

Error estimates for parabolic optimal control problems with control and state constraints

Wei Gong · Michael Hinze

Received: 1 July 2011 / Published online: 20 February 2013
© Springer Science+Business Media New York 2013

Abstract The numerical approximation to a parabolic control problem with control and state constraints is studied in this paper. We use standard piecewise linear and continuous finite elements for the space discretization of the state, while the dG(0) method is used for time discretization. A priori error estimates for control and state are obtained by an improved maximum error estimate for the corresponding discretized state equation. Numerical experiments are provided which support our theoretical results.

Keywords Optimal control problem · Finite element method · A priori error estimate · Parabolic equation · Control and state constraints

1 Introduction

In this paper we consider the optimal control problem

$$\min_{y \in K_Y, u \in K_U} J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega_T)}^2 + \frac{\alpha}{2} \|u\|_U^2 \quad (1.1)$$

W. Gong (✉)

LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science,
Chinese Academy of Sciences, Beijing 100190, China
e-mail: wgong@lsec.cc.ac.cn

W. Gong · M. Hinze

Bereich Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg,
Germany

M. Hinze

e-mail: michael.hinze@uni-hamburg.de

subject to

$$\left. \begin{aligned} y_t - \Delta y + y &= \mathcal{B}u && \text{in } \Omega_T, \\ \frac{\partial y}{\partial n} &= 0 && \text{on } \Gamma_T, \\ y(0) &= y_0 && \text{in } \Omega, \end{aligned} \right\} \tag{1.2}$$

where $\Omega_T = \Omega \times (0, T]$, $\Gamma_T = \partial\Omega \times (0, T]$, Ω is an open bounded domain in \mathbb{R}^2 with sufficiently smooth boundary $\Gamma = \partial\Omega$, and $\alpha > 0$, $T > 0$, $y_d \in L^2(\Omega_T)$ are fixed. The precise smoothness requirements on Γ are given in the next section. The initial value y_0 is specified in Sect. 2. The constraints on control and state are specified through the closed and convex subsets

$$u \in K_U := \{u \in L^2(\Omega_T) : a \leq u(x, t) \leq b, \text{ for a.a. } (x, t) \in \Omega_T\} \subset U \tag{1.3}$$

for controls with $U := L^2(\Omega_T)$ and

$$y \in K_Y := \{y \in L^\infty(\Omega_T) : y(x, t) \leq \phi, \text{ for a.a. } (x, t) \in \Omega_T\} \tag{1.4}$$

for states, where for simplicity we assume that a, b and ϕ denote constants with $a < b$. Furthermore, $\mathcal{B} : L^2(\Omega_T) \rightarrow L^2(0, T; H^1(\Omega)^*)$ denotes the injection.

State constrained optimal control problems are important from the practical point of view. The numerical analysis for these problems is involved since the multipliers associated to constraints on the state in general are Borel measures. To the best of the authors' knowledge, there are only a few contributions to numerical analysis of parabolic optimal control problems with state constraints. Lavrentiev regularization of state constrained parabolic optimal control problems is studied in [24]. Recently, error estimates for state constrained parabolic control problem with controls of the form

$$(\mathcal{B}u)(x, t) := \sum_{i=1}^m u_i(t) f_i(x) \quad (x, t) \in \Omega_T, \tag{1.5}$$

are derived in [9], where $f_1, \dots, f_m \in H^1(\Omega) \cap L^\infty(\Omega)$ are given functions. Error analysis for optimal control problems with final state constraints and control constraints is considered in [29]. Finally, in [22] a priori error estimates for parabolic optimal control problems with pointwise state constraints in time are considered. Other related work on state constrained optimal control problems and parabolic control problems can be found in, e.g., [1, 3, 4, 7, 18, 21].

In this paper we consider an optimal control problem for the heat equation with distributed control and pointwise control and state constraints. The optimization problem is approximated using variational discretization proposed in [14] combined with linear finite elements in space and the dG(0) scheme in time for the discretization of the state equation. Based on an improved maximum error estimate for the state equation, we among other things derive in Theorem 4 the L^2 -norm error estimate

$$\begin{aligned} &\alpha \|u - u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \|y - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 \\ &\leq Cs^2 |\log h|^2 \left(h^{2-\frac{4}{s}} + k^{1-\frac{2}{s}} \right), \end{aligned} \tag{1.6}$$

where (u, y) and $(u_{h,k}, Y_{h,k})$ denote the continuous and discrete optimal controls and states with $4 < s < \infty$ denoting a positive real number related to the regularity of y , and h, k denote the space and time discretization parameters which have to be coupled appropriately. In the analysis we use results of [26] concerning the parabolic projection, which are only proved for two spatial dimensions. Therefore we restrict our analysis to the case $\Omega \subset \mathbb{R}^2$.

The rest of this paper is organized as follows. In Sect. 2 we present the state constrained optimal control problem and the corresponding optimality conditions. In Sect. 3 we establish the fully discrete approximation for the state equation and derive uniform estimates for the discretization error of the state. We obtain the a priori error estimates for the optimal control problem in Sect. 4. We also present some numerical experiments to support our theoretical findings.

2 Optimal control problem

Let $\Omega \subset \mathbb{R}^2$ be a convex domain which matches the smoothness requirements in [17, Chap. 4, § 4]. Specifically, we require that $\Gamma \in O^2$. We say that a surface $\Gamma \subset \mathbb{R}^n$ belongs to class O^l ($l \geq 1$) if there exists a number $\rho > 0$ such that the intersection of Γ with a ball K_ρ of radius ρ with center at an arbitrary point $x^0 \in \Gamma$ is a connected surface, which locally can be represented as the graph of a function ω of class O^l , i.e., in the local coordinate system (y_1, \dots, y_n) with origin at x^0 one has $y_n = \omega(y_1, \dots, y_{n-1})$ for $(y_1, \dots, y_n) \in \Gamma \cap K_\rho$, and ω is a function of class O^l on the projection of $\Gamma \cap K_\rho$ onto the subspace $y_n = 0$. Here $O^l(\bar{\Omega})$ is the set of all continuous functions in $\bar{\Omega}$ having continuous derivatives in $\bar{\Omega}$ up to order $l - 1$, with the derivatives of order $l - 1$ having a first differential at each point of $\bar{\Omega}$ and the derivative of order l being bounded in $\bar{\Omega}$ (see [17, Chap. 1, pp. 9–10] for more details). For example, convex domains Ω with $C^{2,1}$ boundary Γ meet the above mentioned requirements. Here we note that convex polygonal domains do not meet our regularity assumptions on the domain. However, convex polygonal domains could be included in our analysis by assuming that the states appearing in our analysis satisfy the required regularity assumptions.

For nonnegative integer m we adopt the standard notation $W^{m,s}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{m,s,\Omega}$ and seminorm $|\cdot|_{m,s,\Omega}$, and the standard modification for $s = \infty$. We denote by $H^m(\Omega)$ with norm $\|\cdot\|_{m,\Omega}$ and seminorm $|\cdot|_{m,\Omega}$ for $s = 2$. Note that $H^0(\Omega) = L^2(\Omega)$. For real numbers m we define the Sobolev space $W^{m,s}(\Omega)$ by interpolation, i.e., $W^{m,s}(\Omega) := (W^{k,s}(\Omega), W^{k+1,s}(\Omega))_{\theta,s}$, $k \in \mathbb{N}$, $m \in (k, k + 1)$, $\theta = m - [m]$. We denote the L^2 -inner products on $L^2(\Omega)$ by

$$(v, w) = \int_{\Omega} v w dx \quad \forall v, w \in L^2(\Omega).$$

With $\Omega_T = \Omega \times (0, T]$ let $H^{s,r}(\Omega_T) = L^2(0, T; H^s(\Omega)) \cap H^r(0, T; L^2(\Omega))$ equipped with the norm

$$\|w\|_{s,r,\Omega_T} = \left(\int_0^T \|w(\cdot, t)\|_{s,\Omega}^2 dt + \int_{\Omega} \|w(x, \cdot)\|_{r,[0,T]}^2 dx \right)^{\frac{1}{2}},$$

where $\|\cdot\|_{r,[0,T]}$ denotes the norm on $H^r([0, T])$. For a real positive number l , $C^{2l,l}(\bar{\Omega}_T)$ is the space of functions which are continuous in $\bar{\Omega}_T$, together with all derivatives of the form $D_t^r D_x^s$ for $2r + s < l$ (see [17], p. 7), where D_t and D_x denote the derivatives w.r.t time and space, respectively. Throughout the presentation c and C denote generic positive constants.

The variational form of problem (1.2) reads: Given $u \in U$ and $y_0 \in L^2(\Omega)$, find $y \in L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ such that

$$\left. \begin{aligned} (y_t, v) + (\nabla y, \nabla v) + (y, v) &= (\mathcal{B}u, v) \quad \forall v \in H^1(\Omega), \text{ a.e. } t \in (0, T], \\ y(x, 0) &= y_0(x) \quad x \in \Omega. \end{aligned} \right\} \quad (2.1)$$

We denote $y = \mathcal{G}(\mathcal{B}u)$ the solution to problem (2.1). It is well-known that if $\mathcal{B}u \in L^2(\Omega_T)$, $y_0 \in H^1(\Omega)$, problem (2.1) admits a unique solution $y = \mathcal{G}(\mathcal{B}u) \in H^{2,1}(\Omega_T) := L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)) \hookrightarrow C([0, T]; H^1(\Omega))$.

We define $W_s^{2,1}(\Omega_T)$ ($1 \leq s < \infty$) as

$$W_s^{2,1}(\Omega_T) := \{y \in L^s(0, T; W^{2,s}(\Omega)), y_t \in L^s(0, T; L^s(\Omega))\}$$

and use $\|\cdot\|_{2,1,s}$ to denote the norm defined on $W_s^{2,1}(\Omega_T)$. Thanks to the control constraints given by (1.3), we have $\mathcal{B}u \in L^\infty(\Omega_T)$. If in addition $y_0 \in W^{2-\frac{2}{s},s}(\Omega)$ for some $2 < s < \infty$, from [17, Chap. 4, Thm. 9.1] we with our assumptions on the domain Ω have the improved regularity $y \in W_s^{2,1}(\Omega_T)$ such that

$$\|y\|_{2,1,s} \leq C(\|\mathcal{B}u\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},s,\Omega})$$

with a positive constant C depending on Γ but not on s . We note that since $\mathcal{B}u \in L^\infty(\Omega_T)$, we obtain for $y_0 \in W^{2,\infty}(\Omega)$ that $y \in W_s^{2,1}(\Omega_T)$, $\forall 1 < s < \infty$. In what follows we fix $y_0 \in W^{2-\frac{2}{s},s}(\Omega)$ for some $2 < s < \infty$ which will be specified later, $y_0 < \phi$ and $\frac{\partial y_0}{\partial n} = 0$ on Γ throughout the paper.

Our optimal control problem reads:

$$\left. \begin{aligned} \min \quad J(y, u) &= \frac{1}{2} \|y - y_d\|_{L^2(\Omega_T)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega_T)}^2 \\ \text{s.t.} \quad y &= \mathcal{G}(\mathcal{B}u), \quad \text{and } y \in K_Y, u \in K_U. \end{aligned} \right\} \quad (2.2)$$

Since J is quadratic and K_U and K_Y are closed and convex, problem (2.2) admits a unique solution $(y, u) \in W_s^{2,1}(\Omega_T) \times K_U$. Note that $W_s^{2,1}(\Omega_T) \hookrightarrow C(\bar{\Omega}_T)$ holds for $s > 2$, so that it is meaningful to require

Assumption 1 (Slater condition) There exists $\hat{u} \in K_U$ such that the associated state \hat{y} fulfills (1.4) strictly, i.e. $\hat{y}(x, t) < \phi$ holds for all $(x, t) \in \bar{\Omega}_T$.

With this assumption it then follows from e.g., [5, 10, 24] that the first order optimality conditions for the optimal control problem (2.2) are given by

Theorem 1 Assume that $u \in L^\infty(\Omega_T)$ is the solution of problem (1.1) and let y be the corresponding state given by (2.1). Let $\mathcal{M}(\bar{\Omega}_T)$ denote the space of regular Borel measures on $\bar{\Omega}_T$. Then there exists an adjoint state $p \in L^q(0, T; W^{1,\sigma}(\Omega))$ for all $q, \sigma \in [1, 2)$ with $\frac{2}{q} + \frac{2}{\sigma} > 3$, and a Lagrange multiplier $\mu \in \mathcal{M}(\bar{\Omega}_T)$ such that

$$\left. \begin{aligned} -p_t - \Delta p + p &= y - y_d + \mu_{\Omega_T} && \text{in } \Omega_T, \\ \frac{\partial p}{\partial n} &= \mu_{\Gamma_T} && \text{on } \Gamma_T, \\ p(T) &= \mu_T && \text{in } \Omega \end{aligned} \right\} \tag{2.3}$$

is satisfied in the sense of transposition, and

$$\int_{\Omega_T} (\alpha u + \mathcal{B}^* p)(v - u) \geq 0 \quad \forall v \in K_U, \tag{2.4}$$

$$\mu \geq 0, \quad y(x, t) \leq \phi, \quad (x, t) \in \bar{\Omega}_T, \quad \text{and} \quad \int_{\bar{\Omega}_T} (\phi - y) d\mu = 0 \tag{2.5}$$

holds. Here $\mu_{\Omega_T} := \mu|_{\Omega_T}$, $\mu_{\Gamma_T} := \mu|_{\Gamma_T}$ and $\mu_T := \mu|_{\bar{\Omega} \times \{T\}}$.

Proof For a proof we refer to e.g., [5, 9]. We note that in [9] a proof for a slightly different setting is provided whose adaption to the present situation is obvious. \square

Let us note that (2.3) is satisfied in the sense of transposition (see [19]), if

$$\int_{\Omega_T} (w_t - \Delta w + w)p + \int_{\Gamma_T} \frac{\partial w}{\partial n} p = \int_{\Omega_T} (y - y_d)w + \int_{\bar{\Omega}_T} w d\mu \tag{2.6}$$

$$\forall w \in W_0^\infty$$

holds, where

$$W_0^\infty := \left\{ w \in H^{2,1}(\Omega_T) \cap C(\bar{\Omega}_T) : w(\cdot, 0) = 0 \text{ in } \bar{\Omega}, \right. \\ \left. w_t - \Delta w + w \in L^\infty(\Omega_T), \quad \frac{\partial w}{\partial n} \in L^\infty(\Gamma_T) \right\}.$$

3 Finite element discretization and error estimates for the state equation

Let $\Omega^h \subset \Omega$ be a polygonal approximation to Ω with a boundary $\Gamma_h = \partial\Omega^h$. From the assumption on the boundary $\partial\Omega$ we have $|\Omega \setminus \Omega^h| \leq Ch^2$. For the spatial discretization let \mathcal{T}^h be a quasi-uniform partitioning of Ω^h into disjoint regular triangles or rectangles τ , so that $\bar{\Omega}^h = \bigcup_{\tau \in \mathcal{T}^h} \bar{\tau}$. We assume that all vertices of \mathcal{T}^h that are on the boundary Γ_h stay on the boundary $\partial\Omega$. Let h_τ denote the diameter of τ . Set $h = \max_{\tau \in \mathcal{T}^h} h_\tau$. Associated with \mathcal{T}^h is a finite dimensional space V^h consisting of piecewise linear and continuous polynomials. We note that functions in V^h can

continuously be extended to Ω such that $V^h \subset C(\bar{\Omega})$ holds, i.e, for a boundary element τ , the function v_h defined on the part of $\Omega \setminus \Omega^h$ which shares one edge with the element τ is the linear extension of $v_h|_\tau$. It is easy to see that $V^h \subset H^1(\Omega)$. Since \mathcal{T}^h is quasi-uniform, the following inverse estimates (see [6])

$$\|v_h\|_{s,\Omega} \leq Ch^{l-s} \|v_h\|_{l,\Omega}, \quad 0 \leq l \leq s \leq 1, \tag{3.1}$$

$$\|v_h\|_{0,\infty,\Omega} \leq Ch^{-1} \|v_h\|_{0,\Omega} \tag{3.2}$$

hold for all $v_h \in V^h$.

Remark 1 In [8] a quasi-uniform partition with curved boundary elements is proposed for the approximation of elliptic optimal control problems with state constraints. This approach also would be applicable in the present situation.

Let $\Pi_h : C(\bar{\Omega}) \rightarrow V^h$ denote the standard Lagrange interpolation operator. Interpolation error estimates imply that for $y \in W^{m,r}(\Omega)$, $r > 2$ (see, e.g., [6])

$$\|y - \Pi_h y\|_{0,r,\Omega} + h \|y - \Pi_h y\|_{1,r,\Omega} \leq Ch^m \|y\|_{m,r,\Omega} \quad 1 \leq m \leq 2. \tag{3.3}$$

Let $\mathcal{R}_h : H^1(\Omega) \rightarrow V^h$ denote the Ritz projection operator defined as

$$(\nabla \mathcal{R}_h y, \nabla v_h) + (\mathcal{R}_h y, v_h) = (\nabla y, \nabla v_h) + (y, v_h) \quad \forall v_h \in V^h. \tag{3.4}$$

Lemma 1 *Let \mathcal{R}_h be the Ritz projection operator defined above. Then there holds:*

$$\|y - \mathcal{R}_h y\|_{0,r,\Omega} \leq Crh \inf_{v_h \in V^h} \|y - v_h\|_{1,r,\Omega} \quad r \geq 2. \tag{3.5}$$

$$\|y - \mathcal{R}_h y\|_{0,\infty,\Omega} \leq C |\log h| \inf_{v_h \in V^h} \|y - v_h\|_{0,\infty,\Omega}. \tag{3.6}$$

Proof A result related to (3.5) is proved by Rannacher and Scott in [28] for Dirichlet boundary conditions, but the arguments can be adapted to the present situation, we omit the details here. The result of (3.6) can be found in [30]. \square

Then the semi-discrete finite element approximation of (2.1) reads: Given $u \in U$ and $y_0^h \in V^h$, find $y_h(t) \in H^1(0, T; V^h)$ such that

$$\left. \begin{aligned} \left(\frac{\partial y_h}{\partial t}, w_h \right) + (\nabla y_h, \nabla w_h) + (y_h, w_h) &= (Bu, w_h) \quad \forall w_h \in V^h, \quad t \in (0, T], \\ y_h(x, 0) &= y_0^h(x) \quad x \in \Omega \end{aligned} \right\} \tag{3.7}$$

is satisfied. Here $y_0^h = \Pi_h y_0 \in V^h$ is an approximation to y_0 .

We next consider the fully discrete approximation for above semidiscrete problem. Let $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$ be a time grid with $t_n = nk$, $n = 1, 2, \dots, N$, where $k := \frac{T}{N}$. Let $I_n = (t_{n-1}, t_n]$ and

$$V_{h,k} := \{ \phi : \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}, \phi(\cdot, t)|_{\bar{\Omega}} \in V^h, \phi(x, \cdot)|_{I_n} \in \mathbb{P}_0 \text{ for } n = 1, \dots, N \},$$

i.e. $\phi \in V_{h,k}$ is a piecewise constant polynomial w.r.t. time. For $Y, \Phi \in V_{h,k}$ we set

$$A(Y, \Phi) := \sum_{n=1}^N ((Y^n - Y^{n-1}, \Phi^n) + k(\nabla Y^n, \nabla \Phi^n) + k(Y^n, \Phi^n)) + (Y_+^0, \Phi_+^0),$$

where $\Phi^n := \Phi_-^n, \Phi_{\pm}^n = \lim_{s \rightarrow 0^{\pm}} \Phi(t_n + s)$.

The fully discrete dG(0)–cG(1) approximation scheme for (3.7) now reads: Given $u \in U$, find $Y_{h,k} \in V_{h,k}$ such that

$$A(Y_{h,k}, \Phi) = \sum_{n=1}^N \int_{I_n} (\mathcal{B}u, \Phi^n) + (y_0, \Phi_+^0), \quad \forall \Phi \in V_{h,k}. \tag{3.8}$$

Note that on each time interval I_n , the solution $Y_{h,k}^n \in V^h$ satisfies

$$\left. \begin{aligned} & \left(\frac{Y_{h,k}^n - Y_{h,k}^{n-1}}{k}, w_h \right) + (\nabla Y_{h,k}^n, \nabla w_h) + (Y_{h,k}^n, w_h) = ((\overline{\mathcal{B}u})^n, w_h), \\ & \forall w_h \in V^h; \quad n = 1, \dots, N, \quad Y_{h,k}^0(x) = y_0^h(x) \quad x \in \Omega, \end{aligned} \right\} \tag{3.9}$$

where $\overline{\mathcal{B}u} = (\frac{1}{k} \int_{t_{n-1}}^{t_n} \mathcal{B}u)_{n=1}^N$ and $y_0^h \in V^h$ is an approximation to y_0 .

Now we are in a position to estimate the error between the solutions of problem (2.1) and (3.8). The following result is a standard consequence of error estimates for parabolic equation (see, e.g., [12]).

Theorem 2 *Let $\mathcal{B}u \in L^2(\Omega_T)$, let $y \in H^{2,1}(\Omega_T)$ be the solution to problem (2.1), and let $Y_{h,k} \in V_{h,k}$ be the solution to problem (3.8). Then we have*

$$\|y - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))} \leq C(h^2 + k)\|y\|_{2,1,\Omega_T}. \tag{3.10}$$

We also need the maximum norm estimates for the state equation. Using ideas of [26] it is convenient to introduce the weighted-norm technique. For this purpose let

$$\rho(x) := (|x - z|^2 + \omega^2)^{\frac{1}{2}} \quad \forall x \in \Omega,$$

where $z \in \Omega$ and $\omega = Ch|\log h|$. The choice of $C \geq 1$ and z will be specified latter. It is easy to verify (see, e.g., Lemma 4.3 in [26] or p. 216 in [2]) that

$$\int_{\Omega} \rho(x)^{-m} dx \leq C((m - 2)\omega^{m-2})^{-1} \quad \text{for } m > 2. \tag{3.11}$$

Theorem 3 *Let $\mathcal{B}u \in L^\infty(\Omega_T)$, $y \in W_s^{2,1}(\Omega_T)$ ($4 < s < \infty$) be the solution of problem (2.1), and $Y_{h,k} \in V_{h,k}$ be the solution of problem (3.8). There exists $C^* \geq 1$ such that if $\omega = C^*h|\log h|$ and $k \geq C^*h^2|\log h|^3$, then*

$$\begin{aligned} & \max_{1 \leq n \leq N} \|y(\cdot, t_n) - Y_{h,k}^n\|_{0,\infty,\Omega} \\ & \leq Cs^2 |\log h|^2 (h^{2-4/s} + k^{1-2/s}) (\|\mathcal{B}u\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},2,\Omega}) \end{aligned} \quad (3.12)$$

holds, where $Y_{h,k}^n$ ($n = 1, \dots, N$) is the solution of (3.8).

Proof The proof follows [26] (see, e.g., pp. 501–504), where an error estimate for the parabolic projection in the case of Dirichlet boundary conditions is presented. The proof in the present situation is slightly different. We sketch it for the convenience of the reader. We also note that we impose Neumann boundary conditions for the state.

With

$$\bar{y}^n := \frac{1}{k} \int_{t_{n-1}}^{t_n} y(\cdot, t) dt$$

we have

$$\begin{aligned} y(\cdot, t_n) - Y_{h,k}^n &= y(\cdot, t_n) - \mathcal{R}_h \bar{y}^n + \mathcal{R}_h \bar{y}^n - Y_{h,k}^n \\ &= \xi^n - \eta^n. \end{aligned}$$

Thus

$$\|y(\cdot, t_n) - Y_{h,k}^n\|_{0,\infty,\Omega} \leq \|y(\cdot, t_n) - \mathcal{R}_h \bar{y}^n\|_{0,\infty,\Omega} + \|\mathcal{R}_h \bar{y}^n - Y_{h,k}^n\|_{0,\infty,\Omega}. \quad (3.13)$$

Note that (see [17, 26])

$$W_s^{2,1}(\Omega_T) \hookrightarrow C^{2-\frac{m}{s}, 1-\frac{m}{2s}}(\bar{\Omega}_T)$$

with $m = 4 < s < \infty$. From (3.3) and (3.6) we deduce

$$\begin{aligned} & \|y(\cdot, t_n) - \mathcal{R}_h \bar{y}^n\|_{0,\infty,\Omega} \\ & \leq \|y(\cdot, t_n) - \bar{y}^n\|_{0,\infty,\Omega} + \|\bar{y}^n - \mathcal{R}_h \bar{y}^n\|_{0,\infty,\Omega} \\ & \leq \frac{1}{k} \int_{t_{n-1}}^{t_n} \|y(t_n) - y(t')\|_{0,\infty,\Omega} dt' + C |\log h| \|\bar{y}^n - \Pi_h \bar{y}^n\|_{0,\infty,\Omega} \\ & \leq C (k^{1-\frac{2}{s}} + h^{2-\frac{4}{s}} |\log h|) \|y\|_{2,1,s}, \end{aligned} \quad (3.14)$$

where interpolation error estimates in space and time are used (see e.g., [2]).

It remains to estimate $\|\mathcal{R}_h \bar{y}^n - Y_{h,k}^n\|_{0,\infty,\Omega}$. Suppose that $z \in \Omega$ is the point where $\|\mathcal{R}_h \bar{y}^n - Y_{h,k}^n\|_{0,\infty,\Omega}$ achieves its maximal value. Using Thm. 3.3.3 (p. 151) in [6] (see also [25]) we have

$$\|\mathcal{R}_h \bar{y}^n - Y_{h,k}^n\|_{0,\infty,\Omega} \leq C \frac{\omega}{h} \|\rho^{-1} \eta^n\|_{0,\Omega}. \quad (3.15)$$

Integration of (2.1) from t_{i-1} to t_i yields

$$(y^i - y^{i-1}, v) + \int_{t_{i-1}}^{t_i} (\nabla y, \nabla v) + \int_{t_{i-1}}^{t_i} (y, v) = \int_{t_{i-1}}^{t_i} (\mathcal{B}u, v),$$

where $y^i = y(x, t_i)$, $i = 0, 1, \dots, N$, so that using (3.9) we have for all $v \in V^h$

$$\begin{aligned}
 & (\eta^i - \eta^{i-1}, v) + k(\nabla \eta^i, \nabla v) + k(\eta^i, v) \\
 &= (Y_{h,k}^i - Y_{h,k}^{i-1}, v) + k(\nabla Y_{h,k}^i, \nabla v) + k(Y_{h,k}^i, v) \\
 &\quad - (\mathcal{R}_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h \nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h y, v) \\
 &= \int_{t_{i-1}}^{t_i} (\mathcal{B}u, v) - (\mathcal{R}_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h \nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h y, v) \\
 &= (y^i - y^{i-1}, v) + \int_{t_{i-1}}^{t_i} (\nabla y, \nabla v) + \int_{t_{i-1}}^{t_i} (y, v) \\
 &\quad - (\mathcal{R}_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h \nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (\mathcal{R}_h y, v) \\
 &= (\xi^i - \xi^{i-1}, v). \tag{3.16}
 \end{aligned}$$

Let $Z^{i-1} \in V^h$, $i = 1, 2, \dots, n$ be the unique solution of following backward fully discrete problem

$$(Z^{i-1} - Z^i, w_h) + k(\nabla Z^{i-1}, \nabla w_h) + k(Z^{i-1}, w_h) = 0 \quad \forall w_h \in V^h \tag{3.17}$$

with $Z^n = \zeta \in V^h$, $1 \leq n \leq N$. Now let $v = Z^{i-1}$ in (3.16). Summing from 1 to n we find

$$(\eta^n, \zeta) = (\xi^n, \zeta) + \sum_{i=1}^n (\xi^i, Z^{i-1} - Z^i) + (Y_h^0 - y_0, Z^0), \tag{3.18}$$

where we have used (3.17) and the fact that $\eta^i \in V^h$. Setting $\zeta = \mathcal{P}_h(\rho^{-2}\eta^n)$ in (3.18), where $\mathcal{P}_h : L^2(\Omega) \rightarrow V^h$ is the L^2 -projection operator, we deduce from $\eta^n \in V^h$ that

$$\begin{aligned}
 \|\rho^{-1}\eta^n\|_{0,\Omega}^2 &= (\eta^n, \rho^{-2}\eta^n) = (\eta^n, \mathcal{P}_h(\rho^{-2}\eta^n)) \\
 &= (\xi^n, \mathcal{P}_h(\rho^{-2}\eta^n)) + \sum_{i=1}^n (\xi^i, Z^{i-1} - Z^i) + (Y_h^0 - y_0, Z^0). \tag{3.19}
 \end{aligned}$$

Since $\|\rho \mathcal{P}_h v\|_{0,\Omega} \leq C \|\rho v\|_{0,\Omega}$ we have

$$\begin{aligned}
 (\xi^n, \mathcal{P}_h(\rho^{-2}\eta^n)) &\leq \|\rho^{-1}\xi^n\|_{0,\Omega} \|\rho \mathcal{P}_h(\rho^{-2}\eta^n)\|_{0,\Omega} \\
 &\leq \|\rho^{-1}\xi^n\|_{0,\Omega} \|\rho^{-1}\eta^n\|_{0,\Omega}.
 \end{aligned}$$

Thus Young’s inequality implies

$$\begin{aligned} \|\rho^{-1}\eta^n\|_{0,\Omega}^2 &\leq C\|\rho^{-1}\xi^n\|_{0,\Omega}^2 + C\sum_{i=1}^n\|\rho^{-1}\xi^i\|_{0,\Omega}\|\rho(Z^{i-1} - Z^i)\|_{0,\Omega} \\ &\quad + C\|\rho^{-1}(\Pi_h y_0 - y_0)\|_{0,\Omega}\|\rho Z^0\|_{0,\Omega}. \end{aligned} \tag{3.20}$$

Now we need a priori estimates for $Z^{i-1} - Z^i$ and Z^0 in weighted norms. According to [11] there exists $C^* \geq 1$ such that if $\omega = C^*h|\log h|$ and $k \geq C^*h^2|\log h|^3$, the estimate

$$\begin{aligned} \|\rho Z^0\|_{0,\Omega} + \left(k^{-1}\sum_{i=1}^n t_{n-i+1}\|\rho(Z^{i-1} - Z^i)\|_{0,\Omega}^2\right)^{\frac{1}{2}} \\ \leq C|\log h|\|\rho\xi\|_{0,\Omega} \leq C|\log h|\|\rho^{-1}\eta^n\|_{0,\Omega} \end{aligned} \tag{3.21}$$

holds. We note that the estimate (3.21) was proved in [11] for problems with Dirichlet boundary conditions defined on polygonal domain. The technique used there was to exploit the properties of the L^2 -projections in weighted norms, and can be adapted to our cases where smooth domain and homogeneous Neumann boundary conditions are employed.

By exploiting the property of the weight function ρ , the following estimates can be found in, e.g., [26], and generalized to our case:

$$\left(k\sum_{i=1}^n t_{n-i+1}^{-1}\|\rho^{-1}\xi^i\|_{0,\Omega}^2\right)^{\frac{1}{2}} \leq Csh^{-\frac{2}{s}}k^{-\frac{1}{s}}\left(k\sum_{i=1}^n\|\xi^i\|_{0,s,\Omega}^s\right)^{\frac{1}{s}}, \tag{3.22}$$

$$\|\rho^{-1}\xi^n\|_{0,\Omega} \leq Csh^{-\frac{2}{s}}k^{-\frac{1}{s}}\left(k\sum_{i=1}^n\|\xi^i\|_{0,s,\Omega}^s\right)^{\frac{1}{s}}, \tag{3.23}$$

where Hölder’s inequality and property (3.11) are used. Then the interpolation error estimate (3.3), (3.11) and Hölder’s inequality with $q = \frac{s}{s-2}$ lead to

$$\begin{aligned} \|\rho^{-1}(\Pi_h y_0 - y_0)\|_{0,\Omega} &\leq \left(\int_{\Omega}\rho^{-2q}\right)^{1/2q}\|\Pi_h y_0 - y_0\|_{0,s,\Omega} \\ &\leq C((2q - 2)\omega^{2q-2})^{-\frac{1}{2q}}h^{2-\frac{2}{s}}\|y_0\|_{2-\frac{2}{s},s,\Omega} \\ &\leq C\left(\frac{s-2}{4}\right)^{\frac{1}{2}-\frac{2}{s}}\omega^{-\frac{2}{s}}h^{2-\frac{2}{s}}\|y_0\|_{2-\frac{2}{s},s,\Omega} \\ &\leq Cs^{\frac{1}{2}}h^{2-\frac{4}{s}}\|y_0\|_{2-\frac{2}{s},s,\Omega} \end{aligned} \tag{3.24}$$

as well as to

$$\left(k\sum_{i=1}^n\|\xi^i\|_{0,s,\Omega}^s\right)^{\frac{1}{s}} \leq Cs(h^2 + k)\|y\|_{2,1,s}, \tag{3.25}$$

see [26, Lemma 4.9]. The later one used the L^s -norm error estimates (3.5) for the Ritz-projection \mathcal{R}_h . From (3.20)–(3.25) we then have that

$$\|\rho^{-1}\eta^n\|_{0,\Omega} \leq Cs^2|\log h|(h^{2-\frac{4}{s}} + k^{1-\frac{1}{s}})(\|\mathcal{B}u\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},2,\Omega}). \tag{3.26}$$

With (3.15) and (3.26) we conclude that

$$\begin{aligned} &\|\mathcal{R}_h\bar{y}^n - Y_{h,k}^n\|_{0,\infty,\Omega} \\ &\leq Cs^2|\log h|^2(h^{2-\frac{4}{s}} + k^{1-\frac{1}{s}})(\|\mathcal{B}u\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},2,\Omega}). \end{aligned} \tag{3.27}$$

Combining (3.14) and (3.27) we complete the proof of the theorem. □

Remark 2 A uniform error estimate for the discretized error of (2.1) and (3.8) is derived in [9] under the condition that the right hand side and hence the time derivative of the solution is only square integrable in time. Here the right hand side is uniformly bounded w.r.t. space and time, which guarantees an improved regularity of the solution and thus an improved error estimate.

4 Error estimates for optimal control problem

In this section we consider the finite element approximation and error estimates for optimal control problem (1.1)–(1.2).

We consider the variational discretization approach proposed in [9, 14]. Then the fully discrete optimization problem reads

$$\min_{u \in K_U} J_h(Y_{h,k}, u) = \sum_{i=1}^N \frac{1}{2}k \int_{\Omega} (Y_{h,k}^i - \bar{y}_d^i)^2 + \frac{\alpha}{2} \int_0^T \int_{\Omega} u^2 \tag{4.1}$$

subject to

$$\left. \begin{aligned} A(Y_{h,k}, \Phi) &= \sum_{n=1}^N \int_{I_n} (\mathcal{B}u, \Phi) + (y_0, \Phi_+^0), \quad \forall \Phi \in V_{h,k}. \\ Y_{h,k}^i(x_j) &\leq \phi, \quad i = 1, \dots, N, \quad j = 1, \dots, m, \end{aligned} \right\} \tag{4.2}$$

where m denotes the number of nodes in the triangulation \mathcal{T}^h . Let \hat{u} denote the control satisfying Assumption 1, i.e., there exists $\delta > 0$ such that the corresponding state $\hat{y} = \mathcal{G}(\mathcal{B}\hat{u})$ satisfies

$$\hat{y} = \mathcal{G}(\mathcal{B}\hat{u}) \leq \phi - \delta \quad \text{in } \bar{\Omega}_T. \tag{4.3}$$

It follows from Theorem 3 that there exist $h_0, k_0 > 0$ such that $\hat{Y}_{h,k} := \mathcal{G}_{h,k}(\mathcal{B}\hat{u}) \in V_{h,k}$ satisfies

$$\hat{Y}_{h,k}^n(x_j) \leq \phi - \frac{\delta}{2} < \phi, \quad 1 \leq n \leq N, \quad 0 < h \leq h_0, \quad 0 < k \leq k_0. \tag{4.4}$$

Thus the pair $(\hat{u}, \hat{Y}_{h,k})$ is a discrete feasible Slater point for problem (4.1) for h and k small enough.

As a minimization problem for a quadratic functional over a closed convex set, the variational discrete optimization problem (4.1)–(4.2) admits a unique solution $u_{h,k} \in K_U$ with corresponding state $Y_{h,k} \in V_{h,k}$. Furthermore, it follows from [5] again that the discrete Slater condition (4.4) guarantees the existence of a discrete co-state $P_{h,k} \in V_{h,k}$ and discrete Lagrange multiplier $\mu_j^i \in \mathbb{R}$, $i = 1, \dots, N$, $j = 1, \dots, m$, such that the triplet $(Y_{h,k}, P_{h,k}, u_{h,k}) \in V_{h,k} \times V_{h,k} \times K_U$, satisfies the following optimality conditions:

$$A(Y_{h,k}, \Phi) = \sum_{n=1}^N \int_{I_n} (\mathcal{B}u_{h,k}, \Phi^n) + (y_0, \Phi_+^0), \quad \forall \Phi \in V_{h,k}, \tag{4.5}$$

$$A(\Phi, P_{h,k}) = \sum_{n=1}^N \int_{I_n} (Y_{h,k}^n - y_d, \Phi^n) + \sum_{i=1}^N \sum_{j=1}^m \Phi^i(x_j) \mu_j^i, \quad \forall \Phi \in V_{h,k}, \tag{4.6}$$

$$\int_0^T \int_{\Omega} (\alpha u_{h,k} + \mathcal{B}^* P_{h,k})(v - u_{h,k}) dx dt \geq 0 \quad \forall v \in K_U, \tag{4.7}$$

$$\mu_j^i \geq 0; \quad Y_{h,k}^i(x_j) \leq \phi, \quad \text{and} \quad \sum_{i=1}^N \sum_{j=1}^m (\phi - Y_{h,k}^i(x_j)) \mu_j^i = 0. \tag{4.8}$$

It follows from (4.7) that $u_{h,k}$ is piecewise constant w.r.t. time, but in general $u_{h,k}$ is not a finite element function w.r.t. space. It is easy to show that

$$u_{h,k} = \mathcal{P}_{K_U} \left(-\frac{1}{\alpha} \mathcal{B}^* P_{h,k} \right),$$

where \mathcal{P}_{K_U} denotes the orthogonal projection in U onto K_U . Let us define measure $\mu_{h,k} \in \mathcal{M}(\bar{\Omega}_T)$ by

$$\int_{\bar{\Omega}_T} f d\mu_{h,k} := \sum_{i=1}^N \sum_{j=1}^m f(x_j, t_i) \mu_j^i \quad \forall f \in C(\bar{\Omega}_T). \tag{4.9}$$

For $v_{h,k} \in V_{h,k}$ we use the notation

$$\int_{\bar{\Omega}_T} v_{h,k} d\mu_{h,k} := \sum_{i=1}^N \sum_{j=1}^m v_{h,k}^i(x_j) \mu_j^i. \tag{4.10}$$

As a first result for (4.1)–(4.2) we prove that the sequence of optimal controls, states and measures μ_h are uniformly bounded.

Lemma 2 *Let $(Y_{h,k}, u_{h,k}) \in V_{h,k} \times K_U$ be the solutions of problem (4.1)–(4.2), $P_{h,k} \in V_{h,k}$ and $\mu_{h,k} \in \mathcal{M}(\bar{\Omega}_T)$ be the corresponding adjoint state and measure,*

respectively. Then there exists $h_0 > 0$ such that

$$\|u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \|Y_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \sum_{i=1}^N \sum_{j=1}^m \mu_j^i \leq C \quad \text{for all } 0 < h \leq h_0.$$

Proof The proof follows [9]. From (4.4) and (4.8) we obtain

$$\begin{aligned} & \frac{\delta}{2} \sum_{i=1}^N \sum_{j=1}^m \mu_j^i \\ & \leq \sum_{i=1}^N \sum_{j=1}^m (\phi - \hat{Y}_{h,k}^i(x_j)) \mu_j^i = \sum_{i=1}^N \sum_{j=1}^m (Y_{h,k}^i(x_j) - \hat{Y}_{h,k}^i(x_j)) \mu_j^i \\ & = A(Y_{h,k} - \hat{Y}_{h,k}, P_{h,k}) - \sum_{i=1}^N k(Y_{h,k}^i - (\bar{y}_d)^i, Y_{h,k}^i - \hat{Y}_{h,k}^i) \\ & = \sum_{i=1}^N \int_{I_i} (\mathcal{B}u_{h,k} - \mathcal{B}\hat{u}, P_{h,k}) - \sum_{i=1}^N k(Y_{h,k}^i - (\bar{y}_d)^i, Y_{h,k}^i - \hat{Y}_{h,k}^i) \\ & \leq \sum_{i=1}^N \int_{I_i} (\alpha u_{h,k}, \hat{u} - u_{h,k}) + k \int_{\Omega} (-(Y_{h,k}^i)^2 + Y_{h,k}^i \hat{Y}_{h,k}^i + Y_{h,k}^i (\bar{y}_d)^i - \hat{Y}_{h,k}^i (\bar{y}_d)^i) \\ & \leq -\frac{1}{2} \sum_{i=1}^N \int_{I_i} \|Y_{h,k}^i\|_{0,\Omega}^2 - \frac{\alpha}{2} \sum_{i=1}^N \int_{I_i} \|u_{h,k}^i\|_{0,\Omega}^2 + C, \end{aligned} \tag{4.11}$$

where we have used (4.5) and (4.6). This completes the proof of the Lemma. □

Now we are in a position to prove the main result of this paper. We use a proof technique developed in Chap. 3 of [15] which only relies on uniform a priori error estimates of the state approximation.

Theorem 4 *Let $(y, u) \in W_s^{2,1}(\Omega_T) \times L^\infty(\Omega_T)$, $4 < s < \infty$ and $(Y_{h,k}, u_{h,k}) \in V_{h,k} \times K_U$ be the solutions to problem (2.2) and (4.1)–(4.2). Assume that the assumptions of Theorem 3 are all satisfied. Then we have the following a priori error estimate:*

$$\begin{aligned} & \alpha \|u - u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \|y - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 \\ & \leq Cs^2 |\log h|^2 (h^{2-\frac{4}{s}} + k^{1-\frac{2}{s}}). \end{aligned} \tag{4.12}$$

Proof From (2.4) and (4.7) we have

$$\int_{\Omega_T} (\alpha u + \mathcal{B}^* p)(v - u) \geq 0 \quad \forall v \in K_U$$

and

$$\int_{\Omega_T} (\alpha u_{h,k} + \mathcal{B}^* P_{h,k})(v_h - u_{h,k}) dx \geq 0 \quad \forall v_h \in K_U.$$

Choosing $v = u_{h,k}$ as well as $v_h = u$, and adding these inequalities gives

$$\alpha \|u - u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 \leq \int_0^T (p - P_{h,k}, \mathcal{B}(u_{h,k} - u)). \tag{4.13}$$

In the following we need to introduce some auxiliary problems. Let $y^h := \mathcal{G}(\mathcal{B}u_{h,k}) \in W_s^{2,1}(\Omega_T)$ be the variational solution of

$$\left. \begin{aligned} y_t^h - \Delta y^h + y^h &= \mathcal{B}u_{h,k} && \text{in } \Omega_T, \\ \frac{\partial y^h}{\partial n} &= 0 && \text{on } \Gamma_T, \\ y^h(0) &= y_0 && \text{in } \Omega, \end{aligned} \right\} \tag{4.14}$$

i.e.

$$\left(\frac{\partial y^h}{\partial t}, v \right) + (\nabla y^h, \nabla v) + (y^h, v) = (\mathcal{B}u_{h,k}, v) \quad \text{for a.a. } t \in (0, T], \quad \forall v \in H^1(\Omega)$$

with $y^h(\cdot, 0) = y_0$, and let $Y_{h,k}(u) := \mathcal{G}_{h,k}(\mathcal{B}u) \in V_{h,k}$ be the solution of

$$A(Y_{h,k}(u), \Phi) = \sum_{n=1}^N \int_{I_n} (\mathcal{B}u, \Phi^n) + (y_0, \Phi_+^0), \quad \forall \Phi \in V_{h,k}. \tag{4.15}$$

Note that $Y_{h,k}$ and $Y_{h,k}(u)$ are the fully discrete approximations to y^h and y , respectively. Then from (2.1) and (4.14) we have $y^h - y \in H^{2,1}(\Omega_T) \cap C(\bar{\Omega}_T)$, $(y^h - y)_t - \Delta(y^h - y) + (y^h - y) = \mathcal{B}(u_{h,k} - u) \in L^\infty(\Omega_T)$, $(y^h - y)(x, 0) = 0$ and $\frac{\partial(y^h - y)}{\partial n} = 0$. Thus $y^h - y \in W_0^\infty$. Now (2.6) and (4.14) imply that

$$\begin{aligned} \int_0^T (p, \mathcal{B}(u_{h,k} - u)) &= \int_0^T \left(\frac{\partial(y^h - y)}{\partial t} - \Delta(y^h - y) + (y^h - y), p \right) \\ &= \int_0^T (y - y_d, y^h - y) + \int_{\bar{\Omega}_T} (y^h - y) d\mu. \end{aligned} \tag{4.16}$$

Similarly, from (4.6) and (4.15) we have

$$\begin{aligned} \int_0^T (P_{h,k}, \mathcal{B}(u - u_{h,k})) &= A(Y_{h,k}(u) - Y_{h,k}, P_{h,k}) \\ &= \int_0^T (Y_{h,k} - y_d, Y_{h,k}(u) - Y_{h,k}) \\ &\quad + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k}. \end{aligned} \tag{4.17}$$

Thus

$$\begin{aligned}
 & \alpha \|u - u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 \\
 & \leq \int_0^T (y - y_d, y^h - y) + \int_{\bar{\Omega}_T} (y^h - y) d\mu \\
 & \quad + \int_0^T (Y_{h,k} - y_d, Y_{h,k}(u) - Y_{h,k}) + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k} \\
 & = \int_0^T (y - y_d, y^h - Y_{h,k}) + \int_0^T (y - y_d, Y_{h,k} - y) \\
 & \quad + \int_0^T (Y_{h,k} - y_d, Y_{h,k}(u) - y) + \int_0^T (Y_{h,k} - y_d, y - Y_{h,k}) \\
 & \quad + \int_{\bar{\Omega}_T} (y^h - y) d\mu + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k} \\
 & = -\|y - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \int_0^T (y - y_d, y^h - Y_{h,k}) \\
 & \quad + \int_0^T (Y_{h,k} - y_d, Y_{h,k}(u) - y) \\
 & \quad + \int_{\bar{\Omega}_T} (y^h - y) d\mu + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k}.
 \end{aligned}$$

This implies

$$\begin{aligned}
 & \alpha \|u - u_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 + \|y - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))}^2 \\
 & \leq \int_0^T (y - y_d, y^h - Y_{h,k}) + \int_0^T (Y_{h,k} - y_d, Y_{h,k}(u) - y) \\
 & \quad + \int_{\bar{\Omega}_T} (y^h - y) d\mu + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k} \\
 & \leq C(\|y^h - Y_{h,k}\|_{L^2(0,T;L^2(\Omega))} + \|Y_{h,k}(u) - y\|_{L^2(0,T;L^2(\Omega))}) \\
 & \quad + \int_{\bar{\Omega}_T} (y^h - y) d\mu + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k} \\
 & \leq C(h^2 + k)(\|\mathcal{B}u\|_{0,\infty,\Omega} + \|\mathcal{B}u_{h,k}\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},s,\Omega}) \\
 & \quad + \int_{\bar{\Omega}_T} (y^h - y) d\mu + \int_{\bar{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k}, \tag{4.18}
 \end{aligned}$$

where we have used Theorem 2. From (2.5) we have

$$\begin{aligned}
 \int_{\bar{\Omega}_T} (y^h - y) d\mu & = \int_{\bar{\Omega}_T} (y^h - \phi + \phi - y) d\mu \\
 & = \int_{\bar{\Omega}_T} (y^h - \phi) d\mu.
 \end{aligned}$$

Since $\mu \geq 0$, with $(y^h)^+ = \max(y^h, 0)$ we have

$$\int_{\tilde{\Omega}_T} (y^h - \phi) d\mu \leq \int_{\tilde{\Omega}_T} (y^h - \phi)^+ d\mu.$$

Define $y^{h,n} = y^h(x, t_n)$, so that for $t_{n-1} < t \leq t_n$ we have

$$\begin{aligned} & |(y^h - \phi)^+(x, t)| \\ & \leq |(y^h - \phi)^+(x, t) - (\overline{y^{h^n}} - \phi)^+(x)| + |(\overline{y^{h^n}} - \phi)^+(x) - (Y_{h,k}^n - \phi)^+(x)| \\ & \quad + |(Y_{h,k}^n - \phi)^+(x)| \\ & \leq |y^h(x, t) - \overline{y^{h^n}}(x)| + |\overline{y^{h^n}}(x) - Y_{h,k}^n(x)| + |(Y_{h,k}^n - \phi)^+(x)| \\ & \leq C \frac{1}{k} \int_{t_{n-1}}^{t_n} \|y^h(t) - y^h(t')\|_{0,\infty,\Omega} dt' + C \frac{1}{k} \int_{t_{n-1}}^{t_n} \|y^h(t_n) - y^h(t')\|_{0,\infty,\Omega} dt' \\ & \quad + \|y^{h,n}(x) - Y_{h,k}^n(x)\|_{0,\infty,\Omega} + |(Y_{h,k}^n - \phi)^+(x)| \\ & \leq Cs^2 |\log h|^2 (k^{1-\frac{2}{s}} + Ch^{2-\frac{4}{s}}) (\|Bu_{h,k}\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},s,\Omega}) \\ & \quad + |(Y_{h,k}^n - \phi)^+(x)|, \end{aligned} \tag{4.19}$$

where we have used an interpolation estimate and Theorem 3. From $Y_{h,k}^n(x_j) \leq \phi$ ($j = 1, \dots, m$) we conclude $(Y_{h,k}^n - \phi)^+(x) = 0$ on Ω^h . Theorem 3 implies that $Y_{h,k}^n$ is uniformly bounded in Ω , so we conclude from $|\Omega \setminus \Omega^h| \leq Ch^2$ that

$$\begin{aligned} \int_{\tilde{\Omega}_T} (Y_{h,k}^n - \phi)^+ d\mu & \leq C |\Omega \setminus \Omega^h| \max_{x \in \Omega \setminus \Omega^h} |(Y_{h,k}^n - \phi)^+(x)| \\ & \leq Ch^2 (\|Y_{h,k}^n\|_{0,\infty,\Omega} + \|\phi\|_{0,\infty,\Omega}) \\ & \leq Ch^2. \end{aligned} \tag{4.20}$$

Similarly, from (4.8), (4.9), (4.10), Theorem 3 and the fact that $\mu_{h,k} \geq 0, y \leq \phi$ and $\int_{\tilde{\Omega}_T} (\phi - Y_{h,k}) d\mu_{h,k} = 0$ we have

$$\begin{aligned} & \int_{\tilde{\Omega}_T} (Y_{h,k}(u) - Y_{h,k}) d\mu_{h,k} \\ & = \int_{\tilde{\Omega}_T} (Y_{h,k}(u) - y + y - \phi + \phi - Y_{h,k}) d\mu_{h,k} \\ & \leq \int_{\tilde{\Omega}_T} (Y_{h,k}(u) - y) d\mu_{h,k} \\ & \leq C \max_{1 \leq n \leq N} \|Y_{h,k}^n(u) - y^n\|_{0,\infty,\Omega} \left(\sum_{i=1}^N \sum_{j=1}^m \mu_j^i \right) \\ & \leq Cs^2 |\log h|^2 (k^{1-\frac{2}{s}} + Ch^{2-\frac{4}{s}}) (\|Bu\|_{0,\infty,\Omega} + \|y_0\|_{2-\frac{2}{s},s,\Omega}). \end{aligned} \tag{4.21}$$

We note that $\|Bu_{h,k}\|_{0,\infty,\Omega} \leq C|\Omega| \cdot \max\{|a|, |b|\}$. Combining (4.18), (4.19)–(4.21) the claim is proved. \square

In [9], the authors obtain the convergence order of $O(|\log h|^{\frac{1}{4}}(h^{\frac{1}{2}} + k^{\frac{1}{4}}))$ in 2d and $O(h^{\frac{1}{4}} + h^{-\frac{1}{4}}k^{\frac{1}{4}})$ in 3d for problems with state constraints pointwise in space and time where controls only act in time as in (1.5). In [22], the authors obtain the convergence order of $O(\log(\frac{T}{k})^{\frac{1}{2}}(k^{\frac{1}{2}} + h))$ for problems with state constraints acting only pointwise in time and distributed control. In the present paper for some fixed $s \in (4, \infty)$ which depends on the regularity of $y_0 \in W^{2-\frac{2}{s},s}(\Omega)$ and the regularity of the domain Ω , we obtain the convergence order $s|\log h|(k^{\frac{1}{2}-\frac{1}{s}} + h^{1-\frac{2}{s}})$ in two space dimensions. If the assumption of Theorem 3 are satisfied for all $4 < s < \infty$ we can get the order $h + k^{\frac{1}{2}}$ up to a logarithmic factor by setting, e.g., $s = |\log h|$. This compares to the results of [22] in the 2d case. Thus, if the regularity of the boundary of domain Ω allows for $y \in W_s^{2,1}(\Omega_T)$ for all $4 < s < \infty$ we in this case obtain an improvement over the results of [9], which appears to be quasi-optimal for problems with state constraints pointwisely in space and time with distributed control when compared with the elliptic case, see e.g., [8, 9, 23].

5 Numerical examples

In this section we will carry out some numerical experiments to support our theoretical findings. We consider the following parabolic optimal control problem:

$$\min J(y, u) = \frac{1}{2}\|y - y_d\|_{L^2(\Omega_T)}^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega_T)}^2$$

subject to

$$\begin{cases} y_t - \Delta y + y = f + u & \text{in } \Omega_T, \\ \frac{\partial y}{\partial n} = 0 & \text{on } \Gamma_T, \\ y(0) = y_0 & \text{in } \Omega \end{cases}$$

with box type control constraints

$$u \in K_U := \{u \in L^\infty(\Omega_T) : a \leq u(x, t) \leq b, (x, t) \in \Omega_T\}$$

and state constraints

$$y \in K_Y := \{y \in L^\infty(\Omega_T) : y(x, t) \leq \phi(x, t) \text{ or } y(x, t) \geq \phi(x, t), (x, t) \in \Omega_T\}.$$

For constructing an example with exact solution we allow additional data f .

The numerical solution of the two examples is performed with the method proposed in [13, 20], which goes back to an idea of Pierre and Sokolowsky [27].

Table 1 Error of control u and state y for Example 1

h	N	$\ u - u_{h,k}\ _{L^2(\Omega_T)}$	Order	$\ y - Y_{h,k}\ _{L^2(\Omega_T)}$	Order
$\sqrt{2}/4$	8	0.032272253305	\	0.056604772253	\
$\sqrt{2}/6$	18	0.015478395776	1.8121	0.027718367216	1.7609
$\sqrt{2}/8$	32	0.009257272003	1.7868	0.016383380301	1.8278
$\sqrt{2}/10$	50	0.006339627571	1.6966	0.010770666548	1.8797
$\sqrt{2}/12$	72	0.004790681261	1.5366	0.007613893931	1.9024
$\sqrt{2}/14$	98	0.003886436096	1.3570	0.005657242145	1.9269
$\sqrt{2}/16$	128	0.003332771780	1.1510	0.004370258689	1.9330
$\sqrt{2}/18$	162	0.002977342034	0.9575	0.003479452205	1.9353

Example 1 Let $\Omega_T = [0, 1]^2 \times [0, 1]$, $\alpha = 1$. Following the ideas of [9] we set

$$y(x, t) = \cos(\pi x_1) \cos(\pi x_2) \sin(\pi t), \quad \phi(x, t) = \max(0.7, y(x, t)),$$

while control and adjoint state are given by

$$u(x, t) = t(t - 1) \cos(\pi x_1) \cos(\pi x_2),$$

$$p(x, t) = \cos(\pi x_1) \cos(\pi x_2)t(1 - t).$$

We omit control constraints since $u \in L^\infty(\Omega_T)$. We note that $y \in W_s^{2,1}(\Omega_T)$ for all $1 \leq s < \infty$ although the domain is a polygon. In this example we have a regular multiplier associated to the state constraint, namely

$$\mu(x, t) = \max(y - 0.7, 0).$$

A simple calculation shows

$$y_d(x, t) = y(x, t) + \mu(x, t) + (1 - 2t) \cos(\pi x_1) \cos(\pi x_2)$$

$$+ t(1 - t)(-2\pi^2 \cos(\pi x_1) \cos(\pi x_2) - \cos(\pi x_1) \cos(\pi x_2)),$$

and

$$f(x, t) = \pi \cos(\pi t) \cos(\pi x_1) \cos(\pi x_2) + 2\pi^2 \cos(\pi x_1) \cos(\pi x_2) \sin(\pi t)$$

$$+ y(t, p) - u(x, t).$$

To support our theoretical results we test the convergence order with respect to space and time discretization. We choose the time step $k = O(h^2)$ where h denotes the mesh size of space triangulation. The results are listed in Table 1.

It is observed that in this numerical example the convergence orders for the optimal control u and the state y are better than the expected, which may be caused by the fact that the multiplier associated to the state constraints is continuous. However, in this context it should be mentioned that it is often difficult in control of parabolic PDEs to construct numerical examples with exact known solutions which deliver the exact predicted order of convergence, see e.g., [22, 24].

Table 2 Error of control u and state y for Example 2

DOF	N	$\ u - u_{h,k}\ _{L^2(\Omega_T)}$	Order	$\ y - Y_{h,k}\ _{L^2(\Omega_T)}$	Order
289	4	0.128013975900	\	0.020778308176	\
1089	8	0.059342708399	1.1092	0.009776691642	1.0877
4225	16	0.028680813384	1.0490	0.004449004173	1.1359
16641	32	0.016028859291	0.8394	0.002159618979	1.0427

Example 2 This example is chosen from [16]. We set $\Omega_T = B_1(0) \times [0, 1]$,

$$y(x, t) = \left(\|x\|^2 - \left(\frac{t}{2}\right)^2 \right)^2 \exp\left(\frac{8\|x\|^2}{t^2 - 4}\right) + \beta(t), \quad \beta(t) = \begin{cases} 1 - 4t & 0 \leq t \leq \frac{1}{4}, \\ 0 & \frac{1}{4} < t \leq \frac{3}{4}, \\ 4t - 3 & \frac{3}{4} < t \leq 1. \end{cases}$$

Then again $y \in W_s^{2,1}(\Omega_T)$ for all $1 \leq s < \infty$. We impose the state constraint $y(x, t) \geq 0$ such that the active set $\mathcal{A} = \{(x, t) \in \Omega_T : y = 0\}$ has the form

$$\mathcal{A} = \left\{ (x, t) \in \Omega_T : \frac{1}{4} \leq t \leq \frac{3}{4}, \|x(t)\| = \frac{t}{2} \right\}.$$

We define the adjoint state p as

$$p(x, t) = \begin{cases} \left((t - \frac{1}{2})^2 - \frac{1}{16} \right) (2 - \frac{t}{2} - \|x\|) (\|x\| - \frac{t}{2}) & \|x\| > \frac{t}{2}, \frac{1}{4} < t < \frac{3}{4}, \\ 0 & \text{otherwise.} \end{cases}$$

The Lagrange multiplier μ associated with the state constraints satisfies

$$\begin{aligned} \mu &\leq 0, \quad \int_{\bar{\Omega}_T} y d\mu = 0, \\ \langle \mu, w \rangle_{\mathcal{M}, \mathcal{C}} &= \int_{\frac{1}{4}}^{\frac{3}{4}} \int_{\partial B_{\frac{t}{2}}(0)} -2 \left(1 - \frac{t}{2} \right) \left(\left(t - \frac{1}{2} \right)^2 - \frac{1}{16} \right) w ds \end{aligned}$$

for any $w \in C(\bar{\Omega}_T)$.

To accurately resolve the quadrature for this example we need to construct special meshes. This is done by congruent refinement of the initial grid formed by 8 sectors of the unit circle. For each $i = 1, 2, \dots, N$ with $t_i \in [\frac{1}{4}, \frac{3}{4}]$ we need to ensure that the circle with radius $\|x\| = \frac{t_i}{2}$ is well triangulated. Here t_i ($i = 1, 2, \dots, N$) denote the time grid points. To do so, the time stepping has to be coupled with the grid size such that $h = O(k)$. Figure 1 depicts two such meshes with 289 (left) and 1089 (right) nodes.

We observe first order convergence for the optimal control and state from Table 2. Since the time step k is coupled with the mesh size h like $h = O(k)$, the numerical results mainly show the convergence order related to the time discretization. This

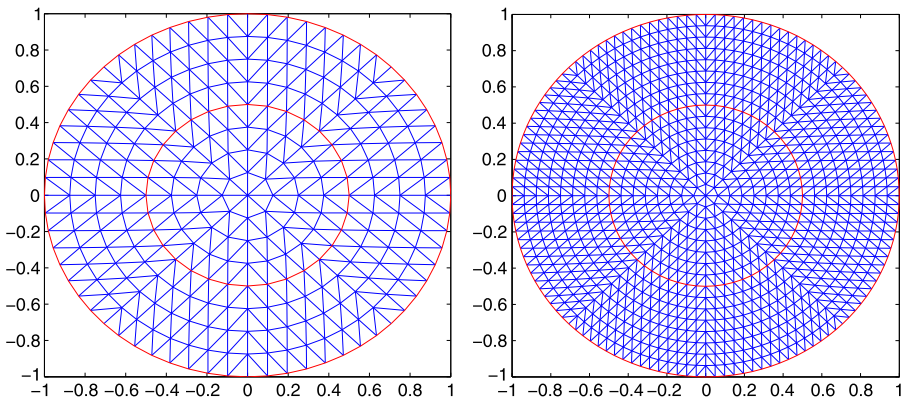


Fig. 1 The congruent refined meshes for Example 2 with nodes 289 (left) and 1089 (right)

order is better than $O(k^{\frac{1}{2}})$, which we would expect from our estimates. Although this numerical example does not exactly meet the predictions of our theory if certainly shows, when compared to Example 1, that less regularity of the exact solution results in a decrease of the numerically observed convergence order.

Acknowledgements The authors would like to thank Christian Kahle for the help on numerical experiments and Klaus Deckelnick for the discussion on Example 2. The first author also gratefully acknowledges the support of the Alexander von Humboldt Foundation during his stay at the University of Hamburg, Germany. He is also very grateful to the Department of Mathematics of the University of Hamburg for the hospitality and support. The first author was partially supported by the National Basic Research Program of China under the Grant 2012CB821204. The second author also gratefully acknowledges support of the priority program SPP1253 of the German research foundation.

References

1. Bonnans, J.F., Jaisson, P.: Optimal control of a parabolic equation with time-dependent state constraints. *SIAM J. Control Optim.* **48**, 4550–4571 (2010)
2. Brenner, S.C., Scott, L.R.: *The Mathematical Theory of Finite Element Methods*, 3rd edn. Springer, Berlin (2007)
3. Casas, E.: Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.* **24**, 1309–1318 (1986)
4. Casas, E.: Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control Optim.* **31**, 993–1006 (1993)
5. Casas, E.: Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.* **35**, 1297–1327 (1997)
6. Ciarlet, P.G.: *The Finite Element Methods for Elliptic Problems*. North-Holland, Amsterdam (1978)
7. Deckelnick, K., Günther, A., Hinze, M.: Finite element approximation of elliptic control problems with constraints on the gradient. *Numer. Math.* **111**, 335–350 (2009)
8. Deckelnick, K., Hinze, M.: Convergence of a finite element approximation to a state constrained elliptic control problem. *SIAM J. Numer. Anal.* **45**, 1937–1953 (2007)
9. Deckelnick, K., Hinze, M.: Variational discretization of parabolic control problems in the presence of pointwise state constraints. *J. Comput. Math.* **29**, 1–16 (2011)
10. de los Reyes, J.C., Merino, J.C., Rehberg, P., Tröltzsch, F.: Optimality conditions for state constrained PDE control problems with time-dependent controls. *Control and Cybernetics* **37**, 5–38 (2008)
11. Eriksson, K., Johnson, C.: Adaptive finite element methods for parabolic problems II: optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.* **32**, 706–740 (1995)

12. French, D.A., King, J.T.: Analysis of a robust finite element approximation for a parabolic equation with rough boundary data. *Math. Comput.* **60**, 79–104 (1993)
13. Gong, W., Yan, N.N.: A mixed finite element scheme for optimal control problems with pointwise state constraints. *J. Sci. Comput.* **46**, 182–203 (2011)
14. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30**, 45–63 (2005)
15. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: *Optimization with PDE Constraints MMTA 23*. Springer, Berlin (2009)
16. Kahle, C.: *Relaxierungsverfahren zur numerischen Lösung von parabolischen Kontrollproblemen mit Zustandsschranken*. Diploma Thesis, Department Mathematik, Universität Hamburg (2009)
17. Ladyženskaja, O.A., Solonnikov, V., Ural'ceva, N.: *Linear and Quasilinear Equations of Parabolic Type*. Translation Mathematical Monographs, vol. 23. AMS, Providence (1968)
18. Lions, J.L.: *Optimal Control of Systems Governed by Partial Differential Equations*. Springer, Berlin (1971)
19. Lions, J.L., Magenes, E.: *Non-Homogeneous Boundary Value Problems and Applications*. Springer, Berlin (1972)
20. Liu, W.B., Gong, W., Yan, N.N.: A new finite element approximation of a state-constrained optimal control problem. *J. Comput. Math.* **27**, 97–114 (2009)
21. Liu, W.B., Yan, N.N.: A posteriori error estimates for optimal control problems governed by parabolic equations. *Numer. Math.* **93**, 497–521 (2003)
22. Meidner, D., Rannacher, R., Vexler, B.: A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time. *SIAM J. Control Optim.* **49**, 1961–1997 (2011)
23. Meyer, C.: Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control Cybern.* **37**, 51–85 (2008)
24. Neitzel, I., Tröltzsch, F.: On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints. *ESAIM Control Optim. Calc. Var.* **15**, 426–453 (2009)
25. Nitsche, J.A.: L^∞ -Convergence of finite element approximations. In: *Mathematical Aspects of Finite Element Methods*, Lecture Notes in Mathematics, vol. 606, pp. 261–274. Springer, Berlin (1977)
26. Nochetto, R.H., Verdi, C.: Convergence past singularities for a fully discrete approximation of curvature-driven interfaces. *SIAM J. Numer. Anal.* **34**, 490–512 (1997)
27. Pierre, M., Sokolowsky, J.: *Differentiability of Projection and Applications*. Control of Partial Differential Equations and Applications (Laredo, 1994). Lecture Notes in Pure and Appl. Math., vol. 174, pp. 231–240. Dekker, New York (1996)
28. Rannacher, R., Scott, R.: Some optimal error estimates for piecewise linear finite element approximations. *Math. Comput.* **38**, 437–445 (1982)
29. Wang, G.S., Yu, X.: Error estimates for an optimal control problem governed by the heat equation with state and control constraints. *Int. J. Numer. Anal. Model.* **7**, 30–65 (2010)
30. Schatz, A.H.: Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids. I: Global estimates. *Math. Comput.* **67**, 877–899 (1998)