CrossMark

Finite Element Method and A Priori Error Estimates for Dirichlet Boundary Control Problems Governed by Parabolic PDEs

Wei Gong[1] · **Michael Hinze**[2] · **Zhaojie Zhou**[3]

Abstract Finite element approximations of Dirichlet boundary control problems governed by parabolic PDEs on convex polygonal domains are studied in this paper. The existence of a unique solution to optimal control problems is guaranteed based on very weak solution of the state equation and $L^2(0, T; L^2(\Gamma))$ as control space. For the numerical discretization of the state equation we use standard piecewise linear and continuous finite elements for the space discretization of the state, while a dG(0) scheme is used for time discretization. The Dirichlet boundary control is realized through a space–time L^2-projection. We consider both piecewise linear, continuous finite element approximation and variational discretization for the controls and derive a priori L^2-error bounds for controls and states. We finally present numerical examples to support our theoretical findings.

✉ Wei Gong
wgong@lsec.cc.ac.cn

Michael Hinze
michael.hinze@uni-hamburg.de

Zhaojie Zhou
zzj534@amss.ac.cn

1 LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

2 Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

3 School of Mathematics Sciences, Shandong Normal University, Jinan 250014, China

⚛ Springer

1 Introduction

In this paper we study the following parabolic optimal control problem:

$$\min_{u \in U_{ad}} \quad J(y, u) = \tfrac{1}{2} \|y - y_d\|_{L^2(0,T;L^2(\Omega))}^2 + \tfrac{\alpha}{2} \|u\|_{L^2(0,T;L^2(\Gamma))}^2 \tag{1.1}$$

subject to

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y = f & \text{in } \Omega_T, \\ y = u & \text{on } \Sigma, \\ y(0) = y_0 & \text{in } \Omega, \end{cases} \tag{1.2}$$

where $\Omega_T = \Omega \times (0, T]$, $\Sigma = \partial \Omega \times (0, T]$ with Ω denoting an open bounded domain with boundary $\Gamma := \partial \Omega$, U_{ad} is the admissible control set which is assumed to be of box type

$$U_{ad} := \left\{ u \in L^2(0, T; L^2(\Gamma)) : u_a \le u(x, t) \le u_b, \text{ a.e. on } \Sigma \right\}, \tag{1.3}$$

with $u_a < u_b$ denoting constants. For the convenience we make the following assumption on the domain Ω and the given data which shall be valid throughout the paper without explicit mentioning:

Assumption 1 We assume that Ω is an open bounded, convex polygonal domain in $\mathbb{R}^2$. $\alpha > 0$, $f \in L^2(0, T; L^2(\Omega))$, $y_0 \in L^2(\Omega)$, $y_d \in L^2(0, T; L^2(\Omega))$ and $T > 0$ are fixed data.

Dirichlet boundary control is important in many practical applications such as the active boundary control of flows, see e.g. [13,18,20]. If one is, e.g. interested in blowing and suction as control on part of the boundary, controls with low regularity should be admissible, which could have jumps and satisfy pointwise bounds. In the mathematical theory one has to use the concept of very weak solutions in this situation, see [4] for a more detailed discussion of this fact.

In the present work we consider a parabolic Dirichlet boundary control problem of tracking type, which may be regarded as prototype problem to study Dirichlet boundary control for time-dependent PDEs. For parabolic optimal boundary control problems of Dirichlet type, only few contributions can be found in the literature [2,3,23]. Kunisch and Vexler [23] considered a semi-smooth Newton method for the numerical solution of parabolic Dirichlet boundary control problems. A Robin penalization method using Robin-type boundary conditions applied to parabolic Dirichlet boundary control problems is investigated in [3]. However, to the best of the authors' knowledge no error analysis is available for the finite element approximation of this kind of problems. With the present paper we intend to fill this gap and derive a priori error estimates for parabolic Dirichlet boundary control problems. Compared to the elliptic case, parabolic Dirichlet boundary control problems are more involved in both the definition of discrete schemes and the a priori error analysis, since the regularity of the involved state variable is low.

Finite element approximations of optimal control problems are important for the numerical treatment of optimal control problems related to practical applications, see e.g. [22, Ch. 4]. An overview on the numerical a priori and a posteriori analysis for elliptic control problems can be found in [22, Ch. 3] and [28], respectively. To the best of the authors' knowledge the first contribution to parabolic optimal control problems is given in [36]. The state of the art in the numerical a priori analysis of distributed parabolic optimal control problems can be found in [30,31]. More recent contributions with higher order in time Galerkin schemes can be found in [1,32,35]. Residual-based a posteriori error estimates are presented in [26] and [27]. For

boundary control problems with parabolic equations we refer to [15]. There is a long list of contributions to boundary control of elliptic PDEs, see e.g. [7,8,10,11,14,16,21,29,34]. Further references can be found in [22, Ch. 3].

In this paper we use the very weak solution concept for the state equation and $L^2(0, T; L^2(\Gamma))$ as control space to argue the existence of a unique solution to the optimal control problems (1.1)–(1.2). For the numerical discretization of the optimal control problem we discretize the state using standard piecewise linear and continuous finite elements in space and dG(0) scheme in time. The Dirichlet boundary conditions are approximated based on the space–time L^2-projection. The control is discretized in space either by piecewise linear finite elements or implicitly through the discretization of the adjoint state, the so-called variational discretization (see [19]). For both cases we derive a priori error bounds for the state and control in the L^2-norm for problems posed on polygonal domains. As main result we obtain the error bound

$$\|u - U_{hk}\|_{L^2(L^2(\Gamma))} + \|y - Y_{hk}\|_{L^2(L^2(\Omega))} \leq Ch^{\frac{1}{2}} \tag{1.4}$$

under the coupling $k = O(h^2)$ with both full control discretisation and variational control discretisation for the optimal solution (y, u) of (1.1), where Y_{hk} and U_{hk} denote the optimal discrete state and control, see also Corollary 1. We present several numerical examples which support our theoretical findings.

The rest of our paper is organized as follows. In Sect. 2 we present the analytical setting of the parabolic Dirichlet boundary control problem and argue the existence of a unique solution. In Sect. 3 we establish the fully discrete finite element approximation to the state equation and the corresponding stability results. Then we formulate the fully discrete approximation for parabolic Dirichlet boundary control problems. The a priori error analysis for the finite element approximation and the variational discretization of the optimal control problems posed on convex, polygonal domains is studied in Sect. 4. Furthermore, we present some numerical experiments in Sect. 5 to support our theoretical results.

2 Optimal Control Problem

For $m \geq 0$ and $1 \leq s \leq \infty$, we adopt the standard notation $W^{m,s}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{m,s,\Omega}$ and seminorm $|\cdot|_{m,s,\Omega}$, where $H^m(\Omega) = W^{m,2}(\Omega)$, $\|\cdot\|_{m,\Omega} = \|\cdot\|_{m,2,\Omega}$ and $|\cdot|_{m,\Omega} = |\cdot|_{m,2,\Omega}$ for $s = 2$. Note that $H^0(\Omega) = L^2(\Omega)$ and $H_0^1(\Omega) = \{v \in H^1(\Omega); \ v = 0 \text{ on } \partial\Omega\}$. We denote by $L^r(0, T; W^{m,s}(\Omega))$ the Banach space of all L^r integrable functions from $[0, T]$ into $W^{m,s}(\Omega)$ with norm $\|v\|_{L^r(0,T;W^{m,s}(\Omega))} = \left(\int_0^T \|v\|_{m,s,\Omega}^r dt\right)^{\frac{1}{r}}$ for $1 \leq r < \infty$, and with the standard modification for $r = \infty$. For a Banach space Y, we use the abbreviations $L^2(Y) = L^2(0, T; Y)$, $H^s(Y) = H^s(0, T; Y)$, $s = [0, \infty)$, and $C(Y) = C([0, T]; Y)$. We denote the L^2-inner products on $L^2(\Omega)$, $L^2(\Omega_T)$ and $L^2(\Gamma)$ by $(\cdot, \cdot)$, $(\cdot, \cdot)_{\Omega_T}$ and $\langle\cdot, \cdot\rangle$, respectively. In addition c and C denote generic positive constants.

Let

$$a(y, w) = \int_\Omega \nabla y \cdot \nabla w \quad \forall \, y, w \in H^1(\Omega).$$

The standard weak form for the parabolic equation (1.2) is to find $y \in L^2(H^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ with $y|_\Sigma = u$ and $y(\cdot, 0)|_\Omega = y_0(\cdot)$ such that

$$\left(\frac{\partial y}{\partial t}, v\right) + a(y, v) = (f, v) \text{ a.a. } t \in (0, T], \ \forall \, v \in H_0^1(\Omega). \tag{2.1}$$

This setting requires $u \in L^2(H^{\frac{1}{2}}(\Gamma))$. Motivated by practical considerations (see e.g. the discussion in [4]) we are interested in controls $u \in U_{ad}$ defined in (1.3). For a proper treatment of the state equation in this case we use the transposition technique introduced by Lions and Magenes (see [24, Ch. 2, Sec. 5.2] and [25, Ch. 2]) to argue the existence of a unique solution to the state equation (1.2) in the present paper. The very weak form of (1.2) that we shall utilize reads: Find $y \in L^2(L^2(\Omega))$ such that

$$\int_{\Omega_T} y(-z_t - \Delta z)dxdt = -\int_\Sigma u\partial_n z \, dsdt + \int_{\Omega_T} f z \, dxdt + \int_\Omega y_0 z(\cdot, 0)dx$$
$$\forall \, z \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega)) \tag{2.2}$$

with $z(\cdot, T) = 0$ holds, where $\partial_n v := \nabla v \cdot n$ with n denoting the unit outward normal to Γ. Then the existence and uniqueness of a very weak solution of (2.2), which we denote by $y = \mathcal{G}(u)$, is shown in the following lemma (see, e.g. [25])

Lemma 1 *For each* $u \in L^2(L^2(\Gamma))$, *there exists a unique very weak solution* $y \in L^2(L^2(\Omega))$ *of* (2.2) *satisfying*

$$\|y\|_{L^2(L^2(\Omega))} \leq C\big(\|y_0\|_{L^2(\Omega)} + \|f\|_{L^2(L^2(\Omega))} + \|u\|_{L^2(L^2(\Gamma))}\big). \tag{2.3}$$

Proof For $y_0 \in L^2(\Omega)$, $f \in L^2(L^2(\Omega))$ and $u \equiv 0$ it is straightforward to show that (1.2) admits a unique solution $y \in L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ in the sense of (2.1), which also satisfies (2.3). To prove the lemma in the case $u \neq 0$ it is sufficient to consider the case $f \equiv 0$, $y_0 \equiv 0$, where we follow the constructive approach of [10]. For each $g \in L^2(L^2(\Omega))$ we denote by $z \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega))$ the solution of

$$\begin{cases} -\frac{\partial z}{\partial t} - \Delta z = g & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma, \\ z(T) = 0 & \text{in } \Omega. \end{cases} \tag{2.4}$$

Then we have $\partial_n z \in H^{\frac{1}{4}}(L^2(\Gamma))$ according to [23, Th. 3.2]. Moreover, from the fact that $z \in L^2(H^2(\Omega))$ and $z = 0$ on Σ we obtain that $\partial_n z \in L^2(H^{\frac{1}{2}}(\Gamma))$ according to Lemma A.2 in [6]. We denote by $T : L^2(L^2(\Omega)) \rightarrow L^2(L^2(\Gamma))$ the continuous linear operator which is defined by $Tg := -\partial_n z|_\Sigma$ and denote its adjoint by T^*. Then with $y = T^*u$ we have

$$\int_{\Omega_T} yg \, dxdt = \int_{\Omega_T} y(-z_t - \Delta z)dxdt = \int_{\Omega_T} T^*ug \, dxdt = -\int_\Sigma u\partial_n z \, dsdt,$$

which verifies that y satisfies (2.2). The estimate (2.3) follows by observing that

$$|\int_{\Omega_T} yg \, dxdt| \leq C\|u\|_{L^2(L^2(\Gamma))}\|\partial_n z\|_{L^2(L^2(\Gamma))} \leq C\|u\|_{L^2(L^2(\Gamma))}\|g\|_{L^2(L^2(\Omega))}.$$

$\square$

Now we are ready to formulate the optimal control problem considered in the present paper. It reads

$$\begin{cases} \min \ J(y, u) = \frac{1}{2}\|y - y_d\|_{L^2(L^2(\Omega))}^2 + \frac{\alpha}{2}\|u\|_{L^2(L^2(\Gamma))}^2 \\ \qquad \text{over } (y, u) \in L^2(L^2(\Omega)) \times L^2(L^2(\Gamma)) \\ \text{subject to } y = \mathcal{G}(u) \text{ defined in (2.2) and } u \in U_{ad}. \end{cases} \tag{2.5}$$

By standard arguments (see, e.g. [24, Ch. 2, Sec. 1.2]), there exists a unique solution (y, u) for problem (2.5). Let $J(u) := J(y(u), u)$ denote the reduced cost functional, where for each $u \in L^2(L^2(\Gamma))$ the state $y(u)$ is the unique very weak solution of (2.2). Then J is infinitely often Fréchet differentiable. Moreover, the first order sufficient and necessary optimality conditions for problem (2.5) are given by

Theorem 1 *Assume that $u \in L^2(L^2(\Gamma))$ is the unique solution of problem (2.5) and let y be the associated state. Then there exists a unique adjoint state $z \in L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ such that*

$$\begin{cases} -\frac{\partial z}{\partial t} - \Delta z = y - y_d & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma, \\ z(T) = 0 & \text{in } \Omega, \end{cases} \tag{2.6}$$

and

$$J'(u)(v - u) = \int_\Sigma (\alpha u - \partial_n z)(v - u) ds dt \geq 0, \quad \forall v \in U_{ad}. \tag{2.7}$$

We note that (2.7) is equivalent to

$$J'(u)(v - u) = \int_\Sigma \alpha u(v - u) ds dt + \int_{\Omega_T} (y - y_d)(y(v) - y) dx dt$$
$$\geq 0, \quad \forall v \in U_{ad} \tag{2.8}$$

or

$$u = P_{U_{ad}}\left(\frac{1}{\alpha} \partial_n z\right), \tag{2.9}$$

where for each $v \in L^2(L^2(\Gamma))$, $y(v)$ is the solution of problem (2.2) with u replaced by v, and $P_{U_{ad}} : L^2(L^2(\Gamma)) \to U_{ad}$ denotes the orthogonal projection.

We now turn to the regularity properties of optimal controls u on Σ. The proof of the following theorem can be found in, e.g. [23, Th. 3.4].

Theorem 2 *Let $(y, u, z) \in L^2(L^2(\Omega)) \times L^2(L^2(\Gamma)) \times L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ be the solution of optimal control problem (2.5)–(2.8). Then we have*

$$u \in L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma)), \quad y \in L^2(H^1(\Omega)) \cap H^{\frac{1}{2}}(L^2(\Omega)), \tag{2.10}$$

and

$$z \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega)). \tag{2.11}$$

Proof From $f \in L^2(L^2(\Omega))$, $y_0 \in L^2(\Omega)$ and $u \in L^2(L^2(\Gamma))$ we conclude that $y \in L^2(L^2(\Omega))$ according to Lemma 1. Thus, $y_d \in L^2(L^2(\Omega))$ implies $z \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega))$, which in turn implies $\partial_n z \in L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))$ (see [6,17, 23]). From (2.9) we obtain that $u \in L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))$ and thus $y \in L^2(H^1(\Omega)) \cap H^{\frac{1}{2}}(L^2(\Omega))$ (see [25, Vol. II, p. 78]). This completes the proof. $\square$

In our analysis we frequently use results of the following backward in time parabolic problem:

$$\begin{cases} -w_t - \Delta w = g & \text{in } \Omega_T, \\ w = 0 & \text{on } \Sigma, \\ w(T) = 0 & \text{in } \Omega. \end{cases} \tag{2.12}$$

If $g \in L^2(L^2(\Omega))$, then (2.12) has a unique solution $w \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega))$ satisfying

$$\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \leq C\|g\|_{L^2(L^2(\Omega))}, \tag{2.13}$$

$$\|w(0)\|_{1,\Omega} \leq C\|g\|_{L^2(L^2(\Omega))}. \tag{2.14}$$

3 Finite Element Discretization of the State Equation and Optimal Control Problems

At first let us consider the finite element approximation of the state Eq. (1.2). For the spatial discretization we consider conforming Lagrange triangular elements.

Let $\mathcal{T}^h$ be a quasi-uniform partitioning of Ω into disjoint regular triangles τ, so that $\bar{\Omega} = \bigcup_{\tau \in \mathcal{T}^h} \bar{\tau}$. Associated with $\mathcal{T}^h$ is a finite dimensional subspace V^h of $C(\bar{\Omega})$, such that for $\chi \in V^h$ and $\tau \in \mathcal{T}^h$, $\chi|_\tau$ are piecewise linear polynomials. We set $V_0^h = V^h \cap H_0^1(\Omega)$.

Let $\mathcal{T}_U^h$ be a partitioning of Γ into disjoint regular segments s, so that $\Gamma = \bigcup_{s \in \mathcal{T}_U^h} \bar{s}$. Associated with $\mathcal{T}_U^h$ is another finite dimensional subspace U^h of $L^2(\Gamma)$, such that for $\chi \in U^h$ and $s \in \mathcal{T}_U^h$, $\chi|_s$ are piecewise linear polynomials. Here we suppose that $\mathcal{T}_U^h$ is the restriction of $\mathcal{T}^h$ on the boundary Γ and $U^h = V^h(\Gamma)$, where $V^h(\Gamma)$ is the restriction of V^h on the boundary Γ.

For the standard Lagrange interpolation operator $I_h : C(\bar{\Omega}) \to V^h$, we have the following error estimate (see, e.g. [9, Sec. 3.1])

$$\|w - I_h w\|_{l,\Omega} \leq Ch^{m-l}\|w\|_{m,\Omega}, \quad 0 \leq l \leq 1 \leq m \leq 2. \tag{3.1}$$

To define our discrete scheme, we need to introduce some projection operators. Here $Q_h : L^2(\Gamma) \to V^h(\Gamma)$ and $\tilde{Q}_h : L^2(\Omega) \to V_0^h$ denote the orthogonal projection operators. Furthermore, $R_h : H_0^1(\Omega) \to V_0^h$ denotes the Ritz projection operator defined as

$$a(R_h w, v_h) = a(w, v_h), \quad w \in H_0^1(\Omega), \ \forall \, v_h \in V_0^h. \tag{3.2}$$

It is well known that the Ritz projection satisfies (see, e.g. [9, Sec. 3.1])

$$\|w - R_h w\|_{s,\Omega} \leq Ch^{l-s}\|w\|_{l,\Omega}, \, w \in H_0^1(\Omega) \cap H^l(\Omega), \forall \, 0 \leq s \leq 1 \leq l \leq 2. \tag{3.3}$$

For the $L^2(\Gamma)$ projection operator Q_h we also have (see [9] and [12, pp. 85–86, Eq. (25) and (28)])

$$\|w - Q_h w\|_{0,\Gamma} \leq Ch^{s-\frac{1}{2}}\|w\|_{s,\Omega} \ \text{for } w \in H^s(\Omega), \ \frac{1}{2} \leq s \leq 2, \tag{3.4}$$

and

$$\|(I - Q_h)\partial_n w\|_{0,\Gamma} \leq Ch^{\frac{1}{2}}\|w\|_{2,\Omega} \ \text{for } w \in H^2(\Omega). \tag{3.5}$$

In our following analysis we need estimates for discrete harmonic functions.

Lemma 2 *Let $v_h \in V^h(\Gamma)$, and suppose that $w \in H^1(\Omega)$ is the solution of*

$$a(w, \phi) = 0, \quad \forall \, \phi \in H_0^1(\Omega), \quad w = v_h \text{ on } \Gamma \tag{3.6}$$

and $w_h \in V^h$ is the solution of

$$a(w_h, \phi_h) = 0, \quad \forall \, \phi_h \in V_0^h, \quad w_h = v_h \text{ on } \Gamma. \tag{3.7}$$

Then

$$\|w - w_h\|_{1,\Omega} \le C\|v_h\|_{\frac{1}{2},\Gamma} \le Ch^{-\frac{1}{2}}\|v_h\|_{0,\Gamma}, \tag{3.8}$$

$$\|w_h\|_{0,\Omega} + h^{\frac{1}{2}}\|w_h\|_{1,\Omega} \le C\|v_h\|_{0,\Gamma}, \tag{3.9}$$

$$\|w - w_h\|_{\frac{1}{2},\Omega} \le C\|v_h\|_{0,\Gamma}. \tag{3.10}$$

Proof The proof of (3.8) and (3.9) can be found in [5, Lm. 3.2], [7, Th. 5.4] and [11, Lm. 1]. Here we provide a proof of (3.10). For each $g \in H^{-\frac{1}{2}}(\Omega)$ let $\psi_g \in H^{\frac{3}{2}}(\Omega) \cap H_0^1(\Omega)$ be the solution of

$$a(\phi, \psi_g) = \langle g, \phi \rangle_{H^{-\frac{1}{2}}, H^{\frac{1}{2}}}, \quad \forall \phi \in H_0^1(\Omega). \tag{3.11}$$

Then we have $\|\psi_g\|_{\frac{3}{2},\Omega} \le C\|g\|_{-\frac{1}{2},\Omega}$. Note that from (3.6) and (3.7) we have

$$\langle g, w - w_h \rangle_{H^{-\frac{1}{2}}, H^{\frac{1}{2}}} = a(w - w_h, \psi_g) = a(w - w_h, \psi_g - I_h\psi_g),$$

where $I_h\psi_g$ is the linear Lagrange interpolation of ψ_g [9]. Then standard error estimates lead to

$$\begin{aligned}
\langle g, w - w_h \rangle_{H^{-\frac{1}{2}}, H^{\frac{1}{2}}} &= a(w - w_h, \psi_g - I_h\psi_g) \\
&\le \|w - w_h\|_{1,\Omega}\|\psi_g - I_h\psi_g\|_{1,\Omega} \\
&\le Ch^{\frac{1}{2}}\|v_h\|_{\frac{1}{2},\Gamma}\|\psi_g\|_{\frac{3}{2},\Omega} \\
&\le C\|v_h\|_{0,\Gamma}\|g\|_{-\frac{1}{2},\Omega}, \tag{3.12}
\end{aligned}$$

where we have used the estimate (3.8). This implies

$$\|w_h - w\|_{\frac{1}{2},\Omega} \le C\|v_h\|_{0,\Gamma},$$

which proves (3.10). $\qquad\square$

The semi-discrete finite element approximation of (1.2) reads: Find $y_h(u) \in L^2(V^h)$ such that

$$\begin{cases}
-(y_h(u), \partial_t v_h)_{\Omega_T} + a(y_h(u), v_h)_{\Omega_T} = (f, v_h)_{\Omega_T} + \left(y_0^h, v_h(\cdot, 0)\right) \quad \forall v_h \in H^1(V_0^h), \\
y_h(u) = Q_h(u) \text{ on } \Sigma
\end{cases} \tag{3.13}$$

with $v_h(\cdot, T) = 0$, $y_0^h = \tilde{Q}_h y_0 \in V^h$ an approximation of y_0 using the L^2-projection, and Q_h the projection operator from $L^2(\Gamma)$ to $V^h(\Gamma)$. Note that the above semi-discrete scheme is well-defined and admits a unique solution $y_h(u) \in L^2(H^1(\Omega))$, which we denote by $y_h(u) = \mathcal{G}_h(u)$, since $Q_h(u) \in L^2(H^{\frac{1}{2}}(\Gamma))$, thus we use a standard bilinear form $a(\cdot, \cdot)$ compared to the very weak form (2.2).

The semi-discrete finite element approximation of (1.1)–(1.2) reads as follows:

$$\begin{cases}
\min_{u_h \in U_{ad}^h, y_h \in L^2(V^h)} J_h(y_h, u_h) = \dfrac{1}{2}\|y_h - y_d\|_{L^2(L^2(\Omega))}^2 + \dfrac{\alpha}{2}\|u_h\|_{L^2(L^2(\Gamma))}^2 \\
\qquad\qquad \text{subject to } y_h = \mathcal{G}_h(u_h) \text{ defined in (3.13)},
\end{cases} \tag{3.14}$$

where $y_0^h \in V^h$ is an approximation of y_0, and U_{ad}^h is an appropriate approximation to U_{ad} depending on the discretization scheme for the control.

It follows that the control problem (3.14) has a unique solution (y_h, u_h) and that a pair (y_h, u_h) is the solution of the problem (3.14) if and only if there is a co-state $z_h \in L^2(V_0^h)$ such that the triplet (y_h, z_h, u_h) satisfies (3.13) and the following optimality conditions:

$$\begin{cases} -\left(\frac{\partial z_h}{\partial t}, q_h\right) + a(q_h, z_h) = (y_h - y_d, q_h), & \forall q_h \in V_0^h, \\ z_h = 0 \text{ on } \Sigma; \ z_h(T) = 0 \text{ in } \Omega, \end{cases} \tag{3.15}$$

$$\int_{\Omega_T} (y_h - y_d)(y_h(v_h) - y_h) dx dt + \alpha \int_{\Sigma} u_h(v_h - u_h) ds dt \geq 0, \quad \forall v_h \in U_{ad}^h, \tag{3.16}$$

where $y_h(v_h) \in L^2(V^h)$ is the solution of state equation (3.13) with Dirichlet boundary condition $Q_h(v_h)$.

We next consider the fully discrete approximation for above semi-discrete problem by using the dG(0) scheme in time [33]. We note that the dG(0) scheme is equivalent to the backward Euler method with the right hand side approximated by the averaged integral.

Let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a time domain partitioning with $k_n = t_n - t_{n-1}$, $n = 1, 2, \ldots, N$ and $k = \max_{1 \leq n \leq N} k_n$. We assume that the time partitioning is quasi-uniform, i.e., there exist positive constants c_1 and c_2 such that $c_1 k_n \leq k \leq c_2 k_n$ holds for each $n = 1, 2, \ldots, N$. We also set $I_n := (t_{n-1}, t_n]$. For $n = 1, 2, \ldots, N$, we construct the finite element spaces $V^h \subset H^1(\Omega)$ with the mesh $\mathcal{T}^h$. Similarly, we construct the finite element spaces $U^h \subset L^2(\Gamma)$ with the mesh $\mathcal{T}_U^h$. In our case we have $U^h = V^h(\Gamma)$. Then we denote by V^h and U^h the finite element spaces defined on $\mathcal{T}^h$ and $\mathcal{T}_U^h$ on each time step.

Let V_k denote the space of piecewise constant functions on the time partition. We define the L^2 projection operator $P_k : L^2(0, T) \to V_k$ on I_n through

$$P_k^n w := (P_k w)(t)|_{I_n} = \frac{1}{k_n} \int_{I_n} w(s) ds \text{ for } t \in I_n \ (n = 1, \ldots, N).$$

Then we have the following estimate

$$\|(I - P_k)w\|_{L^2(0,T;H)} \leq Ck\|w_t\|_{L^2(0,T;H)}, \quad \forall w \in H^1(0, T; H), \tag{3.17}$$

where H denotes some separable Hilbert space .

We consider a dG(0) scheme for the time discretization and set

$$V_{hk} := \left\{ \phi : \overline{\Omega} \times [0, T] \to \mathbb{R}, \ \phi(\cdot, t)|_{\overline{\Omega}} \in V^h, \ \phi(x, \cdot)|_{I_n} \in \mathbb{P}_0 \text{ for } n = 1, \ldots, N \right\},$$

i.e. $\phi \in V_{hk}$ is a piecewise constant polynomial w.r.t. time. We also set $V_{hk}(\Gamma)$ as the restriction of V_{hk} on $L^2(L^2(\Gamma))$. We set $Q = Q_h P_k = P_k Q_h$. Thus, we have $Q : L^2(L^2(\Gamma)) \to V_{hk}(\Gamma)$. For $Y, \Phi \in V_{hk}$ we set

$$A(Y, \Phi) := \sum_{n=1}^{N} k_n a(Y^n, \Phi^n) + \sum_{n=2}^{N} (Y^n - Y^{n-1}, \Phi^n) + (Y_+^0, \Phi_+^0),$$

where $\Phi^n := \Phi_-^n = \lim_{s \to 0^+} \Phi(t_n - s)$, $\Phi^{n+1} := \Phi_+^n = \lim_{s \to 0^+} \Phi(t_n + s)$.

For each $u \in L^2(L^2(\Gamma))$ the fully discrete dG(0)-cG(1) finite element approximation of (3.13) now reads: Find $Y_{hk} \in V_{hk}$ such that

$$\begin{cases} A(Y_{hk}, \Phi) = (f, \Phi)_{\Omega_T} + (y_0, \Phi_+^0), \ \forall \Phi \in V_{hk}^0, \\ \qquad\qquad\qquad Y_{hk} = Q(u) \text{ on } \Sigma, \end{cases} \tag{3.18}$$

where V_{hk}^0 denotes the subspace of V_{hk} with functions vanishing on the boundary Γ.

It is easy to see that on each time interval I_n, $Y_{hk}^n \in V^h$ solves the following problem:

$$\begin{cases} \left(\dfrac{Y_{hk}^n - Y_{hk}^{n-1}}{k_n}, w_h\right) + a(Y_{hk}^n, w_h) = (P_k^n f, w_h), \ \forall \, w_h \in V_0^h, \ \ n = 1, \dots, N, \\ Y_{hk}^0 = y_0^h, \ \text{in } \Omega; \ \ Y_{hk}^n = Q_h(P_k^n u), \ n = 1, \dots, N \ \text{on } \Gamma. \end{cases} \tag{3.19}$$

Here we use the L^2-projection to approximate the non-smooth Dirichlet boundary condition in (3.18).

In the following we need to investigate the stability behavior of the fully discrete scheme (3.18) with respect to the initial value y_0, the right hand side f and the Dirichlet boundary conditions u.

Lemma 3 *There exists a constant C independent of h, k and the data (f, y_0) such that*

$$\sum_{n=1}^N \left(\|Y_{hk}^n - Y_{hk}^{n-1}\|_{0,\Omega}^2 + k_n \|Y_{hk}^n\|_{1,\Omega}^2 \right) + \|Y_{hk}^N\|_{0,\Omega}^2$$
$$\le C(h^{-1} + hk^{-1}) \|u\|_{L^2(L^2(\Gamma))}^2 \tag{3.20}$$

and

$$\sum_{n=1}^N \left(\|Y_{hk}^n - Y_{hk}^{n-1}\|_{0,\Omega}^2 + k_n \|Y_{hk}^n\|_{1,\Omega}^2 \right) + \|Y_{hk}^N\|_{0,\Omega}^2$$
$$\le C(1 + h^2 k^{-1}) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))}^2 \tag{3.21}$$

hold in case $f \equiv 0$, $y_0 \equiv 0$. In the case $u \equiv 0$ the estimate

$$\sum_{n=1}^N \|Y_{hk}^n - Y_{hk}^{n-1}\|_{0,\Omega}^2 \le C\left(h^{-2}k \|y_0\|_{0,\Omega}^2 + k \|f\|_{L^2(L^2(\Omega))}^2\right) \tag{3.22}$$

is valid.

Proof Let us first assume that $f \equiv 0$, $y_0 \equiv 0$. The proof follows the idea of [12] completely and here we give a sketch for the case with variable time steps. To begin with we introduce the following problem: Find $y_u \in V_{hk}$ with

$$\begin{cases} \left(\dfrac{y_u^n - y_u^{n-1}}{k_n}, w_h\right) + a(y_u^n, w_h) = 0, \ \forall \, w_h \in V_0^h, \ \ n = 1, \dots, N, \\ y_u^0 = 0 \ \text{in } \Omega; \ \ y_u^n = Q_h(P_k^n u), \ n = 1, \dots, N \ \text{on } \Gamma. \end{cases} \tag{3.23}$$

For arbitrary $y_h \in V^h$ we have the splitting

$$y_h = y_1 + R_h y_h \quad \text{and} \quad y_h = y_2 + \tilde{Q}_h y_h,$$

where $\tilde{Q}_h y_h \in V_0^h$ and $R_h y_h \in V_0^h$ are the L^2-projection and Ritz-projection of y_h, respectively. Then we have $y_2|_\Gamma = y_h$, $y_1|_\Gamma = y_h$ and

$$(y_2, v_h) = 0 \quad \text{and} \quad a(y_1, v_h) = 0, \ \ \forall \, v_h \in V_0^h.$$

Let $y_u^n = y_2^n + \tilde{Q}_h y_u^n$. Then (3.23) delivers

$$(\tilde{Q}_h y_u^n - \tilde{Q}_h y_u^{n-1}, w_h) + k_n a(\tilde{Q}_h y_u^n, w_h) = -k_n a(y_2^n, w_h), \ \ \forall \, w_h \in V_0^h. \tag{3.24}$$

Similar to the proof of Proposition 1 in [12, P. 88] we conclude from (3.24) that

$$\sum_{n=1}^{N} \left(\|y_u^n - y_u^{n-1}\|_{0,\Omega}^2 + k_n a(y_u^n, y_u^n) \right) + \|y_u^N\|_{0,\Omega}^2$$
$$\leq C \sum_{n=1}^{N} \left(k_n a(y_2^n, y_2^n) + \|y_2^n\|_{0,\Omega}^2 \right). \tag{3.25}$$

For $y_u^n \in V^h$ we also have the splitting $y_u^n = y_1^n + R_h y_u^n$. It follows from the proof of Lemma 3 in [12, P. 87] that

$$\|y_2^n\|_{0,\Omega} \leq C h^{\frac{1}{2}} \|y_1^n\|_{\frac{1}{2},\Omega}.$$

Similarly, we also have

$$\|y_2^n\|_{0,\Omega} \leq C h \|y_1^n\|_{1,\Omega}.$$

We note that $y_1^n|_\Gamma = y_u^n = Q_h(P_k^n u) \in H^{\frac{1}{2}}(\Gamma)$ and

$$a(y_1^n, \phi_h) = 0, \quad \forall \, \phi_h \in V_0^h.$$

Let $w^n \in H^1(\Omega)$ be the solution of (3.6) with v_h substituted by $Q_h(P_k^n u)$. Then

$$\|w^n\|_{\frac{1}{2},\Omega} \leq C \|Q_h(P_k^n u)\|_{0,\Gamma}, \quad \|w^n\|_{1,\Omega} \leq C \|Q_h(P_k^n u)\|_{\frac{1}{2},\Gamma}$$

and y_1^n is the finite element approximation to w^n. So we deduce from Lemma 2 that

$$\|y_1^n\|_{\frac{1}{2},\Omega} \leq \|y_1^n - w^n\|_{\frac{1}{2},\Omega} + \|w^n\|_{\frac{1}{2},\Omega}$$
$$\leq C \|Q_h(P_k^n u)\|_{0,\Gamma}$$

and

$$\|y_1^n\|_{1,\Omega} \leq \|y_1^n - w^n\|_{1,\Omega} + \|w^n\|_{1,\Omega}$$
$$\leq C \|Q_h(P_k^n u)\|_{\frac{1}{2},\Gamma},$$

which in turn give

$$\|y_2^n\|_{0,\Omega} \leq C h^{\frac{1}{2}} \|Q_h(P_k^n u)\|_{0,\Gamma}, \quad \|y_2^n\|_{0,\Omega} \leq C h \|Q_h(P_k^n u)\|_{\frac{1}{2},\Gamma}.$$

Inverse estimates also yield

$$\|y_2^n\|_{1,\Omega} \leq C h^{-\frac{1}{2}} \|Q_h(P_k^n u)\|_{0,\Gamma}.$$

With the help of above estimates and norm interpolation we are led to

$$\|y_2^n\|_{s,\Omega} \leq C h^{\frac{1}{2}-s} \|Q_h(P_k^n u)\|_{0,\Gamma}, \quad \|y_2^n\|_{s,\Omega} \leq C h^{1-s} \|Q_h(P_k^n u)\|_{\frac{1}{2},\Gamma} \tag{3.26}$$

for all $0 \leq s \leq 1$ and $n = 1, 2, \ldots, N$. Thus, from the quasi-uniformality of time partioning we have

$$
\sum_{n=1}^{N} \left(\|y_u^n - y_u^{n-1}\|_{0,\Omega}^2 + k_n a(y_u^n, y_u^n) \right) + \|y_u^N\|_{0,\Omega}^2
$$

$$
\leq C \sum_{n=1}^{N} \left(k_n a(y_2^n, y_2^n) + \|y_2^n\|_{0,\Omega}^2 \right)
$$

$$
\leq C \sum_{n=1}^{N} \left(k_n h^{-1} \|Q_h(P_k^n u)\|_{0,\Gamma}^2 + h \|Q_h(P_k^n u)\|_{0,\Gamma}^2 \right)
$$

$$
\leq C(h^{-1} + hk^{-1}) \sum_{n=1}^{N} \int_{I_n} \|Q_h(P_k^n u)\|_{0,\Gamma}^2
$$

$$
\leq C(h^{-1} + hk^{-1}) \|Q(u)\|_{L^2(L^2(\Gamma))}^2 \leq C(h^{-1} + hk^{-1}) \|u\|_{L^2(L^2(\Gamma))}^2. \tag{3.27}
$$

This gives

$$
\sum_{n=1}^{N} \left(\|y_u^n - y_u^{n-1}\|_{0,\Omega}^2 + k_n \|y_u^n\|_{1,\Omega}^2 \right) + \|y_u^N\|_{0,\Omega}^2
$$

$$
\leq C(h^{-1} + hk^{-1}) \|u\|_{L^2(L^2(\Gamma))}^2. \tag{3.28}
$$

Similarly, we can derive from (3.26) and the $W^{s,q}(\Gamma)$ ($0 \leq s \leq 1$ and $1 \leq q \leq \infty$) stability of L^2-projection operator Q_h (see [7, P. 1601]) that

$$
\sum_{n=1}^{N} \left(\|y_u^n - y_u^{n-1}\|_{0,\Omega}^2 + k_n \|y_u^n\|_{1,\Omega}^2 \right) + \|y_u^N\|_{0,\Omega}^2
$$

$$
\leq C(1 + h^2 k^{-1}) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))}^2. \tag{3.29}
$$

Combining (3.28) and (3.29) we prove the case of $f \equiv 0$ and $y_0 \equiv 0$ with $Y_{hk}^n = y_u^n$.

For the case $u \equiv 0$, let $y_f \in L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ be the solution of following problem

$$
\begin{cases}
(\frac{\partial y_f}{\partial t}, w) + a(y_f, w) = (f, w), \ \ \forall w \in H_0^1(\Omega), \ \ t \in (0, T], \\
y_f = y_0 \ \text{in} \ \Omega; \ \ y_f = 0 \ \text{on} \ \Sigma.
\end{cases} \tag{3.30}
$$

Then we have

$$
\|y_f\|_{L^2(H^1(\Omega))} + \|\frac{\partial y_f}{\partial t}\|_{L^2(H^{-1}(\Omega))} \leq C \left(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} \right). \tag{3.31}
$$

Let $y_f^n \in V^h$, $n = 1, 2, \ldots, N$ be the solutions of following problems:

$$
\begin{cases}
\left(\frac{y_f^n - y_f^{n-1}}{k_n}, w_h \right) + a(y_f^n, w_h) = (P_k^n f, w_h), \forall w_h \in V_0^h, n = 1, \ldots, N, \\
y_f^0 = y_0^h, \ \text{in} \ \Omega; \ y_f^n = 0, \ n = 1, \ldots, N, \ \text{on} \ \Gamma.
\end{cases} \tag{3.32}
$$

Then y_f^n is the standard fully discrete approximation of y_f. Let $w_h = k_n(y_f^n - y_f^{n-1})$ in (3.32) we get

$$\left(y_f^n - y_f^{n-1}, y_f^n - y_f^{n-1}\right) + k_n a\left(y_f^n, y_f^n - y_f^{n-1}\right) = k_n(P_k^n f, y_f^n - y_f^{n-1}),$$

thus we have

$$\|y_f^n - y_f^{n-1}\|_{0,\Omega}^2 + k_n\|y_f^n\|_{1,\Omega}^2$$
$$= k_n a(y_f^n, y_f^{n-1}) + \int_{I_n}(f, y_f^n - y_f^{n-1})dt$$
$$\leq \frac{1}{2}k_n\|y_f^n\|_{1,\Omega}^2 + \frac{1}{2}k_n\|y_f^{n-1}\|_{1,\Omega}^2 + \|y_f^n - y_f^{n-1}\|_{0,\Omega}\int_{I_n}\|f\|_{0,\Omega}dt$$
$$\leq \frac{1}{2}k_n\|y_f^n\|_{1,\Omega}^2 + \frac{1}{2}k_n\|y_f^{n-1}\|_{1,\Omega}^2 + \frac{1}{2}k_n\int_{I_n}\|f\|_{0,\Omega}^2dt + \frac{1}{2}\|y_f^n - y_f^{n-1}\|_{0,\Omega}^2.$$

Summing the above equations over n from 1 to N we obtain

$$\sum_{n=1}^{N}\|y_f^n - y_f^{n-1}\|_{0,\Omega}^2 + k_N\|y_f^N\|_{1,\Omega}^2 \leq k_1\|\tilde{Q}_h y_0\|_{1,\Omega}^2 + k\sum_{n=1}^{N}\int_{I_n}\|f\|_{0,\Omega}^2dt$$
$$\leq k\|\tilde{Q}_h y_0\|_{1,\Omega}^2 + k\|f\|_{L^2(L^2(\Omega))}^2$$
$$\leq Ckh^{-2}\|y_0\|_{0,\Omega}^2 + k\|f\|_{L^2(L^2(\Omega))}^2, \quad (3.33)$$

where we used the estimate $\|\tilde{Q}_h y_0\|_{1,\Omega}^2 \leq Ch^{-2}\|\tilde{Q}_h y_0\|_{0,\Omega}^2 \leq Ch^{-2}\|y_0\|_{0,\Omega}^2$. This proves the case $u \equiv 0$ with $Y_{hk}^n = y_f^n$. □

We next consider the fully discrete approximation for above semi-discrete optimal control problems by using the dG(0) scheme in time. The fully discrete approximation scheme of (3.14) is to find $(Y_{hk}, U_{hk}) \in V_{hk} \times U_{ad}^{hk}$, such that

$$\min_{U_{hk}\in U_{ad}^{hk}, Y_{hk}\in V_{hk}} J_{hk}(Y_{hk}, U_{hk}) = \sum_{i=1}^{N}k_n\left\{\frac{1}{2}\int_{\Omega}(Y_{hk}^i - P_k^i y_d)^2dx + \frac{\alpha}{2}\int_{\Gamma}(U_{hk}^i)^2ds\right\} \quad (3.34)$$

subject to

$$\begin{cases} A(Y_{hk}, \Phi) = (f, \Phi)_{\Omega_T} + (y_0, \Phi_+^0), \ \forall \Phi \in V_{hk}^0, \\ Y_{hk} = Q(U_{hk}) \text{ on } \Sigma. \end{cases} \quad (3.35)$$

Here U_{ad}^{hk} is an appropriate approximation to U_{ad}. We set $U_{ad}^{hk} = V_{hk}(\Gamma) \cap U_{ad}$ for the full discretization of the control problem (1.1)–(1.2) and $U_{ad}^{hk} \equiv U_{ad}$ for its variational discretization.

It follows from standard arguments (see [24]) that the above control problem has a unique solution (Y_{hk}, U_{hk}), and that a pair $(Y_{hk}, U_{hk}) \in V_{hk} \times U_{ad}^{hk}$ is the solution of (3.34)–(3.35) if and only if there is a co-state $Z_{hk} \in V_{hk}^0$, such that the triplet $(Y_{hk}, Z_{hk}, U_{hk}) \in V_{hk} \times V_{hk}^0 \times U_{ad}^{hk}$ satisfies (3.35) and the following optimality conditions:

$$\begin{cases} A(\Phi, Z_{hk}) = \sum_{i=1}^{N}\int_{I_i}(Y_{hk}^i - y_d, \Phi)dt, \ \forall \Phi \in V_{hk}^0, \\ Z_{hk} = 0 \text{ on } \Sigma, \end{cases} \quad (3.36)$$

$$\int_{\Omega_T} (Y_{hk} - y_d)(Y_{hk}(v_{hk}) - Y_{hk})dxdt + \alpha \int_\Sigma U_{hk}(v_{hk} - U_{hk})dsdt \geq 0, \ \forall \ v_{hk} \in U_{ad}^{hk},$$

$$(3.37)$$

where $Y_{hk}(v_{hk})$ is the solution of problem (3.35) with Dirichlet boundary conditions $Q(v_{hk})$.

To derive an expression for the derivative of $J_{hk} : L^2(L^2(\Gamma)) \to \mathbb{R}$ analogous to the one of J given by formula (2.7) we have to define a discrete normal derivative $\partial_n^{hk} Z_{hk} \in V_{hk}(\Gamma)$ satisfying

$$\int_\Sigma \partial_n^{hk} Z_{hk} \Phi dsdt = A(\Phi, Z_{hk}) - \int_{\Omega_T} (Y_{hk} - y_d)\Phi dxdt, \ \forall \ \Phi \in V_{hk}. \quad (3.38)$$

It is easy to verify that the linear form

$$L(\Phi) := A(\Phi, Z_{hk}) - \int_{\Omega_T} (Y_{hk} - y_d)\Phi dxdt$$

is well defined on $V_{hk}(\Gamma)$ and is also continuous. Thus from Riesz representation theorem the equation (3.38) admits a unique solution $\partial_n^{hk} Z_{hk}$ in $V_{hk}(\Gamma)$. For an analogous reconstruction of discrete normal derivatives for elliptic Dirichlet boundary control problems we refer to [7]. With the help of (3.38) it is not difficult to show that

$$\begin{aligned}
0 \leq J_{hk}'(U_{hk})(v_{hk} - U_{hk}) \\
= \alpha \int_\Sigma U_{hk}(v_{hk} - U_{hk})dsdt + \int_{\Omega_T} (Y_{hk} - y_d)(Y_{hk}(v_{hk}) - Y_{hk})dxdt \\
= \alpha \int_\Sigma U_{hk}(v_{hk} - U_{hk})dsdt + A(Y_{hk}(v_{hk}) - Y_{hk}, Z_{hk}) \\
- \int_\Sigma \partial_n^{hk} Z_{hk}(Y_{hk}(v_{hk}) - Y_{hk})dsdt \\
= \alpha \int_\Sigma U_{hk}(v_{hk} - U_{hk})dsdt - \int_\Sigma \partial_n^{hk} Z_{hk} \cdot Q(v_{hk} - U_{hk})dsdt \\
= \int_\Sigma (\alpha U_{hk} - \partial_n^{hk} Z_{hk})(v_{hk} - U_{hk})dsdt
\end{aligned}$$

for $v_{hk} \in U_{ad}^{hk}$, which in turn implies

$$U_{hk} = P_{U_{ad}^{hk}} \left(\frac{1}{\alpha} \partial_n^{hk} Z_{hk} \right), \quad (3.39)$$

where $P_{U_{ad}^{hk}} : L^2(L^2(\Gamma)) \to U_{ad}^{hk}$ denotes the orthogonal projection in $L^2(L^2(\Gamma))$ onto U_{ad}^{hk}.

4 Error Estimates for the Optimal Control Problems

As a preliminary result we first estimate the error introduced by the discretization of the state equation, i.e., the error between the solutions of problems (2.2) and (3.18).

Theorem 3 *Let $y \in L^2(L^2(\Omega))$ and $Y_{hk}(u) \in V_{hk}$ with $Y_{hk}(u)|_\Sigma = Q(u)$ be the solutions of problems (2.2) and (3.18), respectively. Then for $u \in L^2(L^2(\Gamma))$ we have*

$$\begin{aligned}
\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))} \ &\leq C \left(h^{\frac{1}{2}} + k^{\frac{1}{4}} + h^{\frac{3}{2}}k^{-\frac{1}{2}} + h^{-\frac{1}{2}}k^{\frac{1}{2}} + h^{\frac{5}{2}}k^{-1} \right) \\
&\quad \left(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} + \|u\|_{L^2(L^2(\Gamma))} \right)
\end{aligned} \quad (4.1)$$

and for $u \in L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))$ we have

$$\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))} \leq C(h + k^{\frac{1}{2}} + h^2 k^{-\frac{1}{2}} + h^3 k^{-1} + h^{-1}k)$$

$$\left(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} + \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))} + \|u\|_{H^{\frac{1}{4}}(L^2(\Gamma))} \right). \tag{4.2}$$

Proof In view of the linearity of the problem it is sufficient to consider the problems with either $f \equiv 0$, $y_0 \equiv 0$ or $u \equiv 0$.

Let us first assume that $f \equiv 0$, $y_0 \equiv 0$ and $u \in L^2(L^2(\Gamma))$. We first note that according to [12] $y \in L^2(H^{\frac{1}{2}}(\Omega))$ holds. Let $w \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega))$ be the solution of problem (2.12) with right hand side $g = y - Y_{hk}(u)$. Since $w(T) = 0$, we from (2.2) and (3.18) deduce that

$$\begin{aligned}
\|y - Y_{hk}(u)\|^2_{L^2(L^2(\Omega))} &= \int_{\Omega_T} (-w_t - \Delta w)(y - Y_{hk}(u))dxdt \\
&= \int_{\Omega_T} (-w_t y - \Delta w y)dxdt + \int_{\Sigma} Q(u)\frac{\partial w}{\partial n}dsdt \\
&\quad + \int_{\Omega_T} (w_t Y_{hk}(u) - \nabla w \nabla Y_{hk}(u))dxdt \\
&= \int_{\Sigma} (Q(u) - u)\frac{\partial w}{\partial n}dsdt \\
&\quad + \int_{\Omega_T} (w_t Y_{hk}(u) - \nabla w \nabla Y_{hk}(u))dxdt \\
&:= E_1 + E_2. \tag{4.3}
\end{aligned}$$

We treat E_1 by exploiting the properties of P_k and Q_h:

$$\begin{aligned}
E_1 &= \int_{\Sigma} (Q(u) - u)\frac{\partial w}{\partial n}dsdt \\
&= \int_0^T \langle (P_k - I)u, \frac{\partial w}{\partial n}\rangle dt + \int_0^T \langle (Q_h - I)P_k u, \frac{\partial w}{\partial n}\rangle dt \\
&= \int_0^T \langle (P_k - I)u, \frac{\partial}{\partial n}(I - P_k)w\rangle dt + \int_0^T \langle (Q_h - I)P_k u, (I - Q_h)\frac{\partial w}{\partial n}\rangle dt.
\end{aligned}$$

From the Young's inequality, the trace inequality and a norm interpolation inequality we derive (see, e.g. [12])

$$\begin{aligned}
&\left\| \frac{\partial}{\partial n}(I - P_k)w \right\|^2_{L^2(L^2(\Gamma))} \\
&\leq \frac{C}{\epsilon}\|(I - P_k)w\|^2_{L^2(H^2(\Omega))} + \epsilon\|(I - P_k)w\|^2_{L^2(H^1(\Omega))} \\
&\leq \frac{C}{\epsilon}\|(I - P_k)w\|^2_{L^2(H^2(\Omega))} + \epsilon\|(I - P_k)w\|_{L^2(H^2(\Omega))}\|(I - P_k)w\|_{L^2(L^2(\Omega))} \\
&\leq \frac{2C}{\epsilon}\|(I - P_k)w\|^2_{L^2(H^2(\Omega))} + \epsilon^3\|(I - P_k)w\|^2_{L^2(L^2(\Omega))}.
\end{aligned}$$

Setting $\epsilon = k^{-\frac{1}{2}}$ and using the approximation property (3.17) of P_k gives

$$\left\| \frac{\partial}{\partial n}(I - P_k)w \right\|^2_{L^2(L^2(\Gamma))} \leq Ck^{1/2}\left(\|w\|^2_{L^2(H^2(\Omega))} + \|w_t\|^2_{L^2(L^2(\Omega))} \right).$$

We also have

$$\left\| (I - Q_h)\frac{\partial w}{\partial n} \right\|_{L^2(L^2(\Gamma))} \leq Ch^{1/2}\|w\|_{L^2(H^2(\Omega))}.$$

Using the Cauchy-Schwarz inequality and stability results for Q_h and P_k we estimate

$$
\begin{aligned}
|E_1| &\leq Ck^{1/4}\|u\|_{L^2(L^2(\Gamma))}\big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))}\big) \\
&\quad + Ch^{1/2}\|P_k u\|_{L^2(L^2(\Gamma))}\|w\|_{L^2(H^2(\Omega))} \\
&\leq C(h^{1/2} + k^{1/4})\|u\|_{L^2(L^2(\Gamma))}\|g\|_{L^2(L^2(\Omega))}.
\end{aligned}
\tag{4.4}
$$

Next we estimate E_2. Considering (3.18) and $w^N = w(\cdot, T) = 0$ we calculate

$$
\begin{aligned}
E_2 &= \int_{\Omega_T} \big(w_t Y_{hk}(u) - \nabla w \nabla Y_{hk}(u)\big)dxdt \\
&= \sum_{n=1}^{N} \big(Y_{hk}^n(u), w^n - w^{n-1}\big) - k_n(\nabla Y_{hk}^n(u), \nabla P_k^n w) \\
&= -\sum_{n=1}^{N} \big(Y_{hk}^n(u) - Y_{hk}^{n-1}(u), w^{n-1}\big) + k_n(\nabla Y_{hk}^n(u), \nabla P_k^n w) \\
&= -\sum_{n=1}^{N} \Big((Y_{hk}^n(u) - Y_{hk}^{n-1}(u), w^{n-1} - R_h P_k^n w) \\
&\quad + k_n\big(\nabla Y_{hk}^n(u), \nabla(P_k^n w - R_h P_k^n w)\big)\Big).
\end{aligned}
\tag{4.5}
$$

By the Cauchy-Schwarz inequality we have

$$|E_2| \leq F_1 \cdot F_2,$$

where

$$F_1 = \left(\sum_{n=1}^{N} \big(\|Y_{hk}^n(u) - Y_{hk}^{n-1}(u)\|_{0,\Omega}^2 + k_n(\nabla Y_{hk}^n(u), \nabla Y_{hk}^n(u))\big)\right)^{\frac{1}{2}}$$

and

$$F_2 = \left(\sum_{n=1}^{N} \big(\|w^{n-1} - R_h P_k^n w\|_{0,\Omega}^2 + k_n(\nabla(I - R_h)P_k^n w, \nabla(I - R_h)P_k^n w)\big)\right)^{\frac{1}{2}}.$$

In view of the stability result (3.20) of Lemma 3 we have

$$|F_1| \leq C(h^{-\frac{1}{2}} + h^{\frac{1}{2}}k^{-\frac{1}{2}})\|u\|_{L^2(L^2(\Gamma))}.$$
$$\tag{4.6}$$

It remains to estimate F_2. To begin with we note that

$$
\begin{aligned}
\|w^{n-1} - R_h P_k^n w\|_{0,\Omega} &\leq \|w^{n-1} - P_k^n w\|_{0,\Omega} + \|(I - R_h)P_k^n w\|_{0,\Omega} \\
&\leq \|w^{n-1} - P_k^n w\|_{0,\Omega} + Ch^2\|P_k^n w\|_{2,\Omega},
\end{aligned}
\tag{4.7}
$$

and

$$(\nabla(I - R_h)P_k^n w, \nabla(I - R_h)P_k^n w) \leq Ch^2\|P_k^n w\|_{2,\Omega}^2. \tag{4.8}$$

It is straightforward to show that

$$\|w^{n-1} - P_k^n w\|_{0,\Omega} \le k_n^{1/2} \|w_t\|_{L^2(I_n, L^2(\Omega))} \tag{4.9}$$

and

$$\|P_k^n w\|_{2,\Omega} \le k_n^{-1/2} \|w\|_{L^2(I_n, H^2(\Omega))}. \tag{4.10}$$

Combining (4.7)–(4.10) we get

$$
\begin{aligned}
F_2 &\le C \Big(\sum_{n=1}^{N} \big((h^4 + k_n h^2) \|P_k^n w\|_{2,\Omega}^2 + k_n \|w_t\|_{L^2(I_n, L^2(\Omega))}^2 \big) \Big)^{\frac{1}{2}} \\
&\le C(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}}) \Big(\sum_{n=1}^{N} \big(\|w\|_{L^2(I_n, H^2(\Omega))}^2 + \|w_t\|_{L^2(I_n, L^2(\Omega))}^2 \big) \Big)^{\frac{1}{2}} \\
&\le C \Big(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}} \Big) \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big).
\end{aligned}
\tag{4.11}
$$

Using the stability estimate (3.20) of Lemma 3 we conclude

$$
\begin{aligned}
|E_2| &\le C(h^{-\frac{1}{2}} + h^{\frac{1}{2}} k^{-\frac{1}{2}})(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}}) \\
&\quad \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big) \|u\|_{L^2(L^2(\Gamma))} \\
&\le C(h^{\frac{1}{2}} + h^{\frac{3}{2}} k^{-\frac{1}{2}} + h^{-\frac{1}{2}} k^{\frac{1}{2}} + h^{\frac{3}{2}} k^{-1}) \\
&\quad \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big) \|u\|_{L^2(L^2(\Gamma))}.
\end{aligned}
\tag{4.12}
$$

From the estimates (4.3)–(4.12) we conclude the desired result in the case $f \equiv 0$, $y_0 \equiv 0$ and $u \in L^2(L^2(\Gamma))$ that

$$
\begin{aligned}
&\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))} \\
&\le C(h^{\frac{1}{2}} + k^{\frac{1}{4}} + h^{\frac{3}{2}} k^{-\frac{1}{2}} + h^{-\frac{1}{2}} k^{\frac{1}{2}} + h^{\frac{5}{2}} k^{-1}) \|u\|_{L^2(L^2(\Gamma))}.
\end{aligned}
\tag{4.13}
$$

If $u \in L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))$ we can estimate E_1 as

$$
\begin{aligned}
|E_1| &\le C k^{1/2} \|u\|_{H^{\frac{1}{4}}(L^2(\Gamma))} \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big) \\
&\quad + C h \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))} \|w\|_{L^2(H^2(\Omega))} \\
&\le C(h + k^{1/2}) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))} \|g\|_{L^2(L^2(\Omega))}.
\end{aligned}
\tag{4.14}
$$

Combining the estimate (4.11) of F_2 and the stability result (3.21) in Lemma 3 we are led to

$$
\begin{aligned}
|E_2| &\le C(1 + h k^{-\frac{1}{2}})(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}}) \\
&\quad \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))} \\
&\le C(h + k^{\frac{1}{2}} + h^2 k^{-\frac{1}{2}} + h^3 k^{-1}) \\
&\quad \big(\|w\|_{L^2(H^2(\Omega))} + \|w_t\|_{L^2(L^2(\Omega))} \big) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))}.
\end{aligned}
\tag{4.15}
$$

This combining with (4.3) and (4.14) gives

$$
\begin{aligned}
&\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))} \\
&\le C(h + k^{\frac{1}{2}} + h^2 k^{-\frac{1}{2}} + h^3 k^{-1}) \|u\|_{L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))}.
\end{aligned}
\tag{4.16}
$$

If $u \equiv 0$, $f \in L^2(L^2(\Omega))$ and $y_0 \in L^2(\Omega)$ we have $y \in L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ (see, e.g. [25]). Then similar to the above error estimate and using the stability estimate (3.22) of Lemma 3, it is straightforward to prove that (see also [12])

$$\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))} \le C(h + h^{-1}k)(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega}). \tag{4.17}$$

Actually, by using the duality argument it follows from (4.3) and (4.5) that

$$
\begin{aligned}
& \|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))}^2 \\
& = \int_{\Omega_T} (-w_t - \Delta w)(y - Y_{hk}(u)) dx dt \\
& = \int_{\Omega_T} \left(w_t Y_{hk}(u) - \nabla w \nabla Y_{hk}(u) \right) dx dt + (y_0 - Y_{hk}^0(u), w(\cdot, 0)) \\
& = -\sum_{n=1}^{N} \left(\left(Y_{hk}^n(u) - Y_{hk}^{n-1}(u), w^{n-1} - R_h P_k^n w \right) \right) \\
& \quad + (y_0 - Y_{hk}^0(u), w(\cdot, 0))
\end{aligned}
\tag{4.18}
$$

where we used the fact that the second term in (4.5) vanishes because of $Y_{hk}^n(u) \in V_0^h$. Note that

$$
\begin{aligned}
(y_0 - Y_{hk}^0(u), w(\cdot, 0)) & \le \|y_0 - \tilde{Q}_h y_0\|_{-1,\Omega} \|w(\cdot, 0)\|_{1,\Omega} \\
& \le Ch\|y_0\|_{0,\Omega} \|g\|_{L^2(L^2(\Omega))}.
\end{aligned}
\tag{4.19}
$$

It follows from (3.22), (4.7), (4.9), (4.10) that

$$
\begin{aligned}
& \left| -\sum_{n=1}^{N} \left(\left(Y_{hk}^n(u) - Y_{hk}^{n-1}(u), w^{n-1} - R_h P_k^n w \right) \right) \right| \\
& \le \left(\sum_{n=1}^{N} \|w^{n-1} - R_h P_k^n w\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \left(\sum_{n=1}^{N} \|Y_{hk}^n(u) - Y_{hk}^{n-1}(u)\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \\
& \le C \left(\sum_{n=1}^{N} \left(h^4 k_n^{-1} \|w\|_{L^2(I_n, H^2(\Omega))}^2 + k_n \|w_t\|_{L^2(I_n, L^2(\Omega))}^2 \right) \right)^{\frac{1}{2}} \\
& \quad \left(h^{-2} k \|y_0\|_{0,\Omega}^2 + k \|f\|_{L^2(L^2(\Omega))}^2 \right)^{\frac{1}{2}} \\
& \le C(h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}}) \left(\sum_{n=1}^{N} (\|w\|_{L^2(I_n, H^2(\Omega))}^2 + \|w_t\|_{L^2(I_n, L^2(\Omega))}^2) \right)^{\frac{1}{2}} \\
& \quad \left(h^{-2} k \|y_0\|_{0,\Omega}^2 + k \|f\|_{L^2(L^2(\Omega))}^2 \right)^{\frac{1}{2}} \\
& \le C(h + h^{-1}k) \|g\|_{L^2(L^2(\Omega))} \left(\|y_0\|_{0,\Omega} + \|f\|_{L^2(L^2(\Omega))} \right),
\end{aligned}
$$

this together with (4.18), (4.19) gives (4.17). Combining both cases we complete the proof. $\square$

Now we are in a position to derive our main result of this section: the a priori error estimates for optimal control problems. At first we consider the fully discrete case, i.e., $U_{ad}^{hk} = U_{ad} \cap V_{hk}(\Gamma)$.

Theorem 4 *Let* $(y, u, z) \in L^2(L^2(\Omega)) \times L^2(L^2(\Gamma)) \times L^2(H^2(\Omega)) \cap H^1(L^2(\Omega))$ *and* $(Y_{hk}, U_{hk}, Z_{hk}) \in V_{hk} \times U_{ad}^{hk} \times V_{hk}^0$ *be the solutions of problem* (2.5)–(2.8) *and* (3.34)–(3.35) *with* $U_{ad}^{hk} = U_{ad} \cap V_{hk}(\Gamma)$, *respectively. Then we have the a priori error estimate*

$$\|u - U_{hk}\|_{L^2(L^2(\Gamma))} + \|y - Y_{hk}\|_{L^2(L^2(\Omega))}$$
$$\leq C(h^{\frac{1}{2}} + k^{\frac{1}{4}} + h^{\frac{3}{2}}k^{-\frac{1}{2}} + h^{-\frac{1}{2}}k^{\frac{1}{2}} + h^{\frac{5}{2}}k^{-1})$$
$$\left(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} + \|y_d\|_{L^2(L^2(\Omega))} + \|u\|_{L^2(L^2(\Gamma))}\right) \tag{4.20}$$

with a constant $C > 0$ *independent of* h *and* k.

Proof Let us recall the continuous and discrete optimality conditions

$$\int_{\Omega_T} (y - y_d)(y(v) - y)dxdt + \alpha \int_\Sigma u(v - u)dsdt \geq 0, \quad \forall\, v \in U_{ad} \tag{4.21}$$

and

$$\int_{\Omega_T} (Y_{hk} - y_d)(Y_{hk}(v_{hk}) - Y_{hk})dxdt + \alpha \int_\Sigma U_{hk}(v_{hk} - U_{hk})dsdt$$
$$\geq 0, \quad \forall\, v_{hk} \in U_{ad}^{hk}. \tag{4.22}$$

Setting $v = U_{hk} \in U_{ad}$ and $v_{hk} = Q(u) \in U_{ad}^{hk}$ we have

$$\alpha \|u - U_{hk}\|_{L^2(L^2(\Gamma))}^2 = \alpha \int_\Sigma (u - U_{hk})^2 dsdt$$
$$= \alpha \int_\Sigma u(u - U_{hk})dsdt - \alpha \int_\Sigma U_{hk}(u - U_{hk})dsdt$$
$$\leq \int_{\Omega_T} (y - y_d)(y(U_{hk}) - y)dxdt - \alpha \int_\Sigma U_{hk}(u - Q(u))dsdt$$
$$\quad -\alpha \int_\Sigma U_{hk}(Q(u) - U_{hk})dsdt$$
$$\leq \int_{\Omega_T} (y - y_d)(y(U_{hk}) - y)dxdt + \int_{\Omega_T} (Y_{hk} - y_d)(Y_{hk}(Qu) - Y_{hk})dxdt$$
$$\quad -\alpha \int_\Sigma U_{hk}(u - Q(u))dsdt, \tag{4.23}$$

where $y(U_{hk}) \in L^2(L^2(\Omega))$ with $y(U_{hk})|_\Sigma = U_{hk}$ solves

$$\int_{\Omega_T} y(U_{hk})(-v_t - \Delta v)dxdt = -\int_\Sigma U_{hk}\partial_n v dsdt + \int_{\Omega_T} fv dxdt + \int_\Omega y_0 v(\cdot, 0)dx$$
$$\forall\, v \in L^2(H^2(\Omega) \cap H_0^1(\Omega)) \cap H^1(L^2(\Omega)) \tag{4.24}$$

with $v(\cdot, T) = 0$, and $Y_{hk}(Qu) \in V_{hk}$ solves

$$\begin{cases} A(Y_{hk}(Qu), \Phi) = (f, \Phi)_{\Omega_T} + (y_0, \Phi_+^0), & \forall\, \Phi \in V_{hk}^0, \\ Y_{hk}(Qu) = Q(u) & \text{on } \Sigma. \end{cases} \tag{4.25}$$

With Young's inequality we deduce

$$
\begin{aligned}
&\int_{\Omega_T} (y - y_d)(y(U_{hk}) - y)dxdt + \int_{\Omega_T} (Y_{hk} - y_d)(Y_{hk}(Qu) - Y_{hk})dxdt \\
&= (y - y_d, y(U_{hk}) - y)_{\Omega_T} + (Y_{hk} - y_d, Y_{hk}(Qu) - Y_{hk})_{\Omega_T} \\
&= (y - y_d, y(U_{hk}) - y)_{\Omega_T} + (Y_{hk} - y, y - Y_{hk})_{\Omega_T} \\
&\quad + (Y_{hk} - y, Y_{hk}(Qu) - y)_{\Omega_T} + (y - y_d, Y_{hk}(Qu) - Y_{hk})_{\Omega_T} \\
&= -\|y - Y_{hk}\|^2_{L^2(L^2(\Omega))} + (Y_{hk} - y, Y_{hk}(Qu) - y)_{\Omega_T} \\
&\quad + \big(y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(Qu))\big)_{\Omega_T} \\
&\leq -\|y - Y_{hk}\|^2_{L^2(L^2(\Omega))} + \big(y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(Qu))\big)_{\Omega_T} \\
&\quad + \sigma\|y - Y_{hk}\|^2_{L^2(L^2(\Omega))} + C(\sigma)\|Y_{hk}(Qu) - y\|^2_{L^2(L^2(\Omega))}.
\end{aligned}
\tag{4.26}
$$

Taking $\sigma > 0$ small enough, we from (4.23)–(4.26) obtain

$$
\begin{aligned}
&\alpha\|u - U_{hk}\|^2_{L^2(L^2(\Gamma))} + \|y - Y_{hk}\|^2_{L^2(L^2(\Omega))} \\
&\leq -\alpha \int_{\Sigma} U_{hk}(u - Q(u))dsdt + C\|Y_{hk}(Qu) - y\|^2_{L^2(L^2(\Omega))} \\
&\quad + (y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(Qu)))_{\Omega_T} \\
&:= I_1 + I_2 + I_3.
\end{aligned}
\tag{4.27}
$$

Note that from the standard error estimates for the L^2-projection and the regularity of u we have

$$
\|Q(u) - u\|^2_{L^2(L^2(\Gamma))} \leq C(h + k^{\frac{1}{2}})\big(\|u\|^2_{L^2(H^{\frac{1}{2}}(\Gamma))} + \|u\|^2_{H^{\frac{1}{4}}(L^2(\Gamma))}\big).
\tag{4.28}
$$

Thus we are led to

$$
\begin{aligned}
|I_1| &= \Big| -\alpha \int_{\Sigma} U_{hk}(u - Q(u))dsdt \Big| \\
&= \Big| \alpha \int_{\Sigma} u(Q(u) - u)dsdt + \alpha \int_{\Sigma} (U_{hk} - u)(Q(u) - u)dsdt \Big| \\
&= \Big| \alpha \int_{\Sigma} (u - Q(u))(Q(u) - u)dsdt + \alpha \int_{\Sigma} (U_{hk} - u)(Q(u) - u)dsdt \Big| \\
&\leq \sigma\|u - U_{hk}\|^2_{L^2(L^2(\Gamma))} + C\|u - Q(u)\|^2_{L^2(L^2(\Gamma))} \\
&\leq \sigma\|u - U_{hk}\|^2_{L^2(L^2(\Gamma))} + C(h + k^{\frac{1}{2}})\left(\|u\|^2_{L^2(H^{\frac{1}{2}}(\Gamma))} + \|u\|^2_{H^{\frac{1}{4}}(L^2(\Gamma))}\right).
\end{aligned}
\tag{4.29}
$$

Since $Y_{hk}(Qu)$ is the fully discrete finite element approximation of y, the error estimate (4.2) of Theorem 3 gives

$$
\begin{aligned}
I_2 &= \|Y_{hk}(Qu) - y\|^2_{L^2(L^2(\Omega))} \leq C(h^2 + k + h^4 k^{-1} + h^6 k^{-2} + h^{-2}k^2) \\
&\quad \left(\|f\|^2_{L^2(L^2(\Omega))} + \|y_0\|^2_{0,\Omega} + \|u\|^2_{L^2(H^{\frac{1}{2}}(\Gamma)) \cap H^{\frac{1}{4}}(L^2(\Gamma))}\right).
\end{aligned}
\tag{4.30}
$$

Then it remains to estimate I_3. From (2.2), (2.6), (4.24) and (4.25) we have

$$
\begin{aligned}
I_3 &= (y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(Qu)))_{\Omega_T} \\
&= \int_{\Omega_T} (y(U_{hk}) - y)\left(-\frac{\partial z}{\partial t} - \Delta z\right) dxdt \\
&\quad - \int_{\Omega_T} (Y_{hk} - Y_{hk}(Qu))\left(-\frac{\partial z}{\partial t} - \Delta z\right) dxdt \\
&= \int_{\Omega_T} \left(-(y(U_{hk}) - y)z_t - (y(U_{hk}) - y)\Delta z\right) dxdt \\
&\quad + \int_{\Omega_T} \left(z_t(Y_{hk} - Y_{hk}(Qu)) - \nabla(Y_{hk} - Y_{hk}(Qu))\nabla z\right) dxdt \\
&\quad + \int_{\Sigma} (U_{hk} - Q(u))\partial_n z \, dsdt \\
&= -\int_{\Sigma} (U_{hk} - u)\partial_n z \, dsdt + \int_{\Sigma} (U_{hk} - Q(u))\partial_n z \, dsdt \\
&\quad + \int_{\Omega_T} \left(z_t(Y_{hk} - Y_{hk}(Qu)) - \nabla(Y_{hk} - Y_{hk}(Qu))\nabla z\right) dxdt \\
&= H_1 + H_2 + H_3.
\end{aligned}
\tag{4.31}
$$

Note that

$$
\begin{aligned}
H_1 + H_2 &= -\int_{\Sigma} (U_{hk} - u)\partial_n z \, dsdt + \int_{\Sigma} (U_{hk} - Q(u))\partial_n z \, dsdt \\
&= \int_{\Sigma} (u - Qu)\partial_n z \, dsdt \\
&= \int_{\Sigma} (u - Qu)\left(\partial_n z - Q(\partial_n z)\right) dsdt.
\end{aligned}
$$

It is straightforward to estimate

$$
\begin{aligned}
|H_1 + H_2| &\le C(h + k^{\frac{1}{2}})\left(\|z\|_{L^2(H^2(\Omega))} + \|z_t\|_{L^2(L^2(\Omega))}\right) \\
&\quad \left(\|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))} + \|u\|_{H^{\frac{1}{4}}(L^2(\Gamma))}\right).
\end{aligned}
\tag{4.32}
$$

Define $E_{hk} := Y_{hk} - Y_{hk}(Qu)$. Using the proof technique of Theorem 3 we from (3.35), (4.25) obtain

$$
\begin{aligned}
H_3 &= \int_{\Omega_T} \left(z_t E_{hk} - \nabla E_{hk}\nabla z\right) dxdt \\
&= \sum_{n=1}^{N} \left((E_{hk}^n, z^n - z^{n-1}) - k_n(\nabla E_{hk}^n, \nabla P_k^n z)\right) \\
&= -\sum_{n=1}^{N} \left((E_{hk}^n - E_{hk}^{n-1}, z^{n-1}) + k_n(\nabla E_{hk}^n, \nabla P_k^n z)\right) \\
&= -\sum_{n=1}^{N} (E_{hk}^n - E_{hk}^{n-1}, z^{n-1} - R_h P_k^n z) + k_n(\nabla E_{hk}^n, \nabla(P_k^n z - R_h P_k^n z)).
\end{aligned}
$$

With the help of projection error estimate and proceeding as in the estimate of (4.5) we have

$$|H_3| \le C(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}})\left(\|z\|_{L^2(H^2(\Omega))} + \|z\|_{H^1(L^2(\Omega))}\right) \cdot$$
$$\left(\sum_{n=1}^N \|E_{hk}^n - E_{hk}^{n-1}\|_{0,\Omega}^2 + k_n\|E_{hk}^n\|_{1,\Omega}^2\right)^{\frac{1}{2}}.$$

From Lemma 3 we conclude

$$|H_3| \le C(h + h^2 k^{-\frac{1}{2}} + k^{\frac{1}{2}})\left(\|z\|_{L^2(H^2(\Omega))} + \|z_t\|_{L^2(L^2(\Omega))}\right) \cdot$$
$$(h^{-\frac{1}{2}} + h^{\frac{1}{2}}k^{-\frac{1}{2}})\|Q(u) - U_{hk}\|_{L^2(L^2(\Gamma))}$$
$$\le C(h^{\frac{1}{2}} + h^{\frac{3}{2}}k^{-\frac{1}{2}} + h^{-\frac{1}{2}}k^{\frac{1}{2}} + h^{\frac{5}{2}}k^{-1})$$
$$\left(\|z\|_{L^2(H^2(\Omega))} + \|z_t\|_{L^2(L^2(\Omega))}\right)\|u - U_{hk}\|_{L^2(L^2(\Gamma))}. \tag{4.33}$$

Since the projection operator $P_{U_{ad}}$ is continuous on $L^2(H^{\frac{1}{2}}(\Gamma))$ and $H^{\frac{1}{4}}(L^2(\Gamma))$ (see, e.g. [23, Lm. 3.3]), we have from Theorem 2 that

$$\|u\|_{L^2(H^{\frac{1}{2}}(\Gamma))} + \|u\|_{H^{\frac{1}{4}}(L^2(\Gamma))} \le C(\|z\|_{L^2(H^2(\Omega))} + \|z_t\|_{L^2(L^2(\Omega))}). \tag{4.34}$$

From standard regularity result for parabolic equation [25] and Lemma 1 we have

$$\|z\|_{L^2(H^2(\Omega))} + \|z_t\|_{L^2(L^2(\Omega))} \le C(\|y\|_{L^2(L^2(\Omega))} + \|y_d\|_{L^2(L^2(\Omega))})$$
$$\le C(\|f\|_{L^2(L^2(\Omega))} + \|y_d\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} + \|u\|_{L^2(L^2(\Gamma))}). \tag{4.35}$$

Combining above results and using the Cauchy-Schwarz inequality and Young's inequality completes the proof. □

If we use variational discretization concept introduced in [19], i.e., $U_{ad}^{hk} = U_{ad}$, we can prove the following error estimates in a similar way.

Theorem 5 *Let $(y, u, z) \in L^2(L^2(\Omega)) \times L^2(L^2(\Gamma)) \times L^2(H^2(\Omega)) \cap H^1(L^2(\Omega))$ and $(Y_{hk}, U_{hk}, Z_{hk}) \in V_{hk} \times U_{ad} \times V_{hk}^0$ be the solutions of problem (2.5)–(2.8), and (3.34)–(3.35) with $U_{ad}^{hk} = U_{ad}$, respectively. Then we have following a priori error estimate*

$$\|u - U_{hk}\|_{L^2(L^2(\Gamma))} + \|y - Y_{hk}\|_{L^2(L^2(\Omega))}$$
$$\le C(h^{\frac{1}{2}} + k^{\frac{1}{4}} + h^{\frac{3}{2}}k^{-\frac{1}{2}} + h^{-\frac{1}{2}}k^{\frac{1}{2}} + h^{\frac{5}{2}}k^{-1})$$
$$\left(\|f\|_{L^2(L^2(\Omega))} + \|y_0\|_{0,\Omega} + \|y_d\|_{L^2(L^2(\Omega))} + \|u\|_{L^2(L^2(\Gamma))}\right) \tag{4.36}$$

with a constant $C > 0$ independent of h and k.

Proof In the proof of Theorem 4 it suffices to set $v = U_{hk}$ in (4.21) and $v_{hk} = u$ in (4.22) and add the corresponding inequalities. This directly gives

$$\alpha\|u - U_{hk}\|_{L^2(L^2(\Gamma))}^2 \le (y - y_d, y(U_{hk}) - y)_{\Omega_T} + (Y_{hk} - y_d, Y_{hk}(u) - Y_{hk})_{\Omega_T}$$
$$\le -\|y - Y_{hk}\|_{L^2(L^2(\Omega))}^2 + (y - Y_{hk}, y - Y_{hk}(u))$$
$$+ (y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(u)))_{\Omega_T}.$$

Thus

$$\alpha\|u - U_{hk}\|_{L^2(L^2(\Gamma))}^2 + \|y - Y_{hk}\|_{L^2(L^2(\Omega))}^2$$
$$\le C\|y - Y_{hk}(u)\|_{L^2(L^2(\Omega))}^2 + (y - y_d, y(U_{hk}) - y - (Y_{hk} - Y_{hk}(u)))_{\Omega_T}.$$

The rest of proof is along the lines of the estimation of the terms I_2 and I_3 in the proof of Theorem 4.   □

As the final result, we can conclude from Theorems 4 and 5 the explicit convergence rate with respect to h and k for the fully discrete finite element approximation of the optimal control problems under the assumption $k = O(h^2)$.

Corollary 1 *Assume that the spatial mesh size h and time step k satisfy the coupling $k = O(h^2)$. Then we have the following a priori error estimate*

$$\|u - U_{hk}\|_{L^2(L^2(\Gamma))} + \|y - Y_{hk}\|_{L^2(L^2(\Omega))} \leq Ch^{\frac{1}{2}} \tag{4.37}$$

for both full control discretization and variational control discretization with a constant $C > 0$ independent of h and k.

Remark 1 The error estimates we obtained in Theorems 4, 5 and Corollary 1 reflect the worst cases we can expect for parabolic Dirichlet boundary control problems defined on convex polygonal domains.

Since the regularity of parabolic equations depends on the maximum interior angle of the domain, the state admits the improved regularity $y \in L^2(W^{1,p}(\Omega))$ for $2 \leq p < p_*$ with $p_* = \frac{2\omega}{2\omega - \pi}$ depending on the maximum interior angle ω of the polygonal domain Ω and also the data (see [29] for more details). Moreover, for problems defined on curved domains with smooth boundary, we have higher regularity for the optimal control problems as stated in Theorem 3.4 of [23]. Improved space regularity leads to better approximation properties of the state and thus to better convergence rates for space discretization, as is reported in our numerical results. For the elliptic case with polygonal boundaries we refer to [7] where an approximation order for the controls of $O(h^{1 - \frac{1}{p}})$ is derived for some $2 < p \leq p_*$ with p_* depending on the data and the maximum interior angle of the domain. For the error estimates of elliptic Dirichlet boundary control problems defined on curved domains we refer to [8], [10] and [16] for more details.

Note that $y \in L^2(H^1(\Omega)) \cap H^{\frac{1}{2}}(L^2(\Omega))$, if in addition $y_d \in L^2(H^1(\Omega)) \cap H^{\frac{1}{2}}(L^2(\Omega))$ we may derive that $z \in L^2(H^3(\Omega)) \cap H^{\frac{3}{2}}(L^2(\Omega))$ and $\partial_n z \in H^{\frac{3}{4}}(L^2(\Gamma))$ under appropriate assumption on the domain Ω, and thus $u \in H^{\frac{1}{2}}(L^2(\Gamma))$ (see the proof of Theorem 2 and [23]). This improved time regularity may deliver higher order convergence for time discretization, compared to the estimates (4.29) and (4.32). This may explain the higher order convergence for the time discretization observed in the numerical experiments.

5 Numerical Experiments

In this section we will carry out some numerical experiments to support our theoretical findings. We consider the optimal control problems (1.1)–(1.2) of tracking type with control set U_{ad} defined as follows

$$U_{ad} := \{u \in L^2(0, T; L^2(\Gamma)) : 0 \leq u(x, t) \leq 1, \text{ a.a. } (t, x) \in [0, T] \times \Gamma\}.$$

In the numerical experiments we illustrate the convergence orders with respect to the spatial and time discretizations separately by setting h and k small enough respectively, although we derived a priori error estimate with coupling $k = O(h^2)$. The numerical tests indicate that such a coupling of k and h seems not to be needed. We expect that an according analysis is possible with adapting the techniques of [30] and [31] to the present setting.

Table 1 Error of control u, state y and adjoint state z for Example 1 with fixed time step $N = 8192$

Dof	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
31	0.028782014118	\	0.008430762122	\	0.001829667132	\
105	0.010849673318	1.4075	0.002136393772	1.9805	0.000467553402	1.9684
385	0.004042069175	1.4245	0.000553164133	1.9494	0.000117011926	1.9985
1473	0.001462083061	1.4671	0.000139259530	1.9899	0.000029815477	1.9725
5761	0.000522887919	1.4835	0.000034996583	1.9925	0.000010623328	1.4888

Although we do not consider problems defined on curved domains in our numerical analysis, we include some numerical examples on both polygonal and curved domains using full discretization and variational discretisation of the control. For the numerical approximations of Dirichlet boundary control problems defined on curved domains we refer to [8, 10] and [16]. We use $\| \cdot \|_{L^2}$ to denote the $L^2(L^2(\Gamma))$-norm error for the optimal control u and the $L^2(L^2(\Omega))$-norm errors for the state y and adjoint state z.

Example 1 The first example is a unconstrained problem defined on the unit square $\Omega = [0, 1] \times [0, 1]$, $T = 1$. The data is chosen as

$$f = -\frac{4.0}{\alpha} \sin(\pi t) - \frac{\pi}{\alpha}(x_1(1 - x_1) + x_2(1 - x_2)) \cos(\pi t),$$
$$y_d = -(2 + 1.0/\alpha)(x_1(1 - x_1) + x_2(1 - x_2)) \sin(\pi t) + \pi x_1 x_2 (1 - x_1)(1 - x_2) \cos(\pi t),$$

with $\alpha = 1$, so that the optimal solution is given by

$$u = -\frac{1}{\alpha}(x_1(1 - x_1) + x_2(1 - x_2)) \sin(\pi t),$$
$$y = -\frac{1}{\alpha}(x_1(1 - x_1) + x_2(1 - x_2)) \sin(\pi t),$$
$$z = x_1 x_2 (1 - x_1)(1 - x_2) \sin(\pi t).$$

At first we consider the error with respect to spatial mesh size h. We fix the time step to $k = \frac{1}{8192}$ and present the errors of optimal control u, state y and adjoint state z in Table 1. Then we consider the convergence order of error with respect to time step size k. We fixed the space mesh with $DOF = 22785$ and present the errors of optimal control u, state y and adjoint state z in Table 2. We observe an order of convergence $\frac{3}{2}$ for the control and order 2 for the state and adjoint state for spatial discretization, and order 1 convergence for both of them for the time discretization. This is the best result we can expect for linear finite elements and dG(0) approximations.

Example 2 The second example is an unconstrained problem defined in a polygonal domain with maxminum interior angle $\omega_{max} = \frac{5}{6}\pi$ (see [29]), so that the optimal solution may have only reduced regularity. The data is chosen as

$$y_d = \begin{cases} t^2 g(x) & 0 \le t < 0.5, \\ -t^2 g(x) & 0.5 \le t \le 1 \end{cases}$$

with $f = 1$, $g = \frac{1}{(x_1^2 + x_2^2)^{\frac{1}{3}}}$.

Table 2 Error of control u, state y and adjoint state z for Example 1 with fixed mesh $DOF = 22785$

N	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
2	0.030398571028	\	0.025518889725	\	0.030398678795	\
4	0.018595263504	0.7091	0.015608671639	0.7092	0.016205193500	0.9076
8	0.011269239439	0.7225	0.008709737707	0.8416	0.008295451927	0.9661
16	0.006647997579	0.7614	0.004633730272	0.9105	0.004196415068	0.9832
32	0.003745029192	0.8279	0.002392405119	0.9537	0.002112281353	0.9904
64	0.002021680314	0.8894	0.001213660252	0.9791	0.001060211691	0.9944
128	0.001064326895	0.9256	0.000610081508	0.9923	0.000531194879	0.9970
256	0.000563475024	0.9175	0.000305197979	0.9993	0.000265839733	0.9987
512	0.000320272343	0.8151	0.000152181075	1.0040	0.000132937121	0.9998

Table 3 Error of control u, state y and adjoint state z for Example 2 with fixed time step $N = 4096$

Dof	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
53	0.021578034106	\	0.006632500684	\	0.002058710969	\
182	0.013947673670	0.7074	0.002850064176	1.3693	0.000543409130	2.1593
671	0.010619549501	0.4179	0.001537497778	0.9460	0.000141300731	2.0647
2573	0.008212461016	0.3825	0.000866946011	0.8525	0.000036700709	2.0060
10073	0.006098931924	0.4360	0.000467699526	0.9044	0.000009388603	1.9978
Average	\	0.4859	\	1.0180	\	2.0570

There is no exact solution for this example. We take the solution with $k = \frac{1}{4096}$ and $Dof = 158561$ in the spatial discretization as reference solution. Similarly as in the previous example, we investigate the convergence order with respect to the spatial and time discretization separately. Although the assumption $k = O(h^2)$ is not satisfied in this example, the analysis and numerical results in [30] and [31] suggest $O(h^{\frac{1}{2}})$ convergence for spatial discretization and $O(k^{\frac{1}{4}})$ convergence for time discretization in our case. We can observe in Table 3 nearly $O(h^{\frac{1}{2}})$-order convergence for the spatial discretization of the control. The convergence order for the time discretization reported in Table 4 is higher than $O(k^{\frac{1}{4}})$ which might be caused by a higher regularity of the control w.r.t time, compare Remark 1 and [23, Th. 3.4]. Caused by y_d we may expect a regularity loss w.r.t. time at $t = 0.5$, which in our opinion might only have a mild influence on the convergence order of the numerical computations.

Example 3 This example is a control constrained problem defined in a smooth domain (see [10]). The domain is the unit circle $\Omega = B(0, 1)$ with center at zero and radius 1, $T = 1$. The data is presented in polar coordinates. We set

$$f(r, \theta; t) = -6r \max(0, \cos\theta \sin^3(\pi t)) - \frac{\pi}{2}\sin(\pi t)r^3 \max(0, \cos^3\theta),$$

$$y_d(r, \theta; t) = (7r^2\cos^2\theta + 6r^2 - 6r)\cos\theta\sin^3(\pi t) + y(r, \theta)$$
$$- \frac{\pi}{2}\sin(\pi t)r^3(r - 1)\max(0, \cos^3\theta),$$

Table 4 Error of control u, state y and adjoint state z for Example 2 with fixed mesh $DOF = 158561$

N	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
4	0.167327134403	\	0.088804542387	\	0.020552016247	\
8	0.107460725354	0.6389	0.051579167598	0.7838	0.013212388713	0.6374
16	0.070415333055	0.6098	0.030045232355	0.7797	0.007759076126	0.7679
32	0.044502497910	0.6620	0.016906812048	0.8295	0.004258097400	0.8657
64	0.026671216305	0.7386	0.009227011579	0.8737	0.002231368045	0.9323
Average	\	0.6623	\	0.8167	\	0.8008

Table 5 Error of control u, state y and adjoint state z for Example 3 with fixed time step $N = 4096$ and full discretisation

Dof	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
25	0.080438933409	\	0.050499450897	\	0.027537485412	\
81	0.052462945814	0.6166	0.018390173000	1.4573	0.009386608073	1.5527
289	0.025693482841	1.0299	0.005955737135	1.6266	0.002546961550	1.8818
1089	0.010836772478	1.2455	0.001775580269	1.7460	0.000654383533	1.9606
4225	0.004214559039	1.3625	0.000499943184	1.8285	0.000180980668	1.8543

Table 6 Error of control u, state y and adjoint state z for Example 3 with fixed time step $N = 4096$ and variational discretisation

Dof	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
25	0.080080823101	\	0.050452860374	\	0.027529915188	\
81	0.052889327427	0.5985	0.018972070528	1.4111	0.009398192721	1.5506
289	0.025433412568	1.0563	0.005947170731	1.6736	0.002550336513	1.8817
1089	0.010686498089	1.2509	0.001738641208	1.7742	0.000652324545	1.9670
4225	0.004164564394	1.3596	0.000485052340	1.8417	0.000167939460	1.9576

Table 7 Error of control u, state y and adjoint state z for Example 3 with fixed mesh $DOF = 16641$ and full discretization

N	$\|u - U_{hk}\|_{L^2}$	Order	$\|y - Y_{hk}\|_{L^2}$	Order	$\|z - Z_{hk}\|_{L^2}$	Order
2	0.018744907559	\	0.027982932366	\	0.071397322443	\
4	0.024808047492	−0.4043	0.020226462340	0.4683	0.034454664183	1.0512
8	0.013491456808	0.8788	0.012915023887	0.6472	0.019032557741	0.8562
16	0.006930734029	0.9610	0.007288727815	0.8253	0.009953295271	0.9352
32	0.003529211695	0.9737	0.003870945517	0.9130	0.005085050476	0.9689
64	0.001839283189	0.9402	0.001994283984	0.9568	0.002569466386	0.9848
128	0.001046515255	0.8136	0.001012041776	0.9786	0.001291388317	0.9925
256	0.000720808717	0.5379	0.000510064738	0.9885	0.000647313735	0.9964

so that the optimal solution is given by

$$u(r, \theta; t) = \max(0, \cos^3 \theta \sin^3(\pi t)),$$
$$y(r, \theta; t) = r^3 \max(0, \cos^3 \theta \sin^3(\pi t)),$$
$$z(r, \theta; t) = r^3(r - 1)\cos^3 \theta \sin^3(\pi t).$$

We set $\alpha = 1$.

First we consider the error with respect to spatial mesh size h. We fix the time step $k = \frac{1}{4096}$ and present the error of the optimal control u, the state y and the adjoint state z in Table 5 with full discretisation, and in Table 6 with variational discretisation. We as expected observe that both approaches deliver similar results. Then we consider the convergence order of the time error. We fix the space mesh with $DOF = 16641$ and present the error of the optimal control u, the state y and the adjoint state z in Table 7. We observe higher order convergence w.r.t. the spatial discretization for both the control u and the state y.

References

1. Apel, T., Flaig, T.G.: Crank–Nicolson schemes for optimal control problems with evolution equations. SIAM J. Numer. Anal. **50**, 1484–1512 (2012)
2. Arada, N., Raymond, J.P.: Dirichlet boundary control of semilinear parabolic equations: part I. Appl. Math. Optim. **45**, 125–143 (2002)
3. Belgacem, F.B., Bernardi, C., Fekih, H.E.: Dirichlet boundary control for a parabolic equation with a final observation I: a space–time mixed formulation and penalization. Asymptot Anal. **71**, 101–121 (2011)
4. Berggren, M.: Approximation of very weak solutions to boundary value problems. SIAM J. Numer. Anal. **42**, 860–877 (2004)
5. Bramble, J.H., Pasciak, J.E., Schatz, A.H.: The construction of preconditioners for elliptic problems by substructing I. Math. Comput. **47**, 103–134 (1986)
6. Casas, E., Mateos, M., Raymond, J.P.: Penalization of Dirichlet optimal control problems. ESAIM Control Optim. Calc. Var. **15**, 782–809 (2009)
7. Casas, E., Raymond, J.P.: Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. SIAM J. Control Optim. **45**, 1586–1611 (2006)
8. Casas, E., Sokolowski, J.: Approximation of boundary control problems on curved domains. SIAM J. Control Optim. **48**, 3746–3780 (2010)
9. Ciarlet, P.G.: The Finite Element Methods for Elliptic Problems. Elsevier, North-Holland (1978)
10. Deckelnick, K., Günther, A., Hinze, M.: Finite element approximation of Dirichlet boundary control for elliptic PDEs on two- and three-dimensional curved domains. SIAM J. Control Optim. **48**, 2798–2819 (2009)
11. French, D.A., King, J.T.: Approximation of an elliptic control problem by the finite element method. Numer. Funct. Anal. Optim. **12**, 299–314 (1991)
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. Fursikov, A.V., Gunzburger, M.D., Hou, L.S.: Boundary value problems and optimal boundary control for the Navier-Stokes systems: the two-dimensional case. SIAM J. Control Optim. **36**, 852–894 (1998)
14. Geveci, T.: On the approximation of the solution of an optimal control problem governed by an elliptic equation. RAIRO Anal. Numer. **13**, 313–328 (1979)

15. Gong, W., Yan, N.N.: A posteriori error estimate for boundary control problems governed by the parabolic partial differential equations. J. Comput. Math. **27**, 68–88 (2009)
16. Gong, W., Yan, N.N.: Mixed finite element method for Dirichlet boundary control problem governed by elliptic PDEs. SIAM J. Control Optim. **49**, 984–1014 (2011)
17. Grisvard, P.: Singularities in Boundary Value Problems. Springer, Berlin (1992)
18. Gunzburger, M.D., Hou, L.S.: Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses. SIAM J. Numer. Anal. **29**, 390–424 (1992)
19. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. Comput. Optim. Appl. **30**, 45–63 (2005)
20. Hinze, M., Kunisch, K.: Second order methods for boundary control of the instationary Navier–Stokes system. ZAMM Z. Angew. Math. Mech. **84**, 171–187 (2004)
21. Hinze, M., Matthes, U.: A note on variational discretization of Neumann boundary control problems. Control Cybern. **38**, 577–591 (2009)
22. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE Constraints, Mathematical Modelling: Theory and Applications, vol. 23. Springer, Berlin (2009)
23. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in L^2 for a class of evolution equations. SIAM J. Control Optim. **46**, 1726–1753 (2007)
24. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
25. Lions, J.L., Magenes, E.: Non-homogeneous Boundary Value Problems and Applications, vol. I, II. Springer, Berlin (1972)
26. Liu, W.B., Ma, H.P., Tang, T., Yan, N.N.: A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations. SIAM J. Numer. Anal. **42**, 1032–1061 (2004)
27. Liu, W.B., Yan, N.N.: A posteriori error estimates for optimal control problems governed by parabolic equations. Numer. Math. **93**, 497–521 (2003)
28. Liu, W.B., Yan, N.N.: Adaptive Finite Element Methods for Optimal Control Governed by PDEs. Science press, Beijing (2008)
29. May, S., Rannacher, R., Vexler, B.: Error analysis fo a finite element approximation of elliptic Dirichlet boundary control problems. SIAM J. Control Optim. **51**(3), 2585–2611 (2013)
30. Meidner, D., Vexler, B.: A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part I: problems without control constraints. SIAM J. Control Optim. **47**(3), 1150–1177 (2008)
31. Meidner, D., Vexler, B.: A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part II: problems with control constraints. SIAM J. Control Optim. **47**(3), 1301–1329 (2008)
32. Meidner, D., Vexler, B.: A priori error analysis of the Petrov-Galerkin Crank–Nicolson scheme for parabolic optimal control problems. SIAM J. Control Optim. **49**(5), 2183–2211 (2011)
33. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (2006)
34. Vexler, B.: Finite element approximation of elliptic Dirichlet optimal control problems. Numer. Funct. Anal. Optim. **28**, 957–973 (2007)
35. von Daniels, N., Hinze, M., Vierling, M.: Crank-Nicolson time stepping and variational discretization of control-constrained parabolic optimal control problems. SIAM J. Control Optim. **53**(3), 1182–1198 (2015)
36. Winther, R.: Error estimates for a Galerkin approximation of a parabolic control problem. Ann. Math. Pura App. **117**(4), 173–206 (1978)