

ERROR ESTIMATES FOR THE FINITE ELEMENT APPROXIMATION OF A SEMILINEAR ELLIPTIC CONTROL PROBLEM WITH STATE CONSTRAINTS AND FINITE DIMENSIONAL CONTROL SPACE

P. MERINO[‡] F. TRÖLTZSCH[†] B. VEXLER^{*}

ABSTRACT. The finite element approximation of optimal control problems for semilinear elliptic partial differential equation is considered, where the control belongs to a finite-dimensional set and state constraints are given in finitely many points of the domain. Under the standard linear independency condition on the active gradients and a strong second-order sufficient optimality condition, optimal error estimates are derived for locally optimal controls.

1. INTRODUCTION

In this paper, we derive error estimates for a finite element based approximation of a semilinear elliptic optimal control problem with control out of a finite-dimensional set and constraints on the state which are required only in finitely many points of the spatial domain. This type of problems is interesting for several reasons. Finite-dimensional controls appear in many real applications of optimal control theory, because it is rather difficult to practically implement control functions that can vary arbitrarily in space. We refer also to [10], where examples of problems with practical importance are mentioned. If pointwise state constraints are given with finite-dimensional control space, then they are often active in finitely many points only. Therefore, finitely many point constraints turn out to be interesting for numerical computations, if the location of active points has been detected.

Our problem is equivalent to a nonlinear programming problem in a finite-dimensional space. Therefore, the existence of optimal controls, first-or second-order optimality conditions are more or less standard results since they follow from classical results of mathematical programming after transforming the control problem into a finite-dimensional one. We refer to [10], where optimality conditions for semilinear control problems with finite-dimensional control space are discussed with pointwise state constraints given in the whole spatial domain.

The question of error estimates, however, is still interesting. Numerical computations indicate that the numerical approximation of the optimal controls and states is of the order h^2 , where $h > 0$ is the mesh size of the finite element scheme. We can show that the order of the error is equal to $h^2|\log h|$, which corresponds to the error estimate for the finite-element approximation of the state in L^∞ -norm. This is quite clear intuitively, since the control space is finite-dimensional and the state constraints are posed pointwise

in a finite number of points. Then, no interpolation error of the optimal control occurs that in the case of control functions limits the order of convergence to h for a cellwise constant approximation of the control variable, see, e.g., [3], and to $h^{3/2}$ for a piecewise linear approximation, see, e.g., [30] and [13].

For pointwise state constraints in the whole domain, we know only a few results. In Casas [7], the convergence of finite element approximations to optimal control problems for semilinear elliptic equations with finitely many state constraints is proven for piecewise constant approximations of the control function. This result was extended by Casas and Mateos [8] to boundary control problems.

Error estimates were derived by Deckelnick and Hinze [11] and Meyer [23], who consider linear-quadratic problems with pointwise state constraints and obtain the order $h^{1-\varepsilon}$ for domains of the dimension $d = 2$ and $h^{1/2-\varepsilon}$ for $d = 3$ under the assumption that the control function is not discretized. Recently, Deckelnick and Hinze [12] estimated the error by $h|\log h|$ for $d = 2$ and $h^{1/2}$ for $d = 3$ with piecewise constant approximation of the control.

Extending L^∞ -error estimates from [27] for linear elliptic equations to the semilinear case, we show that the order is precisely $h^2|\log h|$. At first glance, this is surprising since the Lagrange multipliers associated with the (finitely many) pointwise state constraints are regular Borel measures, which appear in the adjoint equation of the control system. It is the finite-dimensional nature of our problem that explains the high order of the error.

This problem of error estimates leads to investigating a perturbed nonlinear programming problem, where the perturbation parameter is the mesh size h – seemingly a standard problem of perturbation analysis. However, we were not able to find approximation results in the literature, which can be applied directly to our problem. In view of this, we include a section on a perturbation result for perturbed nonlinear programming problems, where we also comment on related papers.

The paper is organized as follows: In Section 2 we define the optimal control problem (OCP), introduce the notation and state the assumptions. In Section 3, the finite element discretization is introduced and error estimates for the partial differential equations are derived. Moreover, this section is devoted to estimating the error for the solutions of the optimal control problem. Numerical examples are presented in Section 4, and Section 5 contains the error analysis for perturbed nonlinear programming problems.

2. OPTIMAL CONTROL PROBLEM AND OPTIMALITY CONDITIONS

2.1. Definition of the problem and main assumptions. We consider the following optimal control problem:

$$(\text{OCP}) \quad \begin{cases} \min_{u \in \mathcal{U}_{ad}} J(y_u, u) = \int_{\Omega} L(x, y_u(x), u) dx \\ \text{subject to} \\ g_i(y_u(x_i)) = 0, & \text{for all } i = 1, \dots, k, \\ g_i(y_u(x_i)) \leq 0, & \text{for all } i = k + 1, \dots, \ell, \end{cases}$$

where y_u is the solution to the state equation

$$\begin{aligned} Ay(x) + d(x, y(x), u) &= 0 \quad \text{in } \Omega \\ y(x) &= 0 \quad \text{on } \Gamma \end{aligned} \tag{2.1}$$

and \mathcal{U}_{ad} is the set of box constraints defined by $\mathcal{U}_{ad} = \{u \in \mathbb{R}^m : u_a \leq u \leq u_b\}$ with given vectors $u_a \leq u_b$ of \mathbb{R}^m . We assume $l \geq 1$ and set $k = 0$, if only inequality constraints are given and $k = l$, if only equality constraints are given.

The set Ω is a convex bounded open set in \mathbb{R}^2 with boundary Γ . For simplicity, we assume that Γ is polygonal. The extension of some of our results to curved boundaries is possible along the lines of Arada et al. [3]. Moreover, functions $L, d: \Omega \times \mathbb{R}^{m+1} \rightarrow \mathbb{R}$, and $g_i: \mathbb{R} \mapsto \mathbb{R}$ of class C^2 are given together with points $x_i \in \Omega$, $i = 1 \dots \ell$. Notice that these points are not located at the boundary of Ω .

The operator A is a symmetric uniformly elliptic differential operator of the form

$$Ay(x) = - \sum_{i,j=1}^n \partial_j(a_{ij}(x)) \partial_i y(x)$$

with coefficients $a_{ij} \in C^{1+\alpha}(\Omega)$, $0 < \alpha < 1$. The control $u \in \mathbb{R}^m$ is allowed to occur nonlinearly in the state equation, since the controls are vectors. For control functions, the proof of existence would be a delicate issue.

Notation: If not stated otherwise, we consider our vectors as column vectors. By $|\cdot|$, the Euclidean norm of vectors in spaces \mathbb{R}^n is denoted, while $\|\cdot\|$ stands for the norm of matrices (defined by the Euclidean norm of the vector of all entries). The open ball around u in \mathbb{R}^m with radius ρ is denoted by $B(u, \rho)$. If $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a twice differentiable function, then f' is its derivative and ∇f is the associated gradient, i.e. the representation of f' by a column vector. Moreover, $f''(z)$ denotes the Hessian matrix of f at the vector z . We denote by c a generic constant, while C is a constant appearing in all estimates of our theorems; it is the maximum of all associated constants.

Main assumptions:

(A.1) (Carathéodory type assumptions) For each fixed $x \in \Omega$, the functions $L = L(x, y, u)$ and $d = d(x, y, u)$ are of class C^2 with respect to (y, u) . For all fixed (y, u) or fixed y , respectively, they are Hölder continuous with respect to the variable $x \in \Omega$.

(A.2) (Monotonicity) For all $x \in \Omega$, all $u \in \mathcal{U}_{ad}$ and $y \in \mathbb{R}$, it holds that

$$\frac{\partial d}{\partial y}(x, y, u) \geq 0.$$

(A.3) (Boundedness and Lipschitz properties) There is a constant C and, for all $M > 0$,

a constant $c_L(M) > 0$ such that the estimates

$$|d(x, 0, 0)| + |d'(x, 0, 0)| + \|d''(x, 0, 0)\| \leq C$$

$$\|d''(x, y_1, u_1) - d''(x, y_2, u_2)\| + \|g_j''(y_1) - g_j''(y_2)\| \leq C_L(M)(|y_1 - y_2| + |u_1 - u_2|)$$

hold for all $x \in \Omega$, all $u_i \in \mathcal{U}_{ad}$ and all $|y_i| \leq M$, $i = 1, 2$, $j = 1, \dots, \ell$. Here, d' and d'' denote the derivative and the Hessian matrix of $d(x, y, u)$ with respect to (y, u) , respectively.

The function L is assumed to satisfy **(A.3)** accordingly.

Lemma 1. *Under the assumptions **(A.1)** - **(A.3)**, to each $u \in \mathcal{U}_{ad}$ there exists a unique solution $y_u \in H_0^1(\Omega) \cap C(\bar{\Omega})$. There is a constant C such that $\|y_u\|_{L^\infty(\Omega)} \leq C$ is satisfied for all $u \in \mathcal{U}_{ad}$. If Ω is convex or of class C^2 , then y_u belongs to $H^2(\Omega)$.*

The first part of this result and the uniform boundedness are standard, we refer to Casas [6]. The H^2 -regularity follows from results by Grisvard [17] on linear equations after taking the nonlinear terms with $d(\cdot, y_u, u) \in L^\infty(\Omega)$ to the right-hand side of the equation.

The state y that is associated to the control u by the PDE in (OCP), is denoted by y_u . We denote the mapping $u \mapsto y_u$ by $S: \mathbb{R}^m \rightarrow H_0^1(\Omega) \cap C(\bar{\Omega})$, i.e. $y_u = S(u)$; S is our control-to-state mapping.

Lemma 2. *Under the assumptions **(A.1)** - **(A.3)**, the control-to-state mapping $S: \mathbb{R}^m \rightarrow H_0^1(\Omega) \cap C(\bar{\Omega})$ is twice continuously Fréchet-differentiable. For arbitrary elements u and v of \mathbb{R}^m , the function $z_v(u) := S'(u)v$ is given by $z_v(u) = z$, where z is the unique solution to the problem*

$$\begin{cases} Az + \frac{\partial d}{\partial y}(x, y_u, u)z = -\frac{\partial d}{\partial u}(x, y_u, u)v & \text{in } \Omega \\ z = 0 & \text{on } \Gamma, \end{cases} \quad (2.2)$$

and the inequality

$$\|z_v(u)\|_{H^1(\Omega)} + \|z_v(u)\|_{C(\bar{\Omega})} \leq C|v| \quad (2.3)$$

is satisfied with some constant C that is independent of $u \in \mathcal{U}_{ad}$.

The function $z_{v_1 v_2}(u) := S''(u)[v_1, v_2]$ is obtained by $z_{v_1 v_2} = z$, where z is the unique solution of

$$\begin{cases} Az + \frac{\partial d}{\partial y}(x, y_u, u)z = -(\mathbf{y}'_{v_1}, v_1^\top) d''(x, y_u, u) (\mathbf{y}'_{v_2}, v_2^\top)^\top & \text{in } \Omega \\ z = 0 & \text{on } \Gamma \end{cases} \quad (2.4)$$

with $\mathbf{y}'_{v_1} = S'(u)(v_1)$ and $\mathbf{y}'_{v_2} = S'(u)(v_2)$.

For later use of the theory of nonlinear optimization in finite-dimensional spaces, we convert problem **(OCP)** into a finite-dimensional nonlinear programming problem. We introduce the reduced objective function $f: \mathbb{R}^m \mapsto \mathbb{R}$ of class $C^{2,1}$ by

$$f(u) = J(y_u, u) = J(S(u), u).$$

By our assumptions, in particular the Lipschitz properties of the second derivatives of the given nonlinear functions and the boundedness of \mathcal{U}_{ad} , it follows that S'' is Lipschitz on \mathcal{U}_{ad} , i.e. there exists a constant $C > 0$ such that

$$\|S''(u_1)[v_1, v_2] - S''(u_2)[v_1, v_2]\|_{H_0^1(\Omega) \cap C(\bar{\Omega})} \leq C|u_1 - u_2| |v_1| |v_2| \quad \forall u_i \in \mathcal{U}_{ad}, v_i \in \mathbb{R}^m. \quad (2.5)$$

Moreover, we define the restriction function G by

$$G(u) = [g_1(y_u(x_1)), \dots, g_\ell(y_u(x_\ell))]^\top. \quad (2.6)$$

An equivalent formulation is

$$G(u) = [g_1(\delta_1 S(u)), \dots, g_\ell(\delta_\ell S(u))]^\top, \quad (2.7)$$

where δ_i denotes the Dirac measure concentrated at the point x_i , $i = 1, \dots, \ell$. It shows that Dirac measures will naturally appear in the adjoint equation associated with (OCP).

By these definitions, (OCP) becomes equivalent to the finite-dimensional nonlinear programming problem (NP),

$$(NP) \quad \begin{cases} \min f(u) \\ G_i(u) = 0, & i = 1, \dots, k, \\ G_i(u) \leq 0, & i = k + 1, \dots, \ell, \\ u \in \mathcal{U}_{ad}. \end{cases}$$

Remark 1. *This problem admits at least one optimal solution \bar{u} , if its admissible set is non-empty. This follows immediately by the Weierstraß theorem, since f and G are continuous and \mathcal{U}_{ad} is compact. Therefore, by its equivalence to (NP), also (OCP) is solvable in this case.*

3. THE DISCRETIZED OPTIMAL CONTROL PROBLEM

3.1. Finite element scheme and discretized optimal control problem. To discretize the optimal control problem, we introduce a finite element approximation of the state equation (2.1). We consider a family of meshes $\{\mathcal{T}_h\}_{h>0}$ consisting of triangles $T \in \mathcal{T}_h$ such that

$$\cup_{T \in \mathcal{T}_h} T = \bar{\Omega}.$$

Notice that Ω was supposed to be polygonal for simplicity. For each triangle $T \in \mathcal{T}_h$, we introduce the diameter $\rho(T)$ of T , and the diameter $\sigma(T)$ of the largest circle contained in T . The mesh size h is defined by $h = \max_{T \in \mathcal{T}_h} \rho(T)$. We impose the following regularity assumption on the grid:

(A.4) *There exist two positive constants ρ and σ such that*

$$\frac{\rho(T)}{\sigma(T)} \leq \sigma \quad \text{and} \quad \frac{h}{\rho(T)} \leq \rho, \quad \forall T \in \mathcal{T}_h,$$

are fulfilled for all $h > 0$.

Associated with the given triangulation \mathcal{T}_h , we introduce the set of piecewise linear and continuous functions

$$Y_h = \{y_h \in C(\bar{\Omega}) : y_h|_T \in P_1(T) \quad \forall T \in \mathcal{T}_h, y_h = 0 \text{ on } \Gamma\},$$

where $P_1(T)$ denotes the set of affine real-valued functions defined on T .

For convenience, we introduce the bilinear form

$$a(y, \eta) = \int_{\Omega} \sum_{i,j=1}^n a_{ij}(x) \partial_{x_i} y \partial_{x_j} \eta \, dx.$$

The discrete state y_h is defined as the (unique) element of Y_h that satisfies the following finite element scheme associated with (2.1):

$$a(y_h, \eta_h) + \int_{\Omega} d(x, y_h, u) \eta_h \, dx = 0 \quad \forall \eta_h \in Y_h. \quad (3.1)$$

The following existence and approximation result was proven in [9] for distributed control functions in Lipschitz domains. It is obvious that it remains valid without changes for our case with control vectors. It shows that the discretized equation is well posed.

Theorem 1. *Let the assumptions (A.1)–(A.3) and (A.4) be satisfied. Then, for all $u \in \mathcal{U}_{ad}$, the equation (3.1) has a unique solution $y_{h,u} \in Y_h$. There exists a constant C independent of h and $u \in \mathcal{U}_{ad}$ such that, with $n = \dim \Omega$,*

$$\|y_u - y_{h,u}\|_{L^2(\Omega)} \leq C h^2 \|y_u\|_{H^2(\Omega)} \quad (3.2)$$

$$\|y_u - y_{h,u}\|_{L^\infty(\Omega)} \leq C h^{2-n/2} \|y_u\|_{H^2(\Omega)}. \quad (3.3)$$

To indicate the correspondence of y_h to u , we will denote this solution also by $y_{h,u}$. The mapping $u \mapsto y_{h,u}$ is denoted by S_h . We will improve estimate (3.3) in an interior subdomain Ω_0 containing the points x_i , where the state constraints are required.

We approximate (OCP) by substituting the discretized equation (3.1) for the state equation. We study the following (discretized) finite-dimensional control problem

$$(OCP_h) \quad \begin{cases} \min_{u \in \mathcal{U}_{ad}} J(y_{h,u}, u) = \int_{\Omega} L(x, y_{h,u}(x), u) \, dx \\ g_i(y_{h,u}(x_i)) = 0 \quad \text{for all } i = 1, \dots, k, \\ g_i(y_{h,u}(x_i)) \leq 0 \quad \text{for all } i = k+1, \dots, \ell. \end{cases} \quad (3.4)$$

Let us define

$$f_h(u) = J(y_{h,u}, u), \quad G_h(u) = [g_1(y_{h,u}(x_1)), \dots, g_\ell(y_{h,u}(x_\ell))]^\top.$$

By these terms, we transform the problem (3.4) to the finite-dimensional nonlinear programming problem

$$(NP_h) \quad \begin{cases} \min f_h(u) \\ G_{h,i}(u) = 0, & i = 1, \dots, k, \\ G_{h,i}(u) \leq 0, & i = k + 1, \dots, \ell, \\ u \in \mathcal{U}_{ad}. \end{cases} \quad (3.5)$$

3.2. Improved L^∞ -error estimate. Let Ω_0 with $\bar{\Omega}_0 \subset \Omega$ be a subdomain of Ω that contains all given points x_i , $i = 1, \dots, \ell$. The main goal of this subsection is to estimate the L^∞ -norm of the difference $y_u - y_{h,u}$ in Ω_0 . We rely on an associated estimate for linear equations from [27] and on Theorem 1 by Casas and Mateos [9]. An estimate of the order $h^2 |\log(h)|$ was derived in [27] for compact subdomains of polygonal domains. We also mention the related papers [26] and [15], where the error estimate $h^2 |\log(h)|$ was shown for domains with sufficiently smooth boundary. Another technique for proving local error estimates with respect to the L^∞ -norm is presented in [31, 32]. Here, we extend the result of [27] to the semilinear case.

First, we show that the state y_u belongs to $W^{2,\infty}(\Omega_1)$, where $\bar{\Omega}_0 \subset \Omega_1$ and $\bar{\Omega}_1 \subset \Omega$.

Lemma 3. *Under the assumptions (A.1)-(A.3), the state y_u belongs to $W^{2,\infty}(\Omega_1)$ for all $u \in \mathcal{U}_{ad}$.*

Proof. We already know that y_u is bounded and measurable, hence the same is true for $d(\cdot, y_u, u)$. Writing the state equation in the form $Ay_u = -d(\cdot, y_u, u)$, we see that y_u solves a linear elliptic equation with right-hand side in $L^2(\Omega)$. Since Ω is convex and bounded, $y_u \in H^2(\Omega)$ follows from Grisvard [17], hence y_u is Hölder continuous, because Ω is two-dimensional. The smoothness and Lipschitz assumptions on d ensure $d(\cdot, y_u, u) \in C^{0,\kappa}(\bar{\Omega})$ with some $\kappa \in (0, 1)$. Therefore, y_u solves a linear elliptic boundary value problem

$$Ay_u = f, \quad y_\gamma = 0$$

with Hölder continuous right-hand side f and continuous boundary data. As a convex bounded domain, Ω satisfies the exterior sphere condition. In view of this, we can apply Theorem 6.13 by Gilbarg and Trudinger [16] and obtain $y_u \in C^{2,\alpha}(\Omega)$ with some $\alpha \in (0, 1)$. This implies in particular $y_u \in C^{2,\alpha}(\bar{\Omega}_1)$, hence the statement of the Lemma is true. \square

Now we are going to derive a uniform error estimate of the order $h^2 |\log h|$ in Ω_0 . The following estimate is known, e.g., from [27], for the linear case, cf. also [31, 32]:

$$\|y_{h,u} - y_u\|_{C(\Omega_0)} \leq C h^2 |\log h| (\|y_u\|_{W^{2,\infty}(\Omega_1)} + \|y_u\|_{H^2(\Omega)}), \quad (3.6)$$

where $\bar{\Omega}_0 \subset \Omega_1$ and $\bar{\Omega}_1 \subset \Omega$. We shall extend this estimate to the semilinear equation (2.1).

To establish our error estimate, we rely on the following stability result for linear elliptic equations from [27]:

Theorem 2. Let Ω_1 be a subdomain of Ω with $\bar{\Omega}_0 \subset \Omega_1$ and $\bar{\Omega}_1 \subset \Omega$ and assume that $B \in (W^{1,\infty}(\Omega))^{2 \times 2}$ and $\psi \in H_0^1(\Omega) \cap C(\bar{\Omega})$ are given. If $z_h \in Y_h$ is the solution of

$$a(z_h, \eta_h) = (B \nabla \psi, \nabla \eta_h) \quad \forall \eta_h \in Y_h,$$

then there holds

$$\|z_h\|_{L^\infty(\Omega_0)} \leq c \|B\|_{W^{1,\infty}(\Omega)} (|\log(h)| \|\psi\|_{L^\infty(\Omega_1)} + \|\psi\|_{L^2(\Omega)} + h \|\psi\|_{H^1(\Omega)}).$$

Our main error estimate for the state equation is expressed by the following result:

Theorem 3. Let Ω_0 and Ω_1 be defined as in Theorem 2. Then there exists a constant $C > 0$, independent of h and $u \in \mathcal{U}_{ad}$ such that it holds

$$\|y_u - y_{h,u}\|_{L^\infty(\Omega_0)} \leq C (h^2 |\log h| \|y_u\|_{W^{2,\infty}(\Omega_1)} + h^2 \|y_u\|_{H^2(\Omega)}).$$

Proof. In the proof, we keep the control u in (2.1) fixed. In view of this, we denote y_u by y and $y_{h,u}$ by y_h to simplify the notation.

(i) The state y_u is bounded, hence the L^∞ -estimate (3.3) implies uniform boundedness of y_h , independently of h . Therefore, we get from the Lipschitz continuity of d with respect to y and from the estimate (3.2)

$$\|d(\cdot, y, u) - d(\cdot, y_h, u)\|_{L^2(\Omega)} \leq c \|y - y_h\|_{L^2(\Omega)} \leq c h^2 \|y\|_{H^2(\Omega)}. \quad (3.7)$$

We consider an intermediate solution $\tilde{y}_h \in Y_h$ determined by

$$a(\tilde{y}_h, \eta_h) = -(d(\cdot, y, u), \eta_h) \quad \forall \eta_h \in Y_h,$$

where (\cdot, \cdot) denotes the inner product of $L^2(\Omega)$. It is clear that y satisfies the same equation, hence

$$a(\tilde{y}_h, \eta_h) = a(y, \eta_h) \quad \forall \eta_h \in Y_h.$$

Inserting the interpolate $i_h y$ of y , we obtain

$$a(i_h y - \tilde{y}_h, \eta_h) = a(i_h y - y, \eta_h) \quad \forall \eta_h \in Y_h.$$

Now we apply Theorem 2 to $\psi = i_h y - y$ and deduce

$$\|i_h y - \tilde{y}_h\|_{L^\infty(\Omega_0)} \leq c (|\log h| \|i_h y - y\|_{L^\infty(\Omega_1)} + \|i_h y - y\|_{L^2(\Omega)} + h \|i_h y - y\|_{H^1(\Omega)}). \quad (3.8)$$

Due to known properties of the interpolation operator, cf. Brenner and Scott [5], we have

$$\begin{aligned} \|i_h y - y\|_{L^2(\Omega)} + h \|i_h y - y\|_{H^1(\Omega)} &\leq c h^2 \|y\|_{H^2(\Omega)} \\ \|i_h y - y\|_{L^\infty(\Omega_i)} + h \|i_h y - y\|_{W^{1,\infty}(\Omega_i)} &\leq c h^2 \|y\|_{W^{2,\infty}(\Omega_i)}, \quad i = 0, 1. \end{aligned}$$

Using these interpolation estimates in (3.8), we arrive at

$$\|y - \tilde{y}_h\|_{L^\infty(\Omega_0)} \leq c (h^2 |\log h| \|y\|_{W^{2,\infty}(\Omega_1)} + h^2 \|y\|_{H^2(\Omega)}). \quad (3.9)$$

(ii) It remains to estimate the error $v_h = \tilde{y}_h - y_h \in Y_h$. Subtracting the equations for \tilde{y}_h and y_h , we find that v_h solves

$$a(v_h, \eta_h) = (d(\cdot, y_h, u) - d(\cdot, y, u), \eta_h) \quad \forall \eta_h \in Y_h.$$

Now consider the function $v \in H_0^1(\Omega) \cap C(\bar{\Omega})$ defined by

$$a(v, \eta) = (d(\cdot, y_h, u) - d(\cdot, y, u), \eta) \quad \forall \eta \in H_0^1(\Omega).$$

Obviously, v_h is the discrete solution belonging to v . From standard estimates for linear elliptic equations and from inequality (3.7), it follows that

$$\|v\|_{L^\infty(\Omega)} + \|v\|_{H^1(\Omega)} \leq c \|d(\cdot, y, u) - d(\cdot, y_h, u)\|_{L^2(\Omega)} \leq c h^2 \|y\|_{H^2(\Omega)}.$$

Now, we apply Theorem 2 to the linear equation

$$a(v_h, \eta_h) = a(v, \eta_h) \quad \forall \eta_h \in Y_h$$

and obtain

$$\|v_h\|_{L^\infty(\Omega_0)} \leq c \left(|\log(h)| \|v\|_{L^\infty(\Omega_1)} + \|v\|_{L^2(\Omega)} + h \|v\|_{H^1(\Omega)} \right) \leq c |\log(h)| h^2 \|y\|_{H^2(\Omega)}.$$

In view of (3.9), the inequality $\|y - y_h\|_{L^\infty(\Omega_0)} \leq \|y - \tilde{y}_h\|_{L^\infty(\Omega_0)} + \|v_h\|_{L^\infty(\Omega_0)}$ yields the statement of the theorem. \square

3.3. Application to the optimal control problem. Let us first summarize the error estimates of the preceding subsection in the following lemma:

Lemma 4. *Let y_u and $y_{h,u}$ be the solutions of the equations (2.1) and (3.1), respectively. Let $z_v(u)$ solve (2.2) and $z_{v,h}(u)$ satisfy the corresponding discretized equation. Then there exists a constant $C > 0$, independent of h , such that the estimates*

$$\|y_u - y_{h,u}\|_{C(\bar{\Omega}_0)} \leq C h^2 |\log h| \quad (3.10)$$

$$\|z_v(u) - z_{v,h}(u)\|_{C(\bar{\Omega}_0)} \leq C h^2 |\log h| |v| \quad (3.11)$$

hold for all $u \in \mathcal{U}_{ad}$, $v \in \mathbb{R}^m$ and $h > 0$.

Proof. Lemma 1 and the Lipschitz properties of d ensure the existence of a bound $M > 0$ with $\|d(\cdot, y_u, u)\|_{L^2(\Omega)} \leq M$ for all $u \in \mathcal{U}_{ad}$. In view of this and by Lemma 3, $\|y_u\|_{H^2(\Omega)}$ and $\|y_u\|_{W^{2,\infty}(\Omega_1)}$ are uniformly bounded for all $u \in \mathcal{U}_{ad}$ and (3.10) is an immediate conclusion of Theorem 3. Estimate (3.11) follows from (3.10) as in [3], Proposition 7.1. Notice that, due to the boundedness of \mathcal{U}_{ad} , the functions y_u and $y_{h,u}$ belong to a bounded set. This is used for the proof in [3]. \square

Lemma 4 implies an estimate of $G - G_h$ in the C^2 -norm:

Lemma 5. *Under the assumptions (A.1)–(A.3) and (A.4), there is a constant $C > 0$ such that*

$$\begin{aligned} & |G(u) - G_h(w)| + \|G'(u) - G'_h(w)\| + \\ & + \sum_{i=1}^{\ell} \|G''_i(u) - G''_{h,i}(w)\| \leq C(|u - w| + h^2 |\log h|) \end{aligned} \quad (3.12)$$

holds for all u, w in \mathcal{U}_{ad} .

Proof. Let c denote a generic constant. Taking into account that all g_i are of class $C^{2,1}$, hence also the functions G_i , in view of Lemma 4 we have

$$\begin{aligned} |G_i(u) - G_{h,i}(w)| &= |g_i(y_u(x_i)) - g_i(y_{h,w}(x_i))| \leq c \|y_u - y_{h,w}\|_{C(\bar{\Omega}_0)} \\ &\leq c \|y_u - y_w\|_{C(\bar{\Omega}_0)} + c \|y_w - y_{h,w}\|_{C(\bar{\Omega}_0)} \end{aligned}$$

$$\leq c|u - w| + ch^2|\log h| \quad \forall i = 1, \dots, \ell,$$

where the last inequality is obtained since the mapping $u \mapsto y_u$ is Lipschitz. For the derivative G' , we argue similarly. Let w be an arbitrary unit vector of \mathbb{R}^m . Then

$$\begin{aligned} |G'_i(u)v - G'_{h,i}(w)v| &\leq |G'_i(u)v - G'_i(w)v| + |G'_i(w)v - G'_{h,i}(w)v| \\ &\leq |g'_i(y_u(x_i))(S'(u)v)(x_i) - g'_i(y_w(x_i))(S'(w)v)(x_i)| \\ &\quad + |g'_i(y_w(x_i))(S'(w)v)(x_i) - g'_i(y_{h,w}(x_i))(S'_h(w)v)(x_i)| \\ &\leq c \|g'_i(y_u)z_v(u) - g'_i(y_w)z_v(w)\|_{C(\bar{\Omega}_0)} \\ &\quad + \|g'_i(y_w)z_v(w) - g'_i(y_{h,w})z_{h,v}(w)\|_{C(\bar{\Omega}_0)}, \end{aligned}$$

where $z_v(u)$ and $z_v(w)$ are defined as in Theorem 2 and $z_{h,v}(w)$ is the discrete state associated with $z_v(w)$. With a constant C we obtain that

$$\begin{aligned} \|y_u - y_w\|_{C(\bar{\Omega}_0)} &\leq C|u - w|, \quad \|y_w - y_{h,w}\|_{C(\bar{\Omega}_0)} \leq Ch^2|\log h| \\ \|z_v(w) - z_{h,v}(w)\|_{C(\bar{\Omega}_0)} &\leq Ch^2|\log h||v| \end{aligned}$$

for all $u, w \in \mathcal{U}_{ad}$ and all $v \in \mathbb{R}^m$. The first inequality expresses the Lipschitz continuity of $u \mapsto y_u$, while the second and third one follow from Lemma 4. The estimates for G and G' are deduced now immediately from the estimates above. For G'' , the technique is analogous to the one for G' , since $\|z_{v_1 v_2}(w) - z_{h,v_1 v_2}(w)\|_{C(\bar{\Omega}_0)} \leq Ch^2|\log h||v_1||v_2|$ follows from equation (2.2) that has the same structure as (2.4). \square

Lemma 6. *Under the assumptions of Lemma 5, the estimate*

$$|f(u) - f_h(u)| + |f'(u) - f'_h(u)| + \|f''(u) - f''_h(u)\| \leq Ch^2 \quad \forall u \in \mathcal{U}_{ad}$$

is satisfied with some constant C not depending on h and u .

Proof. The function L is Lipschitz on bounded sets and, in view of Lemma 1 and (3.3), $\|y_u\|_{C(\bar{\Omega})}$ and $\|y_{h,u}\|_{C(\bar{\Omega})}$ are uniformly bounded for all $u \in \mathcal{U}_{ad}$ and all h . Therefore, we find

$$|f_h(u) - f(u)| \leq \int_{\Omega} |L(x, y_{h,u}, u) - L(x, y_u, u)| dx \leq c \|y_{h,u} - y_u\|_{L^2(\Omega)} \leq Ch^2$$

by (3.2). Similarly, we can estimate the second part

$$\begin{aligned} (f'(u) - f'_h(u))v &= \int_{\Omega} \left(\frac{\partial L}{\partial y}(x, y_u, u)z_v(u) - \frac{\partial L}{\partial y}(x, y_{h,u}, u)z_{h,v}(u) \right) dx \\ &\quad + \int_{\Omega} \left(\frac{\partial L}{\partial u}(x, y, u)v - \frac{\partial L}{\partial u}(x, y_{h,u}, u)v \right) dx \end{aligned}$$

by (3.2) and the associated L^2 -version of (3.11),

$$\|z_v(u) - z_{v,h}(u)\|_{L^2(\Omega)} \leq Ch^2|v|$$

for all unit vectors v . As in the last proof, the estimation of $\|f''(u) - f''_h(u)\|$ is analogous to that of $|f'(u) - f'_h(u)|$. \square

Let us now state the main result of our paper, an error estimate for optimal controls of the discretized optimal control problem (OCP_h) . This estimate follows from our error analysis for nonlinear mathematical programming problems in Section 5 as a corollary of Theorem 8 applied to (NP) and (NP_h) . The essential assumptions (LICQ) and (5.24) are formulated later in Section 5, but we found it reasonable to present our main result already here.

Theorem 4. *Assume that the assumptions (A.1)-(A.4) are fulfilled and let \bar{u} be a locally optimal control of problem (OCP) satisfying the linear independence condition (LICQ) of Definition 3 along with the strong second-order condition (5.24). Then \bar{u} is locally unique and there exists a sequence \bar{u}_h of locally optimal controls of the corresponding FEM-approximated problem (OCP_h) and a constant $C > 0$ independent of h such that the estimate*

$$|\bar{u} - \bar{u}_h| \leq C h^2 |\log h|. \tag{3.13}$$

is satisfied for all sufficiently small h .

Proof. The optimal control problems (OCP) and (OCP_h) are equivalent to the nonlinear programming problems (NP) and (NP_h) , which are particular cases of the nonlinear programming problems (P) and (P_h) , which will be discussed in Section 5. Due to the previous error estimates, assumption (A.5) is fulfilled with $\alpha(h) = h^2 |\log h|$. Therefore, Theorem 8 is applicable to (NP) and (NP_h) and Theorem 4 is a direct conclusion. \square

Remark 2. *The same estimate holds true for the vectors of Lagrange multipliers \bar{v} and \bar{v}_h associated with \bar{u} and \bar{u}_h , i.e. $|\bar{v} - \bar{v}_h| \leq C h^2 |\log h|$. These multipliers are associated with the state and control constraints, cf. Theorem 8.*

4. NUMERICAL EXAMPLES

In this section we present numerical test examples for which we confirm error estimates of an order close to $\alpha(h) = h^2$. This does not contradict our theory, since the $|\log h|$ -term can hardly be detected numerically. Notice that $h^2 \log h \leq c h^{2-\varepsilon}$ holds for all $\varepsilon > 0$.

In the test runs, regular triangulations of the domain were generated by the Matlab pde-toolbox command `initmesh` and successively refined using the command `refmesh`. The discretized nonlinear optimal control problems were solved by an SQP method where the associated nonlinear programming problems were treated on using the *OOQP*¹ solver. The right-hand side of the state equation was assembled in an exact fashion, i.e. we do not approximate the integral term involving functions e_i , $i = 1, \dots, 3$, because our theory does not include approximation errors for integrals. Therefore, the corresponding right-hand side variational terms $\int_{\Omega} e_i \eta_k dx$, where η_k , $k = 1, \dots, n_h$ are the piecewise linear basis

¹OOQP: Object Oriented Software for Quadratic Programming, Mike Gertz, Steve Wright. <http://pages.cs.wisc.edu/swright/ooqp/>. OOQP is copyrighted to the University of Chicago.

functions of Y_h , were assembled computing the corresponding primitive over its limits in each triangle. The experimental error of convergence is computed by

$$EOC = \frac{\log(|u - u_{h_1}|) - \log(|u - u_{h_2}|)}{\log(h_1) - \log(h_2)},$$

for two consecutive mesh sizes h_1 and h_2 .

Example 1. *This is a slight modification of an example in [24], where the optimal state is active in one single point. The semilinear state equation is considered in the open unit disk $\Omega = B_1(0)$.*

$$(E1) \quad \begin{cases} \min_{u \in \mathbb{R}^3} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2} |u - u_d|^2 \\ \text{subject to} \\ -\Delta y(x) + y(x) + y(x)^3 = \sum_{i=1}^3 u_i e_i(x) & \text{in } \Omega \\ y(x) = 0 & \text{on } \Gamma, \\ y(0, 0) \leq 1, \end{cases} \quad (4.1)$$

where $e_1(x) = 1$, $e_2(x) = x_1^2 + x_2^2$, $e_3(x) = (1 - x_1^2 - x_2^2)^3$. In this case, we construct the optimal functions

$$\bar{y} = 1 - (x_1^2 + x_2^2) \quad \text{and} \quad \bar{u} = [5, -1, 1]^\top.$$

The active set consists of one single point $x = (0, 0)$. The Lagrange multiplier $\mu = \delta_0$ satisfies the complementarity condition, where δ_0 is the Dirac measure concentrated at the origin. If we define $\bar{\varphi} = -\frac{1}{2\pi} \log |x|$, then we have that

$$-\Delta \bar{\varphi} + \bar{\varphi} + 3\bar{y}^2 \bar{\varphi} = \delta_0 - \frac{1}{2\pi} \log |x| (1 + 3(1 - |x|^2)^2),$$

and it is easy to confirm that the right-hand side of this identity is equal to $\bar{y} - y_d + \bar{\mu}$, if we set

$$y_d = 1 - |x|^2 + \frac{1}{2\pi} \log |x| (1 + 3(1 - |x|^2)^2).$$

The gradient equation for \bar{u} is satisfied, if we define

$$u_d = [5, -1, 1]^\top + \left[\frac{1}{4}, \frac{1}{16}, \frac{25}{192} \right]^\top.$$

The numerically computed optimal state and the desired state are shown in Figure 1. Notice that the state is active in the point $(0, 0)$. In Figure 2, the computed Lagrange multiplier and the computed adjoint state are presented. The multiplier has a peak shape concentrated at $(0, 0)$. This reflects the approximation of the Dirac measure δ_0 , which is the exact Lagrange multiplier.

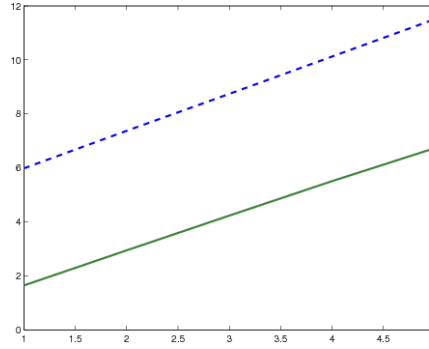


FIGURE 1. Example 1; $-\log(h)$ versus $-\log(|\bar{u} - \bar{u}_h|)$ (solid line) compared with $-2\log(h)$ (dashed line)

h	$ \bar{u} - \bar{u}_h $	EOC
0.08192427	0.96438176	
0.05016470	0.19273740	3.28
0.02508235	0.05252443	1.88
0.01263883	0.01450992	1.86
0.00633700	0.00405185	1.85
0.00317013	0.00119422	1.76

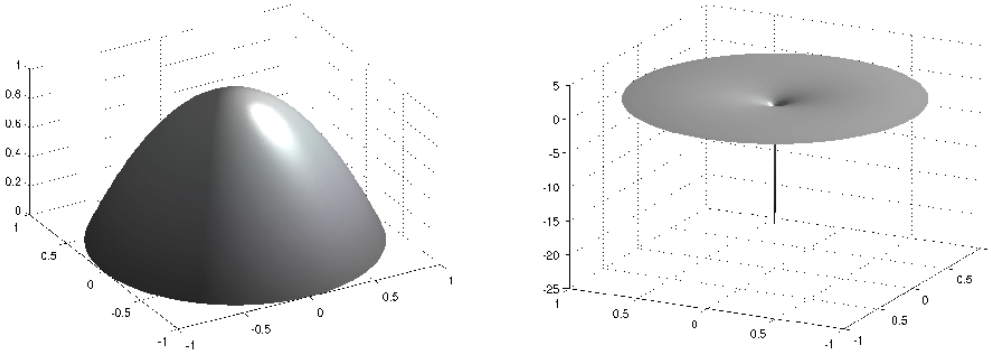


FIGURE 2. Example 1; Numerically computed optimal state \bar{y}_h and shift y_d for $h \approx 0.0006$

Example 2. Here, the state equation contains a the nonlinearity of the type y^2 . Since y^2 is not monotone, it is replaced by the monotone function $y|y|$ for formal reasons. In the numerical computations, this does not matter, because the states turned out to be non-negative. We require that the optimal state \bar{y} is active at five given points inside the

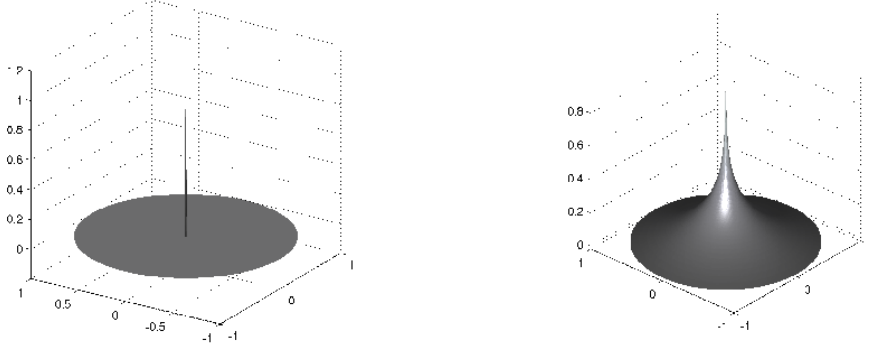


FIGURE 3. Example 1; Numerically computed Lagrange multiplier $\bar{\mu}_h$ and adjoint state $\bar{\varphi}_h$ for $h \approx 0.0006$

domain $\Omega = (-1, 1) \times (-1, 1)$. The problem reads

$$(\mathbf{E2}) \quad \left\{ \begin{array}{l} \min_{u \in \mathbb{R}^5} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2} |u - u_d|^2 \\ \text{subject to} \\ -\Delta y(x) + y(x)(15 + |y(x)|) = \sum_{i=1}^5 u_i e_i(x) \quad \text{in } \Omega \\ y(x) = 0 \quad \text{on } \Gamma, \\ y(x_i) \leq 8/27, \quad i = 1, \dots, 4, \quad \text{and} \quad y(x_5) \geq 0, \\ \text{with} \\ x_1 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), x_2 = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), x_3 = (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}), x_4 = (-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}), x_5 = (0, 0), \end{array} \right. \quad (4.2)$$

and

$$\begin{aligned} e_1(x) &= 12x_1^2 x_2^2 - 2(x_1^4 + x_2^4), & e_2(x) &= x_1^2 + x_2^2, & e_3(x) &= 1, \\ e_4(x) &= (x_1^2 - 1)(x_2^2 - 1)(x_1^2 + x_2^2), & e_5(x) &= (x_1^2 - 1)^2 (x_2^2 - 1)^2 (x_1^2 + x_2^2)^2. \end{aligned}$$

We define

$$\bar{y} = (x_1^2 - 1)(x_2^2 - 1)(x_1^2 + x_2^2), \quad \bar{u} = [-2, 16, -4, 15, 1]^\top$$

so that \bar{y} is the state associated with \bar{u} . Moreover, we set $y_d := \bar{y}$, $u_d := \bar{u}$ so that \bar{y} and \bar{u} are the optimal state and the optimal control, respectively. Obviously, the associated adjoint state is $\bar{\varphi} = 0$. Consider the set

$$\left\{ \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right), \left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right), (0, 0) \right\},$$

where the function \bar{y} attains its maximal value $\frac{8}{27}$ in the first four points, while \bar{y} reaches its minimum value 0 at the point (0, 0). In view of this choice, \bar{y} is active at the five points,

if the state constraints are chosen as above. Therefore, \bar{y} and \bar{u} are also the solution for the state-constrained problem. The gradient equation of the optimality system is

$$\bar{u} - u_d + \left[\int_{\Omega} \bar{\varphi} e_1 dx, \dots, \int_{\Omega} \bar{\varphi} e_5 dx \right]^{\top} = 0. \quad (4.3)$$

In the computations, we chose an initial mesh that contains the five points where the state is active. Then all refined meshes automatically contain these 5 points. The computed rates for the error are listed in the following table.

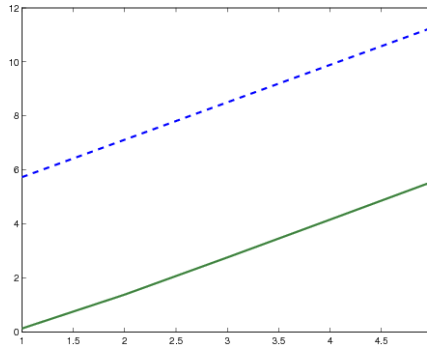


FIGURE 4. Example 2; $-\log(h)$ versus $-\log(|\bar{u} - \bar{u}_h|)$ (solid line) compared with $-2\log(h)$ (dashed line)

h	$ \bar{u} - \bar{u}_h $	EOC
0.057005028172390	0.880415806738625	-
0.028502514086195	0.251486866293358	1.81
0.014251257043098	0.063324994190674	1.99
0.007125628521549	0.015606473275112	2.02
0.003562814260775	0.003822447895272	2.03

Figure 5 shows the numerically computed optimal state, and the Lagrange multiplier is represented in Figure 6. It exhibits peaks in the points where the state is active, this is what we expected from the right-hand side in the adjoint equation (5.5) where Dirac measures are concentrated at the active points.

5. ERROR ESTIMATES FOR A PERTURBED NONLINEAR PROGRAMMING PROBLEM

5.1. The programming problem and its perturbation. In the preceding section, we have transformed our elliptic optimal control problem into the nonlinear programming problem (NP). Associated with (NP), a discretized control problem was constructed that is equivalent to the programming problem (NP_h) . We are interested in estimating the difference between local solutions of (NP) and associated solutions of (NP_h) .

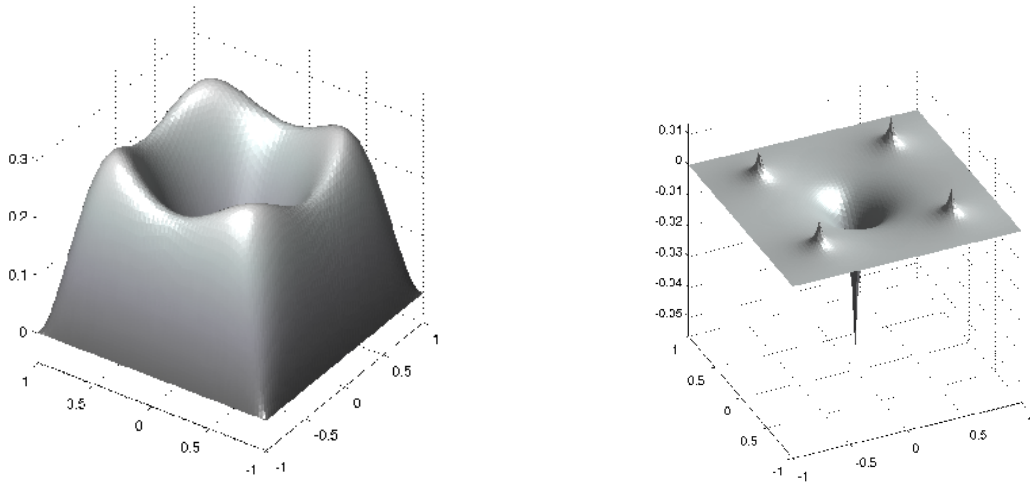


FIGURE 5. Example 2; Numerically computed optimal state, and adjoint state at $h \approx 0.005$

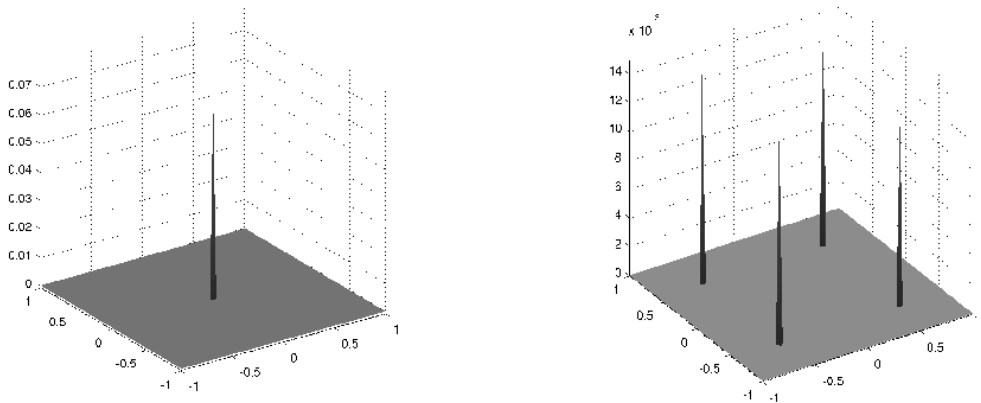


FIGURE 6. Example 2; Numerically computed Lagrange multipliers at $h \approx 0.005$

This question leads to the sensitivity analysis of nonlinear programming problems with respect to perturbations. Although this issue was the subject of various papers, in most of the references we found, the dependence of the given functions with respect to h is assumed to be Lipschitz or even smoother, because Lipschitz stability or differentiability of the solutions with respect to h was the main interest.

In our case, h is a mesh parameter, and the geometry of the grid depends on the distance and on the number of mesh points. The only property we have at our disposal is the continuity of the data at $h = 0$ with some associated rate. Results contained in

Bonnans and Shapiro [4], Klatte und Kummer [19], Malanowski [21] or in many of the references cited therein, cannot be directly used, although their techniques need only a slight modification to be applied to our setting. Similarly, the results presented in Alt [2], Klatte [18] or Malanowski et al. [22], would need some adaptation and the verification of the assumptions stated therein. These papers only assume a rate of continuity w.r. to h at $h = 0$.

Therefore, although our arguments are close to the ones of [2], [18] or [22], we decided to work out the error analysis in a fairly selfcontained way for the convenience of the reader. We underline that our perturbation analysis is inspired by the cited references and is based on standard conclusions. In particular, we were partially guided by the work [2]. The error analysis is discussed under assumptions, which are satisfied by our problems (NP) and (NP_h) related to (OCP) and (OCP_h) .

We discuss the following standard class of nonlinear programming problems that is not necessarily related to (OCP) or (OCP_h) , although we use the same notation. We consider

$$(P) \quad \begin{cases} \min f(u) \\ G_i(u) = 0, & i = 1, \dots, k, \\ G_i(u) \leq 0, & i = k + 1, \dots, \ell, \\ u \in \mathcal{U}_{ad} \end{cases}$$

and its perturbed version

$$(P_h) \quad \begin{cases} \min f_h(u) \\ G_{h,i}(u) = 0, & i = 1, \dots, k, \\ G_{h,i}(u) \leq 0, & i = k + 1, \dots, \ell, \\ u \in \mathcal{U}_{ad}, \end{cases}$$

where $\mathcal{U}_{ad} \in \mathbb{R}^m$ is the convex and bounded box-shaped set introduced in Section 1, and $f, f_h, G_i, G_{h,i} : \mathbb{R}^m \rightarrow \mathbb{R}, i = 1, \dots, \ell, h > 0$ are functions of $C^{2,1}(\mathcal{U}_{ad})$.

We assume that the following approximation assumption is satisfied:

(A.5) There exist a constant $C > 0$ and a function $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha(h) \rightarrow 0$ as $h \rightarrow 0$ such that

$$\begin{aligned} & |f(u) - f_h(w)| + |f'(u) - f'_h(w)| + \|f''(u) - f''_h(w)\| + \sum_{i=1}^{\ell} (|G_i(u) - G_{h,i}(w)| + \\ & + |G'_i(u) - G'_{h,i}(w)| + \|G''_i(u) - G''_{h,i}(w)\|) \leq C (|u - w| + \alpha(h)) \end{aligned}$$

holds for all $h > 0$ and all $u, w \in \mathcal{U}_{ad}$.

Notice that **(A.5)** is satisfied for the problems (NP) and (NP_h) by $\alpha(h) = h^2 |\log h|$.

To cover the given equality and inequality constraints in a unified way, we introduce the cone K of "nonnegative" vectors by

$$K = \{z \in \mathbb{R}^{\ell} : z_i = 0, i = 1, \dots, k, z_i \geq 0, i = k + 1, \dots, \ell\}.$$

Then the constraints $G_i(u) = 0$, $i = 1, \dots, k$, $G_i(u) \leq 0$, $i = k + 1, \dots, \ell$, can be expressed by $G(u) \leq_K 0$, where $z \leq_K 0 \Leftrightarrow -z \in K$. Moreover, we introduce $G(u) = [G_1(u), \dots, G_\ell(u)]^\top$ and define G_h analogously. By these definitions, (P) and (P_h) can be written as

$$(P) \quad \begin{cases} \min f(u) \\ G(u) \leq_K 0, \end{cases} \quad u \in \mathcal{U}_{ad} \quad \text{and} \quad (P_h) \quad \begin{cases} \min f_h(u) \\ G_h(u) \leq_K 0, \end{cases} \quad u \in \mathcal{U}_{ad}.$$

5.2. First- and second-order optimality conditions.

5.2.1. *First-order necessary conditions.* Let us assume once and for all that \bar{u} is a locally optimal solution to (P). According to Robinson [28], we require the following regularity condition at \bar{u} :

$$0 \in \text{int} \{G(\bar{u}) + G'(\bar{u})(\mathcal{U}_{ad} - \bar{u}) + K\}, \quad (5.1)$$

where the set in braces is defined as $\cup\{G(\bar{u}) + G'(\bar{u})(u - \bar{u}) + k \mid u \in \mathcal{U}_{ad}, k \in K\}$. A vector \bar{u} that satisfies all constraints of (P) and the condition (5.1) is said to be *regular*.

Definition 1. *The Lagrange function $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ associated with (P) is defined by*

$$\mathcal{L}(u, \mu) = f(u) + \sum_{i=1}^{\ell} \mu_i G_i(u). \quad (5.2)$$

The well known first-order conditions for (P) are stated in the following theorem.

Theorem 5. *Suppose that \bar{u} is locally optimal for (P) and regular. Then there exists a vector $\bar{\mu} \in \mathbb{R}^\ell$ with $\mu_i \geq 0$, $i = k + 1, \dots, \ell$, such that the variational inequality*

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})(u - \bar{u}) \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (5.3)$$

is satisfied together with the complementary slackness conditions

$$\bar{\mu}_i G_i(\bar{u}) = 0 \quad \text{for } i = k + 1, \dots, \ell. \quad (5.4)$$

These relations express the standard necessary first-order conditions of nonlinear programming. The existence of Lagrange multipliers follows from the Robinson regularity condition (5.1). This condition is equivalent to the well known regularity condition by Kurcyusz and Zowe [34] that is sufficient for the existence of Lagrange multipliers (and optimality conditions in qualified form).

Remark 3. *In the particular case of (NP) that is related to the optimal control problem (OCP), the necessary conditions (5.3)–(5.4) can be expressed in terms of optimal control: We define an adjoint state $\varphi \in W_0^{1,\sigma}(\Omega)$, $1 \leq \sigma < n/(n-1)$, as the unique solution of the adjoint equation*

$$\begin{cases} A\varphi + \frac{\partial d}{\partial y}(x, \bar{y}, \bar{u})\varphi = \frac{\partial L}{\partial y}(x, \bar{y}, \bar{u}) + \sum_{i=1}^{\ell} \bar{\mu}_i g'_i(\bar{y}(x_i))\delta_{x_i} & \text{in } \Omega \\ \varphi = 0 & \text{on } \Gamma. \end{cases} \quad (5.5)$$

Then the variational inequality (5.3) can be written as

$$d^\top(u - \bar{u}) \geq 0 \quad \forall u \in \mathcal{U}_{ad},$$

where $d \in \mathbb{R}^m$ is defined by $d = (d_i)$ with

$$d_i := \int_{\Omega} \left(\frac{\partial L}{\partial u_i}(x, \bar{y}, \bar{u}) - \varphi(x) \frac{\partial d}{\partial u_i}(x, \bar{y}(x), \bar{u}) \right) dx, \quad i = 1, \dots, m. \quad (5.6)$$

5.2.2. Second-order sufficient conditions. Let us also briefly recall the known second-order conditions. We assume that \bar{u} satisfies the first-order necessary optimality conditions (5.3) and (5.4) together with the constraints of (P).

To formulate the second-order conditions, we need the critical cone $C_{\bar{u}}$. For this purpose we define the index sets associated with the inequality constraints that are included in $G(u) \leq_K 0$ or are defining \mathcal{U}_{ad} :

$$\begin{aligned} I_0 &:= \{i \in \{k+1, \dots, \ell\} : G_i(\bar{u}) = 0\} \\ I_+ &:= \{i \in \{k+1, \dots, \ell\} : \bar{\mu}_i > 0\} \\ I_{0,a} &:= \{i \in \{1, \dots, k\} : \bar{u}_i = u_{ai}\} \\ I_{+,a} &:= \{i \in \{1, \dots, k\} : d_i > 0\} \\ I_{0,b} &:= \{i \in \{1, \dots, k\} : \bar{u}_i = u_{bi}\} \\ I_{+,b} &:= \{i \in \{1, \dots, k\} : d_i < 0\}, \end{aligned}$$

where $d_i := \nabla_u \mathcal{L}(\bar{u}, \bar{\mu})_i$. It is a standard conclusion from the variational inequality (5.3) that $\bar{u}_i = u_{ai}$ must hold for those i with $d_i > 0$ and $\bar{u}_i = u_{bi}$ for the i with $d_i < 0$. The set $I_{+,a} \cup I_{+,b}$ is often said to be the set of strongly active control constraints.

Definition 2. The critical cone $C_{\bar{u}}$ is the set of all vectors $u \in \mathbb{R}^m$ that satisfy the following relations:

$$\begin{aligned} G'_i(\bar{u})u &= 0 & \text{for all } i &\in \{1, \dots, k\} \cup I_+ \\ G'_i(\bar{u})u &\leq 0 & \text{for all } i &\in I_0 \setminus I_+ \\ u_i &= 0 & \text{for all } i &\in I_{+,a} \cup I_{+,b} \\ u_i &\geq 0 & \text{for all } i &\in I_{0,a} \setminus I_{+,a} \\ u_i &\leq 0 & \text{for all } i &\in I_{0,b} \setminus I_{+,b}. \end{aligned}$$

Remark 4. In the case of problem (NP), it holds that $G'_i(\bar{u})u = g'_i(\bar{y}(x_i))z(x_i)$ for all $i = 1, \dots, \ell$, where z solves the linearized equation (2.2) with $v = u$, so that all derivatives above can be expressed in terms of the partial differential equation.

Following the theory of second-order conditions in finite-dimensional spaces, cf. Bonnans and Shapiro [4], Luenberger [20] or Fiacco and McCormick [14], we require the following condition:

(A.6) The coercivity property

$$u^\top \mathcal{L}''(\bar{u}, \bar{\mu})u > 0$$

is fulfilled for all $u \in C_{\bar{u}} \setminus \{0\}$.

Here and in the following, $\mathcal{L}''(u, \mu)$ denotes the Hessian matrix of \mathcal{L} at (u, μ) with respect to u .

Theorem 6. *Let \bar{u} satisfy the constraints of (P), the first-order necessary conditions (5.3) – (5.4), and the second-order sufficient condition (A.6). Then there exist real numbers $\omega > 0$ and $\varepsilon > 0$ such that*

$$f(u) - f(\bar{u}) \geq \omega |u - \bar{u}|^2 \quad (5.7)$$

holds for all admissible u with $|u - \bar{u}| \leq \varepsilon$.

This is a standard result of nonlinear programming theory, cf. [14] or [20].

To perform our perturbation analysis, we also need the well known linear independence condition. We extend the vector $G(u)$ to $\hat{G}(u) \in \mathbb{R}^{\ell+2m}$ by

$$\begin{aligned} G_{\ell+i}(u) &= u_{a,i} - u_i, & \text{for } i = 1, \dots, m \\ G_{\ell+m+i}(u) &= u_i - u_{b,i}, & \text{for } i = 1, \dots, m. \end{aligned}$$

Definition 3. *The index set $\mathcal{A}(\bar{u})$ of active constraints at \bar{u} is defined by*

$$\mathcal{A}(\bar{u}) = \{i \in \{1, \dots, \ell + 2m\} : G_i(\bar{u}) = 0\}.$$

We say that the linear independence constraint qualification (LICQ) holds at \bar{u} if the vectors $\nabla G_i(\bar{u})$ of the set of active gradients $\{\nabla G_i(\bar{u}) \mid i \in \mathcal{A}(\bar{u})\}$ are linearly independent.

Remark 5. *For all $i > \ell$, the gradients $\nabla G_i(\bar{u})$ are unit vectors with respect to the standard basis multiplied by 1 or -1 . Therefore, it is easy to see that the associated active gradients form a system of linearly independent vectors. In view of this, the condition (LICQ) amounts to the requirement that set of vectors defined by the first ℓ components of the active gradients $\nabla G(\bar{u})_i$, $i \leq \ell$ is linearly independent.*

5.3. Perturbation analysis. To set up the first order-necessary conditions for (P_h) , we introduce the associated Lagrange function $\mathcal{L}_h : \mathbb{R}^{m+\ell} \rightarrow \mathbb{R}$ by $\mathcal{L}_h(u, \mu) = f_h(u) + G_h(u)^\top \mu$. Then the first-order conditions can be formulated as follows:

If \bar{u}_h is locally optimal for (P_h) and the Robinson regularity condition (5.1) is satisfied at \bar{u} with $G_h(\bar{u}_h)$ and $G'_h(\bar{u}_h)$ substituted for $G(\bar{u})$ and $G'(\bar{u})$, respectively, then there exists a Lagrange multiplier $\bar{\mu}_h \in \mathbb{R}^\ell$ with $\bar{\mu}_{h,i} \geq 0$, $i = k + 1, \dots, \ell$, such that

$$\frac{\partial \mathcal{L}_h}{\partial u}(\bar{u}_h, \bar{\mu}_h)(u - \bar{u}_h) \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (5.8)$$

and

$$\bar{\mu}_{h,i} G_{h,i}(\bar{u}) = 0 \quad \text{for all } i = 1, \dots, \ell. \quad (5.9)$$

Lemma 7. *If \bar{u} satisfies the Robinson regularity condition for (P), then this condition is also fulfilled for all u with $|u - \bar{u}| \leq \rho$ and all sufficiently small $h > 0$, if ρ is taken sufficiently small.*

Proof. Obviously, the Robinson regularity condition holds if and only if a small cuboid with the corners c_1, \dots, c_n and the origin in its interior is covered by $G(\bar{u}) + G'(\bar{u})(\mathcal{U}_{ad} - \bar{u}) + K$. We have

$$c_i = G(\bar{u}) + G'(\bar{u})(u_i - \bar{u}) + k_i, \quad i = 1, \dots, n,$$

with $u_i \in \mathcal{U}_{ad}$ and $k_i \in K$. The cuboid is the convex hull of the corners c_i . We consider now the points

$$c_{h,i} = G_h(u) + G'_h(u)(u_i - u) + k_i, \quad i = 1, \dots, n.$$

Let $\varepsilon > 0$ be given. Then for sufficiently small ρ and h , we have $\sum_{i=1}^n |c_i - c_{h,i}| < \varepsilon$. If ε is small enough, then the convex hull $\text{co}\{c_{h,1}, \dots, c_{h,n}\}$ contains the origin in its interior. Therefore,

$$\begin{aligned} 0 \in \text{int } \text{co}\{c_{h,1}, \dots, c_{h,n}\} &= G_h(u) + G'_h(u)(\text{co}\{u_1, \dots, u_n\} - u) + \text{co}\{k_1, \dots, k_n\} \\ &\subset G_h(u) + G'_h(u)(\mathcal{U}_{ad} - u) + K \end{aligned}$$

so that the Robinson constraint qualification is satisfied. \square

In view of this Lemma, to each local solution \bar{u}_h of (P_h) sufficiently close to \bar{u} , there exists a Lagrange multiplier, if h is small enough. However, we have to show the existence of a locally optimal \bar{u}_h in a neighborhood of \bar{u} and we want to get an estimate for the difference. We proceed as follows: In a preparatory step, via Lemma 7 – Lemma 9, we show the existence of \bar{u}_h and the estimate (5.19) of the order $\sqrt{\alpha(h)}$. This step is inspired by techniques used in Alt [2] and Allgower et al. [1]. Next, we invoke the stability theory for generalized equations to derive an estimate of the order $\alpha(h)$.

First, we approximate admissible vectors for (P) by admissible ones for (P_h) having a distance bounded by $c\alpha(h)$ and vice versa.

Lemma 8. *Suppose that \bar{u} is feasible for (P) and satisfies the Robinson regularity condition (5.1). Then, under the assumption **(A.5)**, there are positive constants C and h_0 such that, for each $h \in (0, h_0)$ an admissible u_h for problem (P_h) exists that satisfies the estimate*

$$|\bar{u} - u_h| \leq C\alpha(h). \quad (5.10)$$

Proof. To be consistent with the notation of [28], we re-write $G, G_h : \mathbb{R}_+ \times \mathbb{R}^m \rightarrow \mathbb{R}^\ell$ by

$$\mathcal{G}(h, u) = \begin{cases} G_h(u), & \text{if } u \in \mathcal{U}_{ad} \text{ and } h > 0, \\ G(u), & \text{if } u \in \mathcal{U}_{ad} \text{ and } h = 0, \\ \emptyset, & \text{if } u \notin \mathcal{U}_{ad}. \end{cases}$$

Thanks to Lemma 5, \mathcal{G} and $\partial\mathcal{G}/\partial u$ are continuous at the point $(h, u) = (0, \bar{u})$. Moreover, we have $\mathcal{G}(0, \bar{u}) \leq_K 0$. In view of the Robinson constraint qualification (5.1), the assumptions of the generalized implicit function in [28] are fulfilled. We obtain the existence of neighborhoods \mathcal{N} of $h = 0$ and \mathcal{O} of \bar{u} such that, for all $h \in \mathcal{N}$, the inequality $\mathcal{G}(h, u) \leq_K 0$ has a solution $u \in \mathcal{O}$, and it holds

$$\text{dist}[v, \Sigma(h)] \leq c|\mathcal{G}(h, v)_+|, \quad \forall h \in \mathcal{N}, \forall v \in \mathcal{O}, \quad (5.11)$$

where $\Sigma(h) = \{u \in \mathcal{U}_{ad} \mid \mathcal{G}(h, u) \leq_K 0\}$ is the solution set of the inequality and $dist$ denotes the Euclidean distance of a point to a set. The value $|\mathcal{G}(h, v)_+|$ is the distance of the set $\mathcal{G}(h, v) + K$ to the origin and measures the residual of v with respect to the inequality $\mathcal{G}(h, v) \leq_K 0$, cf. [28], p. 498. Inserting $v = \bar{u}$ in (5.11), we deduce

$$dist[\bar{u}, \Sigma(h)] \leq c |\mathcal{G}(h, \bar{u})_+| \leq c(|\mathcal{G}(0, \bar{u})_+| + |\mathcal{G}(h, \bar{u})_+ - \mathcal{G}(0, \bar{u})_+|) \leq 0 + c\alpha(h).$$

Therefore, there exists $u_h \in \Sigma(h)$ with $|\bar{u} - u_h| \leq c\alpha(h)$. \square

With the last results at hand we are able to prove the following

Lemma 9. *Let the reference solution \bar{u} satisfy the linear independence condition (LICQ). Then, for all given $\rho > 0$, the auxiliary problem*

$$(P_{h,\rho}) \quad \begin{cases} \min f_h(u) \\ G_h(u) \leq_K 0, \\ u \in \mathcal{U}_{ad} \cap cl B(\bar{u}, \rho) \end{cases} \quad (5.12)$$

is solvable for all sufficiently small h . If \bar{u}_h is any optimal solution to this problem, then there exists an element $v_h \in \mathcal{U}_{ad}$ that is admissible for (P) and satisfies the estimate

$$|\bar{u}_h - v_h| \leq C\alpha(h) \quad (5.13)$$

with some constant $C > 0$ that is independent of h .

Proof. (i) *Solvability of $(P_{h,\rho})$:* For a positive h_0 and all $h \in (0, h_0)$, the admissible set of $(P_{h,\rho})$ is not empty, because u_h constructed in Lemma 8 satisfies all constraints. Therefore, the existence of an optimal \bar{u}_h is an immediate conclusion. We have to find a v_h in \mathcal{U}_{ad} with $G(v_h) \leq_K 0$ and $|v_h - \bar{u}_h| \leq c\alpha(h)$.

For the following, it is more convenient to include also explicitly the inequality constraints defining \mathcal{U}_{ad} . Hence we have to construct v_h such that

$$G_i(v_h) = 0, \quad i = 1, \dots, k, \quad G_i(v_h) \leq 0, \quad i = k + 1, \dots, l + 2m.$$

(ii) *Construction of an equation for v_h :* Notice that $\bar{u}_h \in cl B(\bar{u}, \rho)$ for all $h \leq h_0$. Therefore, if ρ is taken small enough, all inactive components $G_i(\bar{u})$ are inactive for \bar{u}_h as well and there exists $\varepsilon > 0$ such that

$$G_i(\bar{u}_h) \leq -\varepsilon < 0 \quad \forall i \in I, \quad \forall h \leq h_0, \quad (5.14)$$

where I is the set of all inactive indices of \bar{u} in $\{k + 1, \dots, l + 2m\}$.

Suppose that r constraints are active at \bar{u} , $k \leq r \leq m$. After renumbering, if necessary, we can assume that those with the numbers $1 \leq i \leq r$ are active, hence $G_1(\bar{u}) = \dots = G_r(\bar{u}) = 0$. By (LICQ), the associated gradients $\nabla G_i(\bar{u})$ are linearly independent. If ρ is small enough, also $\nabla G_1(\bar{u}_h), \dots, \nabla G_r(\bar{u}_h)$ are linearly independent. Consider the matrix

$$B_h = [\nabla G_1(\bar{u}_h), \dots, \nabla G_r(\bar{u}_h)]^\top.$$

Since B_h has full rank r , we find an invertible submatrix D_h such that (after renumbering of the components of u , if necessary) $B_h = [D_h, E_h]$ holds with some other matrix E_h . We

define $F_h : \mathbb{R}^r \rightarrow \mathbb{R}^r$ by

$$F_{h,i}(w) := G_i(w, \bar{u}_{h,r+1}, \dots, \bar{u}_{h,m}) - G_{h,i}(\bar{u}_h), \quad i = 1, \dots, r.$$

To find v_h , we fix its $m-r$ last components by $v_{h,i} := \bar{u}_{h,i}$, $i = r+1, \dots, m$, and determine the first r components as the solution w of the system

$$F_h(w) = 0, \tag{5.15}$$

i.e. we set $v_{h,i} := w_i$, $i = 1, \dots, r$.

(iii) *Solvability of (5.15)*: In this part of the proof, we follow a technique used by Allgöwer et al. [1]. We define for convenience $\bar{w}_h := (\bar{u}_{h,1}, \dots, \bar{u}_{h,r})^\top$, $\bar{w} := (\bar{u}_1, \dots, \bar{u}_r)^\top$ and have

$$|F_h(\bar{w}_h)| \leq c\alpha(h), \tag{5.16}$$

since $|G_i(\bar{u}_h) - G_{h,i}(\bar{u}_h)| \leq c\alpha(h)$ holds for all $1 \leq i \leq r$.

Thanks to Lemma 5, there exist constants $\gamma > 0$, $\beta > 0$ such that

$$\|F'_h(w_1) - F'_h(w_2)\| \leq \gamma |w_1 - w_2| \quad \forall w_i \in B(\bar{w}, \rho)$$

and

$$\|(F'_h(w))^{-1}\| \leq \beta \quad \forall w \in B(\bar{w}, \rho)$$

holds for all $0 \leq h \leq h_0$, if ρ is taken sufficiently small. Notice that $\partial\mathcal{G}(w)/\partial w$ is then close to $\partial\mathcal{G}(\bar{w})/\partial w$, and this matrix is invertible. We define $\eta > 0$ by

$$\eta := \beta |F_h(\bar{w}_h)|.$$

Then (5.16) implies $\beta\gamma\eta/2 \leq 1$ for all $0 < h < h_0$, if h_0 is sufficiently small. Proceeding as in [1], the Mysovskij theorem, cf. Ortega and Rheinboldt [25], p. 412, ensures that the Newton method starting at $w_0 := \bar{w}_h$ generates a solution w of (5.15) in the ball $cl B(\bar{w}_h, c_0\eta)$, where c_0 is a certain constant. It follows from our construction that

$$G_i(v_h) = G_{h,i}(\bar{u}_h) \begin{cases} = 0, & i = 1, \dots, k, \\ \leq 0, & i = k+1, \dots, r. \end{cases}$$

Moreover, if h is small, $G_i(v_h) < 0$ holds for $r < i \leq \ell+2m$ (according to our renumbering). Therefore, we have that $G(v_h) \leq_K 0$ and $v_h \in \mathcal{U}_{ad}$. From $w \in cl B(\bar{w}_h, c_0\eta)$ it follows $|w - \bar{w}_h| \leq c_0\eta \leq c\alpha(h)$, hence also $|v_h - \bar{u}_h| \leq c\alpha(h)$. \square

Lemma 10. *If $\rho > 0$ is taken sufficiently small, $h \in (0, h_0(\rho))$, and \bar{u} is a local solution of (P) satisfying (LICQ) and the second-order condition **(A.6)**, then all solutions \bar{u}_h of the auxiliary problem $(P_{h,\rho})$ belong to $B(\bar{u}, \rho)$. Therefore, they are also locally optimal for the problem (P_h) .*

Proof. First, we compare the solution \bar{u}_h of $(P_{h,\rho})$ defined in Lemma 8 with u_h that is admissible for $(P_{h,\rho})$ and approximates \bar{u} with the order $\alpha(h)$. We get

$$f_h(\bar{u}_h) \leq f_h(u_h) \leq |f_h(u_h) - f_h(\bar{u})| + |f_h(\bar{u}) - f(\bar{u})| + f(\bar{u}).$$

By

$$|f_h(\bar{u}) - f(\bar{u})| + |u_h - \bar{u}| + |f_h(\bar{u}_h) - f(\bar{u}_h)| \leq c\alpha(h)$$

and by the uniform Lipschitz property of f_h , we find

$$f(\bar{u}_h) \leq f(\bar{u}) + c_1 \alpha(h). \quad (5.17)$$

Next, we compare \bar{u} with v_h taken from Lemma 9 that is admissible for (P) and approximates \bar{u}_h with the order $\alpha(h)$. From the quadratic growth condition we obtain

$$f(v_h) \geq f(\bar{u}) + \omega |\bar{u} - v_h|^2.$$

Notice that v_h is close enough to \bar{u} , if h is sufficiently small. From $|\bar{u}_h - v_h| \leq c \alpha(h)$ we deduce

$$f(\bar{u}_h) + c_2 \alpha(h) \geq f(\bar{u}) + \omega |\bar{u} - \bar{u}_h|^2. \quad (5.18)$$

Combining the inequalities (5.17)–(5.18), it follows that

$$f(\bar{u}) + c_1 \alpha(h) \geq f(\bar{u}) + \omega |\bar{u} - \bar{u}_h|^2 - c_2 \alpha(h)$$

and hence

$$|\bar{u} - \bar{u}_h| \leq c \sqrt{\alpha(h)}. \quad (5.19)$$

For all sufficiently small h , this estimate implies $|\bar{u} - \bar{u}_h| < \rho$ so that \bar{u}_h does not touch the boundary of $B(\bar{u}, \rho)$. In view of this, \bar{u}_h is locally optimal for (P_h) . \square

The estimate (5.19) is not optimal. We are able to get rid of the square root. Moreover, we would like to show local uniqueness of \bar{u}_h , i.e. uniqueness of local optima of (P_h) in $B(\bar{u}, \rho)$, if ρ is small enough. Both tasks can be accomplished by the stability theory for optimality systems written as generalized equations. Let us introduce the necessary notions.

We define the *dual cone* K_+ associated with K by

$$K_+ = \{z \in \mathbb{R}^\ell \mid z^\top v \geq 0 \quad \forall v \in K\}.$$

It is easy to verify that

$$K_+ = \{z \in \mathbb{R}^\ell \mid z_i \geq 0, \quad i = k+1, \dots, \ell\} = \mathbb{R}^k \times \mathbb{R}_+^{\ell-k}.$$

Moreover, we need the *normal cone* $\partial\psi_E(x)$ to a convex set $E \subset \mathbb{R}^n$ at a point $x \in \mathbb{R}^n$ defined by

$$\partial\psi_E(x) = \begin{cases} z \in \mathbb{R}^n \text{ with } z^\top(e - x) \leq 0 \quad \forall e \in E, & \text{if } x \in E \\ \emptyset, & \text{if } x \notin E. \end{cases}$$

By the dual cone K_+ , the complementarity conditions $G(u) \leq_K 0$, $\mu \in K_+$, and $G(u)^\top \mu = 0$ can be compressed to

$$G(u)^\top(\eta - \mu) \leq 0 \quad \forall \eta \in K_+.$$

In terms of the normal cone, this reads $G(u) \in \partial\psi_{K_+}(\mu)$. Therefore, we can express the first-order necessary conditions (5.3)–(5.4) in form of the generalized equation

$$0 \in F(u, \mu) + \mathcal{T}(u, \mu), \quad (5.20)$$

where F is a mapping in $\mathbb{R}^m \times \mathbb{R}^\ell$ defined by

$$F(u, \mu) = \begin{bmatrix} \nabla_u \mathcal{L}(u, \mu) \\ -G(u) \end{bmatrix},$$

and \mathcal{T} is a set-valued mapping defined by $\mathcal{T}(u, \mu) = \partial\psi_{\mathcal{U}_{ad}}(u) \times \partial\psi_{K_+}(\mu)$. Analogously, the optimality system for problem (P_h) is equivalent to

$$0 \in F_h(u_h, \mu_h) + \mathcal{T}(u_h, \mu_h),$$

where F_h is the mapping

$$F_h(u, \mu) = \begin{bmatrix} \nabla_u \mathcal{L}_h(u, \mu) \\ -G_h(u) \end{bmatrix}.$$

For different reasons, we extend the generalized equation by associating Lagrange multipliers also to the inequality constraints defining \mathcal{U}_{ad} . We recall our definition

$$\hat{G}(u) := (G_i(u))_{i=1}^{\ell+2m}.$$

Moreover, we define $\hat{K} = K \times \mathbb{R}_+^{2m}$ and

$$\hat{\mathcal{L}}(u, \nu) := f(u) + \sum_{i=1}^{\ell+2m} \nu_i G_i(u).$$

Then the optimality system for (P) is equivalent to the conditions

$$\nabla_u \hat{\mathcal{L}}(u, \nu) = 0, \quad \hat{G}(u) \leq_{\hat{K}} 0, \quad \nu \geq_{\hat{K}_+} 0, \quad \hat{G}(u)^\top \nu = 0, \quad (5.21)$$

where ν belongs to $\mathbb{R}^{\ell+2m}$. The first ℓ components of this extended ν are equal to those of μ . Given μ , it is not difficult to find the last $2m$ components of ν . They are obtained by

$$\begin{aligned} \nu_i &= \nabla_u \mathcal{L}(u, \mu)_+, \quad i = \ell + 1, \dots, \ell + m, \\ \nu_i &= \nabla_u \mathcal{L}(u, \mu)_-, \quad i = \ell + m + 1, \dots, \ell + 2m, \end{aligned} \quad (5.22)$$

where $a_+ = (a + |a|)/2$, $a_- = (-a + |a|)/2$. It is easy to verify that the complementarity conditions are satisfied by this choice and that these Lagrange multipliers are unique, if $u_a < u_b$. For this simple fact, we refer, for instance, to [33], Section 2.

Written in terms of a generalized equation, (5.21) reads

$$0 \in \begin{pmatrix} \nabla_u \hat{\mathcal{L}}(u, \nu) \\ -\hat{G}(u) \end{pmatrix} + \begin{pmatrix} \{0\} \\ \partial\psi_{\hat{K}_+}(\nu) \end{pmatrix} =: \hat{\mathcal{F}}(u, \nu) + \hat{\mathcal{T}}(u, \nu). \quad (5.23)$$

In addition to the linear independence condition (LICQ), we require now the following *strong sufficient second order condition*:

(A.7) For the pair $(\bar{u}, \bar{\nu})$, it holds

$$v^\top \frac{\partial^2 \hat{\mathcal{L}}(\bar{u}, \bar{\nu})}{\partial u^2} v > 0 \quad \forall v \in \hat{C}_{\bar{u}}, v \neq 0, \quad (5.24)$$

where

$$\hat{C}_{\bar{u}} = \{v \in \mathbb{R}^m \mid G'_i(\bar{u})v = 0 \ \forall i \in \{1, \dots, k\} \cup \{i \in \{k+1, \dots, 2\ell+m\} : \bar{\nu}_i > 0\}\}.$$

The extended Lagrange multiplier vector $\bar{\nu}$ is defined by $\bar{\mu}$ according to (5.22).

We assume that the reader is familiar with the notion of *strong regularity* of generalized equations, cf. Robinson [29]. It follows from Robinson [29] that the generalized equation (5.23) is strongly regular at $(\bar{u}, \bar{\nu})$, if the linear independence condition (LICQ) and the strong second-order sufficient condition (5.24) are fulfilled for $(\bar{u}, \bar{\nu})$.

For convenience, we recall the Robinson implicit function theorem.

Theorem 7. *Let \mathcal{O} be an open subset of a normed linear space X , P be a topological space, $F : \mathcal{O} \times P \rightarrow X'$ be a mapping, and $C \subset X$ be a convex set. Suppose that the partial Fréchet-derivative $F'(\cdot, \cdot)$ of F with respect to the first variable exists on $\mathcal{O} \times P$, that both $F(\cdot, \cdot)$ and $F'(\cdot, \cdot)$ are continuous at $(x_0, p_0) \in \mathcal{O} \times P$ and that x_0 solves*

$$0 \in F(x, p_0) + \partial\psi_C(x). \quad (5.25)$$

If (5.25) is strongly regular at x_0 with associated Lipschitz constant C_L , then, for any $\varepsilon > 0$, there exist neighborhoods $\partial\psi_\varepsilon$ of p_0 and W_ε of x_0 , and a single valued function $x : \partial\psi_\varepsilon \rightarrow W_\varepsilon$, such that, for any $p \in \partial\psi_\varepsilon$, $x(p)$ is the unique solution in W_ε of the inclusion

$$0 \in F(x, p) + \partial\psi_C(x).$$

Further, for each p and q in $\partial\psi_\varepsilon$ one has

$$\|x(p) - x(q)\| \leq (C_L + \varepsilon)\|F(x(q), p) - F(x(q), q)\|.$$

We are going to apply this implicit function theorem to the generalized equation

$$\delta \in \hat{\mathcal{F}}(u, \nu) + \hat{\mathcal{T}}(u, \nu) \quad (5.26)$$

with given perturbation $\delta \in \mathbb{R}^{\ell+2m}$. We know that the pair $(\bar{u}, \bar{\nu})$ is a solution for $\delta = 0$. If (LICQ) and the strong second-order sufficient condition **(A.7)** are fulfilled, then the property of strong regularity is satisfied, cf. our comments before Theorem 7, and we can apply this implicit function Theorem. In this case, there are positive values r, σ such that, for all $\delta \in B(0, r)$, the generalized equation (5.26) has exactly one solution (u, ν) in $B((\bar{u}, \bar{\nu}), \sigma)$, i.e. we have local uniqueness. Moreover, there is a constant $C > 0$ such that the Lipschitz property

$$|u - \bar{u}| + |\nu - \bar{\nu}| \leq C |\delta| \quad (5.27)$$

is satisfied. By the linearity of the Lagrange function with respect to ν and by (LICQ), the Lagrange multipliers $\bar{\nu}$ and ν associated with \bar{u} and u , respectively, are unique.

Lemma 11. *Assume that the linear independence condition (LICQ) is satisfied at \bar{u} , and that \bar{u}_h is a sequence of local solutions to problem (P_h) converging to \bar{u} as $h \downarrow 0$. Then the Lagrange multipliers $\bar{\nu}_h$ associated with \bar{u}_h are uniformly bounded for all sufficiently small $h > 0$.*

Proof. From $\bar{u}_h \rightarrow \bar{u}$ as $h \rightarrow 0$ we obtain the boundedness of the sequence \bar{u}_h and the information that all inactive constraints at \bar{u} must be inactive at \bar{u}_h for all sufficiently small h . Therefore, the pair (\bar{u}_h, \bar{v}_h) satisfies the equation of necessary conditions

$$\nabla_u \mathcal{L}(\bar{u}_h, \bar{v}_h) = \nabla f_h(\bar{u}_h) + \sum_{i \in \mathcal{A}(\bar{u})} \bar{v}_{h,i} \nabla G_{h,i}(\bar{u}_h) = 0.$$

The result of the Lemma is a simple consequence of the linear independence condition. \square

Next, we consider the solution \bar{u}_h of the approximate problem $(P_{h,\rho})$ according to Lemma 9. With an associated Lagrange multiplier \bar{v}_h , it satisfies the generalized equation

$$0 \in \hat{\mathcal{F}}_h(\bar{u}_h, \bar{v}_h) + \hat{\mathcal{T}}(\bar{u}_h, \bar{v}_h). \quad (5.28)$$

This inclusion is unperturbed in its left-hand side but contains the "approximate" function $\hat{\mathcal{F}}_h$ depending on h . We show that the pair (\bar{u}_h, \bar{v}_h) solves the perturbed generalized equation

$$\delta_h \in \hat{\mathcal{F}}(\bar{u}_h, \bar{v}_h) + \hat{\mathcal{T}}(\bar{u}_h, \bar{v}_h). \quad (5.29)$$

with the "exact" function $\hat{\mathcal{F}}$ but a perturbation δ_h of the order $\alpha(h)$.

Lemma 12. *Let the sequence (\bar{u}_h, \bar{v}_h) be a solution to the inclusion (5.28). Assume that $\bar{u}_h \rightarrow \bar{u}$, as $h \downarrow 0$, where \bar{u} is a local solution to (P) satisfying the linear independence condition (LICQ). Then the pair (\bar{u}_h, \bar{v}_h) solves the generalized equation (5.29) with a perturbation δ_h of the order α_h .*

Proof. The generalized equation (5.28) is the h -version of (5.23),

$$0 \in \begin{pmatrix} \nabla_u \hat{\mathcal{L}}_h(\bar{u}, \bar{v}) & + & \{0\} \\ -\hat{G}_h(\bar{u}) & + & \partial \psi_{\hat{K}_+}(\bar{v}) \end{pmatrix}.$$

We consider first the upper equation,

$$\begin{aligned} 0 &= \nabla_u \hat{\mathcal{L}}_h(\bar{u}_h, \bar{v}_h) = \nabla f_h(\bar{u}_h) + \sum_{i=1}^{\ell+2m} \bar{v}_{h,i} \nabla G_{h,i}(\bar{u}_h) \\ &= \nabla f(\bar{u}_h) + r_{h,1} + \sum_{i=1}^{\ell+2m} \bar{v}_{h,i} \nabla G_i(\bar{u}_h) + \sum_{i=1}^{\ell+2m} \bar{v}_{h,i} (\nabla G_{h,i}(\bar{u}_h) - \nabla G_i(\bar{u}_h)) \\ &= \nabla f(\bar{u}_h) + \sum_{i=1}^{\ell+2m} \bar{v}_{h,i} \nabla G_i(\bar{u}_h) + r_{h,1} + r_{h,2} \\ &= \nabla_u \hat{\mathcal{L}}_h(\bar{u}_h, \bar{v}_h) - \delta_{h,1}, \end{aligned}$$

where $|r_{h,1}| + |r_{h,2}| + |\delta_{h,1}| \leq c \alpha(h)$. The estimation of $r_{h,2}$ makes use of the boundedness of the multiplier sequence \bar{v}_h stated in Lemma 11. Analogously, we deduce from the lower inclusion

$$\begin{aligned} 0 &\in -\hat{G}_h(\bar{u}_h) + \partial \psi_{\hat{K}_+}(\bar{v}_h) \\ &\in -\hat{G}(\bar{u}_h) + (\hat{G}(\bar{u}_h) - \hat{G}_h(\bar{u}_h)) + \partial \psi_{\hat{K}_+}(\bar{v}_h) \\ &\in -\hat{G}(\bar{u}_h) - \delta_{h,2} + \partial \psi_{\hat{K}_+}(\bar{v}_h) \end{aligned}$$

with $|\delta_{h,2}| \leq c \alpha(h)$. Collecting both results, we see that (5.29) is verified with perturbation $\delta_h^\top = (\delta_{h,1}^\top, \delta_{h,2}^\top)$ of order $\alpha(h)$. \square

Now we are able to prove our main stability result on nonlinear programming problems.

Theorem 8. *Let a locally optimal solution \bar{u} for problem (P) satisfy, together with the Lagrange multiplier $\bar{\nu}$, the linear independence condition (LICQ) and the strong second-order sufficient optimality condition (5.24). Then \bar{u} is locally unique and there exist $\rho > 0$ and $h_0 > 0$ such that, for all $h \in (0, h_0)$, problem (P_h) admits in $B(\bar{u}, \rho)$ a unique locally optimal solution \bar{u}_h . Associated with \bar{u}_h , there exists a unique associated Lagrange multiplier $\bar{\nu}_h$ and a constant $C > 0$ not depending on h such that*

$$|\bar{u} - \bar{u}_h| + |\bar{\nu} - \bar{\nu}_h| \leq C \alpha(h).$$

Proof. The existence of the sequence \bar{u}_h follows from Lemma 9 and Lemma 10, based on the assumptions (LICQ) and (5.24). In particular, we have $\bar{u}_h \rightarrow \bar{u}$ as $h \rightarrow 0$. The same assumptions permit to apply the Robinson implicit function theorem to the inclusion (5.29). Lemma 12 shows that $\delta_h \rightarrow 0$ as $h \rightarrow 0$. Therefore, for all sufficiently small $h > 0$, δ_h belongs to a neighborhood \mathcal{O} of 0 and $(\bar{u}_h, \bar{\nu}_h)$ is contained in a neighborhood W_ε of $(\bar{u}, \bar{\nu})$ where (5.26) is uniquely solvable and the dependence on δ is Lipschitz. In W_ε , there is a unique element $(\tilde{u}, \tilde{\nu})$ that satisfies the generalized equation

$$\delta_h \in \hat{\mathcal{F}}(\tilde{u}, \tilde{\nu}) + \hat{\mathcal{T}}(\tilde{u}, \tilde{\nu}).$$

Moreover, we have the estimate $|\bar{u} - \tilde{u}| + |\bar{\nu} - \tilde{\nu}| \leq C |\delta|$. By uniqueness, $\tilde{u} = \bar{u}_h$ and $\tilde{\nu} = \bar{\nu}_h$ must hold. The local uniqueness of \bar{u} follows in the same way by the implicit function theorem applied to the inclusion (5.26) with $\delta = 0$ at $(\bar{u}, \bar{\nu})$. \square

Remark 6. *Extracting the first ℓ components of the multipliers $\bar{\nu}$ and ν_h , we obtain the error estimate*

$$|\bar{u} - \bar{u}_h| + |\bar{\mu} - \bar{\mu}_h| \leq C \alpha(h). \quad (5.30)$$

that includes only the Lagrange multiplier associated with the state constraints. Conversely, given μ and μ_h , estimate (3.13) follows via the formula (5.22).

Acknowledgement. The authors are grateful to W. Alt (Jena) and D. Klatte (Zurich) for their helpful comments and suggestions.

REFERENCES

- [1] E. L. Allgower, K. Böhmer, F. A. Potra, and W. C. Rheinboldt. A mesh-independence principle for operator equations and their discretizations. *SIAM Journal on Numerical Analysis*, 23:160–169, 1986.
- [2] W. Alt. On the approximation of infinite optimization problems with an application to optimal control problems. *Appl. Math. Opt.*, 12:15–27, 1984.
- [3] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Computational Optimization and Applications*, 23:201–229, 2002.
- [4] F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.
- [5] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 1994.
- [6] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control and Optimization*, 31:993–1006, 1993.

- [7] E. Casas. Error estimates for the numerical approximation of semilinear elliptic control problems with finitely many state constraints. *ESAIM: Control, Optimization and Calculus of Variations*, 8:345–374, 2002.
- [8] E. Casas and M. Mateos. Second order sufficient optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control and Optimization*, 40:1431–1454, 2002.
- [9] E. Casas and M. Mateos. Uniform convergence of the FEM. Applications to state constrained control problems. *J. of Computational and Applied Mathematics*, 21:67–100, 2002.
- [10] J. C. de los Reyes, P. Merino, J. Rehberg, and F. Tröltzsch. Optimality conditions for state-constrained PDE control problems with finite-dimensional control space]. *submitted*, 2006.
- [11] M. Deckelnick and M. Hinze. Convergence of a finite element approximation to a state constrained elliptic control problem. *SIAM J. Numerical Analysis*, to appear, 2006.
- [12] M. Deckelnick and M. Hinze. Numerical analysis of a control and state constrained elliptic control problem with piecewise constant control approximations. *submitted*, 2008.
- [13] Casas. E. Using piecewise linear functions in the numerical approximation of semilinear elliptic control problems. *Adv. Comput. Math.*, 26(1-3):137–153, 2007.
- [14] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. J. Wiley and Sons, Inc., New York, 1968.
- [15] J. Frehse and R. Rannacher. Eine l^1 -Fehlerabschätzung diskreter Grundlösungen in der Methode der finiten Elemente. *Bonner Math. Schriften*, 89:92–114, 1976.
- [16] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin, 1998.
- [17] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.
- [18] D. Klatte. A note on quantitative stability results in nonlinear optimization. Seminarbericht 90, Humboldt-Universität zu Berlin, Sektion Mathematik, 1987.
- [19] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*. Kluwer Academic Publishers, Dordrecht, 2002.
- [20] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison Wesley, Reading, Massachusetts, 1984.
- [21] K. Malanowski. *Stability of solutions to convex problems of optimization*, volume 93 of *Lecture Notes Contr. Inf. Sci.* Springer-Verlag, Berlin, 1987.
- [22] K. Malanowski, Ch. Büskens, and H. Maurer. Convergence of approximations to nonlinear optimal control problems. In A. V. Fiacco, editor, *Mathematical Programming with Data Perturbations*, volume 195 of *Lecture Notes to Pure and Applied Mathematics, vol. 1998*, pages 253–284, New York, 1998. Marcel Dekker.
- [23] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *WIAS, Preprint 1159*, 2006.
- [24] C. Meyer, U. Prüfert, and F. Tröltzsch. On two numerical methods for state-constrained elliptic control problems. Technical report, Institut für Mathematik, Technische Universität Berlin, 2005. Report 5-2005, To appear in *Optimization Methods and Software*.
- [25] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. SIAM Publ., Philadelphia, 2000.
- [26] R. Rannacher. Zur l^∞ -Konvergenz linearer finiter Elemente beim Dirichlet-Problem. *Mathematische Zeitschrift*, 149:69–77, 1976.
- [27] R. Rannacher and B. Vexler. A priori error estimates for the finite element discretization of elliptic parameter identification problems with pointwise measurements. *SIAM J. Control Optim.*, 44(5):1844–1863, 2005.
- [28] S. M. Robinson. Stability theory for systems of inequalities, part ii: differentiable nonlinear systems. *SIAM J. Numer. Analysis*, 13:497–513, 1976.

- [29] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.
- [30] A. Rösch. Error estimates for linear-quadratic control problems with control constraints. *Optim. Methods Softw.*, 21(1):121–134, 2006.
- [31] A. H. Schatz and L. B. Wahlbin. Interior maximum norm estimates for finite element methods. *Math. Comp.*, 31(138):414–442, 1977.
- [32] A. H. Schatz and L. B. Wahlbin. Interior maximum-norm estimates for finite element methods, part ii. *Math. Comp.*, 64(211):907–928, 1995.
- [33] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen – Theorie, Verfahren und Anwendungen*. Vieweg, Wiesbaden, 2005.
- [34] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optimization*, 5:49–62, 1979.

[‡]DEPARTMENT OF MATHEMATICS, EPN QUITO, ECUADOR

[†]INSTITUT FÜR MATHEMATIK, TU BERLIN, GERMANY

^{*}INSTITUT FÜR MATHEMATIK, TU MÜNCHEN, GERMANY