# AN $h p$-DISCONTINUOUS GALERKIN METHOD FOR THE OPTIMAL CONTROL PROBLEM OF LASER SURFACE HARDENING OF STEEL 

Gupta Nupur ${ }^{1}$ and Nataraj Neela ${ }^{1}$


#### Abstract

In this paper, we discuss an $h p$-discontinuous Galerkin finite element method ( $h p$-DGFEM) for the laser surface hardening of steel, which is a constrained optimal control problem governed by a system of differential equations, consisting of an ordinary differential equation for austenite formation and a semi-linear parabolic differential equation for temperature evolution. The space discretization of the state variable is done using an $h p$-DGFEM, time and control discretizations are based on a discontinuous Galerkin method. A priori error estimates are developed at different discretization levels. Numerical experiments presented justify the theoretical order of convergence obtained.


Mathematics Subject Classification. 65N12, 65N30, 65M12, 93C20.
Received January 25, 2010. Revised March 19, 2011.
Published online June 28, 2011.

## 1. Introduction

In most structural components in mechanical engineering, the surface is stressed. The purpose of surface hardening is to increase the hardness of the boundary layer of a work piece by rapid heating and subsequent quenching (see Fig. 1). The desired hardening effect is achieved as the heat treatment leads to a change in micro structure. A few applications include cutting tools, wheels, driving axles, gears, etc.

The mathematical model for the laser surface hardening of steel has been studied in [20,26]. For an extensive survey on mathematical models for laser material treatments, we refer to [27]. In this article, we follow the Leblond-Devaux model [26]. In $[1,20]$, the mathematical model for the laser hardening problem which gives rise to an optimal control problem governed by a system of nonlinear parabolic equations and a set of ordinary differential equations with a non differentiable right hand side function is discussed. The authors have regularised the right hand side function and have established results on existence, regularity and stability. This approach seems to be common in all subsequent literature not only for existence results but also for numerical approximations. In [15], the convergence of the solution of the regularized problem to that of the original problem has been established. In [19], laser and induction hardening has been used to explain the model and then a finite volume method has been used for the space discretization in the numerical approximation. In [21], the optimal control problem is analyzed and error estimates for proper orthogonal decomposition (POD) Galerkin method for the state system are derived. Also a penalized problem has been considered for the purpose of numerical

[^0]

Figure 1. Laser hardening process.
simulations. In [36], a finite element scheme combined with a nonlinear conjugate gradient method has been used to solve the optimal control problem and a finite element method has been used for the purpose of space discretization. In [16], a priori error estimates are developed for a finite element scheme in which the space discretization is done using conformal finite elements, whereas the time and control discretizations are based on a discontinuous Galerkin method.

In literature, a substantial amount of work on the a priori error estimates for linear and non linear parabolic problems are available, see for example $[7,9,10,35]$ to mention a few. For optimal control problems governed by linear parabolic equations without control constraints, a priori error bounds are developed in [28].

In recent years, there has been a renewed interest in DGFEM for the numerical solution of a wide range of partial differential equations. This is due to their flexibility in local mesh adaptivity and in handling nonuniform degrees of approximation for solutions whose smoothness exhibit variation over the computational domain. Besides, they are elementwise conservative.

The use of DGFEM for elliptic and parabolic problems started with the work of Douglas and Dupont [8] and Wheeler [37] in the 70's. These methods are generalization of work by Nitsche [29] for treating Dirichlet boundary condition by introduction of a penalty term on the boundary. In 1973, Babuška [4] introduced another penalty method to impose the Dirichlet boundary condition weakly. Interior Penalty (IP) methods by Arnold [2] and Wheeler [37] arose from the observations that just as Dirichlet boundary conditions, interior element continuity can be imposed weakly instead being built into the finite element space. This makes it possible and easier to use the space of discontinuous piecewise polynomials of higher degree. The IP methods are presently called as Symmetric Interior Penalty Galerkin (SIPG) methods. The variational form of SIPG method is symmetric and adjoint consistent, but the stabilizing penalty parameter in these methods depends on the bounds of the coefficients of the problem and various constants in the inverse inequalities which are not known explicitly. To overcome this, Oden et al. [30] proposed a DGFEM which is based on a non-symmetric formulation for advection diffusion problems. This method is known to be stable when the degree of approximation is greater or equal to 2 , see $[30,34]$. In Houston et al. [23], $h p$ discontinuous finite element methods are studied for diffusion reaction problems. For a review of work on DG methods for elliptic problems, we refer to [3,31]. [12-14] discuss DG methods for quasilinear and strongly non-linear elliptic problems. In [33,34], a non-symmetric interior penalty DGFEM is analyzed for elliptic and non-linear parabolic problems, respectively. An $h p$-version of interior penalty discontinuous Galerkin method for semilinear parabolic equation with mixed Dirichlet and Neumann boundary conditions has been analyzed in [24]. Error estimates are derived under hypothesis on regularity of the solution. DGFEM and corresponding error estimates for continuous and discrete time, for non-linear parabolic equations, have been developed in [33]. For a detailed description of DGFEM for elliptic and parabolic problems, we refer to [32].

In this paper, we discuss a DGFEM for the optimal control problem of laser surface hardening of steel. Since the temperature around the boundary of the computational domain of the laser surface hardening of steel problem is higher than at the other parts of the domain, a non-uniformity in the triangulation of domain becomes
relevant (see Fig. 1). DGFEM is effective here because of the ease in the choice of finite element spaces with discontinuous polynomials of higher degrees. Also, a finer triangulation near the boundary region permitting hanging nodes helps to yield better results. The laser surface hardening of steel problem being an optimal control problem, adjoint consistency becomes important. Therefore, in this paper, a symmetric version of $h p$-DGFEM has been introduced and analyzed. To state more precisely, we apply an $h p$-DGFEM for the discretization of space and a DGFEM for time and control variables. A priori error estimates have been developed for the temperature and austenite variables at different discretization levels and numerical experiments are performed to justify the theoretical results obtained.

The outline of this paper is as follows. This section is introductory in nature. In Section 2, a weak formulation of the regularized laser surface hardening of steel problem is presented. In Section 3, an hp-DGFEM weak formulation for the laser surface hardening of steel problem with its adjoint system is presented. Also, error estimates are developed for the state and the adjoint variables. In Section 4, a space-time discretization using DGFEM in time and an $h p$-DGFEM in space has been done. Also, a completely discrete formulation is derived using DGFEM for control variable. Error estimates are developed for space-time and completely discrete schemes. In Section 5, results of numerical experiments are presented.

## 2. The laser surface hardening of steel problem

Let $\Omega \subset \mathbb{R}^{2}$, denoting the workpiece, be a convex, bounded domain with piecewise Lipschitz continuous boundary $\partial \Omega, Q=\Omega \times I$ and $\Sigma=\partial \Omega \times I$, where $I=(0, T), T<\infty$. Following Leblond and Devaux [26], the evolution of volume fraction of austenite $a(t)$ for a given temperature evolution $\theta(t)$ is described by the following initial value problem:

$$
\begin{align*}
\partial_{t} a=f_{+}(\theta, a) & =\frac{1}{\tau(\theta)}\left[a_{e q}(\theta)-a\right]_{+} \quad \text { in } Q,  \tag{2.1}\\
a(0) & =0 \quad \text { in } \Omega \tag{2.2}
\end{align*}
$$

where $a_{e q}(\theta(t))$, denoted as $a_{e q}(\theta)$ for notational convenience, is the equilibrium volume fraction of austenite and $\tau$ depends only on the temperature $\theta$. The term $\left[a_{e q}(\theta)-a\right]_{+}=\left(a_{e q}(\theta)-a\right) \mathcal{H}\left(a_{e q}(\theta)-a\right)$, where $\mathcal{H}$ is the Heaviside function

$$
\mathcal{H}(s)=\left\{\begin{array}{lc}
1 & s>1 \\
0 & s \leq 0
\end{array}\right.
$$

denotes the non-negative part of $a_{e q}(\theta)-a$, that is, $\left[a_{e q}(\theta)-a\right]_{+}=\frac{\left(a_{e q}(\theta)-a\right)+\left|a_{e q}(\theta)-a\right|}{2}$.
Neglecting the mechanical effects and using the Fourier law of heat conduction, the temperature evolution can be obtained by solving the non-linear energy balance equation given by

$$
\begin{align*}
\rho c_{p} \partial_{t} \theta-\mathcal{K} \Delta \theta & =-\rho L f_{+}(\theta, a)+\alpha u \quad \text { in } Q  \tag{2.3}\\
\theta(0) & =\theta_{0} \quad \text { in } \Omega  \tag{2.4}\\
\frac{\partial \theta}{\partial n} & =0 \quad \text { on } \Sigma \tag{2.5}
\end{align*}
$$

where the density $\rho$, the heat capacity $c_{p}$, the thermal conductivity $\mathcal{K}$ and the latent heat $L$ are assumed to be positive constants. The term $u(t) \alpha(x, t)$ describes the volumetric heat source due to laser radiation, $u(t)$ being the time dependent control variable. Since the main cooling effect is the self cooling of the workpiece, homogeneous Neumann conditions are assumed on the boundary. Also, $\theta_{0}$ denotes the initial temperature.

To maintain the quality of the workpiece surface, it is important to avoid the melting of surface. In the case of laser hardening, it is a quite delicate problem to obtain parameters that avoid melting but nevertheless lead to the right amount of hardening. Mathematically, this corresponds to an optimal control problem in which we


Figure 2. Regularized heaviside $\left(\mathcal{H}_{\epsilon}(s)\right)$ and heaviside $(\mathcal{H}(s))$ functions.
minimize the cost functional defined by:

$$
\begin{equation*}
J(\theta, a, u)=\frac{\beta_{1}}{2} \int_{\Omega}\left|a(T)-a_{d}\right|^{2} \mathrm{~d} x+\frac{\beta_{2}}{2} \int_{0}^{T} \int_{\Omega}\left[\theta-\theta_{m}\right]_{+}^{2} \mathrm{~d} x \mathrm{~d} s+\frac{\beta_{3}}{2} \int_{0}^{T}|u|^{2} \mathrm{~d} s \tag{2.6}
\end{equation*}
$$

subject to the state equations (2.1)-(2.5) in the set of admissible controls $U_{a d}$,
where $U_{a d}=\left\{u \in U:\|u\|_{L^{2}(I)} \leq M\right\}$ is the closed, bounded and convex subset of $U=L^{2}(I)$, denoting the admissible intensities, $\beta_{1}, \beta_{2}$ and $\beta_{3}$ being positive constants and $a_{d}$ being the given desired fraction of the austenite. The second term in (2.6) is a penalizing term that penalizes the temperature above the melting temperature $\theta_{m}$. For theoretical, as well as computational reasons, the term $\left[a_{e q}-a\right]_{+}$in (2.1) is regularized (see Fig. 2) and the regularized laser surface hardening problem is given by:

$$
\begin{align*}
& \min _{u_{\epsilon} \in U_{a d}} J\left(\theta_{\epsilon}, a_{\epsilon}, u_{\epsilon}\right) \text { subject to }  \tag{2.7}\\
& \partial_{t} a_{\epsilon}=f_{\epsilon}\left(\theta_{\epsilon}, a_{\epsilon}\right)=\frac{1}{\tau\left(\theta_{\epsilon}\right)}\left(a_{e q}\left(\theta_{\epsilon}\right)-a_{\epsilon}\right) \mathcal{H}_{\epsilon}\left(a_{e q}\left(\theta_{\epsilon}\right)-a_{\epsilon}\right) \quad \text { in } Q,  \tag{2.8}\\
& a_{\epsilon}(0)=0 \quad \text { in } \Omega  \tag{2.9}\\
& \rho c_{p} \partial_{t} \theta_{\epsilon}-\mathcal{K} \Delta \theta_{\epsilon}=-\rho L f_{\epsilon}\left(\theta_{\epsilon}, a_{\epsilon}\right)+\alpha u_{\epsilon} \quad \text { in } Q  \tag{2.10}\\
& \theta_{\epsilon}(0)=\theta_{0} \quad \text { in } \Omega  \tag{2.11}\\
& \frac{\partial \theta_{\epsilon}}{\partial n}=0 \quad \text { on } \Sigma \tag{2.12}
\end{align*}
$$

where $\mathcal{H}_{\epsilon} \in C^{1,1}(\mathbb{R})$ is a monotone approximation of the Heaviside function satisfying $\mathcal{H}_{\epsilon}(x)=0$ for $x \leq 0$.
We now make the following assumptions [21]:
(A1) $a_{e q}(x) \in(0,1)$ for all $x \in \mathbb{R}$ and $\left\|a_{e q}\right\|_{C^{1}(\mathbb{R})} \leq c_{a}$;
(A2) $0<\tau \leq \tau(x) \leq \bar{\tau}$ for all $x \in \mathbb{R}$ and $\|\tau\|_{C^{1}(\mathbb{R})} \leq c_{\tau}$;
(A3) $\theta_{0} \in H^{1}(\Omega), \theta_{0} \leq \theta_{m}$ a.e. in $\Omega$, where the constant $\theta_{m}>0$ denotes the melting temperature of steel;
(A4) $\alpha \in L^{\infty}(Q)$;
(A5) $u \in L^{2}(I)$;
(A6) $a_{d} \in L^{\infty}(\Omega)$ with $0 \leq a_{d} \leq 1$ a.e. in $\Omega$.

Since we will be discretizing the regularized problem in this paper, $\left(\theta_{\epsilon}, a_{\epsilon}, u_{\epsilon}\right)$ and $f_{\epsilon}$ will be replaced by $(\theta, a, u)$ and $f$, respectively, for the sake of notational simplicity.

Let $X=\left\{v \in L^{2}(I ; V): v_{t} \in L^{2}\left(I ; V^{*}\right)\right\}$ and $Y=H^{1}\left(I ; L^{2}(\Omega)\right)$, where $V=H^{1}(\Omega)$ [11]. Together with $H=L^{2}(\Omega)$, the Hilbert space $V$ and its dual $V^{*}$ build a Gelfand triplet $V \hookrightarrow H \hookrightarrow V^{*}$. The duality pairing between $V$ and its dual $V^{*}$ is denoted by $\langle\cdot, \cdot\rangle=\langle\cdot, \cdot\rangle_{V^{*} \times V}$. Also, let $(\cdot, \cdot)$ (resp. $(\cdot, \cdot)_{I, \Omega}$ ), and $\|\cdot\|\left(\right.$ resp. $\left.\|\cdot\|_{I, \Omega}\right)$ denote the inner product and norm in $L^{2}(\Omega)$ (resp. $L^{2}\left(I, L^{2}(\Omega)\right)$ ). The inner product and norm in $L^{2}(I)$ are denoted by $(\cdot, \cdot)_{L^{2}(I)}$ and $\|\cdot\|_{L^{2}(I)}$, respectively.

The weak formulation corresponding to (2.8)-(2.12), for a fixed $u \in U_{a d}$, reads as: Find $(\theta, a) \in X \times Y$ such that

$$
\begin{align*}
\left(\partial_{t} a, w\right) & =(f(\theta, a), w) \quad \forall w \in H,  \tag{2.13}\\
a(0) & =0,  \tag{2.14}\\
\rho c_{p}\left\langle\partial_{t} \theta, v\right\rangle+\mathcal{K}(\nabla \theta, \nabla v) & =-\rho L(f(\theta, a), v)+(\alpha u, v) \quad \forall v \in V,  \tag{2.15}\\
\theta(0) & =\theta_{0} . \tag{2.16}
\end{align*}
$$

Therefore, the weak formulation for the optimal control problem can be stated as

$$
\begin{equation*}
\min J(\theta, a, u) \quad \text { subject to the constraints }(2.13)-(2.16) \text { and } u \in U_{a d} . \tag{2.17}
\end{equation*}
$$

The existence of a unique solution to the state equation (2.13)-(2.16) ensures the existence of a control-to-state mapping $u \longmapsto(\theta, a)=(\theta(u), a(u))$ through (2.13)-(2.16). By means of this mapping, we introduce the reduced cost functional $j: U_{a d} \longrightarrow \mathbb{R}$ as

$$
\begin{equation*}
j(u)=J(\theta(u), a(u), u) \tag{2.18}
\end{equation*}
$$

Then the optimal control problem (2.17) can be equivalently reformulated as

$$
\begin{equation*}
\min _{u \in U_{a d}} j(u) . \tag{2.19}
\end{equation*}
$$

The following theorem ([36], Thm. 2.1) ensures the existence of a unique solution of the system (2.13)-(2.16).
Theorem 2.1. Suppose (A1)-(A6) are satisfied. Then, the system (2.13)-(2.16) has a unique solution

$$
(\theta, a) \in H^{1,1} \times W^{1, \infty}\left(I ; L^{\infty}(\Omega)\right)
$$

where $H^{1,1}=L^{2}\left(I ; H^{1}(\Omega)\right) \cap H^{1}\left(I ; L^{2}(\Omega)\right)$. Moreover, a satisfies

$$
0 \leq a<1 \text { a.e. in } Q .
$$

Remark 2.2 [36]. Due to (A1)-(A2), Theorem 2.1 and the definition of the regularized Heaviside function $\mathcal{H}_{\epsilon}$, there exists a constant $c_{f}>0$ independent of $\theta$ and $a$ such that

$$
\max \left(\|f(\theta, a)\|_{L^{\infty}(Q)},\left\|f_{a}(\theta, a)\right\|_{L^{\infty}(Q)},\left\|f_{\theta}(\theta, a)\right\|_{L^{\infty}(Q)}\right) \leq c_{f}
$$

for all $(\theta, a) \in L^{2}(Q) \times L^{\infty}(Q)$ and $(\theta, a)$ satisfying (2.13)-(2.16) for a fixed control $u \in U_{a d}$.
The existence of the optimal control is guaranteed by the following theorem ([36], Thm. 2.3).
Theorem 2.3. Suppose that (A1)-(A6) hold true. Then the optimal control problem (2.7)-(2.12) has at least one (global) solution.

In [15], convergence of the solution of regularized problem to the solution of the original problem has been established.

Now, we state a lemma [21] which establishes the existence and uniqueness of the solution of the adjoint problem.

Lemma 2.4. Let (A1)-(A6) hold true and $\left(\theta^{*}, a^{*}, u^{*}\right) \in X \times Y \times U_{\text {ad }}$ be a solution to (2.17). Then there exists a unique solution $\left(z^{*}, \lambda^{*}\right) \in H^{1,1} \times H^{1}\left(I, L^{2}(\Omega)\right)$ of the adjoint system

$$
\begin{align*}
-\left(\psi, \partial_{t} \lambda^{*}\right)+\left(\psi, f_{a}\left(\theta^{*}, a^{*}\right) g\left(z^{*}, \lambda^{*}\right)\right) & =0 \quad \forall \psi \in H, \text { a.e. in } I,  \tag{2.20}\\
\lambda^{*}(T) & =\beta_{1}\left(a^{*}(T)-a_{d}\right)  \tag{2.21}\\
-\rho c_{p}\left(\phi, \partial_{t} z^{*}\right)+\mathcal{K}\left(\nabla \phi, \nabla z^{*}\right)+\left(\phi, f_{\theta}\left(\theta^{*}, a^{*}\right) g\left(z^{*}, \lambda^{*}\right)\right)= & \beta_{2}\left(\phi,\left[\theta^{*}-\theta_{m}\right]_{+}\right)  \tag{2.22}\\
& \forall \phi \in V, \text { a.e. in } I, \\
z^{*}(T)= & 0 . \tag{2.23}
\end{align*}
$$

Here, $g\left(z^{*}, \lambda^{*}\right)=\rho L z^{*}-\lambda^{*}$. Moreover, $z^{*}$ satisfies the following variational inequality

$$
\begin{equation*}
\left(\beta_{3} u^{*}+\int_{\Omega} \alpha z^{*} \mathrm{~d} x, p-u^{*}\right)_{L^{2}(I)} \geq 0 \quad \forall p \in U_{a d} \tag{2.24}
\end{equation*}
$$

We now establish a regularity result for $\theta$.
Lemma 2.5. Under the assumptions (A1)-(A6), the solution $(\theta, a)$ of (2.1)-(2.5) satisfies:

$$
\|\Delta \theta\|_{I, \Omega} \leq C
$$

where $C>0$ is a constant.
Proof. Multiply (2.3) by $-\Delta \theta$ and then integrate over $\Omega \times[0, T]$ to obtain

$$
\begin{equation*}
-\rho c_{p} \int_{0}^{T}\left(\partial_{t} \theta, \Delta \theta\right) \mathrm{d} s+\mathcal{K} \int_{0}^{T}\|\Delta \theta\|^{2} \mathrm{~d} s=\rho L \int_{0}^{T}\left(\partial_{t} a, \Delta \theta\right) \mathrm{d} s-\int_{0}^{T}(\alpha u, \Delta \theta) \mathrm{d} s . \tag{2.25}
\end{equation*}
$$

Use Cauchy Schwarz and Young's inequality to obtain

$$
\|\Delta \theta\|_{I, \Omega}^{2} \leq C\left(\left\|\partial_{t} a\right\|_{I, \Omega}^{2}+\|u\|_{L^{2}(I)}^{2}+\left\|\partial_{t} \theta\right\|_{I, \Omega}^{2}+\sigma\|\Delta \theta\|_{I, \Omega}^{2}\right)
$$

where $C=\max \left\{\rho c_{p}, \rho L, \max _{Q}|\alpha(x, t)|\right\}$. Choosing Young's constant $\sigma>0$ appropriately and using $(\theta, a) \in$ $H^{1,1} \times W^{1, \infty}\left(I, L^{\infty}(\Omega)\right)$ we obtain the required result and this completes the rest of the proof.

Remark 2.6. The constant $C>0$ will be used to denote different values at different steps throughout the paper and is a generic one.

## 3. An $h p$-Discontinuous Galerkin formulation

First we state some preliminaries which are essential in the sequel.

## Broken spaces

Let $\mathcal{T}_{h}=\{K, K \subset \Omega\}$ be a shape regular finite element subdivision of $\Omega$ in the sense that there exists $\gamma>0$ such that if $h_{K}$ is the diameter of $K$, then $K$ contains a ball of radius $\gamma h_{K}$ in its interior [6]. Each element $K$ is a rectangle/triangle defined as follows. Let $\hat{K}$ be a shape regular master rectangle/triangle in $\mathbb{R}^{2}$, and let $\left\{F_{K}\right\}$ be a family of invertible maps such that each $F_{K}$ maps from $\hat{K}$ to $K$. Let $h=\max _{K \in \mathcal{T}_{h}} h_{K}$.

Let $\mathcal{E}, \mathcal{E}_{\text {int }}$ and $\mathcal{E}_{\partial}$ be the set of all the edges, interior edges and boundary edges of the elements, respectively, defined as follows:

$$
\begin{aligned}
\mathcal{E} & =\left\{e: e=\partial K \cap \partial K^{\prime} \text { or } e=\partial K \cap \partial \Omega, K, K^{\prime} \in \mathcal{T}_{h}\right\} \\
\mathcal{E}_{\text {int }} & =\left\{e \in \mathcal{E}: e=\partial K \cap \partial K^{\prime}, K, K^{\prime} \in \mathcal{T}_{h}\right\} \\
\mathcal{E}_{\partial} & =\left\{e \in \mathcal{E}: e=\partial K \cap \partial \Omega, K \in \mathcal{T}_{h}\right\}
\end{aligned}
$$

For $e_{K} \in \mathcal{E}_{i n t}$, the average and jump of $w \in H^{1}\left(\Omega, \mathcal{T}_{h}\right)$ are defined by:

$$
\{w\}=\frac{1}{2}\left(\left.\left(\left.w\right|_{K}\right)\right|_{e_{K}}+\left.\left(\left.w\right|_{K^{\prime}}\right)\right|_{e_{K}}\right) \quad, \quad[w]=\left.\left(\left.w\right|_{K}\right)\right|_{e_{K}}-\left.\left(\left.w\right|_{K^{\prime}}\right)\right|_{e_{K}}, \quad \text { respectively. }
$$

The jump and average on $e_{K} \in \mathcal{E}_{\partial}$ are defined by

$$
\{w\}=\left.\left(\left.w\right|_{K}\right)\right|_{e_{K}}=[w] .
$$

Also, we assume

$$
\begin{equation*}
c_{1}(\kappa) h_{K} \leq\left|e_{K}\right| \leq c_{2}(\kappa) h_{K}, \quad c_{3}(\varrho) p_{K} \leq p_{e_{K}} \leq c_{4}(\varrho) p_{K} \tag{3.1}
\end{equation*}
$$

where $p_{e_{K}}$ is the degree of the polynomial used for the approximation of the unknown variables over the edge $e_{K}$. Note that the definition of the triangulation $\mathcal{T}_{h}$ admits atmost one hanging node along each side of $K$.
On the subdivision $\mathcal{T}_{h}$, we define the required broken Sobolev spaces for $s=1,2$ as

$$
H^{s}\left(\Omega, \mathcal{T}_{h}\right)=\left\{w \in L^{2}(\Omega):\left.w\right|_{K} \in H^{s}(K), K \in \mathcal{T}_{h}\right\}
$$

The associated broken norm and semi-norm are defined by:

$$
\|w\|_{H^{s}\left(\Omega, \mathcal{T}_{h}\right)}=\left(\sum_{K \in \mathcal{T}_{h}}\|w\|_{H^{s}(K)}^{2}\right)^{1 / 2} \text { and }|w|_{H^{s}\left(\Omega, \mathcal{T}_{h}\right)}=\left(\sum_{K \in \mathcal{T}_{h}}|w|_{H^{s}(K)}^{2}\right)^{1 / 2}, \text { respectively. }
$$

Also, let $\mathcal{U}=\left\{w \in H^{2}\left(\Omega, \mathcal{T}_{h}\right): w, \nabla w \cdot \mathbf{n}\right.$ are continuous along each $\left.e \in \mathcal{E}_{\text {int }}\right\}$.

## Finite element spaces

Let $Q_{p_{K}}(\hat{K})$ be the set of polynomials of degree less than or equal to $p_{K}$ in each coordinate on the reference element $\hat{K}$. Now consider a finite element subspace of $H^{1}\left(\Omega, \mathcal{T}_{h}\right)$,

$$
S^{p}=\left\{w \in L^{2}(\Omega):\left.w\right|_{K} \circ F_{K} \in Q_{p_{K}}(\hat{K}), K \in \mathcal{T}_{h}\right\}
$$

where $p=\left\{p_{K}: K \in \mathcal{T}_{h}\right\}$ and $F=\left\{F_{K}: K \in \mathcal{T}_{h}\right\}, F_{K}$ being the affine map from $\hat{K}$ to $K$.
We define the broken energy norm for $w \in H^{1}\left(\Omega, \mathcal{T}_{h}\right)$ as

$$
\left|\|w \mid\|=\left(\sum_{K \in \mathcal{T}_{h}}\|w\|_{H^{1}(K)}^{2}+\mathcal{J}^{\gamma}(w, w)\right)^{1 / 2}\right.
$$

where $\mathcal{J}^{\gamma}(w, v)=\sum_{e \in \mathcal{E}_{\text {int }}} \frac{\gamma}{|e|} \int_{e}[w][v] \mathrm{d} e, \gamma>0$ being the penalty parameter to be chosen later.

## Approximation properties of finite element spaces

Lemma 3.1 [24]. Let $\left.w\right|_{K} \in H^{s^{\prime}}(K), s^{\prime} \in \mathbb{Z}^{+}$. Then there exists a sequence $z_{p_{k}}^{h_{K}} \in Q_{p_{K}}(K)$, $p_{K}=1,2 \ldots$, such that for $0 \leq l \leq s^{\prime}$,

$$
\begin{aligned}
\left\|w-z_{p_{k}}^{h_{K}}\right\|_{H^{l}(K)} \leq C \frac{h_{K}^{s-l}}{p_{K}^{s^{\prime}-l}}\|w\|_{H^{s^{\prime}}(K)} & \forall K \in \mathcal{T}_{h} \\
\left\|w-z_{p_{k}}^{h_{K}}\right\|_{L^{2}(e)} \leq C \frac{h_{K}^{s-\frac{1}{2}}}{p_{K}^{s^{\prime}-\frac{1}{2}}}\|w\|_{H^{s^{\prime}}(K)} & \forall e \in \mathcal{E}_{i n t}
\end{aligned}
$$

and

$$
\left\|\nabla\left(w-z_{p_{k}}^{h_{K}}\right)\right\|_{L^{2}(e)} \leq C \frac{h_{K}^{s-\frac{3}{2}}}{p_{K}^{s^{\prime}-\frac{3}{2}}}\|w\|_{H^{s^{\prime}}(K)} \quad \forall e \in \mathcal{E}_{i n t},
$$

where $1 \leq s \leq \min \left(p_{K}+1, s^{\prime}\right), p_{K} \geq 1, s \in \mathbb{Z}^{+}$, and $C$ is a constant independent of $w, h_{K}$, and $p_{K}$, but dependent on $s^{\prime}$.
Remark 3.2. Given $w \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)$, define the interpolant $I_{h} w=\hat{w} \in S^{p}$ by

$$
\begin{equation*}
\left.I_{h} w\right|_{K}=\left.\hat{w}\right|_{K}=z_{p_{k}}^{h_{K}}\left(\left.w\right|_{K}\right) \quad \forall K \in \mathcal{T}_{h} . \tag{3.2}
\end{equation*}
$$

Lemma 3.3 [34], Lemma 2.1. Let $v_{h} \in Q_{p_{K}}(K)$. Then, there exists a constant $C>0$ such that

$$
\begin{equation*}
\left\|\nabla^{l} v_{h}\right\|_{e_{k}} \leq C p_{K} h_{K}^{-1 / 2}\left\|\nabla^{l} v_{h}\right\|_{K}, \quad l=0,1 \tag{3.3}
\end{equation*}
$$

The $h p$-DGFEM formulation corresponding to (2.13)-(2.16) can be stated as:
Find $\left(\theta_{h}(t), a_{h}(t)\right) \in S^{p} \times S^{p}$, a.e. in $\bar{I}$ such that

$$
\begin{align*}
\sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} a_{h}, w\right)_{K} & =\sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta_{h}, a_{h}\right), w\right)_{K} \quad \forall w \in S^{p},  \tag{3.4}\\
a_{h}(0) & =0,  \tag{3.5}\\
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta_{h}, v\right)_{K}+B\left(\theta_{h}, v\right) & =-\rho L \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} a_{h}, v\right)_{K}+\sum_{K \in \mathcal{T}_{h}}\left(\alpha u_{h}, v\right)_{K} \forall v \in S^{p},  \tag{3.6}\\
\theta_{h}(0)=\theta_{0}, & \tag{3.7}
\end{align*}
$$

where $B(\theta, v)=\mathcal{K} \sum_{K \in \mathcal{T}_{h}}(\nabla \theta, \nabla v)_{K}-\mathcal{K} \sum_{e \in \mathcal{E}_{\text {int }}} \int_{e}\{\nabla \theta \cdot \mathbf{n}\}[v] \mathrm{d} e-\mathcal{K} \sum_{e \in \mathcal{E}_{\text {int }}} \int_{e}\{\nabla v \cdot \mathbf{n}\}[\theta] \mathrm{d} e+\mathcal{J}^{\gamma}(\theta, v)$ and
 normal to the edge $e$.
Remark 3.4. Note that the bilinear form $B(\cdot, \cdot)$ is symmetric. Therefore, (3.4)-(3.7) corresponds to the symmetric interior penalty Galerkin formulation for the regularized laser surface hardening of steel problem.

Let $\left\{\phi_{1}, \phi_{2}, \ldots, \phi_{M}\right\}$ be the basis functions for $S^{p}$. Substituting $a_{h}=\sum_{i=1}^{M} a_{i}(t) \phi_{i}$ and $\theta_{h}=\sum_{i=1}^{M} \theta_{i}(t) \phi_{i}$ for $v=\phi_{j}, w=\phi_{j}, j=1,2, \ldots, M$ in (3.4)-(3.7), we obtain

$$
\begin{align*}
\mathbf{A} \partial_{t} \overline{\mathbf{a}} & =\overline{\mathbf{F}}(\bar{\theta}, \overline{\mathbf{a}}),  \tag{3.8}\\
\overline{\mathbf{a}}(0) & =0  \tag{3.9}\\
\rho c_{p} \mathbf{A} \partial_{t} \bar{\theta}+\mathbf{B} \bar{\theta} & =-\rho L \overline{\mathbf{F}}(\bar{\theta}, \overline{\mathbf{a}})+u_{h}(t) \bar{\alpha},  \tag{3.10}\\
\bar{\theta}(0) & =\theta_{0} \tag{3.11}
\end{align*}
$$

$$
\begin{gather*}
\text { where } \overline{\mathbf{a}}=\left(a_{i}(t)\right)_{1 \leq i \leq M}, \bar{\theta}=\left(\theta_{i}(t)\right)_{1 \leq i \leq M}, \mathbf{A}=\left(\sum_{K \in \mathcal{T}_{h}}\left(\phi_{i}, \phi_{j}\right)_{K}\right)_{1 \leq i, j \leq M} \\
\begin{array}{r}
\mathbf{B}=\left(B\left(\phi_{i}, \phi_{j}\right)\right)_{1 \leq i, j \leq M}, \overline{\mathbf{F}}(\bar{\theta}, \overline{\mathbf{a}})=\left(\sum_{K \in \mathcal{T}_{h}}\left(f\left(\sum_{i=1}^{M} \theta_{i}(t) \phi_{i}, \sum_{i=1}^{M} a_{i}(t) \phi_{i}\right), \phi_{j}\right)_{K}\right)_{1 \leq j \leq M} \\
\bar{\alpha}=\left(\left(\alpha(t), \phi_{j}\right)_{K}\right)_{1 \leq j \leq M}
\end{array}
\end{gather*}
$$

(3.8)-(3.11) is a system of ordinary nonlinear differential equations in independent variable $t$, with Lipschitz continuous right hand side in $(\bar{\theta}, \overline{\mathbf{a}})$ and hence admits a unique solution in a neighbourhood of $t=0$.

The $h p$-DGFEM scheme corresponding to the optimal control problem is

$$
\begin{equation*}
\min J\left(\theta_{h}, a_{h}, u_{h}\right) \quad \text { subject to the constraints (3.4)-(3.7) and } u_{h} \in U_{a d} . \tag{3.13}
\end{equation*}
$$

The adjoint system of (3.13) determined from the Karush-Kuhn-Tucker (KKT) system is defined by: Find $\left(z_{h}^{*}(t), \lambda_{h}^{*}(t)\right) \in S^{p} \times S^{p}$, a.e. in $\bar{I}$ such that

$$
\begin{align*}
-\sum_{K \in \mathcal{T}_{h}}\left(\chi, \partial_{t} \lambda_{h}^{*}\right)_{K}= & -\sum_{K \in \mathcal{T}_{h}}\left(\chi, f_{a}\left(\theta_{h}^{*}, a_{h}^{*}\right) g\left(z_{h}^{*}, \lambda_{h}^{*}\right)\right)_{K}  \tag{3.14}\\
\lambda_{h}^{*}(T)= & \beta_{1}\left(a_{h}^{*}(T)-a_{d}\right)  \tag{3.15}\\
-\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\phi, \partial_{t} z_{h}^{*}\right)_{K}+B\left(\phi, z_{h}^{*}\right)= & -\sum_{K \in \mathcal{T}_{h}}\left(\phi,\left(f_{\theta}\left(\theta_{h}^{*}, a_{h}^{*}\right) g\left(z_{h}^{*}, \lambda_{h}^{*}\right)\right)_{K}\right. \\
& \left.+\beta_{2}\left(\phi,\left[\theta_{h}^{*}-\theta_{m}\right]_{+}\right)_{K}\right)  \tag{3.16}\\
z_{h}^{*}(T)= & 0 \tag{3.17}
\end{align*}
$$

for all $(\chi, \phi) \in S^{p} \times S^{p}$. Moreover, $z^{*}$ satisfies the following variational inequality

$$
\left(\beta_{3} u_{h}^{*}+\int_{\Omega} \alpha z_{h}^{*} \mathrm{~d} x, p-u_{h}^{*}\right)_{L^{2}(I)} \geq 0 \quad \forall p \in U_{a d}
$$

## Continuous time a priori error estimates

For deriving a priori error estimates for the $h p$-DGFEM formulation of the laser surface hardening of steel problem, we would like to define the broken projector $\Pi: H^{2}\left(\Omega, \mathcal{T}_{h}\right) \longrightarrow S^{p}$ satisfying;

$$
\begin{equation*}
B(\Pi v-v, w)+\nu(\Pi v-v, w)=0 \quad \forall w \in S^{p} \tag{3.18}
\end{equation*}
$$

where $\nu>0$ is a constant. Now we state the following lemmas, the proofs of which are in the similar lines as in [24].
Lemma 3.5. There exists a constant $C>0$, independent of $h$ such that

$$
\left|B_{\nu}(v, w)\right| \leq C\left|\left\|v \left|\left\|\left|\|w \mid\| \quad \forall v, w \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)\right.\right.\right.\right.\right.
$$

where $B_{\nu}(v, w)=B(v, w)+\nu(v, w) \quad \forall v, w \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)$.
Lemma 3.6. For a sufficiently large choice of the penalty parameter $\gamma$, there exists $C>0$ such that

$$
B_{\nu}(w, w) \geq C\left|\|w \mid\|^{2} \quad \forall w \in S^{p}\right.
$$

Using Lemmas 3.5 and $3.6, \Pi v$ is well defined for $v \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)$. Now, we state an estimate for $\|v-\Pi v\|$.
Lemma 3.7 [24]. Let $\Pi v$ be the projection of $v \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)$ onto $S^{p}$ defined by (3.18), then the following error estimate holds true:

$$
\|v-\Pi v\|^{2} \leq C\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \sum_{K \in \mathcal{T}_{h}} \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}\|v\|_{H^{s^{\prime}}(K)}^{2}
$$

where $s=\min \left(p_{K}+1, s^{\prime}\right), s^{\prime} \geq 2, p_{K} \geq 2$.
Let $\theta-\theta_{h}=\eta^{\theta}+\zeta^{\theta}$ and $a-a_{h}=\eta^{a}+\zeta^{a}$, where $\eta^{\theta}=\Pi \theta-\theta_{h}, \zeta^{\theta}=\theta-\Pi \theta, \eta^{a}=\hat{a}-a_{h}, \zeta^{a}=a-\hat{a}$ and $\hat{a}$ is the interpolant of $a$ as defined in (3.2). Using the triangle inequality, we have

$$
\left\|\theta-\theta_{h}\right\| \leq\left\|\eta^{\theta}\right\|+\left\|\zeta^{\theta}\right\|, \quad\left\|a-a_{h}\right\| \leq\left\|\eta^{a}\right\|+\left\|\zeta^{a}\right\|
$$

In the next theorem, we develop an a priori error estimate for $\left\|\theta(t)-\theta_{h}(t)\right\|$ and $\left\|a(t)-a_{h}(t)\right\|, \quad t \in \bar{I}$, for a fixed $u \in U_{a d}$.
Theorem 3.8. Let $(\theta(t), a(t))$ and $\left(\theta_{h}(t), a_{h}(t)\right)$ be the solutions of (2.8)-(2.12) and (3.4)-(3.7), respectively, for a fixed $u \in U_{a d}$. Then,

$$
\begin{aligned}
\| \theta(t)- & \theta_{h}(t)\left\|^{2}+\right\| a(t)-a_{h}(t) \|^{2} \leq C\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \sum_{K \in \mathcal{T}_{h}} \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}\left(\left\|\theta_{0}\right\|_{H^{s^{\prime}(K)}}^{2}+\|\theta\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\left\|\partial_{t} \theta\right\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right. \\
& \left.+\|a\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t} a\right\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|\theta\|_{L^{\infty}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|a\|_{L^{\infty}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right), t \in \bar{I},
\end{aligned}
$$

where $C>0$ is independent of $p_{K}, h_{K}$ and $(\theta, a)$, also $s=\min \left(p_{K}+1, s^{\prime}\right)$ and $s^{\prime}, p_{K} \geq 2$.
Proof. A solution $(\theta, a)$ of (2.8)-(2.12), under the regularity assumption that $\theta(t) \in \mathcal{U}, t \in \bar{I}$, satisfies the broken weak formulation

$$
\begin{equation*}
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta, v\right)_{K}+B_{\nu}(\theta, v)=-\rho L \sum_{K \in \mathcal{T}_{h}}(f(\theta, a), v)_{K}+\sum_{K \in \mathcal{T}_{h}}(\alpha u, v)_{K}+\nu \sum_{K \in \mathcal{T}_{h}}(\theta, v)_{K} . \tag{3.19}
\end{equation*}
$$

Subtracting (3.6) from (3.19) and using $B_{\nu}\left(\zeta^{\theta}, v\right)=0 \quad \forall v \in S^{p}$ (see (3.18)), we obtain
$\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \eta^{\theta}, v\right)_{K}+B_{\nu}\left(\eta^{\theta}, v\right)=-\rho L \sum_{K \in \mathcal{T}_{h}}\left(f(\theta, a)-f\left(\theta_{h}, a_{h}\right), v\right)_{K}-\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\left(\partial_{t} \zeta^{\theta}, v\right)_{K}+\nu\left(\eta^{\theta}+\zeta^{\theta}, v\right)_{K}\right)$.

Choose $v=\eta^{\theta}$, use Lemma 3.6 and integrate from 0 to $t$ to obtain

$$
\begin{align*}
\frac{1}{2}\left\|\eta^{\theta}(t)\right\|^{2}+\int_{0}^{t}\left|\left\|\eta^{\theta} \mid\right\|^{2} \mathrm{~d} s \leq\right. & C\left(\left\|\eta^{\theta}(0)\right\|^{2}+\sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(f(\theta, a)-f\left(\theta_{h}, a_{h}\right), \eta^{\theta}\right)_{K}\right| \mathrm{d} s\right. \\
& \left.+\sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(\partial_{t} \zeta^{\theta}, \eta^{\theta}\right)_{K}\right| \mathrm{d} s+\sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(\zeta^{\theta}, \eta^{\theta}\right)_{K}\right| \mathrm{d} s+\sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left\|\eta^{\theta}\right\|_{K}^{2} \mathrm{~d} s\right) \\
= & C\left\|\eta^{\theta}(0)\right\|^{2}+I_{1}+I_{2}+I_{3}+\int_{0}^{t}\left\|\eta^{\theta}\right\|^{2} \mathrm{~d} s, \text { say. } \tag{3.20}
\end{align*}
$$

Now we estimate $I_{1}, I_{2}$ and $I_{3}$ in the right hand side of (3.20). Using Cauchy-Schwarz inequality, Young's inequality and Remark 2.2, we obtain

$$
\begin{equation*}
I_{1}=\sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(f(\theta, a)-f\left(\theta_{h}, a_{h}\right), \eta^{\theta}\right)_{K}\right| \mathrm{d} s \leq C \int_{0}^{t}\left(\left\|\eta^{\theta}\right\|^{2}+\left\|\zeta^{\theta}\right\|^{2}+\left\|\eta^{a}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}\right) \mathrm{d} s \tag{3.21}
\end{equation*}
$$

Using Cauchy-Schwarz inequality and Young's inequality, we have

$$
\begin{align*}
& I_{2} \leq \sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(\partial_{t} \zeta^{\theta}, \eta^{\theta}\right)_{K}\right| \mathrm{d} s \leq C \int_{0}^{t}\left(\left\|\partial_{t} \zeta^{\theta}\right\|^{2}+\left\|\eta^{\theta}\right\|^{2}\right) \mathrm{d} s  \tag{3.22}\\
& I_{3} \leq \sum_{K \in \mathcal{T}_{h}} \int_{0}^{t}\left|\left(\zeta^{\theta}, \eta^{\theta}\right)_{K}\right| \mathrm{d} s \leq C \int_{0}^{t}\left(\left\|\zeta^{\theta}\right\|^{2}+\left\|\eta^{\theta}\right\|^{2}\right) \mathrm{d} s \tag{3.23}
\end{align*}
$$

Using (3.21)-(3.23) in (3.20), we obtain

$$
\begin{align*}
\frac{1}{2}\left\|\eta^{\theta}(t)\right\|^{2}+\int_{0}^{t}\left|\left\|\eta^{\theta} \mid\right\|^{2} \mathrm{~d} s \leq\right. & C\left(\left\|\eta^{\theta}(0)\right\|^{2}+\int_{0}^{t}\left(\left\|\zeta^{\theta}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}+\left\|\partial_{t} \zeta^{\theta}\right\|^{2}\right) \mathrm{d} s\right. \\
& \left.+\int_{0}^{t}\left(\left\|\eta^{\theta}\right\|^{2}+\left\|\eta^{a}\right\|^{2}\right) \mathrm{d} s\right) \\
\text { That is, } \quad\left\|\eta^{\theta}(t)\right\|^{2} \leq & C\left(\left\|\eta^{\theta}(0)\right\|^{2}+\int_{0}^{t}\left(\left\|\zeta^{\theta}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}+\left\|\partial_{t} \zeta^{\theta}\right\|^{2}\right) \mathrm{d} s\right. \\
& \left.+\int_{0}^{t}\left(\left\|\eta^{\theta}\right\|^{2}+\left\|\eta^{a}\right\|^{2}\right) \mathrm{d} s\right) \tag{3.24}
\end{align*}
$$

Now subtracting (3.4) from (2.13), we obtain

$$
\sum_{K \in \mathcal{T}_{h}}\left(\partial_{t}\left(a-a_{h}\right), w\right)_{K}=\sum_{K \in \mathcal{T}_{h}}\left(f(\theta, a)-f\left(\theta_{h}, a_{h}\right), w\right)_{K} \quad \forall w \in S^{p}
$$

Using $a-a_{h}=\eta^{a}+\zeta^{a}$ and substituting $w=\eta^{a}$, we obtain

$$
\sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \eta^{a}, \eta^{a}\right)_{K}=\sum_{K \in \mathcal{T}_{h}}\left(\left(f(\theta, a)-f\left(\theta_{h}, a_{h}\right), \eta^{a}\right)_{K}-\left(\partial_{t} \zeta^{a}, \eta^{a}\right)_{K}\right)
$$

Now integrating from 0 to $t$, using Cauchy-Schwarz inequality, Young's inequality and Remark 2.2, we obtain

$$
\begin{equation*}
\left\|\eta^{a}(t)\right\|^{2} \leq C \int_{0}^{t}\left(\left\|\eta^{\theta}\right\|^{2}+\left\|\zeta^{\theta}\right\|^{2}+\left\|\eta^{a}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}+\left\|\partial_{t} \zeta^{a}\right\|^{2}\right) \mathrm{d} s \tag{3.25}
\end{equation*}
$$

Adding (3.24) and (3.25), we obtain

$$
\begin{aligned}
\left\|\eta^{\theta}(t)\right\|^{2}+\left\|\eta^{a}(t)\right\|^{2} \leq & C\left(\left\|\eta^{\theta}(0)\right\|^{2}+\int_{0}^{T}\left(\left\|\zeta^{\theta}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}+\left\|\partial_{t} \zeta^{\theta}\right\|^{2}+\left\|\partial_{t} \zeta^{a}\right\|^{2}\right) \mathrm{d} s\right. \\
& \left.+\int_{0}^{t}\left(\left\|\eta^{\theta}\right\|^{2}+\left\|\eta^{a}\right\|^{2}\right) \mathrm{d} s\right)
\end{aligned}
$$

Use Gronwall's lemma to obtain

$$
\left\|\eta^{\theta}(t)\right\|^{2}+\left\|\eta^{a}(t)\right\|^{2} \leq C\left(\left\|\eta^{\theta}(0)\right\|^{2}+\int_{0}^{T}\left(\left\|\zeta^{\theta}\right\|^{2}+\left\|\zeta^{a}\right\|^{2}+\left\|\partial_{t} \zeta^{\theta}\right\|^{2}+\left\|\partial_{t} \zeta^{a}\right\|^{2}\right) \mathrm{d} s\right)
$$

From Lemmas 3.1 and 3.7, we have

$$
\begin{aligned}
\left\|\eta^{\theta}(t)\right\|^{2}+\left\|\eta^{a}(t)\right\|^{2} \leq & C\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \sum_{K \in \mathcal{T}_{h}} \frac{h_{K}^{2 s-2}}{p_{K}^{2 s_{k}^{\prime}-3}}\left(\left\|\theta_{0}\right\|_{H^{s^{\prime}}(K)}^{2}+\|\theta\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right. \\
& \left.+\left\|\partial_{t} \theta\right\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\|a\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\left\|\partial_{t} a\right\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right) .
\end{aligned}
$$

Using triangle inequality we obtain the required estimate. This completes the proof.
Next we state the error estimates for the system (2.20)-(2.23), which is the adjoint system corresponding to $(2.13)-(2.16)$. The proof has been omitted as it is on the similar lines as Theorem 3.8. We denote $\left(z_{h}^{*}, \lambda_{h}^{*}\right)$ as $\left(z_{h}, \lambda_{h}\right)$ for notational convenience.

Theorem 3.9. Let $(z(t), \lambda(t))$ and $\left(z_{h}(t), \lambda_{h}(t)\right)$ be the solutions for (2.20)-(2.23) and (3.14)-(3.17), respectively. Then, there exists a positive constant $C$ such that

$$
\begin{aligned}
\| z(t)- & z_{h}(t)\left\|^{2}+\right\| \lambda(t)-\lambda_{h}(t) \|^{2} \leq C\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \sum_{K \in \mathcal{T}_{h}} \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}\left(\left\|a_{d}\right\|_{H^{2}(K)}^{2}+\left\|\theta_{0}\right\|_{H^{s^{\prime}(K)}}^{2}+\|\theta\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right. \\
& +\left\|\partial_{t} \theta\right\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|a\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t} a\right\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|z\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2} \\
& +\left\|\partial_{t} z\right\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|\lambda\|_{L^{2}\left(I, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t} \lambda\right\|_{L^{2}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\|\theta\|_{L^{\infty}\left(I, H^{s^{\prime}}(K)\right)}^{2} \\
& \left.+\|z\|_{L^{\infty}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\|a\|_{L^{\infty}\left(I, H^{s^{\prime}}(K)\right)}^{2}+\|\lambda\|_{L^{\infty}\left(I, H^{s^{\prime}}(K)\right)}^{2}\right), \quad t \in \bar{I},
\end{aligned}
$$

where $C$ is independent of $p_{K}, h_{K},(\theta, a)$ and $(z, \lambda)$, also $s=\min \left(p_{K}+1, s^{\prime}\right)$ and $s^{\prime}, p_{K} \geq 2$.
Remark 3.10. Note that Theorems 3.8 and 3.9 hold true under the following minimum extra-regularity assumptions on the data, the continuous and the adjoint solutions:

$$
\theta_{0}, a_{d} \in H^{2}\left(\Omega, \mathcal{T}_{h}\right), \theta, a, z, \lambda \in H^{1}\left(I, H^{2}\left(\Omega, \mathcal{T}_{h}\right)\right) \cap L^{\infty}\left(I, H^{2}\left(\Omega, \mathcal{T}_{h}\right)\right), \theta(t), z(t) \in \mathcal{U}
$$

where $H^{1}\left(I, H^{2}\left(\Omega, \mathcal{I}_{h}\right)\right)=\left\{v: v \in H^{2}\left(\Omega, \mathcal{T}_{h}\right), v_{t} \in H^{2}\left(\Omega, \mathcal{T}_{h}\right)\right\}, \mathcal{U}=\left\{w \in H^{2}\left(\Omega, \mathcal{T}_{h}\right): w, \nabla w \cdot \mathbf{n}\right.$ are continuous along each $\left.e \in \mathcal{E}_{\text {int }}\right\}$ and $L^{\infty}\left(I, H^{2}\left(\Omega, \mathcal{T}_{h}\right)\right)=\operatorname{ess}_{\sup _{t \in I}}\|v(t)\|_{H^{2}\left(\Omega, \mathcal{T}_{h}\right)}<\infty$.

## 4. $h p$-DGFEM-DG space-time-control discretization

In this section, first of all, a temporal discretization is done using a DGFEM with piecewise constant approximation and a priori error estimates are proved in Theorems 4.1 and 4.2. The control is then discretized using
piecewise constants in each discrete interval $I_{n}, n=1,2, \ldots, N$ and a convergence result is established. In order to discretize (3.4)-(3.7) in time, we consider the following partition of $I$ :

$$
0=t_{0}<t_{1}<\ldots<t_{N}=T
$$

Set $I_{1}=\left[t_{0}, t_{1}\right], I_{n}=\left(t_{n-1}, t_{n}\right], k_{n}=t_{n}-t_{n-1}$, for $n=2, \ldots, N$ and $k=\max _{1 \leq n \leq N} k_{n}$. We define the space

$$
\begin{equation*}
X_{h k}^{q}=\left\{\phi: I \rightarrow S^{p} ;\left.\quad \phi\right|_{I_{n}}=\sum_{j=0}^{q} \psi_{j} t^{j}, \psi_{j} \in S^{p}\right\}, q \in \mathbb{N} \tag{4.1}
\end{equation*}
$$

The space time $h p$-DGFEM scheme corresponding to (3.4)-(3.7) reads as;
Find $\left(\theta_{h k}, a_{h k}\right) \in X_{h k}^{q} \times X_{h k}^{q}$ such that

$$
\begin{align*}
\sum_{n=1}^{N}\left(\left(\partial_{t} a_{h k}, w\right)_{I_{n}, \Omega}+\left(<a_{h k}>_{n-1}, w_{n-1}^{+}\right)\right) & =\sum_{n=1}^{N}\left(f\left(\theta_{h k}, a_{h k}\right), w\right)_{I_{n}, \Omega}  \tag{4.2}\\
a_{h k}(0) & =0  \tag{4.3}\\
\sum_{n=1}^{N}\left(\rho c_{p}\left(\partial_{t} \theta_{h k}, v\right)_{I_{n}, \Omega}+\int_{I_{n}} B\left(\theta_{h k}, v\right) \mathrm{d} t+\rho c_{p}\left(\left\langle\theta_{h k}\right\rangle_{n-1}, v_{n-1}^{+}\right)\right)= & \sum_{n=1}^{N}\left(-\rho L\left(f\left(\theta_{h k}, a_{h k}\right), v\right)_{I_{n}, \Omega}\right. \\
& \left.+\left(\alpha u_{h k}, v\right)_{I_{n}, \Omega}\right)  \tag{4.4}\\
\theta_{h k}(0)= & \theta_{0} . \tag{4.5}
\end{align*}
$$

for all $(w, v) \in X_{h k}^{q} \times X_{h k}^{q}$ and the jump $\langle\cdot\rangle$ is defined by

$$
\langle v\rangle_{n}=v\left(t_{n}^{+}\right)-v\left(t_{n}\right)
$$

with $v_{n-1}^{+}$denoting $v\left(t_{n-1}^{+}\right)$.
For $q=0$, the space-time $h p$-DGFEM scheme corresponding to (3.4)-(3.7) reads as; Find $\left(\theta_{h k}^{n}, a_{h k}^{n}\right) \in S^{p} \times S^{p}, n=1,2, \ldots, N$ such that

$$
\begin{align*}
\sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} a_{h k}^{n}, w\right)_{K}= & \sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), w\right)_{K},  \tag{4.6}\\
a_{h k}(0)= & 0,  \tag{4.7}\\
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} \theta_{h k}^{n}, v\right)_{K}+B\left(\theta_{h k}^{n}, v\right)= & -\rho L \sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), v\right)_{K} \\
& +\sum_{K \in \mathcal{T}_{h}}\left(\frac{1}{k_{n}} \int_{I_{n}} \alpha u_{h k}(t) \mathrm{d} s, v\right)_{K}  \tag{4.8}\\
\theta_{h k}(0)= & \theta_{0} \tag{4.9}
\end{align*}
$$

$\forall(w, v) \in S^{p} \times S^{p}$, where $\bar{\partial} \phi^{n}=\frac{\phi^{n}-\phi^{n-1}}{k_{n}} \forall \phi \in S^{p}$. Expanding $a_{h k}^{n}=\sum_{i=1}^{M} a_{i}^{n} \phi_{i}$ and $\theta_{h k}^{n}=\sum_{i=1}^{M} \theta_{i}^{n} \phi_{i}$, where $\left\{\phi_{i}\right\}_{i=1}^{N}$ is the basis for $S^{p}$, we obtain the system

$$
\begin{align*}
\mathbf{A} \overline{\mathbf{a}}^{\mathbf{n}} & =k_{n} \overline{\mathbf{F}}\left(\bar{\theta}^{\mathbf{n}}, \overline{\mathbf{a}}^{\mathbf{n}}\right)+\mathbf{A} \overline{\mathbf{a}}^{\mathbf{n}-\mathbf{1}},  \tag{4.10}\\
\overline{\mathbf{a}}^{0} & =0,  \tag{4.11}\\
\left(\rho c_{p} \mathbf{A}+k_{n} \mathbf{B}\right) \bar{\theta}^{\mathbf{n}} & =-k_{n} \rho L \overline{\mathbf{F}}\left(\bar{\theta}^{\mathbf{n}}, \overline{\mathbf{a}}^{\mathbf{n}}\right)+k_{n} u_{h k}\left(t_{n}\right) \bar{\alpha}^{\mathbf{n}}+\mathbf{A} \bar{\theta}^{\mathbf{n}-\mathbf{1}},  \tag{4.12}\\
\bar{\theta}(0) & =\theta_{0}, \tag{4.13}
\end{align*}
$$

where $\mathbf{A}, \overline{\mathbf{a}}, \overline{\mathbf{F}}, \mathbf{B}, \bar{\theta}$ and $\bar{\alpha}$ are defined in (3.12). (4.10)-(4.13) form a system of non-linear equations with a Lipschitz continuous right hand side and hence admits a unique local solution in the neighbourhood of $t=0$. The time discrete $h p$-DGFEM scheme for the optimal control problem is

$$
\begin{equation*}
\min J\left(\theta_{h k}, a_{h k}, u_{h k}\right) \quad \text { subject to the constraints (4.6)-(4.9) and } u_{h k} \in U_{a d} \tag{4.14}
\end{equation*}
$$

The adjoint system of (4.14) determined by the KKT system is defined by: find $\left(z_{h k}^{n, *}, \lambda_{h k}^{n, *}\right) \in S^{p} \times S^{p}$ such that

$$
\begin{align*}
-\sum_{K \in \mathcal{T}_{h}}\left(\chi, \tilde{\partial} \lambda_{h k}^{n-1, *}\right)_{K} & =-\sum_{K \in \mathcal{T}_{h}}\left(\chi, f_{a}\left(\theta_{h k}^{n-1, *}, a_{h k}^{n-1, *}\right) g\left(z_{h k}^{n-1, *}, \lambda_{h k}^{n-1, *}\right)\right)_{K},  \tag{4.15}\\
\lambda_{h k}^{*}(T)= & \beta_{1}\left(a_{h k}^{*}(T)-a_{d}\right),  \tag{4.16}\\
-\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\phi, \tilde{\partial} z_{h k}^{n-1, *}\right)_{K}+B\left(\phi, z_{h k}^{n-1, *}\right)= & -\sum_{K \in \mathcal{T}_{h}}\left(\left(\phi, f_{\theta}\left(\theta_{h k}^{n-1, *}, a_{h k}^{n-1, *}\right) g\left(z_{h k}^{n-1, *}, \lambda_{h k}^{n-1, *}\right)\right)_{K}\right. \\
& \left.+\beta_{2}\left(\phi,\left[\theta_{h k}^{n-1, *}-\theta_{m}\right]_{+}\right)_{K}\right)  \tag{4.17}\\
z_{h k}^{*}(T)= & 0 \tag{4.18}
\end{align*}
$$

for all $(\chi, \phi) \in S^{p} \times S^{p}$, where $\tilde{\partial} \phi^{n-1}=\frac{\phi^{n}-\phi^{n-1}}{k_{n}}$.

## Discrete time a priori error estimates

Before estimating the a priori error estimates for space-time discretization, we define the interpolant $\pi_{k}$ : $C\left(\bar{I}, S^{p}\right) \longrightarrow X_{h k}^{1},\left.\pi_{k} v\right|_{I_{n}} \in \mathcal{P}_{0}\left(I_{n}, S^{p}\right)$, (see [35]) as:

$$
\begin{equation*}
\pi_{k} v(t)=v\left(t_{n}\right) \quad \forall t \in \bar{I}_{n}, n=1,2, \ldots, N, \tag{4.19}
\end{equation*}
$$

where $\mathcal{P}_{0}\left(I_{n}, S^{p}\right)$ is the space of all functions in $S^{p}$ which are constants with respect to the variable $t$ in each interval $I_{n}$. Note that

$$
\begin{equation*}
\left\|v-\pi_{k} v\right\|_{I_{n}, K} \leq C k_{n}\left\|\partial_{t} v\right\|_{I_{n}, K} \tag{4.20}
\end{equation*}
$$

Theorem 4.1. Let $(\theta(t), a(t))$ and $\left(\theta_{h k}^{n}, a_{h k}^{n}\right), n=1,2, \ldots, N$ be the solutions of (2.13)-(2.16) and (4.6)-(4.9), respectively, for a fixed $u \in U_{a d}$. Then,

$$
\begin{aligned}
\| \theta\left(t_{n}\right)- & \theta_{h k}^{n}\left\|^{2}+\right\| a\left(t_{n}\right)-a_{h k}^{n} \|^{2} \leq C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}+k_{n}^{2}\right)\left(\left\|\theta_{0}\right\|_{H^{s^{\prime}(K)}}^{2}+\|\theta\|_{L^{\infty}\left(I_{n} ; H^{s^{\prime}}(K)\right)}^{2}\right. \\
& +\left\|\partial_{t} \theta\right\|_{L^{\infty}\left(I_{n} ; H^{\left.s^{\prime}(K)\right)}\right.}^{2} \\
& +\|a\|_{L^{\infty}\left(I_{n} ; H^{s^{\prime}}(K)\right)}^{2}+\left\|\partial_{t} a\right\|_{L^{\infty}\left(I_{n} ; H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I ; L^{2}(K)\right)}^{2}+\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2} \\
& \left.+\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}\right), t_{n} \in \bar{I}_{n}
\end{aligned}
$$

where $C>0$ is independent of $p_{K}, h_{K}$ and $(\theta, a)$, also $s=\min \left(p_{K}+1, s^{\prime}\right)$ and $s^{\prime}, p_{K} \geq 2$.
Proof. Subtracting (4.8) from (3.19) at $t=t_{n}$, we obtain

$$
\begin{aligned}
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta\left(t_{n}\right)-\bar{\partial} \theta_{h k}^{n}, v\right)_{K}+B_{\nu}\left(\theta\left(t_{n}\right)-\theta_{h k}^{n}, v\right)= & -\rho L \sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), v\right)_{K} \\
& +\sum_{K \in \mathcal{T}_{h}}\left(\alpha\left(x, t_{n}\right) u\left(t_{n}\right)-\frac{1}{k_{n}} \int_{I_{n}} \alpha u \mathrm{~d} s, v\right)_{K} \\
& +\nu \sum_{K \in \mathcal{T}_{h}}\left(\theta\left(t_{n}\right)-\theta_{h k}^{n}, v\right)_{K}
\end{aligned}
$$

where $v \in S^{p}$. Using (4.19), we find that

$$
\begin{aligned}
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta\left(t_{n}\right)-\bar{\partial} \theta_{h k}^{n}, v\right)_{K}+B_{\nu}\left(\theta\left(t_{n}\right)-\theta_{h k}^{n}, v\right) \leq & -\sum_{K \in \mathcal{T}_{h}} \rho L\left(f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), v\right)_{K} \\
& +\sum_{K \in \mathcal{T}_{h}} \max _{K \times I_{n}} \frac{1}{k_{n}}|\alpha|\left(\int_{I_{n}}\left(\pi_{k} u\left(t_{n}\right)-u\right) \mathrm{d} s, v\right)_{K} \\
& \left.+\nu \sum_{K \in \mathcal{T}_{h}}\left(\theta\left(t_{n}\right)-\theta_{h k}^{n}\right), v\right)_{K}
\end{aligned}
$$

Writing $\theta\left(t_{n}\right)-\theta_{h k}^{n}=\left(\theta\left(t_{n}\right)-\Pi \theta\left(t_{n}\right)\right)+\left(\Pi \theta\left(t_{n}\right)-\theta_{h k}^{n}\right)=\zeta^{\theta, n}+\eta^{\theta, n}$ and $a\left(t_{n}\right)-a_{h k}^{n}=\left(a\left(t_{n}\right)-\hat{a}\left(t_{n}\right)\right)+\left(\hat{a}\left(t_{n}\right)-\right.$ $\left.a_{h k}^{n}\right)=\zeta^{a, n}+\eta^{a, n}$ and using $B_{\nu}\left(\zeta^{\theta, n}, v\right)=0$, we obtain

$$
\begin{align*}
\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} \eta^{\theta, n}, v\right)_{K}+B_{\nu}\left(\eta^{\theta, n}, v\right) \leq & -\rho L \sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), v\right)_{K} \\
& +\frac{\max _{\Omega \times I_{n}}|\alpha|}{k_{n}} \sum_{K \in \mathcal{T}_{h}}\left(\int_{I_{n}}\left(\pi_{k} u\left(t_{n}\right)-u\right) \mathrm{d} s, v\right)_{K}-\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta\left(t_{n}\right)-\bar{\partial} \theta\left(t_{n}\right), v\right)_{K} \\
& -\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} \zeta^{\theta, n}, v\right)_{K}+\nu \sum_{K \in \mathcal{T}_{h}}\left(\eta^{\theta, n}+\zeta^{\theta, n}, v\right)_{K} \tag{4.21}
\end{align*}
$$

Now,

$$
\begin{aligned}
\frac{1}{2 k_{n}}\left(\left\|\eta^{\theta, n}\right\|^{2}-\left\|\eta^{\theta, n-1}\right\|^{2}\right) & =\frac{1}{2 k_{n}}\left(\left(\eta^{\theta, n}, \eta^{\theta, n}\right)-\left(\eta^{\theta, n-1}, \eta^{\theta, n-1}\right)\right) \\
& =\frac{1}{2 k_{n}}\left(\left(\eta^{\theta, n}-\eta^{\theta, n-1}, \eta^{\theta, n}\right)-\left(\eta^{\theta, n-1}, \eta^{\theta, n-1}-\eta^{\theta, n}\right)\right)
\end{aligned}
$$

Adding and subtracting $\eta^{\theta, n}$ in the first argument of the second term in the right hand side of the above expression, we obtain

$$
\begin{align*}
\frac{1}{2 k_{n}}\left(\left\|\eta^{\theta, n}\right\|^{2}-\left\|\eta^{\theta, n-1}\right\|^{2}\right) & =\left(\bar{\partial} \eta^{\theta, n}, \eta^{\theta, n}\right)-\frac{1}{2 k_{n}}\left(\eta^{\theta, n}-\eta^{\theta, n-1}, \eta^{\theta, n}-\eta^{\theta, n-1}\right) \\
& \leq\left(\bar{\partial} \eta^{\theta, n}, \eta^{\theta, n}\right) \tag{4.22}
\end{align*}
$$

Substituting $v=\eta^{\theta, n}$ in (4.21), using the coercivity of $B_{\nu}(\cdot, \cdot)$ and using (4.22) in (4.21), we obtain

$$
\begin{align*}
\left\|\eta^{\theta, n}\right\|^{2}-\left\|\eta^{\theta, n-1}\right\|^{2} \leq & C k_{n}\left(\sum_{K \in \mathcal{T}_{h}}\left|\left(f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), \eta^{\theta, n}\right)_{K}\right|\right. \\
& +\frac{1}{k_{n}} \sum_{K \in \mathcal{T}_{h}}\left|\left(\int_{I_{n}}\left(\pi_{k} u\left(t_{n}\right)-u\right) \mathrm{d} s, \eta^{\theta, n}\right)_{K}\right|+\sum_{K \in \mathcal{T}_{h}}\left|\left(\partial_{t} \theta\left(t_{n}\right)-\bar{\partial} \theta\left(t_{n}\right), \eta^{\theta, n}\right)_{K}\right| \\
& \left.+\sum_{K \in \mathcal{T}_{h}}\left|\left(\bar{\partial} \zeta^{\theta, n}, \eta^{\theta, n}\right)_{K}\right|+\sum_{K \in \mathcal{T}_{h}}\left|\left(\zeta^{\theta, n}, \eta^{\theta, n}\right)_{K}\right|+\sum_{K \in \mathcal{T}_{h}}\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \\
= & C k_{n}\left(J_{1}+\frac{1}{k_{n}} J_{2}+J_{3}+J_{4}+J_{5}+\left\|\eta^{\theta, n}\right\|^{2}\right), \quad \text { say } \tag{4.23}
\end{align*}
$$

From Cauchy-Schwarz inequality, Young's inequality and Remark 2.2, we have

$$
\begin{align*}
J_{1} & \leq \sum_{K \in \mathcal{T}_{h}}\left\|f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right)\right\|_{K}\left\|\eta^{\theta, n}\right\|_{K} \\
& \leq C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\zeta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{a, n}\right\|_{K}^{2}+\left\|\eta^{a, n}\right\|_{K}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) . \tag{4.24}
\end{align*}
$$

For $J_{2}$, use Cauchy-Schwarz inequality, Young's inequality and (4.20) to obtain

$$
\begin{align*}
J_{2} \leq \sum_{K \in \mathcal{T}_{h}}\left\|\pi_{k} u-u\right\|_{L^{2}\left(I_{n}\right)}\left\|\eta^{\theta, n}\right\|_{K} & \leq C\left(\left\|\pi_{k} u-u\right\|_{L^{2}\left(I_{n}\right)}^{2}+\sum_{K \in \mathcal{T}_{h}}\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \\
& \leq C \sum_{K \in \mathcal{T}_{h}}\left(k_{n}^{2}\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \tag{4.25}
\end{align*}
$$

Now consider $J_{3}$. Using Cauchy-Schwarz inequality and Young's inequality, we obtain

$$
\begin{equation*}
J_{3} \leq C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\partial_{t} \theta\left(t_{n}\right)-\bar{\partial} \theta\left(t_{n}\right)\right\|_{K}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \tag{4.26}
\end{equation*}
$$

For the first term on the right hand side of (4.26), we have

$$
\begin{aligned}
\left\|\bar{\partial} \theta\left(t_{n}\right)-\partial_{t} \theta\left(t_{n}\right)\right\|_{K} & =\left\|k_{n}^{-1} \int_{t_{n-1}}^{t_{n}}\left(t-t_{n-1}\right) \partial_{t t} \theta \mathrm{~d} t\right\|_{K} \leq k_{n}^{-1} \int_{t_{n-1}}^{t_{n}}\left(t-t_{n-1}\right)\left\|\partial_{t t} \theta\right\|_{K} \mathrm{~d} t \\
& \leq\left. C k_{n}^{-1} \frac{\left(t-t_{n-1}\right)^{2}}{2}\right|_{t_{n-1}} ^{t_{n}}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)} \leq C k_{n}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}
\end{aligned}
$$

Therefore, we have

$$
\begin{equation*}
J_{3} \leq C \sum_{K \in \mathcal{T}_{h}}\left(k_{n}^{2}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \tag{4.27}
\end{equation*}
$$

Also for $J_{4}$, using Cauchy-Schwarz inequality and Young's inequality, we have

$$
\begin{gathered}
J_{4} \leq C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\bar{\partial} \zeta^{\theta, n}\right\|_{K}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \\
\text { Also, }\left\|\bar{\partial} \zeta^{\theta, n}\right\|_{K}
\end{gathered}=\left\|k_{n}^{-1} \int_{t_{n-1}}^{t_{n}} \partial_{t} \zeta^{\theta} \mathrm{d} t\right\|_{K} \leq C\left\|\partial_{t} \zeta^{\theta, n}\right\|_{L^{\infty}\left(I_{n} ; L^{2}(K)\right)} .
$$

From Cauchy-Schwarz inequality and Young's inequality, we have

$$
\begin{equation*}
J_{5}=\sum_{K \in \mathcal{T}_{h}}\left|\left(\zeta^{\theta, n}, \eta^{\theta, n}\right)_{K}\right| \leq C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\eta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{\theta, n}\right\|_{K}^{2}\right) \tag{4.29}
\end{equation*}
$$

Using (4.24)-(4.29) in (4.23), we have

$$
\begin{align*}
\left\|\eta^{\theta, n}\right\|^{2}-\left\|\eta^{\theta, n-1}\right\|^{2} \leq & C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\zeta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{a, n}\right\|_{K}^{2}+k_{n}^{2}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}\right. \\
& \left.+\left\|\partial_{t} \zeta^{\theta}\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}+\left\|\eta^{a, n}\right\|_{K}^{2}+k_{n}^{2}\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}\right) . \tag{4.30}
\end{align*}
$$

Subtracting (2.13) from (4.6), we obtain

$$
\sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} \eta^{a, n}, w\right)_{K}=\sum_{K \in \mathcal{T}_{h}}\left(f\left(\theta\left(t_{n}\right), a\left(t_{n}\right)\right)-f\left(\theta_{h k}^{n}, a_{h k}^{n}\right), w\right)_{K}-\sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} a\left(t_{n}\right)-\partial_{t} a\left(t_{n}\right), w\right)_{K}-\sum_{K \in \mathcal{T}_{h}}\left(\bar{\partial} \zeta^{a, n}, w\right)_{K}
$$

where $w \in S^{p}$. Substituting $w=\eta^{a, n}$, proceeding as in (4.21)-(4.22) and using Remark 2.2, we obtain

$$
\begin{align*}
\left\|\eta^{a, n}\right\|^{2}- & \left\|\eta^{a, n-1}\right\|^{2} \\
\leq & C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\eta^{\theta, n}\right\|_{K}^{2}+\left\|\eta^{a, n}\right\|_{K}^{2}+\left\|\zeta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{a, n}\right\|_{K}^{2}+\left\|\bar{\partial} a\left(t_{n}\right)-\partial_{t} a\left(t_{n}\right)\right\|_{K}^{2}+\left\|\partial_{t} \zeta^{a, n}\right\|_{K}^{2}\right) \\
\leq & C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\zeta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{a, n}\right\|_{K}^{2}+k_{n}^{2}\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\partial_{t} \zeta^{a}\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right. \\
& \left.+\left\|\eta^{a, n}\right\|_{K}^{2}\right) \tag{4.31}
\end{align*}
$$

Adding (4.30) and (4.31), we obtain

$$
\begin{aligned}
\left\|\eta^{a, n}\right\|^{2}+\left\|\eta^{\theta, n}\right\|^{2}- & \left\|\eta^{a, n-1}\right\|^{2}-\left\|\eta^{\theta, n-1}\right\|^{2} \\
\leq & C \sum_{K \in \mathcal{T}_{h}}\left(\left\|\zeta^{\theta, n}\right\|_{K}^{2}+\left\|\zeta^{a, n}\right\|_{K}^{2}+k_{n}^{2}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+k_{n}^{2}\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}\right. \\
& \left.+\left\|\partial_{t} \zeta^{\theta}\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\partial_{t} \zeta^{a}\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}+\left\|\eta^{a, n}\right\|_{K}^{2}+k_{n}^{2}\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}\right)
\end{aligned}
$$

Summing from 1 to $n$ and using the fact that $\theta(0)=\theta_{0}$ and $a(0)=0$, we obtain

$$
\begin{aligned}
\sum_{K \in \mathcal{T}_{h}}\left(\left\|\eta^{a, n}\right\|_{K}^{2}+\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \leq & C\left(\sum_{K \in \mathcal{T}_{h}}\left\|\eta^{\theta, 0}\right\|_{K}^{2}+\sum_{m=1}^{n} \sum_{K \in \mathcal{T}_{h}}\left(\left\|\zeta^{\theta, m}\right\|_{K}^{2}+\left\|\zeta^{a, m}\right\|_{K}^{2}\right.\right. \\
& +\left\|\partial_{t} \zeta^{\theta}\right\|_{L^{\infty}\left(I_{m}, L^{2}(K)\right)}^{2}+\left\|\partial_{t} \zeta^{a}\right\|_{L^{\infty}\left(I_{m}, L^{2}(K)\right)}^{2} \\
& \left.+k_{m}^{2}\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{m}, L^{2}(K)\right)}^{2}+k_{m}^{2}\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{m}, L^{2}(K)\right)}^{2}+k_{m}^{2}\left\|\partial_{t} u\right\|_{L^{2}\left(I_{m}\right)}^{2}\right) \\
& \left.+\sum_{m=1}^{n} \sum_{K \in \mathcal{T}_{h}}\left(\left\|\eta^{\theta, m}\right\|_{K}^{2}+\left\|\eta^{a, m}\right\|_{K}^{2}\right)\right)
\end{aligned}
$$

Now using Gronwall's lemma, Lemmas 3.1 and 3.7, we obtain

$$
\begin{aligned}
\sum_{K \in \mathcal{T}_{h}}\left(\left\|\eta^{a, n}\right\|_{K}^{2}+\right. & \left.\left\|\eta^{\theta, n}\right\|_{K}^{2}\right) \leq C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}+k_{n}^{2}\right)\left(\left\|\theta_{0}\right\|_{H^{s^{\prime}(K)}}^{2}+\|\theta\|_{L^{\infty}\left(I_{n} ; H^{s^{\prime}}(K)\right)}^{2}\right. \\
& +\left\|\partial_{t} \theta\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|a\|_{L^{\infty}\left(I_{n} ; H^{s^{\prime}}(K)\right)}^{2}+\left\|\partial_{t} a\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2} \\
& \left.+\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}\right) .
\end{aligned}
$$

Using triangle inequality, we obtain the required result. This completes the proof.

Next we state the discrete time error for the adjoint equation (2.20)-(2.23).
Theorem 4.2. Let $(z(t), \lambda(t))$ and $\left(z_{h k}^{n}, \lambda_{h k}^{n}\right), n=1,2, \ldots, N$ be the solutions for (2.20)-(2.23) and (4.15)(4.18), respectively. Then,

$$
\begin{aligned}
\| z\left(t_{n-1}\right)- & z_{h k}^{n-1}\left\|^{2}+\right\| \lambda\left(t_{n-1}\right)-\lambda_{h k}^{n-1} \|^{2} \\
\leq & C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left(\max _{K \in \mathcal{T}_{h}} \frac{h_{K}^{2}}{p_{K}}\right) \frac{h_{K}^{2 s-2}}{p_{K}^{2 s^{\prime}-3}}+k_{n}^{2}\right)\left(\left\|\theta_{0}\right\|_{H^{s^{\prime}}(K)}^{2}+\left\|a_{d}\right\|_{H^{s^{\prime}}(K)}^{2}+\|\theta\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}\right. \\
& +\left\|\partial_{t} \theta\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|a\|_{L^{\infty}\left(I_{n}, H^{s^{\prime}}(K)\right)}^{2}+\left\|\partial_{t} a\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t t} \theta\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2} \\
& +\left\|\partial_{t t} a\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\|z\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t} z\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\|\lambda\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2} \\
& \left.+\left\|\partial_{t} \lambda\right\|_{L^{\infty}\left(I_{n}, H^{\left.s^{\prime}(K)\right)}\right.}^{2}+\left\|\partial_{t t} z\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\partial_{t t} \lambda\right\|_{L^{\infty}\left(I_{n}, L^{2}(K)\right)}^{2}+\left\|\partial_{t} u\right\|_{L^{2}\left(I_{n}\right)}^{2}\right), t_{n} \in \bar{I}_{n},
\end{aligned}
$$

where $C>0$ is independent of $p_{K}, h_{K},(\theta, a)$ and $(z, \lambda)$, also $s=\min \left(p_{K}+1, s^{\prime}\right)$ and $s^{\prime}, p_{K} \geq 2$.
Remark 4.3. In addition to the extra-regularity assumptions given in Remark 3.10, we need to assume that the continuous and adjoint solutions and the control $u$ satisfy

$$
\partial_{t t} \theta, \partial_{t t} z, \partial_{t t} a, \partial_{t t} \lambda \in L^{\infty}\left(I, L^{2}(\Omega)\right), \partial_{t} u \in L^{2}(I)
$$

for Theorems 4.1 and 4.2 to hold true.

## Complete discretization

Now, we will discretize the control by using a DGFEM. In order to completely discretize the problem (2.17), we choose a discontinuous Galerkin piecewise constant approximation of the control variable. Let $U_{d}$ be the finite dimensional subspace of $U$ defined by

$$
U_{d}=\left\{v_{d} \in L^{2}(I):\left.v_{d}\right|_{I_{n}}=\text { constant }\right\} \quad \forall n=1,2, \ldots, N
$$

Let $U_{d, a d}=U_{d} \cap U_{a d}$ and $\sigma$ collects all the three discretization parameters $h, k, d$. The completely discretized problem reads as: find $\left(\theta_{\sigma}, a_{\sigma}\right) \in X_{h k}^{q} \times X_{h k}^{q}$ such that

$$
\begin{align*}
\sum_{n=1}^{N}\left(\left(\partial_{t} a_{\sigma}, w\right)_{I_{n}, \Omega}+\left(\left\langle a_{\sigma}\right\rangle_{n-1}, w_{n-1}^{+}\right)\right) & =\sum_{n=1}^{N}\left(f\left(\theta_{\sigma}, a_{\sigma}\right), w\right)_{I_{n}, \Omega}  \tag{4.32}\\
a_{\sigma}(0) & =0  \tag{4.33}\\
\sum_{n=1}^{N}\left(\rho c_{p}\left(\partial_{t} \theta_{\sigma}, v\right)_{I_{n}, \Omega}+\int_{I_{n}} B\left(\theta_{\sigma}, v\right) \mathrm{d} t+\rho c_{p}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, v_{n-1}^{+}\right)\right)= & \sum_{n=1}^{N}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), v\right)_{I_{n}, \Omega}\right. \\
& \left.+\left(\alpha u_{\sigma}, v\right)_{I_{n}, \Omega}\right)  \tag{4.34}\\
\theta_{\sigma}(0)= & \theta_{0} \tag{4.35}
\end{align*}
$$

for all $(w, v) \in X_{h k}^{q} \times X_{h k}^{q}$. Next we establish stability estimates for $\theta_{\sigma}$ and $a_{\sigma}$.
Lemma 4.4. For a fixed control $u_{\sigma} \in U_{d, a d}$, the solution $\left(\theta_{\sigma}, a_{\sigma}\right) \in X_{h k}^{q} \times X_{h k}^{q}$ of (4.32)-(4.35), satisfies the following a priori bounds:

$$
\begin{equation*}
\int_{I} \sum_{K \in \mathcal{T}_{h}}\left\|\theta_{\sigma}\right\|_{H^{1}\left(\Omega, \mathcal{T}_{h}\right)}^{2} \mathrm{~d} s \leq C \text {, i.e., } \quad \theta_{\sigma} \in L^{2}\left(I, H^{1}\left(\Omega, \mathcal{T}_{h}\right)\right) \tag{4.36}
\end{equation*}
$$

Also,

$$
\begin{equation*}
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left\|\partial_{t} \theta_{\sigma}\right\|_{I_{n}, K}^{2}+\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2}\right) \leq C, \quad \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left\|\partial_{t} a_{\sigma}\right\|_{I_{n}, K}^{2} \leq C \tag{4.37}
\end{equation*}
$$

where $\Delta_{h}: S^{p} \rightarrow S^{p}$ is the discrete Laplacian defined by

$$
\begin{equation*}
-\sum_{K \in \mathcal{T}_{h}}\left(\Delta_{h} v, w\right)_{K}=B(v, w), \quad \forall v, w \in S^{p} \tag{4.38}
\end{equation*}
$$

that is, $\left(\theta_{\sigma}, a_{\sigma}\right) \in \mathcal{Y} \times \mathcal{Y}$, where

$$
\mathcal{Y}=\left\{v: v \in L^{2}\left(I, L^{2}(\Omega)\right),\left.v_{t}\right|_{I_{n}} \in L^{2}\left(I_{n}, L^{2}(\Omega)\right) \forall n=1, \ldots, N\right\}
$$

Proof. From (4.34) and definition of $B_{\nu}$, we have

$$
\begin{aligned}
\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}} \rho c_{p}\left(\partial_{t} \theta_{\sigma}, v\right)_{I_{n}, K} \mathrm{~d} s\right. & \left.+\int_{I_{n}} B_{\nu}\left(\theta_{\sigma}, v\right) \mathrm{d} s+\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, v_{n-1}^{+}\right)_{K}\right) \\
& =\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), v\right)_{I_{n}, K}+\left(\alpha u_{\sigma}, v\right)_{I_{n}, K}+\nu\left(\theta_{\sigma}, v\right)_{I_{n}, K}\right)
\end{aligned}
$$

Substituting $v=\theta_{\sigma}$, using the coercivity of $B_{\nu}$ and the fact that $\left\|\left|\theta_{\sigma}\|\mid \geq\| \theta_{\sigma} \|_{H^{1}\left(\Omega, \mathcal{T}_{h}\right)}\right.\right.$, we obtain

$$
\begin{align*}
& \sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}} \frac{\rho c_{p}}{2} \int_{I_{n}} \frac{d}{\mathrm{~d} t}\left\|\theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} s+C \int_{I_{n}}\left\|\theta_{\sigma}\right\|_{H^{1}\left(\Omega, \mathcal{T}_{h}\right)}^{2} \mathrm{~d} s+\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, \theta_{\sigma, n-1}^{+}\right)_{K}\right) \\
& \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), \theta_{\sigma}\right)_{I_{n}, K}+\left(\alpha u_{\sigma}, \theta_{\sigma}\right)_{I_{n}, K}+\nu\left\|\theta_{\sigma}\right\|_{I_{n}, K}^{2}\right) . \tag{4.39}
\end{align*}
$$

Using

$$
\begin{align*}
\int_{I_{n}} \frac{1}{2} \frac{\mathrm{~d}}{\mathrm{~d} t}\left\|\theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} s & =\frac{1}{2}\left(\left\|\theta_{\sigma, n}\right\|_{K}^{2}-\left\|\theta_{\sigma, n-1}^{+}\right\|_{K}^{2}\right) \text { and }  \tag{4.40}\\
\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, \theta_{\sigma, n-1}^{+}\right)_{K} & =\frac{1}{2}\left(\left\|\theta_{\sigma, n-1}^{+}\right\|_{K}^{2}+\left\|\left\langle\theta_{\sigma}\right\rangle_{n-1}\right\|_{K}^{2}-\left\|\theta_{\sigma, n-1}\right\|_{K}^{2}\right), \tag{4.41}
\end{align*}
$$

in (4.39), we obtain

$$
\begin{aligned}
\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}} \frac{\rho c_{p}}{2}\left(\left\|\theta_{\sigma, n}\right\|_{K}^{2}-\left\|\theta_{\sigma, n-1}\right\|_{K}^{2}\right)+\right. & \left.C \int_{I_{n}}\left\|\theta_{\sigma}\right\|_{H^{1}\left(\Omega, \mathcal{T}_{h}\right)}^{2} \mathrm{~d} s\right) \leq \sum_{n=1}^{N}\left(\sum _ { K \in \mathcal { T } _ { h } } \left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), \theta_{\sigma}\right)_{I_{n}, K}\right.\right. \\
& \left.\left.+\left(\alpha u_{\sigma}, \theta_{\sigma}\right)_{I_{n}, K}+\left\|\theta_{\sigma}\right\|_{I_{n}, K}^{2}\right)-\frac{\rho c_{p}}{2}\left\|\left\langle\theta_{\sigma}\right\rangle_{n-1}\right\|_{K}^{2}\right)
\end{aligned}
$$

Using Cauchy-Schwartz inequality, Remark 2.2 and Young's inequality with appropriately chosen Young's constant, we obtain $\theta_{\sigma} \in L^{2}\left(I, H^{1}\left(\Omega, \mathcal{T}_{h}\right)\right)$.

Now we proceed to prove (4.37). Using (4.38) in (4.34), we have

$$
\begin{array}{r}
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\rho c_{p}\left(\partial_{t} \theta_{\sigma}, v\right)_{I_{n}, K}-\left(\Delta_{h} \theta_{\sigma}, v\right)_{I_{n}, K}+\rho c_{p}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, v_{n-1}^{+}\right)_{K}\right) \\
=\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), v\right)_{I_{n}, K}+\left(\alpha u_{\sigma}, v\right)_{I_{n}, K}\right) \tag{4.42}
\end{array}
$$

Put $v=-\Delta_{h} \theta_{\sigma}$ in (4.42) to obtain

$$
\begin{align*}
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\rho c _ { p } \left(\partial_{t} \theta_{\sigma},-\right.\right. & \left.\left.\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}-\left(\Delta_{h} \theta_{\sigma},-\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}+\rho c_{p}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1},-\Delta_{h} \theta_{\sigma, n-1}^{+}\right)_{K}\right) \\
& =\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right),-\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}+\left(\alpha u_{\sigma},-\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}\right) \tag{4.43}
\end{align*}
$$

Again using (4.38) in first and third terms on the left hand side of (4.43), we obtain

$$
\begin{array}{r}
\sum_{n=1}^{N}\left(\rho c_{p} \int_{I_{n}} B\left(\partial_{t} \theta_{\sigma}, \theta_{\sigma}\right) \mathrm{d} t+\sum_{K \in \mathcal{T}_{h}}\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2}+\rho c_{p} B\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, \theta_{\sigma, n-1}^{+}\right)\right) \\
=\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right),-\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}+\left(\alpha u_{\sigma},-\Delta_{h} \theta_{\sigma}\right)_{I_{n}, K}\right) \tag{4.44}
\end{array}
$$

Now we find estimates for the terms in (4.44) one by one. Consider

$$
\begin{align*}
\int_{I_{n}} B\left(\partial_{t} \theta_{\sigma}, \theta_{\sigma}\right) \mathrm{d} t & =\mathcal{K} \sum_{K \in \mathcal{T}_{h}} \int_{I_{n}}\left(\partial_{t} \nabla \theta_{\sigma}, \nabla \theta_{\sigma}\right)_{K} \mathrm{~d} t-\mathcal{K} \sum_{e \in \mathcal{E}_{i n t}} \int_{I_{n}}\left(\left\{\nabla\left(\partial_{t} \theta_{\sigma}\right) \cdot \mathbf{n}\right\},\left[\theta_{\sigma}\right]\right)_{e} \mathrm{~d} t \\
& -\mathcal{K} \sum_{e \in \mathcal{E}_{i n t}} \int_{I_{n}}\left(\left\{\nabla \theta_{\sigma} \cdot \mathbf{n}\right\},\left[\partial_{t} \theta_{\sigma}\right]\right)_{e} \mathrm{~d} t+\sum_{e \in \mathcal{E}_{i n t}} \frac{\gamma}{|e|} \int_{I_{n}}\left(\left[\partial_{t} \theta_{\sigma}\right],\left[\theta_{\sigma}\right]\right)_{e} \mathrm{~d} t \\
& =\sum_{K \in \mathcal{T}_{h}} \mathcal{K} I_{1}-\sum_{e \in \mathcal{E}_{i n t}} \mathcal{K}\left(I_{2}+I_{3}\right)-\sum_{e \in \mathcal{E}_{i n t}} I_{4}, \quad \text { say. } \tag{4.45}
\end{align*}
$$

For $I_{1}$, we have

$$
\begin{equation*}
I_{1}=\int_{I_{n}}\left(\partial_{t} \nabla \theta_{\sigma}, \nabla \theta_{\sigma}\right)_{K} \mathrm{~d} t=\int_{I_{n}} \frac{1}{2} \frac{d}{\mathrm{~d} t}\left\|\nabla \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t=\frac{1}{2}\left(\left\|\nabla \theta_{\sigma, n}\right\|_{K}^{2}-\left\|\nabla \theta_{\sigma, n-1}^{+}\right\|_{K}^{2}\right) \tag{4.46}
\end{equation*}
$$

Using integration by parts for $I_{2}$, we have

$$
\begin{equation*}
I_{2}=\int_{I_{n}}\left(\partial_{t}\left\{\nabla \theta_{\sigma} \cdot \mathbf{n}\right\},\left[\theta_{\sigma}\right]\right)_{e} \mathrm{~d} t=\left.\left(\left\{\nabla \theta_{\sigma} \cdot \mathbf{n}\right\},\left[\theta_{\sigma}\right]\right)_{e}\right|_{I_{n}}-\int_{I_{n}}\left(\left\{\nabla \theta_{\sigma} \cdot \mathbf{n}\right\},\left[\partial_{t} \theta_{\sigma}\right]\right)_{e} \mathrm{~d} t=\left.\left(\left\{\nabla \theta_{\sigma} \cdot \mathbf{n}\right\},\left[\theta_{\sigma}\right]\right)_{e}\right|_{I_{n}}-I_{3} . \tag{4.47}
\end{equation*}
$$

For $I_{4}$, we have

$$
\begin{equation*}
I_{4}=\int_{I_{n}}\left(\partial_{t}\left[\theta_{\sigma}\right],\left[\theta_{\sigma}\right]\right)_{e} \mathrm{~d} t=\int_{I_{n}} \frac{1}{2} \frac{d}{\mathrm{~d} t}\left\|\left[\theta_{\sigma}\right]\right\|_{e}^{2} \mathrm{~d} t=\left.\frac{1}{2}\left\|\left[\theta_{\sigma}\right]\right\|_{e}^{2}\right|_{I_{n}} . \tag{4.48}
\end{equation*}
$$

Using (4.46)-(4.48) in (4.45), we obtain

$$
\begin{align*}
\int_{I_{n}} B\left(\partial_{t} \theta_{\sigma}, \theta_{\sigma}\right) \mathrm{d} t= & \frac{1}{2} \sum_{K \in \mathcal{T}_{h}} \mathcal{K}\left(\left\|\nabla \theta_{\sigma, n}\right\|_{K}^{2}-\left\|\nabla \theta_{\sigma, n-1}^{+}\right\|_{K}^{2}\right)-\left.\sum_{e \in \mathcal{E}_{i n t}} \mathcal{K}\left(\left\{\nabla \theta_{\sigma}\right\},\left[\theta_{\sigma}\right]\right)_{e}\right|_{I_{n}} \\
& +\left.\frac{1}{2} \sum_{e \in \mathcal{E}_{i n t}} \frac{\gamma}{|e|}\left\|\left[\theta_{\sigma}\right]\right\|_{e}^{2}\right|_{I_{n}} \tag{4.49}
\end{align*}
$$

Using the definition of $B(\cdot, \cdot)$ in the third term on the left hand side of the (4.44), we obtain

$$
\begin{align*}
B\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, \theta_{\sigma, n-1}^{+}\right)= & \sum_{K \in \mathcal{T}_{h}} \mathcal{K}\left(\langle\nabla \theta\rangle_{n-1}, \nabla \theta_{\sigma, n-1}^{+}\right)_{K}-\sum_{e \in \mathcal{E}_{\text {int }}}\left(\mathcal{K}\left(\left\{\nabla\left\langle\theta_{\sigma}\right\rangle_{n-1} \cdot \mathbf{n}\right\},\left[\theta_{\sigma, n-1}^{+}\right]\right)_{e}\right. \\
& \left.+\mathcal{K}\left(\left\{\nabla \theta_{\sigma, n-1}^{+} \cdot \mathbf{n}\right\},\left[\left\langle\theta_{\sigma}\right\rangle_{n-1}\right]\right)_{e}-\frac{\gamma}{|e|}\left(\left[\left\langle\theta_{\sigma}\right\rangle_{n-1}\right],\left[\theta_{\sigma, n-1}^{+}\right]\right)_{e}\right) . \tag{4.50}
\end{align*}
$$

Using $\left(\left\langle\nabla \theta_{\sigma}\right\rangle_{n-1}, \nabla \theta_{\sigma, n-1}^{+}\right)_{K}=\frac{1}{2}\left(\left\|\nabla \theta_{\sigma, n-1}^{+}\right\|_{K}^{2}+\left\|\left\langle\nabla \theta_{\sigma}\right\rangle_{n-1}\right\|_{K}^{2}-\left\|\nabla \theta_{\sigma, n-1}\right\|_{K}^{2}\right)$,
in (4.50), we have

$$
\begin{align*}
B\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, \theta_{\sigma, n-1}^{+}\right)= & \sum_{K \in \mathcal{T}_{h}} \frac{\mathcal{K}}{2}\left(\left\|\nabla \theta_{\sigma, n-1}^{+}\right\|_{K}^{2}+\left\|\left\langle\nabla \theta_{\sigma}\right\rangle_{n-1}\right\|_{K}^{2}-\left\|\nabla \theta_{\sigma, n-1}\right\|_{K}^{2}\right) \\
& -\sum_{e \in \mathcal{E}_{i n t}}\left(\mathcal{K}\left(\left\{\nabla\left\langle\theta_{\sigma}\right\rangle_{n-1} \cdot \mathbf{n}\right\},\left[\theta_{\sigma, n-1}^{+}\right]\right)_{e}\right. \\
& \left.+\mathcal{K}\left(\left\{\nabla \theta_{\sigma, n-1}^{+} \cdot \mathbf{n}\right\},\left[\left\langle\theta_{\sigma}\right\rangle_{n-1}\right]\right)_{e}-\frac{\gamma}{|e|}\left(\left[\left\langle\theta_{\sigma}\right\rangle_{n-1}\right],\left[\theta_{\sigma, n-1}^{+}\right]\right)_{e}\right) . \tag{4.52}
\end{align*}
$$

Using (4.49), (4.52), Cauchy-Schwarz and Young's inequalities in (4.44), we have

$$
\begin{aligned}
\sum_{K \in \mathcal{T}_{h}}\left(\left\|\nabla \theta_{\sigma, N}\right\|_{K}^{2}-\left\|\nabla \theta_{0}\right\|_{K}^{2}\right)+ & \sum_{n=1} \sum_{K \in \mathcal{T}_{h}}\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2} \\
\leq & C \sum_{n=1}^{N}\left(\sum _ { K \in \mathcal { T } _ { h } } \left(\left\|f\left(\theta_{\sigma}, a_{\sigma}\right)\right\|_{I_{n}, K}^{2}+\left\|\alpha u_{\sigma}\right\|_{I_{n}, K}^{2}+\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2}\right.\right. \\
& +\sum_{e \in \mathcal{E}_{i n t}}\left(\left\|\nabla \theta_{\sigma, n} \cdot \mathbf{n}\right\|_{e}^{2}+\left\|\nabla \theta_{\sigma, n-1}^{+} \cdot \mathbf{n}\right\|_{e}^{2}+\left\|\theta_{\sigma, n}\right\|_{e}^{2}+\left\|\theta_{\sigma, n-1}\right\|_{e}^{2}\right. \\
& \left.\left.+\left\|\theta_{\sigma, n-1}^{+}\right\|_{e}^{2}\right)\right) .
\end{aligned}
$$

Choosing Young's constant appropriately, using Remark 2.2 and $\theta_{\sigma} \in L^{2}\left(I, H^{1}\left(\Omega, \mathcal{T}_{h}\right)\right)$, we obtain

$$
\begin{equation*}
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2} \text { is bounded. } \tag{4.53}
\end{equation*}
$$

Put $v=\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}$ in (4.34), use $\left(\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}\right)_{n-1}^{+}=0$ and (4.38) to obtain

$$
\begin{align*}
\rho c_{p} \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\partial_{t} \theta_{\sigma},\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}\right)_{I_{n}, K} & -\sum_{n=1}^{N} \int_{I_{n}}\left(\Delta_{h} \theta_{\sigma},\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}\right)_{I_{n}, K} \\
= & \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right),\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}\right)_{I_{n}, K}\right. \\
& \left.+\left(\alpha u_{\sigma},\left(t-t_{n-1}\right) \partial_{t} \theta_{\sigma}\right)_{I_{n}, K}\right) \tag{4.54}
\end{align*}
$$

Use Cauchy-Schwarz inequality and Young's inequality to obtain

$$
\begin{aligned}
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}} \int_{I_{n}}\left(t-t_{n-1}\right)\left\|\partial_{t} \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t \leq & C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left\|f\left(\theta_{\sigma}, a_{\sigma}\right)\right\|_{I_{n}, K}^{2}+\left\|\alpha u_{\sigma}\right\|_{I_{n}, K}^{2}+\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2}\right. \\
& \left.+\int_{I_{n}}\left(t-t_{n-1}\right)\left\|\partial_{t} \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t\right)
\end{aligned}
$$

Choosing Young's constant appropriately, using (4.53) and Remark 2.2, we obtain

$$
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}} \int_{I_{n}}\left(t-t_{n-1}\right)\left\|\partial_{t} \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t \text { is bounded. }
$$

From inverse estimate, we have

$$
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}} \int_{I_{n}}\left\|\partial_{t} \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t \leq C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}} k_{n}^{-1} \int_{I_{n}}\left(t-t_{n-1}\right)\left\|\partial_{t} \theta_{\sigma}\right\|_{K}^{2} \mathrm{~d} t .
$$

Therefore,

$$
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left(\left\|\partial_{t} \theta_{\sigma}\right\|_{I_{n}, K}^{2}+\left\|\Delta_{h} \theta_{\sigma}\right\|_{I_{n}, K}^{2}\right) \leq C
$$

Similarly putting $w=\left(t-t_{n-1}\right) \partial_{t} a_{\sigma}$ in (4.32) and using inverse estimate, we obtain

$$
\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_{h}}\left\|\partial_{t} a_{\sigma}\right\|_{I_{n}, K}^{2} \leq C
$$

The discrete space-time-control DGFEM scheme for the optimal control problem is

$$
\begin{equation*}
\min J\left(\theta_{\sigma}, a_{\sigma}, u_{\sigma}\right) \quad \text { subject to the constraints }(4.32)-(4.35) \text { and } u_{\sigma} \in U_{d, a d} \tag{4.55}
\end{equation*}
$$

where $\left(\theta_{\sigma}(t), a_{\sigma}(t), u_{\sigma}(t)\right)=\left(\theta_{\sigma}^{n}, a_{\sigma}^{n}, u_{\sigma}^{n}\right), \quad$ for $t \in I_{n}$.
Lemma 4.5. Let assumptions (A1)-(A6) hold true. Let $u^{*}$ and $u_{\sigma}^{*}$ be the optimal solutions of (2.7) and (4.55), respectively. Also, let $u_{\sigma}^{*} \longrightarrow u^{*}$ weakly in $L^{2}(I)$. Then, under the assumptions of Theorem 4.1, we have

$$
\begin{array}{ll}
\theta_{\sigma}^{*} \longrightarrow \theta^{*} & \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right), \\
a_{\sigma}^{*} \longrightarrow a^{*} & \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right), \tag{4.57}
\end{array}
$$

as the discretization parameters $h, k$ (and hence $\sigma$ ) $\rightarrow 0$, where $\left(\theta^{*}, a^{*}\right)$ and $\left(\theta_{\sigma}^{*}, a_{\sigma}^{*}\right)$ are the solutions of (2.8)(2.12) and (4.32)-(4.35), respectively.

Proof. The solution $\left(\theta^{*}, a^{*}\right)$ of (2.8)-(2.12), under the regularity assumption that $\theta^{*}(t) \in \mathcal{U}, t \in \bar{I}$, satisfies

$$
\begin{align*}
\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}} \rho c_{p}\left(\partial_{t} \theta^{*}, v\right)_{I_{n}, K}+\int_{I_{n}}\right. & \left.B\left(\theta^{*}, v\right) \mathrm{d} s+\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\left\langle\theta^{*}\right\rangle_{n-1}, v_{n-1}^{+}\right)_{K}\right) \\
& =\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}}-\rho L\left(f\left(\theta^{*}, a^{*}\right), v\right)_{I_{n}, K}+\sum_{K \in \mathcal{T}_{h}}\left(\alpha u^{*}, v\right)_{I_{n}, K}\right) \quad \forall v \in S^{p} \tag{4.58}
\end{align*}
$$

Subtracting (4.34) from (4.58), we obtain

$$
\begin{align*}
\sum_{n=1}^{N}\left(\sum _ { K \in \mathcal { T } _ { h } } \rho c _ { p } \int _ { I _ { n } } \left(\partial _ { t } \left(\theta^{*}\right.\right.\right. & \left.\left.\left.-\theta_{\sigma}^{*}\right), v\right)_{K} \mathrm{~d} s+\int_{I_{n}} B\left(\theta^{*}-\theta_{\sigma}^{*}, v\right) \mathrm{d} s+\rho c_{p} \sum_{K \in \mathcal{I}_{h}}\left(\left\langle\theta^{*}-\theta_{\sigma}^{*}\right\rangle_{n-1}, v_{n-1}^{+}\right)_{K}\right) \\
& =\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}}-\rho L \int_{I_{n}}\left(f\left(\theta^{*}, a^{*}\right)-f\left(\theta_{\sigma}^{*}, a_{\sigma}^{*}\right), v\right)_{K} \mathrm{~d} s+\sum_{K \in \mathcal{T}_{h}} \int_{I_{n}}\left(\alpha\left(u^{*}-u_{\sigma}^{*}\right), v\right)_{K} \mathrm{~d} s\right) \tag{4.59}
\end{align*}
$$

Using (A4), $v \in S_{p}$ and the fact that $u_{\sigma}^{*} \longrightarrow u^{*}$ weakly in $L^{2}(I)$, we obtain

$$
\begin{align*}
\sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}} \rho c_{p} \int_{I_{n}}\left(\partial_{t}\left(\theta^{*}-\theta_{\sigma}^{*}\right), v\right)_{K} \mathrm{~d} s\right. & \left.+\int_{I_{n}} B\left(\theta^{*}-\theta_{\sigma}^{*}, v\right) \mathrm{d} s+\rho c_{p} \sum_{K \in \mathcal{T}_{h}}\left(\left\langle\theta^{*}-\theta_{\sigma}^{*}\right\rangle_{n-1}, v_{n-1}^{+}\right)_{K}\right) \\
& \leq \sum_{n=1}^{N}\left(\sum_{K \in \mathcal{T}_{h}}-\rho L \int_{I_{n}}\left(f\left(\theta^{*}, a^{*}\right)-f\left(\theta_{\sigma}^{*}, a_{\sigma}^{*}\right), v\right)_{K} \mathrm{~d} s\right) \tag{4.60}
\end{align*}
$$

Now proceeding in similar lines as in the proof of Theorem 4.1, we obtain the required result.

Theorem 4.6. Let $u_{\sigma}^{*}$ be the optimal control of (4.55), for $0<\epsilon<1$. Then, there exists a subsequence of $\left\{u_{\sigma}^{*}\right\}$ (still denoted by $\left.\left\{u_{\sigma}^{*}\right\}\right)$ such that $\lim _{\sigma \rightarrow 0} u_{\sigma}^{*}=u^{*}$ exists in $L^{2}(I)$ and $u^{*}$ is an optimal control of (2.17).

Proof. Since $u_{\sigma}^{*}$ is an optimal control, we obtain

$$
\left\|u_{\sigma}^{*}\right\|_{L^{2}(I)} \leq C
$$

that is, $\left\{u_{\sigma}^{*}\right\}_{\sigma>0}$ is uniformly bounded in $L^{2}(I)$. Thus, it is possible to extract a subsequence say $\left\{u_{\sigma}^{*}\right\}_{\sigma>0}$ in $L^{2}(I)$ such that

$$
\begin{equation*}
u_{\sigma}^{*} \longrightarrow u^{*} \text { weakly in } L^{2}(I) \tag{4.61}
\end{equation*}
$$

Since $U_{a d} \subset L^{2}(I)$ is a closed and convex set, we have $u^{*} \in U_{a d}$. Now corresponding to each $u_{\sigma}^{*}$ there exists solution $\left(\theta_{\sigma}^{*}, a_{\sigma}^{*}\right)$ to (4.32)-(4.35) converging to $\theta^{*}$ and $a^{*}$ strongly in $L^{\infty}\left(I, L^{2}(\Omega)\right)$ from Lemma 4.5. Thus from Lemma 4.4 and Lemma 4.5, we have

$$
\begin{align*}
& a_{\sigma}^{*} \longrightarrow a^{*} \text { weakly in } \mathcal{Y},  \tag{4.62}\\
& \theta_{\sigma}^{*} \longrightarrow \theta^{*} \text { weakly in } \mathcal{Y} \cap L^{2}\left(I, H^{1}\left(\Omega, \mathcal{T}_{h}\right)\right),  \tag{4.63}\\
& \theta_{\sigma}^{*} \longrightarrow \theta^{*}  \tag{4.64}\\
& a_{\sigma}^{*} \longrightarrow a^{*}  \tag{4.65}\\
& \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right) \\
& \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right)
\end{align*}
$$

Now passing limit as $\sigma \rightarrow 0(h \rightarrow 0, k \rightarrow 0)$, using (4.63)-(4.65) and Remark 2.2 in (4.32)-(4.35), we obtain that $\left(u^{*}, \theta^{*}, a^{*}\right)$ is an admissible solution for the optimal control problem (2.17). It now remains to show that $\left(u^{*}, \theta^{*}, a^{*}\right)$ is an optimal solution.

If possible, let $\left(\bar{u}^{*}, \bar{\theta}^{*}, \bar{a}^{*}\right)$ be another optimal solution of (2.17). Consider the auxiliary problem

$$
\begin{align*}
\sum_{n=1}^{N}\left(\left(\partial_{t} a_{\sigma}, w\right)_{\Omega, I_{n}}+\left(\left\langle a_{\sigma}\right\rangle_{n-1}, w_{n-1}^{+}\right)\right) & =\sum_{n=1}^{N}\left(f\left(\theta_{\sigma}, a_{\sigma}\right), w\right)_{\Omega, I_{n}}  \tag{4.66}\\
a_{\sigma}(0) & =0  \tag{4.67}\\
\sum_{n=1}^{N}\left(\rho c_{p}\left(\partial_{t} \theta_{\sigma}, v\right)_{\Omega, I_{n}}+\int_{I_{n}} B\left(\theta_{\sigma}, v\right) \mathrm{d} t+\rho c_{p}\left(\left\langle\theta_{\sigma}\right\rangle_{n-1}, v_{n-1}^{+}\right)\right)= & \sum_{n=1}^{N}\left(-\rho L\left(f\left(\theta_{\sigma}, a_{\sigma}\right), v\right)_{\Omega, I_{n}}\right. \\
& \left.+\left(\alpha \pi_{k} \bar{u}^{*}, v\right)_{\Omega, I_{n}}\right)  \tag{4.68}\\
\theta_{\sigma}(0)= & \theta_{0} \tag{4.69}
\end{align*}
$$

for all $(w, v) \in X_{h k}^{q} \times X_{h k}^{q}$. Then, there exists a solution to (4.66)-(4.69), say $\left(\bar{\theta}_{\sigma}, \bar{a}_{\sigma}\right)$ converging to $(\bar{\theta}, \bar{a})$ as $\sigma \rightarrow 0$. Similar to (4.63)-(4.65), we arrive at

$$
\begin{align*}
& \bar{a}_{\sigma} \longrightarrow \bar{a} \text { weakly in } \mathcal{Y},  \tag{4.70}\\
& \bar{\theta}_{\sigma} \longrightarrow \bar{\theta} \text { weakly in } \mathcal{Y} \cap L^{2}\left(I, H^{1}\left(\Omega, \mathcal{T}_{h}\right)\right),  \tag{4.71}\\
& \bar{\theta}_{\sigma} \longrightarrow \bar{\theta} \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right),  \tag{4.72}\\
& \bar{a}_{\sigma} \longrightarrow \bar{a} \text { strongly in } L^{\infty}\left(I, L^{2}(\Omega)\right) . \tag{4.73}
\end{align*}
$$

Now letting $\sigma \rightarrow 0$ in (4.66)-(4.69), we obtain that $(\bar{\theta}, \bar{a})$ is a unique solution of (2.13)-(2.16) with respect to the control $\bar{u}^{*}$. Since the solution to (2.13)-(2.16) for a fixed control is unique, we find that $\bar{\theta}=\bar{\theta}^{*}$ and $\bar{a}=\bar{a}^{*}$. Since $u_{\sigma}^{*}$ is the optimal control for (4.55), we have

$$
\begin{equation*}
j\left(u_{\sigma}^{*}\right) \leq j\left(\pi_{k} \bar{u}^{*}\right) \tag{4.74}
\end{equation*}
$$

Now letting $\sigma \rightarrow 0$ in (4.74) and using (4.61), we obtain

$$
\begin{equation*}
j\left(u^{*}\right) \leq j\left(\bar{u}^{*}\right) \tag{4.75}
\end{equation*}
$$

Hence $u^{*}$ is the optimal control. Next we need to show that $\lim _{\sigma \rightarrow 0}\left\|u_{\sigma}^{*}-u\right\|_{L^{2}(I)}=0$. Since $u_{\sigma}^{*} \longrightarrow u^{*}$ weakly in $L^{2}(\Omega)$, it is enough show that $\lim _{\sigma \rightarrow 0}\left\|u_{\sigma}^{*}\right\|_{L^{2}(I)}=\left\|u^{*}\right\|_{L^{2}(I)}$.

Using (4.63)-(4.65), we find that

$$
\begin{aligned}
\lim _{\sigma \rightarrow 0} \frac{\beta_{3}}{2}\left\|u_{\sigma}^{*}\right\|_{L^{2}(I)}^{2} & =\lim _{\sigma \rightarrow 0}\left(J\left(\theta_{\sigma}^{*}, a_{\sigma}^{*}, u_{\sigma}^{*}\right)-\frac{\beta_{1}}{2}\left\|a_{\sigma}^{*}(T)-a_{d}\right\|^{2}-\frac{\beta_{2}}{2}\left\|\left[\theta_{\sigma}^{*}-\theta_{m}\right]_{+}\right\|_{I, \Omega}^{2}\right) \\
& =J\left(\theta^{*}, a^{*}, u^{*}\right)-\frac{\beta_{1}}{2}\left\|a^{*}(T)-a_{d}\right\|^{2}-\frac{\beta_{2}}{2}\left\|\left[\theta^{*}-\theta_{m}\right]_{+}\right\|_{I, \Omega}^{2} \\
& =\frac{\beta_{3}}{2}\left\|u^{*}\right\|_{L^{2}(I)}^{2}
\end{aligned}
$$

Therefore, we have $\lim _{\sigma \rightarrow 0}\left\|u_{\sigma}^{*}\right\|_{L^{2}(I)}=\left\|u^{*}\right\|_{L^{2}(I)}$ and $\lim _{\sigma \rightarrow 0}\left\|u_{\sigma}^{*}-u^{*}\right\|=0$. This completes the rest of the proof.


Figure 3. (a) Goal $a_{d}$ to be achieved for the volume fraction of austenite; (b) triangulation permitting hanging nodes.

## 5. Numerical experiments

In this section, we consider two examples and observe the performance of $h p$-DGFEM for the laser surface hardening of steel problem. An attempt has been made to achieve (i) a constant hardening of 1 mm near the boundary in the first example (ii) a non uniform hardening near the boundary in the second example. We have used non-linear conjugate gradient method [36] to evaluate the optimal control for the complete discretized problem.

## Physical data [36]

The parameters in the heat equation used are given by $\rho c_{p}=4.91 \frac{\mathrm{~J}}{\mathrm{~cm}^{3} \mathrm{~K}}, \mathcal{K}=0.64 \frac{\mathrm{~J}}{\mathrm{~cm}^{3} \mathrm{~K}}$ and $\rho L=627.9 \frac{\mathrm{~J}}{\mathrm{~cm}^{3} \mathrm{~K}}$. The regularized monotone function $\mathcal{H}_{\epsilon}$ is chosen as

$$
\mathcal{H}_{\epsilon}(s)=\left\{\begin{array}{lr}
1 & s \geq \epsilon \\
10\left(\frac{s}{\epsilon}\right)^{6}-24\left(\frac{s}{\epsilon}\right)^{5}+15\left(\frac{s}{\epsilon}\right)^{4} & 0<s \leq \epsilon \\
0 & s \leq 0
\end{array}\right.
$$

where $\epsilon=0.15$. The initial temperature $\theta_{0}$ and the melting temperature $\theta_{m}$ are chosen as 20 and 1800 , respectively. Pointwise data for $a e q(\theta)$ and $\tau(\theta)$ are given by

| $\theta$ | 730 | 830 | 840 | 930 |
| :---: | :---: | :---: | :---: | :---: |
| $a e q(\theta)$ | 0 | 0.91 | 1 | 1 |
| $\tau(\theta)$ | 1 | 0.2 | 0.18 | 0.05 |

The shape function $\alpha(x, y, t)$ is given by $\alpha(x, y, t)=\frac{4 k_{1} A}{\pi D^{2}} \exp \left(-\frac{2(x-v t)^{2}}{D^{2}}\right) \exp \left(k_{1} y\right)$, where $D=0.47 \mathrm{~cm}$, $k_{1}=60 / \mathrm{cm}, A=0.3 \mathrm{~cm}$ and $v=1 \mathrm{~cm} / \mathrm{s}$. In the nonlinear conjugate gradient method tolerance is chosen as $10^{-7}$. We choose $\beta_{1}=7500, \beta_{2}=1000$ and $\beta_{3}=10^{-3}$.

Example 1. The main aim of this experiment is to achieve a constant hardening depth of 1 mm near the boundary, see Figure 3a. We choose a triangulation which permits hanging nodes, see Figure 3b. To apply non-linear conjugate gradient method for the optimal control problem, we take $u_{0}$ (initial control) as 1404 .

We investigate the convergence of $h p$-DGFEM on a sequence of meshes with polynomial degree of approximation $p=1$ and 2. Similarly, convergence has been established by enriching the polynomial degree $p$ for a fixed mesh.


Figure 4. (a) Convergence of $h p$-DGFEM with $h$-refinement: temperature and austenite, convergence of $h p$-DGFEM with $p$-refinement: (b) austenite (c) temperature.

For the purpose of computation, penalty parameter is taken as $\gamma=10$. In Figure 4a, we plot the $L^{2}$-norm of the error against the discretization parameter $h$ for polynomial degrees $p=1,2$. Here, we observe that $\left\|\theta-\theta_{\sigma}\right\|$ and $\left\|a-a_{\sigma}\right\|$ converges to zero at the rate of $\mathcal{O}\left(h^{p}\right)$ as the mesh is refined. The numerical results justify the theoretical results obtained. In Figures 4 b and 4 c , we present the convergence of the solution in $L^{2}$-norm as the degree of polynomials increases for a fixed mesh. Figure 5 a shows the convergence as $k$ is refined. We plot the error in $\theta$ and $a$ in $L^{2}$-norm against the time mesh parameter $k$. In Figure 5 b , we plot the error in control function $u$ computed in $L^{2}$-norm against the time discretization parameter $k$. The computational order of convergence for the control function is approximately equal to 2 .


Figure 5. Convergence of DGFEM with $k$-refinement.


Figure 6. (a) Graph of computed austenite variable at $T=1$; (b) graph of computed temperature variable at $T=1$.

Figures 6 a and 6 b show the graphs of the austenite and temperature variables at the final time $T$ after using $h p$-DGFEM for the discretization in space and a DGFEM for space and control variables.

Figure 7 shows the evolution of control variable (laser energy) in time. At first the laser energy has increased and then during the long term it can be kept a constant. Towards the end of the process it has to be reduced to cope the accumulation of the heat at the end of the plate.

Example 2. In this example we implement the optimal control problem of laser surface hardening of steel aiming at a hardening of depth 1 mm from $x=0$ to $x=2.5$ and of 0.5 mm from $x>2.5$ to $x=5$, near the boundary (see Fig. 8). The physical data for this example is same as the one used for Example 1.

The convergence of $h p$-DGFEM as the discretization parameter $h$ tends to 0 for polynomial degrees of approximation $p=1$ and 2 is illustrated in Figure 9a. We obtain that the $L^{2}$-norm of the error in $a$ and $\theta$


Figure 7. Laser energy.


Figure 8. Goal $a_{d}$ to be achieved for the volume fraction of austenite.
converges to zero at the rate of $\mathcal{O}\left(h^{p}\right)$ as the mesh is refined. Figures 9 b and 9 c represent the convergence of the solution in $L^{2}$-norm as the degree of polynomial is increased.

Figures 10a and 10b represent the hardening of steel at time $t=2.625$ and $t=5.25$, respectively. It shows that a hardening of 1 mm is achieved as the laser beam moves from $x=0$ to $x=2.5$ and after that a hardening of 0.5 mm of hardening is achieved. Figures 11a and 11b illustrate that the temperature is higher when 1 mm of hardening is needed and then it lowers for 0.5 mm of hardening.

Figure 12a shows the convergence of 'a' and ' $\theta$ ' as $k$ is refined. Figure 12b shows the numerical order of convergence obtained for the control variable $u$. Figure 13 shows the graph of the laser beam, that is the control variable. Since hardening depth of steel in the first half is more than that in the second half, the intensity of laser beam is higher in the beginning. As similar to Example 1, it increases at first and then it can be kept constant till $t=2.625$ and then it goes down to perform the hardening of 0.5 mm . The latter half represents that again the laser beam can be kept constant till it reaches $t=5.25$.

(a)

(b)

(c)

Figure 9. (a) Convergence of $h p$-DGFEM with $h$-refinement: temperature and austenite, convergence of $h p$-DGFEM with $p$-refinement: (b) austenite (c) temperature.


Figure 10. Graph of computed hardening of austenite at time (a) $t=2.625$; (b) $t=5.25$.


Figure 11. Graph of computed hardening of temperature and time (a) $t=2.625$; (b) $t=5.25$.


Figure 12. Convergence of DGFEM with $k$-refinement.


Figure 13. Laser energy.

Acknowledgements. The authors are thankful to the referees for their excellent comments and suggestions which have improved the quality of the work.

## References

[1] V. Arnăutu, D. Hömberg and J. Sokołowski, Convergence results for a nonlinear parabolic control problem. Numer. Funct. Anal. Optim. 20 (1999) 805-824.
[2] D.N. Arnold, An interior penalty method for discontinuous elements. SIAM J. Numer. Anal. 19 (1982) 742-760.
[3] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. 39 (2002) 1749-1779.
[4] I. Babuška, The finite element method with penalty. Math. Comput. 27 (1973) 221-228.
[5] C. Bernardi, M. Dauge and Y. Maday, Polynomials in the Sobolev world. Preprint of the Laboratoire Jacques-Louis Lions R03038 (2003).
[6] P.G. Ciarlet, The finite element method for elliptic problems. North-Holland, Amsterdam (1978).
[7] M. Crouzeix, V. Thomee, L.B. Wahlbin, Error estimates for spatial discrete approximation of semilinear parabolic equation with initial data of low regularity. Math. Comput. 53187 (1989) 25-41.
[8] J. Douglas, Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, Computing Methods in Applied Sciences, Lecture Notes in Phys. 58. Springer-Verlag, Berlin (1976).
[9] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems I: a linear model problem. SIAM J. Numer. Anal. 28 (1991) 43-77.
[10] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems II: optimal error estimates in $L_{\infty} L_{2}$ and $L_{\infty} L_{\infty} . S I A M$ J. Numer. Anal. 32 (1995) 706-740.
[11] L.C. Evans, Partial Differential Equations. American Mathematics Society, Providence, Rhode Island (1998).
[12] T. Gudi, N. Nataraj and A.K. Pani, Discontinuous Galerkin methods for quasi-linear elliptic problems of nonmonotone type. SIAM J. Numer. Anal. 45 (2007) 163-192.
[13] T. Gudi, N. Nataraj and A.K. Pani, $h p$-discontinuous Galerkin methods for strongly nonlinear elliptic boundary value problems. Numer. Math. 109 (2008) 233-268.
[14] T. Gudi, N. Nataraj and A.K. Pani, An $h p$-local discontinuous Galerkin method for some quasilinear elliptic boundary value problems of nonmonotone type. Math. Comput. 77 (2008) 731-756.
[15] N. Gupta, N. Nataraj and A.K. Pani, An optimal control problem of laser surface hardening of steel. Int. J. Numer. Anal. Model. 7 (2010).
[16] N. Gupta, N. Nataraj and A.K. Pani, A priori error estimates for the optimal control of laser surface hardening of steel. Paper communicated.
[17] D. Hömberg, A mathematical model for the phase transitions in eutectoid carbon steel. IMA J. Appl. Math. 54 (1995) 31-57.
[18] D. Hömberg, Irreversible phase transitions in steel. Math. Methods Appl. Sci. 20 (1997) 59-77.
[19] D. Hömberg and J. Fuhrmann, Numerical simulation of surface hardening of steel. Int. J. Numer. Meth. Heat Fluid Flow 9 (1999) 705-724.
[20] D. Hömberg and J. Sokolowski, Optimal control of laser hardening. Adv. Math. Sci. 8 (1998) 911-928.
[21] D. Hömberg and S. Volkwein, Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition. Math. Comput. Model. 37 (2003) 1003-1028.
[22] D. Hömberg and W. Weiss, PID-control of laser surface hardening of steel. IEEE Trans. Control Syst. Technol. 14 (2006) 896-904.
[23] P. Houston, C. Schwab and E. Süli, Discontinuous $h p$-finite element methods for advection-diffusion-reaction problems. SIAM J. Numer. Anal. 39 (2002) 2133-2163.
[24] A. Lasis and E. Süli, hp-version discontinuous Galerkin finite element methods for semilinear parabolic problems. Report No. 03/11, Oxford university computing laboratory (2003).
[25] A. Lasis and E. Süli, Poincaré-type inequalities for broken Sobolev spaces. Isaac Newton Institute for Mathematical Sciences, Preprint No. NI03067-CPD (2003).
[26] J.B. Leblond and J. Devaux, A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size. Acta Metall. 32 (1984) 137-146.
[27] V.I. Mazhukin and A.A. Samarskii, Mathematical modelling in the technology of laser treatments of materials. Surveys Math. Indust. 4 (1994) 85-149.
[28] 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 Optim. 47 (2007) 1150-1177.
[29] J.A. Nitsche, Über ein Variationprinzip zur Lösung Dirichlet-Problemen bei Verwen-dung von Teilräumen, die keinen Randbedingungen unterworfen sind. Abh. Math. Sem. Univ. Hamburg 36 (1971) 9-15.
[30] J.T. Oden, I. Babuška and C.E. Baumann, A discontinuous hp finite element method for diffusion problems. J. Comput. Phys. 146 (1998) 491-519.
[31] S. Prudhomme, F. Pascal and J.T. Oden, Review of error estimation for discontinuous Galerkin method. TICAM-report 00-27 (2000).
[32] B. Rivière, Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation. Frontiers in Mathematics 35. SIAM 2008. ISBN: 978-0-898716-56-6.
[33] B. Rivière and M.F. Wheeler, A discontinuous Galerkin method applied to nonlinear parabolic equations. The Center for Substance Modeling, TICAM, The University of Texas, Austin TX 78712, USA.
[34] 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. 39 (2001) 902-931.
[35] V. Thomée, Galerkin finite element methods for parabolic problems. Springer (1997).
[36] S. Volkwein, Non-linear conjugate gradient method for the optimal control of laser surface hardening, Optim. Methods Softw. 19 (2004) 179-199.
[37] M.F. Wheeler, An elliptic collocation-finite element method with interior penalties. SIAM J. Numer. Anal. 15 (1978) $152-161$.


[^0]:    Keywords and phrases. Laser surface hardening of steel, semi-linear parabolic equation, optimal control, error estimates, discontinuous Galerkin finite element method.
    ${ }^{1}$ Department of Mathematics, Indian Institute of Technology Bombay, Powai, 400076 Mumbai, India. nupur@math.iitb.ac.in; neela@math.iitb.ac.in

