

Error estimates for the finite-element approximation of  
a semilinear elliptic control problem

by

Eduardo Casas<sup>1</sup> and Fredi Tröltzsch<sup>2</sup><sup>1</sup> Departamento de Matemática Aplicada y Ciencias de la Computación  
Universidad de Cantabria, 39005 Santander, Spain.<sup>2</sup> Institut für Mathematik, Technische Universität Berlin  
10623 Berlin, Germany.

**Abstract:** We consider the finite-element approximation of a distributed optimal control problem governed by a semilinear elliptic partial differential equation, where pointwise constraints on the control are given. We prove the existence of local approximate solutions converging to a given local reference solution. Moreover, we derive error estimates for local solutions in the maximum norm.

**Keywords:** Distributed control, semilinear elliptic equation, numerical approximation, finite element method, error estimates.

## 1. Introduction

In this paper, we consider the finite-element discretization of the optimal control problem

$$(P) \quad \min J(u) = \frac{1}{2} \int_{\Omega} \{(y(x) - y_d(x))^2 + \nu u(x)^2\} dx, \\ \text{subject to } (y, u) \in (C(\bar{\Omega}) \cap H^1(\Omega)) \times L^\infty(\Omega),$$

$$Ay + f(y) = u \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma, \\ u \in U^{ad} = \{u \in L^\infty(\Omega) \mid \alpha \leq u(x) \leq \beta \text{ for a.a. } x \in \Omega\}, \quad (1.1)$$

where  $\Omega \subset \mathbb{R}^n$  is a convex bounded domain,  $\Gamma$  is the boundary of  $\Omega$ , and  $A$  denotes a second-order elliptic operator of the form

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

Here,  $D_i$  denotes the partial derivative with respect to  $x_i$ ,  $u$  is the *control*, and  $y = y(u)$  is said to be the *associated state*. The function  $y_d$  is given in  $L^\infty(\Omega)$ , and  $\alpha < \beta$ ,  $\nu > 0$  are real constants.

Based on a standard finite-element approximation, we set up an approximate optimal control problem  $(P_h)$ . Our main aim is to estimate the error  $\|\bar{u} - \bar{u}_h\|$  in the maximum norm, where  $\bar{u}$  stands for a fixed locally optimal control of  $(P)$  and  $\bar{u}_h$  is an associated one of  $(P_h)$ . Error estimates for optimal controls certainly cannot improve those known for the solutions of elliptic equations. However, one should expect that they reflect the order of the associated estimates for equations. Due to the non-convexity of  $(P)$  and the presence of control-constraints, this is not an easy task. Optimal  $L^2$ -estimates are known since long time for linear-quadratic elliptic control problems, see Falk (1973) or Geveci (1979). Recently,  $L^\infty$ -error estimates being optimal in that sense have been derived for the case of nonlinear equations in Arada, Casas and Tröltzsch (2002).

Moreover, we mention two further papers related to the semilinear elliptic case. Recently, Arnautu and Neittaanmäki (1998) contributed error estimates to this class of problems. Their technique, however, slightly overestimates the order of the error. We also mention the paper by Casas and Mateos (2001), who carefully study error estimates for semilinear elliptic equations. In contrast to the elliptic case, quite a number of papers was devoted to parabolic problems, although the associated theory is far from being complete. We refer to the references in Arada et al. (2002).

Our paper complements the theory presented in Arada et al. (2002), where error estimates have been derived for a subsequence  $(\bar{u}_h)_h$  of *globally* optimal controls for  $(P_h)$  that converges to an optimal control  $\bar{u}$  of  $(P)$  as  $h \downarrow 0$ . The existence of this sequence has been obtained by weak compactness arguments.

The main difference of our paper with respect the one cited concerns the existence part. Here, we concentrate on *locally* optimal controls, since they are the natural result of numerical optimization algorithms. Suppose that a *locally* optimal control  $\bar{u}$  of  $(P)$  is given. Then we expect to have a sequence  $(\bar{u}_h)_h$  of *locally* optimal controls for  $(P_h)$  converging to  $\bar{u}$ . This should be true for *each* fixed local solution  $\bar{u}$ . We prove that each locally optimal control of  $(P)$  can be approximated by locally optimal controls of  $(P_h)$ , while Arada et al. (2002) only guarantee that the computed global solutions contain a *subsequence* that converges to a certain globally optimal control.

Therefore, we start from a fixed reference control  $\bar{u}$  being locally optimal for  $(P)$ . Next we prove the existence of a sequence  $(\bar{u}_h)_h$  of locally optimal controls for  $(P_h)$  converging to  $\bar{u}$ . We do not use compactness arguments. Finally, the order of convergence is quantified by estimating the error  $\bar{u}_h - \bar{u}$ . The error analysis is similar to that of our paper Arada et al. (2002).

However, our problem  $(P)$  is simplified to shorten the presentation. In our former paper, the objective functional and the nonlinearity  $f$  are more general. Following the lines of Arada et al. (2002), the results of this paper can be

## 2. Assumptions and notation

The domain  $\Omega$  is assumed to be a convex, bounded, and open subset in  $\mathbb{R}^n$ , where  $n = 2$  or  $n = 3$ . We also assume that  $\Omega$  has a boundary  $\Gamma$  of class  $C^{1,1}$ . The coefficients  $a_{ij}$  of the operator  $A$  are assumed to be in  $C^{0,1}(\bar{\Omega})$ , and to satisfy the ellipticity condition

$$m_0|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \quad \forall (\xi, x) \in \mathbb{R}^n \times \bar{\Omega}, \quad m_0 > 0.$$

On  $f$ , we impose the following assumption

**(A1)** *The function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is of class  $C^2$  and its first derivative  $f'$  is nonnegative. For all  $M > 0$ , there exists  $C_M > 0$  such that*

$$|f''(y_1) - f''(y_2)| \leq C_M|y_1 - y_2|$$

for all  $(y_1, y_2) \in [-M, +M]^2$ .

Assumption **(A1)** permits to deal with highly nonlinear functions. For instance,  $f(y) = \exp(y)$  satisfies **(A1)**.

**THEOREM 2.1** (Bonnans and Casas, 1995) *Let  $u$  in  $L^\infty(\Omega)$  satisfy  $\|u\|_{\infty, \Omega} \leq M$ . Then, for every  $p > n$ , equation (1.1) admits a unique solution  $y = y(u) \in H_0^1(\Omega) \cap W^{2,p}(\Omega)$ . There exists a positive constant  $C = C(\Omega, n, p, M)$ , independent of  $u$ , such that*

$$\|y(u)\|_{W^{2,p}(\Omega)} \leq C.$$

In what follows,  $\|\cdot\|_2$  and  $\|\cdot\|_\infty$  denote the natural norms in  $L^2(\Omega)$  and  $L^\infty(\Omega)$ , respectively, and  $c$  is a generic constant.

## 3. Optimality conditions for local solutions of (P)

The existence of a (global) solution to (P) can be proved by classical arguments. However, we concentrate on *local* solutions. Therefore, we just *assume* that a locally optimal reference control  $\bar{u}$  is given for (P) that satisfies the standard first-order necessary and second-order sufficient optimality conditions.

A control  $\bar{u} \in U^{ad}$  is said to be locally optimal or a *local solution* of (P), if there is an  $r > 0$  such that

$$J(u) \geq J(\bar{u}) \quad \forall u \in U^{ad} \text{ with } \|u - \bar{u}\|_\infty \leq r.$$

In what follows, we denote by  $y(u)$  the solution  $y$  of (1.1) that is associated with  $u$ . Let  $\bar{y}$  be the state corresponding to  $\bar{u}$ , i.e.  $\bar{y} = y(\bar{u})$ .

Next we recall the known first-order necessary optimality conditions for (P). To this aim, we introduce the adjoint equation. Let  $u$  be in  $L^\infty(\Omega)$  with state  $y(u)$ . The *adjoint equation* has the following form:

Here,  $A^*$  is the formal adjoint operator of  $A$ . The solution  $\varphi = \varphi(u)$  is the *adjoint state* associated with  $u$ .

**THEOREM 3.1** *If  $\bar{u}$  is a local solution of (P), then there exists an adjoint state  $\bar{\varphi} = \varphi(\bar{u}) \in H_0^1(\Omega) \cap W^{2,p}(\Omega)$  such that*

$$A^* \bar{\varphi} + f'(\bar{y}) \bar{\varphi} = \bar{y} - y_d \quad \text{in } \Omega, \quad (3.2)$$

$$\int_{\Omega} (\bar{\varphi} + \nu \bar{u})(u - \bar{u}) \, dx \geq 0 \quad \forall u \in U^{ad}. \quad (3.3)$$

The classical proof is omitted. By a further discussion, the variational inequality (3.3) is seen to be equivalent to the following known relation:

$$\bar{u}(x) = \text{Proj}_{[\alpha, \beta]} \left( -\frac{1}{\nu} \varphi(\bar{u})(x) \right), \quad (3.4)$$

where  $\text{Proj}_{[\alpha, \beta]}$  denotes the projection from  $\mathbb{R}$  onto  $[\alpha, \beta]$ . Since (P) is non-convex, the optimality conditions above are not sufficient for (local) optimality. To have this, in addition the following *second-order sufficient optimality condition* is assumed:

**(SSC)** *There are  $\delta > 0$  and  $\tau > 0$  such that*

$$J''(\bar{u})v^2 \geq \delta \|v\|_2^2 \quad (3.5)$$

*holds for all  $v \in L^\infty(\Omega)$  satisfying*

$$v(x) \begin{cases} \geq 0 & \text{if } \bar{u}(x) = \alpha, \\ \leq 0 & \text{if } \bar{u}(x) = \beta, \\ = 0 & \text{if } |\bar{\varphi}(x) + \nu \bar{u}(x)| \geq \tau > 0. \end{cases} \quad (3.6)$$

All functions  $v$  satisfying the conditions of (3.6) form a cone that we shall call the  $\tau$ -critical cone. The set

$$A_\tau = \{x \in \Omega \mid |\bar{\varphi}(x) + \nu \bar{u}(x)| \geq \tau\}$$

is the set of all points where the control constraints are *strongly active*. This notion was introduced by Dontchev, Hager, Poore and Yang (1995).

Notice that  $J$  is defined as a functional on  $L^\infty(\Omega)$ . It is this space, where the derivatives  $J'$  and  $J''$  are defined. The concrete expression for the second derivative can be formulated by the Lagrange function

$$L(y, u, \varphi) = \frac{1}{2} \int_{\Omega} \{(y(x) - y_d(x))^2 + \nu u(x)^2\} \, dx$$

$$- \int_{\Omega} (\lambda \nu u + f(y) - u) \varphi \, dx$$

which is here only formally defined (in our setting,  $\Delta y$  is not a function; by selecting a slightly different state space for  $y$ , this can be made precise). Then, see Casas and Tröltzsch (2000),

$$\begin{aligned} J''(u)(u_1, u_2) &= D_{yy}L(y, u, \varphi)(y_1, y_2) + D_{uu}L(y, u, \varphi)(u_1, u_2) \\ &= \int_{\Omega} (1 - f''(y)\varphi(u)) y_1 y_2 \, dx + \nu \int_{\Omega} u_1 u_2 \, dx, \end{aligned}$$

where  $y_i \in H_0^1(\Omega)$  solve the linearized equation  $-\Delta y_i + f'(y)y_i = u_i$ . Therefore, (SSC) requires the coercivity of  $L''$  on the cone defined by the controls  $u$  of the  $\tau$ -critical cone and the associated solutions  $y(u)$  of the linearized equation.

## 4. Finite-element approximation of $(P)$ : Basic results

### 4.1. The approximate problem $(P_h)$

Here we define a finite-element based approximation of the optimal control problem  $(P)$ . To this aim, we consider a family of triangulations  $(\mathcal{T}_h)_{h>0}$  of  $\bar{\Omega}$ . With each element  $T \in \mathcal{T}_h$ , we associate two parameters  $\rho(T)$  and  $\sigma(T)$ , where  $\rho(T)$  denotes the diameter of the set  $T$  and  $\sigma(T)$  is the diameter of the largest ball contained in  $T$ . Define the mesh size of the grid by  $h = \max_{T \in \mathcal{T}_h} \rho(T)$ . We suppose that the following regularity assumptions are satisfied.

(A2) There exist two positive constants  $\rho$  and  $\sigma$  such that

$$\frac{\rho(T)}{\sigma(T)} \leq \sigma, \quad \frac{h}{\rho(T)} \leq \rho$$

hold for all  $T \in \mathcal{T}_h$  and all  $h > 0$ .

Let us take  $\bar{\Omega}_h = \cup_{T \in \mathcal{T}_h} T$ , and let  $\Omega_h$  and  $\Gamma_h$  denote its interior and its boundary, respectively. We assume that  $\bar{\Omega}_h$  is convex and that the vertices of  $\mathcal{T}_h$  placed on the boundary of  $\Gamma_h$  are points of  $\Gamma$ . It is known that

$$|\Omega \setminus \Omega_h| \leq Ch^2. \quad (4.1)$$

Now, for  $n = 2$ , to every boundary triangle  $T$  of  $\mathcal{T}_h$ , we associate another triangle  $\hat{T} \subset \bar{\Omega}$  with curved boundary as follows: The edge between the two boundary nodes of  $T$  is substituted by the part of  $\Gamma$  connecting these nodes and forming a triangle with the remaining interior sides of  $T$ . We denote by  $\hat{\mathcal{T}}_h$  the union of these curved boundary triangles with the interior triangles to  $\Omega$  of  $\mathcal{T}_h$ , so that  $\bar{\Omega} = \cup_{\hat{T} \in \hat{\mathcal{T}}_h} \hat{T}$ . For  $n = 3$ , tetrahedra are analogously. Let us set

$$U_h = \{u \in L^\infty(\Omega) \mid u|_{\hat{T}} \text{ is constant on all } \hat{T} \in \hat{\mathcal{T}}_h\}, \quad U_h^{ad} = U_h \cap U^{ad},$$

$$V_h = \{y_h \in C(\bar{\Omega}) \mid y_h|_T \in \mathcal{P}_1, \text{ for all } T \in \mathcal{T}_h, \text{ and } y_h = 0 \text{ on } \bar{\Omega} \setminus \Omega_h\},$$

where  $\mathcal{P}_1$  is the space of polynomials of degree less or equal than 1. For each  $u_h \in U_h$ , we denote by  $y_h = y_h(u_h)$  the unique element of  $V_h$  that satisfies

$$a(u_h, v_h) = \int_{\Omega} (u_h - f(u_h)) v_h \, dx \quad \forall v_h \in V_h. \quad (4.2)$$

where  $a : V_h \times V_h \rightarrow \mathbb{R}$  is the bilinear form defined by

$$a(y, \eta) = \int_{\Omega} \left( \sum_{i,j=1}^n a_{ij}(x) D_i y(x) D_j \eta(x) \right) dx.$$

In other words,  $y_h(u_h)$  is the approximate state associated with  $u_h$ . In the two integrals above, the test function  $\eta_h$  vanishes outside  $\Omega_h$  so that there is no difference between integration on  $\Omega$  and  $\Omega_h$ . Existence and uniqueness of this solution  $y_h(u_h)$  can be shown under our assumption (A1), see Casas and Mateos (2001) and Mateos (2000). The finite-dimensional *approximate optimal control problem*  $(P_h)$  is defined by

$$(P_h) \quad \min J_h(u_h) = \frac{1}{2} \int_{\Omega_h} \{ (y_h(u_h) - y_d)^2 + \nu u_h^2 \} dx, \quad u_h \in U_h^{ad}.$$

The existence of at least one global solution for  $(P_h)$  follows from the continuity of  $J_h$  and the compactness of  $U_h^{ad}$ . However, this global solution need not be unique. Moreover, it can be far from the reference solution  $\bar{u}$ . Therefore, we do not concentrate on global solutions of  $(P_h)$ . Again, we consider certain local solutions.

REMARK: We tacitly assume that we are able to evaluate the integrals in (4.2) and  $(P_h)$  exactly. In general, numerical integration has to be used, which generates another sort of errors. We do not include them in our analysis.

#### 4.2. Characterization of local solutions of $(P_h)$

Local solutions of the approximate problem  $(P_h)$  are defined analogously to  $(P)$ : A control  $\bar{u}_h \in U_h^{ad}$  is a *local solution* of  $(P_h)$ , if

$$J_h(u_h) \geq J_h(\bar{u}_h) \quad \forall u_h \in U_h^{ad} \text{ with } \|u_h - \bar{u}_h\|_{\infty} \leq r$$

holds for a certain  $r > 0$ . Associated necessary optimality conditions are similar to those for  $(P)$  in Section 3.: With the solution  $\bar{u}_h$  we associate the *discrete adjoint equation* for  $\varphi_h \in V_h$

$$\begin{aligned} & \int_{\Omega} \sum_{i,j=1}^n a_{ij} D_j \varphi_h D_i \eta_h dx + \int_{\Omega} f'(y_h(\bar{u}_h)) \varphi_h \eta_h dx \\ & = \int_{\Omega} (y_h(\bar{u}_h) - y_d) \eta_h dx \quad \forall \eta_h \in V_h. \end{aligned} \quad (4.3)$$

THEOREM 4.1 *Suppose that assumption (A1) is satisfied. If  $\bar{u}_h$  is a local solution of  $(P_h)$ , then there exists a unique solution  $\bar{\varphi}_h = \varphi_h(\bar{u}_h) \in H_0^1(\Omega) \cap C^{0,1}(\bar{\Omega})$  of the discrete adjoint equation (4.3) such that the variational inequality*

$$\int_{\Omega_h} (\bar{\varphi}_h + \nu \bar{u}_h)(u - \bar{u}_h) dx \geq 0 \quad \forall u \in U_h^{ad} \quad (4.4)$$

The standard proof of this result is omitted. Throughout the sequel, for  $v$  fixed in  $L^\infty(\Omega)$ , we denote by  $y_h(v)$  and  $\varphi_h(v)$  the solutions of (4.2) and (4.3), respectively, associated with  $v$ . The next result is the discrete counterpart of (3.4). The discrete local solution  $\bar{u}_h$  satisfies

$$\bar{u}_h|_T = \text{Proj}_{[\alpha, \beta]} \left( -\frac{1}{\nu|T|} \int_T \varphi_h(\bar{u}_h)(x) dx \right) \quad \forall T \in \mathcal{T}_h. \quad (4.5)$$

In this paper, we frequently use an interpolation operator  $\Pi_h : L^2(\Omega) \rightarrow U_h$  that assigns piecewise constant functions on  $\Omega$  to functions of  $L^2(\Omega)$ . To define  $\Pi_h$ , we first introduce the interpolation operator  $\pi_h : L^2(\Omega) \rightarrow L^2(\Omega_h)$  by

$$(\pi_h v)|_T = \frac{1}{|T|} \int_T v(x) dx.$$

We extend  $\pi_h$  to  $\Pi_h$  by

$$(\Pi_h v)(x) = \begin{cases} (\pi_h v)(x) & \text{if } x \in T \\ (\pi_h v)(x_o) & \text{if } x \in \hat{T} \setminus T. \end{cases}$$

Here,  $x_o$  is the projection of  $x$  onto the boundary of the triangle (or tetrahedron)  $T$  that is covered by  $\hat{T}$ . Let us mention an important property of  $\Pi_h$ : If  $v$  is a Lipschitz function, then

$$\|v - \Pi_h v\|_\infty \leq ch.$$

This is seen as follows: On triangles  $T \in \mathcal{T}_h$  we have  $\max_{x \in T} |v(x) - (\Pi_h v)(x)| = \max_{x \in T} |v(x) - (\pi_h v)(x)| \leq ch$  by the known properties of the interpolation operator  $\pi_h$  and the Lipschitz property of  $v$ . If  $x \in \hat{T} \setminus T$ , then

$$\begin{aligned} |v(x) - (\Pi_h v)(x)| &\leq |v(x) - v(x_o)| + |v(x_o) - (\Pi_h v)(x)| \\ &\leq ch + |v(x_o) - (\pi_h v)(x_o)| \leq ch. \end{aligned}$$

Here, we have used the fact that  $\text{dist}(x_o, T) \leq ch$ . The same estimate follows for the  $L^2$ -norm using (4.1). With this interpolation operator, (4.5) admits the form

$$\bar{u}_h = \text{Proj}_{[\alpha, \beta]} \left( -\frac{1}{\nu} \Pi_h \varphi_h(\bar{u}_h) \right), \quad (4.6)$$

since the extension of (4.5) from boundary triangles (or tetrahedra)  $T$  to  $\hat{T}$  is the same on the left and right hand sides of (4.6).

#### 4.3. Error-estimates for the state and the adjoint state

Here we provide some known results on the finite element approximation of the state equation (1.1) and its adjoint equation (3.1). They are basic for

Recall that  $y(v)$  and  $y_h(v_h)$  are the solutions of (1.1) and (4.2) corresponding to  $v$  and  $v_h$ . Analogously,  $\varphi(v)$  and  $\varphi_h(v_h)$  are the solutions of (3.1) and (4.3) corresponding to  $v$  and  $v_h$ .

In all what follows we tacitly assume that (A1) and (A2) are satisfied. Moreover, we fix once and for all a local reference solution  $\bar{u}$  for (P) that satisfies (SSC). Therefore, we do not mention (A1), (A2), and (SSC) in further statements.

All controls  $u$ ,  $v$ ,  $u_h$ ,  $v_h$  etc. used below are contained in  $U^{ad}$ . Therefore, they are uniformly bounded, and the same holds true for all associated states and adjoint states so that all  $y$ ,  $\varphi$ ,  $y_h$ ,  $\varphi_h$  are bounded by the same constant  $M$ .

**THEOREM 4.2** *Let  $v$  and  $v_h$  belong to  $U^{ad}$ . Then the estimates*

$$\|y(v) - y_h(v_h)\|_{H^1(\Omega)} + \|\varphi(v) - \varphi_h(v_h)\|_{H^1(\Omega)} \leq C(h + \|v - v_h\|_2), \quad (4.7)$$

$$\|y(v) - y_h(v_h)\|_2 + \|\varphi(v) - \varphi_h(v_h)\|_2 \leq C(h^2 + \|v - v_h\|_2), \quad (4.8)$$

$$\|y(v) - y_h(v_h)\|_\infty + \|\varphi(v) - \varphi_h(v_h)\|_\infty \leq C(h^\lambda + \|v - v_h\|_2), \quad (4.9)$$

hold, where  $C = C(\Omega, n)$  is a positive constant independent of  $h$ , and  $\lambda = 2 - n/2$ . Moreover, if the triangulation is of nonnegative type, then

$$\|y(v) - y_h(v_h)\|_\infty + \|\varphi(v) - \varphi_h(v_h)\|_\infty \leq C(h + \|v - v_h\|_2), \quad (4.10)$$

holds independently of  $h$ .

For the proof of this theorem the reader is referred to Arada et al. (2002). In all what follows, let us fix

$$\lambda = \begin{cases} 2 - n/2 & \text{for regular triangulations} \\ 1 & \text{for triangulations of nonnegative type.} \end{cases}$$

#### 4.4. Convergence results

Aiming to derive error estimates, we have to find a sequence  $(\bar{u}_h)_h$  of local solutions of  $(P_h)$  tending to  $\bar{u}$  as  $h \downarrow 0$ . To solve this nontrivial problem, we proceed as follows: For  $\varepsilon > 0$  we consider the auxiliary control problem

$$(P_h^\varepsilon) \quad \min J_h(u_h) = \frac{1}{2} \int_{\Omega_h} \{(y_h(u_h) - y_d)^2 + \nu u_h^2\} dx, \quad u_h \in U_{h,\varepsilon}^{ad},$$

where

$$U_{h,\varepsilon}^{ad} = \{u \in U_h^{ad} \mid (\Pi_h \bar{u})(x) - \varepsilon \leq u(x) \leq (\Pi_h \bar{u})(x) + \varepsilon \text{ in } \Omega\}.$$

The interpolate  $\Pi_h \bar{u}$  belongs to  $U_{h,\varepsilon}^{ad}$ , and so the admissible set of  $(P_h^\varepsilon)$  is not empty. This problem has a global solution  $u_h^\varepsilon$ , hence it is also a local solution for  $(P_h^\varepsilon)$ . We show that this solution is even a local solution of  $(P_h)$  and tends



It is known that the second-order condition (SSC) implies the existence of positive constants  $\kappa$  and  $r$  such that the quadratic growth condition

$$J(u) \geq J(\bar{u}) + \kappa \|u - \bar{u}\|_2^2 \quad (4.11)$$

is satisfied for all  $u \in U^{ad}$  with  $\|u - \bar{u}\|_\infty \leq r$ , see Casas, Tröltzsch and Unger (2000). Now take  $\bar{\varepsilon} := r/2$ . Then for all  $\varepsilon \leq \bar{\varepsilon}$  and all sufficiently small  $h$ , say  $0 < h \leq \bar{h}$ ,

$$u \in U_{h,\varepsilon}^{ad} \Rightarrow \|u - \bar{u}\|_\infty \leq r, \quad (4.12)$$

because  $\|u - \bar{u}\|_\infty \leq \|u - \Pi_h \bar{u}\|_\infty + \|\Pi_h \bar{u} - \bar{u}\|_\infty$ , the first term is not greater than  $r/2$  by the definition of  $U_{h,\varepsilon}^{ad}$ , and the second term tends to zero as  $h \downarrow 0$ . Notice that (4.11) and (4.12) imply

$$J(u) \geq J(\bar{u}) + \kappa \|u - \bar{u}\|_2^2 \quad \forall u \in U_{h,\bar{\varepsilon}}^{ad}. \quad (4.13)$$

LEMMA 4.1 For all  $\varepsilon \leq \bar{\varepsilon}$ , the objective values  $J_h(u_h^\varepsilon)$  converge to  $J(\bar{u})$ , i.e.

$$\lim_{h \downarrow 0} J_h(u_h^\varepsilon) = J(\bar{u}).$$

*Proof.* We have

$$J_h(u_h^\varepsilon) = J(u_h^\varepsilon) + (J_h(u_h^\varepsilon) - J(u_h^\varepsilon)) \geq J(\bar{u}) - ch,$$

since  $\|u_h^\varepsilon\|_\infty$  is uniformly bounded, hence  $|J_h(u_h^\varepsilon) - J(u_h^\varepsilon)| \leq ch$ . Moreover,  $J(u_h^\varepsilon) \geq J(\bar{u})$  follows from (4.13). On the other hand, we know that  $\Pi_h \bar{u} \in U_{h,\varepsilon}^{ad}$ , and the optimality of  $u_h^\varepsilon$  for  $(P_h^\varepsilon)$  gives

$$\begin{aligned} J_h(u_h^\varepsilon) &\leq J_h(\Pi_h \bar{u}) = J(\bar{u}) + (J(\Pi_h \bar{u}) - J(\bar{u})) + (J_h(\Pi_h \bar{u}) - J(\Pi_h \bar{u})) \\ &\leq J(\bar{u}) + ch, \end{aligned}$$

since  $\|\Pi_h \bar{u} - \bar{u}\|_\infty \leq ch$  and  $|J_h(v) - J(v)| \leq ch$  for all  $v \in U^{ad}$ . Both inequalities imply the statement of the Lemma.  $\blacksquare$

LEMMA 4.2 There are  $0 < \varepsilon_\tau \leq \bar{\varepsilon}$  and  $0 < h_\tau \leq \bar{h}$  such that

$$|\varphi_h(u_h^\varepsilon)(x) + \nu u_h^\varepsilon(x)| \geq \tau/4 \quad (4.14)$$

$$u_h^\varepsilon(x) = \bar{u}(x) \quad (4.15)$$

hold for all  $\varepsilon \leq \varepsilon_\tau$ , all  $h \leq h_\tau$ , and all  $x \in T$ , if  $T$  has a non-empty intersection with  $A_\tau$ .

*Proof.* We know that on  $A_\tau$  either  $\varphi(\bar{u})(x) + \nu \bar{u}(x) \geq \tau$ , where  $\bar{u}(x) = \alpha$  or  $\varphi(\bar{u})(x) + \nu \bar{u}(x) \leq -\tau$ , where  $\bar{u}(x) = \beta$ . Now take an arbitrary but fixed  $T$  having a non-empty intersection with  $A_\tau$ . If  $h$  is sufficiently small, then we can assume that one of these two cases holds for all  $x \in A_\tau \cap T$ , since the function  $\varphi(\bar{u}) + \nu \bar{u}$  is Lipschitz continuous. We consider the case

where  $\bar{u}(x) \equiv \alpha$  on  $A_\tau \cap T$ . The arguments for  $\bar{u}(x) \equiv \beta$  are analogous. Therefore, if  $h$  is sufficiently small, then

$$\varphi(\bar{u})(x) + \nu \bar{u}(x) \geq 3\tau/4 \quad \forall x \in T,$$

thus also  $\bar{u}(x) \equiv \alpha$  on  $T$ . If  $\varepsilon$  is sufficiently small, say  $\varepsilon \leq \varepsilon_\tau$ , then  $\|u_h^\varepsilon - \bar{u}\|_\infty$  is so small that

$$\begin{aligned} \varphi_h(u_h^\varepsilon) + \nu u_h^\varepsilon &= \varphi(\bar{u}) + \nu \bar{u} + (\varphi(u_h^\varepsilon) - \varphi(\bar{u})) + \nu(u_h^\varepsilon - \bar{u}) \\ &+ (\varphi_h(u_h^\varepsilon) - \varphi(u_h^\varepsilon)) \geq 2\tau/4 - ch^\lambda \geq \tau/4 \end{aligned}$$

holds on  $T$  for all sufficiently small  $h \leq h_\tau$ . On  $T$ , the variational inequality for  $u_h^\varepsilon$  reads

$$\int_T (\varphi_h(u_h^\varepsilon) + \nu u_h^\varepsilon) (u - u_h^\varepsilon) dx \geq 0$$

for all  $u \in \mathbb{R}$  such that  $u \in [\alpha, \beta] \cap [\Pi_h \bar{u}|_T - \varepsilon, \Pi_h \bar{u}|_T + \varepsilon]$ . We know that on  $T$ ,  $\bar{u}(x) \equiv \alpha$ , hence  $\Pi_h \bar{u}|_T = \alpha$ , and therefore  $u$  varies in  $[\alpha, \alpha + \varepsilon]$ . The positivity of  $\varphi_h(u_h^\varepsilon) + \nu u_h^\varepsilon$  in the variational inequality above implies that  $u_h^\varepsilon$  must admit the left end of  $[\alpha, \alpha + \varepsilon]$ , i.e.  $u_h^\varepsilon|_T = \alpha = \bar{u}(x)$ . ■

By our construction, this Lemma is also true for boundary triangles (or tetrahedra)  $\hat{T}$ .

LEMMA 4.3 *If  $\varepsilon \leq \bar{\varepsilon}$ , then  $\lim_{h \downarrow 0} \|u_h^\varepsilon - \bar{u}\|_2 = 0$ .*

*Proof.* By  $u_h^\varepsilon \in U_{h,\varepsilon}^{ad}$ ,  $\varepsilon \leq \bar{\varepsilon}$ ,  $h \downarrow 0$ , and (4.12) we know  $\|u_h^\varepsilon - \bar{u}\|_\infty \leq r$ , hence (4.11) applies,

$$J(u_h^\varepsilon) \geq J(\bar{u}) + \kappa \|u_h^\varepsilon - \bar{u}\|_2^2,$$

thus

$$J_h(u_h^\varepsilon) = J(u_h^\varepsilon) + (J_h(u_h^\varepsilon) - J(u_h^\varepsilon)) \geq J(\bar{u}) + \kappa \|u_h^\varepsilon - \bar{u}\|_2^2 - ch^\lambda$$

and therefore

$$J_h(u_h^\varepsilon) - J(\bar{u}) + ch^\lambda \geq \kappa \|u_h^\varepsilon - \bar{u}\|_2^2.$$

Lemma 4.1 yields  $J_h(u_h^\varepsilon) \rightarrow J(\bar{u})$  as  $h \downarrow 0$  and the assertion of Lemma 4.3 follows immediately. ■

THEOREM 4.3 *If  $\varepsilon \leq \bar{\varepsilon}$ , then*

$$\lim \|u_h^\varepsilon - \bar{u}\|_\infty = 0. \quad (4.16)$$

*Proof.* We start with the result of Lemma 4.3. From Theorem 4.2, (4.9), we deduce that  $u_h^\varepsilon \rightarrow \bar{u}$  in  $L^2(\Omega)$  implies  $\|\varphi_h(u_h^\varepsilon) - \varphi(\bar{u})\|_\infty \rightarrow 0$ . We have the projection formulas

$$\bar{u}(x) = \text{Proj}_{[\alpha, \beta]} \left( -\frac{1}{\nu} \bar{\varphi}(x) \right) \quad (4.17)$$

$$u_h^\varepsilon(x) = \text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} \left( -\frac{1}{\nu} \Pi_h \varphi_h(u_h^\varepsilon(x)) \right), \quad (4.18)$$

where

$$\alpha_h^\varepsilon(x) = \max(\alpha, \Pi_h \bar{u}(x) - \varepsilon), \quad \beta_h^\varepsilon(x) = \min(\beta, \Pi_h \bar{u}(x) + \varepsilon).$$

Notice that  $\alpha_h^\varepsilon$  and  $\beta_h^\varepsilon$  are step functions on  $\Omega$ . Define analogously

$$\alpha^\varepsilon(x) = \max(\alpha, \bar{u}(x) - \varepsilon), \quad \beta^\varepsilon(x) = \min(\beta, \bar{u}(x) + \varepsilon).$$

It is quite obvious that  $\bar{u}$  also satisfies the projection formula

$$\bar{u}(x) = \text{Proj}_{[\alpha^\varepsilon(x), \beta^\varepsilon(x)]} \left( -\frac{1}{\nu} \bar{\varphi}(x) \right). \quad (4.19)$$

Indeed,  $\bar{u}$  solves (P) with the additional restrictions  $u(x) \leq \bar{u}(x) + \varepsilon$ ,  $u(x) \geq \bar{u}(x) - \varepsilon$ , and both of these inequalities are not active at  $\bar{u}$ . Therefore the equations (4.17) and (4.19) are equivalent. Of course, (4.19) can also be directly derived from (4.17). We leave this to the reader. With these prerequisites, the proof can be easily completed. In view of (4.18) and (4.19)

$$\begin{aligned} & |\bar{u}(x) - u_h^\varepsilon(x)| \\ &= |\text{Proj}_{[\alpha^\varepsilon(x), \beta^\varepsilon(x)]} \left( -\frac{1}{\nu} \varphi(\bar{u}(x)) \right) - \text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} \left( -\frac{1}{\nu} \Pi_h \varphi_h(u_h^\varepsilon(x)) \right)| \\ &\leq |\text{Proj}_{[\alpha^\varepsilon(x), \beta^\varepsilon(x)]} \left( -\frac{1}{\nu} \varphi(\bar{u}(x)) \right) - \text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} \left( -\frac{1}{\nu} \varphi(\bar{u}(x)) \right)| \\ &\quad + |\text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} \left( -\frac{1}{\nu} \varphi(\bar{u}(x)) \right) - \text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} \left( -\frac{1}{\nu} \Pi_h \varphi_h(u_h^\varepsilon(x)) \right)|. \end{aligned}$$

The first difference tends uniformly to zero, as

$$\text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} v(x) = \min(\beta_h^\varepsilon(x), \max(\alpha_h^\varepsilon(x), v(x)))$$

is a composition based on continuous functions, if  $v \in C(\bar{\Omega})$ . Therefore

$$\text{Proj}_{[\alpha_h^\varepsilon(x), \beta_h^\varepsilon(x)]} v(x) \rightarrow \text{Proj}_{[\alpha^\varepsilon(x), \beta^\varepsilon(x)]} v(x)$$

in  $C(\bar{\Omega})$ , since  $\alpha_h^\varepsilon(x) \rightarrow \alpha^\varepsilon(x)$  and  $\beta_h^\varepsilon(x) \rightarrow \beta^\varepsilon(x)$  in  $C(\bar{\Omega})$ . The second difference tends uniformly to zero, as the projection operator is Lipschitz continuous with constant 1 and  $\Pi_h \varphi_h(u_h^\varepsilon(x))$  tends uniformly to  $\varphi(\bar{u}(x))$  by Lemma 4.3 and (4.9). ■

Finally, we show that  $u_h^\varepsilon$  is a local solution of  $(P_h)$ . Intuitively, this follows from  $u_h^\varepsilon \rightarrow \bar{u}$ . Therefore  $u_h^\varepsilon$  cannot be located at the boundary of the ball

LEMMA 4.4 *Suppose that  $\varepsilon \leq \bar{\varepsilon}$ . Then,  $u_h^\varepsilon$  is a local solution of  $(P_h)$  for all sufficiently small  $h$ .*

*Proof.* We have to show that

$$J_h(u_h) \geq J_h(u_h^\varepsilon) \quad (4.20)$$

holds for all  $u_h \in U_h^{ad}$  such that  $\|u_h - u_h^\varepsilon\|_\infty \leq \varepsilon/2$ . By the definition of  $u_h^\varepsilon$  we know (4.20) only for all  $u_h \in U_h^{ad}$  with  $\|u_h - \Pi_h \bar{u}\|_\infty \leq \varepsilon$ . Let  $u_h \in U_h^{ad}$  satisfy  $\|u_h - u_h^\varepsilon\|_\infty \leq \varepsilon/2$ . Then, if  $h$  is sufficiently small,

$$\begin{aligned} \|u_h - \Pi_h \bar{u}\|_\infty &\leq \|u_h - u_h^\varepsilon\|_\infty + \|u_h^\varepsilon - \bar{u}\|_\infty + \|\bar{u} - \Pi_h \bar{u}\|_\infty \\ &\leq \varepsilon/2 + \varepsilon/4 + \varepsilon/4 = \varepsilon, \end{aligned}$$

since  $u_h^\varepsilon$  tends to  $\bar{u}$  by Theorem 4.3 and  $\Pi_h \bar{u} \rightarrow \bar{u}$  as  $h \downarrow 0$ . Therefore,  $u_h^\varepsilon$  belongs to  $U_{h,\varepsilon}^{ad}$ , where (4.20) is satisfied. The optimality of  $u_h^\varepsilon$  is proved in the intersection of  $U_h^{ad}$  with a ball of radius  $\varepsilon/2$  around  $u_h^\varepsilon$ . This is local optimality. ■

One can also show that  $u_h^\varepsilon$  is the unique local solution of  $(P_h)$  in a certain neighborhood of  $\bar{u}$ . However, we do not discuss this here. In what follows, let us fix  $(P_h^\varepsilon)$  by  $\varepsilon = \min(\bar{\varepsilon}, \varepsilon_\tau)$  and put  $\bar{u}_h := u_h^\varepsilon$ . In this way, a sequence of local approximate solutions  $(\bar{u}_h)_h$  is found that tends to  $\bar{u}$  as  $h \downarrow 0$ . In the next section we estimate the error  $\|\bar{u}_h - \bar{u}\|$ .

## 5. FEM-approximation of $(P)$ : Error-estimates for local solutions

In this section, we prove the error estimates for local approximate solutions in the norms of  $L^2$  and  $L^\infty$ . As outlined in the preceding subsection, we start our investigations by the sequence  $(\bar{u}_h)_{h>0}$  of local solutions for  $(P_h)$ ,  $h > 0$ , converging to the fixed local reference solution  $\bar{u}$  of  $(P)$  that satisfies  $(SSC)$ .

To perform our analysis, we need an element  $u_h$  admissible for  $(P_h)$  so that it can serve as a test function in the variational inequality and has an optimal distance  $O(h)$  to  $\bar{u}$ . The idea is to take  $u_h = \text{Proj}_{[\alpha,\beta]}(-\frac{1}{\nu}\Pi_h\varphi(\bar{u}))$ . This element is admissible and close to  $\bar{u}$ , but we cannot expect that  $\bar{u}_h - u_h$  be in the  $\tau$ -critical cone where our second-order sufficient condition holds. To overcome this difficulty, we apply a splitting  $\bar{u}_h - u_h = e_h + d_h$ , where

$$\begin{aligned} e_h &= \begin{cases} 0 & \text{on } \Omega \setminus \Omega_h \\ \bar{u}_h - u_h & \text{on } (A_\tau \cup A^i) \cap \Omega_h \\ \bar{u}_h - \bar{u} & \text{on } \Omega_h \setminus (A_\tau \cup A^i), \end{cases} \\ d_h &= \begin{cases} \bar{u}_h - u_h & \text{on } \Omega \setminus \Omega_h \\ 0 & \text{on } (A_\tau \cup A^i) \cap \Omega_h \end{cases} \end{aligned}$$

Here,  $A^i$  denotes the inactive set of  $\bar{u}$ , i.e.  $A^i = \{x \in \Omega \mid \alpha < \bar{u}(x) < \beta\}$ . We have taken  $e_h = 0$  outside  $\Omega_h$  to apply later Lemma 5.2.

Then  $e_h$  belongs to the  $\tau$ -critical cone for all sufficiently small  $h$ :

If  $h$  is small, then for all  $T$  with  $T \cap A_\tau \neq \emptyset$  we know that  $\bar{u}(x) \equiv \alpha$  or  $\bar{u}(x) \equiv \beta$ , hence on  $T$  also  $u_h = \Pi_h \bar{u} \equiv \bar{u}(x)$  holds. Moreover, Lemma 4.2 yields  $\bar{u}_h = \bar{u}(x)$  on  $T$ . Therefore,  $e_h = 0$  is true on  $A_\tau$ . On  $A^i$ , the  $\tau$ -critical cone does not restrict the functions. On the remaining set  $\Omega_h \setminus (A_\tau \cup A^i)$ , the function  $\bar{u}$  is active, while  $\bar{u}_h$  belongs to  $U^{ad}$ . This ensures that the difference  $\bar{u}_h - \bar{u}$  has the right sign required by the  $\tau$ -critical cone.

The part  $d_h$  can be estimated by the optimal order  $\|d_h\|_2 \leq O(h)$ . Notice that  $|\Omega \setminus \Omega_h| \leq ch$  and  $d_h$  is uniformly bounded. The part  $\|e_h\|_2$  must be estimated yet.

REMARK: In the case  $A_\tau = \Omega$ , the  $\tau$ -critical cone consists of the zero element. Here, the second-order condition (SSC) is trivially satisfied and does not contribute to the error estimation. However, in this case, the continuity of the function  $\bar{\varphi} + \nu \bar{u}$  implies that the sign is constant and then  $\bar{\varphi} + \nu \bar{u} \geq \tau$  in  $\Omega$  or, conversely,  $\bar{\varphi} + \nu \bar{u} \leq -\tau$  in  $\Omega$ . In the first case, (3.2) implies that  $\bar{u} \equiv \alpha$  in  $\Omega$ . In the second case, the identity  $\bar{u} \equiv \beta$  in  $\Omega$  holds. On the other hand, the uniform convergence  $\bar{\varphi}_h + \nu \bar{u}_h \rightarrow \bar{\varphi} + \nu \bar{u}$  implies that  $\bar{\varphi}_h + \nu \bar{u}_h$  has the same sign as  $\bar{\varphi} + \nu \bar{u}$  for every  $h$  small enough. Then (4.4) leads to  $\bar{u}_h \equiv u_h \equiv \bar{u}$  in  $\Omega$  for every  $h$  small enough. Consequently, also  $e_h = d_h = 0$  holds true for  $h$  small.

The following auxiliary statements express important properties of  $J''$  and  $J''_h$ , which are more or less intuitively clear. For their proofs we refer to Arada et al. (2002). First, since  $e_h$  belongs to the  $\tau$ -critical cone for sufficiently small  $h$ , we obtain:

LEMMA 5.1 It holds

$$J''(\bar{u})(e_h)^2 \geq \delta \|e_h\|_2^2$$

for all sufficiently small  $h$ .

The next result concerns the approximation of  $J''$  by  $J''_h$ .

LEMMA 5.2 Suppose that  $w$  belongs to  $U_h^{ad}$ . Then

$$|J''(w)v^2 - J''_h(w)v^2| \leq Ch^\lambda \|v\|_2^2$$

holds for all  $v \in L^2(\Omega)$  vanishing on  $\Omega \setminus \Omega_h$ , where the constant  $C = C(\Omega, n)$  does not depend on  $v$  and  $h$ .

LEMMA 5.3 For all sufficiently small  $h > 0$ ,

$$J''_h(\bar{u})(e_h)^2 \geq \frac{\delta}{2} \|e_h\|_2^2$$

*Proof.* This is a direct consequence of Lemma 5.1 and Lemma 5.2.  $\blacksquare$

Moreover,  $J_h''(w)$  is in some sense Lipschitz with respect to  $w$ :

LEMMA 5.4 *Let  $w_1$  and  $w_2$  belong to  $U^{ad}$ . Then*

$$|J_h''(w_1)v^2 - J_h''(w_2)v^2| \leq C(\|w_1 - w_2\|_\infty + h^\lambda)\|v\|_2^2 \quad (5.1)$$

is satisfied for all  $v \in L^2(\Omega)$  with a constant  $C = C(\Omega, n)$  independent of  $v$  and  $h$ .

The term  $h^\lambda$  in (5.1) can be avoided, if the so-called discrete maximum principle holds for the finite-element approximation of (1.1).

By (4.4) the approximate local solution  $\bar{u}_h$  satisfies

$$\int_{\Omega_h} (\varphi_h(\bar{u}_h) + \nu \bar{u}_h)(v - \bar{u}_h)(x) dx \geq 0 \quad \forall v \in U_h^{ad}.$$

The auxiliary control  $u_h$  will not fulfill the analogous inequality

$$\int_{\Omega_h} (\varphi_h(u_h) + \nu u_h)(v - u_h)(x) dx \geq 0 \quad \forall v \in U_h^{ad}.$$

Instead of this, we are able to show that  $u_h$  satisfies an associated perturbed variational inequality with perturbation  $\zeta_h$ , namely

$$\int_{\Omega_h} (\varphi_h(u_h) + \nu u_h + \zeta_h)(v - u_h)(x) dx \geq 0 \quad \forall v \in U_h^{ad}. \quad (5.2)$$

To this aim, we introduce  $\zeta_h \in U_h$  by

$$\zeta_h|_T = \begin{cases} \left\{ -\frac{1}{|T|} \int_T (\varphi_h(\bar{u}_h) + \nu \bar{u}_h) dx \right\}^+ & \text{if } u_h|_T = \alpha, \\ -\left\{ \frac{1}{|T|} \int_T (\varphi_h(\bar{u}_h) + \nu \bar{u}_h) dx \right\}^+ & \text{if } u_h|_T = \beta, \\ -\frac{1}{|T|} \int_T (\varphi_h(\bar{u}_h) + \nu \bar{u}_h) dx & \text{otherwise,} \end{cases}$$

for all  $T \in \mathcal{T}_h$ . We extend  $\zeta_h$  up to the boundary of  $\Omega$  analogously to the definition of the controls in  $U_h$ . As we shall verify below, the function  $\zeta_h$  is constructed such that the auxiliary function  $u_h$  satisfies the first-order necessary optimality condition of the problem

$$\min J_h(v) + \int_{\Omega_h} \zeta_h v dx, \quad v \in U_h^{ad}, \quad (5.3)$$

which is a perturbation of  $(P_h)$  by the linear functional  $(\zeta_h, v)$ . We have adopted the idea to work with this type of perturbation from Malanowski, Büskens and Maurer (1997). It was introduced there for the optimal control of ODEs and can be transferred to our case. We shall not exactly follow that method, but

LEMMA 5.5 *The auxiliary control  $u_h$  satisfies the variational inequality (5.2).*

*Proof.* How can we define  $\zeta_h$  to fulfill (5.2)? Select an arbitrary  $T \in \mathcal{T}_h$ . First, observe that (5.2) can be equivalently written as

$$\left( \int_T (\varphi_h(u_h) + \nu u_h) dx + |T| \zeta_h|_T \right) (v - u_h|_T) \geq 0 \quad (5.4)$$

for all  $T \in \mathcal{T}_h$  and all  $v \in [\alpha, \beta]$ .

(i) If  $u_h|_T = \alpha$  then  $v - u_h|_T \geq 0$  holds in (5.4) for all  $v \in [\alpha, \beta]$ . Therefore,  $\zeta_h$  must be chosen such that  $\int_T (\varphi_h(u_h) + \nu u_h) dx + |T| \zeta_h|_T \geq 0$  holds. Obviously,

$$|T| \zeta_h = \left( \int_T (\varphi_h(u_h) + \nu u_h) dx \right)^- = \left( - \int_T (\varphi_h(u_h) + \nu u_h) dx \right)^+$$

meets that requirement.

(ii) If  $u_h|_T = \beta$ , then  $v - u_h|_T \leq 0$ , and  $\int_T (\varphi_h(u_h) + \nu u_h) dx + |T| \zeta_h|_T \leq 0$  must hold. This is accomplished by

$$|T| \zeta_h = \left( - \int_T (\varphi_h(u_h) + \nu u_h) dx \right)^+.$$

(iii) If  $\alpha < u_h|_T < \beta$ , then  $v - u_h|_T$  can be positive or negative, hence  $\zeta_h$  must be taken such that  $\int_T (\varphi_h(u_h) + \nu u_h) dx + |T| \zeta_h|_T = 0$ . We have found the function  $\zeta_h$  as defined above. ■

LEMMA 5.6 *There exists a positive constant  $C$ , independent of  $h$ , such that*

$$\|\zeta_h\|_2 \leq Ch. \quad (5.5)$$

For the proof, the reader is referred to Arada et al. (2002).

THEOREM 5.1 *For all sufficiently small  $h > 0$*

$$\|\bar{u} - \bar{u}_h\|_2 \leq Ch,$$

*holds with a positive constant  $C$  independent of  $h$ .*

*Proof.* From the optimality conditions for the problem  $(P_h)$ , and since  $u_h$  satisfies the optimality conditions of (5.3), we deduce that

$$J'_h(\bar{u}_h)(u_h - \bar{u}_h) \geq 0 \quad \text{and} \quad J'_h(u_h)(\bar{u}_h - u_h) + \int_{\Omega_h} \zeta_h(\bar{u}_h - u_h) dx \geq 0.$$

Therefore,

$$(J'_h(\bar{u}_h) - J'_h(u_h))(\bar{u}_h - u_h) \leq \int_{\Omega_h} \zeta_h(\bar{u}_h - u_h) dx$$

On the other hand, we have

$$\begin{aligned} (J'_h(\bar{u}_h) - J'_h(u_h))(\bar{u}_h - u_h) &= J''_h((1-\theta)\bar{u}_h + \theta u_h)(\bar{u}_h - u_h)^2 \\ &= J''_h(\bar{u})(\bar{u}_h - u_h)^2 + (J''_h((1-\theta)\bar{u}_h + \theta u_h) - J''_h(\bar{u}))(\bar{u}_h - u_h)^2 \\ &= I_1 + I_2, \end{aligned}$$

with some  $\theta \in (0, 1)$ . Now we estimate  $I_1$  and  $I_2$  separately and apply the splitting  $\bar{u}_h \perp u_h = e_h + d_h$  introduced at the beginning of this section. In view of Lemma 5.3 and the Young inequality we obtain for sufficiently small  $h$

$$\begin{aligned} I_1 &= J''_h(\bar{u})(e_h + d_h)^2 = J''_h(\bar{u})e_h^2 + 2J''_h(\bar{u})(e_h, d_h) + J''_h(\bar{u})d_h^2 \\ &\geq \frac{\delta}{2}\|e_h\|_2^2 - c\|e_h\|_2\|d_h\|_2 - c\|d_h\|_2^2 \\ &\geq \frac{\delta}{3}\|e_h\|_2^2 - c\|d_h\|_2^2 = \frac{\delta}{3}\|e_h + d_h - d_h\|_2^2 - c\|d_h\|_2^2 \\ &\geq \frac{\delta}{3}\|\bar{u}_h - u_h\|_2^2 - \frac{2\delta}{3}\|\bar{u}_h - u_h\|_2\|d_h\|_2 - c\|d_h\|_2^2 \\ &\geq \frac{\delta}{4}\|\bar{u}_h - u_h\|_2^2 - c\|d_h\|_2^2. \end{aligned}$$

For  $I_2$  we obtain by Lemma 5.4

$$|I_2| = |J''_h((1-\theta)\bar{u}_h + \theta u_h) - J''_h(\bar{u})(\bar{u}_h - u_h)^2| \leq \frac{\delta}{8}\|\bar{u}_h - u_h\|_2^2$$

for all sufficiently small  $h$ , since  $\bar{u}_h \rightarrow \bar{u}$  and  $u_h = \Pi_h \bar{u} \rightarrow \bar{u}$  as  $h \downarrow 0$ . Summarizing up, we have

$$I_1 + I_2 \geq \frac{\delta}{8}\|\bar{u}_h - u_h\|_2^2 - c\|d_h\|_2^2 \geq \frac{\delta}{8}\|\bar{u}_h - u_h\|_2^2 - ch^2,$$

hence (5.6) yields

$$\|\zeta_h\|_2\|u_h - \bar{u}_h\|_2 \geq \frac{\delta}{8}\|\bar{u}_h - u_h\|_2^2 - ch^2.$$

By the Young inequality

$$\|\zeta_h\|_2\|u_h - \bar{u}_h\|_2 \leq \delta/16\|u_h - \bar{u}_h\|_2^2 + c\|\zeta_h\|_2^2$$

is obtained. Now from the estimate (5.5),

$$ch^2 \geq \frac{\delta}{16}\|u_h - \bar{u}_h\|_2^2,$$

follows, hence  $\|u_h - \bar{u}_h\|_2 \leq ch$ . This, together with  $\|u_h - \bar{u}\|_2 = \|\Pi_h \bar{u} - \bar{u}\|_2 \leq ch$ , gives the desired estimate  $\|\bar{u}_h - \bar{u}\|_2 \leq ch$ . ■

Now it is an easy task to improve this  $L^2$ -estimate by the one in  $L^\infty$ . Here,



THEOREM 5.2 *The estimate*

$$\|\bar{u} - \bar{u}_h\|_\infty \leq C h^\lambda$$

holds for all sufficiently small  $h$ . Here,  $C$  is a positive constant independent of  $h$ ,  $\lambda = 1$  if  $n = 2$  or if  $n = 3$  and the triangulation is of nonnegative type, and  $\lambda = 1/2$  otherwise.

*Proof.* Invoking Theorem 4.2 and the projection formulas (3.4), (4.6) we get

$$\begin{aligned} \|\bar{u} - \bar{u}_h\|_\infty &= \|\text{Proj}_{[\alpha,\beta]}(-\frac{1}{\nu}\varphi(\bar{u})) - \text{Proj}_{[\alpha,\beta]}(-\frac{1}{\nu}\Pi_h\varphi_h(\bar{u}_h))\|_\infty \\ &\leq C(h + \|\varphi(\bar{u}) - \varphi_h(\bar{u}_h)\|_\infty) \leq C(h + \|\bar{u} - \bar{u}_h\|_2 + h^\lambda). \end{aligned}$$

Therefore we obtain

$$\|\bar{u} - \bar{u}_h\|_\infty \leq C(h^\lambda + \|\bar{u} - \bar{u}_h\|_2).$$

The conclusion follows from Theorem 5.1. ■

## References

- ARADA, N., CASAS, E. and TRÖLTZSCH, F. (2002) Error estimates for the numerical approximation of a semilinear elliptic control problem. *Computational Optimization and Appl.*, **23**, 201–229.
- ARNAUTU, V. and NEITTAANMÄKI, P. (1998) Discretization estimates for an elliptic control problem. *Numer. Funct. Anal. and Optimization*, **19**, 431–464.
- BONNANS, J.F. and CASAS, E. (1995) An extension of Pontryagin's principle for state constrained optimal control of semilinear elliptic equation and variational inequalities. *SIAM J. Control and Optimization*, **33**, 274–298.
- CASAS and MATEOS, E.M. (2001) Uniform convergence of the FEM. Applications to state constrained control problems. *Comp. Appl. Math.*, to appear.
- CASAS, E. and TRÖLTZSCH, F. (2000) Second order necessary and sufficient optimality conditions for optimization problems and applications to control theory. *SIAM J. on Optimization*, to appear (2002).
- CASAS, E., TRÖLTZSCH, F. and UNGER, A. (2000) Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control and Optimization*, **38**, 1369–1391.
- DONTCHEV, A.L., HAGER, W.W., POORE, A.B. and YANG, B. (1995) Optimality, stability, and convergence in nonlinear control. *Appl. Math. Optimization*, **31**, 297–326.
- FALK R. (1973) Approximation of a class of optimal control problems with order of convergence estimates. *J. Math. Anal. Appl.*, **44**, 28–47.
- GEVECI, T. (1979) On the approximation of the solution of an optimal control problem governed by an elliptic equation. *R.A.I.R.O. Analyse*

- MALANOWSKI, K., BÜSKENS, C. and MAURER, H. (1997) Convergence of approximations to nonlinear control problems. In: *Mathematical Programming with Data Perturbation*, A. V. Fiacco, ed., Marcel Dekker, Inc., New York, 253–284.
- MATEOS, M. (2000) Problemas de control óptimo gobernados por ecuaciones semilineales con restricciones de tipo integral sobre el gradiente del estado. *Thesis*, University of Cantabria, Santander.