ESAIM: COCV 24 (2018) 1181–1206
https://doi.org/10.1051/cocv/2017031

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

A MORLEY FINITE ELEMENT METHOD FOR AN ELLIPTIC DISTRIBUTED OPTIMAL CONTROL PROBLEM WITH POINTWISE STATE AND CONTROL CONSTRAINTS *

Susanne C. Brenner[1,a], Thirupathi Gudi[2], Kamana Porwal[3] and Li-yeng Sung[3]

Abstract. We design and analyze a Morley finite element method for an elliptic distributed optimal control problem with pointwise state and control constraints on convex polygonal domains. It is based on the formulation of the optimal control problem as a fourth order variational inequality. Numerical results that illustrate the performance of the method are also presented.

Mathematics Subject Classification. 49J20, 65K15, 65N30.

Received September 19, 2016. Revised March 22, 2017. Accepted April 17, 2017.

1. Introduction

Let $\Omega \in \mathbb{R}^2$ be a convex polygon, $y_d \in L_2(\Omega)$, and β be a positive constant. We will consider the following elliptic distributed optimal control problem:

$$\text{Find} \quad (\bar{y}, \bar{u}) = \underset{(y,u)\in\mathbb{K}}{\operatorname{argmin}} \left[\frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)} \right], \tag{1.1}$$

where (y, u) belongs to $\mathbb{K} \subset H^1_0(\Omega) \times L_2(\Omega)$ if and only if

$$\int_\Omega \nabla y \cdot \nabla z \, \mathrm{d}x = \int_\Omega uz \, \mathrm{d}x \qquad \forall\, z \in H^1_0(\Omega), \tag{1.2}$$

$$\psi_1 \le y \le \psi_2 \qquad \text{on} \qquad \Omega, \tag{1.3}$$

$$\phi_1 \le u \le \phi_2 \qquad \text{on} \qquad \Omega. \tag{1.4}$$

Keywords and phrases. Elliptic distributed optimal control problem, pointwise state and control constraints, fourth order variational inequality, Morley element.

[*] *The work of the first and fourth authors was supported in part by the National Science Foundation under Grant Nos. DMS-13-19172 and Grant No. DMS-16-20273. The second author acknowledges the financial support from the DST Fast Track Project DST01302.*

[1] Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803 and Indian Institute of Science, Bangalore 560 01, India.
[a] Corresponding author: brenner@math.lsu.edu

[2] Thirupathi Gudi, Department of Mathematics, Indian Institute of Science, Bangalore 560 01, India.
gudi@math.iisc.ernet.in

[3] Department of Mathematics and Center for Computation and Technology Louisiana State University, Baton Rouge, LA 70803.
kporwal@math.lsu.edu; kamana.porwal@iitgn.ac.in; sung@math.lsu.edu

We assume that the functions ψ_1, ψ_2, ϕ_1 and ϕ_2 satisfy (i) $\psi_1, \psi_2 \in W^{2,\infty}(\Omega)$, (ii) $\psi_1 < \psi_2$ on $\bar{\Omega}$, (iii) $\psi_1 < 0 < \psi_2$ on $\partial\Omega$, (iv) $\phi_1, \phi_2 \in W^{1,\infty}(\Omega)$ and (v) $\phi_1 < \phi_2$ on $\bar{\Omega}$. Here and below we follow the standard notation for differential operators, Sobolev spaces and norms that can be found for example in [1, 11, 21], and the inequalities involving L_2 functions are to be interpreted in the almost everywhere sense.

The literature on finite element methods for elliptic distributed optimal control problem with pointwise state and control constraints is quite small. We refer the readers to [37, 39] for an approach that is based on the first order optimality conditions for a reduced minimization problem involving only the control u, and to [18, 19] for an approach that is based on a regularization of the constraints.

Here we follow a different approach that is based on a minimization problem involving only the state y. Since Ω is convex, we have $y \in H^2(\Omega)$ by elliptic regularity [22, 29], and hence the optimal control problem (1.1)–(1.4) can be reformulated as

$$\begin{aligned} \text{Find} \quad \bar{y} &= \underset{y\in K}{\operatorname{argmin}} \left[\frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|\Delta y\|^2_{L_2(\Omega)}\right] \\ &= \underset{y\in K}{\operatorname{argmin}} \left[\frac{1}{2}a(y,y) - (y_d, y)\right], \end{aligned} \tag{1.5}$$

where

$$K = \left\{y \in H^2(\Omega)\cap H^1_0(\Omega) : \psi_1 \le y \le \psi_2 \quad \text{and} \quad \phi_1 \le -\Delta y \le \phi_2 \qquad \text{in} \qquad \Omega\right\}, \tag{1.6}$$

$$a(y,z) = \beta\int_\Omega (\Delta y)(\Delta z)\mathrm{d}x + \int_\Omega yz\,\mathrm{d}x, \tag{1.7}$$

and $(y,z) = \int_\Omega yz\,\mathrm{d}x$.

Remark 1.1. Since the identity

$$\int_\Omega (\Delta y)(\Delta z)\,\mathrm{d}x = \sum_{i,j=1}^{2}\int_\Omega \left(\frac{\partial^2 y}{\partial x_i \partial x_j}\right)\left(\frac{\partial^2 z}{\partial x_i \partial x_j}\right)\mathrm{d}x$$

holds for $y, z \in H^2(\Omega)\cap H^1_0(\Omega)$ (*cf.* [30], Lem. 2.2.2), the bilinear form $a(\cdot,\cdot)$ in (1.7) has an alternative expression given by

$$a(y,z) = \beta\sum_{i,j=2}^{2}\int_\Omega \left(\frac{\partial^2 y}{\partial x_i \partial x_j}\right)\left(\frac{\partial^2 z}{\partial x_i \partial x_j}\right)\mathrm{d}x + \int_\Omega yz\,\mathrm{d}x. \tag{1.8}$$

We assume the following Slater condition:

$$\begin{aligned} &\text{There exists } y \in H^2(\Omega)\cap H^1_0(\Omega) \text{ such that (i) } \psi_1 < y < \psi_2 \text{ in } \Omega, \text{ and} \\ &\text{(ii) } u = -\Delta y \text{ satisfies the constraint (1.4).} \end{aligned} \tag{1.9}$$

Note that (1.9) implies in particular that the closed convex set K is nonempty. It then follows from the classical theory of calculus of variations [24, 34] that (1.5)–(1.7) has a unique solution $\bar{y} \in K$ characterized by the fourth order variational inequality

$$\beta\int_\Omega (-\Delta\bar{y})(-\Delta(y-\bar{y}))\,\mathrm{d}x + \int_\Omega (\bar{y}-y_d)(y-\bar{y})\,\mathrm{d}x \ge 0 \qquad \forall\, y \in K. \tag{1.10}$$

In this paper we will develop and analyze a finite element method for (1.5)–(1.7) (or equivalently (1.10)) that is based on the Morley element [38]. This approach of solving elliptic optimal control problems as fourth order variational inequalities was introduced in [28, 35], where a Morley finite element method and a mixed finite element method for state constrained problems were analyzed under additional assumptions on the active sets from [5]. C^0 interior penalty methods and partition of unity methods for two dimensional state constrained problems were studied later in [8, 13] without additional assumptions on the active sets. A new convergence analysis for state constrained optimal control problems was presented in the recent paper [12] that simplifies the analysis in [8, 13] and is also applicable to the finite element methods in [10] for three dimensional domains. The current paper extends the results in [12, 35] to the problem defined by (1.1)–(1.4) that involves both state and control constraints.

The rest of the paper is organized as follows. We collect the results for the variational inequality (1.10) in Section 2 and introduce the Morley finite element method in Section 3. The convergence analysis is carried out in Section 4, followed by numerical results in Section 5. We end the paper with some concluding remarks in Section 6. Appendix A contains the construction and analysis of an operator that plays a key role in the convergence analysis.

2. Results for the continuous problem

Under the Slater condition (1.9), we have (*cf.* [36], Thm. 9.4.1 and [33], Thm. 1.6) the following (generalized) Karush-Kuhn-Tucker conditions for (1.10):

$$\beta \int_\Omega (-\Delta \bar{y})(-\Delta z)\mathrm{d}x + \int_\Omega (\bar{y} - y_d) z \,\mathrm{d}x = \int_\Omega \lambda(-\Delta z)\mathrm{d}x + \int_\Omega z \,\mathrm{d}\mu \tag{2.1}$$

for all $z \in H^2(\Omega) \cap H_0^1(\Omega)$, where $\lambda \in L_2(\Omega)$ and $\mu \in \mathcal{M}(\Omega)$ (the space of regular Borel measures) such that

$$\lambda \geq 0 \qquad \text{if} \quad -\Delta \bar{y} = \phi_1, \tag{2.2}$$
$$\lambda \leq 0 \qquad \text{if} \quad -\Delta \bar{y} = \phi_2, \tag{2.3}$$
$$\lambda = 0 \qquad \text{otherwise}, \tag{2.4}$$
$$\mu \geq 0 \qquad \text{if} \quad \bar{y} = \psi_1, \tag{2.5}$$
$$\mu \leq 0 \qquad \text{if} \quad \bar{y} = \psi_2, \tag{2.6}$$
$$\mu = 0 \qquad \text{otherwise}. \tag{2.7}$$

Remark 2.1. Since $\psi_1 < 0 < \psi_2$ on $\partial\Omega$, the conditions (2.5)–(2.7) imply that the support of μ is a compact subset of Ω and hence μ is a bounded measure.

2.1. The adjoint state

Let the adjoint state $p \in L_2(\Omega)$ be defined by

$$\int_\Omega p(-\Delta z)\,\mathrm{d}x = \int_\Omega (\bar{y} - y_d) z \,\mathrm{d}x - \int_\Omega z \,\mathrm{d}\mu \tag{2.8}$$

for all $z \in H^2(\Omega) \cap H_0^1(\Omega)$. Then we have (*cf.* [16], Thm. 1)

$$p \in W_0^{1,s}(\Omega) \qquad \forall\, s < 2. \tag{2.9}$$

Remark 2.2. Since the support of μ is disjoint from $\partial\Omega$ (*cf.* Rem. 2.1), the adjoint state p actually belongs to H^1 in a neighborhood of $\partial\Omega$.

It follows from (2.1) and (2.8) that

$$\lambda = \beta(-\Delta\bar{y}) + p, \tag{2.10}$$

which together with (2.2)–(2.4) implies that $\bar{y} \in K$ satisfies

$$-\Delta\bar{y} = P_{[\phi_1,\phi_2]}(-p/\beta), \tag{2.11}$$

where

$$P_{[\phi_1,\phi_2]}v = \max(\phi_1, \min(\phi_2, v))$$

is the orthogonal projection from $L_2(\Omega)$ onto the closed convex subset $\{v \in L_2(\Omega) : \phi_1 \le v \le \phi_2\}$.

2.2. Regularity of $\bar{u}$

From ([17], Lem. 3.3) we have

$$P_{[\phi_1,\phi_2]}(-p/\beta) \in H^1(\Omega) \cap L_\infty(\Omega), \tag{2.12}$$

which together with (2.11) implies that

$$\bar{u} = -\Delta\bar{y} \in H^1(\Omega) \cap L_\infty(\Omega).$$

2.3. Regularity of λ

It follows from (2.9)–(2.12) that

$$\lambda \in W^{1,s}(\Omega) \qquad \text{for any} \quad s < 2. \tag{2.13}$$

Remark 2.3. In view of Remark 2.2, λ belongs to H^1 in a neighborhood of $\partial\Omega$.

2.4. Regularity of $\bar{y}$

According to (2.11), the solution $\bar{y} \in K \subset H^2(\Omega) \cap H_0^1(\Omega)$ of (1.10) also satisfies

$$-\Delta\bar{y} = P_{[\phi_1,\phi_2]}(-p/\beta) \qquad \text{in} \qquad \Omega, \tag{2.14a}$$

$$\bar{y} = 0 \qquad \text{on} \qquad \partial\Omega. \tag{2.14b}$$

2.4.1. Interior regularity of $\bar{y}$

It follows from (2.12), (2.14) and interior elliptic regularity (*cf.* [25], Thm. 6.2 and [32], Lem. 17.1.1) that

$$\bar{y} \in H^3_{\mathrm{loc}}(\Omega) \cap W^{2,s}_{\mathrm{loc}}(\Omega) \tag{2.15}$$

for any $s \in (1, \infty)$.

2.4.2. Global regularity of $\bar{y}$

It follows from (2.12), (2.14), and the elliptic regularity theory on polygonal domains [22, 29] that

$$\bar{y} \in H^{2+\alpha}(\Omega) \tag{2.16}$$

for some $\alpha \in (0, 1]$, where the index of elliptic regularity α is determined by the angles at the corners of Ω.

2.5. Regularity of μ

It follows from (2.1), Remark 2.1, (2.13), (2.15) and integration by parts that

$$\mu \in W^{-1,s}(\Omega) \tag{2.17}$$

for all $s < 2$.

2.6. Results under additional assumptions

We can obtain stronger results for $\bar{y}$, $\bar{u}$, λ and μ under some additional assumptions.

2.6.1. Additional assumption on λ and μ

If we assume

$$\operatorname{supp}\lambda \cap \operatorname{supp}\mu = \emptyset, \tag{2.18}$$

then it follows from (2.1), Remark 2.1, (2.15) and integration by parts that

$$\mu \in H^{-1}(\Omega) = [H_0^1(\Omega)]', \tag{2.19}$$

which together with (2.8) implies

$$p \in H_0^1(\Omega). \tag{2.20}$$

Combining (2.10)–(2.12), and (2.20), we have

$$\lambda \in H^1(\Omega). \tag{2.21}$$

Moreover we can apply the result in [26] to conclude that

$$\bar{y} \text{ belongs to } W^{2,\infty}(G) \text{ in a neighborhood } G \text{ of } \operatorname{supp}\mu. \tag{2.22}$$

2.6.2. Additional assumption on ϕ_1 and ϕ_2

If we assume

$$\phi_1 \leq 0 \leq \phi_2 \qquad \text{in a neighborhood of } \partial\Omega, \tag{2.23}$$

then

$$P_{[\phi_1,\phi_2]}(-p/\beta) \in H_0^1(\Omega) \tag{2.24}$$

because p belongs to H^1 near $\partial\Omega$ (*cf.* Rem. 2.2) and p vanishes on $\partial\Omega$.

It follows from (2.11) and (2.24) that

$$\bar{u} = -\Delta\bar{y} \in H_0^1(\Omega).$$

Remark 2.4. The stronger condition (2.24) on the right-hand side of (2.14a) yields a better regularity for $\bar{y}$ (*cf.* [29], Sect. 5.1). In the case of a rectangular Ω, we have $\bar{y} \in H^3(\Omega)$ under (2.24) and $\bar{y} \in H^{3-\epsilon}(\Omega)$ for any $\epsilon > 0$ under (2.12).

Remark 2.5. Under both additional Assumptions (2.18) and (2.23), the Lagrange multiplier λ belongs to $H_0^1(\Omega)$ by (2.10), (2.11), (2.20) and (2.24).

Remark 2.6. The results in this section are also valid for the case where there is one (or no) state constraint.

3. The Morley finite element method

Let $\mathcal{T}_h$ be a quasi-uniform simplicial triangulation of Ω, where the mesh parameter h is proportional to $\max_{T\in\mathcal{T}_h} \operatorname{diam} T$. We denote by I_h the nodal interpolation for the conforming P_1 finite element space associated with $\mathcal{T}_h$, and denote by Q_h the orthogonal projection operator from $L_2(\Omega)$ onto the space of piecewise constant functions associated with $\mathcal{T}_h$.

The Morley finite element space V_h is defined as follows. A function v belongs to V_h if and only if (i) v is a quadratic polynomial on each $T \in \mathcal{T}_h$, (ii) v is continuous at the vertices of $\mathcal{T}_h$, (iii) the normal derivative of v is continuous at the midpoints of the edges of $\mathcal{T}_h$, and (iv) v vanishes at the vertices of $\mathcal{T}_h$ that belong to $\partial\Omega$.

A function in V_h is determined by its values at the vertices of $\mathcal{T}_h$ and the integrals of its normal derivatives over the edges of $\mathcal{T}_h$. (The degrees of freedom for the Morley finite element are depicted in Fig. A.1 in Appendix A).

The Morley finite element method is to find

$$\bar{y}_h = \underset{y\in K_h}{\operatorname{argmin}} \left[\frac{1}{2}a_h(y,y) - (y_d, y)\right] \tag{3.1}$$

where

$$K_h = \{y \in V_h : I_h\psi_1 \le I_h y \le I_h\psi_2 \quad \text{and} \quad Q_h\phi_1 \le -\Delta_h y \le Q_h\phi_2\}, \tag{3.2}$$

$$a_h(y,z) = \beta \sum_{T\in\mathcal{T}_h} \int_T D^2y : D^2z \,\mathrm{d}x + \int_\Omega yz \,\mathrm{d}x, \tag{3.3}$$

Δ_h is the piecewise defined Laplacian operator, and

$$D^2y : D^2z = \sum_{i,j=1}^{2} \left(\partial^2 y/\partial x_i \partial x_j\right)\left(\partial^2 z/\partial x_i \partial x_j\right)$$

is the Frobenius inner product of the Hessian matrices D^2y and D^2z.

Remark 3.1. We use the piecewise version of the alternative expression (1.8) for $a(\cdot,\cdot)$ because the Morley finite element method is nonconforming (*cf.* [21], Rem. 6.2.1).

The error analysis for the Morley method will be carried out in terms of the mesh-dependent energy norm $\|\cdot\|_h$ defined by

$$\|y\|_h^2 = a_h(y,y) = \beta \sum_{T\in\mathcal{T}_h} |y|^2_{H^2(T)} + \|y\|^2_{L_2(\Omega)}. \tag{3.4}$$

The convergence analysis in Section 4 requires certain operators that connect the continuous space and the discrete space.

3.1. The interpolation operator Π_h

The interpolation operator $\Pi_h : H^2(\Omega) \cap H^1_0(\Omega) \longrightarrow V_h$ is defined by

$$(\Pi_h\zeta)(p) = \zeta(p) \qquad \text{for all vertices } p \text{ of } \mathcal{T}_h, \tag{3.5a}$$

$$\int_e \frac{\partial(\Pi_h\zeta)}{\partial n_e}\mathrm{d}s = \int_e \frac{\partial \zeta}{\partial n_e}\mathrm{d}s \qquad \text{for all edges } e \text{ of } \mathcal{T}_h, \tag{3.5b}$$

where n_e is a unit normal vector of e, and we have a standard interpolation error estimate

$$\sum_{k=0}^{2} h_T^k |\zeta - \Pi_h\zeta|_{H^k(T)} \le C h_T^{2+\alpha} |\zeta|_{H^{2+\alpha}(T)} \qquad \forall T \in \mathcal{T}_h,\ 0 \le \alpha \le 1, \tag{3.6}$$

where h_T is the diameter of T.

Note that (3.6) implies

$$\|\zeta - \Pi_h\zeta\|_h \le Ch^\alpha |\zeta|_{H^{2+\alpha}(\Omega)}. \tag{3.7}$$

It follows from (3.5b) and integration by parts that

$$\int_T \Delta(\Pi_h\zeta)\mathrm{d}x = \sum_{e\in\mathcal{E}_T} \int_e \frac{\partial(\Pi_h\zeta)}{\partial n}\mathrm{d}s = \sum_{e\in\mathcal{E}_T} \int_e \frac{\partial\zeta}{\partial n}\mathrm{d}s = \int_T (\Delta\zeta)\mathrm{d}x \qquad \forall T \in \mathcal{T}_h,$$

where $\mathcal{E}_T$ is the set of the three edges of T, and hence

$$\Delta_h(\Pi_h\zeta) = Q_h(\Delta\zeta) \qquad \forall\,\zeta \in H^2(\Omega)\cap H_0^1(\Omega). \tag{3.8}$$

In view of (1.6), (3.2), (3.5a), and (3.8), we have

$$\Pi_h K \subset K_h. \tag{3.9}$$

3.2. The enriching operator E_h

In the other direction we have an operator $E_h : V_h \longrightarrow H^2(\Omega)\cap H_0^1(\Omega)$ with the following properties:

$$(E_h v)(p) = v(p) \qquad \text{for all vertices } p \text{ of } \mathcal{T}_h, \tag{3.10}$$

$$\int_e \frac{\partial(E_h v)}{\partial n_e}\mathrm{d}s = \int_e \frac{\partial v}{\partial n_e}\mathrm{d}s \qquad \text{for all edges } e \text{ of } \mathcal{T}_h, \tag{3.11}$$

$$\sum_{k=0}^{2} h^{2k} \sum_{T\in\mathcal{T}_h} |v - E_h v|^2_{H^k(T)} \le Ch^4 \sum_{T\in\mathcal{T}_h} |v|^2_{H^2(T)} \qquad \forall\, v \in V_h, \tag{3.12}$$

$$\sum_{k=0}^{2} h^k |\zeta - E_h\Pi_h\zeta|_{H^k(\Omega)} \le Ch^{2+\alpha}|\zeta|_{H^{2+\alpha}(\Omega)} \tag{3.13}$$

for all $\zeta \in H^{2+\alpha}(\Omega)\cap H_0^1(\Omega)$ and $0 \le \alpha \le 1$,

$$\|\zeta - E_h\Pi_h\zeta\|_{W^{1,2/(1-\epsilon)}(\Omega)} \le Ch^{1+\alpha-\epsilon}|\zeta|_{H^{2+\alpha}(\Omega)} \tag{3.14}$$

for all $\zeta \in H^{2+\alpha}(\Omega)\cap H_0^1(\Omega)$ and $0 \le \epsilon < 1$, and

$$|a_h(\Pi_h\zeta, v) - a(\zeta, E_h v)| \le Ch^\alpha |\zeta|_{H^{2+\alpha}(\Omega)}\|v\|_h \tag{3.15}$$

for all $\zeta \in H^{2+\alpha}(\Omega)\cap H_0^1(\Omega)$, $v \in V_h$ and $0 \le \alpha \le 1$.

Details for E_h are given in Appendix A.

Remark 3.2. The estimates (3.13) and (3.14) indicate that the operator $E_h\Pi_h$ behaves like a quasi-local interpolation operator. The estimate (3.15) implies that E_h behaves like the transpose of Π_h with respect to the bilinear forms $a(\cdot,\cdot)$ and $a_h(\cdot,\cdot)$.

Note that (3.11) leads to the following analog of (3.8):

$$Q_h(\Delta E_h v) = \Delta_h v \qquad \forall\, v \in V_h. \tag{3.16}$$

4. Convergence analysis

Since the closed convex set K_h is nonempty by (3.9) and the bilinear form $a_h(\cdot,\cdot)$ is symmetric positive definite, the discrete problem (3.1) has a unique solution characterized by the discrete variational inequality

$$a_h(\bar{y}_h, y - \bar{y}_h) - (y_d, y - \bar{y}_h) \ge 0 \qquad \forall\, y \in K_h. \tag{4.1}$$

We begin with a preliminary estimate. Using (3.4), (3.7), (3.9) and (4.1), we find

$$\begin{aligned}
\|\bar{y} - \bar{y}_h\|_h^2 &\le 2\|\bar{y} - \Pi_h\bar{y}\|_h^2 + 2\|\Pi_h\bar{y} - \bar{y}_h\|_h^2 \\
&\le C_1 h^{2\alpha} + 2a_h(\Pi_h\bar{y} - \bar{y}_h, \Pi_h\bar{y} - \bar{y}_h) \\
&\le C_1 h^{2\alpha} + 2[a_h(\Pi_h\bar{y}, \Pi_h\bar{y} - \bar{y}_h) - (y_d, \Pi_h\bar{y} - \bar{y}_h)].
\end{aligned} \tag{4.2}$$

From here on we will use C (with or without subscript) to denote a (generic) positive constant that is independent of the mesh parameter h. In view of (3.4), (3.12) and (3.15), we have

$$a_h(\Pi_h\bar{y}, \Pi_h\bar{y} - \bar{y}_h) - (y_d, \Pi_h\bar{y} - \bar{y}_h) \le \left[a(\bar{y}, E_h(\Pi_h\bar{y} - \bar{y}_h)) - (y_d, E_h(\Pi_h\bar{y} - \bar{y}_h))\right] + C_2 h^\alpha \|\Pi_h\bar{y} - \bar{y}_h\|_h, \quad (4.3)$$

and it only remains to estimate the first term on the right-hand side of (4.3).

Note that this term only involves the bilinear form $a(\cdot,\cdot)$ and functions in $H^2(\Omega) \cap H_0^1(\Omega)$, and we can use (1.7) and (2.1) to rewrite it as

$$a(\bar{y}, E_h(\Pi_h\bar{y} - \bar{y}_h)) - (y_d, E_h(\Pi_h\bar{y} - \bar{y}_h)) = \int_\Omega \lambda[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x + \int_\Omega E_h(\Pi_h\bar{y} - \bar{y}_h)\mathrm{d}\mu. \quad (4.4)$$

Below we will use ϵ to denote an arbitrary number in $(0, 1)$.

4.1. Discretization error due to the control constraints

The first integral on the right-hand side of (4.4), which measures the effects of the constraint (1.4), can be estimated as follows.

First we can write, by (2.2)–(2.4),

$$\int_\Omega \lambda[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x = \int_\Omega \lambda_1[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x + \int_\Omega \lambda_2[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x, \quad (4.5)$$

where

$$\lambda_1 = \begin{cases} \lambda & \text{if} \quad -\Delta y = \phi_1 \\ 0 & \text{otherwise} \end{cases}, \quad (4.6)$$

$$\lambda_2 = \begin{cases} \lambda & \text{if} \quad -\Delta y = \phi_2 \\ 0 & \text{otherwise} \end{cases}. \quad (4.7)$$

Note that $\lambda_1 = \max(\lambda, 0)$ and $\lambda_2 = \min(\lambda, 0)$. Therefore $\lambda_j \in W^{1,s}(\Omega)$ for any $s < 2$ by (2.13) and ([27], Lem. 7.6). It then follows from a standard interpolation error estimate [11,21] that

$$\|\lambda_j - Q_h\lambda_j\|_{L_2(\Omega)} \le C_\epsilon h^{1-\epsilon} \qquad \text{for } j = 1, 2 \text{ and any positive } \epsilon. \quad (4.8)$$

Remark 4.1. Under the additional Assumption (2.18) we can take ϵ to be 0 in (4.8) (*cf.* (2.21)).

Next we use (4.6) to rewrite the first integral on the right-hand side of (4.5) as the sum of four integrals:

$$\begin{aligned}\int_\Omega \lambda_1[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x = &\int_\Omega \lambda_1[-\Delta(E_h\Pi_h\bar{y} - \bar{y})]\mathrm{d}x + \int_\Omega \lambda_1(\phi_1 - Q_h\phi_1)\mathrm{d}x \\ &+ \int_\Omega \lambda_1 Q_h(\phi_1 + \Delta E_h\bar{y}_h)\mathrm{d}x + \int_\Omega \lambda_1[(\Delta E_h\bar{y}_h) - Q_h(\Delta E_h\bar{y}_h)]\mathrm{d}x. \end{aligned} \quad (4.9)$$

Note that (3.8) and (3.16) imply

$$Q_h\Delta(E_h\Pi_h\bar{y} - \bar{y}) = 0,$$

and hence, by (3.13) and (4.8),

$$\int_\Omega \lambda_1[-\Delta(E_h\Pi_h\bar{y} - \bar{y})]\mathrm{d}x = \int_\Omega (\lambda_1 - Q_h\lambda_1)[-\Delta(E_h\Pi_h\bar{y} - \bar{y})]\mathrm{d}x \le C_\epsilon h^{1+\alpha-\epsilon}. \quad (4.10)$$

Similarly we find

$$\int_\Omega \lambda_1(\phi_1 - Q_h\phi_1)\mathrm{d}x = \int_\Omega (\lambda_1 - Q_h\lambda_1)(\phi_1 - Q_h\phi_1)\mathrm{d}x \le C_\epsilon h^{2-\epsilon}. \tag{4.11}$$

In view of (2.2), (3.2), (3.16) and (4.6), we have

$$\int_\Omega \lambda_1 Q_h(\phi_1 + \Delta E_h\bar{y}_h)\mathrm{d}x = \int_\Omega \lambda_1(Q_h\phi_1 + \Delta_h\bar{y}_h)\mathrm{d}x \le 0. \tag{4.12}$$

The last integral on the right-hand side of (4.9) can be written as

$$\int_\Omega \lambda_1[(\Delta E_h\bar{y}_h) - Q_h(\Delta E_h\bar{y}_h)]\mathrm{d}x = \int_\Omega \lambda_1[\Delta(E_h\bar{y}_h - \bar{y}) - Q_h\Delta(E_h\bar{y}_h - \bar{y})]\mathrm{d}x + \int_\Omega \lambda_1(\Delta\bar{y} - Q_h\Delta\bar{y})\mathrm{d}x, \tag{4.13}$$

and we have, by (4.8) and a standard error estimate for Q_h,

$$\int_\Omega \lambda_1(\Delta\bar{y} - Q_h\Delta\bar{y})\mathrm{d}x = \int_\Omega (\lambda_1 - Q_h\lambda_1)(\Delta\bar{y} - Q_h\Delta\bar{y})\mathrm{d}x \le C_\epsilon h^{1+\alpha-\epsilon} \tag{4.14}$$

since $\Delta\bar{y} \in H^\alpha(\Omega)$.

Finally, by (3.6), (3.12), (3.13) and (4.8), we obtain

$$\begin{aligned}\int_\Omega \lambda_1[\Delta(E_h\bar{y}_h - \bar{y}) - Q_h\Delta(E_h\bar{y}_h - \bar{y})]\mathrm{d}x &= \int_\Omega (\lambda_1 - Q_h\lambda_1)[\Delta(E_h\bar{y}_h - \bar{y})]\mathrm{d}x \\ &\le C_\epsilon h^{1-\epsilon}(|E_h(\bar{y}_h - \Pi_h\bar{y})|_{H^2(\Omega)} + |E_h\Pi_h\bar{y} - \bar{y}|_{H^2(\Omega)}) \\ &\le C_\epsilon(h^{1-\epsilon}\|\Pi_h\bar{y} - \bar{y}_h\|_h + h^{1+\alpha-\epsilon}).\end{aligned} \tag{4.15}$$

Combining (4.9)–(4.15), we arrive at the estimate

$$\int_\Omega \lambda_1[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x \le C_\epsilon(h^{1+\alpha-\epsilon} + h^{1-\epsilon}\|\Pi_h\bar{y} - \bar{y}_h\|_h). \tag{4.16}$$

Similarly we have

$$\int_\Omega \lambda_2[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x \le C_\epsilon(h^{1+\alpha-\epsilon} + h^{1-\epsilon}\|\Pi_h\bar{y} - \bar{y}_h\|_h). \tag{4.17}$$

It follows from (4.5), (4.16) and (4.17) that

$$\int_\Omega \lambda[-\Delta E_h(\Pi_h\bar{y} - \bar{y}_h)]\mathrm{d}x \le C_\epsilon(h^{1+\alpha-\epsilon} + h^{1-\epsilon}\|\Pi_h\bar{y} - \bar{y}_h\|_h). \tag{4.18}$$

4.2. Discretization error due to the state constraints

Next we turn to the second integral on the right-hand side of (4.4) that measures the effects of the constraint (1.3). First we can write, by using (2.5)–(2.7) and (3.10),

$$\begin{aligned}\int_\Omega E_h(\Pi_h\bar{y} - \bar{y}_h)\mathrm{d}\mu &= \int_\Omega (E_h\Pi_h\bar{y} - \bar{y})\mathrm{d}\mu \\ &\quad + \int_\Omega [(\psi_1 - I_h\psi_1) + (I_h\psi_1 - I_h\bar{y}_h) + (I_hE_h\bar{y}_h - E_h\bar{y}_h)]\mathrm{d}\mu_1 \\ &\quad + \int_\Omega [(\psi_2 - I_h\psi_2) + (I_h\psi_2 - I_h\bar{y}_h) + (I_hE_h\bar{y}_h - E_h\bar{y}_h)]\mathrm{d}\mu_2,\end{aligned} \tag{4.19}$$

where

$$\mu_1 = \begin{cases} \mu & \text{if} \quad y = \psi_1 \\ 0 & \text{otherwise,} \end{cases} \tag{4.20}$$

$$\mu_2 = \begin{cases} \mu & \text{if} \quad y = \psi_2 \\ 0 & \text{otherwise.} \end{cases} \tag{4.21}$$

Since the active sets $\{x \in \Omega : \bar{y}(x) = \psi_1(x)\}$ and $\{x \in \Omega : \bar{y}(x) = \psi_2(x)\}$ are disjoint compact subsets of Ω, we have $\mu_j \in W^{-1,s}$ for any $s < 2$ because of (2.17). It follows that

$$\int_\Omega v \, \mathrm{d}\mu_j \leq C_\epsilon \|v\|_{W^{1,2/(1-\epsilon)}(\Omega)} \qquad \text{for } j = 1, 2 \text{ and any } \epsilon \in (0, 1). \tag{4.22}$$

Remark 4.2. Under the additional Assumption (2.18) we can take ϵ to be 0 in (4.22) (*cf.* (2.19)).

According to (2.16), (3.14) and (4.22), the first integral on the right-hand side of (4.19) is bounded by

$$\int_\Omega (E_h \Pi_h \bar{y} - \bar{y}) \mathrm{d}\mu \leq C_\epsilon h^{1+\alpha-\epsilon}. \tag{4.23}$$

The second integral on the right-hand side of (4.19) can be bounded by using the following estimates. Since $\psi_1 \in W^{2,\infty}(\Omega)$, we have,

$$\int_\Omega (\psi_1 - I_h \psi_1) \mathrm{d}\mu_1 \leq C \|\psi - I_h \psi\|_{L_\infty(\Omega)} \leq C h^2 \tag{4.24}$$

by a standard interpolation error estimate [11,21].

Since $I_h \psi_1 \leq I_h \bar{y}_h$ by (3.2), and $\mu_1 \geq 0$ by (2.5) and (4.20), we have

$$\int_\Omega (I_h \psi_1 - I_h \bar{y}_h) \mathrm{d}\mu_1 \leq 0. \tag{4.25}$$

We can write

$$\int_\Omega (I_h E_h \bar{y}_h - E_h \bar{y}_h)] \mathrm{d}\mu_1 = \int_\Omega [I_h(E_h \bar{y}_h - \bar{y}) - (E_h \bar{y}_h - \bar{y})] \mathrm{d}\mu_1 + \int_\Omega (I_h \bar{y} - \bar{y}) \mathrm{d}\mu_1. \tag{4.26}$$

Let G be a neighborhood of the active set $\mathcal{A}_1 = \{x \in \Omega : \bar{y}(x) = \psi_1(x)\}$ such that $\bar{G}$ is a compact subset of Ω. Since $\bar{y} \in W^{2,s}_{\mathrm{loc}}(\Omega)$ for any $s \in (1, \infty)$ (*cf.* (2.15)), we have

$$\int_\Omega (I_h \bar{y} - \bar{y}) \mathrm{d}\mu_1 \leq C \|I_h \bar{y} - \bar{y}\|_{L_\infty(G)} \leq C_\epsilon h^{2-\epsilon}, \tag{4.27}$$

by a standard interpolation error estimate [11,21].

Remark 4.3. Under the additional Assumption (2.18) we can take ϵ to be 0 in (4.27) (*cf.* (2.22)).

Finally we apply (3.12), (3.13), (4.22) and a standard interpolation error estimate to obtain

$$\begin{aligned}\int_\Omega [I_h(E_h\bar{y}_h-\bar{y})-(E_h\bar{y}_h-\bar{y})]\mathrm{d}\mu_1 &\le C_\epsilon \|I_h(E_h\bar{y}_h-\bar{y})-(E_h\bar{y}_h-\bar{y})\|_{W^{1,2/(1-\epsilon)}(\Omega)}\\ &\le C_\epsilon h^{1-\epsilon}|E_h\bar{y}_h-\bar{y}|_{H^2(\Omega)}\\ &\le C_\epsilon h^{1-\epsilon}\big[|E_h(\bar{y}_h-\Pi_h\bar{y})|_{H^2(\Omega)}+|E_h\Pi_h\bar{y}-y|_{H^2(\Omega)}\big]\\ &\le C_\epsilon h^{1-\epsilon}\big[\|\Pi_h\bar{y}-\bar{y}_h\|_h+h^\alpha\big].\end{aligned}\tag{4.28}$$

Putting (4.24)–(4.28) together, we arrive at the estimate

$$\int_\Omega [(\psi_1-I_h\psi_1)+(I_h\psi_1-I_h\bar{y}_h)+(I_hE_h\bar{y}_h-E_h\bar{y}_h)]\mathrm{d}\mu_1 \le C_\epsilon(h^{1+\alpha-\epsilon}+h^{1-\epsilon}\|\Pi_h\bar{y}-\bar{y}_h\|_h).\tag{4.29}$$

Similar arguments based on (2.6) lead to the following estimate for the third integral on the right-hand side of (4.19):

$$\int_\Omega [(\psi_2-I_h\psi_2)+(I_h\psi_2-I_h\bar{y}_h)+(I_hE_h\bar{y}_h-E_h\bar{y}_h)]\mathrm{d}\mu_2 \le C_\epsilon(h^{1+\alpha-\epsilon}+h^{1-\epsilon}\|\Pi_h\bar{y}-\bar{y}_h\|_h).\tag{4.30}$$

Combining (4.19), (4.23), (4.29) and (4.30), we have

$$\int_\Omega E_h(\Pi_h\bar{y}-\bar{y}_h)\mathrm{d}\mu \le C_\epsilon(h^{1+\alpha-\epsilon}+h^{1-\epsilon}\|\Pi_h\bar{y}-\bar{y}_h\|_h).\tag{4.31}$$

4.3. Convergence Results

We are now ready to establish the following main result of the paper.

Theorem 4.4. *For any $\epsilon\in(0,1)$, there exists a positive constant C_ϵ independent of h such that*

$$\|\bar{y}-\bar{y}_h\|_h \le C_\epsilon h^{\min(\alpha,1-\epsilon)},\tag{4.32}$$

where α is the index of elliptic regularity in (2.16).

Proof. It follows from (3.7), (4.2)–(4.4), (4.18) and (4.31) that

$$\begin{aligned}\|\bar{y}-\bar{y}_h\|_h^2 &\le C\big[h^{2\alpha}+h^\alpha\|\Pi_h\bar{y}-\bar{y}_h\|_h\big]+C_\epsilon\big[h^{1+\alpha-\epsilon}+h^{1-\epsilon}\|\Pi_h\bar{y}-\bar{y}_h\|_h\big]\\ &\le C_\epsilon\big[h^{2\min(\alpha,1-\epsilon)}+h^{\min(\alpha,1-\epsilon)}(\|\Pi_h\bar{y}-\bar{y}\|_h+\|\bar{y}-\bar{y}_h\|_h)\big]\\ &\le C_\epsilon\big[h^{2\min(\alpha,1-\epsilon)}+h^{\min(\alpha,1-\epsilon)}\|\bar{y}-\bar{y}_h\|_h\big].\end{aligned}$$

The estimate (4.32) is then obtained by invoking the inequality of arithmetic and geometric means. □

The following corollary is immediate.

Corollary 4.5. *There exists a positive constant C, independent of h, such that*

$$\|\bar{y}-\bar{y}_h\|_h \le Ch^\alpha$$

provided the index of elliptic regularity α is strictly less than 1.

Under the additional Assumption (2.18), we can take ϵ to be 0 in Theorem 4.4 (*cf.* Rems. 4.1–4.3).

Theorem 4.6. *Under the additional assumption that the supports of λ and μ are disjoint, we have*

$$\|\bar{y}-\bar{y}_h\|_h \leq Ch^{\alpha},$$

where the positive constant C is independent of h.

Remark 4.7. The error estimates in Theorems 4.4 and 4.6 are for a H^2-like norm, which provide more information on the optimal state $\bar{y}$ than the H^1 error estimates in [18,19,37,39].

Remark 4.8. The estimates in Corollary 4.5 and Theorem 4.6 are the best possible since we are using quasi-uniform meshes.

Let $|\cdot|_{H^1(\Omega;\mathcal{T}_h)}$ be defined by

$$|v|^2_{H^1(\Omega;\mathcal{T}_h)} = \sum_{T\in\mathcal{T}_h} |v|^2_{H^1(T)}. \tag{4.33}$$

Corollary 4.9. *For any $\epsilon \in (0,1)$, there exists a positive constant C_ϵ independent of h such that*

$$\|\bar{y}-\bar{y}_h\|_{L_2(\Omega)} + \|\bar{y}-\bar{y}_h\|_{H^1(\Omega;\mathcal{T}_h)} + \|\bar{y}-\bar{y}_h\|_{L_\infty(\Omega)} \leq C_\epsilon h^{\min(\alpha,1-\epsilon)},$$

where α is the index of elliptic regularity in (2.16). Under the additional assumption that the supports of λ and μ are disjoint, we have

$$\|\bar{y}-\bar{y}_h\|_{L_2(\Omega)} + \|\bar{y}-\bar{y}_h\|_{H^1(\Omega;\mathcal{T}_h)} + \|\bar{y}-\bar{y}_h\|_{L_\infty(\Omega)} \leq Ch^{\alpha},$$

where the positive constant C is independent of h.

Proof. According to ([15], Example 5.1 and [9], Appendix A), there exists a positive constant C independent of h such that

$$\|\bar{y}-\bar{y}_h\|_{L_2(\Omega)} + \|\bar{y}-\bar{y}_h\|_{H^1(\Omega;\mathcal{T}_h)} + \|\bar{y}-\bar{y}_h\|_{L_\infty(\Omega)} \leq C\|\bar{y}-\bar{y}_h\|_h.$$

Therefore the corollary follows from Theorems 4.4 and 4.6. □

Remark 4.10. Since $\|\cdot\|_{L_2(\Omega)}$, $\|\cdot\|_{H^1(\Omega;\mathcal{T}_h)}$ and $\|\cdot\|_{L_\infty(\Omega)}$ are lower order norms, the error estimates in Corollary 4.9 are not expected to be sharp. This is confirmed by the numerical results in Section 5.

We can take the piecewise constant function $\bar{u}_h = -\Delta_h \bar{y}_h$ as an approximation of the optimal control $\bar{u}_h$, and we immediately have the following result by Theorems 4.4 and 4.6.

Corollary 4.11. *For any $\epsilon \in (0,1)$, there exists a positive constant C_ϵ independent of h such that*

$$\|\bar{u}-\bar{u}_h\|_{L_2(\Omega)} \leq C_\epsilon h^{\min(\alpha,1-\epsilon)},$$

where α is the index of elliptic regularity in (2.16). Under the additional assumption that the supports of λ and μ are disjoint, we have

$$\|\bar{u}-\bar{u}_h\|_{L_2(\Omega)} \leq Ch^{\alpha},$$

where the positive constant C is independent of h.

Remark 4.12. In the case where the state constraints are absent, the condition (2.18) is automatically satisfied, and hence Theorem 4.6, Corollary 4.9 and Corollary 4.11 provide the error estimates for an optimal control problem with only control constraints.

TABLE 1. Results for Example 1.

h	$\|y - y_h\|_h$	Order	$\|y - y_h\|_{H^1(\Omega;\mathcal{T}_h)}$	Order	$\|y - y_h\|_{L_2(\Omega)}$	Order	$\|y - y_h\|_{\ell_\infty}$	Order
2^{-1}	3.374×10^{-1}	–	1.045×10^{0}	–	1.486×10^{-1}	–	2.547×10^{-1}	–
2^{-2}	3.422×10^{-1}	-0.02	6.932×10^{-1}	0.71	5.897×10^{-2}	1.61	1.198×10^{-1}	1.31
2^{-3}	4.892×10^{-1}	-0.57	4.721×10^{-1}	0.61	5.031×10^{-2}	0.25	1.141×10^{-1}	0.08
2^{-4}	2.766×10^{-1}	0.86	1.199×10^{-1}	2.07	1.586×10^{-2}	1.74	3.148×10^{-2}	1.95
2^{-5}	1.438×10^{-1}	0.97	2.958×10^{-2}	2.07	4.175×10^{-3}	1.97	7.571×10^{-3}	2.10
2^{-6}	7.304×10^{-2}	0.99	7.415×10^{-3}	2.02	1.143×10^{-3}	1.89	1.970×10^{-3}	1.96
2^{-7}	3.668×10^{-2}	1.00	1.898×10^{-3}	1.98	3.050×10^{-4}	1.92	5.077×10^{-4}	1.97
2^{-8}	1.836×10^{-2}	1.00	4.654×10^{-4}	2.03	7.565×10^{-5}	2.02	1.246×10^{-4}	2.03

5. Numerical results

In this section we present five numerical examples to illustrate the performance of the Morley finite element method. The first three examples concern the optimal control problem defined by (1.1)–(1.4), where the active sets exhibit different behaviors. The other two examples involve optimal control problems from the literature that do not fit our framework exactly. Nevertheless they can be solved by straightforward modifications of the discrete problem (3.1), which demonstrates the robustness of the Morley finite element method.

The computational domain for all the examples is the unit square $(0,1) \times (0,1)$ and we use uniform meshes. The discrete problems are solved by a primal-dual active set method [3,4,31,33]. The discrete ℓ_∞ norm in the tables is computed by using the vertices of the triangles, the midpoints of the edges and the centers of the triangles.

5.1. Example 1

We solve the optimal control problem defined by (1.1)–(1.4). The data are given by $\beta = 10^{-3}$, $y_d = 2$, $\psi_1 = -\infty$, $\psi_2 = 1$, $\phi_1 = -1$ and $\phi_2 = 25$. The Slater condition (1.9) is satisfied by $y = 0$. The numerical results are reported in Table 1, where the errors are estimated by comparing solutions from consecutive refinement levels.

The discrete active sets for the upper state constraint and the upper control constraint with $h = 2^{-8}$ are displayed in Figure 1. (The active set for the lower control constraint is empty). Since the active sets for the state are disjoint from the active sets for the control and the condition (2.23) is satisfied, the error estimate in Theorem 4.6 holds with $\alpha = 1$ (*cf.* Rem. 2.4). This is confirmed by the results in Table 1 for the energy norm error. On the other hand the results in Table 1 also indicate that the performance of the Morley finite element method in lower order norms is better than the one predicted by Corollary 4.9.

The discrete optimal state $\bar{y}_h$ and optimal control $\bar{u}_h$ for $h = 2^{-8}$ are plotted in Figure 2.

5.2. Example 2

We solve the optimal control problem defined by (1.1)–(1.4). The data are given by $\beta = 10^{-3}$, $\psi_1 = -\infty$, $\psi_2 = 4(x_1 - x_1^2)(x_2 - x_2^2) + 0.5$, $\phi_2 = 100$,

$$\phi_1 = \begin{cases} 8\exp\left(\dfrac{|x-(0.5,0.5)|^2}{|x-(0.5,0.5)|^2 - 0.25}\right) & \text{if} \quad |x-(0.5,0.5)| \le 0.5 \\ 0 & \text{otherwise,} \end{cases}$$

and $y_d = 1$. It is straightforward to check that the Slater condition is satisfied by $y = 9(x_1 - x_1^2)(x_2 - x_2^2)$. Since the condition (2.23) is satisfied, the index of elliptic regularity α equals 1 (*cf.* Rem. 2.4). The numerical results are displayed in Table 2, where the errors are estimated by comparing solutions from consecutive refinement levels.

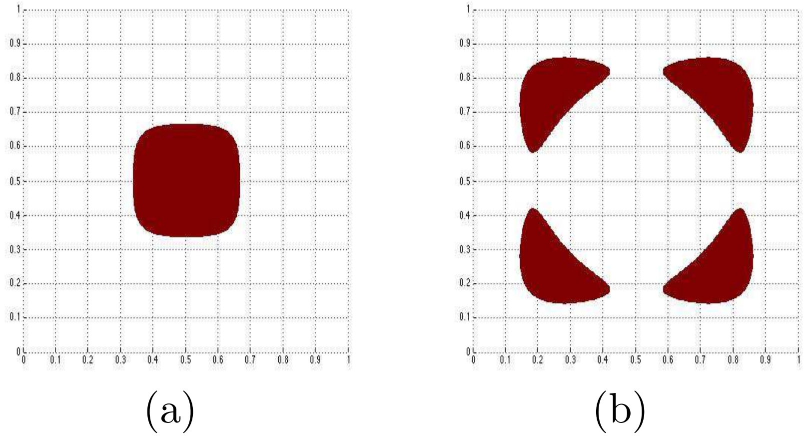


FIGURE 1. (a) discrete active set for the upper state constraint and (b) discrete active set for the upper control constraint for Example 1 with $h = 2^{-8}$.

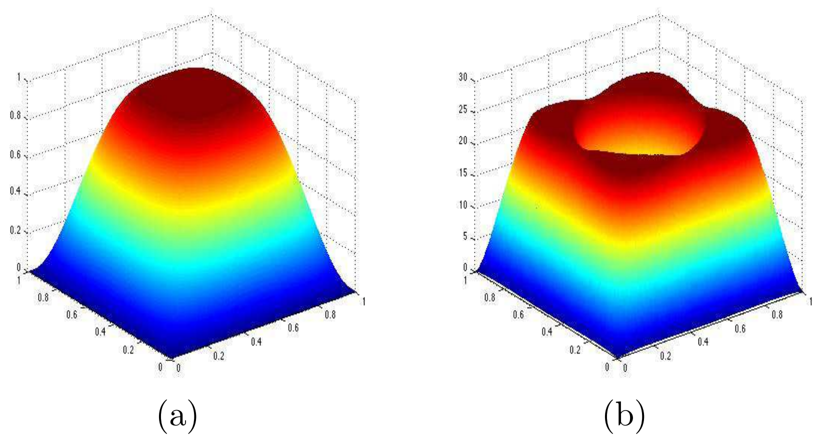
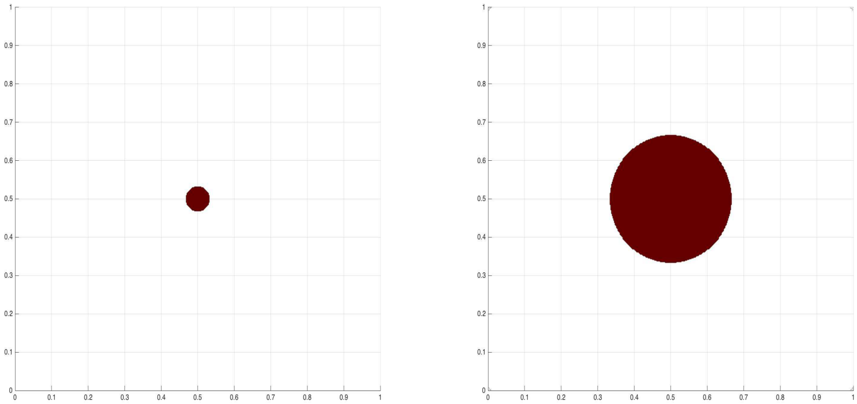


FIGURE 2. (a) discrete optimal state $\bar{y}_h$ and (b) discrete optimal control $\bar{u}_h$ for Example 1 with $h = 2^{-8}$.

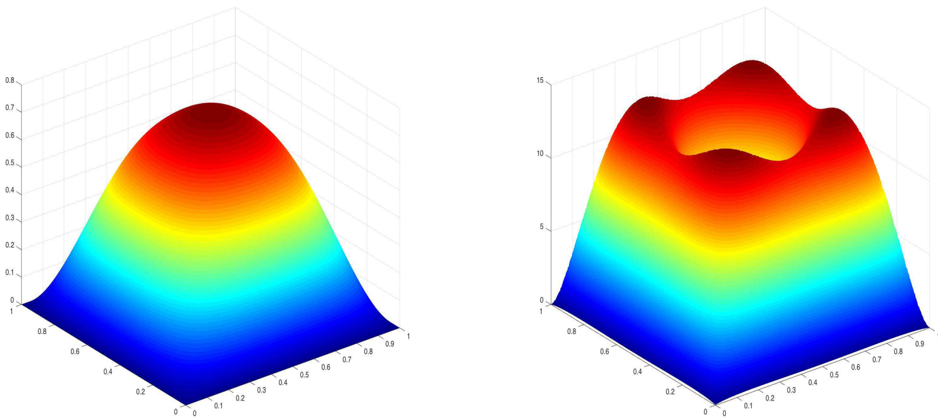
TABLE 2. Results for Example 2.

h	$\|y - y_h\|_h$	Order	$\|y - y_h\|_{H^1(\Omega;\mathcal{T}_h)}$	Order	$\|y - y_h\|_{L_2(\Omega)}$	Order	$\|y - y_h\|_{\ell_\infty}$	Order
2^{-1}	9.947×10^{-2}	–	2.887×10^{-1}	–	4.010×10^{-2}	–	8.600×10^{-2}	–
2^{-2}	2.479×10^{-1}	-1.59	5.052×10^{-1}	-0.97	4.269×10^{-2}	-0.11	9.241×10^{-2}	-0.13
2^{-3}	2.597×10^{-1}	-0.07	2.487×10^{-1}	1.12	2.838×10^{-2}	0.65	6.385×10^{-2}	0.59
2^{-4}	1.482×10^{-1}	0.85	7.088×10^{-2}	1.90	1.344×10^{-2}	1.13	2.155×10^{-2}	1.64
2^{-5}	8.110×10^{-2}	0.89	2.452×10^{-2}	1.57	5.509×10^{-3}	1.32	8.064×10^{-3}	1.45
2^{-6}	4.420×10^{-2}	0.89	8.931×10^{-3}	1.47	2.047×10^{-3}	1.44	3.149×10^{-3}	1.37
2^{-7}	2.381×10^{-2}	0.90	2.986×10^{-3}	1.59	6.828×10^{-4}	1.59	1.078×10^{-3}	1.55
2^{-8}	1.269×10^{-2}	0.91	9.346×10^{-4}	1.68	2.130×10^{-4}	1.69	3.423×10^{-4}	1.66



(a) (b)

FIGURE 3. (a) discrete active set for the upper state constraint and (b) discrete active set for the lower control constraint for Example 2 with $h = 2^{-8}$.



(a) (b)

FIGURE 4. (a) discrete optimal state $\bar{y}_h$ and (b) discrete optimal control $\bar{u}_h$ for Example 2 with $h = 2^{-8}$.

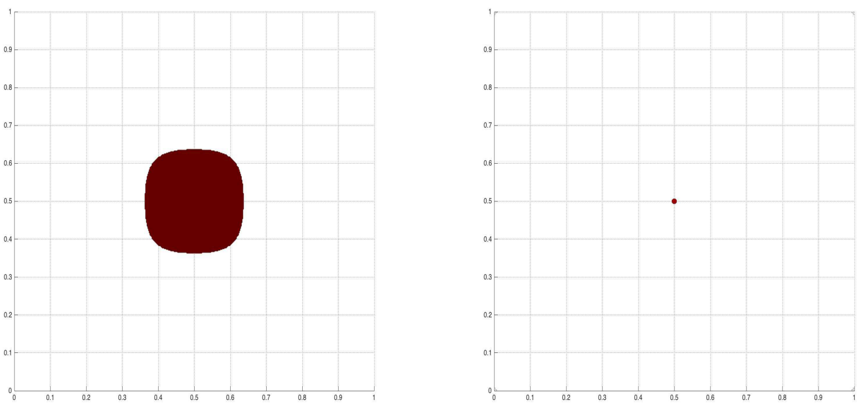
The discrete active sets for the state constraint and lower control constraint with $h = 2^{-8}$ are depicted in Figure 3. (The active set for the upper control constraint is empty.) In view of the symmetry of the data, it appears that the active set for the state constraint should contain just the center of the square.

Since μ appears to be a point measure supported at the center of the square, which is inside the active set of the lower control constraint, we expect that the convergence rate of the Morley finite element method is determined by Theorem 4.4 (with $\alpha = 1$) instead of Theorem 4.6. Hence for this example it should take more refinements to reach the asymptotic regime, which is observed in Table 2. The performance in the lower order norms is still better than the one predicted by Corollary 4.9.

The discrete optimal state $\bar{y}_h$ and optimal control $\bar{u}_h$ for $h = 2^{-8}$ are plotted in Figure 4.

TABLE 3. Results for Example 3.

h	$\|y-y_h\|_h$	Order	$\|y-y_h\|_{H^1(\Omega;\mathcal{T}_h)}$	order	$\|y-y_h\|_{L_2(\Omega)}$	Order	$\|y-y_h\|_{\ell_\infty}$	Order
2^{-1}	9.947×10^{-2}	–	2.887×10^{-1}	–	4.010×10^{-2}	–	8.600×10^{-2}	–
2^{-2}	2.461×10^{-1}	-1.58	5.080×10^{-1}	-0.98	4.307×10^{-2}	-0.12	9.160×10^{-2}	-0.11
2^{-3}	2.605×10^{-1}	-0.09	2.494×10^{-1}	1.13	2.666×10^{-2}	0.76	6.204×10^{-2}	0.62
2^{-4}	1.431×10^{-1}	0.90	6.209×10^{-2}	2.10	8.188×10^{-3}	1.78	1.714×10^{-2}	1.94
2^{-5}	7.308×10^{-2}	0.99	1.523×10^{-2}	2.07	2.098×10^{-3}	2.01	4.126×10^{-3}	2.10
2^{-6}	3.672×10^{-2}	1.00	3.776×10^{-3}	2.04	5.784×10^{-4}	1.88	1.049×10^{-3}	2.00
2^{-7}	1.836×10^{-2}	1.01	9.333×10^{-4}	2.03	1.374×10^{-4}	2.09	2.505×10^{-4}	2.01
2^{-8}	9.177×10^{-3}	1.00	2.332×10^{-4}	2.01	3.659×10^{-5}	1.91	6.413×10^{-5}	1.97



(a) (b)

FIGURE 5. (a) discrete active set for the state constraint and (b) discrete active set for the lower control constraint for Example 3 with $h = 2^{-8}$.

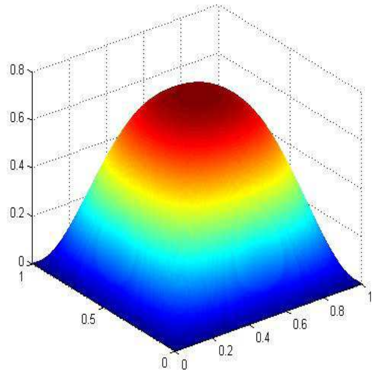
5.3. Example 3

We solve the optimal control problem defined by (1.1)–(1.4). The data are identical to the ones in Example 2 except that ϕ_1 is now given by

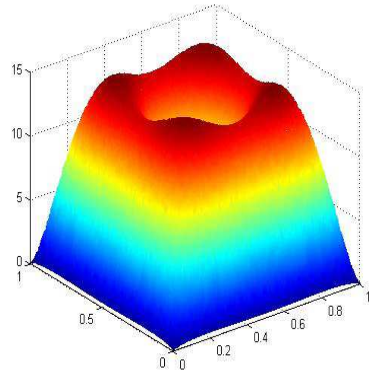
$$\phi_1 = \begin{cases} 4\exp\left(\dfrac{|x-(0.5,0.5)|^2}{|x-(0.5,0.5)|^2-0.25}\right) & \text{if} \quad |x-(0.5,0.5)|\le 0.5 \\ 0 & \text{otherwise} \end{cases}.$$

The Slater condition is satisfied by $y = 5(x_1 - x_1^2)(x_2 - x_2^2)$ and the index of elliptic regularity α again equals 1. The numerical results are displayed in Table 3, where the errors are estimated by comparing solutions from consecutive refinement levels.

For this example (*cf.* Fig. 5), the active set for the lower control constraint appears to be the center of the square, and the active set for the upper control constraint is empty. Since $\lambda \in L_2(\Omega)$ (*cf.* (2.13)) and λ is supported on the active sets of the control constraints (*cf.* (2.2)–(2.4)), we conclude that $\lambda = 0$. Therefore the condition (2.18) is satisfied even though the active set for the state contains the active set for the lower



(a)



(b)

FIGURE 6. (a) discrete optimal state $\bar{y}_h$ and (b) discrete optimal control $\bar{u}_h$ for Example 3 with $h = 2^{-8}$.

control constraint. The behavior of the energy error in Table 3 agrees with Theorem 4.6 with $\alpha = 1$. Again the performance in the lower order norms is better than the one predicted by Corollary 4.9.

Remark 5.1. In view of (2.1) and the fact that $\lambda = 0$, the solution $\bar{y}$ of this example is also the solution of an optimal control problem with only state constraints. This has been verified numerically.

The discrete optimal state $\bar{y}_h$ and optimal control $\bar{u}_h$ for $h = 2^{-8}$ are plotted in Figure 6.

5.4. Example 4

In this example we consider the following optimal control problem:

$$\text{Find} \quad (\bar{y}, \bar{u}) = \underset{(y,u)\in\mathbb{K}}{\operatorname{argmin}} \left[\frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u - u_d\|^2_{L_2(\Omega)} \right], \tag{5.1}$$

where $\mathbb{K}$ is the subset of $H^1_0(\Omega) \times L_2(\Omega)$ defined by (1.3), (1.4) and

$$\int_\Omega \nabla y \cdot \nabla z \, \mathrm{d}x = \int_\Omega (u + f) z \, \mathrm{d}x \qquad \forall\, z \in H^1_0(\Omega). \tag{5.2}$$

The following data are from [18]. We take $\beta = 0.1$, $\phi_1 = 0$, $\phi_2 = 100$ and $\psi_2 = +\infty$. The exact solution $\bar{y} = \sin(\pi x_1)\sin(\pi x_2)$, and the lower bound ψ_1 for the state is given by

$$\psi_1(x) = \begin{cases} \bar{y}(x) & \text{if} \quad \bar{y}(x) \geq c \\ 2\bar{y}(x) - c & \text{if} \quad \bar{y}(x) \leq c, \end{cases}$$

where $c = 0.6$. The function

$$f = -\Delta\bar{y} - \max(-\Delta\bar{y} - \kappa, 0),$$

where $\kappa = 5$, and the functions u_d and y_d are defined by

$$u_d(x) = \begin{cases} \max(-\Delta\bar{y}(x) - \kappa, 0) & \text{if} \quad \bar{y}(x) > c \\ -\kappa & \text{if} \quad \bar{y}(x) < c, \end{cases} \quad y_d(x) = \begin{cases} \bar{y}(x) - 1 & \text{if} \quad \bar{y}(x) > c \\ \beta\Delta^2\bar{y}(x) + \bar{y}(x) & \text{if} \quad \bar{y}(x) < c. \end{cases}$$

TABLE 4. Results for Example 4.

h	$\Vert y-y_h\Vert_h$	Order	$\Vert y-y_h\Vert_{H^1(\Omega;\mathcal{T}_h)}$	Order	$\Vert y-y_h\Vert_{L_2(\Omega)}$	Order	$\Vert y-y_h\Vert_{\ell_\infty}$	Order
2^{-1}	1.733×10^{0}	–	5.932×10^{-1}	–	7.022×10^{-2}	–	1.794×10^{-1}	–
2^{-2}	1.138×10^{0}	0.73	2.452×10^{-1}	1.54	1.435×10^{-2}	2.77	4.061×10^{-2}	2.59
2^{-3}	6.320×10^{-1}	0.93	7.671×10^{-2}	1.84	2.755×10^{-3}	2.61	9.420×10^{-3}	2.31
2^{-4}	3.334×10^{-1}	0.97	2.498×10^{-2}	1.69	1.211×10^{-3}	1.24	3.130×10^{-3}	1.66
2^{-5}	1.688×10^{-1}	1.00	7.955×10^{-3}	1.69	4.733×10^{-4}	1.39	9.930×10^{-4}	1.69
2^{-6}	8.428×10^{-2}	1.01	2.329×10^{-3}	1.79	1.332×10^{-4}	1.85	3.031×10^{-4}	1.73
2^{-7}	4.190×10^{-2}	1.01	6.115×10^{-4}	1.94	3.494×10^{-5}	1.94	8.488×10^{-5}	1.85
2^{-8}	2.089×10^{-2}	1.01	1.593×10^{-4}	1.95	9.252×10^{-6}	1.92	2.125×10^{-5}	2.00

Remark 5.2. It is straightforward to check that $\bar{y}$ satisfies the following first order optimality conditions for the optimal control problem defined by (1.3), (1.4), (5.1) and (5.2):

$$\beta\int_\Omega(-\Delta\bar{y}-f-u_d)(-\Delta z)\mathrm{d}x+\int_\Omega(\bar{y}-y_d)z\,\mathrm{d}x=\int_\Omega\lambda(-\Delta z)\mathrm{d}x+\int_\Omega z\,\mathrm{d}\mu\qquad\forall\,z\in H^2(\Omega)\cap H_0^1(\Omega),$$

where

$$\lambda=\begin{cases}-\beta(f+u_d)\,(\geq 0) & \text{if } -\Delta\bar{y}+\kappa\leq 0\\ 0 & \text{otherwise,}\end{cases}\qquad \langle\mu,z\rangle=\int_{\bar{y}\geq c}z\,\mathrm{d}x-2\pi^2\beta\int_{\bar{y}=c}\frac{\partial\bar{y}}{\partial n}z\,\mathrm{d}s,$$

and n is the unit outer normal on the boundary of the set defined by $\bar{y}\geq c$. (Note that $\partial\bar{y}/\partial n$ is negative on $\bar{y}=c$.) Thus there is no need for a Slater condition.

The discrete problem is to find

$$\bar{y}_h=\underset{y\in K_h}{\operatorname{argmin}}\left[\frac{1}{2}a_h(y,y)-(y_d,y)-\beta\sum_{T\in\mathcal{T}_h}\int_T(f+u_d)(\Delta y)\mathrm{d}x\right],$$

where $a_h(\cdot,\cdot)$ is defined by (3.3) and

$$K_h=\{y\in V_h:\,I_h\psi_1\leq I_hy\leq I_h\psi_2\quad\text{and}\quad Q_h\phi_1\leq-\Delta_hy-Q_hf\leq Q_h\phi_2\}.$$

The numerical results are reported in Table 4.

For this example the active set for the upper control constraint is empty. The active set for the state constraint, defined by the condition $\bar{y}\geq c$, is disjoint from the active set for the lower control constraint defined by the condition $-\Delta\bar{y}-\kappa\leq 0$ (*cf.* Fig. 7). The behavior of the energy norm error is consistent with the error estimate in Theorem 4.6 with $\alpha=1$, and the errors in the lower order norms are better than the ones predicted by Corollary 4.9.

The discrete optimal state $\bar{y}_h$ and optimal control $\bar{y}_h$ for h^{-8} are plotted in Figure 8. We observe good agreements with the exact optimal state $\bar{y}$ and the exact optimal control $\bar{u}=-\Delta\bar{y}-f$.

5.5. Example 5

In this example we solve the optimal control problem defined by (1.1), (1.3), (1.4) and the constraint

$$\int_\Omega\nabla y\cdot\nabla z\,\mathrm{d}x+\int_\Omega yz\,\mathrm{d}x=\int_\Omega uz\,\mathrm{d}x\qquad\forall\,z\in H_0^1(\Omega).$$

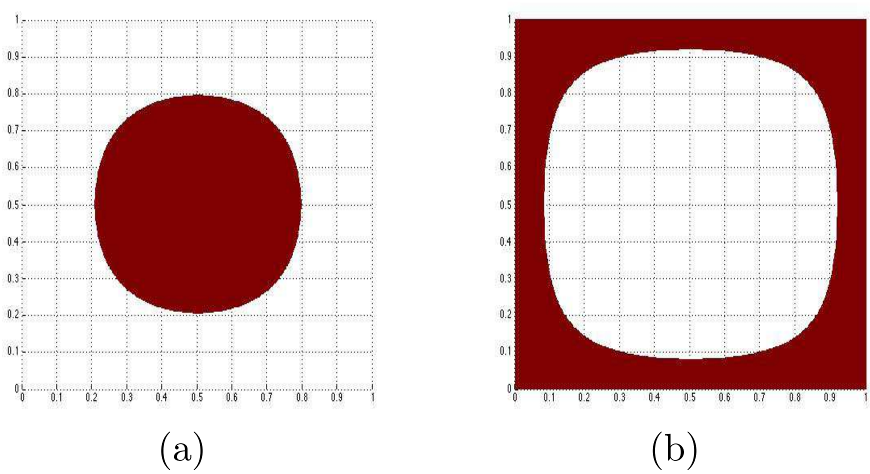


FIGURE 7. (a) active set for the state constraint and (b) active set for the lower control constraint for Example 4.

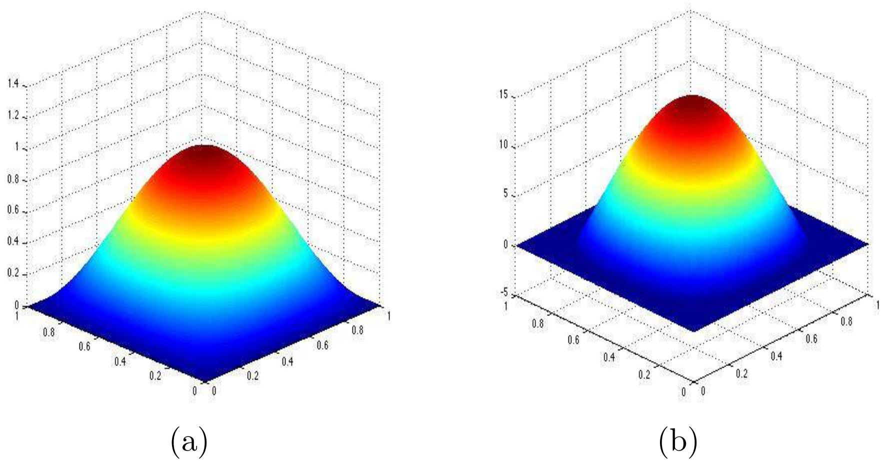


FIGURE 8. (a) discrete optimal state $\bar{y}_h$ and (b) discrete optimal control $\bar{u}_h$ for Example 4 with $h = 2^{-8}$.

The data $\beta = 10^{-3}$, $y_d = 10(1 - x_1 - x_2)^3$, $\psi_1 = -0.35$, $\psi_2 = 0.4$, $\phi_1 = -0.35$ and $\phi_2 = 20$ are taken from ([39], Example 5.2). The Slater condition (1.9) (with $u = -\Delta y + y$) is satisfied by $y = 0$. We use the bilinear form

$$a_h(y, z) = \beta \sum_{T \in \mathcal{T}_h} \int_T \left(D^2 y : D^2 z + 2\nabla y \cdot \nabla z\right) \mathrm{d}x + (1 + \beta) \int_\Omega yz \, \mathrm{d}x$$

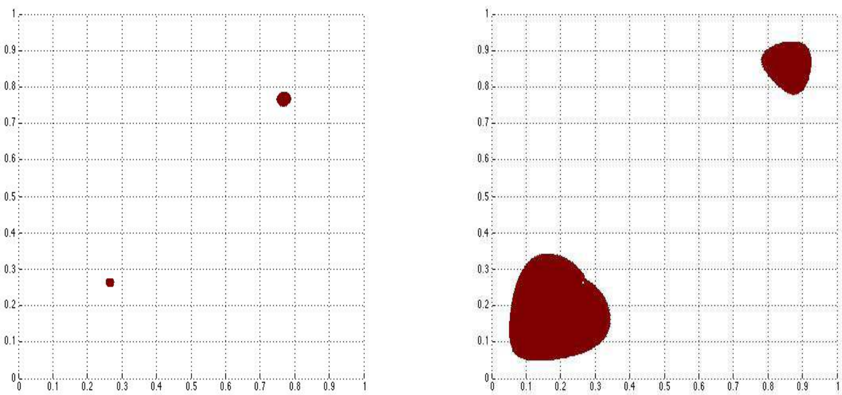
for the discrete problem (3.1), where

$$K_h = \{y \in V_h : I_h\psi_1 \leq I_h y \leq I_h\psi_2 \quad \text{and} \quad Q_h\phi_1 \leq -\Delta_h y + Q_h y \leq Q_h\phi_2\}.$$

The numerical results are reported in Table 5, where $\|\cdot\|_h = \sqrt{a_h(\cdot,\cdot)}$ and the errors are estimated by comparing solutions from consecutive refinement levels.

TABLE 5. Results for Example 5

h	$\|y-y_h\|_h$	Order	$\|y-y_h\|_{H^1(\Omega;\mathcal{T}_h)}$	Order	$\|y-y_h\|_{L_2(\Omega)}$	Order	$\|y-y_h\|_{\ell_\infty}$	Order
2^{-1}	5.063×10^{-1}	–	1.660×10^{0}	–	1.676×10^{-1}	–	3.748×10^{-1}	–
2^{-2}	4.108×10^{-1}	0.36	8.898×10^{-1}	1.09	7.286×10^{-2}	1.45	1.628×10^{-1}	1.45
2^{-3}	2.940×10^{-1}	0.53	3.228×10^{-1}	1.61	1.910×10^{-2}	2.12	1.196×10^{-1}	0.49
2^{-4}	2.144×10^{-1}	0.48	1.124×10^{-1}	1.59	8.428×10^{-3}	1.24	4.061×10^{-2}	1.63
2^{-5}	1.104×10^{-1}	0.98	2.826×10^{-2}	2.04	2.463×10^{-3}	1.82	9.767×10^{-3}	2.10
2^{-6}	5.571×10^{-2}	1.00	7.316×10^{-3}	1.97	8.066×10^{-4}	1.63	2.503×10^{-3}	1.99
2^{-7}	2.783×10^{-2}	1.01	1.819×10^{-3}	2.02	1.941×10^{-4}	2.07	6.738×10^{-3}	1.90
2^{-8}	1.390×10^{-2}	1.01	4.758×10^{-4}	1.94	4.716×10^{-5}	2.05	1.613×10^{-4}	2.07



(a) (b)

FIGURE 9. (a) discrete active sets for the state constraints and (b) discrete active sets for the control constraints for Example 5 with $h = 2^{-8}$.

The discrete active sets for the state constraints and the control constraints with $h = 2^{-8}$ are displayed in Figure 7, which match the ones in ([39], Fig. 5.3).

For this example the active set for the upper state constraint in the lower left part of (a) overlaps with the active set for the upper control constraint in the lower left part of (b). The results for the energy norm error in Table 4 agree with the error estimate (4.32) with $\alpha = 1$. Again the performance in the lower order norms is better than the one predicted by Corollary 4.9.

The plots for the discrete optimal state $\bar{y}_h$ and the discrete optimal control $\bar{u}_h$ with $h = 2^{-8}$ are given in Figure 10, which agree with the ones in ([39], Fig. 5.2).

6. CONCLUDING REMARKS

We have developed and analyzed a Morley finite element method for an elliptic distributed optimal control problem with pointwise state and control constraints on convex polygons.

The error estimate (4.32) is essentially the best possible for quasi-uniform meshes. It can be improved to

$$\|\bar{y}-\bar{y}_h\|_h \le C_\epsilon h^{1-\epsilon}$$

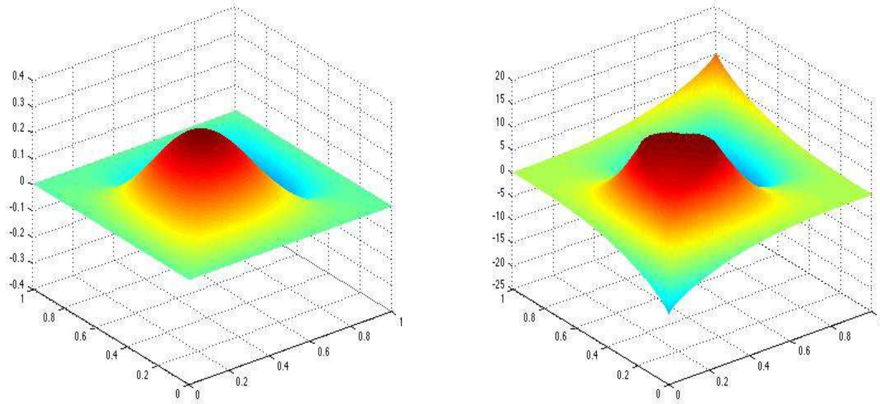


FIGURE 10. (a) discrete optimal state $\bar{y}_h$ and (b) discrete optimal control $\bar{u}_h$ for Example 5 with $h = 2^{-8}$.

for an arbitrary convex polygon Ω if properly graded meshes [2,7,13] are employed. For concreteness we only consider the Dirichlet boundary condition in the partial differential equation constraint. But the results can be extended to the case of the Neumann boundary condition.

For simplicity we take $\bar{u}_h = -\Delta_h \bar{y}_h$ to be the approximation for the optimal control $\bar{u}$. We can also apply more sophisticated post-processing techniques to generate better approximations for $\bar{u}$ (*cf.* [14]). This will be addressed elsewhere.

The convergence analysis in this paper can also be applied to other finite element methods as long as the estimates in Section 3.1 and Section 3.2 are satisfied by appropriate interpolation operator Π_h and enriching operator E_h. For example, it can be directly applied to a finite element method based on the (more complicated) Hsieh-Clough-Tocher macro element [20].

Appendix A. An enriching operator for the Morley finite element

In this appendix we construct an operator $E_h : V_h \longrightarrow H^2(\Omega) \cap H_0^1(\Omega)$ with the properties stated in Section 3.2. The construction involves the Hsieh-Clough-Tocher (HCT) element. (The degrees of freedom for the Morley element and HCT element are depicted in Fig. A.1).

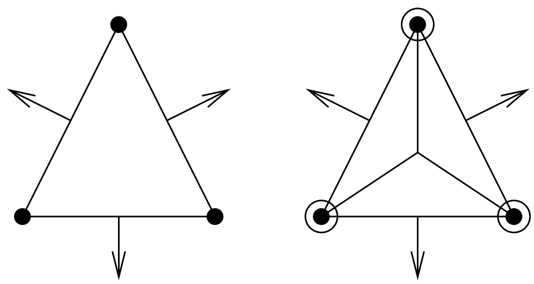


FIGURE A.1. Morley element (*left*) and HCT element (*right*).

A.1. Construction of E_h

Let $\tilde{W}_h \subset H^2(\Omega)$ be the HCT finite element space associated with $\mathcal{T}_h$. A function $\tilde{w}$ belongs to $\tilde{W}_h$ if and only if (i) $\tilde{w} \in C^1(\bar{\Omega})$ and (ii) on each triangle $T \in \mathcal{T}_h$, $\tilde{w}$ is a piecewise cubic polynomial with respect to the subdivision determined by the center and the vertices of T. A function in $\tilde{W}_h$ is determined by (i) its values and the values of its first order derivatives at the vertices of $\mathcal{T}_h$ and (ii) the integrals of its normal derivatives over the edges of $\mathcal{T}_h$. The space $W_h \subset H^2(\Omega) \cap H_0^1(\Omega)$ consists of all the members of $\tilde{W}_h$ that vanish on $\partial\Omega$.

We will use the following notation.

- h_T is the diameter of the triangle $T \in \mathcal{T}_h$.
- v_T is the restriction of v to $T \in \mathcal{T}_h$.
- $\mathcal{V}_h$ is the set of the vertices of $\mathcal{T}_h$.
- $\mathcal{V}_h^b$ is the subset of $\mathcal{V}_h$ consisting of the vertices that belong to $\partial\Omega$.
- $\mathcal{E}_h$ is the set of the edges of $\mathcal{T}_h$.
- $\mathcal{E}_h^i$ is the set of edges interior to Ω.
- $\mathcal{E}_h^b = \mathcal{E}_h \setminus \mathcal{E}_h^i$ is the set of the boundary edges.
- $|e|$ is the length of the edge $e \in \mathcal{E}_h$.
- For an edge $e \in \mathcal{E}_h^i$, the set $\mathcal{T}_e$ contains the two triangles that share e as a common edge.
- $\mathcal{V}_T$ (resp., $\mathcal{E}_T$) is the set of the three vertices (resp., edges) of $T \in \mathcal{T}_h$.
- $\mathcal{T}_p$ is the set of the triangles in $\mathcal{T}_h$ that share p as a common vertex and $|\mathcal{T}_p|$ is the number of triangles in $\mathcal{T}_p$.
- n_e (resp., t_e) is a unit normal (resp., tangent) of $e \in \mathcal{E}_h$.
- $\mathcal{E}_p^i$ is the set of edges in $\mathcal{E}_h^i$ that share p as a common endpoint.
- S_T is the set of triangles in $\mathcal{T}_h$ that share a common vertex with T.
- For an interior edge $e \in \mathcal{E}_h^i$ shared by the two triangles $T_{e,1}$ and $T_{e,2}$, we define the jump of the gradient across e by
$$[\![\nabla v]\!]_e = v_1 \otimes n_1 + v_2 \otimes n_2,$$
where v_j is the restriction of v to $T_{e,j}$ and n_j is the unit normal of e pointing towards the outside of $T_{e,j}$.

We begin by defining an operator operator $\tilde{E}_h : V_h \longrightarrow \tilde{W}_h$. First we take

$$(\tilde{E}_h v)(p) = v(p) \qquad \forall\, p \in \mathcal{V}_h. \tag{A.1}$$

This means in particular that $\tilde{E}_h v$ vanishes at all the vertices in $\mathcal{V}_h^b$, since v vanishes at these vertices as part of the definition of V_h.

Next we define

$$\nabla(\tilde{E}_h v)(p) = \frac{1}{|\mathcal{T}_p|} \sum_{T \in \mathcal{T}_p} (\nabla v_T)(p) \qquad \forall\, p \in \mathcal{V}_h. \tag{A.2}$$

We complete the definition of $\tilde{E}_h v$ by specifying that

$$\int_e \frac{\partial(\tilde{E}_h v)}{\partial n_e} \mathrm{d}s = \int_e \frac{\partial v}{\partial n_e} \mathrm{d}s \qquad \forall\, e \in \mathcal{E}_h. \tag{A.3}$$

The operator $E_h : V_h \longrightarrow W_h$ is defined by modifying $\tilde{E}_h$ as follows:

$$\nabla(E_h v)(p) = 0 \tag{A.4}$$

at any corner p of Ω, and at a vertex $p \in \mathcal{V}_h^b$ that is not a corner of Ω, we define

$$\frac{\partial(E_h v)}{\partial n} = \frac{\partial(\tilde{E}_h v)}{\partial n}, \tag{A.5}$$

$$\frac{\partial(E_h v)}{\partial t} = 0, \tag{A.6}$$

where n (resp., t) is a unit normal (resp., tangent) along $\partial\Omega$.

Note that (A.1), (A.4) and (A.6) imply $E_h v$ vanishes on $\partial\Omega$.

By the construction of E_h (*cf.* (A.1) and (A.3)), the invariance properties (3.10) and (3.11) are satisfied.

Remark A.1. The analysis of $\tilde{E}_h$ is simpler since the definition of $\tilde{E}_h$ does not involve any boundary condition. Below we will establish the estimates for E_h by obtaining their analogs for $\tilde{E}_h$ and by bounding the difference between $\tilde{E}_h$ and E_h.

A.2. Proof of (3.12)

Let $T \in \mathcal{T}_h$ and $v \in V_h$ be arbitrary. We have, by equivalence of norms on finite dimensional spaces and scaling,

$$\|v - \tilde{E}_h v\|_{L_2(T)}^2 \approx \sum_{p \in \mathcal{V}_T} \left(h_T^2 [v_T(p) - \tilde{E}_h(p)]^2 + h_T^4 [(\nabla v_T)(p) - (\nabla \tilde{E}_h v)(p)]^2\right) + \sum_{e \in \mathcal{E}_T} h_T^3 \int_e \left[\frac{\partial v_T}{\partial n_e} - \frac{\partial(\tilde{E}_h v)}{\partial n_e}\right]^2 \mathrm{d}s,$$

where the constants in the equivalence relation depend only on the shape regularity of T, and hence

$$\|v - \tilde{E}_h v\|_{L_2(T)}^2 \le C \left(\sum_{p \in \mathcal{V}_T} \sum_{e \in \mathcal{E}_p^i} h_T^3 \|[\![\nabla v]\!]_e\|_{L_2(e)}^2 + \sum_{e \in \mathcal{E}_T} h_T^3 \|\partial(v_T - \tilde{E}_h v))/\partial n_e\|_{L_2(e)}^2 \right) \tag{A.7}$$

by (A.1) and (A.2).

Note that

$$\|[\![\nabla v]\!]_e\|_{L_2(e)}^2 \le C|e| \sum_{T \in \mathcal{T}_e} |v|_{H^2(T)}^2 \qquad \forall\, e \in \mathcal{E}_h^i \tag{A.8}$$

because $[\![\nabla v]\!]_e = 0$ at the midpoint of e, and

$$\|\partial(v_T - \tilde{E}_h v))/\partial n_e\|_{L_2(e)}^2 \le C h_T |v|_{H^2(T)}^2 \qquad \forall\, e \in \mathcal{E}_T \tag{A.9}$$

because of (A.3).

Combining (A.7)–(A.9) we obtain

$$\|v - \tilde{E}_h v\|_{L_2(T)}^2 \le C \sum_{T' \in S_T} h_{T'}^4 |v|_{H^2(T')},$$

which together with standard inverse estimates implies

$$\sum_{k=0}^{2} h^{2k} \sum_{T \in \mathcal{T}_h} |v - \tilde{E}_h v|_{H^k(T)}^2 \le C h^4 \sum_{T \in \mathcal{T}_h} |v|_{H^2(T)}^2 \qquad \forall\, v \in V_h. \tag{A.10}$$

Now we consider the difference between $\tilde{E}_h v$ and $E_h v$. A direct calculation using (A.2) and (A.4)–(A.6) yields

$$\begin{aligned} \|\tilde{E}_h v - E_h v\|_{L_2(\Omega)}^2 &\le C\left(\sum_{p\in\mathcal{V}_h^b}\sum_{e\in\mathcal{E}_p^i}|e|^3\|[\![\nabla v]\!]_e\|_{L_2(e)}^2 + \sum_{e\in\mathcal{E}_h^b}|e|^3\|\partial v/\partial t_e\|_{L_2(e)}^2\right) \\ &\le C\sum_{p\in\mathcal{V}_h^b}\sum_{T\in\mathcal{T}_p} h_T^4|v|_{H^2(T)}^2, \end{aligned} \tag{A.11}$$

where we have used (A.8) and the fact that v vanishes at the boundary vertices.

The estimate (3.12) follows from (A.10), (A.11) and standard inverse estimates.

A.3. Proofs of (3.13) and (3.14)

Observe that $[(\tilde{E}_h\Pi_h)\zeta]\big|_T$ only depends on $\zeta\big|_{S_T}$ and $[(\tilde{E}_h\Pi_h)\zeta]\big|_T = \zeta\big|_T$ if $\zeta\big|_{S_T}$ is a quadratic polynomial. (This observation relies on the fact that the definition of $\tilde{E}_h$ does not involve any boundary condition.)

It then follows from the Bramble-Hilbert lemma [6,23] and scaling that

$$\begin{aligned} &\sum_{k=0}^{2} h_T^k|\zeta - \tilde{E}_h\Pi_h\zeta|_{H^k(T)} \le Ch_T^{2+\alpha}|\zeta|_{H^{2+\alpha}(S_T)}, \\ &|\zeta - \tilde{E}_h\Pi_h\zeta|_{W^{1,2/(1-\epsilon)}(T)} \le C_\epsilon h_T^{1+\alpha-\epsilon}|\zeta|_{H^{2+\alpha}(S_T)}, \end{aligned}$$

for $0\le\alpha\le 1$ and $0\le\epsilon<1$.

Consequently we have

$$\sum_{k=0}^{2} h^k|\zeta - \tilde{E}_h\Pi_h\zeta|_{H^k(\Omega)} \le Ch^{2+\alpha}|\zeta|_{H^{2+\alpha}(\Omega)}, \tag{A.12}$$

$$\|\zeta - \tilde{E}_h\Pi_h\zeta\|_{W^{1,2/(1-\epsilon)}(\Omega)} \le C_\epsilon h^{1+\alpha-\epsilon}|\zeta|_{H^{2+\alpha}(\Omega)}, \tag{A.13}$$

for $0\le\alpha\le 1$ and $0\le\epsilon<1$.

On the other hand, for $\zeta\in H^{2+\alpha}(\Omega)\cap H_0^1(\Omega)$, we have by (3.6) and a direct calculation

$$\begin{aligned} \|\tilde{E}_h\Pi_h\zeta - E_h\Pi_h\zeta\|_{L_2(\Omega)}^2 &\le C\left(\sum_{p\in\mathcal{V}_h^b}\sum_{e\in\mathcal{E}_p^i}|e|^3\|[\![\nabla(\zeta-\Pi_h\zeta)]\!]_e\|_{L_2(e)}^2 \sum_{e\in\mathcal{E}_h^b}|e|^3\|\partial(\zeta-\Pi_h\zeta)/\partial t_e\|_{L_2(e)}^2\right) \\ &\le C\sum_{p\in\mathcal{V}_h^b}\sum_{T\in\mathcal{T}_p} h_T^{4+2\alpha}|\zeta|_{H^{2+\alpha}(T)}^2, \end{aligned}$$

which together with standard inverse estimates implies

$$\sum_{k=0}^{2} h^k|\tilde{E}_h\Pi_h\zeta - E_h\Pi_h\zeta|_{H^k(\Omega)} \le Ch^{2+\alpha}|\zeta|_{H^{2+\alpha}(\Omega)}, \tag{A.14}$$

$$|\tilde{E}_h\Pi_h\zeta - E_h\Pi_h\zeta|_{W^{1,2/(1-\epsilon)}(\Omega)} \le Ch^{1+\alpha-\epsilon}|\zeta|_{H^{2+\alpha}(\Omega)}. \tag{A.15}$$

The estimates (3.13) and (3.14) follow from (A.12)–(A.13) and (A.14)–(A.15).

A.4. Proof of (3.15)

It suffices to show that

$$|a_h(\zeta, v - E_h v)| \le C h^\alpha |\zeta|_{H^{2+\alpha}(\Omega)} \|v\|_h \tag{A.16}$$

for all $\zeta \in H^{2+\alpha}(\Omega)$ and $v \in V_h$. Then we can immediately obtain the estimate (3.15) from (3.6), (A.16) and the identity

$$a(\zeta, E_h v) - a_h(\Pi_h \zeta, v) = a_h(\zeta - \Pi_h \zeta, v) - a_h(\zeta, v - E_h v).$$

In view of (3.3) and (3.12), the estimate (A.16) follows from

$$\left| \sum_{T \in \mathcal{T}_h} \int_T D^2 \zeta : D^2 (v - E_h v)\, \mathrm{d}x \right| \le C h^\alpha |\zeta|_{H^{2+\alpha}(\Omega)} \|v\|_h, \tag{A.17}$$

whose proof can be found in ([12], Appendix A).

References

[1] R.A. Adams and J.J.F. Fournier, Sobolev Spaces (2nd Edition). Academic Press, Amsterdam (2003).

[2] Th. Apel, A.-M. Sändig and J.R. Whiteman, Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Math. Methods Appl. Sci.* **19** (1996) 63–85.

[3] M. Bergounioux, K. Ito and K. Kunisch, Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37** (1999) 1176–1194.

[4] M. Bergounioux and K. Kunisch, Primal-dual strategy for state-constrained optimal control problems. *Comput. Optim. Appl.* **22** (2002) 193–224.

[5] M. Bergounioux and K. Kunisch, On the structure of Lagrange multipliers for state-constrained optimal control problems. *Systems Control Lett.* **48** (2003) 169–176.

[6] J.H. Bramble and S.R. Hilbert, Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation. *SIAM J. Numer. Anal.* **7** (1970) 113–124.

[7] J.J. Brannick, H. Li and L.T. Zikatanov, Uniform convergence of the multigrid V-cycle on graded meshes for corner singularities. *Numer. Linear Algebra Appl.* **15** (2008) 291–306.

[8] S.C. Brenner, C.B. Davis and L.-Y. Sung, A partition of unity method for a class of fourth order elliptic variational inequalities. *Comput. Methods Appl. Mech. Engrg.* **276** (2014) 612–626.

[9] S.C. Brenner, M. Neilan, A. Reiser and L.-Y. Sung, A C^0 interior penalty method for a von Kármán plate. *Numer. Math.* **135** (2017) 803–832.

[10] S.C. Brenner, M. Oh, S. Pollock, K. Porwal, M. Schedensack and N. Sharma, A C^0 interior penalty method for elliptic distributed optimal control problems in three dimensions with pointwise state constraints. In *Topics in Numerical Partial Differential Equations and Scientific Computing*, edited by S.C. Brenner. vol. 160 of *The IMA Volumes in Mathematics and its Applications*, Cham-Heidelberg-New York-Dordrecht-London (2016) 1–22. Springer.

[11] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods (3rd Edition). Springer Verlag, New York (2008).

[12] S.C. Brenner and L.-Y. Sung, A new convergence analysis of finite element methods for elliptic distributed optimal control problems with pointwise state constraints. *SICON*, to appear.

[13] S.C. Brenner, L.-Y. Sung and Y. Zhang, A quadratic C^0 interior penalty method for an elliptic optimal control problem with state constraints. In *Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations*, edited by O. Karakashian X. Feng and Y. Xing. vol. 157 of *The IMA Volumes in Mathematics and its Applications* (2012 John H. Barrett Memorial Lectures). Cham-Heidelberg-New York-Dordrecht-London. Springer (2013) 97–132.

[14] S.C. Brenner, L.-Y. Sung and Y. Zhang, Post-processing procedures for a quadratic C^0 interior penalty method for elliptic distributed optimal control problems with pointwise state constraints. *Appl. Numer. Math.* **95** (2015) 99–117.

[15] S.C. Brenner, K. Wang and J. Zhao, Poincaré-Friedrichs inequalities for piecewise H^2 functions. *Numer. Funct. Anal. Optim.* **25** (2004) 463–478.

[16] E. Casas, L^2 estimates for the finite element method for the Dirichlet problem with singular data. *Numer. Math.* **47** (1985) 627–632.

[17] E. Casas, M. Mateos and B. Vexler, New regularity results and improved error estimates for optimal control problems with state constraints. *ESAIM: COCV* **20** (2014) 803–822.

[18] S. Cherednichenko and A. Rösch, Error estimates for the regularization of optimal control problems with pointwise control and state constraints. *Z. Anal. Anwend.* **27** (2008) 195–212.

[19] S. Cherednichenko and A. Rösch, Error estimates for the discretization of elliptic control problems with pointwise control and state constraints. *Comput. Optim. Appl.* **44** (2009) 27–55.

[20] P.G. Ciarlet, Sur l'élément de Clough et Tocher. *RAIRO Anal. Numér.* **8** (1974) 19–27.

[21] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978).
[22] M. Dauge, Elliptic Boundary Value Problems on Corner Domains. *Lect. Notes Math.* Springer Verlag, Berlin Heidelberg **1341** (1988).
[23] T. Dupont and R. Scott, Polynomial approximation of functions in Sobolev spaces. *Math. Comput.* **34** (1980) 441–463.
[24] I. Ekeland and R. Témam, Convex Analysis and Variational Problems. *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1999).
[25] L.C. Evans, Partial Differential Equations (Second Edition). *Amer. Math. Soc.*, Providence, RI (2010).
[26] J. Frehse, On the regularity of the solution of the biharmonic variational inequality. *Manuscripta Math.* **9** (1973) 91–103.
[27] D. Gilbarg and N.S. Trudinger, Elliptic Partial Differential Equations of Second Order. *Classics in Mathematics.* Springer Verlag, Berlin (2001).
[28] W. Gong and N. Yan, A mixed finite element scheme for optimal control problems with pointwise state constraints. *J. Sci. Comput.* **46** (2011) 182–203.
[29] P. Grisvard, Elliptic Problems in Non Smooth Domains. Pitman, Boston (1985).
[30] P. Grisvard, Singularities in Boundary Value Problems. Masson, Paris (1992).
[31] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2003) 865–888.
[32] L. Hörmander, The Analysis of Linear Partial Differential Operators. III. Springer Verlag, Berlin (1985).
[33] K. Ito and K. Kunisch, Lagrange Multiplier Approach to Variational Problems and Applications. Society for Industrial and Applied Mathematics, Philadelphia, PA (2008).
[34] D. Kinderlehrer and G. Stampacchia. An Introduction to Variational Inequalities and Their Applications. Society for Industrial and Applied Mathematics, Philadelphia (2000).
[35] W. Liu, W. Gong and N. Yan, A new finite element approximation of a state-constrained optimal control problem. *J. Comput. Math.* **27** (2009) 97–114.
[36] D.G. Luenberger, Optimization by Vector Space Methods. John Wiley and Sons Inc., New York (1969).
[37] C. Meyer, Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control Cybernet.* **37** (2008) 51–83.
[38] L.S.D. Morley, The triangular equilibrium problem in the solution of plate bending problems. *Aero. Quart.* **19** (1968) 149–169.
[39] A. Rösch and D. Wachsmuth, A posteriori error estimates for optimal control problems with state and control constraints. *Numer. Math.* **120** (2012) 733–762.