INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 13, Number 4, Pages 545–568

ON DISCONTINUOUS FINITE VOLUME APPROXIMATIONS FOR SEMILINEAR PARABOLIC OPTIMAL CONTROL PROBLEMS

RUCHI SANDILYA AND SARVESH KUMAR

Abstract. In this article, we discuss and analyze discontinuous finite volume approximations of the distributed optimal control problems governed by a class of semilinear parabolic partial differential equations with control constraints. For the spatial discretization of the state and costate variables, piecewise linear elements are used and an implicit finite difference scheme is used for time derivatives; whereas, for the approximation of the control variable, three different strategies are used: variational discretization, piecewise constant and piecewise linear discretization. *A priori* error estimates (for these three approaches) in suitable L^2-norm are derived for state, co-state and control variables. Numerical experiments are presented in order to assure the accuracy and rate of the convergence of the proposed scheme.

Key words. Semilinear parabolic optimal control problems, variational discretization, piecewise constant and piecewise linear discretization, discontinuous finite volume methods, *a priori* error estimates, numerical experiments.

1. Introduction

1.1. Scope. The purpose of this paper is to introduce discontinuous finite volume methods for the approximations of control, state and co-state variables involved in a semilinear parabolic optimal control problems. As it is well known that optimal control problems governed by a class of partial differential equations (introduced in [20], [34]) have various applications in scientific and engineering-related problems. For instance, heat conduction, diffusion, electromagnetic waves, fluid flows, freezing processes, and many other physical phenomena can be put forward as models based on partial differential equations. In particular, parabolic optimal control problems are used in describing a controlled heat transfer process for optimal cooling of steel profiles. The optimization of semilinear heat equations represent mathematical model for many physical applications, e.g. laser hardening, welding of steel, laser thermotherapy (used for cancer treatment) etc.

Due to the computational simplicity, efficiency and robustness of finite element methods, these methods are extensively employed for the approximation of optimal control problems. For instance, the finite element error analysis for elliptic optimal control problems has been established in [8, 9, 13, 32, 35] and references therein. In addition to that finite element approximations for parabolic optimal control problems have been discussed in [26, 27, 33, 36] and references cited in these articles. In most of these articles, the state and costate variables were approximated by conforming (continuous) finite element methods in which piecewise linear polynomials are used and control variable by piecewise constant or piecewise linear polynomials. For control variable, the rate of convergence is of $\mathcal{O}(h)$ and $\mathcal{O}(h^{3/2})$ for piecewise constant and piecewise linear discretization, respectively. For discretization of the

Received by the editors January 25, 2015 and, in revised form, November 24, 2015.
1991 *Mathematics Subject Classification.* 65Q05.

control which is the primary variable, a variational approach is proposed by Hinze in which control set is not discretized explicitly but discretized by a projection (to be defined later) and obtained improved convergence of $\mathcal{O}(h^2)$, for more details, we refer to [14].

The typical inter-element continuity criteria which is usually imposed on finite dimensional trial spaces involved in conforming and even nonconforming finite element methods in order to make the resulting system well-posed is no longer required for discontinuous Galerkin (DG) methods. Apart from this, other intrinsic attractive features of DG methods are: suitability for local mesh adaptivity, element-wise conservative, they allow high degree polynomials in different elements and can easily handle non-standard boundary conditions. For more details regarding DG methods, we refer to [1, 2, 30, 31] and references therein. In the context of control problems, DG methods have also been employed for parabolic optimal control problems, for instance see [26, 27, 29].

On the other hand, finite volume element (FVE) methods can be considered as Petrov-Galerkin methods in which the finite dimensional trial space consists of piecewise linear polynomials and piecewise constant functions are used in the test space. We can expect the computational advantages of FVE methods over finite element methods, as the test space associated with the dual grid is piecewise constant. In addition, the desirable feature of FVE methods is conservation of a quantity of interest, e.g., mass, momentum or energy. Due to this property of local conservation, finite volume element methods are widely used in computational fluid dynamics. However, the low regularity used in the test space, demands high regularity on given data or exact solution in order to achieve optimal L^2-estimates. For instance, for non-homogeneous elliptic problems, derivation of optimal L^2-estimates requires either an exact solution in H^3 or a source term globally in H^1 (see e.g. [12]). For more details and advantages of FVE methods, kindly see the early work [7, 10] and the recent review [19]. Recently, FVE methods have been employed in [24, 25] for the approximation of the state and costate variables appeared in linear elliptic and parabolic problems. In these articles, for discretization of the control variable, a variational discretization approach is used and optimal order of convergence has been shown.

In order to make use of desirable properties of DG methods and FVE methods, we will focus on a hybrid scheme discontinuous finite volume methods (DFVM) for the approximation of the distributed parabolic semilinear optimal control problems. These methods were originally introduced by [38] for elliptic problems and later with some modifications these methods were applied to elliptic, Stokes and parabolic problems and fluid flow problems, see [4, 6, 16, 17, 22, 37, 40, 39]. Recently, in [18] Kumar proposed a stabilized DFVM formulation for more general Stokes problems. However, up-to to our knowledge, there are hardly any results available on DFVM for the approximation of semilinear parabolic optimal control problems. Therefore, in this article an attempt has been made to introduce a fully discrete discontinuous finite volume methods for the approximation of the parabolic control problems. In addition, we use three different approaches: variational discretization (introduced in [14]), piecewise linear and piecewise constant discretization to approximate the control.

For the solvability of the optimal control problems, in literature, there are two different approaches: one is *discretize-then-optimize* and another one is *optimize-then-discretize*. In the *discretize-then-optimize* approach, one first discretizes the continuous problem and then accordingly derive for the optimality conditions; whereas, in the *optimize-then-discretize* approach, optimality condition on the continuous level is formulated first and then discretized. These two approaches coincide provided the discrete formulation is symmetric; however, this may not be true if the formulation is not symmetric, for more details we refer to [3] and references therein. We would like to mention that in general, FVE formulation is not symmetric (except if the matrix is constant) even if the coefficient matrix is symmetric, see [15]. In [24, 25], authors have used *optimize-then-discretize* approach together with FVE methods to approximate elliptic and parabolic control problems. We stress that our resulted DFVE scheme also leads to a nonsymmetric formulation (see Section 2), and therefore, in view of the articles [24, 25], in this article we have also opted for *optimize-then-discretize* technique.

We have organized this paper in the following manner. The remaining part of this section deals with primary notations used for Sobolev spaces and statement of the governing problems. Section 2 deals with finite dimensional formulation of the proposed control problems with details of DFV formulation for state and co-state variables and discretization approaches for control variable. This section also recalls the optimality conditions and some primary auxiliary results required for subsequent sections. In Section 3, we derive *a priori* error estimates in suitable L^2-norm for state, co-state and control variables. In Section 4, we present some numerical experiments to justify the convergence rates derived in Section 3. Finally, based on computational and theoretical results, some concluding remarks are made in Section 5.

Notations. In this article, we denote $\Omega \subset \mathbb{R}^2$ as a bounded convex polygonal domain with boundary $\partial\Omega$. Also, we adopt the standard notations for the Lebesgue spaces $L^p(\Omega)$ and the Sobolev Spaces $H^s(\Omega)$ defined over Ω with associated norms $\|\cdot\|_{s,\Omega}$ and seminorms $|\cdot|_{s,\Omega}$. Further, as usual, we write $H^0(\Omega) := L^2(\Omega)$ and for simplicity we drop Ω whenever its possible. Throughout this article, C denotes a generic positive constant independent from the mesh size h (to be defined in the next section) but may depend on the size of Ω and can take different values at different places. In addition we denote by $L^p(0,T;H^s(\Omega))$, $1 \le p,\ q \le \infty,\ s \ge 0$, the space of functions $\psi(t) : [0,T] \longrightarrow H^s(\Omega)$ such that $\|\psi(t)\|_{s,p,\Omega} \in L^p(0,T)$ with the following norm

$$\|\psi\|_{L^p([0,T];H^s(\Omega))} := \left(\int_0^T \|\psi\|_s^p\right)^{1/p} \quad s \in [1,\infty).$$

Governing equations. Keeping in mind the applications (mentioned earlier) of parabolic optimal control problems, we are interested in finding the numerical solution of the following semilinear parabolic optimal control problem: For a given desired state y_d and data f, find the state variable y and the control variable u satisfying

$$(1)\quad \min_{u(t,x)\in U_{ad}} J(y,u) := \frac{1}{2}\int_0^T\int_\Omega (y(t,x)-y_d(t,x))^2 dxdt + \frac{\alpha}{2}\int_0^T\int_\Omega u(t,x)^2 dxdt,$$

subject to

$$(2)\quad \begin{aligned} &\partial_t y(t,x) - \nabla\cdot(A\nabla y(t,x)) + \phi(y(t,x)) \\ &= Bu(t,x) + f(t,x), \quad in \quad (0,T)\times\Omega, \end{aligned}$$

$$
\begin{aligned}
(3) \qquad y(t,x) &= 0, \quad on \quad (0,T)\times\partial\Omega,\\
(4) \qquad y(0,x) &= y_0(x), \quad x\in\Omega.
\end{aligned}
$$

Here, $\alpha > 0$ is the regularization parameter, B is a bounded linear operator and $A = (a_{ij}(x))_{2\times 2}$ denotes a real valued, symmetric and uniformly positive definite matrix in Ω, i.e., there exists a positive constant α_0 such that $\xi^T A\xi \geq \alpha_0 \xi^T \xi, \quad \forall \xi \in \mathbf{R}^2$. The space of admissible controls U_{ad} is defined by

(5)
$$U_{ad} := \{u(t,x) \in L^2(I; L^2(\Omega)) : \ a \leq u(t,x) \leq b, \ \ a.e.\,(t,x) \in (0,T)\times\Omega; \ a, b \in \mathbb{R}; \}.$$

In addition for our analysis, we require the following assumptions on the given data: we assume that the desired state y_d and the source term $f \in L^2(I; L^2(\Omega))$ or $L^2(I; H^1(\Omega))$ with $I = (0,T)$. Also, for any $l > 0$ we have $\phi(\cdot) \in W^{2,\infty}(-l,l)$; $\phi'(y) \in L^2(\Omega)$ and $\phi'(y) \geq 0$ for $y \in L^2(I; H_0^1(\Omega))$.
Moreover, with the help of control-to-state mapping G (as introduced in [29]) with $G(u) = y$, the above problem (1)-(4) can be reduced to:

$$(6) \qquad \min_{u\in U_{ad}} j(u) := \min_{u\in U_{ad}} J(G(u), u).$$

Also under some extra assumptions on ϕ, listed in [29], the existence of at least one optimal control $u \in U_{ad}$ with associated state $y = G(u)$ for the optimal control problem (6) has been demonstrated in [29]. However, for our further analysis, we also would require the notion of the local solution in the following sense: A control $\bar{u} \in U_{ad}$ is said to be the local solution of (6), if there exists a constant $\lambda > 0$ such that

$$(7) \qquad j(v) \geq j(\bar{u}), \quad \forall v \in U_{ad} \text{ with } \|v - \bar{u}\|_{L^2(I;L^2(\Omega))} \leq \lambda.$$

We also assume that the local solution $\bar{u}$ satisfies the first-order necessary and second order-sufficient optimality conditions.
The following first-order optimality condition corresponding to the parabolic optimal control problem has been established in [34] (see also [26, 27, 29]):

$$(8) \qquad j'(u)(v-u) \geq 0, \quad \forall v \in U_{ad},$$

which can also be rewritten in the form

$$(9) \qquad \int_0^T (\alpha u + B^* p, v - u) \geq 0, \quad \forall v \in U_{ad}.$$

Here p is called *adjoint state* (or *costate*) associated with u and solves the *adjoint state equation*

$$
\begin{aligned}
(10) \qquad -\partial_t p - \nabla\cdot(A\,\nabla p) + \phi'(y)p &= y - y_d, \quad in \quad I\times\Omega,\\
(11) \qquad p &= 0, \quad in \quad I\times\partial\Omega,\\
(12) \qquad p(T,x) &= 0, \quad x\in\Omega.
\end{aligned}
$$

We assume that the local solution $\bar{u} \in U_{ad}$ satisfies the first-order necessary optimality condition (8) and the following standard second-order sufficient condition (see [29]). There exists a constant $C > 0$ such that

$$(13) \qquad j''(\bar{u})(v,v) \geq C\,\|v\|^2_{L^2(I;L^2(\Omega))}, \quad \forall v \in L^2(I; L^2(\Omega)).$$

If we define the following pointwise projection operator on the admissible set U_{ad} (see, e.g. [14]):

$$P_{[a,b]}(g(t,x)) = max(a, min(b, g(t,x))),$$

then the optimality condition (9) can be expressed as

$$u(t,x) = P_{[a,b]}\left(\frac{-1}{\alpha}B^*p(t,x)\right).$$

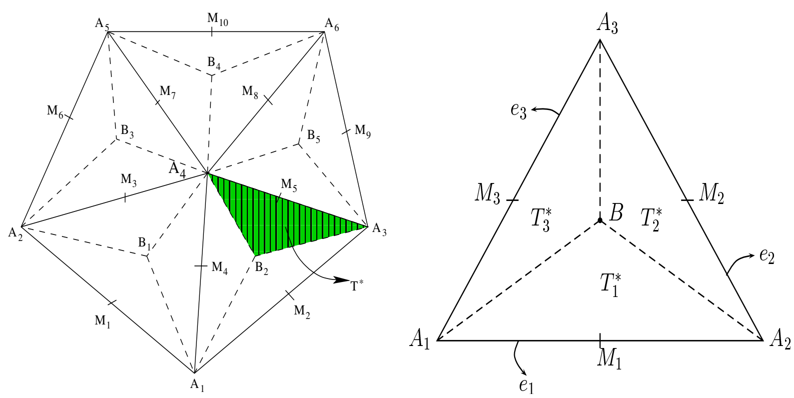
Here, B^* is the adjoint operator of B. We also note that the projection $P_{[a,b]}$ satisfies the regularity property (see [29])

$$\left\|\nabla(P_{[a,b]}(v))(t)\right\|_{L^\infty(\Omega)} \leq \|\nabla v(t)\|_{L^\infty(\Omega)}, \quad \forall v \in L^2(I, W^{1,\infty}(\Omega)), \tag{14}$$

for almost all $t \in I$.

2. Finite dimensional formulation

2.1. Discontinuous finite volume discretization. We would like to seek numerical approximation of the state and costate variables by discontinuous finite volume method and for spatial discretization we proceed as follows. Let τ_h be a regular, quasi-uniform triangulation of $\bar{\Omega}$ into closed triangles K with $h = \max_{K\in\tau_h}(h_K)$, where h_K is the diameter of the triangle K. The set of all interior edges in τ_h is denoted by $\mathcal{E}_h$.



(a) The dual partition of a triangulation. (b) A triangular partition and its dual.

FIGURE 1

The dual partition τ_h^* of τ_h is constructed as follows: divide each triangle $K \in \tau_h$ into three subtriangles, say $(T_i^*)_{i=1}^3$ by joining the barycenter B and the vertices of K as shown in Figure 1, for more details see [38]. Let τ_h^* consists of all these triangles T_i^*. Figure 1 indicates that the elements T_i^* of the the dual partition have the support in the triangle in which they belong, whereas in the case of continuous FVE methods, the elements of the dual partition may have support in the neighboring triangles, see [7]. The high localizability of the dual elements provide an advantage for parallel computing and implementation of adaptive FVE methods. The advantages of DFV methods (in terms of small support of the control volume and other aspects of the computational issues) over the other numerical methods are clearly mentioned in [6, 38]. Now, we define the finite dimensional trial and test spaces associated with τ_h and τ_h^*, respectively as follows:

$$V_h \quad = \quad \{v_h \in L^2(\Omega) : v_h|_K \in P_1(K) \quad \forall K \in \tau_h\},$$

$$W_h = \{w_h \in L^2(\Omega) : w_h|_{T^*} \in P_0(T^*) \quad \forall T^* \in \tau_h^*\},$$

where $P_n(K)$ or $P_n(T^*)$ denotes the space of all polynomials of degree less than or equal to n defined on K or T^*, respectively. Let $V(h) = V_h + H^2(\Omega) \cap H_0^1(\Omega)$.

To connect the trial and test spaces, we define a transfer operator $\gamma : V(h) \longrightarrow W_h$ as:

$$\gamma v|_{T^*} = \frac{1}{h_e} \int_e v|_{T^*} ds, \quad T^* \in \tau_h^*,$$

where e is an edge in K, T^* is the dual element in τ_h^* containing e, and h_e is the length of the edge e. The operator γ satisfies the following technical result, see [16].

Lemma 2.1. *The following results hold true for* $v_h \in V_h$

$$(15) \quad \int\limits_e (v_h - \gamma v_h) ds = 0; \int\limits_K (v_h - \gamma v_h) dx = 0; \; \|v_h - \gamma v_h\|_{0,K} \leq C h_K \|v_h\|_{1,K}.$$

Also, let e be an interior edge shared by two elements $K_1, K_2 \in \tau_h$, and let $\mathbf{n_1}$ and $\mathbf{n_2}$ be unit normal vectors on e pointing exterior to K_1 and K_2, respectively. The average $\langle . \rangle$ and jump $[\![\cdot]\!]$ on e for scalar q and vector $\mathbf{r}$ are defined respectively as:

$$\begin{aligned} \langle q \rangle &= \frac{1}{2}(q_1 + q_2), \qquad [\![q]\!] = q_1 \mathbf{n_1} + q_2 \mathbf{n_2}, \\ \langle \mathbf{r} \rangle &= \frac{1}{2}(\mathbf{r_1} + \mathbf{r_2}), \qquad [\![\mathbf{r}]\!] = \mathbf{r_1} \cdot \mathbf{n_1} + \mathbf{r_2} \cdot \mathbf{n_2}, \end{aligned}$$

where $q_i = (q \restriction_{K_i}) \restriction_e$, $\mathbf{r}_i = (\mathbf{r} \restriction_{K_i}) \restriction_e$.
For boundary edge with outward normal vector $\mathbf{n}$, we define

$$\begin{aligned} \langle q \rangle &= q, \qquad [\![q]\!] = q\mathbf{n}, \\ \langle \mathbf{r} \rangle &= \mathbf{r}, \qquad [\![\mathbf{r}]\!] = \mathbf{r} \cdot \mathbf{n}. \end{aligned}$$

For our future analysis we also define the following natural mesh-dependent norms for all $v_h \in V(h)$:

$$|\!|\!| v_h |\!|\!|_h^2 := \sum_{K \in \mathcal{T}_h} |v_h|_{1,K}^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![v_h]\!]\|_{0,e}^2.$$

The following inequality has been proved in [38]:

$$\|v_h\|^2 \leq C \left[\sum_{K \in \mathcal{T}_h} |v_h|_{1,K}^2 + \sum_{\sigma \in \mathcal{E}_h} [\![\gamma \boldsymbol{v}_h]\!]_\sigma^2 \right] \quad \forall v_h \in V_h.$$

A repeated application of Cauchy-Schwarz inequality yields

(16)

$$[\![\gamma v_h]\!]_e^2 = h_e^{-2} \left(\int_e [\![v_h]\!]_e \, \mathrm{d}s \right)^2 \mathrm{d}s \leq h_e^{-2} \left(\int_e \mathrm{d}s \right) \left(\int_e [\![v_h]\!]_e^2 \, \mathrm{d}s \right) = h_e^{-1} \int_e [\![\boldsymbol{v}_h]\!]_e^2 \, \mathrm{d}s.$$

Therefore, we have the following discrete Poincaré-Friedrichs type inequality

$$(17) \qquad \|v_h\| \leq C \, |\!|\!| v_h |\!|\!|_h \qquad \forall v_h \in V_h.$$

For the approximation of the time derivative, let $0 = t_0 < t_1 < ... < t_M = T$ be a partition of time interval $[0, T]$ into subintervals $I_m = (t_{m-1}, t_m]$ with length

$k_m = t_m - t_{m-1}$ for $m = 1, 2, ...M$ and $k = \max_{1\leq m\leq M} k_m$. Now we use the backward Euler scheme which is defined as follows:

$$\partial_t v^m := \frac{(v^m - v^{m-1})}{k_m},$$

where $v^m = v(t_m, x)$. Then we define the following discrete time-dependent norms to be used for further analysis

$$\|v\|_{L^2(I;L^2(\Omega))} := \left(\sum_{m=1}^{M} k_m \|v^m\|^2\right)^{\frac{1}{2}}, \quad \|v\|_{L^\infty(I;L^2(\Omega))} := \max_{1\leq m\leq M} \|v^m\|.$$

Similarly we denote time and mesh dependent norms as

$$\|v\|_{L^2(I;V(h))} := \left(\sum_{m=1}^{M} k_m \|v^m\|_h^2\right)^{\frac{1}{2}}, \quad \|v\|_{L^\infty(I;V(h))} := \max_{1\leq m\leq M} \|v^m\|_h.$$

Discretization techniques for control variable. For discretization of control variable we describe here three different approaches. Let U_h be a finite dimensional subspace of $L^2(I; L^2(\Omega))$, we introduce the discrete admissible space for control as

$$U_{h,ad} = U_h \cap U_{ad}.$$

(1) **Variational approach.** In this approach, control variables are not discretized explicitly and the discrete admissible space $U_{h,ad}$ coincides with the space U_{ad}.

(2) **Piecewise linear discretization.** Other natural way for seeking approximation of the control variable in the similar space used for the approximation of the state and co-state variables, i.e, piecewise linear subspace on triangulation which is defined as

$$U_h = \{u_h(.,t) \in L^2(I; L^2(\Omega)) : u_h(.,t)|_K \in P_1(K) \quad \forall K \in \tau_h, t \in I\}.$$

We stress that the state space V_h coincides with control space U_h in the case of homogeneous Neumann boundary conditions and is a subspace of it in presence of Dirichlet boundary conditions.

Now, for each time interval I_m, let us define a function $v_h^m \in U_{h,ad}$ on an arbitrary triangle $K \in \tau_h$ by

$$v_h^m = \begin{cases} a & \text{if } \min_{x\in K} u(t_m, x) = a, \\ b & \text{if } \max_{x\in K} u(t_m, x) = b, \\ \tilde{I}_h u^m & \text{else,} \end{cases}$$

where $\tilde{I}_h u^m$ be the linear interpolate of u^m. To avoid the ambiguity, we choose the mesh size h sufficiently small such that $\min_{x\in K} u(t_m, x) = a$ and $\max_{x\in K} u(t_m, x) = b$ cannot happen simultaneously in the same triangle K. Moreover, the triangles $K \in \tau_h$ are grouped into three sets $\tau_h = \tau_{h,m}^1 \cup \tau_{h,m}^2 \cup \tau_{h,m}^3$ with $\tau_{h,m}^i \cap \tau_{h,m}^j = \phi$ for $i \neq j$ according to the value of $u(t^m, x)$ on K. The sets are defined as follows:

$$\begin{aligned} \tau_{h,m}^1 &= \{K \in \tau_h : u(t_m, x) = a \quad or \quad u(t_m, x) = b \quad \forall x \in K\}, \\ \tau_{h,m}^2 &= \{K \in \tau_h : a < u(t_m, x) < b \quad \forall x \in K\}, \\ \tau_{h,m}^3 &= \tau_h \setminus (\tau_{h,m}^1 \cup \tau_{h,m}^2). \end{aligned}$$

For our further analysis, we impose the following assumption on the above described discretization approach:
Assumption 1. $\exists$ a positive constant C independent of k, h and m such that

$$\sum_{K\in\tau_{h,m}^3} |K| \leq Ch, \quad m = 1, 2, ..., M. \tag{18}$$

(3) **Piecewise constant discretization**. Another approach for the discretization of the control variable is to use elementwise constant functions. In this case, the discrete control space is defined as

$$U_h = \{u_h(.,t) \in L^2(I; L^2(\Omega)) : u_h(.,t)|_K \in P_0(K) \quad \forall K \in \tau_h, t \in I\}.$$

2.2. Fully-discrete discontinuous finite volume formulation: On multiplying (2) by γw_h, integrating over the control volumes, applying Gauss divergence methods and following the arguments used in [16] (see also [17]), we can obtain the DFV formulation of the state equation (2)-(4) as:

$$(\partial_t y_h, \gamma w_h) + A_h(y_h, w_h) + (\phi(y_h), \gamma w_h) = (Bu_h + f, \gamma w_h), \quad \forall w_h \in V_h, \tag{19}$$
$$y_h(0,x) = y_{0,h}, \quad x \in \Omega, \tag{20}$$

where, $y_{0,h}$ is a certain approximation of y_0 to be defined later and the bilinear form $A_h(\cdot,\cdot) : V_h \times V_h \longrightarrow \mathbb{R}$ is defined as (see [16])

$$A_h(\Phi_h, \Psi_h) = -\sum_{K\in\tau_h}\sum_{j=1}^{3}\int_{A_{j+1}BA_j}(A\nabla\Phi_h\cdot\mathbf{n})\gamma\Psi_h ds + \theta\sum_{e\in\mathcal{E}_h}\int_e [\![\gamma\Phi_h]\!]\cdot\langle A\nabla\Psi_h\rangle ds$$
$$-\sum_{e\in\mathcal{E}_h}\int_e [\![\gamma\Psi_h]\!]\cdot\langle A\nabla\Phi_h\rangle ds + \sum_{e\in\mathcal{E}_h}\int_e \frac{\alpha_d}{h_e^\beta}[\![\Phi_h]\!]\cdot[\![\Psi_h]\!] ds, \quad \forall\Phi_h, \Psi_h \in V_h,$$

where, $A_4 = A_1$, (see Figure 1(b)), and α_d and β are penalty parameters. In general $\theta \in [-1, 1]$ and different values of β are required for achieving the optimal rate of convergence in the L^2-norm for $\theta \neq -1$, for more details kindly see [16]. But in what follows, we assume $\theta = -1$ which is known as symmetric interior penalty Galerkin (SIPG) method in the context of discontinuous finite element methods.
The approximation of the optimal control problem (1)-(4) using DFV method is given by: Find $(y_h(\cdot,t), p_h(\cdot,t), u_h(\cdot,t)) \in V_h \times V_h \times U_{h,ad}$; $(0 < t < T)$ such that

$$\begin{aligned}(\partial_t y_h, \gamma w_h) &+ A_h(y_h, w_h) + (\phi(y_h), \gamma w_h)\\ &= (Bu_h + f, \gamma w_h), \quad \forall w_h \in V_h,\\ y_h(0,x) &= y_{0,h}, \quad x \in \Omega,\end{aligned} \tag{21}$$

$$\begin{aligned}-(\partial_t p_h, \gamma q_h) + A_h(p_h, q_h) &+ (\phi'(y_h)p_h, \gamma q_h)\\ &= (y_h - y_d, \gamma q_h), \quad \forall q_h \in V_h,\\ p_h(T,x) &= 0, \quad x \in \Omega,\end{aligned} \tag{22}$$

$$\int_0^T (\alpha u_h + B^* p_h, v_h - u_h) \geq 0, \quad \forall v_h \in U_{h,ad}. \tag{23}$$

The above discrete optimal system admits a unique local optimal control u_h with the associated state y_h and the associated costate p_h, see [25, 29] for details. In the light of the above mentioned discretization approaches for spatial and time domain, the backward Euler fully-discrete piecewise linear discontinuous finite volume formulation of the parabolic control problem (1)-(4) read as follows (see also [23, 25]): find $(y_h^m, p_h^{m-1}, u_h^m) \in V_h \times V_h \times U_{h,ad}$ such that $\forall\, w_h, q_h \in V_h$

$$\begin{aligned} &(\partial_t y_h^m, \gamma w_h) + A_h(y_h^m, w_h) + (\phi(y_h^m), \gamma w_h) \\ &= (Bu_h^m + f^m, \gamma w_h), \;\; m = 1, ..., M; \\ &y_h^0(x) = y_{0,h}, \quad x \in \Omega, \end{aligned} \tag{24}$$

$$\begin{aligned} &-(\partial_t p_h^m, \gamma q_h) + A_h(p_h^{m-1}, q_h) + (\phi'(y_h^m)p_h^{m-1}, \gamma q_h) \\ &= (y_h^m - y_d^m, \gamma q_h), \;\; m = M, ..., 1; \\ &p_h^M(x) = 0, \quad x \in \Omega, \end{aligned} \tag{25}$$

$$(\alpha u_h^m + B^* p_h^{m-1}, v_h - u_h^m) \geq 0 \quad \forall v_h \in U_{h,ad}, \;\; m = 1, ..., M. \tag{26}$$

2.3. Some auxiliary results: For our further analysis, we would require the following well known results.

Result 1. With help of technical lemma 15, one can easily show that the bilinear form $A_h(\cdot,\cdot)$ is bounded and coercive with respect to the norm $|\!|\!|\cdot|\!|\!|_h$, i.e, there exist positive constants β_0 and C independent of h such that (kindly see [16])

$$A_h(\phi_h, \phi_h) \;\; \geq \;\; \beta_0 \,|\!|\!| v_h |\!|\!|_h^2, \qquad \forall \phi_h \in V_h, \tag{27}$$

$$|A_h(\phi_h, \psi_h)| \;\; \leq \;\; C \,|\!|\!| \phi_h |\!|\!|_h \,|\!|\!| \psi_h |\!|\!|_h, \qquad \forall \phi_h, \psi_h \in V_h. \tag{28}$$

Result 2. The operator γ is self-adjoint with respect to the L^2-inner product,

$$(\phi_h, \gamma\psi_h) = (\psi_h, \gamma\phi_h), \quad \forall \phi_h, \psi_h \in V_h. \tag{29}$$

Also, if $|\!|\!|\psi_h|\!|\!|_0 := (\psi_h, \gamma\psi_h)$ then $|\!|\!|\cdot|\!|\!|_0$ and $\|\cdot\|$ are equivalent and

$$\|\gamma\psi_h\| = \|\psi_h\|, \quad \psi_h \in V_h. \tag{30}$$

For a proof we refer to [4].

Result 3. For each $\phi_h, \psi_h \in V_h$, we have

$$|A_h(\phi_h, \psi_h) - A_h(\psi_h, \phi_h)| \leq Ch \,|\!|\!|\phi_h|\!|\!|_h \,|\!|\!|\psi_h|\!|\!|_h\,. \tag{31}$$

For details we refer to [4] and also see [38].

Result 4. If $\epsilon_a(\phi_h, \psi_h) := a(\phi_h, \psi_h) - A_h(\phi_h, \psi_h)$ then for all $\phi_h, \psi_h \in V_h$, we have

$$|\epsilon_a(\phi_h, \psi_h)| \leq Ch \,|\!|\!|\phi_h|\!|\!|_h \,|\!|\!|\psi_h|\!|\!|_h\,, \tag{32}$$

where the bilinear form $a(\cdot,\cdot)$ is defined as for $\phi_h, \psi_h \in V_h$:

$$a(\phi_h, \psi_h) = \sum_{K\in\tau_h} A\nabla\phi_h \cdot \nabla\psi_h \; dx + \theta \sum_{e\in\mathcal{E}_h} \int_e [\![\phi_h]\!].\langle A\nabla\psi_h\rangle ds$$

$$- \sum_{e\in\mathcal{E}_h} \int_e [\![\psi_h]\!].\langle A\nabla\phi_h\rangle ds + \sum_{e\in\mathcal{E}_h} \int_e \frac{\alpha_d}{h_e^\beta} [\![\phi_h]\!].[\![\psi_h]\!] ds.$$

For a proof we refer to Lemma 3.2 of [5].

Result 5. For explicit discretization of control, the following holds true for $v_h \in U_{h,ad}$ (for details we refer to Lemma 2.1, [9]).

$$(\alpha u + B^* p, v - v_h)_{L^2(I;L^2(\Omega))} \geq 0, \quad \forall v \in U_{ad}. \tag{33}$$

3. *A priori* error estimates

In this section we derive error estimates for a fixed local (in the sense of (7)) reference solution of the problem (6) which also satisfy first and second order optimality conditions. Since the control and state variables u and y appears in the state and costate equations, respectively, the error estimates for state and costate variables depend on the control variable and state variables, respectively. For deriving these estimates we proceed in the following way. For a given arbitrary $\tilde{u} \in L^2(I; L^2(\Omega))$ and $\tilde{y} = y(\tilde{u}) \in L^2(I; H_0^1(\Omega))$, let $y_h^m(\tilde{u})$ and $p_h^{m-1}(\tilde{y})$ be the solutions of the following equations $\forall w_h, q_h \in V_h$; for $m = 1, ..., M$;

$$(34)\qquad \begin{aligned}&(\partial_t y_h^m(\tilde{u}), \gamma w_h) + A_h(y_h^m(\tilde{u}), w_h) + (\phi(y_h^m(\tilde{u})), \gamma w_h) = (B\tilde{u}^m + f^m, \gamma w_h)\\ &y_h^0(\tilde{u})(x) = y_{0,h}, \quad x \in \Omega,\end{aligned}$$

and for $m = M, ..., 1$;

$$(35)\qquad \begin{aligned}&-(\partial_t p_h^m(\tilde{y}), \gamma q_h) + A_h(p_h^{m-1}(\tilde{y}), q_h) + (\phi'(y^m)p_h^{m-1}(\tilde{y}), \gamma q_h) = (\tilde{y}^m - y_d^m, \gamma q_h),\\ &p_h^M(\tilde{y})(x) = 0, \quad x \in \Omega,\end{aligned}$$

respectively. In order to avoid confusion, in what follows we will use the following notations: $y_h = y_h(u_h)$, $p_h = p_h(y_h)$ and $p_h(u) = p_h(y_h(u))$. Now using similar arguments as in the proof of Lemma 5.1 given in [25], we now prove the following lemma for $\tilde{u} = u$ and $\tilde{y} = y(u)$.

Lemma 3.1. *For sufficiently small k, there exists a positive constant C independent of h and k such that*

$$|||y_h^n(u) - y_h^n|||_h \le C \|u - u_h\|_{L^2(I;L^2(\Omega))}, \quad |||p_h^n(y) - p_h^n|||_h \le C \|y - y_h\|_{L^2(I;L^2(\Omega))}.$$

Proof. Subtracting equations (24) from (34), we obtain for all $w_h \in V_h$ and $m = 1, 2, ..., M$

$$\begin{aligned}&(\partial_t y_h^m(u) - \partial_t y_h^m, \gamma w_h) + A_h(y_h^m(u) - y_h^m, w_h) + (\phi(y_h^m(u)) - \phi(y_h^m), \gamma w_h)\\ =&(B(u^m - u_h^m), \gamma w_h).\end{aligned}$$

Now, by choosing $w_h = \partial_t \eta^m$ and denoting $y_h^m(u) - y_h^m = \eta^m$, the above equation can be rewritten as follows:

$$\begin{aligned}&(\partial_t \eta^m, \gamma(\partial_t \eta^m)) + A_h(\eta^m, \partial_t \eta^m)\\ =&(B(u^m - u_h^m), \gamma(\partial_t \eta^m)) + (\phi(y_h^m) - \phi(y_h^m(u)), \gamma(\partial_t \eta^m)).\end{aligned}$$

Using the definitions of the norm $|||\cdot|||_0$ and $\epsilon_a(\cdot,\cdot)$, we arrive at

$$(36)\qquad \begin{aligned}|||\partial_t \eta^m|||_0^2 + a(\eta^m, \partial_t \eta^m) \;=\;& (B(u^m - u_h^m), \gamma(\partial_t \eta^m)) + \epsilon_a(\eta^m, \partial_t \eta^m)\\ &+(\phi(y_h^m) - \phi(y_h^m(u)), \gamma(\partial_t \eta^m)).\end{aligned}$$

A simple manipulation shows that

$$(37)\qquad a(\eta^m, \partial_t \eta^m) \ge \frac{1}{2k_i}(a(\eta^m, \eta^m) - a(\eta^{m-1}, \eta^{m-1})).$$

An application of (32) and inverse inequality together with Young's inequality, provide us

$$(38)\qquad \begin{aligned}&\epsilon_a(\eta^m, \partial_t \eta^m) \le Ch\, |||\eta^m|||_h\, |||\partial_t \eta^m|||_h\\ \le& C\, |||\eta^m|||_h \|\partial_t \eta^m\| \le C(\epsilon)\, |||\eta^m|||_h^2 + \epsilon \|\partial_t \eta^m\|^2.\end{aligned}$$

The following inequality follows by Cauchy-Schwarz inequality and (30)

$$
\begin{aligned}
(B(u^m - u_h^m), \gamma(\partial_t \eta^m)) &\leq C\,\|u^m - u_h^m\|\,\|\gamma(\partial_t \eta^m)\| \leq C\,\|u^m - u_h^m\|\,\|\partial_t \eta^m\| \\
&\leq C(\epsilon)\,\|u^m - u_h^m\|^2 + \epsilon\,\|\partial_t \eta^m\|^2 . \qquad (39)
\end{aligned}
$$

The Lipschitz continuity of $\phi(\cdot)$ together with (17) implies that

$$
\begin{aligned}
(\phi(y_h^m) - \phi(y_h^m(u)), \gamma(\partial_t \eta^m)) &\leq \|\eta^m\|\,\|\gamma(\partial_t \eta^m)\| \leq |\!|\!|\eta^m|\!|\!|_h\,\|\partial_t \eta^m\| \\
&\leq C(\epsilon)\,|\!|\!|\eta^m|\!|\!|_h^2 + \epsilon\,\|\partial_t \eta^m\|^2 . \qquad (40)
\end{aligned}
$$

Collecting the bounds obtained in (37), (38), (39), (40) and using equivalence of $|\!|\!|\cdot|\!|\!|_0$ and $\|.\|$ with appropriate value of ϵ in relation (36), enable us to write the following

$$
a(\eta^m, \eta^m) - a(\eta^{m-1}, \eta^{m-1}) \leq C\{2k_m\,|\!|\!|\eta^m|\!|\!|_h^2 + 2k_m\,\|u^m - u_h^m\|^2\}.
$$

Summing m from 1 to n, using coercivity of $a(\cdot,\cdot)$ and noting $\eta^0 = 0$, we find that

$$
|\!|\!|\eta^n|\!|\!|_h^2 \leq C\{\sum_{m=1}^{n} k_m\,|\!|\!|\eta^m|\!|\!|_h^2 + \sum_{m=1}^{n} k_m\,\|u^m - u_h^m\|^2\},
$$

and an appeal to the discrete Gronwall's Lemma implies that

$$
|\!|\!|y_h^n(u) - y_h^n|\!|\!|_h \leq C\,\|u - u_h\|_{L^2(I;L^2)} .
$$

For estimating $|\!|\!|p_h^n(y) - p_h^n|\!|\!|_h$, we proceed in the similar way as we have estimated $|\!|\!|y_h^n(u) - y_h^n|\!|\!|_h$. On subtracting (26) from (35), writing $p_h^m(y) - p_h^m = \mu^m$ and choosing $q_h = \partial_t \mu^m$, we infer that for $m = 1, 2, ..., M$

$$
\begin{aligned}
&|\!|\!|\partial_t \mu^m|\!|\!|_0^2 - a(\mu^{m-1}, \partial_t \mu^m) \\
= \;& (y_h^m - y^m, \gamma(\partial_t \mu^m)) - \epsilon_a(\mu^{m-1}, \partial_t \mu^m) \\
&+ (\phi'(y^m) p_h^{m-1}(y) - \phi'(y_h^m) p_h^{m-1}, \gamma(\partial_t \mu^m)). \qquad (41)
\end{aligned}
$$

We note that

$$
-\,a(\mu^{m-1}, \partial_t \mu^m) \geq \frac{1}{2k_i}(a(\mu^{m-1}, \mu^{m-1}) - a(\mu^m, \mu^m)). \qquad (42)
$$

It follows from the assumption $\phi'(\cdot) \geq 0$ and Lipschitz continuity that

$$
\begin{aligned}
&(\phi'(y^m) p_h^{m-1}(y) - \phi'(y_h^m) p_h^{m-1}, \gamma(\partial_t \mu^m)) \\
\leq \;& C(\mu^{m-1}, \gamma(\partial_t \mu^m)) \leq C\,\|\mu^{m-1}\|\,\|\gamma(\partial_t \mu^m)\| \\
\leq \;& C(\epsilon)\,|\!|\!|\mu^{m-1}|\!|\!|_h^2 + \epsilon\,\|\partial_t \mu^m\|^2 . \qquad (43)
\end{aligned}
$$

Also, using the arguments used in derivation of inequalities (38) and (40), we have the following bounds

$$
\begin{aligned}
&\epsilon_a(\mu^{m-1}, \partial_t \mu^m) \leq Ch\,|\!|\!|\mu^{m-1}|\!|\!|_h\,|\!|\!|\partial_t \mu^m|\!|\!|_h \\
&\leq C(\epsilon)\,|\!|\!|\mu^{m-1}|\!|\!|_h^2 + \epsilon\,\|\partial_t \mu^m\|^2 , \qquad (44) \\
&(y_h^m - y^m, \gamma(\partial_t \mu^m)) \leq C\,\|y^m - y_h^m\|\,\|\gamma(\partial_t \mu^m)\| \\
&\leq C(\epsilon)\,\|y^m - y_h^m\|^2 + \epsilon\,\|\partial_t \mu^m\|^2 . \qquad (45)
\end{aligned}
$$

Using inequalities (42), (43), (44), (45) and equivalence of $|\!|\!|\cdot|\!|\!|_0$ and $\|.\|$ with appropriate value of ϵ in relation (41), we can find that

$$
a(\mu^{m-1}, \mu^{m-1}) - a(\mu^m, \mu^m) \leq C\{2k_m\,|\!|\!|\mu^{m-1}|\!|\!|_h^2 + 2k_m\,\|y^m - y_h^m\|^2\}.
$$

Again summing m from $n+1$ to M, using coercivity of $a(\cdot,\cdot)$ and noticing $\mu^M = 0$ we see that

$$\|\mu^n\|_h^2 \leq C\{ \sum_{m=n+1}^{M} k_m \|\mu^m\|_h^2 + \sum_{m=n+1}^{n} k_m \|y^m - y_h^m\|^2\}$$

Now an application to discrete Gronwall's Lemma for sufficiently small k_m implies the required estimate, i.e.,

$$\|p_h^n(y) - p_h^n\|_h \leq C \|y - y_h\|_{L^2(I;L^2(\Omega))}.$$

□

3.1. Error estimates for variational discretization approach. First we define the elliptic projection $R_h : H^2(\Omega) \cap H_0^1(\Omega) \longrightarrow V_h$ by

$$A_h(R_h u, v_h) := A_h(u, v_h) \quad \forall v_h \in V_h, \tag{46}$$

and in what follows, we choose $y_{0,h} = R_h y_0(x)$ for $x \in \Omega$. Now, for a given u, the following estimates can be derived by using the elliptic projection defined in (46) and appealing to duality arguments used for the standard discontinuous finite volume analysis for parabolic problems. Therefore, we refrain ourself for providing this proof and we refer to [4], also see [17] and [6].

Lemma 3.2. *For any $\tilde{u} \in L^2(I;L^2(\Omega))$ and $\tilde{y} = y(\tilde{u}) \in L^2(I;H_0^1(\Omega))$, there exists a positive constant C independent of h and k such that*

$$\begin{aligned} \|y(\tilde{u}) - y_h(\tilde{u})\|_{L^2(I;L^2(\Omega))} &= \mathcal{O}(h^2+k), \quad \|p(\tilde{y}) - p_h(\tilde{y})\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2+k) \\ \|p(\tilde{u}) - p_h(\tilde{u})\|_{L^2(I;L^2(\Omega))} &= \mathcal{O}(h^2+k), \end{aligned}$$

and in particular, for $\tilde{u} = u_h$, we have

$$\|p(u_h) - p_h(u_h)\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2+k). \tag{47}$$

Now, we are in the position to prove the following result of this section.

Theorem 3.3. *Let u be a fixed local optimal control of problem (6) and u_h be the solution of the fully discrete optimal control problem (24)-(26) with variational discretization approach, then the following error estimate holds.*

$$\|u - u_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2+k), \ \|y - y_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2+k),$$
$$\|p - p_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2+k).$$

Proof. For each time interval I_m the continuous variational inequality will be of the form

$$(\alpha u^m + B^* p^{m-1}, w - u^m) \geq 0, \quad \forall w \in U_{ad} \tag{48}$$

and the discrete variational inequality is

$$(\alpha u_h^m + B^* p_h^{m-1}, v - u_h^m) \geq 0, \forall v \in U_{ad}. \tag{49}$$

Choosing $w = u_h^m$ in (48) and $v = u^m$ in (49), we have

$$(\alpha u^m + B^* p^{m-1}, u_h^m - u^m) \geq 0 \leq -(\alpha u_h^m + B^* p_h^{m-1}, u^m - u_h^m). \tag{50}$$

The condition (13) for $u - u_h \in U_{ad} \subset L^2(I;L^2(\Omega))$, implies that

$$\begin{aligned} & C\|u - u_h\|_{L^2(I;L^2(\Omega))}^2 \\ \leq \; & j'(u)(u - u_h) - j'(u_h)(u - u_h) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{m=1}^{M} k_m \left((\alpha u^m + B^* p^{m-1}, u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) \right) \\
&\leq \sum_{m=1}^{M} k_m \left((\alpha u_h^m + B^* p_h^{m-1}, u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) \right) \\
&\leq \|p(u_h) - p_h\|_{L^2(I;L^2(\Omega))} \|u - u_h\|_{L^2(I;L^2(\Omega))}.
\end{aligned}
$$

Using (47) in the above relation yields the required result, i.e.

$$
\|u - u_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k). \tag{51}
$$

Now, decomposing the state and costate error as $y - y_h = y - y_h(u) + y_h(u) - y_h$ and $p - p_h = p - p_h(y) + p_h(y) - p_h$, respectively, using triangle inequality together with the results of Lemmas 3.1, 3.2 and the estimate for u given in (51), we can easily obtain the following

$$
\|y - y_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k), \; \|p - p_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k).
$$

□

3.2. Error estimates for control with piecewise linear discretization: Under the assumption (18), it is not hard to prove the following Lemma which play a vital role for the subsequent analysis of this subsection. An analogous version of this Lemma has been demonstrated in [27] and therefore, we skip the proof.

Lemma 3.4. *There exists a positive constant C independent of h and k such that*

$$
|(\alpha u + B^* p, v_h - u)_{L^2(I,L^2(\Omega))}| \leq \frac{C}{\alpha} h^3 \|\nabla p\|^2_{L^2(I;L^\infty(\Omega))} \quad for \quad v_h \in U_{h,ad}. \tag{52}
$$

Following the same idea used in establishment of the Lemma 5.7 given in [27], we prove our main result.

Theorem 3.5. *Let u be a fixed local optimal control of problem (6) and u_h be the solution of the fully discrete optimal control problem (24)-(26) with piecewise linear discretization of controls then the following estimate holds true*

$$
\|u - u_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^{3/2} + k).
$$

Proof. We proceed in the similar way as we have proved Theorem 3.3. First by testing the continuous and discrete variational inequalities on each subinterval I_m with $u_h^m \in U_{h,ad} \subset U_{ad}$ and $v_h^m \in U_{h,ad}$, we find that

$$
(\alpha u_h^m + B^* p_h^m, v_h^m - u_h^m) \geq 0 \geq (\alpha u^m + B^* p^m, u^m - u_h^m).
$$

Using the condition (13) for $u - u_h \in U_{h,ad}$ and the discrete and continuous variational inequalities, we have

$$
\begin{aligned}
& C \|u - u_h\|^2_{L^2(I;L^2(\Omega))} \\
&\leq j'(u)(u - u_h) - j'(u_h)(u - u_h) \\
&= \sum_{m=1}^{M} k_m \left((\alpha u^m + B^* p^{m-1}, u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) \right) \\
&\leq \sum_{m=1}^{M} k_m ((\alpha u^m + B^* p^{m-1}, u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m)
\end{aligned}
$$

$$
\begin{aligned}
& +(\alpha u_h^m + B^* p_h^{m-1}, v_h^m - u_h^m)) \\
\leq & \sum_{m=1}^{M} k_m((\alpha u^m + B^* p^{m-1}, u^m - v_h^m) + (\alpha u^m + B^* p^{m-1}, v_h^m - u^m) \\
& -(\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) + (\alpha u_h^m + B^* p_h^{m-1}, u^m - u_h^m) \\
& +(\alpha u_h^m + B^* p_h^{m-1}, v_h^m - u^m)) \\
\leq & \sum_{m=1}^{M} k_m(\alpha(u^m - u_h^m, u^m - v_h^m) + (p^{m-1} - p_h^{m-1}, B(u^m - v_h^m)) \\
& +(p_h^{m-1} - p^{m-1}(u_h), B(u^m - u_h^m)) + (\alpha u^m + B^* p^{m-1}, v_h^m - u^m)) \\
\leq & \left(\|u - u_h\|_{L^2(I;L^2(\Omega))} + \|p - p_h\|_{L^2(I;L^2(\Omega))} \right) \|u - v_h\|_{L^2(I;L^2(\Omega))} \\
& +C(h^2 + k) \|u - u_h\|_{L^2(I;L^2(\Omega))} + |(\alpha u + B^* p, v_h - u)_{L^2(I;L^2(\Omega))}|. \qquad (53)
\end{aligned}
$$

Now, we can write $\|u - v_h\|_{L^2(I;L^2(\Omega))} = \sum_{m=1}^{M} k_m \|u^m - v_h^m\|$. Since $v_h^m = u^m$ on $\tau_{h,m}^1$ we have,
$\sum_{K \in \tau_{h,m}^1} \|u^m - v_h^m\|_{L^2(K)}^2 = 0$. Therefore, for each $m = 1, 2, ..M$, we can split

$$
\begin{aligned}
\|u^m - v_h^m\|^2 &= \sum_{K \in \tau_h} \|u^m - v_h^m\|_{L^2(K)}^2 \\
&= \sum_{K \in \tau_{h,m}^2} \|u^m - v_h^m\|_{L^2(K)}^2 + \sum_{K \in \tau_{h,m}^3} \|u^m - v_h^m\|_{L^2(K)}^2 := T_1 + T_2. \qquad (54)
\end{aligned}
$$

To bound T_1, we use the relation $u^m = \frac{-B^*}{\alpha} p^m$ on all triangles $K \in \tau_{h,m}^2$ to obtain

$$
\sum_{K \in \tau_{h,m}^2} \|u^m - I_h u^m\|_{L^2(K)}^2 \leq Ch^4 \sum_{K \in \tau_{h,m}^2} \left\|\nabla^2 u^m\right\|_{L^2(K)}^2 \leq \frac{C}{\alpha^2} h^4 \left\|\nabla^2 p^m\right\|^2.
$$

For T_2, a use of projection property (14) and assumption 18 gives us

$$
\begin{aligned}
\sum_{K \in \tau_{h,m}^3} \|u^m - I_h u^m\|_{L^2(K)}^2 &\leq C \sum_{K \in \tau_{h,m}^3} |K| \, \|u^m - I_h u^m\|_{L^\infty(K)}^2 \\
&\leq Ch^3 \|\nabla u^m\|_{L^\infty(\Omega)}^2 \leq \frac{C}{\alpha^2} h^3 \|\nabla p^m\|_{L^\infty(\Omega)}^2.
\end{aligned}
$$

Substituting the bounds of T_1 and T_2 in (54) and summing m from 1 to M on each intervals I_m, we get the following estimate

$$
\|u - v_h\|_{L^2(I;L^2(\Omega))} \leq \frac{C}{\alpha} \left(h^2 \left\|\nabla^2 p\right\|_{L^2(I;L^2(\Omega))} + h^{3/2} \|\nabla p\|_{L^2(I;L^\infty(\Omega))} \right). \qquad (55)
$$

Using triangle inequality along with property (17) and results of Lemma 3.1 and Lemma 3.2, we find that

$$
\begin{aligned}
\|p - p_h\|_{L^2(I;L^2(\Omega))} &\leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + \|p_h(y) - p_h\|_{L^\infty(I;L^2(\Omega))} \\
&\leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + \|p_h(y) - p_h\|_{L^\infty(I;V(h))} \\
&\leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + C \|y - y_h\|_{L^2(I;L^2(\Omega))} \\
&\leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + C(\|y - y_h(u)\|_{L^2(I;L^2(\Omega))} \\
&\quad + \|y_h(u) - y_h\|_{L^\infty(I;V(h))}) \\
&\leq C(h^2 + k) + C \|u - u_h\|_{L^2(I;L^2(\Omega))}.
\end{aligned}
$$

Inserting the above relation and (55) in (53) and using the results of Lemma 3.4, we can obtain the desired estimate

$$\|u - u_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^{\frac{3}{2}} + k).$$

□

3.3. Error estimates for control with piecewise constant discretization: In this section we will derive the error estimates for $\|u - u_h\|_{L^2(I;L^2(\Omega))}$ when the control variable is discretized by piecewise constants. For the accomplishment of the main result, by following the idea of [9], we introduce an L^2-projection operator $\Gamma_h : L^2(I, L^2(\Omega)) \longrightarrow U_h$ which satisfy the following property.

$$\Gamma_h U_{ad} \subset U_{h,ad}.$$

Now, below we prove the main result of this section.

Theorem 3.6. *Let u be a fixed local optimal control of problem (6) and u_h be the solution of the fully discrete optimal control problem (24)-(26) with piecewise constant discretization of controls, then we have the following discretization error estimate*

$$\|u - u_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h + k).$$

Proof. Since $\Gamma_h U_{ad} \subset U_{h,ad}$ and proceeding in a similar manner as in the proof of Theorem 3.3, we easily see that the following holds with the help of continuous and discrete optimality conditions

$$(\alpha u_h^m + B^* p_h^{m-1}, \Gamma_h u^m - u_h^m) \geq 0 \geq (\alpha u^m + B^* p^{m-1}, u^m - u_h^m).$$

Applying condition (13) for $u - u_h \in U_{ad} \subset L^2(I; L^2(\Omega))$ and using discrete and continuous variational inequalities, we have

$$\begin{aligned}
& C\,\|u - u_h\|^2_{L^2(I;L^2(\Omega))} \\
\leq\ & j'(u)(u - u_h) - j'(u_h)(u - u_h) \\
=\ & \sum_{m=1}^{M} k_m \left((\alpha u^m + B^* p^{m-1}, u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) \right) \\
\leq\ & \sum_{m=1}^{M} k_m \left((\alpha u_h^m + B^* p_h^{m-1}, \Gamma_h u^m - u_h^m) - (\alpha u_h^m + B^* p^{m-1}(u_h), u^m - u_h^m) \right) \\
\leq\ & \underbrace{\sum_{m=1}^{M} k_m (B^* p_h^{m-1} - B^* p^{m-1}(u_h), u^m - u_h^m)}_{J_1}
\end{aligned}$$

$$+ \underbrace{\sum_{m=1}^{M} k_m (\alpha u_h^m + B^* p_h^{m-1}, \Gamma_h u^m - u^m)}_{J_2}. \tag{56}$$

Using (47) and continuity property of operator B, yields

$$\begin{aligned}
J_1 \ &\leq\ \|p(u_h) - p_h\|_{L^2(I;L^2(\Omega))} \|u - u_h\|_{L^2(I;L^2(\Omega))} \\
&\leq\ C(h^2 + k)\,\|u - u_h\|_{L^2(I;L^2(\Omega))}\,.
\end{aligned}$$

To achieve the desried bound for J_2 we use the property of the projection Γ_h to rewrite it as:

$$\begin{aligned} J_2 &= \sum_{m=1}^{M} k_m (B^* p_h^{m-1} - \Gamma_h(B^* p_h^{m-1}), \Gamma_h u^m - u^m) \\ &\leq \sum_{m=1}^{M} k_m \underbrace{\left(\sum_{K \in \tau_h} \left\| B^* p_h^{m-1} - \Gamma_h(B^* p_h^{m-1}) \right\|_{L^2(K)}^2 \right)^{\frac{1}{2}}}_{\leq Ch \left\| p_h^{m-1} \right\|_h} \left\| \Gamma_h u^m - u^m \right\|. \end{aligned}$$

Now let $I_h^0 u^m$ be a peicewise constant interpolant of u^m with the following approximation properties:

$$\left\| u^m - I_h^0 u^m \right\|_{L^2(K)} \leq C h_K \left\| u^m \right\|_{H^1(K)}. \tag{57}$$

Then we first note that

$$\begin{aligned} \left\| I_h^0 u^m - \Gamma_h u^m \right\|_{L^2(K)} &= (I_h^0 u^m - \Gamma_h u^m, I_h^0 u^m - \Gamma_h u^m)_{L^2(K)} \\ &= (I_h^0 u^m - u^m, I_h^0 u^m - \Gamma_h u^m)_{L^2(K)} \\ &\quad + \underbrace{(u^m - \Gamma_h u^m, I_h^0 u^m - \Gamma_h u^m)_{L^2(K)}}_{=0}. \end{aligned}$$

Thus, we have

$$\left\| I_h^0 u^m - \Gamma_h u^m \right\| \leq \left\| I_h^0 u^m - u^m \right\|. \tag{58}$$

Now using the relation (58) together with the property (57) we find that

$$\left\| \Gamma_h u^m - u^m \right\| \leq \left\| u_m - I_h^0 u^m \right\| + \left\| I_h^0 u^m - \Gamma_h u^m \right\| \leq \left\| u_m - I_h^0 u^m \right\| \leq Ch.$$

For completing the proof, we need to show that the p_h is uniformly bounded. This can be easily achieved by making use of coercivity of the bilinear form $A_h(\cdot,\cdot)$ with respect to the norm $\|\!|\cdot|\!\|_h$ and uniform boundedness of $U_{h,ad}$. Therefore, on substituting the bounds for J_1 and J_2 in (56), we complete the rest of the proof. □

3.4. Error estimates for state and costate with both piecewise linear and constant discretization of control. We would like to mention that for variational discretization we are enable to derive optimal error estimates (for state and co-state) with the help of Lemmas 3.1 and 3.2. But if we proceed in the similar way we end up with the order of convergence $(h^{3/2} + k)$ and $(h + k)$ for piecewise linear and constant discretization approaches, respectively. In order to achieve the desired optimal estimates for both piecewise linear and constant discretizations for control, we appeal to duality arguments in the following main theorem of this section. The similar idea also used in [29] and [26].

Theorem 3.7. *Let u be an optimal control of problem (6) with the associated state y and costate p, respectively, and let u_h, y_h and p_h be the solution of the fully discrete optimal control problem (24)-(26), then the following discretization error estimates are satisfied*

$$\|y - y_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k); \qquad \|p - p_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k).$$

Proof. Splitting the error $\|y - y_h\|_{L^2(I;L^2(\Omega))}$ as:

$$\|y - y_h\|_{L^2(I;L^2(\Omega))} \leq \|y - y_h(u)\|_{L^2(I;L^2(\Omega))} + \|y_h(u) - y_h(\Pi_h u)\|_{L^2(I;L^2(\Omega))}$$

$$+\|y_h(\Pi_h u) - y_h\|_{L^2(I;L^2(\Omega))} \tag{59}$$

where Π_h is the L^2 projection onto the discrete control space (piecewise constant or piecewise linear polynomials). For a given time interval I_m, let $\tilde{p}_h^m \in V_h$ be the solution of the auxiliary discrete dual equation

$$-(\gamma\xi, \partial_t \tilde{p}_h^m) + A_h(\tilde{p}_h^m, \xi) = (\gamma\xi, y_h^m(u) - y_h^m(\Pi_h u)) - (\gamma\xi, \hat{\phi}\tilde{p}_h^m), \quad \forall \xi \in V_h \tag{60}$$

with

$$\hat{\phi}(t_m, x) = \begin{cases} \frac{\phi(y_h^m(u)) - \phi(y_h^m(\Pi_h u))}{y_h^m(u) - y_h^m(\Pi_h u)}, & \text{if} \quad y_h^m(u) \neq y_h^m(\Pi_h u)) \\ 0, & \text{else.} \end{cases}$$

Since Ω is a convex polygonal domain and under suitable assumptions on $\hat{\phi}$, we assume that the following regularity result holds true for the above mentioned dual problem (for details kindly see Proposition 2.2 of [29]

$$\|\nabla \tilde{p}_h^m\| \leq C\|\xi\|. \tag{61}$$

By using integration by parts in time we can equivalently express (60) as:

$$(\partial_t \xi, \gamma \tilde{p}_h^m) + A_h(\tilde{p}_h^m, \xi) = (\gamma\xi, y_h^m(u) - y_h^m(\Pi_h u)) - (\gamma\xi, \hat{\phi}\tilde{p}_h^m), \quad \forall \xi \in V_h.$$

Now testing the above expression with $\xi = y_h^m(u) - y_h^m(\Pi_h u)$, we can write

$$\begin{aligned} &(\partial_t(y_h^m(u) - y_h^m(\Pi_h u)), \gamma\tilde{p}_h^m) + A_h(\tilde{p}_h^m, y_h^m(u) - y_h^m(\Pi_h u)) \\ &= (\gamma(y_h^m(u) - y_h^m(\Pi_h u)), y_h^m(u) - y_h^m(\Pi_h u)) - (\gamma(y_h^m(u) - y_h^m(\Pi_h u)), \hat{\phi}\tilde{p}_h^m). \end{aligned} \tag{62}$$

Employing the discrete state equation for $y_h^m(u)$ and $y_h^m(\Pi_h u)$, we obtain

$$\begin{aligned} &(\partial_t(y_h^m(u) - y_h^m(\Pi_h u)), \gamma\tilde{p}_h^m) + A_h(y_h^m(u) - y_h^m(\Pi_h u), \tilde{p}_h^m) \\ = \; &(u^m - \Pi_h u^m, \gamma\tilde{p}_h^m) - (\phi(y_h^m(u)) \\ &-\phi(y_h^m(\Pi_h u)), \gamma\tilde{p}_h^m). \end{aligned} \tag{63}$$

Using (62) and (63), we arrive at

$$\begin{aligned} &(\gamma(y_h^m(u) - y_h^m(\Pi_h u)), y_h^m(u) - y_h^m(\Pi_h u)) - (\gamma(y_h^m(u) - y_h^m(\Pi_h u)), \hat{\phi}\tilde{p}_h^m) \\ &-A_h(\tilde{p}_h^m, y_h^m(u) - y_h^m(\Pi_h u)) = (u^m - \Pi_h u^m, \gamma\tilde{p}_h^m) \\ &-(\phi(y_h^m(u)) - \phi(y_h^m(\Pi_h u)), \gamma\tilde{p}_h^m) - A_h(y_h^m(u) - y_h^m(\Pi_h u), \tilde{p}_h^m). \end{aligned}$$

Using the definition of the norm $|\!|\!|\cdot|\!|\!|_0$ and its equivalence with the norm $\|\cdot\|$ we find that

$$\begin{aligned} &\|y_h^m(u) - y_h^m(\Pi_h u)\|^2 \\ \leq \; &(u^m - \Pi_h u^m, \gamma\tilde{p}_h^m) + ((y_h^m(u) - y_h^m(\Pi_h u))\hat{\phi}, \gamma\tilde{p}_h^m) \\ &-(\phi(y_h^m(u)) - \phi(y_h^m(\Pi_h u)), \gamma\tilde{p}_h^m) + A_h(\tilde{p}_h^m, y_h^m(u) - y_h^m(\Pi_h u)) \\ &-A_h(y_h^m(u) - y_h^m(\Pi_h u), \tilde{p}_h^m). \end{aligned}$$

An application of the definition of $\hat{\phi}$ and property of L^2-projection Π_h in the above inequality gives us

$$\begin{aligned} &\|y_h^m(u) - y_h^m(\Pi_h u)\|^2 \\ \leq \; &(u^m - \Pi_h u^m, \gamma\tilde{p}_h^m) + A_h(\tilde{p}_h^m, y_h^m(u) - y_h^m(\Pi_h u)) \\ &-A_h(y_h^m(u) - y_h^m(\Pi_h u), \tilde{p}_h^m) \\ = \; &\underbrace{(u^m - \Pi_h u^m, \gamma\tilde{p}_h^m - \tilde{p}_h^m)}_{I_1} + \underbrace{(u^m - \Pi_h u^m, \tilde{p}_h^m - \Pi_h\tilde{p}_h^m)}_{I_2} \\ &+\underbrace{(A_h(\tilde{p}_h^m, y_h^m(u) - y_h^m(\Pi_h u)) - A_h(y_h^m(u) - y_h^m(\Pi_h u), \tilde{p}_h^m))}_{I_3}. \end{aligned} \tag{64}$$

First we use the approximation properties of γ (given in Lemma 15) and L^2-projection to bound I_1 and I_2, respectively as

$$\begin{aligned} |I_1| &\leq Ch\,\|u^m - \Pi_h u^m\|\,\|\tilde{p}_h^m\|_h\,, \\ |I_2| &\leq Ch\,\|u^m - \Pi_h u^m\|\,\|\tilde{p}_h^m\|_h\,. \end{aligned}$$

Then an application of (61) yields

$$|I_1| + |I_2| \leq Ch\,\|u^m - \Pi_h u^m\|\,\|y_h^m(u) - y_h^m(\Pi_h u)\|\,. \tag{65}$$

For I_3, we will use the relation (31) and (61) to obtain

$$\begin{aligned} |I_3| &\leq Ch\,\|y_h^m(u) - y_h^m(\Pi_h u)\|_h\,\|\tilde{p}_h^m\|_h \\ &\leq Ch\,\|y_h^m(u) - y_h^m(\Pi_h u)\|_h\,\|y_h^m(u) - y_h^m(\Pi_h u)\|\,. \end{aligned}$$

By following similar steps as in the proof of Lemma 3.1 we can obtain the relation

$$\|y_h^m(u) - y_h^m(\Pi_h u)\|_h \leq \|u - \Pi_h u\|_{L^2(I;L^2(\Omega))}\,, \tag{66}$$

which implies that $|I_3| \leq Ch\,\|u - \Pi_h u\|_{L^2(I;L^2(\Omega))}\,\|y_h^m(u) - y_h^m(\Pi_h u)\|$. Finally substituting the estimates for I_1, I_2 and I_3 in (65) and using Young's inequality, it is easy to see

$$\|y_h^m(u) - y_h^m(\Pi_h u)\| \leq Ch\,\|u^m - \Pi_h u^m\| + Ch\,\|u - \Pi_h u\|_{L^2(I;L^2(\Omega))}\,,$$

which can be equivalently expressed by summing over each interval I_m as

$$\|y_h(u) - y_h(\Pi_h u)\|_{L^2(I;L^2(\Omega))} \leq Ch\,\|u - \Pi_h u\|_{L^2(I;L^2(\Omega))}\,. \tag{67}$$

For the third term in (65), using (17) and proceeding with similar steps in the proof of Lemma 3.1 we can obtain

$$\begin{aligned} &\|y_h(\Pi_h u) - y_h\|_{L^2(I;L^2(\Omega))} \\ \leq\; &\|y_h(\Pi_h u) - y_h\|_{L^\infty(I;L^2(\Omega))} \leq \|y_h(\Pi_h u) - y_h\|_{L^\infty(I;V(h))} \\ \leq\; &C\,\|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))}\,. \end{aligned} \tag{68}$$

Now using the condition (13) for $\Pi_h u - u_h \in U_{h,ad} \subset L^2(I;L^2(\Omega))$, we have

$$\begin{aligned} &C\,\|\Pi_h u - u_h\|^2_{L^2(I;L^2(\Omega))} \\ \leq\; &j'(\Pi_h u)(\Pi_h u - u_h) - j'(u_h)(\Pi_h u - u_h) \\ \leq\; &(\alpha \Pi_h u + B^* p(\Pi_h u), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} - (\alpha u_h + B^* p(u_h), \\ &\Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\; &\alpha\,\|\Pi_h u - u_h\|^2_{L^2(I;L^2(\Omega))} - (B^* p(u_h) - B^* p(\Pi_h u), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))}\,. \end{aligned} \tag{69}$$

On using the variational inequality (26), projection property of Π_h and (33), we have the following relation

$$\begin{aligned} &\alpha\,\|\Pi_h u - u_h\|^2_{L^2(I;L^2(\Omega))} \\ =\; &\alpha(u - u_h, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\; &(B^* p_h - B^* p, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\; &(B^* p_h - B^* p(u_h), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} + (B^* p(u_h) - B^* p(\Pi_h u), \Pi_h u \\ &- u_h)_{L^2(I;L^2(\Omega))} + (B^* p(\Pi_h u) - B^* p, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} &\alpha\,\|\Pi_h u - u_h\|^2_{L^2(I;L^2(\Omega))} - (B^* p(u_h) - B^* p(\Pi_h u), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\; &(B^* p_h - B^* p(u_h), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \end{aligned}$$

$$+(B^*p(\Pi_h u) - B^*p, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))}. \tag{70}$$

For the first term of (70), using (47) and continuity of operator B, gives the estimate

$$\begin{aligned} & (B^*p_h - B^*p(u_h), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\ & C \|p_h - p(u_h)\|_{L^2(I;L^2(\Omega))} \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))} \\ \leq\ & C(h^2 + k) \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))}. \end{aligned} \tag{71}$$

An application of Lemma 3.2 in the second term of (70) gives

$$\begin{aligned} & (B^*p(\Pi_h u) - B^*p, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ =\ & (B^*p(\Pi_h u) - B^*p_h(\Pi_h u), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ & +(B^*p_h(\Pi_h u) - B^*p_h(u), \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ & +(B^*p_h(u) - B^*p, \Pi_h u - u_h)_{L^2(I;L^2(\Omega))} \\ \leq\ & C(h^2 + k) \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))} \\ & + \|p_h(y_h(\Pi_h u)) - p_h(y_h(u)\|_{L^2(I;L^2(\Omega))} \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))} \\ \leq\ & C(h^2 + k) \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))} \\ & + \|y_h(\Pi_h u) - y_h(u)\|_{L^2(I;L^2(\Omega))} \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))} \\ \leq\ & (C(h^2 + k) + Ch \|\Pi_h u - u\|_{L^2(I;L^2(\Omega))}) \|\Pi_h u - u_h\|_{L^2(I;L^2(\Omega))}, \end{aligned} \tag{72}$$

where, the last inequality follows from the proof of Lemma 3.1 and estimate (67). Using the estimates of (70), (71) and (72) in (69) and inserting it in (68), we can obtain

$$\|y_h(\Pi_h u) - y_h\|_{L^2(I;L^2(\Omega))} \leq C(h^2 + k). \tag{73}$$

Plugging (67) and (73) in (59), using the estimates of Lemma 3.2 and approximation of $\Pi_h u$, we can obtain the optimal order for piecewise constant or piecewise linear discretization of control

$$\|y - y_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k). \tag{74}$$

Now, using the results of Lemma 3.1, Lemma 3.2 and (74), we have

$$\begin{aligned} \|p - p_h\|_{L^2(I;L^2(\Omega))} & \leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + \|p_h(y) - p_h\|_{L^2(I;L^2(\Omega))} \\ & \leq \|p - p_h(y)\|_{L^2(I;L^2(\Omega))} + \|p_h(y) - p_h\|_{L^\infty(I;V(h))} \\ & \leq \|p - p_h(y)\| + \|y - y_h\|_{L^2(I;L^2(\Omega))} = \mathcal{O}(h^2 + k). \end{aligned}$$

□

4. Numerical Experiments

In this section, we present our numerical result to validate the theoretical error estimates derived for control, state and costate variables. For this purpose, we consider the following optimal control problem.

$$\min_{u \in U_{ad}} J(y, u) := \frac{1}{2} \int_0^1 \|(y(t, x) - y_d(t, x))\|^2 \, dt + \frac{\alpha}{2} \int_0^1 \|u(t, x)\|^2 \, dt,$$

subject to

$$\begin{aligned} \partial_t y - \nabla \cdot (A \nabla y) + y^3 &= u + f, \quad in \quad (0, 1] \times \Omega, \\ y(t, x) &= 0, \quad on \quad (0, 1] \times \partial\Omega, \end{aligned}$$

$$y(0,x) \quad = \quad x_1x_2(x_1-1)(x_2-1), \quad in \quad \Omega.$$

Here, the coefficient matrix

$$A = \begin{pmatrix} 1+x_1^2 & 0 \\ 0 & 1+x_2^2 \end{pmatrix},$$

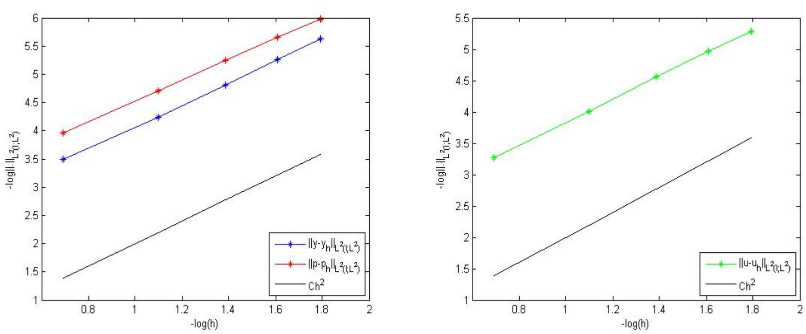
the regularization parameter $\alpha = 0.5$ and the space domain $\Omega = \{x = (x_1, x_2) : 0 \le x_1 \le 1, 0 \le x_2 \le 1\}$. The source term f and the desired state y_d will be of the form

$$\begin{aligned} f(t,x) \quad &= \quad e^t(x_1x_2(x_1-1)(x_2-1) - (2+6x_1^2-2x_1)(x_2^2-x_2) \\ &\quad -(x_1^2-x_1)(2+6x_2^2-2x_2^2)) \\ &\quad +e^{3t}x_1^3x_2^3(x_1-1)^3(x_2-1)^3 - u(t,x), \\ y_d(t,x) \quad &= \quad 2e^t(x_1^2-x_1)(x_2^2-x_2) + (e^t-e)((2+6x_1^2-2x_1)(x_2^2-x_2) \\ &\quad +(x_1^2-x_1)(2+6x_2^2-2x_2^2)) \\ &\quad -3e^{2t}(e^t-e)(x_1^2-x_1)^3(x_2^2-x_2)^3. \end{aligned}$$

For computing the order of convergence, we would require the exact solution of the above mentioned problem. Therefore, with the choice of the source term f and the desired state y_d, the exact state y and the adjoint state p will be given in the following manner

$$y(t,x) = e^t x_1x_2(x_1-1)(x_2-1), \quad p(t,x) = (e^t-e)x_1x_2(x_1-1)(x_2-1).$$

Moreover, the control variable is defined as: $u(t,x) = max(0, min(1, -\frac{1}{\alpha}p(t,x))$.



(a) Convergence of state and costate. (b) Convergence of control.

FIGURE 2. The convergence rates of the DFV approximations of the state, adjoint state and control variables with variational discretization approach computed with $\theta = -1$, $\beta = 1$ and $k = 0.01$.

The convergence of the approximate solutions measured by errors in discrete $L^2(I; L^2(\Omega))$ norm for state, costate and control variables and corresponding observed rates are defined as

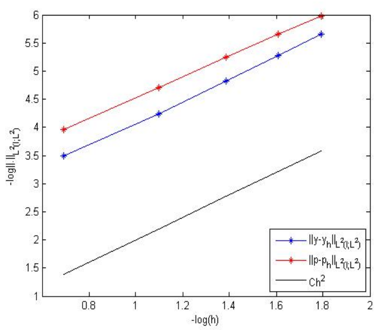
$$\begin{aligned} e_h(y) \quad &:= \quad \|y-y_h\|_{L^2(I;L^2(\Omega))}, e_h(p) := \|p-p_h\|_{L^2(I;L^2(\Omega))}, \\ e_h(u) \quad &:= \quad \|u-u_h\|_{L^2(I;L^2(\Omega))}, r_h(y) := \frac{\log(e_h(y)/\hat{e}_h(y))}{\log(h/\hat{h})}, \\ r_h(p) \quad &:= \quad \frac{\log(e_h(p)/\hat{e}_h(p))}{\log(h/\hat{h})}, r_h(u) := \frac{\log(e_h(u)/\hat{e}_h(u))}{\log(h/\hat{h})}. \end{aligned}$$

TABLE 1. Numerical results for error with $k = 0.01$ for state, adjoint state and control variables using variational discretization method for $\theta = -1$ and $\beta = 1$.

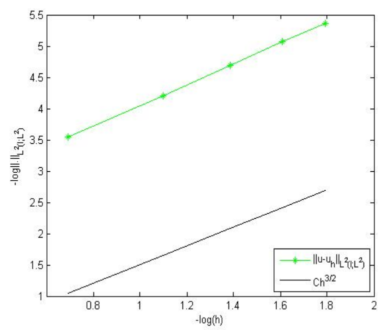
h	$e_h(y)$	$r_h(y)$	$e_h(p)$	$r_h(p)$	$e_h(u)$	$r_h(u)$
0.5000000	0.0306030	-	0.0190899	-	0.0381798	-
0.3333333	0.0143815	1.8624412	0.0090646	1.8368500	0.0181292	1.8368500
0.2500000	0.0081559	1.9716095	0.0052642	1.8890259	0.0105285	1.8890259
0.2000000	0.0052157	2.0034798	0.0034871	1.8457554	0.0069742	1.8457554
0.1666667	0.0036108	2.0169715	0.0025275	1.7651376	0.0050551	1.7651376

Here e and $\hat{e}$ denote errors computed on two consecutive meshes of sizes h and $\hat{h}$, respectively.

For variational discretization approach, the difference between the computed solution and the exact solution with respect to the discrete $L^2(I; L^2(\Omega))$ norm for state, co-state and control variables with a fixed time step $k = 0.01$ have been reported in Table 1. Further from Figure 2, we observe that the rate of convergence for state, co-state and control variables is of order h^2 which matches with the theoretical rate of convergence derived in Theorem 3.3.



(a) Convergence of state and costate.



(b) Convergence of control.

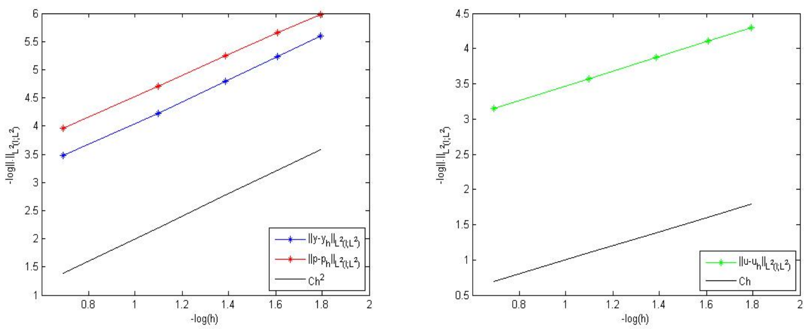
FIGURE 3. The convergence rates of the DFV approximations of the state, adjoint state and control variables with piecewise linear discretization of control for $\theta = -1$, $\beta = 1$ and $k = 0.01$.

TABLE 2. Numerical results for error with $k = 0.01$ for state, adjoint state and control variables using piecewise linear discretization of control for $\theta = -1$ and $\beta = 1$.

h	$e_h(y)$	$r_h(y)$	$e_h(p)$	$r_h(p)$	$e_h(u)$	$r_h(u)$
0.5000000	0.0305763	-	0.0190907	-	0.0287159	-
0.3333333	0.0143466	1.8662814	0.0090657	1.8366436	0.0149358	1.6121916
0.2500000	0.0080601	2.0042228	0.0052682	1.8868466	0.0091847	1.6901356
0.2000000	0.0051291	2.0255513	0.0034908	1.8442640	0.0062928	1.6945581
0.1666667	0.0035176	2.0685532	0.0025319	1.7615460	0.0046740	1.6310945

When piecewise linear discretization is used for control, for a fixed time step $k = 0.01$, the computed order of convergence for state, co-state and control variables has been shown in Figure 3 which is of $\mathcal{O}(h^2)$ (for state and co-state) and

$\mathcal{O}(h^{3/2})$ for control. We note that this also matches with the theoretical rate of convergence derived in Theorem 3.5 and Theorem 3.7. Similarly, the computed order of convergence and difference between errors are shown in Figure 4 and Table 3, respectively, when piecewise constant discretization approach is used for control. In this case also our theoretical order of convergence matches with the computed rate of convergence (for a fixed time step $k = 0.01$) derived in Theorem 3.6 and 3.7.



(a) Convergence of state and costate. (b) Convergence of control.

FIGURE 4. The convergence rates of the DFV approximations of the state, co-state and control variables using piecewise discretization of control which are computed for $\theta = -1$, $\beta = 1$ and $k = 0.01$.

TABLE 3. Computational error with $k = 0.01$ for state, co-state and control variables using piecewise constant discretization of control for $\theta = -1$ and $\beta = 1$.

h	$e_h(y)$	$r_h(y)$	$e_h(p)$	$r_h(p)$	$e_h(u)$	$r_h(u)$
0.5000000	0.0308900	-	0.0190792	-	0.04275929	-
0.3333333	0.0146205	1.8448129	0.0090547	1.8381477	0.02807274	1.0377892
0.2500000	0.0083204	1.9594710	0.0052569	1.8901131	0.02074522	1.0514493
0.2000000	0.0053314	1.9946630	0.0034816	1.8465432	0.01644297	1.0415617
0.1666667	0.0036954	2.0102921	0.0025233	1.7656141	0.01362724	1.0301967

5. Concluding Remarks

In this article, a discontinuous finite volume method is used for the approximation of state and costate variables and three different techniques (variational discretization, piecewise linear and constant discretization) are employed for the approximation of the control variable. We stress that deriving the optimal error estimate for state and co-state variables in $L^2(I; L^2(\Omega))$-norm for variational discretization of control, is not a tough task and one can achieve this by decomposing the error. However, by following the same arguments for deriving the error estimates in $L^2(I; L^2(\Omega))$-norm for state and co-state variables would lead to a suboptimal rate of convergence, when piecewise linear and piecewise constant discretizations are used for the control. This is because in this case the rate of convergence is of $\mathcal{O}(h^{3/2} + k)$ and $\mathcal{O}(h + k)$ for piecewise linear and constant discretizations, respectively. To overcome this difficulty, duality arguments have been used for the establishment of optimal error estimates in $L^2(I; L^2(\Omega))$-norm for state and co-state variables. Further, numerical experiments have been reported to make sure the performance of various proposed numerical schemes and to support the theoretical findings. In the light of our theoretical error estimates derived in Section 3 and discretization approaches used in Section 2, the computational and theoretical advantages and disadvantages over each other for three approaches (variational, piecewise linear and constant) used for discretization of control can be explained as follows: For variational discretization, even though we have $\mathcal{O}(h^2 + k)$ rate of convergence for control in $L^2(I; L^2(\Omega))$-norm but there would be computational difficulties, as in this approach the approximation of the control variable does not lie in the finite dimensional space associated with triangulation and this would lead to a nonstandard numerical algorithms and involvement of more sophisticated stoping criteria. Whereas, for piecewise linear and constant approaches, the convergence is of $\mathcal{O}(h^{3/2} + k)$ and $\mathcal{O}(h + k)$, respectively.

Acknowledgment

The authors would like to thank the referees for their constructive comments and suggestions, which led to a number of improvements to this manuscript.

References

[1] D. N. Arnold, An interior penalty finite element method with discontinous elements, SIAM J. Numer. Anal. (1982), 19:742–760.
[2] D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. (2002), 39:1749–1779.
[3] M. Braack, Optimal control in fluid mechanics by finite elements with symmetric stabilization, SIAM J. Control Optim. (2009), 48:672–687.
[4] C. Bi and J. Geng, Discontinuous finite volume element method for parabolic problems, Numer. Methods Part. Diff. Eq. (2010), 26:367–383.
[5] C. Bi and M. Liu, A Discontinuous finite volume element method for second-order elliptic problems, Numer. Methods Part. Diff. Eq. (2012), 28:425–440.
[6] R. Bürger, S. Kumar, R. Ruiz-Baier, Discontinuous finite volume element discretization for coupled flow-transport problems arising in models of sedimentation, J. Comput. Phys. (2015), 299:446–471.
[7] Z. Cai, On the finite volume element method, Numer. Math. (1991), 58:713–735.
[8] E. Casas and F. Tröltzsch, Error estimates for the finite-element approximation of a semilinear elliptic control problem, Control and Cybernetics (2002), 31:695–712.
[9] E. Casas and F. Tröltzsch, Error estimates for linear-quadratic elliptic control problems, IFIP: Anal. and Optimization of Diff. Systems (2003), 121:89–100.
[10] S.H. Chou, D.Y. Kwak and Q. Li, L^p error estimates and superconvergence for covolume or finite volume element methods, Numer. Meth. Part. Diff. Eqns. (2003), 19:463–486.
[11] S.H. Chou and X. Ye, Unified analysis of finite volume methods for second order elliptic problems, SIAM J. Numer. Anal. (2007), 45:1639–1653.
[12] R.E. Ewing, T. Lin and Y. Lin, On the accuracy of the finite volume element method based on piecewise linear polynomials, SIAM J. Numer. Anal. (2002), 39:1865–1888.
[13] W. Gong and N. Yan, Robust error estimates for the finite element approximation of elliptic optimal control problems, J. Comput. Appl. Math.(2011), 236:1370–1381.
[14] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case , Computational Optimization and Applications (2005), 30:45–61.
[15] J. Huang and S. Xi, On the finite volume element method for general self-adjoint elliptic problems, SIAM J. Numer. Anal. (1998), 35:1762-1774.
[16] S. Kumar, N. Nataraj and A.K. Pani, Discontinuous Galerkin finite volume element methods for second order linear elliptic problems, Numer. Meth. Part. Diff. Eqns. (2009), 25:1402–1424.
[17] S. Kumar, A mixed and discontinuous Galerkin finite volume element method for incompressible miscible displacement problems in porous media, Numer. Meth. Part. Diff. Eqns. (2012), 28:1354–1381.

[18] S. Kumar and R. Ruiz-Baier, Equal order discontinuous finite volume element methods for the Stokes problem, J. Sci. Comput. (2015), in press.
[19] Y. Lin, J. Liu and M. Yang, Finite volume element methods: An overview on recent developments, Int. J. Numer. Anal. Model. B (2013), 4:14–24.
[20] J.L. Lions, Optimal control of systems governed by partial differential equations, Springer Verlag, Berlin(1971).
[21] J. Liu, L. Mu and X. Ye, An adaptive discontinuous finite volume element method for elliptic problems, J. Comput. Appl. Math. (2011), 235:5422–5431.
[22] J. Liu, L. Mu, X. Ye and R. Jari, Convergence of the discontinuous finite volume method for elliptic problems with minimal regularity, J. Comput. Appl. Math. (2012), 236:4537–4546.
[23] Zuliang Lu, Variational discretization and mixed methods for semilinear parabolic optimal control problem., Gen. Math. Notes, ICRS Publication (2011), 4:16–27.
[24] X. Luo, Y. Chen and Y. Huang, Some error estimates of finite volume element approximation for elliptic optimal control problems, Int. J. Num. Anal. Model. (2013), 10:697–711.
[25] X. Luo, Y. Chen, Y. Huang and T. Hou, Some error estimates of finite volume element approximation for parabolic optimal control problems, Optim. Control Appl. Meth. (2014), 35:145–165.
[26] D. Meidner and B. Vexler A priori error estimates for space-time finite element discretization of parabolic optimal control problems part I: problems without control constraints, SIAM J. Control and Optimization. (2008), 47:1150–1177.
[27] D. Meidner and B. Vexler A priori error estimates for space-time finite element discretization of parabolic optimal control problems part II: problems with control constraints, SIAM J. Control and Optimization (2008), 47:1301–1329.
[28] I.D. Mishev Finite volume and finite volume element methods for non symmetric problems, Ph.D thesis, Technical report ISC-96-04-MATH, Institute of Scientific Computation, Texas A & M University, College Station, TX, (1997).
[29] I. Neitzel and B. Vexler A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems, Numer. Math. (2012), 120:345–386.
[30] S. Prudhomme, F. Pascal, J.T. Oden and A. Romkes, A priori error estimation for discontinuous Galerkin methods, TICAM Report, (2000).
[31] B. Rivière, M.F. Wheeler and V. Girault, A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems, SIAM J. Numer. Anal. (2002), 39:902–931.
[32] A. Rösch and S. Steinig, A priori error estimates for a state-constrained elliptic optimal control problems, ESIAM: Mathem. Model. and Numer. Anal. (2012), 46:1107–1120.
[33] A. Rösch, Error estimates for parabolic optimal control problems with control constraints, Z. Anal. Anwendungen (2004), 23:353–376.
[34] F. Tröltzsch, Optimal control of partial differential equations: theory, methods and applications, American Math. Society, book series Graduate Studies in Mathematics (2010).
[35] F. Tröltzsch, On finite element error estimates for optimal control problems with elliptic PDEs, Springer, Large Scale Scientific Computing, Lecture Notes (2010), 5910:40–53.
[36] R. Winther, Error estimates for a Galerkin approximation of a parabolic control problem, Ann. Math. Pura Appl. (1978), 4:173–206.
[37] Q. Yang and Z. Jiang, A discontinuous mixed covolume method for elliptic problems, J. Comput. Appl. Math. (2011), 235:2467–2476.
[38] X. Ye, A new discontinuous finite volume method for elliptic problems, SIAM J. Numer. Anal. (2004), 42:1062–1072.
[39] X. Ye, A discontinuous finite volume method for the Stokes problems, SIAM J. Numer. Anal. (2006), 44:183–198.
[40] M. Cui and X. Ye, Unified analysis of finite volume methods for the Stokes equations, SIAM J. Numer. Anal. (2010), 48:824–839.

Department of Mathematics, Indian Institute of Space Science and Technology, Thiruvananthapuram 695 547, Kerala , India

E-mail: `ruchisandilya.12@iist.ac.in`

Department of Mathematics, Indian Institute of Space Science and Technology, Thiruvananthapuram 695 547, Kerala , India

E-mail: `sarvesh@iist.ac.in`