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

Wei Gong
LSEC, Institute of Computational Mathematics,
Academy of Mathematics and Systems Science,
Chinese Academy of Sciences, Beijing 100190, China.
Email: wgong@lsec.cc.ac.cn
Michael Hinze
Bereich Optimierung und Approximation,
Universität Hamburg, Bundestrasse 55,
20146 Hamburg, Germany.
Email: michael.hinze@uni-hamburg.de

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 backward Euler method is used for time discretization. A priori error estimates for control and state are obtained by an improved maximum error estimate for corresponding discretized state equation. Numerical experiments are provided which confirm our theoretical results.

Key words. optimal control problem, finite element method, a priori error estimate, parabolic equation, control and state constraints.
Subject Classification: 49J20, 49K20, 65N30.

1 Introduction

In this paper we consider the optimal control problem

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

subject to

$$\begin{cases} y_t - \Delta y + y = Bu & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_T, \\ y(0) = y_0 & \text{in } \Omega, \end{cases} \tag{1.2}$$

where $\Omega_T = \Omega \times (0,T]$, $\Gamma_T = \partial\Omega \times (0,T]$, Ω is an open bounded domain in $\mathcal{R}^2$ with boundary $\Gamma = \partial\Omega$, $\alpha > 0$, $T > 0$ and $y_d \in L^2(\Omega_T)$ are fixed, and the initial value y_0 is specified in Section 2. Here $B : U \to L^2(0,T;H^1(\Omega)^*)$ denotes the linear bounded control operator, where U is a Hilbert space. Typical choices include the case $U = L^\infty(\Omega_T) \subset L^2(\Omega_T)$ with $Bu := u$ the injection(see [19] and [20]) or $U = L^\infty(\omega) \subset L^2(\omega)$ with $\omega \subseteq \Omega_T$ and $Bu = \chi_\omega u$, where χ_ω is the characteristic function of the subset ω(see [24]), or defined as $U = L^\infty(0,T)^m \subset L^2(0,T)^m$ with

$$(Bu)(x,t) := \sum_{i=1}^{m} u_i(t) f_i(x), \quad (x,t) \in \Omega_T, \tag{1.3}$$

where $f_1, \cdots, f_m \in H^1(\Omega) \cap L^\infty(\Omega)$ are given functions(see [8]). For our analysis we require $Bu \in L^\infty(\Omega_T)$, so that from here onwards we choose $U = L^2(\Omega_T)$ with B the injection and enforce $Bu \in L^\infty(\Omega_T)$ through box constraints, i.e., we require

$$u \in K_U := \{u \in L^2(\Omega_T) : \quad a \leqslant u(x,t) \leqslant b, \quad \text{for } a.a.\ (x,t) \in \Omega_T\} \tag{1.4}$$

with $a < b$ constants. Furthermore, we also consider state constraints

$$y \in K_Y := \{y \in L^\infty(\Omega_T) : \quad y(x,t) \leqslant \phi, \quad \text{for } a.a.\ (x,t) \in \Omega_T\}. \tag{1.5}$$

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 parabolic control problems with state constraints. Pontryagin's principles for several class of control problems were derived in, e.g., [1], [4] and [9]. Lavrentiev regularization of state constrained parabolic control problems was studied in [20]. Recently, error estimates for state constrained parabolic control problem with controls of type (1.3) were derived in [8]. The error analysis for problem with final state constraints and control constraints was also studied in [24]. In [19] a priori error estimates for problems with pointwise state constraints only in time are considered.

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 [13] combined with linear finite elements in space and the backward Euler scheme in time for the discretization of the state equation. Based on an improved maximum error estimate for the state equation, we derive L^2-norm error estimates for both the control and the state, which seems to be quasi-optimal for optimal control.

The rest of this paper is organized as follows. In Section 2 we present the state constrained optimal control problem and the corresponding optimality conditions. In Section 3 we establish the fully discrete approximation for the state equation and derive uniform estimates for the discretization error of the state in Section 4. We obtain the a priori error estimates for the optimal control problem in Section 5. We also present some numerical experiments to confirm our theoretical findings.

2 Optimal control problem

We denote by $H^m(\Omega)$ the usual Sobolev space of integer order $m \geqslant 0$ with norm $\|\cdot\|_m$. Note that $H^0(\Omega) = L^2(\Omega)$. Similarly, $H^r(\Gamma)$ denotes the Sovolev space of integer order

$r \geqslant 0$ on Γ with norm $|\cdot|_r$. We denote the L^2-inner products on $L^2(\Omega)$ and $L^2(\Gamma)$ by

$$(v,w) = \int_\Omega vwdx, \quad \forall\ v,w \in L^2(\Omega)$$

and

$$<v,w>= \int_\Gamma vwds, \quad \forall\ v,w \in L^2(\Gamma).$$

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

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

where $\|\cdot\|_{r,[0,T]}$ denotes the norm on $H^r([0,T])$. Similarly, $H^{s,r}(\Gamma_T) = L^2(0,T;H^s(\Gamma)) \cap H^r(0,T;L^2(\Gamma))$, and the norm on $H^{s,r}(\Gamma)$ will be denoted by $|\cdot|_{s,r}$. In addition c and C denote generic positive constants.

The weak form of problem (1.2) reads: Find $y \in L^2(0,T;H^1(\Omega)) \cap C(0,T;L^2(\Omega))$ such that

$$\begin{cases} (y_t,v) + (\nabla y, \nabla v) + (y,v) = (Bu,v), \ \ t \in (0,T], \quad \forall v \in H^1(\Omega), \\ \qquad\qquad\qquad\qquad y(x,0) = y_0(x), \ \ x \in \Omega. \end{cases} \tag{2.1}$$

We assume that the initial value $y_0 \in H^2(\Omega)$ throughout the paper. We denote $y = G(Bu)$ the solution to problem (2.1). It is well-known that if $Bu \in L^2(\Omega_T)$, $y_0 \in H^1(\Omega)$, problem (1.2) admits a unique solution $y = G(Bu) \in W := L^2(0,T;H^2(\Omega)) \cap H^1(0,T;L^2(\Omega)) \hookrightarrow C(0,T;H^1(\Omega))$. Thanks to the control constraints given by (1.4), we have $Bu \in L^\infty(\Omega_T)$, thus we have the improved regularity $y \in W_s^{2,1}(\Omega_T)$ for all $s < \infty$, where $W_s^{2,1}(\Omega_T)$ is defined as

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

Our optimal control problem reads:

$$\begin{cases} \min \quad J(y,u) = \frac{1}{2}\|y - y_d\|_{L^2(\Omega_T)}^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega_T)}^2 \\ s.t.\ y = G(Bu), \text{ and } y(x,t) \in K_Y, \quad u(x,t) \in K_U. \end{cases} \tag{2.2}$$

Note that $W_s^{2,1}(\Omega_T) \hookrightarrow C(\bar\Omega_T)$, so the state constrained optimal control problem (2.2) is well defined. Since J is quadratic and K_U and K_Y are closed and convex, problem (2.2) admits a unique soluton $(y,u) \in W_s^{2,1}(\Omega_T) \times K_U$. To ensure existence of Lagrange multipliers we assume the Slater condition:

Assumption 2.1. *(Slater condition): We assume that $y_0 < \phi$ and there exists $\hat u \in L^\infty(\Omega_T)$ satisfying (1.4) such that the associated state $\hat y$ fulfills (1.5) strictly, which means $\hat y(x,t) < \phi$ holds for all $(x,t) \in \bar\Omega_T$.*

It then follows from e.g., [4], [9] and [20] that the first order optimality conditions for optimal problem (2.2) are given by

Theorem 2.2. *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 $\mathcal{M}(\bar\Omega_T)$ be 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 $\frac{2}{q}+\frac{d}{\sigma} > d+1$, d is the dimension of Ω, and a Lagrange multiplier $\mu \in \mathcal{M}(\bar\Omega_T)$ such that*

$$\begin{cases} -p_t - \Delta p + p = y - y_d + \mu_{\Omega_T} & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = \mu_{\Sigma_T} & \text{on } \Sigma_T, \\ p(T) = \mu_T & \text{in } \Omega \end{cases} \tag{2.3}$$

is satisfied in the sense of distributions, and

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

$$\mu \geqslant 0, \quad y(x,t) \leqslant \phi, \quad (x,t) \in \bar\Omega_T, \quad \text{and} \quad \int_{\Omega_T} (\phi - y) d\mu = 0. \tag{2.5}$$

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

It is worth noting that the weak form of (2.3) has to be understood in the sense that

$$\int_0^T \int_\Omega (w_t - \Delta w + w)p + \int_0^T \int_\Gamma \frac{\partial w}{\partial n} p = \int_0^T \int_\Omega (y - y_d) w + \int_{\bar\Omega_T} w d\mu, \quad \forall w \in W_0^\infty, \tag{2.6}$$

where

$$W_0^\infty := \{w \in W \cap C(\bar\Omega_T): \ w(\cdot,0) = 0 \text{ in } \bar\Omega, \ w_t - \Delta w + w \in L^\infty(\Omega_T), \ \frac{\partial w}{\partial n} \in L^\infty(\Gamma_T)\} \tag{2.7}$$

3 Finite element discretization of the state equation

Let Ω^h be a polygonal approximation to Ω with a boundary $\Gamma_h = \partial\Omega^h$. Let $\mathcal{T}^h$ be a partitioning of Ω^h into disjoint regular n-simplices τ, so that $\bar\Omega^h = \bigcup_{\tau\in\mathcal{T}^h} \bar\tau$. For simplicity we assume that Ω is a polygonal convex domain such that $\Omega^h = \Omega$. Associated with $\mathcal{T}^h$ is a finite dimensional subspace V^h of $C(\bar\Omega^h)$, such that $\chi|_\tau$ are polynomials of order 1 for $\forall \chi \in V^h$ and $\tau \in \mathcal{T}^h$. It is easy to see that $V^h \subset V = H^1(\Omega)$.

Then the semi-discrete finite element approximation of (2.1) reads:

$$\begin{cases} (\dfrac{\partial y_h}{\partial t}, w_h) + (\nabla y_h, \nabla w_h) + (y_h, w_h) = (Bu, w_h), \ \forall w_h \in V^h, \ t \in (0,T], \\ y_h(x,0) = y_0^h(x), \quad x \in \Omega, \end{cases} \tag{3.1}$$

where $y_h(t) \in H^1(0,T;V^h)$, and $y_0^h \in V^h$ is an approximation to y_0.

We next consider the fully discrete approximation for above semidiscrete problem by using the backward Euler scheme in time.

Let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$, $t_i = i\frac{T}{N}$, $i = 1,2,\cdots,N$ and $k = \frac{T}{N}$. Let h denote the mesh size for triangulation T^h.

The fully discrete approximation scheme for (3.1) now reads: Find $Y_h^i \in V^h$, $i = 1, 2, \cdots, N$, such that

$$\begin{cases} (\dfrac{Y_h^i - Y_h^{i-1}}{k}, w_h) + (\nabla Y_h^i, \nabla w_h) + (Y_h^i, w_h) = (B\bar{u}, w_h), \ \forall w_h \in V^h, \ \ i = 1, \cdots, N, \\ Y_h^0(x) = y_0^h(x), \quad x \in \Omega, \end{cases} \tag{3.2}$$

where $B\bar{u} = \frac{1}{k}\int_{t_{i-1}}^{t_i} Bu$.

For $i = 1, 2, \cdots, N$, let

$$Y_h|_{(t_{i-1}, t_i]} = ((t - t_{i-1})Y_h^i + (t_i - t)Y_h^{i-1})/k,$$

and for $w \in C(0, T; L^2(\Omega))$, let $\hat{w}(x,t)|_{(t_{i-1},t_i]} = w(x, t_i)$, $\check{w}(x,t)|_{(t_{i-1},t_i]} = w(x, t_{i-1})$. Then (3.2) can be formulated as: Find $Y_h := G_{h,k}(Bu) \in V^h$ such that

$$\begin{cases} (\dfrac{\partial Y_h}{\partial t}, w_h) + (\nabla \hat{Y}_h, \nabla w_h) + (\hat{Y}_h, w_h) = (B\bar{u}, w_h), \ \forall w_h \in V^h, \\ y_h(x, 0) = y_0^h(x), \quad x \in \Omega, \end{cases} \tag{3.3}$$

where $\frac{\partial Y_h}{\partial t}|_{(t_{i-1},t_i]} := \frac{1}{k}(Y_h^i - Y_h^{i-1})$.

4 Error estimates for the state equation

Let $\Pi_h : C(\bar{\Omega}) \to V^h$ denote the standard Lagrange interpolation operator. Interpolation error estimates imply that for $y \in W^{m,r}(\Omega)$, $r > 2$(see, e.g., [5])

$$\|y - \Pi_h y\|_{0,r} + h\|y - \Pi_h y\|_{1,r} \leqslant Ch^m \|y\|_{m,r}, \quad 1 \leqslant m \leqslant 2. \tag{4.1}$$

We choose $y_0^h = \Pi_h y_0$ in (3.3). Let R_h denote the Ritz projection operator defined as

$$(\nabla R_h y, \nabla v_h) + (R_h y, v_h) = (\nabla y, \nabla v_h) + (y, v_h), \quad \forall v_h \in V^h. \tag{4.2}$$

Lemma 4.1. *Let R_h be the Ritz projection operator defined above. Then there holds:*

$$\|y - R_h y\|_{0,r} \leqslant Crh \inf_{v_h \in V^h} \|y - v_h\|_{1,r}, \quad r \geqslant 2. \tag{4.3}$$

$$\|y - R_h y\|_{0,\infty} \leqslant C|\log h| \inf_{v_h \in V^h} \|y - v_h\|_{0,\infty}. \tag{4.4}$$

Proof. A result related to (4.3) is proved by Rannacher and Scott in [23] for Dirichlet boundary conditions, but the arguments can be adapted to the present situation, we omit the details here. The result of (4.4) can be found in [25]. □

Now we are in a position to estimate the error between the solutions of problem (2.1) and (3.3). The following result is a standard consequence of error estimate for parabolic equation(see, e.g., [11]).

Theorem 4.2. *Let $Bu \in L^2(\Omega_T)$, $y \in H^{2,1}(\Omega_T)$ be the solution of problem (2.1), and $Y_h \in V^h$ be the solution of problem (3.3), then we have*

$$\|y - Y_h\|_{L^2(0,T;L^2(\Omega))} \leqslant C(h^2 + k)\|y\|_{2,1}. \tag{4.5}$$

To estimate the error of optimal control problem we need the maximum norm estimate for state equation. Following the idea of [21] we need to introduce the weighted-norm technique. Let

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

where $z \in \Omega$ and $\omega = ch|\log h|$, then it follows from [21](see, e.g, Lemma 4.3) that

$$\int_\Omega \rho(x)^{-m} dx \leqslant C((m-2)\omega^{m-2})^{-1} \quad \text{for } m > 2. \tag{4.6}$$

Theorem 4.3. *Let $Bu \in L^\infty(\Omega_T)$, $y \in W_s^{2,1}(\Omega_T)$ be the solution of problem (2.1), and $Y_h \in V^h$ be the solution of problem (3.3). Then*

$$\max_{1\leqslant n\leqslant N} \|y(\cdot,t_n) - Y_h^n\|_{\infty,\Omega} \leqslant Cs^2|\log h|^2(h^{2-4/s}+k^{1-2/s})\|y\|_{2,1,s}. \tag{4.7}$$

Proof. The proof follows [21]. We present a sketch for the convenience of the reader.

Note that

$$y(\cdot,t) - Y_h(\cdot,t) = \frac{t-t_{n-1}}{k}(y(\cdot,t)-Y_h^n) + \frac{t_n-t}{k}(y(\cdot,t)-Y_h^{n-1}) \tag{4.8}$$

holds for all $t_{n-1} \leqslant t \leqslant t_n$, $1 \leqslant n \leqslant N$. For $y(\cdot,t) - Y_h^n$ we have the splitting

$$\begin{aligned} y(\cdot,t) - Y_h^n &= y(\cdot,t) - R_h\bar{y}^n + R_h\bar{y}^n - Y_h^n \\ &= \xi^n - \eta^n, \end{aligned}$$

where

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

Then

$$\|y(\cdot,t) - Y_h^n\|_{\infty,\Omega} \leqslant \|y(\cdot,t) - R_h\bar{y}^n\|_{\infty,\Omega} + \|R_h\bar{y}^n - Y_h^n\|_{\infty,\Omega}. \tag{4.9}$$

Note that(see [21])

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

with $m = 2+d = 4 < s < \infty$ as s approaches ∞. Then from (4.1) and (4.4) we deduce

$$\begin{aligned} \|y(\cdot,t) - R_h\bar{y}^n\|_{\infty,\Omega} &\leqslant \|y(\cdot,t) - \bar{y}^n\|_{\infty,\Omega} + \|\bar{y}^n - R_h\bar{y}^n\|_{\infty,\Omega} \\ &\leqslant \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \|y(t)-y(t')\|_{\infty,\Omega}dt' + C|\log h|\|\bar{y}^n - \Pi_h\bar{y}^n\|_{0,\infty,\Omega} \\ &\leqslant C(k^{1-\frac{2}{s}} + Ch^{2-\frac{4}{s}}|\log h|)\|y\|_{W_s^{2,1}(\Omega_T)}. \end{aligned} \tag{4.10}$$

It remains to estimate $\|R_h\bar{y}^n - Y_h^n\|_{\infty,\Omega}$. Suppose that $z \in \Omega$ is the point where $\|R_h\bar{y}^n - Y_h^n\|_{\infty,\Omega}$ achieves maximal value. Inverse estimates give

$$\|R_h\bar{y}^n - Y_h^n\|_{\infty,\Omega} \leqslant Ch^{-1}\|R_h\bar{y}^n - Y_h^n\|_{0,\Omega}$$

$$
\begin{aligned}
&\leqslant Ch^{-1}(\int_\Omega \rho^2)^{\frac{1}{2}}(\int_\Omega (\rho^{-1}\eta^n)^2)^{\frac{1}{2}} \\
&\leqslant C\frac{\omega}{h}\|\rho^{-1}\eta^n\|_{0,\Omega}. \qquad (4.11)
\end{aligned}
$$

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

$$
(y^i - y^{i-1}, v) + \int_{t_{i-1}}^{t_i} (\nabla y, \nabla w) + \int_{t_{i-1}}^{t_i} (y, v) = \int_{t_{i-1}}^{t_i} (Bu, v),
$$

so that using (3.2) we have for all $v \in V^h$

$$
\begin{aligned}
& (\eta^i - \eta^{i-1}, v) + k(\nabla\eta^i, \nabla v) + k(\eta^i, v) \\
= \; & (Y_h^i - Y_h^{i-1}, v) + k(\nabla Y_h^i, \nabla v) + k(Y_h^i, v) \\
& -(R_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (R_h\nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (R_h y, v), \\
= \; & \int_{t_{i-1}}^{t_i} (Bu, v) - (R_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (R_h\nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (R_h y, v), \\
= \; & (y^i - y^{i-1}, v) + \int_{t_{i-1}}^{t_i} (\nabla y, \nabla v) + \int_{t_{i-1}}^{t_i} (y, v) \\
& -(R_h(\bar{y}^i - \bar{y}^{i-1}), v) - \int_{t_{i-1}}^{t_i} (R_h\nabla y, \nabla v) - \int_{t_{i-1}}^{t_i} (R_h y, v) \\
= \; & (\xi^i - \xi^{i-1}, v). \qquad (4.12)
\end{aligned}
$$

Let $Z^i \in V^h$, $i = 1, 2, ..., n$ be the solution of following backward fully discrete problem

$$
(Z^{i-1} - Z^i, w_h) + k(\nabla Z^{i-1}, \nabla w_h) + k(Z^{i-1}, w_h) = 0, \;\; \forall w_h \in V^h, \qquad (4.13)
$$

with $Z^n = \zeta$, $1 \leqslant n \leqslant N$. Now let $v = Z^{i-1}$ in (4.12), summing from 1 to n we find

$$
(\eta^n, \zeta) = (\xi^n, \zeta) + \sum_{i=1}^{n} (\xi^i, Z^{i-1} - Z^i) + (Y_h^0 - y_0, Z^0), \qquad (4.14)
$$

where we have used (4.13) and the fact that $\eta^i \in V^h$. Setting $\zeta = P_h(\rho^{-2}\eta^n)$ in (4.14), where $P_h: \; L^2(\Omega) \to V^h$ is the L^2-projection operator, we obtain

$$
\|\rho^{-1}\eta^n\|_{0,\Omega}^2 = (\xi^n, P_h(\rho^{-2}\eta^n)) + \sum_{i=1}^{n} (\xi^i, Z^{i-1} - Z^i) + (Y_h^0 - y_0, Z^0).
$$

Since $\|\rho^t P_h v\|_{0,\Omega} \leqslant C\|\rho^t v\|_{0,\Omega}$ holds for all $t \in R$ we have

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

Thus

$$
\|\rho^{-1}\eta^n\|_{0,\Omega}^2 \;\; \leqslant \;\; C\|\rho^{-1}\xi^n\|_{0,\Omega}^2 + C\sum_{i=1}^{n} \|\rho^{-1}\xi^i\|_{0,\Omega}\|\rho(Z^{i-1} - Z^i)\|_{0,\Omega}
$$

$$+C\|\rho^{-1}(\Pi_h y_0 - y_0)\|_{0,\Omega}\|\rho Z^0\|_{0,\Omega}. \tag{4.15}$$

Now we need a priori estimates for $Z^{i-1}-Z^i$ and Z^0 in weighted norms. It has been proved in [10] and [21], that if there exists $C^* \geqslant 1$ such that $w = C^* h|\log h|$ and $k \geqslant C^* h^2|\log h|^3$ then

$$\begin{aligned} &\|\rho Z^0\|_{0,\Omega} + (k^{-1}\sum_{i=1}^{n} t_{n-i+1}\|\rho(Z^{i-1}-Z^i)\|_{0,\Omega}^2)^{\frac{1}{2}} \\ \leqslant \quad & C|\log h|\|\rho\zeta\|_{0,\Omega} \leqslant C|\log h|\|\rho^{-1}\eta^n\|_{0,\Omega}. \end{aligned} \tag{4.16}$$

The following estimates can also be found in, e.g., [21]:

$$(k\sum_{i=1}^{n} t_{n-i+1}^{-1}\|\rho^{-1}\xi^i\|_{0,\Omega}^2)^{\frac{1}{2}} \leqslant Csh^{-\frac{2}{s}}k^{-\frac{1}{s}}(k\sum_{i=1}^{n}\|\xi^i\|_{0,s,\Omega}^s)^{\frac{1}{s}}, \tag{4.17}$$

$$\|\rho^{-1}\xi^n\|_{0,\Omega} \leqslant Csh^{-\frac{2}{s}}k^{-\frac{1}{s}}(k\sum_{i=1}^{n}\|\xi^i\|_{0,s,\Omega}^s)^{\frac{1}{s}}, \tag{4.18}$$

where Hölder's inequality and property (4.6) are used. Then the interpolation error estimate (4.1) and estimate (4.3) for the Ritz-projection lead to

$$\|\rho^{-1}(\Pi_h y_0 - y_0)\|_{0,\Omega} \leqslant Cs^{\frac{1}{2}}h^{2-\frac{4}{s}}\|y\|_{W_s^{2,1}(\Omega_T)} \tag{4.19}$$

as well as

$$(k\sum_{i=1}^{n}\|\xi^i\|_{0,s,\Omega}^s)^{\frac{1}{s}} \leqslant Cs(h^2+k)\|y\|_{W_s^{2,1}(\Omega_T)}. \tag{4.20}$$

From (4.15)-(4.20) we then have that

$$\|\rho^{-1}\eta^n\|_{0,\Omega} \quad \leqslant \quad Cs^2|\log h|(h^{2-\frac{4}{s}} + k^{1-\frac{1}{s}})\|y\|_{W_s^{2,1}(\Omega_T)}. \tag{4.21}$$

With (4.11) and (4.21) we conclude that

$$\|R_h\bar{y}^n - Y_h^n\|_{\infty,\Omega} \leqslant Cs^2|\log h|^2(h^{2-\frac{4}{s}} + k^{1-\frac{1}{s}})\|y\|_{W_s^{2,1}(\Omega_T)}. \tag{4.22}$$

Combining (4.10) and (4.22) we get the proof for $y(\cdot,t) - Y_h^n$, while an estimate for $y(\cdot,t) - Y_h^{n-1}$ can be derived similarly. This completes the proof of the theorem. □

Remark 4.4. *A uniform error estimate for the discretized error of equation (2.1) and (3.2) is derived in [8] 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.*

5 Error estimates for optimal control problem

In this section we consider the finite element approximation to optimal control problem (1.1)-(1.2).

We consider the variational discretization approach proposed in [8] and [13]. Then the fully discrete optimization problem reads

$$\min_{u\in K_U} J_h(u) = \sum_{i=1}^{N} \frac{1}{2}k\int_\Omega (Y_h^i - \bar{y}_d)^2 + \frac{\alpha}{2}\int_0^T\int_\Omega u^2, \tag{5.1}$$

subject to

$$\begin{cases} (\frac{Y_h^i - Y_h^{i-1}}{k}, w_h) + (\nabla Y_h^i, \nabla w_h) + (Y_h^i, w_h) = (B\bar{u}, w_h), \\ \forall w_h \in V^h, \ \ i = 1, \cdots, N, \ y_h^0(x) = \Pi_h y_0(x), \ x \in \Omega. \\ Y_h^i \leqslant \phi, \ \ i = 1, \cdots, N. \end{cases} \tag{5.2}$$

Let $\hat{u}$ denote the control satisfying Assumption 2.1, i.e., there exists $\delta > 0$ such that the corresponding state $\hat{y} = G(B\hat{u})$ satisfies

$$\hat{y} = G(B\hat{u}) \leqslant \phi - \delta \quad \text{ in } \bar{\Omega}_T. \tag{5.3}$$

It follows from Theorem 4.2 that there exists $h_0 > 0$ such that $\hat{Y}_h := G_{h,k}(B\hat{u}) \in V^h$ satisfies

$$\hat{Y}_h^n(x_j) \leqslant \phi - \frac{\delta}{2} < \phi, \quad 1 \leqslant n \leqslant N, \ 1 \leqslant j \leqslant J, \ 0 < h \leqslant h_0. \tag{5.4}$$

Thus the pair $(\hat{u}, \hat{Y}_h)$ is a discrete feasible Slater point for problem (5.1) as h approaches 0.

As a minimization problem for a quadratic functional over a closed convex set, the discrete optimization problem (5.1)-(5.2) admits a unique solution $U_h \in K_U$. Furthermore, it follows from [4] again that the discrete Slater condition (5.4) guarantees the existence of a discrete co-state $P_h^{i-1} \in V^h$, $i = 1, 2, \cdots, N$ and discrete Lagrange multiplier $\mu_j^i \in \mathcal{R}$, $i = 1, \cdots, N$, $j = 1, \cdots, J$, such that the triplet $(Y_h^i, P_h^{i-1}, U_h^i) \in V^h \times V^h \times K_U$, $i = 1, 2, \cdots, N$, satisfies the following optimality conditions:

$$\begin{cases} (\frac{Y_h^i - Y_h^{i-1}}{k}, w_h) + (\nabla Y_h^i, w_h) = (BU_h^i, w_h), \\ \forall w_h \in V^h, \ \ i = 1, \cdots, N, \ Y_h^0(x) = \Pi_h y_0(x), \ x \in \Omega, \end{cases} \tag{5.5}$$

$$\begin{cases} (\frac{P_h^{i-1} - P_h^i}{k}, q_h) + (\nabla q_h, \nabla P_h^{i-1}) + (q_h, P_h^{i-1}) = (Y_h^i - \bar{y}_d, q_h) + \sum_{j=1}^{J} \mu_j^i q_h(x_j), \\ \forall q_h \in V^h, \ \ i = N, \cdots, 1,; \ \ P_h^N = \sum_{j=1}^{J} \mu_j^N \delta_{x_j}, \ x \in \Omega, \end{cases} \tag{5.6}$$

$$\int_\Omega (aU_h^i + B^* P_h^{i-1})(v_h - U_h^i)dx \geqslant 0, \ \forall v_h \in K_U, \ i = 1, 2, \cdots, N. \tag{5.7}$$

$$\mu_j^i \geqslant 0;\ \ Y_h^i \leqslant \phi,\ \text{ and } \sum_{i=1}^{N}\sum_{j=1}^{J}(\phi - Y_h^i(x_j))\mu_j^i = 0. \tag{5.8}$$

Again the optimality conditions (5.5)-(5.7) can be formulated as

$$(\frac{\partial Y_h}{\partial t}, w_h) + (\nabla \hat{Y}_h, \nabla w_h) + (\hat{Y}_h, w_h) = (B\hat{U}_h, w_h),\ \ \forall w_h \in V^h,\ \ i = 1, \cdots, N, \tag{5.9}$$

$$Y_h(x, 0) = \Pi_h y_0(x),\ x \in \Omega,$$

$$-(\frac{\partial P_h}{\partial t}, q_h) + (\nabla q_h, \nabla \check{P}_h) + (q_h, \check{P}_h) = (\hat{Y}_h - \bar{y}_d, q_h) + \sum_{j=1}^{J}\mu_j^i q_h(x_j), \tag{5.10}$$

$$\forall q_h \in V^h,\ \ i = N, \cdots, 1,\quad P_h(x, T) = \sum_{j=1}^{J}\mu_j^N \delta_{x_j},\ x \in \Omega,$$

$$\int_\Omega (a\hat{U}_h + B^*\check{P}_h)(v_h - \hat{U}_h)dx \geqslant 0,\ \forall v_h \in K_U,\ i = 1, 2, \cdots, N. \tag{5.11}$$

It is easy to show that

$$U_h^i = P_{K_U}(\frac{1}{\alpha}B^* P_h^{i-1}),$$

where P_{K_U} is the projection operator over control set K_U. Here $\mu_h \in \mathcal{M}(\bar{\Omega}_T)$ is given by

$$\int_{\bar{\Omega}_T} f d\mu_h := \sum_{i=1}^{N}\sum_{j=1}^{J} f(x_j, t_i)\mu_j^i,\ \ \forall f \in C(\bar{\Omega}_T). \tag{5.12}$$

As a first result for (5.1)-(5.2) we prove that the sequence of optimal controls, states and measures μ_h are uniformly bounded.

Lemma 5.1. *Let $(Y_h, U_h) \in V^h \times K_U$ be the solutions of problem (5.1)-(5.2), $P_h \in V^h$ and $\mu_h \in \mathcal{M}(\bar{\Omega}_T)$ be the corresponding adjoint state and measure, respectively. Then there exists $h_0 > 0$ such that*

$$\|U_h\|^2_{L^2(0,T;L^2(\Omega))} + \|Y_h\|^2_{L^2(0,T;L^2(\Omega))} + \sum_{i=1}^{N}\sum_{j=1}^{J}\mu_j^i \leqslant C \quad \text{ for all } 0 < h \leqslant h_0.$$

Proof. The proof follows [8]. From (5.4) and (5.8) we obtain

$$\begin{aligned}
\frac{\delta}{2}\sum_{i=1}^{N}\sum_{j=1}^{J}\mu_j^i \ &\leqslant\ \sum_{i=1}^{N}\sum_{j=1}^{J}(\phi - \hat{Y}_h^i(x_j))\mu_j^i = \sum_{i=1}^{N}\sum_{j=1}^{J}(Y_h^i(x_j) - \hat{Y}_h^i(x_j))\mu_j^i \\
&=\ \sum_{i=1}^{N}(P_h^{i-1} - P_h^i, Y_h^i - \hat{Y}_h^i) + k(\nabla(Y_h^i - \hat{Y}_h^i), \nabla P_h^{i-1}) \\
&\quad + k(P_h^{i-1}, Y_h^i - \hat{Y}_h^i) - k(Y_h^i - \bar{y}_d, Y_h^i - \hat{Y}_h^i)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N}(P_h^{i-1}, Y_h^i - \hat{Y}_h^i - (Y_h^{i-1} - \hat{Y}_h^i)) + k(\nabla(Y_h^i - \hat{Y}_h^i), \nabla P_h^{i-1}) \\
&\quad + k(P_h^{i-1}, Y_h^i - \hat{Y}_h^i) - k(Y_h^i - \bar{y}_d, Y_h^i - \hat{Y}_h^i) - (P_h^N, Y_h^N - \hat{Y}_h^N) \\
&= \sum_{i=1}^{N} k(B(U_h^i - \bar{\hat{u}}), P_h^{i-1}) - k(Y_h^i - \bar{y}_d, Y_h^i - \hat{Y}_h^i) + \sum_{j=1}^{J}\mu_j^N(\hat{Y}_h^N - Y_h^N)(x_j) \\
&\leqslant \sum_{i=1}^{N} k(\alpha U_h^i, \bar{\hat{u}} - U_h^i) + k\int_\Omega(-(Y_h^i)^2 + Y_h^i\hat{Y}_h^i + Y_h^i\bar{y}_d - \hat{Y}_h^i\bar{y}_d) + 2\sum_{j=1}^{J}\mu_j^N\phi \\
&\leqslant -\frac{1}{2}\sum_{i=1}^{N}k\|Y_h^i\|_{0,\Omega}^2 - \frac{\alpha}{2}\sum_{i=1}^{N}k\|U_h^i\|_{0,\Omega}^2 + \frac{\delta}{4}\sum_{i=1}^{N}\sum_{j=1}^{J}\mu_j^i + C, \qquad (5.13)
\end{aligned}
$$

where we have used (5.5) and (5.6). This completes the proof of the Lemma. □

Now we are in a position to give the main result of this paper. We use a proof technique developed in Chap. 3 of [14] which only relies on uniform a priori error estimates of the state approximation.

Theorem 5.2. *Let $(y,u) \in W_s^{2,1}(\Omega_T) \times L^\infty(\Omega_T)$ and $(Y_h, U_h) \in V^h \times K_U$ be the solutions to problem (2.2) and (5.1)-(5.2). Then we have the following a priori error estimate*

$$
\|u - U_h\|_{L^2(0,T;L^2(\Omega))}^2 + \|y - Y_h\|_{L^2(0,T;L^2(\Omega))}^2 \leqslant Cs^2|\log h|^2(h^{2-\frac{4}{s}} + k^{1-\frac{1}{s}}). \qquad (5.14)
$$

Proof. From (2.4) and (5.11) we have

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

and

$$
\int_{\Omega_T}(a\hat{U}_h + B^*\check{P}_h)(v_h - \hat{U}_h)dx \geqslant 0, \ \forall v_h \in K_U.
$$

Adding these inequalities gives

$$
\alpha\|u - U_h\|_{L^2(0,T;L^2(\Omega))}^2 \ \leqslant \ \int_0^T (p - \check{P}_h, B(U_h - u)). \qquad (5.15)
$$

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

$$
\begin{cases}
y_t^h - \Delta y^h + y^h = BU_h & \text{in } \Omega_T, \\
\dfrac{\partial y^h}{\partial n} = 0 & \text{on } \Gamma_T, \\
y^h(0) = y_0 & \text{in } \Omega
\end{cases} \qquad (5.16)
$$

i.e.

$$
(\frac{\partial y^h}{\partial t}, v) + (\nabla y^h, \nabla v) + (y^h, v) = (BU_h, v), \quad \forall v \in H^1(\Omega)
$$

with $y^h(\cdot,0)=y_0$, and let $Y_h(u)\in V^h$ be the solution of

$$\begin{cases} (\dfrac{\partial Y_h(u)}{\partial t}, w_h) + (\nabla \hat{Y}_h(u), \nabla w_h) + (\hat{Y}_h(u), w_h) = (B\bar{u}, w_h),\ \forall w_h \in V^h, \\ \hfill Y_h(u)(x,0) = y_0^h(x),\ x\in\Omega. \end{cases} \tag{5.17}$$

Note that Y_h and $Y_h(u)$ are the fully discrete approximations to y^h and y, respectively. Then from (2.1) and (5.16) we have $y^h - y \in W\cap C(\bar{\Omega}_T)$, $(y^h-y)_t - \Delta(y^h-y) + (y^h-y) = B(U_h-u)\in L^\infty(\Omega_T)$, $(y^h-y)(x,0)=0$ and $\frac{\partial(y^h-y)}{\partial n}=0$. Thus $y^h-y\in W_0^\infty$. Now (2.6) implies that

$$\begin{aligned} \int_0^T (p, B(U_h-u)) &= \int_0^T (\frac{\partial(y^h-y)}{\partial t} - \Delta(y^h-y) + (y^h-y), p) \\ &= \int_0^T (y-y_d, y^h-y) + \int_{\bar{\Omega}_T} (y^h-y)d\mu. \end{aligned} \tag{5.18}$$

Similarly, from (5.10) and (5.16) we have

$$\begin{aligned} \int_0^T (\check{P}_h, B(u-U_h)) &= \int_0^T (\frac{\partial(Y_h(u)-Y_h)}{\partial t}, \check{P}_h) + (\nabla(\hat{Y}_h(u)-\hat{Y}_h), \nabla\check{P}_h) + (\hat{Y}_h(u)-\hat{Y}_h, \check{P}_h) \\ &= \int_0^T (Y_h-y_d, Y_h(u)-Y_h) + \int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h. \end{aligned} \tag{5.19}$$

Thus

$$\begin{aligned} & \alpha\|u-U_h\|^2_{L^2(0,T;L^2(\Omega))} \\ \leqslant\ & \int_0^T (y-y_d, y^h-y) + \int_{\bar{\Omega}_T}(y^h-y)d\mu \\ & + \int_0^T (Y_h-y_d, Y_h(u)-Y_h) + \int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h \\ =\ & \int_0^T (y-y_d, y^h-Y_h) + \int_0^T (y-y_d, Y_h-y) \\ & + \int_0^T (Y_h-y_d, Y_h(u)-y) + \int_0^T (Y_h-y_d, y-Y_h) \\ & + \int_{\bar{\Omega}_T}(y^h-y)d\mu + \int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h \\ =\ & -\|y-Y_h\|^2_{L^2(0,T;L^2(\Omega))} + \int_0^T (y-y_d, y^h-Y_h) + \int_0^T (Y_h-y_d, Y_h(u)-y) \\ & + \int_{\bar{\Omega}_T}(y^h-y)d\mu + \int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h. \end{aligned}$$

Therefore we have

$$\begin{aligned} & \alpha\|u-U_h\|^2_{L^2(0,T;L^2(\Omega))} + \|y-Y_h\|^2_{L^2(0,T;L^2(\Omega))} \\ \leqslant\ & \int_0^T (y-y_d, y^h-Y_h) + \int_0^T (Y_h-y_d, Y_h(u)-y) \end{aligned}$$

$$
\begin{aligned}
&+\int_{\bar{\Omega}_T}(y^h-y)d\mu+\int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h\\
\leqslant\;& C(\|y^h-Y_h\|_{L^2(0,T;L^2(\Omega))}+\|Y_h(u)-y\|_{L^2(0,T;L^2(\Omega))})\\
&+\int_{\bar{\Omega}_T}(y^h-y)d\mu+\int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h\\
\leqslant\;& C(h^2+k)+\int_{\bar{\Omega}_T}(y^h-y)d\mu+\int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h, \qquad (5.20)
\end{aligned}
$$

where we have used Theorem 4.1. 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 $\mu \geqslant 0$, with $(y^h)^+ = \max(y^h-\phi,0)$ we have

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

Note that for $t_{n-1} < t \leqslant t_n$

$$
\begin{aligned}
|(y^h)^+(x,t)| &\leqslant |(y^h)^+(x,t)-(y^{h,n})^+(x)|+|(y^{h,n})^+(x)-(Y_h^n)^+(x)|\\
&\leqslant |y^h(x,t)-y^{h,n}(x)|+|y^{h,n}(x)-Y_h^n(x)|\\
&\leqslant C\frac{1}{\tau}\int_{t_{n-1}}^{t_n}\|y^h(t)-y^h(t')\|_{\infty,\Omega}dt'+\|y^{h,n}(x)-Y_h^n(x)\|_{\infty,\Omega}\\
&\leqslant Cs^2|\log h|^2(k^{1-\frac{2}{s}}+Ch^{2-\frac{4}{s}})\|y^h\|_{W_s^{2,1}(\Omega_T)}, \qquad (5.21)
\end{aligned}
$$

where we have used Theorem 4.2. Similarly, from (5.8) and Theorem 4.2 we have

$$
\begin{aligned}
\int_{\bar{\Omega}_T}(Y_h(u)-Y_h)d\mu_h &= \int_{\bar{\Omega}_T}(Y_h(u)-y+y-\phi+\phi-Y_h)d\mu_h\\
&\leqslant \int_{\bar{\Omega}_T}(Y_h(u)-y)d\mu_h\\
&\leqslant C\|Y_h(u)-y\|_{L^\infty(\Omega_T)}\|\mu_h\|_{\mathcal{M}(\bar{\Omega}_T)}\\
&\leqslant Cs^2|\log h|^2(k^{1-\frac{2}{s}}+Ch^{2-\frac{4}{s}})\|y\|_{W_s^{2,1}(\Omega_T)}. \qquad (5.22)
\end{aligned}
$$

Combining (5.20), (5.21) and (5.22) we prove the claim.

□

In [8], 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 (1.3). In [19], 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 pointwise in time and distributed control. In this paper for some $\varepsilon > 0$ we obtain the convergence order of $O(h^{1-\varepsilon}+h^{\frac{1}{2}-\varepsilon})$, which appears to be quasi-optimal for problems with state constraints pointwise in space and time and distributed control.

6 Numerical Examples

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

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

subject to

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

with box type control constraint

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

and state constraint

$$y \in K_Y := \{y \in L^\infty(\Omega_T): \quad y(x,t) \leqslant \phi(x,t), \quad (x,t) \in \Omega_T\}.$$

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

The numerical solution of the optimal control problem is performed with the method proposed in [12] and [17], which goes back to an idea of Pierre and Sokolowsky([22]).

Example 6.1. *Let $\Omega_T = [0,1]^2 \times [0,1]$, $\alpha = 1$. Here we set $a = -2$ and $b = 2$. Following the ideas of [8] we set*

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

while control and adjoint state are given by

$$u(x,t) = t(t-1)\cos(\pi x_1)\cos(\pi x_2) \tag{6.1}$$

and

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

Note that the control constraint is not active in this example, but $u \in L^\infty(\Omega_T)$ which is the crucial ingredient for our analysis. 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

$$\begin{aligned} y_d(x,t) \quad &= \quad y(x,t) + \mu(x,t) + (1-2t)\cos(\pi x_1)\cos(\pi x_2) \\ &\quad +t(1-t)(-2\pi^2\cos(\pi x_1)\cos(\pi x_2) - \cos(\pi x_1)\cos(\pi x_2)), \end{aligned}$$

and

$$\begin{aligned} u_d(x,t) \quad &= \quad u(x,t) + p(x,t) = 0, \\ f(x,t) \quad &= \quad \pi\cos(\pi t)\cos(\pi x_1)\cos(\pi x_2) + 2\pi^2\cos(\pi x_1)\cos(\pi x_2)\sin(\pi t) + y(t,p) - u(x,t). \end{aligned}$$

To confirm our theoretical results we test the convergence order with respect to space and time discretization. We choose the time step $\Delta t = O(h^2)$ where h denotes the mesh size of space triangulation. The results are listed in Table 6.1.

Table 6.1. Error of control u and state y for example 6.1 .

h	N	$\|u - u_h\|_{L^2(\Omega_T)}$	order	$\|y - y_h\|_{L^2(\Omega_T)}$	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

It can be seen that 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.

Acknowledgements The first author would thank the support of Alexander von Humboldt Foundation during the stay in University of Hamburg, Germany. He is also very grateful to the Department of Mathematics, University of Hamburg for the hospitality and support. The second author acknowledges support from the priority program SPP1253.

References

[1] J. F. Bonnans, P. Jaisson, Optimal control of a parabolic equation with time-dependent state constraints, Rapport de Recherche INRIA RR 6784 (2008).

[2] E. Casas, Control of an elliptic problem with pointwise state constraints, SIAM J. Contr. Optim., Vol. 24, 6(1986), pp. 1309-1318.

[3] E. Casas, Boundary control of semilinear elliptic equations with pointwise state constraints, SIAM J. Cont. Optim., 31(1993), pp. 993-1006.

[4] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Cont. Optim., 35(1997), pp. 1297-1327.

[5] P. G. Ciarlet, The finite element methods for elliptic problems, North-Holland, Amsterdam, 1978.

[6] K. Deckelnick, A. Günther, M. Hinze: Finite element approximation of elliptic control problems with constraints on the gradient. Numer. Math., 111 (2009), 335-350.

[7] K. Deckelnick and M. Hinze, Convergence of a fnite element approximation to a state constrained elliptic control problem, SIAM J. Numer. Anal., 45(2007), pp. 1937–1953.

[8] K. Deckelnick, M. Hinze, Variational discretization of parabolic control problems in the presence of pointwise state constraints. J. Comput. Math., 29, No. 1 (2011), 1-16.

[9] de los Reyes, J. C. Merino, P. Rehberg, F. Tröltzsch, Optimality conditions for state constrained pde control problems with time-dependent controls, preprint, TU Berlin(2008).

[10] K. Eriksson and C. Johnson, 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(1995), 706-740.

[11] D. A. French and J. T. King, Analysis of a robust finite element approximation for a parabolic equation with rough boundary data, Math. Compu., Vol 60, 201(1993), pp. 79-104.

[12] W. Gong and N. N. Yan, A mixed finite element scheme for optimal control problems with pointwise state constraints, to appear in J. Sci. Comput.(2011).

[13] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case. Computational Optimization and Application, 30(2005), pp. 45-63.

[14] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, Optimization with PDE constraints MMTA 23, Springer, 2009.

[15] J. L. Lions, Optimal control of systems governed by partial differential equations, Springer-Verlag, Berlin, 1971.

[16] J. L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications, Springer-Verlag, Berlin, 1972.

[17] W. B. Liu, W. Gong, N. N. Yan, A new finite element approximation of a state-constrained optimal control problem. J. Comput. Math., 27(2009), no. 1, 97-114.

[18] W. B. Liu, N. N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations. Numer. Math., 93(2003), no. 3, 497–521.

[19] D. Meidner, R. Rannacher, and B. Vexler, A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time, Preprint-Nr.: SPP1253-098.

[20] I. Neitzel and F. Tröltzsch, On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints, ESAIM: COCV, 15(2009), no. 2, 426–453.

[21] R. H. Nochetto, C. Verdi, Convergence past singularities for a fully discrete approximation of curvature-driven interfaces, SIAM J. Numer. Anal., 34(1997), 490-512.

[22] Michel Pierre, Jan Sokolowsky, Differentiability of projection and applications. Control of partial differential equations and applications (Laredo, 1994), 231–240, Lecture Notes in Pure and Appl. Math., 174, Dekker, New York, 1996.

[23] R. Rannacher and R. Scott, Some optimal error estimates for piecewise linear finite element approximations, Math. Comput., 38(1982), No. 158: 437-445.

[24] G. S. Wang, X. Yu, Error estimates for an optimal control problem governed by the heat equation with state and control constraints. Int. J. Numer. Anal. Model., 7(2010), no. 1, 30-65.

[25] A. H. Schatz, Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids. I: Global estimates, Math. Comp., 67 (1998), pp. 877899.