_h \bar u} - y_{\bar u}\|^2_{L^2(\Omega)} \\
&\le C_{L,M} \check C_p^2 M_1 (1 + \max\{|u_a|,|u_b|\}) |\Omega| \|\Pi_h \bar u - \bar u\|^2_{W^{-1,p}(\Omega_h)} \\
&\le C_{L,M} \check C_p^2 \hat C_p^2 M_1 (1 + \max\{|u_a|,|u_b|\}) |\Omega|^{1+2/p} \max\{|u_a|,|u_b|\}^2 h^2.
\end{aligned}
$$

For the estimation of the term K_4 we use that due to Assumption 1.2, $L_{y,a}$ is locally Lipschitz continuous, with Lipschitz constant denoted by $\mathrm{Lip}_{L_{a,y};M}$. We obtain using again Lemma 2.5, Lemma 2.6 and Lemma 4.2 that

$$
\begin{aligned}
|K_4| &\le \mathrm{Lip}_{L_{y,a};M} |\Omega| \|y_{\Pi_h \bar u} - y_{\bar u}\|^2_{L^\infty(\Omega)} \\
&\le \mathrm{Lip}_{L_{a,y};M} \check C_p^2 \hat C_p^2 |\Omega|^{1+2/p} \max\{|u_a|,|u_b|\}^2 h^2.
\end{aligned}
$$

Denoting the local Lpischitz constant of $L_{b,y}$ by $\mathrm{Lip}_{L_{b,y};M}$, the term K_5 is estimated using the same arguments by

$$
|K_5| \le \mathrm{Lip}_{L_{b,y};M} \check C_p^2 \hat C_p^2 |\Omega|^{1+2/p} \max\{|u_a|,|u_b|\}^3 h^2.
$$

Finally, for the term I_6,using Assumption 1.2, Lemma 2.5, Lemma 2.6 and Lemma 4.2, we estimate

$$
\begin{aligned}
|K_6| &\le C_{L,M} \|y_{\Pi_h \bar u} - y_{\bar u}\|_{L^\infty(\Omega)} \|\Pi_h \bar u - \bar u\|_{L^1(\Omega_h)} \\
&\le C_{L,M} \check C_p \|\Pi_h \bar u - \bar u\|_{W^{-1,p}(\Omega_h)} \|\Pi_h \bar u - \bar u\|_{L^1(\Omega_h)} \\
&\le C_{L,M} \check C_p \hat C_p \max\{|u_a|,|u_b|\} |\Omega|^{1/p} c h^{1+\gamma}.
\end{aligned}
$$

To complete the proof of (4.19), we conclude from the estimates of the terms I_i, $i \in \{1,...,6\}$, Theorem 3.11 and Theorem 3.14, that there exist positive constants κ,k and α such that

$$\kappa\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)}^{1+\frac{1}{\gamma}} \leq J(\bar{u}_h) - J(\bar{u}) \leq kh^{1+\gamma} \text{ for all } \bar{u}_h \text{ with } \|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} < \alpha.$$

This is equivalent to: $(\kappa/k)^{\frac{\gamma}{\gamma+1}} h^\gamma \geq \|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)}$ for all $\bar{u}_h$ with $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} < \alpha$. To estimate the states and adjoint states, we argue as follows. For the states we see by Lemma 2.2 applied to $z_v := \bar{y} - y_{\bar{u}_h}$, $v = \bar{u} - \bar{u}_h$, and Lemma 4.1, (4.4), that

$$\begin{aligned}\|\bar{y} - y_h(\bar{u}_h)\|_{L^2(\Omega)} &\leq \|\bar{y} - y_{\bar{u}_h}\|_{L^2(\Omega)} + \|y_{\bar{u}_h} - y_h(\bar{u}_h)\|_{L^2(\Omega)} \\ &\leq \hat{C}_2\|\bar{u} - \bar{u}_h\|_{L^1(\Omega_h)} + \tilde{C}_2 h^2 \leq \hat{C}_2(\kappa/k)^{\gamma/(\gamma+1)}h^\gamma + \tilde{C}_2 h^2.\end{aligned}$$

For the estimate of the adjoint states, we use that as a consequence of Lemma 2.2 and $\|\bar{p} - p_{\bar{u}_h}\|_{L^2(\Omega)} \leq C_2\|\bar{y} - y_{\bar{u}_h}\|_{L^2(\Omega)}$ that

$$\begin{aligned}\|\bar{\varphi} - \varphi_h(\bar{u}_h)\|_{L^2(\Omega)} &\leq \|\bar{\varphi} - \varphi_{\bar{u}_h}\|_{L^2(\Omega)} + \|\varphi_{\bar{u}_h} - \varphi_h(\bar{u}_h)\|_{L^2(\Omega)} \\ &\leq \hat{C}_2\|\bar{u} - \bar{u}_h\|_{L^1(\Omega_h)} + \tilde{C}_2 h^2 \leq \hat{C}_2(\kappa/k)^{\gamma/(\gamma+1)}h^\gamma + \tilde{C}_2 h^2.\end{aligned}$$

and the proof of (4.19) is complete. Let us briefly comment on the procedure for the other claims (4.17) and (4.18). Let us first consider (4.17). Here, we do not assume that $\bar{u}$ is bang-bang. The terms I_1 and I_3 can be estimated by the same arguments as before thus we infer the existence of a positive constant c such that

$$|I_1|, |I_2| \leq ch^2.$$

The estimation of the term I_4 is substantially easier due to the assumption $L_b = 0$. We only need to consider the terms K_1, K_3 and K_4. For K_1 we estimate using Assumption 1.2, Lemma 2.5, Lemma 2.6 and Lemma 4.2 to obtain

$$\begin{aligned}|K_1| &\leq C_{L,M}\|z_{\bar{u},\Pi_h\bar{u}-\bar{u}}\|_{L^2(\Omega)} \leq 3/2C_{L,M}|\Omega|^{1/2}\|y_{\Pi_h\bar{u}} - \bar{y}\|_{L^\infty(\Omega)} \\ &\leq 2/3C_{L,M}\check{C}_p\hat{C}_p\Omega|^{1/2}\|\Pi_h\bar{u} - \bar{u}\|_{W^{-1,p}(\Omega)} \leq 2/3C_{L,M}\check{C}_p\hat{C}_p\Omega|^{1/2+1/p}\max\{|u_a|, |u_b|\}h.\end{aligned}$$

The terms K_3 and K_4 are estimated in the same way as before. Thus, in total, we proved the existence of a positive constant c such that

$$\kappa\|z_{\bar{u},\bar{u}_h-\bar{u}}\|_{L^2(\Omega)}^2 \leq ch \text{ for all } \bar{u}_h \text{ with } \|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^\infty(\Omega)} < \alpha.$$

This of course using Lemma 2.5 implies that $\|\bar{y} - y_h(\bar{u}_h)\|_{L^2(\Omega)} \leq ch^{1/2}$. Now, the adjoint states can be estimated as argued above. Finally, let us consider (4.18). The terms I_1 and I_3 are estimated as before. Since $\frac{\partial L_b}{\partial y} = 0$ we only have to estimate the terms K_i, $i \in \{1,...,4\}$ in I_4. Since $\bar{u}$ is bang-bang according to Proposition 3.7, we can employ Lemma 4.6 to infer

$$\begin{aligned}|K_1 + K_2| &= \Big|\int_\Omega (\varphi_{\bar{u}} + L_b(x, y_{\bar{u}}))(\Pi_h\bar{u} - \bar{u})\,\mathrm{d}x\Big| \\ &\leq c\mathrm{Lip}_{\bar{H}_u}\|\Pi_h\bar{u} - \bar{u}\|_{L^1(\Omega_h)}h.\end{aligned}$$

The terms K_3 and K_4 are estimated as before. Thus, all in all, we have a positive constant c such that

$$\frac{\kappa}{C_2}\|z_{\bar u,u-\bar u}\|^2_{L^2(\Omega)} \le \kappa\|z_{\bar u,u-\bar u}\|_{L^2(\Omega)}\|\bar u - u\|_{L^1(\Omega)} \le J(u) - J(\bar u) \le ch(h + \|\Pi_h\bar u - \bar u\|_{L^1(\Omega)}),$$

which, estimating the adjoint states as above, completes the proof. □

Under some mild additional assumption on the zero level set of $\bar H_u$ that exclude the appearance of singular arcs, we can significantly improve the result of Theorem 4.8 using the next lemma instead of Lemma 4.7. In what follows, we denote by $\mathcal{H}^m$, the m-dimensional Hausdorff measure, see [38], Chapter 2.

Lemma 4.9. *Let $\bar u \in \mathcal{U}$ be bang-bang. Let us denote by A the points where $\bar H_u = 0$. Assume that A consists of a finite union of C^1 curves $(n = 2)$ or C^1-hypersurfaces $(n = 3)$. Then there exist positive constants c (independent of h) and h_0 such that for all $h \le h_0$:*

$$\|\bar u - \Pi_h\bar u\|_{L^1(\Omega_h)} \le ch. \tag{4.20}$$

Proof. First we notice that $\mathcal{H}^{n-1}(A) < \infty$. Let $\bar u$ be bang-bang and let, as before, denote by $\mathcal{S}_h$ the collection of the elements where there exists $x \in \text{int}\, T$ with $\bar H_u(x) = 0$ and denote, as in Lemma 4.6, by S_h the union of those elements. Then

$$\begin{aligned}
&\|\bar u - \Pi_h\bar u\|_{L^1(\Omega_h)} = \|\bar u - \Pi_h\bar u\|_{L^1(S_h)}\\
&\le \sum_{T\in\mathcal{S}_h}\int_T\Big|\bar u(x) - \frac{1}{\mathcal{H}^n(T)}\int_T \bar u(y)\,\mathrm{d}y\Big|\mathrm{d}x = \sum_{T\in\mathcal{S}_h}\int_T\frac{1}{\mathcal{H}^n(T)}\Big|\int_T \bar u(x) - \bar u(y)\,\mathrm{d}y\Big|\mathrm{d}x\\
&\le \sum_{T\in\mathcal{S}_h}\int_T\frac{1}{\mathcal{H}^n(T)}\int_T |\bar u(x) - \bar u(y)|\,\mathrm{d}y\mathrm{d}x \le \|u_b - u_a\|_{L^\infty(\Omega)}\sum_{T\in\mathcal{S}_h}\mathcal{H}^n(T).
\end{aligned}$$

If the diameter h of the quasi-uniform triangulations is sufficiently small, the number of elements that cover the set A can be estimated by the quotient of the diameter h and $\mathcal{H}^{n-1}(A)$. That is, there exists a constant c (independent of h) and a h_0 such that for all $h \le h_0$

$$\sum_{T\in\mathcal{S}_h} 1 \le c\frac{\mathcal{H}^{n-1}(A)}{h^{n-1}}.$$

This can be seen by the following arguments for the 2-dimensional case. The set $A = \cup_{i=1}^m A_i$ consists of a finite union of C^1-curves. Let us consider a given curve A_i and a triangle $T \in \mathcal{S}_h$; we realize that A_i intersects triangles T such that for any given $x \in A_i$, $\max_{y\in T} d(x,y) \le h$. Now take for each $x \in A_i$ the unit normal $\nu(x)$ to A_i and define the set $E_i^x := \{x + \xi\nu(x) | \xi \in [-h,h]\}$ and $E_i := \cup_{x\in A_i} E_i^x$. Then $\mathcal{H}^2(E_i) = 2h \cdot \mathcal{H}^1(A_i)$.

On the other hand, due to the quasi uniformity of the triangulation, there exists a positive constant $\bar c$ such that the measure of the triangles T is uniformly bounded from below by $\bar c h^2$. Therefore there are at most

$$\frac{\mathcal{H}^2(E_i)}{\bar c h^2} = \frac{2\mathcal{H}^1(A_i)h}{\bar c h^2} = \frac{2\mathcal{H}^1(A_i)}{\bar c h},$$

triangles that intersect the curve A_i. We obtain the claim by applying this to all the arcs A_i. The three-dimensional case follows by straightforward adaptions of the argument. From here, using that for a positive constant $\tilde c$, the measure of the elements T can be uniformly estimated by $|T| \le \tilde c h^n$, we conclude:

$$\begin{aligned}
&\|\bar u - \Pi_h\bar u\|_{L^1(\Omega_h)} = \|\bar u - \Pi_h\bar u\|_{L^1(S_h)} \le \|u_b - u_a\|_{L^\infty(\Omega)}\sum_{T_h\in\mathcal{S}_h}\mathcal{H}^n(T)\\
&\le \|u_b - u_a\|_{L^\infty(\Omega)}\frac{c\mathcal{H}^{n-1}(A)}{h^{n-1}}\max_{T\in\mathcal{S}_h}\mathcal{H}^n(T) \le \|u_b - u_a\|_{L^\infty(\Omega)}\frac{\mathcal{H}^{n-1}(A)}{h^{n-1}}c\tilde c h^n.
\end{aligned}$$

□

Due to the assumption on the optimal control problem in this paper, the adjoint $\bar{\varphi}$ has regularity $W^{2,p}(\Omega)$, $p < \infty$. We remark that due to [33], Section 2, for this result, it is necessary to have a $C^{1,1}$ boundary. Then, due to the $W^{2,p}(\Omega)$ regularity of the adjoint, in dimension $n = 2$, the Morse-Sard theorem for Sobolev functions and the implicit function theorem implies that for almost all t in the image of $\bar{\varphi}$, the level set $[\bar{\varphi} = t]$ consists of finitely many disjoint C^1 simple curves [39–41]. This almost everywhere result can be improved if $\bar{\varphi}$ satisfies

$$\min_{x\in A} |\nabla\bar{\varphi}(x)| > 0, \quad A = \{x \in \Omega | \ \bar{\varphi}(x) = 0\}. \tag{4.21}$$

Indeed, if (4.21) is satisfied, $[\bar{\varphi} = 0]$ consists of a finite union of simple C^1 curves, see [42], Proposition 2.4, Corollary 2.12. This supports the assumption of Lemma 4.9. On the other hand, if $\bar{\varphi} \in C^1(\bar{\Omega})$ satisfies (4.21), it already holds $|\{x \in \Omega | \ |\bar{\varphi}| \leq \varepsilon\}| \leq c\varepsilon$, see [43], Lemma 3.2. We apply Lemma 4.9 when we only expect

$$|\{x \in \Omega| \ |\bar{\varphi}| \leq \varepsilon\}| \leq c\varepsilon^\gamma, \quad \gamma \in (0, 1),$$

which permits $\min_{x\in A} |\nabla\bar{\varphi}(x)| = 0$.

We obtain the following improvement of the estimation in Theorem 4.8.

Theorem 4.10. *Let $\bar{u}$ be a local solution of (1.3). Consider the constant α corresponding to Assumptions 3.4, 3.5 or 3.6. Consider discrete optimal controls $\bar{u}_h \in \mathcal{U}_h$ of (4.6) that satisfy $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} < \alpha$. Further, assume that $\bar{H}_u$ satisfies the assumption of Lemma 4.9.*

1. *Let $L_b = 0$ in the objective functional, let $\bar{u}$ be bang-bang and let $\bar{u}$ satisfy Assumption 3.6($\beta = 1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^2(\Omega)} \leq ch \quad \textit{for all } h \leq h_0. \tag{4.22}$$

2. *Let $\frac{\partial L_b}{\partial y} = 0$ in the objective functional and let $\bar{u}$ satisfy Assumption 3.5($\beta = 1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^2(\Omega)} \leq ch \quad \textit{for all } h \leq h_0. \tag{4.23}$$

3. *Let $\bar{u}$ satisfy Assumption 3.4($\beta = 1/2$). Then, there exists a positive constant c independent of h and a h_0 such that*

$$\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} + \|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^2(\Omega)} \leq ch^{\frac{2\gamma}{\gamma+1}} \quad \textit{for all } h \leq h_0. \tag{4.24}$$

Proof. Most of the steps of the proof are the same as in the proof of Theorem 4.8. What is different is that instead of Lemma 4.7, we apply Lemma 4.6 together with Lemma 4.9, which allows for all bang-bang optimal controls $\bar{u}$, the estimate

$$|K_1 + K_2| = |J'(\bar{u})(\Pi_h\bar{u} - \bar{u})| \leq \mathrm{Lip}_{\bar{\sigma}} h\|\bar{u} - \Pi_h\bar{u}\|_{L^1(\Omega_h)} \leq ch^2. \tag{4.25}$$

From here, we argue as before, to obtain the estimate

$$\kappa\|\bar{u}_h - \bar{u}\|^{1+\frac{1}{\gamma}}_{L^1(\Omega_h)} \leq J(\bar{u}_h) - J(\bar{u}) \leq ch^2 \ \text{ for all } \bar{u}_h \text{ with } \|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} < \alpha,$$

which yields following the same arguments as before the estimate (4.24). The claim under Assumption 3.5, (4.23), follows again using the same estimations as in the proof of Theorem 4.8 together with the estimate (4.25). Finally, (4.22) is also a direct consequence of the estimations in the proof of Theorem 4.8 and (4.25).

□

4.3. Variational discretization

We prove that Assumptions 3.4, 3.5, and 3.6 with $\beta = 1$ are sufficient for error estimates for a variational discretization scheme. We refer to the [44] for the idea and introduction of variational discretization. The assumptions on the objective functional we are considering are weaker than the ones in [21], still the estimates given in Theorem 4.11 below agree with the estimates in [21], Remark 7 for the variational discretization. We come to the error estimates for the variational discretization.

Theorem 4.11. *Let $\bar{u}$ be a local solution of (1.3). There exist positive constant c and h_0 independent of h such that for any sequence of solutions to the first-order optimality condition of the discrete problems, $\{\bar{u}_h\}_h$, the following holds:*

1. *Let Assumption 3.6($\beta = 1$) be satisfied by $\bar{u}$. Then*
$$\|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^\infty(\Omega)} \le ch \quad \text{for all } h \le h_0. \tag{4.26}$$
2. *Let Assumption 3.5($\beta = 1$) be satisfied by $\bar{u}$. Then*
$$\|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^\infty(\Omega)} \le c(h|\log h|)^2 \quad \text{for all } h \le h_0. \tag{4.27}$$
3. *Let Assumption 3.4($\beta = 1$) be satisfied by $\bar{u}$ for some $\gamma \in (n/(2+n), 1]$. Then*
$$\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} + \|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^\infty(\Omega)} \le c(h|\log h|)^{2\gamma} \quad \text{for all } h \le h_0. \tag{4.28}$$

Proof. We consider (4.28). Since $\bar{u}_h$ satisfies the first-order necessary optimality condition of the discrete problem, it holds

$$\begin{aligned}
0 &\ge J_h'(\bar{u}_h)(\bar{u}_h - \bar{u}) = J'(\bar{u}_h)(\bar{u}_h - \bar{u}) + J_h'(\bar{u}_h)(\bar{u}_h - u) - J'(\bar{u}_h)(\bar{u}_h - \bar{u}) \\
&\ge J'(\bar{u})(\bar{u}_h - \bar{u}) + J''(\bar{u})(\bar{u}_h - \bar{u})^2 - |J'(\bar{u}_h)(\bar{u}_h - \bar{u}) - J'(\bar{u})(\bar{u}_h - \bar{u}) - J''(\bar{u})(\bar{u}_h - \bar{u})^2| \\
&+ J_h'(\bar{u}_h)(\bar{u}_h - u) - J'(\bar{u}_h)(\bar{u}_h - \bar{u}).
\end{aligned} \tag{4.29}$$

Utilizing Taylor's theorem, Lemma 3.9 and Assumption 3.4, we obtain

$$J'(\bar{u}_h)(\bar{u}_h - \bar{u}) - J_h'(\bar{u}_h)(\bar{u}_h - \bar{u}) \ge c\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)}^{1+\frac{1}{\gamma}}. \tag{4.30}$$

Since the discrete problem depends only on the values of the optimal control on the set Ω_h, we may define $\bar{u}_h = \bar{u}$ on $\Omega \setminus \Omega_h$ and write

$$\begin{aligned}
J_h'(\bar{u}_h)(\bar{u}_h - \bar{u}) - J'(\bar{u}_h)(\bar{u}_h - \bar{u}) &= \int_{\Omega_h} (\varphi_h(\bar{u}_h) + L_b(x, y_h(\bar{u}_h)))(\bar{u}_h - \bar{u})\,\mathrm{d}x \\
- \int_{\Omega_h} (\varphi_{\bar{u}_h} + L_b(x, y_{\bar{u}_h}))(\bar{u}_h - \bar{u})\,\mathrm{d}x &= I.
\end{aligned}$$

To estimate I, we follow similar reasoning as in [21], using (4.5), Lemma 4.1, Theorem 4.5 and also using the local Lipschitz property of L_b for y, to infer for some positive constant c.

$$\begin{aligned}
I &\le (\|\bar{\varphi}_h(\bar{u}_h) - \varphi_{\bar{u}_h}\|_{L^\infty(\Omega)} + \mathrm{Lip}_{L_b,y;M}\|y_h(\bar{u}_h) - y_{\bar{u}_h}\|_{L^\infty(\Omega)})\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \\
&\le (\|\varphi_h(\bar{u}_h) - \varphi_{\bar{u}_h}^h\|_{L^\infty(\Omega)} + \|\varphi_{\bar{u}_h}^h - \varphi_{\bar{u}_h}\|_{L^\infty(\Omega)})\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \\
&+ \mathrm{Lip}_{L_b,y;M}\|y_h(\bar{u}_h) - y_{\bar{u}_h}\|_{L^\infty(\Omega)}\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le c(h^2 + h^2|\log h|^2)\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)}.
\end{aligned}$$

Altogether, utilizing (4.30) we obtain for a positive constant again denoted by c that

$$\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \leq c(h^2 + 2h^2|\log h|^2)^\gamma. \tag{4.31}$$

For the states, we use (4.5), Lemma 2.2 and Lemma 2.6 to find for a positive constant c

$$\|y_h(\bar{u}_h) - y_{\bar{u}}\|_{L^2(\Omega)} \leq \|y_h(\bar{u}_h) - y_{\bar{u}_h}\|_{L^2(\Omega)} + \|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^2(\Omega)} \leq c(h^2 + \|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)})$$

and the estimate follows from (4.31). The adjoints can be estimated from here using straightforward arguments. Under Assumption 3.5, by (3.16), it holds

$$c\|\bar{u}_h - \bar{u}\|_{L^1(\Omega)}\|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^2(\Omega)} \leq J'(\bar{u}_h)(\bar{u}_h - \bar{u}) - J_h'(\bar{u}_h)(\bar{u}_h - \bar{u}).$$

Estimating as before, we obtain the existence of a positive constant c that satisfies

$$\|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^2(\Omega)} \leq c(h^2 + h^2|\log h|^2).$$

By again (4.4), (4.11) and (4.12) the claim (4.27) holds. Finally, consider Assumption 3.6 and apply Lemma 3.10 to (4.29), then it holds

$$c\|y_{\bar{u}_h} - y_{\bar{u}}\|^2_{L^2(\Omega)} \leq J'(\bar{u}_h)(\bar{u}_h - \bar{u}) - J_h'(\bar{u}_h)(\bar{u}_h - \bar{u}).$$

To estimate I, we use (4.11)–(4.12) to find

$$\begin{aligned} I &\leq \|\varphi_h(\bar{u}_h) - \varphi_{\bar{u}_h}\|_{L^2(\Omega)}\|\bar{u}_h - \bar{u}\|_{L^2(\Omega_h)} \\ &\leq (\|\varphi_h(\bar{u}_h) - \varphi^h_{\bar{u}_h}\|_{L^2(\Omega)} + \|\varphi^h_{\bar{u}_h} - \varphi_{\bar{u}_h}\|_{L^2(\Omega)})\|\bar{u}_h - \bar{u}\|_{L^2(\Omega_h)} \leq ch^2\|u_a - u_b\|_{L^\infty(\Omega)}, \end{aligned}$$

for some positive constant c. This leads to the estimate $\|y_{\bar{u}_h} - y_{\bar{u}}\|_{L^2(\Omega)} \leq ch$, and by (4.4) and (4.11) the claim (4.26) holds. □

For a numerical example supporting the theoretical error estimates achieved in this paper, especially for the case $\gamma < 1$, we refer to [21].

4.4. Solution stability and variational discretization

One of the main results of this paper is that Assumptions 3.4, 3.5, and 3.6 imply error estimates for the numerical approximation. These assumptions appeared first in the study of the solution stability of optimal control and states under perturbations appearing in the objective functional and the constraining PDE [18–20]. We present an application of the solution stability property to obtain error estimates for a variational discretization scheme. In this sense, Theorem 4.13 below shows that a property related to solution stability guarantees the achievement of error estimates for a variational discretization scheme. The intuition we propose is that once solution stability is obtained under a growth condition on the objective functional, we can expect error estimates for the variational discretization scheme.

First, let us fix a positive constant M and define the set of feasible perturbations by

$$\Big\{\zeta := (\xi, \eta, \rho) \in L^2(\Omega) \times L^2(\Omega) \times L^\infty(\Omega) |\ \|\xi\|_{L^2(\Omega)} + \|\eta\|_{L^2(\Omega)} + \|\rho\|_{L^\infty(\Omega)} \leq M\Big\}.$$

Then, we define the perturbed problem by

$$\min_{u \in \mathcal{U}} \Big\{ J_\zeta(u) := \int_\Omega L(x, y_u, u) + \rho u + \eta y_u \,\mathrm{d}x \Big\} \tag{4.32}$$

subject to (1.2) and

$$\begin{cases} Ay + f(\cdot, y) & = u + \xi \quad \text{in } \Omega, \\ y & = 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.33}$$

The existence of a globally optimal solution to (4.32)–(4.33) is guaranteed by the assumptions on the optimal control problem and the direct method in the calculus of variations. Let us define a property that is strongly related to the notion of solution stability.

Definition 4.12. We call the optimal control problem (1.2)–(1.4) to be solution stable at $\bar{u}$ for $V \subset \mathcal{U}$, with parameters κ, γ and α if

$$\|\bar{u} - \bar{u}^\zeta\|_{L^1(\Omega)} + \|\bar{y} - \bar{y}^\zeta\|_{L^2(\Omega)} + \|\bar{\varphi} - \bar{\varphi}^\zeta\|_{L^\infty(\Omega)} \le \kappa\Big(\|\xi\|_{L^2(\Omega)} + \|\eta\|_{L^2(\Omega)} + \|\rho\|_{L^\infty(\Omega)}\Big)^\gamma$$

for all triples $(\bar{u}^\zeta, \bar{y}^\zeta, \bar{\varphi}^\zeta)$ corresponding to the perturbed problem (4.32)–(4.33) that satisfy $\|\bar{u} - \bar{u}^\zeta\|_{L^1(\Omega)} < \alpha$ and $J'_\zeta(\bar{u}^\zeta)(v - \bar{u}^\zeta) \ge 0$ for all $v \in V$.

Assumption 3.4 implies the notion of solution stability in Definition 4.12. This can be observed by investigating the proof of the strong metric subregularity property of the optimality mapping in [20].

Theorem 4.13. *Let the optimal control problem be solution stable at $\bar{u}$ for $\{\bar{u}\}$ with positive constants γ, κ and α. Let $\{\bar{u}_h\}_h$ be a sequence of solutions to the discrete problems (4.6) with $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega)} < \alpha$. Then there exist positive constants $c(\kappa)$ and h_0 such that*

$$\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} + \|y_h(\bar{u}_h) - \bar{y}\|_{L^2(\Omega)} + \|\varphi_h(\bar{u}_h) - \bar{\varphi}\|_{L^\infty(\Omega)} \le c(\kappa)(h|\log h|)^{2\gamma} \quad \forall h \le h_0.$$

Proof. The idea is to construct a perturbed optimal control problem that relates the continuous problem with the discrete. This is done by considering a certain affine perturbation of the control, similar to the discussion in [45], p. 4. Given a minimizer $\bar{u}_h$ of the discrete problem (4.6), let $\zeta := (0, 0, \rho)$, with $\rho := \varphi_h(\bar{u}_h) - \varphi_{\bar{u}_h}$. Then, we define the perturbed optimal control problem

$$\min_{u \in \mathcal{U}} \left\{ J_\zeta(u) := \int_\Omega L(x, y(x), u(x))\, \mathrm{d}x + \int_\Omega (\varphi_h(\bar{u}_h) - \varphi_{\bar{u}_h}) u \, \mathrm{d}x \right\},$$

subject to (1.4). It is easy to see that $J'_\zeta(\bar{u}_h)(\bar{u} - \bar{u}_h) = \int_\Omega \varphi_h(\bar{u}_h)(\bar{u} - \bar{u}_h)\, \mathrm{d}x \ge 0$. But that is all we need of $\bar{u}_h$ to apply the solution stability at $\bar{u}$. That is, we obtain $\|\bar{u}_h - \bar{u}\|_{L^1(\Omega_h)} \le \kappa \|(\varphi(\bar{u}_h) - \varphi_{\bar{u}_h})\|^\gamma_{L^\infty(\Omega)}$. By Theorem 2.1, Lemma 2.2, Lemma 4.1 and Theorem 4.1 the claim follows. □

Acknowledgements

The author is thankful to E. Casas, A. Domínguez Corella, V. M. Veliov, anonymous readers, and the anonymous referees for many helpful comments and suggestions that improved the quality of the manuscript. This work was supported by the FWF grants P-31400-N32 and I4571-N.

References

[1] A. Domínguez Corella and V.M. Veliov, Hölder regularity in bang-bang type affine optimal control problems, in Large-scale Scientific Computing. Vol. 13127 of *Lecture Notes in Comput. Sci.*. Springer, Cham (2022) 306–313.

[2] U. Felgenhauer, On stability of bang-bang type controls. *SIAM J. Control Optim.* **41** (2003) 1843–1867.

[3] H. Maurer and N.P. Osmolovskii, Second order sufficient conditions for time-optimal bang-bang control. *SIAM J. Control Optim.* **42** (2004) 2239–2263.

[4] A.A. Milyutin and N.P. Osmolovskii, Calculus of Variations and Optimal Control. Vol. 180 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI (1998).

[5] N.P. Osmolovskii and H. Maurer, Equivalence of second order optimality conditions for bang-bang control problems. I. Main results. *Control Cybernet.* **34** (2005) 927–950.

[6] N.P. Osmolovskii and H. Maurer, Equivalence of second order optimality conditions for bang-bang control problems. II. Proofs, variational derivatives and representations. *Control Cybernet.* **36** (2007) 5–45.

[7] N.P. Osmolovskii and V.M. Veliov, On the regularity of Mayer-type affine optimal control problems, in Large-scale Scientific Computing. Vol. 11958 of *Lecture Notes in Comput. Sci.* Springer (2020) 56–63.

[8] J. Preininger, T. Scarinci and V.M. Veliov, On the regularity of linear-quadratic optimal control problems with bang-bang solutions, in Large-scale Scientific Computing. Vol. 10665 of *Lecture Notes in Comput. Sci.*. Springer, Cham (2018) 237–245.

[9] J. Preininger, T. Scarinci and V.M. Veliov, Metric regularity properties in bang-bang type linear-quadratic optimal control problems. *Set-Valued Var. Anal.* **27** (2019) 381–404.

[10] M. Quincampoix and V.M. Veliov, Metric regularity and stability of optimal control problems for linear systems. *SIAM J. Control Optim.* **51** (2013) 4118–4137.

[11] E. Casas, Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.* **50** (2012) 2355–2372.

[12] J.F. Bonnans, Optimal control of a semilinear parabolic equation with singular arcs. *Optim. Methods Softw.* **29** (2014) 964–978.

[13] E. Casas, M. Mateos and A. Rösch, Error estimates for semilinear parabolic control problems in the absence of Tikhonov term. *SIAM J. Control Optim.* **57** (2019) 2515–2540.

[14] E. Casas, C. Ryll and F. Tröltzsch, Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation. *SIAM J. Control Optim.* **53** (2015) 2168–2202.

[15] E. Casas and F. Tröltzsch, Second-order and stability analysis for state-constrained elliptic optimal control problems with sparse controls. *SIAM J. Control Optim.* **52** (2014) 1010–1033.

[16] E. Casas and F. Tröltzsch, Second-order optimality conditions for weak and strong local solutions of parabolic optimal control problems. *Vietnam J. Math.* **44** (2016) 181–202.

[17] E. Casas, D. Wachsmuth and G. Wachsmuth, Sufficient second-order conditions for bang-bang control problems. *SIAM J. Control Optim.* **55** (2017) 3066–3090.

[18] E. Casas, A. Domínguez Corella and N. Jork, New assumptions for stability analysis in elliptic optimal control problems. *SIAM J. Control Optim.* **61** (2023) 1394–1414.

[19] A. Domínguez Corella, N. Jork and V. Veliov, Stability in affine optimal control problems constrained by semilinear elliptic partial differential equations. *ESAIM Control Optim. Calc. Var.* **28** (2022) Paper No. 79.

[20] A. Domínguez Corella, N. Jork and V. Veliov, On the solution stability of parabolic optimal control problems. *Comput. Optim. Appl.* **86** (2023) 1035–1079.

[21] E. Casas and M. Mateos, State error estimates for the numerical approximation of sparse distributed control problems in the absence of Tikhonov regularization. *Vietnam J. Math.* **49** (2021) 713–738.

[22] E. Casas, D. Wachsmuth and G. Wachsmuth, Second-order analysis and numerical approximation for bang-bang bilinear control problems. *SIAM J. Control Optim.* **56** (2018) 4203–4227.

[23] K. Deckelnick and M. Hinze, A note on the approximation of elliptic control problems with bang-bang controls. *Comput. Optim. Appl.* **51** (2012) 931–939.

[24] E. Casas and M. Mateos, Critical cones for sufficient second-order conditions in PDE constrained optimization. *SIAM J. Optim.* **30** (2020) 585–603.

[25] I. Mazari, Quantitative inequality for the eigenvalue of a Schrödinger operator in the ball. *J. Diff. Equ.* **269** (2020) 10181–10238.

[26] I. Mazari, Quantitative estimates for parabolic optimal control problems under L^∞ and L^1 constraints in the ball: quantifying parabolic isoperimetric inequalities. *Nonlinear Anal.* **215** (2022) Paper No. 112649.

[27] E. Casas, M. Mateos and A. Rösch, Analysis of control problems of nonmontone semilinear elliptic equations. *ESAIM Control Optim. Calc. Var.* **26** (2020) Paper No. 80.

[28] E. Casas, M. Mateos and A. Rösch, Numerical approximation of control problems of non-monotone and non-coercive semilinear elliptic equations. *Numer. Math.* **149** (2021) 305–340.

[29] D. Gilbarg and N.S. Trudinger, Elliptic Partial Differential Equations of Second Order. Vol. 224 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 2nd edn. Springer-Verlag, Berlin (1983).

[30] G. Stampacchia, Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)* **15** (1965) 189–258.

[31] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods and Applications. Vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Philadelphia (2010).

[32] T. Bayen, J.F. Bonnans and F.J. Silva, Characterization of local quadratic growth for strong minima in the optimal control of semi-linear elliptic equations. *Trans. Amer. Math. Soc.* **366** (2014) 2063–2087.

[33] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Vol. 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA (1985).

[34] E. Casas, Pontryagin's principle for optimal control problems governed by semilinear elliptic equations, in *Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena (Vorau, 1993)*. Vol. 118 of *Internat. Ser. Numer. Math.* Birkhäuser, Basel (1994) 97–114.

[35] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods. Vol. 15 of *Texts in Applied Mathematics*, 2nd edn. Springer-Verlag, New York (2002).

[36] P.A. Raviart and J.M. Thomas, *Introduction à l'analyse numérique des équations aux dérivées partielles.* (1983).

[37] E. Casas and M. Mateos, Uniform Convergence of the FEM. Applications to State Constrained Control Problems. Vol. 21 (2002) 67–100.

[38] L.C. Evans and R.F. Gariepy, Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL (1992).

[39] J. Bourgain, J. Kristensen and M.V. Korobkov, On the Morse-Sard property and level sets of Sobolev and BV functions. *Rev. Mat. Iberoam.* **29** (2013) 1–23.

[40] L. De Pascale, The Morse-Sard theorem in Sobolev spaces. *Indiana Univ. Math. J.* **50** (2001) 1371–1386.

[41] A. Figalli, A simple proof of the Morse-Sard theorem in Sobolev spaces. *Proc. Am. Math. Soc.* **136** (2008) 3675–3681.

[42] C. Clason, V.H. Nhu and A. Rösch, Numerical analysis of a nonsmooth quasilinear elliptic control problem. I. Explicit second-order optimality conditions. (2022) Preprint.

[43] K. Deckelnick and M. Hinze, A note on the approximation of elliptic control problems with bang-bang controls. *Comput. Optim. Appl.* **51** (2012) 931–939.

[44] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30** (2005) 45–61.

[45] E. Kammann, F. Tröltzsch and S. Volkwein, *A posteriori* error estimation for semilinear parabolic optimal control problems with application to model reduction by pod. *ESAIM Math. Model. Numer. Anal.* **47** (2013) 555–581.