

# A priori error analysis of the upwind symmetric interior penalty Galerkin (SIPG) method for the optimal control problems governed by unsteady convection diffusion equations

Tuğba Akman · Hamdullah Yücel ·  
Bülent Karasözen

Received: 12 July 2012 / Published online: 25 September 2013  
© Springer Science+Business Media New York 2013

**Abstract** In this paper, we analyze the symmetric interior penalty Galerkin (SIPG) for distributed optimal control problems governed by unsteady convection diffusion equations with control constraint bounds. A priori error estimates are derived for the semi- and fully-discrete schemes by using piecewise linear functions. Numerical results are presented, which verify the theoretical results.

**Keywords** Unsteady convection diffusion equations · Optimal control · Discontinuous Galerkin methods · A priori error estimates

## 1 Introduction

Optimal control problems (OCPs) governed by convection diffusion partial differential equations (PDEs) arise in environmental modeling, petroleum reservoir simulation and in many other applications. Hence, efficient numerical methods are essential to obtain effective solutions of the such optimal control problems.

It is well known that the standard Galerkin finite element method produces non-physical oscillating solutions for mesh sizes larger than a critical value depending

---

T. Akman · H. Yücel · B. Karasözen  
Department of Mathematics & Institute of Applied Mathematics, Middle East Technical University,  
06800 Ankara, Turkey

T. Akman  
e-mail: [takman@metu.edu.tr](mailto:takman@metu.edu.tr)

B. Karasözen  
e-mail: [bulent@metu.edu.tr](mailto:bulent@metu.edu.tr)

*Present address:*

H. Yücel (✉)  
Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics  
of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany  
e-mail: [yuecel@mpi-magdeburg.mpg.de](mailto:yuecel@mpi-magdeburg.mpg.de)

on the ratio between diffusion and convection terms. To enhance stability and accuracy of the optimal control problems governed by the steady convection diffusion equations, some effective stabilization techniques are used, i.e., the streamline upwind/Petrov Galerkin (SUPG) finite element method [6], the local projection stabilization [1], the edge stabilization [12, 25]. Recently, discontinuous Galerkin (DG) methods have become popular for the optimal control problems governed by convection diffusion equations due the better convergence behavior, local mass conservation, flexibility in approximating rough solutions on complicated meshes and mesh adaptation, see, e.g., [15, 16, 26–28].

However, to the best of our knowledge, a few papers are published so far for unsteady optimal control problems governed by the convection diffusion equations. A characteristic finite element approximation in space and backward Euler method in time are used in [9, 10]. Zhou et al. [29] used local discontinuous Galerkin (LDG) discretization in space, whereas Sun [24] used the nonsymmetric interior penalty Galerkin (NIPG) discretization. In [24], a priori error estimates are only given for semi-discrete scheme, whereas it is investigated for both semi- and fully-discrete schemes with the backward Euler method in [29]. In both papers, numerical results are not given.

In this paper, we will investigate a priori error analysis of the optimal control problems governed by the unsteady convection diffusion equations using the symmetric interior penalty Galerkin (SIPG) method for the semi- and fully-discrete schemes. For time discretization, we apply the backward Euler method. We present the numerical results related to the DG discretization for the unsteady optimal control problems.

The rest of the paper is organized as follows: In Sect. 2, we introduce the control constrained optimal control problems governed by the unsteady convection diffusion equations. The upwind symmetric interior penalty Galerkin (SIPG) discretization and semi-discrete scheme are given in Sect. 3. A priori error estimates of the semi-discrete scheme are derived in Sect. 4. In Sect. 5, we give the fully-discrete scheme of the optimal control problems by using the backward Euler discretization in time. We derive a priori error estimates of the fully-discrete scheme in Sect. 6. Finally, we present the numerical results in Sect. 7.

## 2 The optimal control problem

We adopt the standard notations for Sobolev spaces on computational domains and their norms.  $\Omega$  and  $\Omega_U$  are bounded convex polygon domains in  $\mathbb{R}^2$  with Lipschitz boundaries  $\partial\Omega$  and  $\partial\Omega_U$ , respectively. The inner products in  $L^2(\Omega_U)$  and  $L^2(\Omega)$  are denoted by  $(\cdot, \cdot)_U$  and  $(\cdot, \cdot)$ , respectively. Further, we consider spaces of functions mapping the time interval  $(0, T)$  to a normed space  $X$  in which the norm  $\|\cdot\|_X$  is defined. For  $r \geq 1$ , we define

$$L^r(0, T; X) = \left\{ z : [0, T] \rightarrow X \text{ measurable} : \int_0^T \|z(t)\|_X^r dt < \infty \right\}$$

with

$$\|z(t)\|_{L^r(0,T;X)} = \begin{cases} (\int_0^T \|z(t)\|_X^r dt)^{1/r}, & \text{if } 1 \leq r < \infty, \\ \text{ess sup}_{t \in (0,T]} \|z(t)\|_X, & \text{if } r = \infty. \end{cases}$$

In this paper, we are interested in the following distributed optimal control problem governed by the unsteady diffusion convection reaction equation with control constraints

$$\underset{u \in U_{ad} \subseteq L^2(0,T;L^2(\Omega_U))}{\text{minimize}} \quad J(y, u) := \int_0^T \left( \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega_U)}^2 \right) dt, \tag{2.1}$$

subject to

$$\partial_t y - \varepsilon \Delta y + \beta \cdot \nabla y + r y = f + B u \quad x \in \Omega, \quad t \in (0, T], \tag{2.2a}$$

$$y(x, t) = 0 \quad x \in \partial\Omega, \quad t \in (0, T], \tag{2.2b}$$

$$y(x, 0) = y_0(x) \quad x \in \Omega, \tag{2.2c}$$

where the admissible space of control constraints is given by

$$U_{ad} = \{u \in L^2(0, T; L^2(\Omega_U)) : u_a \leq u \leq u_b, \text{ a.e. in } \Omega_U \times (0, T]\} \tag{2.3}$$

with the constant bounds  $u_a, u_b \in \mathbb{R} \cup \{\pm\infty\}$ , i.e.,  $u_a < u_b$ .  $B$  is a bounded linear continuous operator to ensure the transition from  $\Omega_U$  to  $\Omega$ . Generally,  $\Omega_U$  can be a subset of  $\Omega$ . In the special case,  $\Omega_U = \Omega$  and  $B = I$  is an identity operator.

We make the following assumptions for the functions and parameters on the optimal control problem (2.1), (2.2a)–(2.2c):

- (i) The source function  $f$  and the desired state  $y_d$  belong to  $H^1(0, T; L^2(\Omega))$  with  $f(0), y_d(T) \in H_0^1(\Omega)$ .
- (ii) The initial condition is defined as  $y_0(x) \in H_0^1(\Omega)$  with  $\Delta y_0 \in H_0^1(\Omega)$ .
- (iii) The diffusion and reaction parameters are denoted by  $\varepsilon > 0$  and  $r \in L^\infty(\Omega)$ , respectively.
- (iv)  $\beta$  denotes a velocity field. It belongs to  $(W^{1,\infty}(\Omega))^2$  and satisfies the incompressibility condition, i.e.  $\nabla \cdot \beta = 0$ .

Further, we assume the existence of a constant  $c_0 \equiv c_0(x) \geq 0$  such that

$$r(x) \geq c_0 \geq 0 \quad \text{a.e. in } \Omega \tag{2.4}$$

to ensure the well-posedness of the optimal control problem (2.1), (2.2a)–(2.2c).

Using the assumptions defined above, the following result on regularity of the state solution can be proved.

**Proposition 1** *Under the assumptions defined above and for a given control  $u \in H^1(0, T; L^2(\Omega_U))$ , the state  $y$  satisfies the following regularity condition*

$$y \in H^1(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^2(0, T; L^2(\Omega))$$

and the weak formulation

$$(\partial_t y, v) + a(y, v) + b(u, v) = (f, v) \quad \forall v \in V = H_0^1(\Omega), \quad t \in (0, T], \tag{2.5}$$

$$y(x, 0) = y_0, \tag{2.6}$$

where the (bi)-linear forms are defined by

$$a(y, v) = \int_{\Omega} (\varepsilon \nabla y \cdot \nabla v + \beta \cdot \nabla y v + r y v) dx, \quad b(u, v) = - \int_{\Omega} B u v dx,$$

$$(f, v) = \int_{\Omega} f v dx.$$

*Proof* The regularity of the state  $y \in H^1(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^2(0, T; L^2(\Omega))$  can be proved as done [7] provided that  $f + Bu \in H^1(0, T; L^2(\Omega))$  with  $(f + Bu)(0) \in H_0^1(\Omega)$  is satisfied. This condition is ensured by our assumptions. See, e.g., [20] for details.  $\square$

Then, variational formulation corresponding to (2.1), (2.2a)–(2.2c) can be written as

$$\underset{u \in U_{ad}}{\text{minimize}} \quad J(y, u) := \int_0^T \left( \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega_U)}^2 \right) dt \tag{2.7a}$$

$$\begin{aligned} \text{subject to} \quad & (\partial_t y, v) + a(y, v) + b(u, v) = (f, v) \quad \forall v \in V, \quad t \in (0, T], \tag{2.7b} \\ & y(x, 0) = y_0, \\ & (y, u) \in Y \times U_{ad}. \end{aligned}$$

It is well known that the triple  $(y, u)$  is the unique solution of (2.7a), (2.7b) if and only if there is an adjoint  $p \in H^1(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^2(0, T; L^2(\Omega))$  such that  $(y, u, p)$  satisfies the following optimality system

$$(\partial_t y, v) + a(y, v) + b(u, v) = (f, v) \quad \forall v \in V, \quad y(x, 0) = y_0, \tag{2.8a}$$

$$- (\partial_t y, \psi) + a(\psi, p) = -(y - y_d, \psi) \quad \forall \psi \in V, \quad p(x, T) = 0, \tag{2.8b}$$

$$\int_0^T (\alpha u - B^* p, w - u)_U dt \geq 0 \quad \forall w \in U_{ad}, \tag{2.8c}$$

where  $B^*$  denotes the adjoint of  $B$  [18, 20].

### 3 Discontinuous Galerkin (DG) scheme for optimal control problem

#### 3.1 Discontinuous Galerkin discretization

Let  $\{\mathcal{T}_h\}_h$  be a family of shape regular meshes such that  $\overline{\Omega} = \cup_{K \in \mathcal{T}_h} \overline{K}$ ,  $K_i \cap K_j = \emptyset$  for  $K_i, K_j \in \mathcal{T}_h, i \neq j$ . The diameter of an element  $K$  and the length of an edge

$E$  are denoted by  $h_K$  and  $h_E$ , respectively. Further, the maximum value of element diameter is denoted by  $h = \max_{K \in \mathcal{T}_h} h_K$ .

We only consider discontinuous piecewise linear finite element spaces to define the discrete spaces of the state and test functions

$$V_h = Y_h = \{y \in L^2(\Omega) : y|_K \in \mathbb{P}^1(K) \forall K \in \mathcal{T}_h\}. \tag{3.1}$$

*Remark 1* When the state equation (2.2a)–(2.2c) contains nonhomogeneous Dirichlet boundary conditions, the space of discrete states  $Y_h$  and the space of test functions  $V_h$  can still be taken the same due to the weak treatment of boundary conditions in DG methods. See, [16] for details.

We split the set of all edges  $\mathcal{E}_h$  into the set  $\mathcal{E}_h^0$  of interior edges and the set  $\mathcal{E}_h^\partial$  of boundary edges so that  $\mathcal{E}_h = \mathcal{E}_h^\partial \cup \mathcal{E}_h^0$ . Let  $\mathbf{n}$  denote the unit outward normal to  $\partial\Omega$ . We define the inflow boundary

$$\Gamma^- = \{x \in \partial\Omega : \beta \cdot \mathbf{n}(x) < 0\}$$

and the outflow boundary  $\Gamma^+ = \partial\Omega \setminus \Gamma^-$ . The boundary edges are decomposed into edges  $\mathcal{E}_h^- = \{E \in \mathcal{E}_h^\partial : E \subset \Gamma^-\}$  that correspond to inflow boundary and edges  $\mathcal{E}_h^+ = \mathcal{E}_h^\partial \setminus \mathcal{E}_h^-$  that correspond to outflow boundary. The inflow and outflow boundaries of an element  $K \in \mathcal{T}_h$  are defined by

$$\partial K^- = \{x \in \partial K : \beta \cdot \mathbf{n}_K(x) < 0\}, \quad \partial K^+ = \partial K \setminus \partial K^-,$$

where  $\mathbf{n}_K$  is the unit normal vector on the boundary  $\partial K$  of an element  $K$ .

Let the edge  $E$  be a common edge for two elements  $K$  and  $K^e$ . For a piecewise continuous scalar function  $y$ , there are two traces of  $y$  along  $E$ , denoted by  $y|_E$  from inside  $K$  and  $y^e|_E$  from inside  $K^e$ . Then, the jump and average of  $y$  across the edge  $E$  are defined by:

$$[[y]] = y|_E \mathbf{n}_K + y^e|_E \mathbf{n}_{K^e}, \quad \{\{y\}\} = \frac{1}{2}(y|_E + y^e|_E). \tag{3.2}$$

Similarly, for a piecewise continuous vector field  $\nabla y$ , the jump and average across an edge  $E$  are given by

$$[[\nabla y]] = \nabla y|_E \cdot \mathbf{n}_K + \nabla y^e|_E \cdot \mathbf{n}_{K^e}, \quad \{\{\nabla y\}\} = \frac{1}{2}(\nabla y|_E + \nabla y^e|_E). \tag{3.3}$$

For a boundary edge  $E \in K \cap \Gamma$ , we set  $\{\{\nabla y\}\} = \nabla y$  and  $[[y]] = y \mathbf{n}$  where  $\mathbf{n}$  is the outward normal unit vector on  $\Gamma$ .

We now consider the discretization of the control variable. Let  $\{\mathcal{T}_h^U\}_h$  is also a family of shape regular meshes of  $\Omega_U$  such that  $\overline{\Omega}_U = \bigcup_{K_U \in \mathcal{T}_h^U} \overline{K}_U$ ,  $K_U^i \cap K_U^j = \emptyset$  for  $K_U^i, K_U^j \in \mathcal{T}_h^U$ ,  $i \neq j$ . The maximum diameter is defined by  $h_U = \max_{K_U \in \mathcal{T}_h^U} h_{K_U}$ , where  $h_{K_U}$  denotes the diameter of an element  $K_U$ . The discrete

space of the control variable associated with  $\{\mathcal{T}_h^U\}_h$  is also piecewise linear finite element space

$$U_h = \{u \in L^2(\Omega_U) : u|_{K_U} \in \mathbb{P}^1(K_U) \forall K_U \in \mathcal{T}_h^U\}. \tag{3.4}$$

Note that in general, the sizes of the elements in  $\{\mathcal{T}_h^U\}_h$  are smaller than those in  $\{\mathcal{T}_h\}_h$ , so we assume that  $h_U/h \leq C$  throughout this paper.

We can now give DG discretizations of the state equation (2.2a)–(2.2c) in space for fixed control  $u$ . The DG method proposed here is based on the upwind discretization of the convection term and on the SIPG discretization of the diffusion term [22]. This leads to the following (bi-)linear forms applied to  $y_h \in H^1(0, T; Y_h)$  for  $\forall t \in (0, T]$

$$(\partial_t y_h, v_h) + a_h(y_h, v_h) + b_h(u_h, v_h) = (f_h, v_h) \quad \forall v_h \in V_h, t \in (0, T], \tag{3.5}$$

where

$$a_h(y, v)$$

$$\begin{aligned} &= \underbrace{\sum_{K \in \mathcal{T}_h} \int_K \varepsilon \nabla y \cdot \nabla v \, dx - \sum_{E \in \mathcal{E}_h} \int_E \left( \{\{\varepsilon \nabla y\}\} \cdot \llbracket v \rrbracket + \{\{\varepsilon \nabla v\}\} \cdot \llbracket y \rrbracket - \frac{\sigma \varepsilon}{h_E} \llbracket y \rrbracket \cdot \llbracket v \rrbracket \right) ds}_{a^d(y,v)} \\ &+ \underbrace{\sum_{K \in \mathcal{T}_h} \int_K (\beta \cdot \nabla y v + r y v) \, dx + \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \setminus \Gamma^-} \beta \cdot \mathbf{n} (y^e - y) v \, ds - \sum_{K \in \mathcal{T}_h} \int_{\partial K^- \cap \Gamma^-} \beta \cdot \mathbf{n} y v \, ds}_{a^{cr}(y,v)}, \end{aligned} \tag{3.6a}$$

$$b_h(u, v) = - \sum_{K \in \mathcal{T}_h} \int_K B u v \, dx \tag{3.6b}$$

with a constant interior penalty parameter  $\sigma > 0$ . We choose  $\sigma$  to be sufficiently large, independent of the mesh size  $h$  and the diffusion coefficient  $\varepsilon$  to ensure the stability of the DG discretization as described in [21, Sect. 2.7.1] with a lower bound depending only on the polynomial degree. Large penalty parameters decrease the jumps across element interfaces, which can affect the numerical approximation. Further, the DG approximation can converge to the continuous Galerkin approximation as the penalty parameter goes to infinity. See, e.g., [5] for details.

To make the notation easier for the readers, we introduce the  $L^2$  inner product on the inflow or outflow boundaries as follows

$$(w, v)_{\Gamma^-} = \int_{\Gamma^-} |\beta \cdot \mathbf{n}| w v \, ds$$

with analogous definition of  $(\cdot, \cdot)_{\Gamma^+}$  and associated norms  $\|\cdot\|_{\Gamma^-}$  and  $\|\cdot\|_{\Gamma^+}$ . Further, the standard notation  $W^{m,q}(\Omega)$  is used for the Sobolev space with a norm  $\|\cdot\|_{W^{m,q}(\Omega)}$  and the broken Sobolev spaces used in DG discretization are given by

$$\|v\|_{W^{m,q}(\mathcal{T}_h)} = \left( \sum_{K \in \mathcal{T}_h} \|v\|_{W^{m,q}(K)}^2 \right)^{1/2}.$$

### 3.2 Semi-discrete formulation of optimal control problem

The discretization of admissible set (2.3) is defined by

$$U_h^{ad} = \{u_h \in L^2(0, T; U_h) : u_a \leq u_h \leq u_b \text{ a.e. in } \Omega_U \times (0, T)\}. \tag{3.7}$$

Let  $f_h, y_h^d$  and  $y_h^0$  be approximations of the source function  $f$ , the desired state function  $y_d$  and initial condition  $y_0$ , respectively. Then, the semi-discrete approximation of the optimal control problem (2.8a)–(2.8c) can be defined as follows:

$$\underset{u_h \in U_h^{ad}}{\text{minimize}} \int_0^T \left( \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|y_h - y_h^d\|_{L^2(K)}^2 + \frac{\alpha}{2} \sum_{K_U \in \mathcal{T}_h^U} \|u_h\|_{L^2(K_U)}^2 \right) dt, \tag{3.8a}$$

$$\begin{aligned} \text{subject to } & (\partial_t y_h, v_h) + a_h(y_h, v_h) + b_h(u_h, v_h) = (f_h, v_h) \quad \forall v_h \in V_h, t \in (0, T], \\ & y_h(x, 0) = y_h^0, \quad (y_h, u_h) \in Y_h \times U_h^{ad}. \end{aligned} \tag{3.8b}$$

### 4 A priori error analysis of semi-discrete scheme

In this section, we derive a priori error estimates for the semi-discrete scheme of the optimal control problem (2.1), (2.2a)–(2.2c) by using the upwind symmetric interior penalty Galerkin (SIPG) discretization for the space. By introducing the following norm [21]

$$\|v\|_\varepsilon^2 = \sum_{K \in \mathcal{T}_h} \int_K \|\varepsilon \nabla y\|_{L^2(K)}^2 dx + \sum_{E \in \mathcal{E}_h} \frac{\sigma \varepsilon}{h_E} \int_E \llbracket y \rrbracket \|_{L^2(E)}^2 ds,$$

we obtain the following coercivity result for some positive constant  $\kappa > 0$  independent of the mesh size  $h$  and the diffusion parameter  $\varepsilon$  provided that a sufficiently large penalty parameter  $\sigma$  is chosen based on the polynomial degree as described in [21]:

$$\forall t > 0, \forall v_h \in V_h, \quad \kappa \|v\|_\varepsilon^2 \leq a_h^d(v, v). \tag{4.1}$$

We also need the following trace inequality at the rest of the paper:

$$\|v\|_{L^2(E)} \leq C |h_E|^{1/2} |h_K|^{-1/2} \|v\|_{L^2(K)}, \quad \forall E \subset \partial K, \tag{4.2}$$

where the constant  $C$  is independent of mesh size  $h$ , but depends on polynomial degree. In addition, the generalization of Poincaré-Friedrichs inequality to the broken Sobolev space  $H^1(\mathcal{T}_h)$  [4]

$$\|v\|_{L^2(\Omega)}^2 \leq C \left( \|v\|_{H^0(\mathcal{T}_h)}^2 + \sum_{E \in \mathcal{E}_h} \frac{1}{h_E} \llbracket v \rrbracket \|_{L^2(E)}^2 \right), \quad \forall v \in H^1(\mathcal{T}_h) \tag{4.3}$$

is needed for some of the following proofs.

Now, we turn to derive a semi-discrete stability estimate for the state variable at the following Lemma 1.

**Lemma 1** *Let  $y_h$  be the solution of (3.8b) and let  $c_0$  be a positive constant such that (2.4) holds. Then, there exists a positive constant  $C$  independent of mesh size  $h$  for  $\forall t \in (0, T]$  such that*

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \|y_h(t)\|_{L^2(K)}^2 + \int_0^t \|y_h\|_\varepsilon^2 dt \\ & + \int_0^t \left( \sum_{K \in \mathcal{T}_h} c_0 \|y_h\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_h} \|y_h\|_{L^2(\partial K^- \cap \Gamma^-)}^2 \right) dt \\ & + \int_0^t \left( \sum_{K \in \mathcal{T}_h} \|y_h - y_h^e\|_{L^2(\partial K^- \setminus \Gamma^-)}^2 + \sum_{K \in \mathcal{T}_h} \|y_h\|_{L^2(\partial K^+ \cap \Gamma^+)}^2 \right) dt \\ & \leq C \left( \sum_{K \in \mathcal{T}_h} \|y_h^0\|_{L^2(K)}^2 + \int_0^t \left( \sum_{K \in \mathcal{T}_h} \|f\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_h} \|Bu_h\|_{L^2(K)}^2 \right) dt \right). \end{aligned} \tag{4.4}$$

*Proof* The proof is shown as done in [24, Lemma 3.1]. □

Let  $J(\cdot)$  be a continuous functional in  $L^2(\Omega)$ . Then, there exists at least one solution for the minimization problem (3.8a), (3.8b) since  $\int_0^T \sum_{K \in \mathcal{T}_h} \|y(u_h)\|_{H^1(K)}^2$  is bounded as proven in Lemma 1 (see, e.g., [24] for details). Then, we can deduce that the semi-discrete optimal control problem (3.8a), (3.8b) has a unique solution  $(y_h, u_h) \in H^1(0, T; Y_h) \times U_h^{ad}$ . See, e.g., [18]. The functions  $(y_h, u_h) \in H^1(0, T; Y_h) \times U_h^{ad}$  solve (3.8a), (3.8b) if and only if  $(y_h, u_h, p_h) \in H^1(0, T; Y_h) \times U_h^{ad} \times H^1(0, T; Y_h)$  is a unique solution of the following optimality system:

$$(\partial_t y_h, v_h) + a_h(y_h, v_h) + b(u_h, v_h) = (f_h, v_h) \quad \forall v_h \in V_h, \quad y_h(x, 0) = y_h^0, \tag{4.5a}$$

$$-(\partial_t p_h, \psi_h) + a_h(\psi_h, p_h) = -(y_h - y_h^d, \psi_h) \quad \forall \psi_h \in V_h, \quad p_h(x, T) = 0, \tag{4.5b}$$

$$\int_0^T (\alpha u_h - B^* p_h, w_h - u_h)_U dt \geq 0 \quad \forall w_h \in U_h^{ad}. \tag{4.5c}$$

Similar to Lemma 1, we can obtain the following semi-discrete stability estimate for the adjoint variable in Lemma 2.

**Lemma 2** *Let  $p_h$  be the solution of (4.5b) and let  $c_0$  be a positive constant such that (2.4) holds. Then, there exists a positive constant  $C$  independent of  $h$  such that*

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \|p_h(t)\|_{L^2(K)}^2 + \int_t^T \|p_h\|_\varepsilon^2 dt \\ & + \int_t^T \left( \sum_{K \in \mathcal{T}_h} c_0 \|p_h\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_h} \|p_h\|_{L^2(\partial K^- \cap \Gamma^-)}^2 \right) dt \end{aligned}$$



$$\begin{aligned}
 & + \int_t^T \left( \sum_{K \in \mathcal{T}_h} \|p_h - p_h^e\|_{L^2(\partial K^+ \setminus \Gamma^+)}^2 + \sum_{K \in \mathcal{T}_h} \|p_h\|_{L^2(\partial K^+ \cap \Gamma^+)}^2 \right) dt \\
 & \leq C \int_t^T \sum_{K \in \mathcal{T}_h} \|y_h - y_h^d\|_{L^2(K)}^2 dt.
 \end{aligned} \tag{4.6}$$

*Proof* The proof is similar to (4.4) with  $p_h(x, T) = 0$ . □

In order to derive a priori error estimates for the semi-discrete scheme, we make use of the following definitions and estimates. Firstly, we define an elliptic projection  $\tilde{y}$  of  $y$  onto  $Y_h$  satisfying the Galerkin orthogonality

$$a_h^d(y(t) - \tilde{y}(t), v) = 0 \quad \forall t \geq 0, \forall v \in V_h \tag{4.7}$$

to derive an error estimate for  $y - y_h(u)$ . Then, we use the following estimates that are given in [21]:

$$\|y(t) - \tilde{y}(t)\|_\varepsilon \leq Ch \|y(t)\|_{H^2(\mathcal{T}_h)} \quad \forall t \geq 0, \tag{4.8a}$$

$$\|y(t) - \tilde{y}(t)\|_{L^2(\Omega)} \leq Ch^2 \|y(t)\|_{H^2(\mathcal{T}_h)} \quad \forall t \geq 0. \tag{4.8b}$$

Moreover, the domain  $\Omega_U$  is divided as the active and inactive regions of the control  $u$  for each time interval as firstly introduced in [17]

$$\begin{aligned}
 \Omega_U^* &= \left\{ \bigcup_{K_U} : K_U \subset \Omega_U, u_a < u|_{K_U} < u_b \right\}, \\
 \Omega_U^c &= \left\{ \bigcup_{K_U} : K_U \subset \Omega_U, u|_{K_U} = u_a \text{ or } u|_{K_U} = u_b \right\}, \\
 \Omega_U^b &= \Omega \setminus (\Omega_U^* \cup \Omega_U^c).
 \end{aligned}$$

It is assumed that the intersection of the three sets is empty, i.e.,  $\Omega_U^i \cap \Omega_U^j = \emptyset$  for  $i \neq j$  and  $\Omega_U = \Omega_U^* \cup \Omega_U^c \cup \Omega_U^b$ .  $\Omega_U^b$  consists of elements which lie close to the free boundary between the active and the inactive sets for each time interval. We also hold the following assumption

$$\text{meas}(\Omega_U^b) \leq Ch_U \tag{4.9}$$

on the regularity of  $u$  and  $\mathcal{T}_h^U$ . This assumption is valid if the boundary of the level set  $\Omega_U^c$  consists of a finite number of rectifiable curves [19]. In addition, we set

$$\Omega^* = \{x \in \Omega_U : u_a < u(x) < u_b\},$$

which includes  $\Omega_U^* \subset \Omega^*$  [25].

We finally define

$$(J'_h(u), v - u)_U = \int_0^T (\alpha u - B^* p_h(u), v - u)_U dt, \tag{4.10}$$

in which the auxiliary solution  $p_h(u) \in H^1(0, T; Y_h)$  is the solution of the following system

$$\begin{aligned}
 &(\partial_t y_h(u), v_h) + a_h(y_h(u), v_h) + b_h(u, v_h) = (f_h, v_h) \\
 &\forall v_h \in V_h, \quad y_h(u)(x, 0) = y_0^h,
 \end{aligned}
 \tag{4.11a}$$

$$\begin{aligned}
 &-(\partial_t p_h(u), q_h) + a_h(q_h, p_h(u)) = -(y_h(u) - y_h^d, q_h) \\
 &\forall q_h \in V_h, \quad p_h(u)(x, T) = 0,
 \end{aligned}
 \tag{4.11b}$$

where  $y_h(u) \in H^1(0, T; Y_h)$  is also an auxiliary solution for given  $u \in U_h^{ad}$ .

To complete the a priori error estimate of semi-discrete scheme, we firstly derive convergence estimates between the approximate solutions  $(y_h, p_h)$  and the auxiliary solutions  $(y_h(u), p_h(u))$ .

**Lemma 3** *Let  $(y_h, p_h)$  and  $(y_h(u), p_h(u))$  be the solutions of (4.5a), (4.5b) and (4.11a), (4.11b), respectively. Then, there are positive constants  $C_1$  and  $C_2$  independent of  $h$  such that*

$$\|y_h - y_h(u)\|_{L^\infty(0,T;L^2(\Omega))} \leq C_1 \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))}
 \tag{4.12a}$$

and

$$\|p_h - p_h(u)\|_{L^\infty(0,T;L^2(\Omega))} \leq C_2 \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))}.
 \tag{4.12b}$$

*Proof* By subtracting (4.11a) (respectively, (4.11b)) from (4.5a) (respectively, (4.5b)), taking  $v_h = y_h - y_h(u)$  (respectively,  $v_h = p_h - p_h(u)$ ) and following the approach in the stability estimates of the semi-discrete state equation (respectively, the semi-discrete adjoint equation), the desired results are obtained.  $\square$

Now, we will derive an estimate for the control  $u$  using the discontinuous piecewise linear finite element space by following the approach in [25, 29].

**Lemma 4** *Let  $(y, p, u)$  and  $(y_h, p_h, u_h)$  be the solutions of (2.8a)–(2.8c) and (4.5a)–(4.5c), respectively. Assume that  $u \in L^2(0, T; W^{1,\infty}(\Omega_U)), u|_{\Omega^*} \in L^2(0, T; H^2(\Omega^*))$ . Then, we have*

$$\|u - u_h\|_{L^2(0,T;L^2(\Omega_U))} \leq C(h_U^{3/2} + \|p - p_h(u)\|_{L^2(0,T;L^2(\Omega))}).
 \tag{4.13}$$

*Proof* Let  $(J'_h(u), v - u)_U = \int_0^T (\alpha u - B^* p_h(u), v - u)_U dt$ , where  $p_h(u)$  is the solution of the auxiliary equation (4.11b). Then, we have

$$\begin{aligned}
 (J'_h(v) - J'_h(u), v - u)_U &= \int_0^T (\alpha(v - u), v - u)_U dt \\
 &\quad + \int_0^T (B^* p_h(u) - B^* p_h(v), v - u)_U dt.
 \end{aligned}$$

By using the auxiliary equations (4.11a) and (4.11b), we obtain

$$\begin{aligned}
 & \int_0^T (Bv - Bu, p_h(u) - p_h(v))_U dt \\
 &= \int_0^T (\partial_t(y_h(v) - y_h(u)), p_h(u) - p_h(v)) dt \\
 & \quad + \int_0^T (a_h(y_h(v) - y_h(u)), p_h(u) - p_h(v)) dt \\
 &= \int_0^T (\partial_t(y_h(v) - y_h(u)), p_h(u) - p_h(v)) dt \\
 & \quad + \int_0^T (\partial_t(p_h(u) - p_h(v)), y_h(v) - y_h(u)) dt \\
 & \quad + \int_0^T (y_h(v) - y_h(u), y_h(v) - y_h(u)) dt.
 \end{aligned}$$

Application of integration by parts on the first term by using the fact  $(y_h(v) - y_h(u))|_{t=0} = 0$  and  $(p_h(v) - p_h(u))|_{t=T} = 0$  yields

$$\int_0^T (v - u, B^* p_h(u) - B^* p_h(v))_U dt = \int_0^T (y_h(v) - y_h(u), y_h(v) - y_h(u)) dt \geq 0. \tag{4.14}$$

By using (4.14), we obtain

$$(J'_h(v) - J'_h(u), v - u)_U \geq \alpha \int_0^T \|v - u\|_{L^2(\Omega_U)}^2 dt. \tag{4.15}$$

With the help of the inequalities (4.15), (4.14), (2.8c), (4.5c), the standard Lagrangian interpolation  $\Pi u$  with Young’s inequality and the notation  $p_h = p_h(u_h)$ , we obtain

$$\begin{aligned}
 & \alpha \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))}^2 \\
 & \leq \underbrace{\int_0^T (\alpha u - B^* p, u - u_h)_U dt}_{\geq 0} + \int_0^T (B^* p - B^* p_h(u), u - u_h)_U dt \\
 & \quad + \underbrace{\int_0^T (\alpha u_h - B^* p_h, u_h - \Pi u)_U dt}_{\geq 0} + \int_0^T (\alpha u_h - B^* p_h, \Pi u - u)_U dt \\
 & \leq \int_0^T (B^* p - B^* p_h(u), u - u_h)_U dt \\
 & \quad + \int_0^T (\alpha u_h - B^* p_h - \alpha u + B^* p, \Pi u - u)_U dt
 \end{aligned}$$

$$\begin{aligned}
 & + \int_0^T (\alpha u - B^* p, \Pi u - u)_U dt \\
 = & \int_0^T (B^* p - B^* p_h(u), u - u_h)_U dt + \int_0^T (\alpha u_h - \alpha u, \Pi u - u)_U dt \\
 & + \int_0^T (B^* p - B^* p_h(u), \Pi u - u)_U dt + \int_0^T (B^* p_h(u) - B^* p_h, \Pi u - u)_U dt \\
 & + \int_0^T (\alpha u - B^* p, \Pi u - u)_U dt \\
 \leq & \int_0^T (\alpha u - B^* p, \Pi u - u)_U dt + C_1 \|B^* p_h(u) - B^* p_h\|_{L^2(0,T;L^2(\Omega_U))}^2 \\
 & + C_2 \|u - \Pi u\|_{L^2(0,T;L^2(\Omega_U))}^2 + C_3 \|B^* p - B^* p_h(u)\|_{L^2(0,T;L^2(\Omega_U))}^2 \\
 & + C_4 \|\alpha u - \alpha u_h\|_{L^2(0,T;L^2(\Omega_U))}^2 + C_5 \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))}^2. \tag{4.16}
 \end{aligned}$$

As described in (3.4), we use the discontinuous piecewise linear finite element space for the control variable. Assuming  $\Pi u$  is the standard Lagrangian interpolation satisfying  $\Pi u(x) = u(x)$  for any vertex  $x$ . Then,  $\Pi u$  belongs to  $U_{ad}^h$ . We get

$$\|u - \Pi u\|_{L^2(\Omega_U^*)} \leq Ch_U^2 \|u\|_{H^2(\Omega_U^*)}, \quad \|u - \Pi u\|_{W^{0,\infty}(\Omega_U^b)} \leq Ch_U \|u\|_{W^{1,\infty}(\Omega_U^b)}$$

for  $u \in W^{1,\infty}(\Omega_U)$  and  $u|_{\Omega^*} \subset H^2(\Omega^*)$ . Hence,

$$\begin{aligned}
 \|u - \Pi u\|_{L^2(\Omega_U)}^2 & = \int_{\Omega_U^*} (u - \Pi u)^2 + \int_{\Omega_U^c} (u - \Pi u)^2 + \int_{\Omega_U^b} (u - \Pi u)^2 \\
 & \leq Ch_U^4 \|u\|_{H^2(\Omega_U^*)}^2 + 0 + Ch_U^2 \|u\|_{W^{1,\infty}(\Omega_U^b)}^2 \text{meas}(\Omega_U^b) \\
 & \leq Ch_U^3 (h_U \|u\|_{H^2(\Omega_U^*)}^2 + \|u\|_{W^{1,\infty}(\Omega_U^b)}^2) \\
 & \leq Ch_U^3 (\|u\|_{H^2(\Omega_U^*)}^2 + \|u\|_{W^{1,\infty}(\Omega_U)}^2). \tag{4.17}
 \end{aligned}$$

By the inequality in (4.5c), we have

$$\alpha u - B^* p = 0 \quad \text{on } \Omega_U^* \quad \text{and} \quad \Pi u - u = 0 \quad \text{on } \Omega_U^c.$$

In addition, there exists  $x_0 \in K_U \subset \Omega_U^b$  with  $u_a < u(x_0) < u_b$  satisfying  $(\alpha u - B^* p)(x_0) = 0$ . Then, the following estimate by [25]

$$\begin{aligned}
 \|\alpha u - B^* p\|_{W^{0,\infty}(\Omega_U^b)} & = \|\alpha u - B^* p - (\alpha u - B^* p)(x_0)\|_{W^{0,\infty}(\Omega_U^b)} \\
 & \leq Ch_U \|\alpha u - B^* p\|_{W^{1,\infty}(\Omega_U^b)}
 \end{aligned}$$

results in

$$\begin{aligned}
 (\alpha u - B^* p, \Pi u - u)_U &= \int_{\Omega_U^*} (\alpha u - B^* p)(\Pi u - u) + \int_{\Omega_U^c} (\alpha u - B^* p)(\Pi u - u) \\
 &\quad + \int_{\Omega_U^b} (\alpha u - B^* p)(\Pi u - u) \\
 &= 0 + 0 + \int_{\Omega_U^b} (\alpha u - B^* p)(\Pi u - u) \\
 &\leq \|\alpha u - B^* p\|_{W^{0,\infty}(\Omega_U^b)} \|u - \Pi u\|_{W^{0,\infty}(\Omega_U^b)} \text{meas}(\Omega_U^b) \\
 &\leq Ch_U^3 \|\alpha u - B^* p\|_{W^{1,\infty}(\Omega_U^b)} \|u\|_{W^{1,\infty}(\Omega_U^b)}. \tag{4.18}
 \end{aligned}$$

Finally, the desired result is obtained by inserting (4.17), (4.18) and (4.12b) into (4.16). □

*Remark 2* In Lemma 4, we assume that  $u \in W^{1,\infty}(\Omega_U)$  and  $u \in H^2(\Omega^*)$  in space, instead of  $u \in H^2(\Omega_U)$  due to the regularity issues on the boundary of the control as done [24, 25, 29]. The control variable  $u$  has lower regularity due to the discontinuity of the derivative of  $u$  on the free boundary  $\Omega^b$ . Hence, the convergence rate of the control  $u$  is around  $h^{3/2}$ . However, in numerical experiments, the optimal convergence rate can be obtained if the initial mesh is generated properly. It means that the initial grid aligns with the points where the bounds of control and the values of adjoint coincide. Hence, there happens no kink.

In the following lemma, the connection between the exact solution of the state  $y$  (respectively, the adjoint  $p$ ) and the auxiliary state solution  $y_h(u)$  (respectively, the auxiliary adjoint solution  $p_h(u)$ ) will be established.

**Lemma 5** *Let  $(y, p)$  be the solutions of (2.8a), (2.8b), respectively and  $(y_h(u), p_h(u))$  be the solutions of the auxiliary equations (4.11a), (4.11b), respectively. Then, there is a constant  $C$  independent of  $h$  such that*

$$\|y - y_h(u)\|_{L^\infty(0,T;L^2(\Omega))} \leq Ch^2 \|y\|_{H^1(0,T;H^2(\mathcal{T}_h))} \tag{4.19}$$

and

$$\|p - p_h(u)\|_{L^\infty(0,T;L^2(\Omega))} \leq Ch^2 (\|p\|_{H^1(0,T;H^2(\mathcal{T}_h))} + \|y\|_{H^1(0,T;H^2(\mathcal{T}_h))}). \tag{4.20}$$

*Proof* To show the estimate of the state (4.19), we begin with subtracting (4.11a) from (2.8a),

$$\left( \frac{\partial(y - y_h(u))}{\partial t}, v_h \right) + a_h^d(y - y_h(u), v_h) + a_h^{cr}(y - y_h(u), v_h) = 0, \quad \forall v_h \in V_h.$$

By writing

$$y - y_h(u) = (y - \tilde{y}) - (y_h(u) - \tilde{y}) = \eta - \xi,$$

where  $\tilde{y}$  is the elliptic projection of  $y$  and taking  $v_h = \xi$ , we obtain

$$\left(\frac{\partial \xi}{\partial t}, \xi\right) + a_h^d(\xi, \xi) + a_h^{cr}(\xi, \xi) = \left(\frac{\partial \eta}{\partial t}, \xi\right) + a_h^d(\eta, \xi) + a_h^{cr}(\eta, \xi), \quad \forall t > 0.$$

Coercivity of  $a_h^d(\cdot, \cdot)$  (4.1) and the Galerkin orthogonality (4.7) yield

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\xi\|_{L^2(\Omega)}^2 + \kappa \|\xi\|_\varepsilon^2 + \sum_{K \in \mathcal{T}_h} c_0 \|\xi\|_K^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi\|_{\partial K^- \cap \Gamma^-}^2 \\ & + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi - \xi^e\|_{\partial K^- \setminus \Gamma^-}^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi(t)\|_{\partial K^+ \cap \Gamma^+}^2 \\ & \leq \left| \left(\frac{\partial \eta}{\partial t}, \xi\right) + a_h^{cr}(\eta, \xi) \right|. \end{aligned} \quad (4.21)$$

The bounds in [21] for the first term in right-hand side of (4.21), i.e.,  $(\frac{\partial \eta}{\partial t}, \xi)$ , and the bounds in [8] for the second term, i.e.,  $a_h^{cr}(\eta, \xi)$ , give us

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\xi\|_{L^2(\Omega)}^2 + \kappa \|\xi\|_\varepsilon^2 + \sum_{K \in \mathcal{T}_h} c_0 \|\xi\|_K^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi\|_{\partial K^- \cap \Gamma^-}^2 \\ & + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi - \xi^e\|_{\partial K^- \setminus \Gamma^-}^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \\ & \leq \frac{\kappa}{2} \|\xi\|_\varepsilon^2 + Ch^4 \left\| \left\| \frac{\partial y}{\partial t} \right\| \right\|_{H^2(\mathcal{T}_h)}^2 + \frac{\kappa}{8} \|\xi\|_\varepsilon^2 + C \|\eta\|_{L^2(\Omega)}^2 + C \|\xi\|_{L^2(\Omega)}^2 \\ & + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \|\xi - \xi^e\|_{\partial K^- \setminus \Gamma^-}^2 + \sum_{K \in \mathcal{T}_h} \|\eta^e\|_{\partial K^- \setminus \Gamma^-}^2 + \frac{1}{4} \sum_{K \in \mathcal{T}_h} \|\xi\|_{\partial K^+ \cap \Gamma^+}^2 \\ & + \sum_{K \in \mathcal{T}_h} \|\eta\|_{\partial K^+ \cap \Gamma^+}^2. \end{aligned} \quad (4.22)$$

Now, we eliminate the terms related to  $\eta$  by using the estimate related to  $\|\eta\|_{\partial K^-}$  or  $\|\eta\|_{\partial K^+}$  in [13], trace inequality (4.2) and elliptic projection (4.7)

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \|\eta^e\|_{\partial K^- \setminus \Gamma^-}^2 & \leq \sum_{K \in \mathcal{T}_h} \|\beta\|_{L^\infty(K)} \|\eta\|_{\partial K}^2 \leq \sum_{K \in \mathcal{T}_h} C \|\eta\|_K^2 = C \|\eta\|_{L^2(\Omega)}^2 \\ & \leq Ch^4 \|y\|_{H^2(\mathcal{T}_h)}^2. \end{aligned} \quad (4.23)$$

A bound for  $\|\xi\|_{L^2(\Omega)}^2$  is derived by multiplying (4.22) by 2 and integrating from 0 to  $t$ . Using the continuous Gronwall inequality for  $\xi$ , we complete the proof of (4.19) by noting  $\xi(0) = 0$ .

We proceed the proof of (4.20) by starting with the following equation

$$\left(-\frac{\partial(p - p_h(u))}{\partial t}, q_h\right) + a_h(q_h, p - p_h(u)) = -(y - y_h(u), q_h), \quad \forall q_h \in V_h,$$

as the proof of (4.19). □

Now, we complete the a priori error estimate of the semi-discrete scheme by combining Lemmas 3–5 with triangle inequality.

**Theorem 1** *Let  $(y, u, p)$  and  $(y_h, u_h, p_h)$  be the solutions of (2.8a)–(2.8c) and (4.5a)–(4.5c), respectively. Suppose that the conditions of Proposition 1 and Lemma 4 are valid. Assume that the regularity condition (4.9) is also satisfied. Then, the following estimate holds*

$$\begin{aligned} &\|y - y_h\|_{L^\infty(0,T;L^2(\Omega))} + \|p - p_h\|_{L^\infty(0,T;L^2(\Omega))} + \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))} \\ &\leq Ch^2(\|P\|_{H^1(0,T;H^2(\mathcal{T}_h))} + \|y\|_{H^1(0,T;H^2(\mathcal{T}_h))}) + Ch_U^{3/2}. \end{aligned} \tag{4.24}$$

### 5 Fully-discrete formulation of optimal control problem

We use the standard backward Euler method to discretize the optimal control problem (2.1), (2.2a)–(2.2c) in time. Let  $N_T$  be a positive integer. The discrete time interval  $\bar{I} = [0, T]$  is defined as

$$0 = t_0 < t_1 < \dots < t_{N_T-1} < t_{N_T} = T$$

with size  $k_n = t_n - t_{n-1}$  for  $n = 1, \dots, N_T$  and  $k = \max_{n=1, \dots, N_T} k_n$ .

To prove the a priori error estimate of the fully-discrete scheme, we need the discrete time-dependent norm for  $1 \leq q < \infty$  by [9],

$$\|v\|_{L^q(0,T;L^2(\Omega))} = \left(\sum_{n=1}^{N_T} k_n \|v_n\|_{L^2(\Omega)}^q\right)^{1/q}. \tag{5.1}$$

Let  $f_{h,n}$  and  $y_{h,n}^d$  be approximations of the source function  $f_h$  and the desired state function  $y_h^d$  at time  $t_n$ . Then, the fully-discrete approximate scheme of the semi-discrete problem (3.8a), (3.8b) is

$$\begin{aligned} &\underset{u_{h,n} \in U_h^{ad}}{\text{minimize}} \quad \sum_{n=1}^{N_T} k_n \left( \frac{1}{2} \sum_{K \in \mathcal{T}_h} \|y_{h,n} - y_{h,n}^d\|_{L^2(K)}^2 + \frac{\alpha}{2} \sum_{K_U \in \mathcal{T}_h^U} \|u_{h,n}\|_{L^2(K_U)}^2 \right), \end{aligned} \tag{5.2a}$$

$$\begin{aligned} &\text{subject to} \quad \left(\frac{y_{h,n} - y_{h,n-1}}{k_n}, v\right) + a_h(y_{h,n}, v) + b_h(u_{h,n}, v) = (f_{h,n}, v) \quad \forall v \in V_h, \\ &y_{h,0}(x, 0) = y_h^0. \end{aligned} \tag{5.2b}$$

### 6 A priori error analysis of fully-discrete scheme

As for the semi-discrete scheme, we first give the stability result of the state variable at the following Lemma 6.

**Lemma 6** *Let  $y_{h,n}$  be the solution of (5.2b) and let  $c_0$  be a positive constant such that (2.4) holds. Then, there exists a positive constant  $C$  independent of  $h$  and  $k_m$  for  $m = 1, 2, \dots, N_T$  such that*

$$\begin{aligned} & \|y_{h,m}\|_{L^2(\Omega)}^2 + \sum_{n=1}^m k_n \left( \kappa \|y_{h,n}\|_{\varepsilon}^2 + \sum_K 2c_0 \|y_{h,n}\|_K^2 + \sum_K \|y_{h,n}\|_{\partial K^- \cap \Gamma^-}^2 \right) \\ & + \sum_{n=1}^m k_n \left( \sum_K \|y_{h,n} - y_{h,n}^e\|_{\partial K^- \setminus \Gamma^-}^2 + \sum_K \|y_{h,n}\|_{\partial K^+ \cap \Gamma^+}^2 \right) \\ & \leq C \|y_h^0\|_{L^2(\Omega)}^2 + C \sum_{n=1}^m k_n (\|f_{h,n}\|_{L^2(\Omega)}^2 + \|Bu_{h,n}\|_{L^2(\Omega)}^2). \end{aligned} \tag{6.1}$$

*Proof* Choose  $v_h = y_{h,n}$  in (6.2a). By using the algebraic inequality  $\frac{x^2 - y^2}{2} \leq (x - y)x, \forall x, y \in \mathbb{R}$  and following the steps in Lemma 1, we obtain the desired result. □

The minimization problem (5.2a), (5.2b) has at least one solution due to the boundedness of solution  $y_{h,n}$  as proven in Lemma 6. Then, the fully discretized control problem (5.2a), (5.2b) obtained by using the backward Euler method has a unique solution  $(y_{h,n}, u_{h,n}), n = 1, 2, \dots, N_T$ , and  $(y_{h,n}, u_{h,n}) \in Y_h \times U_h^{ad}, n = 1, 2, \dots, N_T$  is the solution of (5.2a), (5.2b) if and only if  $(y_{h,n}, u_{h,n}, p_{h,n-1}) \in Y_h \times U_h^{ad} \times Y_h$  is a unique solution of the following optimality system:

$$\left( \frac{y_{h,n} - y_{h,n-1}}{k_n}, v \right) + a_h(y_{h,n}, v) + b_h(u_{h,n}, v) = (f_{h,n}, v) \quad \forall v \in V_h, \tag{6.2a}$$

$$y_{h,0} = y_h^0, \quad n = 1, 2, \dots, N_T,$$

$$\left( \frac{p_{h,n-1} - p_{h,n}}{k_n}, q \right) + a_h(q, p_{h,n-1}) = -(y_{h,n} - y_{h,n}^d, q) \quad \forall q \in V_h, \tag{6.2b}$$

$$p_{h,T} = 0, \quad n = N_T, \dots, 2, 1,$$

$$(\omega u_{h,n} - B^* p_{h,n-1}, w - u_{h,n})_U \geq 0 \quad \forall w \in U_h^{ad}, n = 1, 2, \dots, N_T. \tag{6.2c}$$

The stability result of the adjoint variable for the fully-discrete scheme is also given at the following Lemma 7.

**Lemma 7** *Let  $p_{h,n}$  be the solution of (6.2b) and let  $c_0$  be a positive constant such that (2.4) holds. Then, there exists a positive constant  $C$  independent of  $h$  and  $k_m$  for*



$m = N_T - 1, \dots, 2, 1$  such that

$$\begin{aligned} & \|p_{h,m}\|_{L^2(\Omega)}^2 + \sum_{n=1}^m k_n \kappa \|p_{h,n-1}\|_{\varepsilon}^2 \\ & + \sum_{n=1}^m k_n \left( 2 \sum_K c_0 \|p_{h,n-1}\|_K^2 + \sum_K \|p_{h,n-1}\|_{\partial K^- \cap \Gamma^-}^2 \right) \\ & + \sum_{n=1}^m k_n \left( \sum_K \|p_{h,n-1} - p_{h,n-1}^e\|_{\partial K^+ \setminus \Gamma^+}^2 + \sum_K \|p_{h,n-1}\|_{\partial K^+ \cap \Gamma^+}^2 \right) \\ & \leq C \sum_{n=1}^m k_n \|y_{h,n} - y_{h,n}^d\|_{L^2(\Omega)}^2. \end{aligned} \tag{6.3}$$

Now, we derive the a priori error estimates of the fully-discrete scheme by introducing the following auxiliary equations as for the semi-discrete scheme. Let

$$(J'_h(u), v - u)_U = \sum_{n=1}^{N_T} k_n (\alpha u_n - B^* p_{h,n-1}(u), v_n - u_n)_U, \tag{6.4}$$

where  $p_{h,n-1}(u)$  is the solution of the following system

$$\begin{aligned} & \left( \frac{y_{h,n}(u) - y_{h,n-1}(u)}{k_n}, v \right) + a_h(y_{h,n}(u), v) + b_h(u_n, v) = (f_{h,n}, v) \quad \forall v \in V_h, \\ & y_{h,0}(u) = y_h^0, \quad n = 1, 2, \dots, N_T, \end{aligned} \tag{6.5a}$$

$$\begin{aligned} & \left( \frac{p_{h,n-1}(u) - p_{h,n}(u)}{k_n}, q \right) + a_h(q, p_{h,n-1}(u)) = -(y_{h,n}(u) - y_{h,n}^d, q) \quad \forall q \in V_h, \\ & p_{h,T}(u) = 0, \quad n = N_T, \dots, 2, 1. \end{aligned} \tag{6.5b}$$

For the simplicity, we use the following notations,

$$\begin{aligned} \zeta_n &= y_{h,n} - y_{h,n}(u), \quad n = 0, 1, \dots, N_T, \\ \chi_n &= p_{h,n} - p_{h,n}(u), \quad n = N_T, \dots, 1, 0. \end{aligned}$$

Firstly, we establish a connection between the approximation results  $(y_h, p_h)$  and the auxiliary solutions  $(y_h(u), p_h(u))$  as described at the following lemma.

**Lemma 8** *Let  $(y_h, p_h)$  and  $(y_h(u), p_h(u))$  be the solutions of (6.2a), (6.2b) and (6.5a), (6.5b), respectively. Then, there are positive constants  $C_1$  and  $C_2$  independent of  $h$  and  $k$  such that*

$$\| \|y_h - y_h(u)\| \|_{L^2(0,T;L^2(\Omega))} \leq C_1 \| \|u - u_h\| \|_{L^2(0,T;L^2(\Omega_U))}, \tag{6.6}$$

$$\| \| p_h - p_h(u) \| \|_{L^2(0,T;L^2(\Omega))} \leq C_2 \| \| u - u_h \| \|_{L^2(0,T;L^2(\Omega_U))}. \tag{6.7}$$

*Proof* We start the proof of (6.6) by subtracting (6.5a) from (6.2a) to obtain the following equality

$$\left( \frac{\zeta_n - \zeta_{n-1}}{k_n}, v_h \right) + a_h(\zeta_n, v_h) = -b_h(u_{h,n} - u_n, v_h).$$

By choosing  $v_h = \zeta_n$  and following the steps in Lemma 6, we obtain

$$\begin{aligned} & \frac{1}{2k_n} (\| \zeta_n \|_{L^2(\Omega)}^2 - \| \zeta_{n-1} \|_{L^2(\Omega)}^2) + \kappa \| \zeta_n \|_\varepsilon^2 + \sum_K c_0 \| \zeta_n \|_{L^2(K)}^2 + \frac{1}{2} \sum_K \| \zeta_n \|_{\partial K^- \cap \Gamma^-}^2 \\ & + \frac{1}{2} \sum_K \| \zeta_n - \zeta_n^e \|_{\partial K^- \setminus \Gamma^-}^2 + \frac{1}{2} \sum_K \| \zeta_n \|_{\partial K^+ \cap \Gamma^+}^2 \\ & \leq \frac{1}{2} \| u_{h,n} - u_n \|_{L^2(\Omega)}^2 + \frac{1}{2} \| \zeta_n \|_{L^2(\Omega)}^2. \end{aligned}$$

Multiplying the above inequality by  $2k_n$  and summing from  $n = 1$  to  $n = N_T$ , we derive

$$\begin{aligned} & (\| \zeta_{N_T} \|_{L^2(\Omega)}^2 - \| \zeta_0 \|_{L^2(\Omega)}^2) + 2 \sum_{n=1}^{N_T} k_n \left( \kappa \| \zeta_n \|_\varepsilon^2 + \sum_K c_0 \| \zeta_n \|_{L^2(K)}^2 \right) \\ & + \sum_{n=1}^{N_T} k_n \left( \sum_K \| \zeta_n \|_{\partial K^- \cap \Gamma^-}^2 + \sum_K \| \zeta_n - \zeta_n^e \|_{\partial K^- \setminus \Gamma^-}^2 + \sum_K \| \zeta_n \|_{\partial K^+ \cap \Gamma^+}^2 \right) \\ & \leq \sum_{n=1}^{N_T} k_n (\| u_{h,n} - u_n \|_{L^2(\Omega)}^2 + \| \zeta_n \|_{L^2(\Omega)}^2). \end{aligned}$$

Then, we apply discrete Gronwall’s inequality to the terms related to  $\zeta$  and use (4.3) which leads to the inequality  $\| \cdot \|_{L^2(\Omega)} \leq C \| \cdot \|_\varepsilon$  for some positive constant  $C$  and finally use the definition of the norm in (5.1) to obtain (6.6).

To show the second part of the Lemma 8, we subtract (6.5b) from (6.2b) to obtain

$$\left( \frac{\chi_{n-1} - \chi_n}{k_n}, q_h \right) + a_h(q_h, \chi_{n-1}) = -(\zeta_n, q_h).$$

By choosing  $q_h = \chi_{n-1}$  and proceeding as in the first part, we obtain the following inequality

$$\| \| p_h - p_h(u) \| \|_{L^2(0,T;L^2(\Omega))} \leq C \| \| y_h - y_h(u) \| \|_{L^2(0,T;L^2(\Omega))}.$$

This inequality gives us the desired result (6.7). □

To derive an estimate for the control  $u$  in the fully-discrete scheme, we use the discontinuous piecewise linear finite element space by following the approach in [29].

**Lemma 9** *Let  $(y, p, u)$  and  $(y_h, p_h, u_h)$  be the solutions of (2.8a)–(2.8c) and (6.2a)–(6.2c), respectively. Under the assumptions  $u \in L^2(0, T; W^{1,\infty}(\Omega_U))$ ,  $u|_{\Omega^*} \in L^2(0, T; H^2(\Omega^*))$ ,  $p \in L^2(0, T; W^{1,\infty}(\Omega))$ , we have*

$$\begin{aligned} \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))} &\leq C \left( k \left\| \frac{\partial p}{\partial t} \right\|_{L^2(0,T;L^2(\Omega))} + \|p_h(u) - p\|_{L^2(0,T;L^2(\Omega))} \right) \\ &\quad + Ch_U^{3/2}. \end{aligned} \tag{6.8}$$

*Proof* Let

$$(J'_h(u), v - u)_U = \sum_{n=1}^{N_T} k_n (\alpha u_n - B^* p_{h,n-1}(u), v_n - u_n)_U,$$

where  $p_{h,n-1}(u)$  is the solution of the auxiliary solution (6.5b). Then,

$$\begin{aligned} (J'_h(v) - J'_h(u), v - u)_U &= \sum_{n=1}^{N_T} k_n (\alpha v_n - \alpha u_n, v_n - u_n)_U \\ &\quad + \sum_{n=1}^{N_T} k_n (B^* p_{h,n-1}(u) - B^* p_{h,n-1}(v), v_n - u_n)_U \\ &= \alpha \|v - u\|_{L^2(0,T;L^2(\Omega_U))}^2 \\ &\quad + \sum_{n=1}^{N_T} k_n (B^* p_{h,n-1}(u) - B^* p_{h,n-1}(v), v_n - u_n)_U. \end{aligned}$$

By using the auxiliary solutions (6.5a), (6.5b) as done for the semi-discrete scheme, we obtain

$$\sum_{n=1}^{N_T} k_n (B^* p_{h,n-1}(u) - B^* p_{h,n-1}(v), v_n - u_n)_U \geq 0.$$

Hence,

$$(J'_h(v) - J'_h(u), v - u)_U \geq \alpha \|v - u\|_{L^2(0,T;L^2(\Omega_U))}^2. \tag{6.9}$$

Set  $\Pi u_n \in U_h$  be the standard Lagrange interpolation of  $u$  at time  $t_n$  such that  $\Pi u_n(x) = u_n(x)$  for all vertices  $x$ . Then,  $\Pi u_n$  belongs to  $U_{ad}^h$  at time  $t_n$ . With the help of the inequalities (6.9), (2.8c), (6.2c) and an approximation of  $u$  at time  $t_n$ , i.e.,  $\Pi u_n$ , we obtain

$$\begin{aligned} \alpha \|u - u_h\|_{L^2(0,T;L^2(\Omega_U))}^2 &\leq (J'_h(u) - J'_h(u_h), u - u_h)_U \\ &= \sum_{n=1}^{N_T} k_n (\alpha u_n - B^* p_n, u_n - u_{h,n})_U + \sum_{n=1}^{N_T} k_n (B^* p_n - B^* p_{h,n-1}(u), u_n - u_{h,n})_U \end{aligned}$$

$$\begin{aligned}
 & + \sum_{n=1}^{N_T} k_n (\alpha u_{h,n} - B^* p_{h,n-1}, \Pi u_n - u_n)_U \\
 & + \sum_{n=1}^{N_T} k_n (\alpha u_{h,n} - B^* p_{h,n-1}(u), u_{h,n} - \Pi u_n)_U \\
 \leq & \sum_{n=1}^{N_T} k_n (B^* p_n - B^* p_{h,n-1}(u), u_n - u_{h,n})_U \\
 & + \sum_{n=1}^{N_T} k_n (\alpha u_{h,n} - B^* p_{h,n-1}, \Pi u_n - u_n)_U \\
 = & \underbrace{\sum_{n=1}^{N_T} k_n (B^* p_n - B^* p_{n-1}, u_n - u_{h,n})_U}_{T_1} \\
 & + \underbrace{\sum_{n=1}^{N_T} k_n (B^* p_{n-1} - B^* p_{h,n-1}(u), u_n - u_{h,n})_U}_{T_2} \\
 & + \underbrace{\sum_{n=1}^{N_T} k_n (\alpha u_{h,n} - B^* p_{h,n-1}, \Pi u_n - u_n)_U}_{T_3}. \tag{6.10}
 \end{aligned}$$

The following estimates of  $T_1$  and  $T_2$  are derived by using Young’s inequality,

$$\begin{aligned}
 T_1 & \leq C_1 \sum_{n=1}^{N_T} k_n \|p_n - p_{n-1}\|_{L^2(\Omega)}^2 + C_2 \sum_{n=1}^{N_T} k_n \|u_n - u_{h,n}\|_{L^2(\Omega_U)}^2 \\
 & \leq C_1 k^2 \left\| \frac{\partial p}{\partial t} \right\|_{L^2(0,T;L^2(\Omega))}^2 + C_2 \| \|u - u_h \| \|_{L^2(0,T;L^2(\Omega_U))}^2, \\
 T_2 & \leq C_1 \sum_{n=1}^{N_T} k_n \|p_{n-1} - p_{h,n-1}(u)\|_{L^2(\Omega)}^2 + C_2 \sum_{n=1}^{N_T} k_n \|u_n - u_{h,n}\|_{L^2(\Omega_U)}^2 \\
 & \leq C_1 \| \|p - p_h(u) \| \|_{L^2(0,T;L^2(\Omega))}^2 + C_2 \| \|u - u_h \| \|_{L^2(0,T;L^2(\Omega_U))}^2.
 \end{aligned}$$

By considering the discontinuous piecewise linear finite element space for the control  $u$  and following the steps in Lemma 4, we obtain

$$\begin{aligned}
 T_3 & \leq C k^2 \left\| \frac{\partial p}{\partial t} \right\|_{L^2(0,T;L^2(\Omega))}^2 + C \| \|p_h(u) - p \| \|_{L^2(0,T;L^2(\Omega))}^2 \\
 & \quad + \| \|u - u_h \| \|_{L^2(0,T;L^2(\Omega_U))}^2 + C h_U^3.
 \end{aligned}$$

Summing up the estimates of  $T_1 - T_3$ , we complete the proof. □

Now, we establish the connection between the exact solutions and auxiliary solutions.

**Lemma 10** *Let  $(y, p)$  be the solutions of (2.8a), (2.8b) and  $(y_h(u), p_h(u))$  be the solutions of (6.5a), (6.5b), respectively. Suppose that the conditions of Proposition 1 and Lemma 9 are valid. Then, the following estimates hold*

$$\begin{aligned} \| \|y - y_h(u)\| \|_{L^2(0,T;L^2(\Omega))} &\leq Ch^2 \| \|y\| \|_{L^2(0,T;H^2(\mathcal{T}_h))} \\ &\quad + Ch^2 \left\| \frac{\partial y}{\partial t} \right\|_{L^2(0,T;H^2(\mathcal{T}_h))} + Ck \left\| \frac{\partial^2 y}{\partial t^2} \right\|_{L^2(0,T;L^2(\Omega))}, \end{aligned} \tag{6.11}$$

and

$$\begin{aligned} &\| \|p - p_h(u)\| \|_{L^2(0,T;L^2(\Omega))} \\ &\leq Ch^2 \sum_{v=y,p} \| \|v\| \|_{L^2(0,T;H^2(\mathcal{T}_h))} + Ch^2 \sum_{v=y,p} \left\| \frac{\partial v}{\partial t} \right\|_{L^2(0,T;H^2(\mathcal{T}_h))} \\ &\quad + Ck \sum_{v=y,p} \left\| \frac{\partial^2 v}{\partial t^2} \right\|_{L^2(0,T;L^2(\Omega))} + Ck \sum_{v=y,p} \left\| \frac{\partial v}{\partial t} \right\|_{L^2(0,T;L^2(\Omega))}. \end{aligned} \tag{6.12}$$

*Proof* Here, we only prove (6.11) since both cases follow the same procedure. We start with the following equation obtained by using (6.5a) and (2.8a)

$$(\partial_t y_n, v_h) + a_h(y_n, v_h) - \left( \frac{y_{h,n}(u) - y_{h,n-1}(u)}{k_n}, v_h \right) - a_h(y_{h,n}(u), v_h) = 0. \tag{6.13}$$

We decompose  $y - y_h(u)$  as

$$y_n - y_{h,n}(u) = y_n - \tilde{y}_n - (y_h(u) - \tilde{y}_n) = \eta_n - \xi_n,$$

where  $\tilde{y}$  is an elliptic projection of  $y$ . We only need to estimate  $\xi_n$  since the estimate of  $\eta_n$  is given in (4.8b). Hence, we write (6.13) as

$$\begin{aligned} \left( \frac{\xi_n - \xi_{n-1}}{k_n}, v_h \right) + a_h(\xi_n, v_h) &= \left( \frac{\partial y_n}{\partial t} - \frac{y_n - y_{n-1}}{k_n}, v_h \right) + \left( \frac{\eta_n - \eta_{n-1}}{k_n}, v_h \right) \\ &\quad + a_h(\eta_n, v_h). \end{aligned}$$

By choosing  $v_h = \xi_n$  and applying the steps in Lemma 5 to bound the terms on the inflow and outflow boundaries, we obtain

$$\begin{aligned} &\| \xi_{N_T} \|_{L^2(\Omega)}^2 - \| \xi_0 \|_{L^2(\Omega)}^2 + \frac{3\kappa}{4} \sum_{n=1}^{N_T} k_n \| \xi_n \|_{\varepsilon}^2 + 2 \sum_{n=1}^{N_T} k_n \sum_K c_0 \| \xi_n \|_{L^2(K)}^2 \\ &\quad + \frac{1}{2} \sum_{n=1}^{N_T} k_n \left( 2 \sum_K \| \xi_n \|_{\partial K \cap \Gamma^-}^2 + \sum_K \| \xi_n - \xi_n^e \|_{\partial K \setminus \Gamma^-}^2 + \sum_K \| \xi_n \|_{\partial K \cap \Gamma^+}^2 \right) \end{aligned}$$

$$\begin{aligned} &\leq 2 \sum_{n=1}^{N_T} k_n \left( \sum_K \|\eta_n^e\|_{\partial K \setminus \Gamma^-}^2 + \sum_K \|\eta_n\|_{\partial K^+ \cap \Gamma^+}^2 \right) \\ &\quad + C \sum_{n=1}^{N_T} k_n \|\xi_n\|_{L^2(\Omega)}^2 + Ck_n^2 \int_0^T \left\| \frac{\partial^2 y}{\partial t^2} \right\|_{L^2(\Omega)}^2 dt + Ch^4 \int_0^T \left\| \frac{\partial y}{\partial t} \right\|_{H^2(\mathcal{T}_h)} dt. \end{aligned}$$

Then, applying Gronwall’s inequality to the terms related to  $\xi$  in above inequality, the desired result (6.11) is obtained.  $\square$

Now, we finalize the a priori error estimate of the fully-discrete scheme by combining Lemmas 8–10 with the triangle inequality.

**Theorem 2** *Let  $(y, p, u)$  be the solutions of (2.8a)–(2.8c) and  $(y_h, p_h, u_h)$  be the solutions of (6.2a)–(6.2c), respectively. Suppose that the conditions of Proposition 1 and Lemma 9 are valid. Further, the regularity condition (4.9) is satisfied. Then, the following estimate holds*

$$\begin{aligned} &\| \|y - y_h\| \|_{L^2(0,T;L^2(\Omega))} + \| \|p - p_h\| \|_{L^2(0,T;L^2(\Omega))} + \| \|u - u_h\| \|_{L^2(0,T;L^2(\Omega_U))} \\ &\leq Ch^2 \sum_{v=y,p} \| \|v\| \|_{L^2(0,T;H^2(\mathcal{T}_h))} + Ch^2 \sum_{v=y,p} \left\| \frac{\partial v}{\partial t} \right\|_{L^2(0,T;H^2(\mathcal{T}_h))} + Ch_U^{3/2} \\ &\quad + Ck \sum_{v=y,p} \left\| \frac{\partial^2 v}{\partial t^2} \right\|_{L^2(0,T;L^2(\Omega))} + Ck \sum_{v=y,p,y_d} \left\| \frac{\partial v}{\partial t} \right\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

### 7 Numerical results

In this section, we present numerical results for the unsteady control constrained optimal control problems governed by the convection diffusion equation (2.1), (2.2a)–(2.2c). We take  $\Omega = \Omega_U$  and  $B = I$ . To do this, we consider the problem in [10] with the following parameters

$$\begin{aligned} Q &= (0, 1] \times \Omega, \quad \Omega = (0, 1)^2, \quad \varepsilon = 10^{-5}, \quad \beta = (1, 0)^T, \\ r &= 0, \quad \alpha = 1 \quad \text{and} \quad u_a = 0. \end{aligned}$$

The source function  $f$ , the desired state  $y_d$  and the initial condition  $y_0$  are computed from (2.8a)–(2.8c) using the following exact solutions of the state, adjoint and control, respectively,

$$\begin{aligned} y(x, t) &= \exp(-t) \sin(2\pi x_1) \sin(2\pi x_2), \\ p(x, t) &= \exp(-t)(1 - t) \sin(2\pi x_1) \sin(2\pi x_2), \\ u(x, t) &= \max(0, p). \end{aligned}$$

In our numerical example, the control variable is only bounded from below, i.e.,  $u_a = 0$ . The state, the adjoint, and the control variables are discretized by using the

piecewise linear polynomials, i.e.,  $(x, y, 1 - x - y)$ . Discretized control constraint problems are solved by the primal dual active set (PDAS) algorithm as a semi-smooth Newton step, see, e.g., [2, 3]. The algorithm is terminated when two consecutive active sets coincide. The initial guess for the control variable is taken as equal to zero for all discretization levels.

There are two approaches to solve the optimization problem (2.1), (2.2a)–(2.2c) numerically, i.e., the *discretize-then-optimize* (DO) and the *optimize-then-discretize* (OD). It is desirable that both approaches lead to the same discrete optimality system. In the DO approach, one first discretizes the optimal control problem with the objective function (2.1) and the state equation (2.2a)–(2.2c) and then form the optimality system. In the OD approach, one first derives the optimality conditions consisting of the state and adjoint PDEs and the algebraic equation that links the control and the adjoint variable. Afterwards the infinite dimension optimality system is discretized to form the optimality system. It is known that the two approaches for optimal control problems are governed by convection diffusion PDEs lead the same linear optimality systems for some discretization schemes. Although the commutative property is preserved for the steady case using upwind SIPG methods [15, 28], it is not preserved for the unsteady problems when the backward Euler discretization is used in time. A straightforward time discretization will usually not lead to the same optimality system for the DO and OD approaches; the initial condition of the adjoint PDE makes the difference between the DO and OD approaches. It was shown in [23] for the Stokes equation. By adjusting the time discretization for the initial condition of the forward problem, it was possible to show that the OD and DO approaches commute. The same technique was also used in [11]. Further, both approaches lead to the same discrete optimality system when discontinuous Galerkin discretization dG(0) is used in time [14].

Tables 1 and 2 show the errors and converge rates with respect to the discrete time-dependent norm (5.1) by fixing mesh size in space for OD and DO, respectively. The order of convergence for time is  $k$  as expected from the a priori error estimates. For fixed time steps, the errors and convergence rates in terms of space are given at Tables 3 and 4 for OD and DO, respectively. Again, the results confirm the a priori error estimates. Although the rate of the control is  $h^{3/2}$  theoretically, it is observed to be  $h^2$  since there is no kink for the control. It means that our initial grid aligns with the points where the lower bound of control and the value of adjoint coincide, i.e.,  $x_1 = x_2 = 0.5$ .

Figure 1 show the computed solutions of the state and adjoint at  $t = 0.5$  with  $h/\sqrt{2} = 1/32$ ,  $k = 1/128$  by using the OD approach. In addition, the exact and computed solutions of the control is given at Fig. 2 for  $t = 0.5$  with  $h/\sqrt{2} = 1/32$ ,  $k = 1/128$  by using the OD approach.

Discontinuous Galerkin (DG) discretizations exhibit a better convergence behavior for convection dominated optimal control problems since errors in boundary layers are not propagated into the entire domain [16]. Therefore, DG discretization with mesh adaptivity presents better results with respect to stabilized finite element methods such as in [26, 27] for steady convection dominated optimal control problems. Although our example is smooth, we can still see the effect of the DG discretization. The same example was solved in [10] with characteristic finite element method.

**Table 1**  $h/\sqrt{2} = 1/32$  via the OD approach

$k$	$\ y - y_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ p - p_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ u - u_h\ _{L^2(0,1;L^2(\Omega))}$	Rate
$\frac{1}{2}$	1.66e-2	–	4.82e-2	–	1.29e-2	–
$\frac{1}{4}$	9.47e-3	0.81	2.22e-2	1.12	9.98e-3	0.37
$\frac{1}{8}$	5.22e-3	0.86	1.06e-2	1.07	6.00e-3	0.73
$\frac{1}{16}$	2.80e-3	0.90	5.21e-3	1.03	3.28e-3	0.87
$\frac{1}{32}$	1.52e-3	0.89	2.60e-3	1.00	1.73e-3	0.92
$\frac{1}{64}$	9.03e-4	0.75	1.32e-3	0.97	9.28e-4	0.90

**Table 2**  $h/\sqrt{2} = 1/32$  via the DO approach

$k$	$\ y - y_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ p - p_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ u - u_h\ _{L^2(0,1;L^2(\Omega))}$	Rate
$\frac{1}{2}$	1.61e-2	–	3.58e-2	–	2.44e-2	–
$\frac{1}{4}$	8.85e-3	0.86	2.44e-2	0.56	1.68e-2	0.54
$\frac{1}{8}$	4.71e-3	0.91	1.46e-2	0.74	1.01e-2	0.74
$\frac{1}{16}$	2.49e-3	0.92	8.20e-3	0.83	5.67e-3	0.83
$\frac{1}{32}$	1.36e-3	0.87	4.42e-3	0.89	3.06e-3	0.89
$\frac{1}{64}$	8.45e-4	0.69	2.33e-3	0.93	1.62e-3	0.92

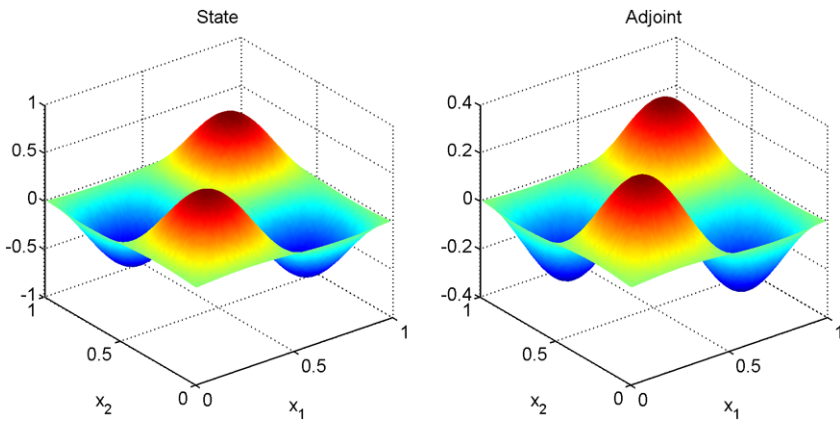
**Table 3**  $k = 1/2048$  via the OD approach

$h/\sqrt{2}$	$\ y - y_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ p - p_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ u - u_h\ _{L^2(0,1;L^2(\Omega))}$	Rate
$\frac{1}{4}$	4.41e-2	–	2.44e-2	–	2.49e-2	–
$\frac{1}{8}$	1.02e-2	2.11	5.68e-3	2.10	6.82e-3	1.87
$\frac{1}{16}$	2.45e-3	2.06	1.40e-3	2.02	1.65e-3	2.05
$\frac{1}{32}$	5.83e-4	2.07	3.44e-4	2.03	3.56e-4	2.20

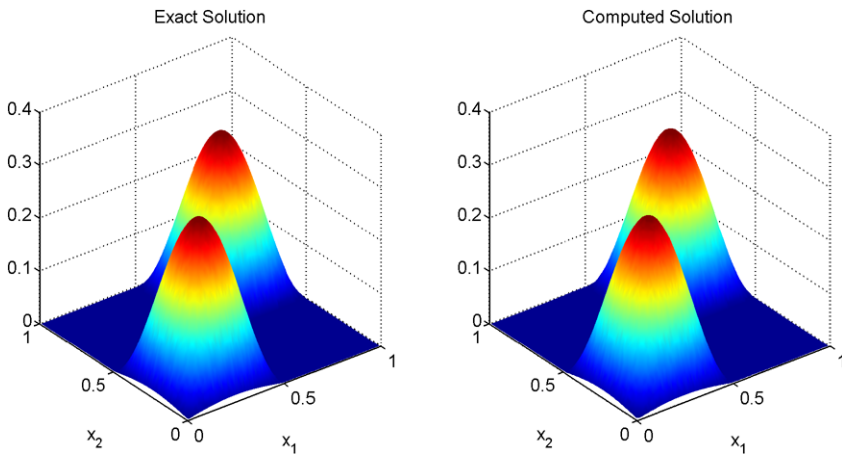
**Table 4**  $k = 1/2048$  via the DO approach

$h/\sqrt{2}$	$\ y - y_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ p - p_h\ _{L^2(0,1;L^2(\Omega))}$	Rate	$\ u - u_h\ _{L^2(0,1;L^2(\Omega))}$	Rate
$\frac{1}{4}$	4.41e-2	–	2.44e-2	–	2.49e-2	–
$\frac{1}{8}$	1.02e-2	2.11	5.68e-3	2.10	6.82e-3	1.87
$\frac{1}{16}$	2.45e-3	2.06	1.41e-3	2.01	1.63e-3	2.06
$\frac{1}{32}$	5.84e-4	2.07	3.49e-4	2.01	3.59e-4	2.19





**Fig. 1** The computed solutions of the state (*left*) and adjoint (*right*) at  $t = 0.5$  with  $h/\sqrt{2} = 1/32$ ,  $k = 1/128$  by using the OD approach



**Fig. 2** The exact and computed solutions of the control at  $t = 0.5$  with  $h/\sqrt{2} = 1/32$ ,  $k = 1/128$  by using the OD approach

Comparing the results in Table 1 with the ones obtained in [10], it turns out that the upwind SIPG discretization yields more accurate results.

## 8 Conclusions

We have derived a priori error estimates for the optimal control problems governed by the unsteady convection diffusion equation using the upwind SIPG discretization in space and the standard backward Euler in time. Although the OD and DO approaches lead two different optimality systems under the backward Euler discretization, there is no remarkable differences in numerical results and convergence rates. Numerical experiments are given to confirm the theoretical results. With adaptive meshes, the

DG methods resolve the boundary and/or interior layers for convection dominated problems better and more efficient than the continuous finite elements in [26, 27] for steady optimal control problems. This issue will be addressed in the coming work with space-time adaptivity for unsteady convection dominated optimal control problem.

**Acknowledgements** The authors wish to thank the referees for their helpful suggestions and comments. This research was supported by the Middle East Technical University Scientific Research Fund (Project: BAP-07-05-2012-102).

## References

1. Becker, R., Vexler, B.: Optimal control of the convection-diffusion equation using stabilized finite element methods. *Numer. Math.* **106**, 349–367 (2007)
2. Bergounioux, M., Ito, K., Kunisch, K.: Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37**(4), 1176–1194 (1999)
3. Bergounioux, M., Haddou, M., Hintermueller, M., Kunisch, K.: A comparison of interior–point methods and a Moreau–Yosida based active set strategy for constrained optimal control problems. *SIAM J. Optim.* **11**, 495–521 (2001)
4. Brenner, S.: Poincaré–Friedrichs inequalities for piecewise  $H^1$  functions. *SIAM J. Numer. Anal.* **41**(1), 306–324 (2003)
5. Cangiani, A., Chapman, J., Georgoulis, E.H., Jensen, M.: On Local Super-Penalization of Interior Penalty Discontinuous Galerkin Methods. *CoRR* [1205.5672](https://arxiv.org/abs/1205.5672) [abs] (2012)
6. Collis, S.S., Heinkenschloss, M.: Analysis of the streamline upwind/Petrov Galerkin method applied to the solution of optimal control problems. Tech. Rep. TR02–01, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005-1892 (2002)
7. Evans, L.C.: *Partial Differential Equations*. Grad. Stud. Math., vol. 19. AMS, Providence (2002)
8. Feistauer, M., Švadlenka, K.: Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems. *J. Numer. Math.* **12**, 97–117 (2004)
9. Fu, H.: A characteristic finite element method for optimal control problems governed by convection-diffusion equations. *J. Comput. Appl. Math.* **235**, 825–836 (2010)
10. Fu, H., Rui, H.: A priori error estimates for optimal control problems governed by transient advection-diffusion equations. *J. Sci. Comput.* **38**(3), 290–315 (2009)
11. Hinze, M.K., Turek, S.: A Space-Time Multigrid Method for Optimal Flow Control. *International Series of Numerical Mathematics*, vol. 160 (2011)
12. Hinze, M., Yan, N., Zhou, Z.: Variational discretization for optimal control governed by convection dominated diffusion equations. *J. Comput. Math.* **27**, 237–253 (2009)
13. Houston, P., Schwab, C., Süli, E.: Discontinuous hp-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.* **39**(6), 2133–2163 (2002)
14. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in  $L^2$  for a class of evolution equations. *SIAM J. Control Optim.* **46**(5), 1726–1753 (2007)
15. Leykekhman, D.: Investigation of commutative properties of discontinuous Galerkin methods in PDE constrained optimal control problems. *J. Sci. Comput.* **53**, 483–511 (2012)
16. Leykekhman, D., Heinkenschloss, M.: Local error analysis of discontinuous Galerkin methods for advection-dominated elliptic linear-quadratic optimal control problems. *SIAM J. Numer. Anal.* **50**, 2012–2038 (2012)
17. Li, R., Liu, W., Ma, H., Tang, T.: Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optim.* **41**, 1321–1349 (2002)
18. Lions, J.L.: *Optimal Control of Systems Governed by Partial Differential Equations*. Springer, Berlin (1971)
19. Meidner, D., Vexler, B.: A priori error estimates for space-time finite element discretization of parabolic optimal control problems. II. Problems with control constraints. *SIAM J. Control Optim.* **47**(3), 1301–1329 (2008)
20. 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)

21. Rivière, B.: *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. Theory and Implementation*. *Frontiers in Applied Mathematics*, vol. 35. Society for Industrial and Applied Mathematics, Philadelphia (2008)
22. Schötzau, D., Zhu, L.: A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations. *Appl. Numer. Math.* **59**, 2236–2255 (2009)
23. Stoll, M., Wathen, A.: All-at-once solution of time-dependent Stokes control. *J. Comput. Phys.* **232**, 498–515 (2013)
24. Sun, T.: Discontinuous Galerkin finite element method with interior penalties for convection diffusion optimal control problem. *Int. J. Numer. Anal. Model.* **7**, 87–107 (2010)
25. Yan, N., Zhou, Z.: A priori and a posteriori error analysis of edge stabilization Galerkin method for the optimal control problem governed by convection-dominated diffusion equation. *J. Comput. Appl. Math.* **223**, 198–217 (2009)
26. Yücel, H., Karasözen, B.: Adaptive Symmetric Interior Penalty Galerkin (SIPG) method for optimal control of convection diffusion equations with control constraints (2013). *Optimization (electronic)*
27. Yücel, H., Heinkenschloss, M., Karasözen, B.: An adaptive discontinuous Galerkin method for convection dominated distributed optimal control problems. *Tech. Rep.*, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892 (2012)
28. Yücel, H., Heinkenschloss, M., Karasözen, B.: Distributed optimal control of diffusion-convection-reaction equations using discontinuous Galerkin methods. In: *Numerical Mathematics and Advanced Applications*, vol. 2011, pp. 389–397. Springer, Berlin (2013)
29. Zhou, Z., Yan, N.: The local discontinuous Galerkin method for optimal control problem governed by convection diffusion equations. *Int. J. Numer. Anal. Model.* **7**, 681–699 (2010)