A priori error estimates for the optimal control of laser surface hardening of steel

Gupta Nupur and Nataraj Neela and Pani Amiya K.

Original Citation:

Nupur, Gupta and Neela, Nataraj and Amiya K., Pani

(2011)

A priori error estimates for the optimal control of laser surface hardening of steel.

(Submitted)

A Priori Error Estimates for the Optimal Control of Laser Surface Hardening of Steel

Nupur Gupta* Neela Nataraj† A.K.Pani‡

October 26, 2010

Abstract

In this paper, we discuss a finite element method 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 in austenite and a semi-linear parabolic differential equation in temperature. The space discretization of the state variable is done using usual conforming finite elements, whereas the time discretization and control discretization are based on a discontinuous Galerkin method. *A priori* error estimates are developed and numerical experiments which justify the theoretical estimates are presented.

Key words. Laser surface hardening of steel, semi-linear parabolic equation, constrained optimal control, regularised problem, *a priori* error estimates, finite element method, discontinuous Galerkin in time, numerical experiments.

1 Introduction

In this paper, we develop *a priori* error estimates for the optimal control problem describing the laser surface hardening of steel. The purpose of surface hardening is to increase the hardness of the boundary layer of a workpiece by rapid heating and subsequent quenching (see Figure 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 [2] and [3]. For an extensive survey on mathematical models for laser material treatments, we refer to [4]. In [3], [5], the mathematical model for the laser hardening problem which gives rise to a system consisting of a nonlinear parabolic equations and a set of ordinary differential equations with non-differentiable right hand side is discussed. Then, the authors have first regularised the non-differentiable right hand side functions and results on existence, regularity and stability are derived for the regularised problem. This seems to be a common approach in all subsequent results not only on existence, but also on numerical approximations. In [6], both laser and induction hardening have been used to explain the model and then a finite volume method has been used for the spatial discretization and finite difference scheme for temporal discretization of the regularised problem. In [7], the optimal

*nupur@math.iitb.ac.in, +91 22 2576 7451, Department of Mathematics, Industrial Mathematics Group, Indian Institute of Technology Bombay, Powai, Mumbai- 400076, India

†neela@math.iitb.ac.in, +91 22 2576 7468, Department of Mathematics, Industrial Mathematics Group, Indian Institute of Technology Bombay, Powai, Mumbai- 400076, India

‡akp@math.iitb.ac.in, +91 22 2576 7481, Department of Mathematics, Industrial Mathematics Group, Indian Institute of Technology Bombay, Powai, Mumbai- 400076, India

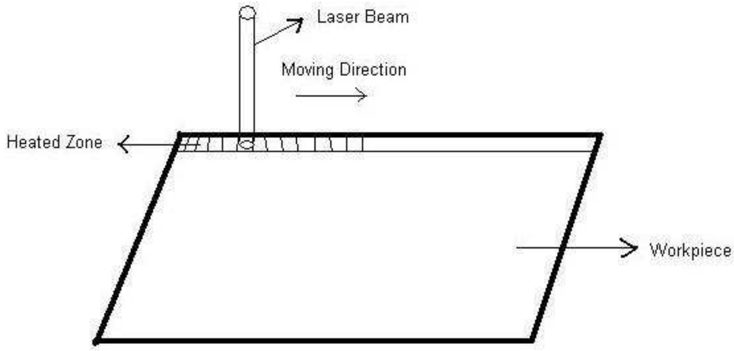


Figure 1: Laser Hardening Process

control problem is analyzed and error estimates for a regularised problem are derived using proper orthogonal decomposition (POD) Galerkin method. Moreover, some results on numerical simulations are also presented. In [8], a finite element scheme combined with a nonlinear conjugate gradient method is applied to approximate the solution of the regularised problem. Although, some of the articles [3]-[6], [8], [9] mentioned above have discussed different numerical methods for approximating the regularised laser surface hardening problem, *a priori* error estimates have not been developed. In the present paper, *a priori* error estimates have been developed for both semi-discrete and complete discrete problem of laser surface hardening of steel. The space discretization of the state variable is achieved by using usual conforming finite elements, whereas discretizations of temporal and control variables are based on a discontinuous Galerkin method. Initially, keeping control fixed, *a priori* error bounds are developed for the state variables, for both semi-discrete and fully discrete schemes. Finally, the convergence of the approximate control to exact control is established. A variant of the non-linear conjugate method [8] is applied to the optimal control problem numerically. The numerical results obtained are in good agreement with the findings of the theoretical results.

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 [10], [11], [14], [15] to mention a few. For optimal control problems governed by linear parabolic equations without control constraints, *a priori* error bounds are developed in [17]. Subsequently, using space-time finite element discretization, optimal parabolic control problems with control constraints are discussed in [18].

The outline of this paper is as follows. Section 2 describes the mathematical model of the problem of laser surface hardening of steel and its regularization using a regularized Heaviside function. In Section 3, a weak formulation is presented and results on existence and uniqueness of the solution of the regularized problem are discussed. Section 4 contains space-time discrete formulation of the laser surface hardening of steel with *a priori* error estimates at different levels of discretization. Section 5 describes the complete discretization with results of convergence for the control. Finally, numerical results are presented in Section 6.

2 Laser Surface Hardening of Steel

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 [2], the evolution of volume fraction of austenite $a(t)$ for a given temperature evolution

$\theta(t)$ is described by the following initial value problem:

$$\begin{aligned} \partial_t a = f_+(\theta, a) &= \frac{1}{\tau(\theta)}[a_{eq}(\theta) - a]_+ \quad \text{in } Q, && (2.1)\\ a(0) &= 0 \quad \text{in } \Omega, && (2.2) \end{aligned}$$

where $a_{eq}(\theta(t))$, denoted as $a_{eq}(\theta)$ for notational convenience, is the equilibrium volume fraction of austenite and τ is a time constant. The term

$$[a_{eq}(\theta) - a]_+ = (a_{eq}(\theta) - a)\mathcal{H}(a_{eq}(\theta) - a),$$

where $\mathcal{H}$ is the Heaviside function

$$\mathcal{H}(s) = \begin{cases} 1 & s > 1 \\ 0 & s \le 0, \end{cases}$$

denotes the non-negative part of $a_{eq}(\theta) - a$, that is,

$$[a_{eq}(\theta) - a]_+ = \frac{(a_{eq}(\theta) - a) + |a_{eq}(\theta) - a|}{2}.$$

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

$$\begin{aligned} \rho c_p \partial_t \theta - K \triangle \theta &= -\rho L a_t + \alpha u \quad \text{in } Q, && (2.3)\\ \theta(0) &= \theta_0 \quad \text{in } \Omega, && (2.4)\\ \frac{\partial \theta}{\partial n} &= 0 \quad \text{on } \Sigma, && (2.5) \end{aligned}$$

where the density ρ, the heat capacity c_p, the thermal conductivity 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, θ_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 minimize the cost functional defined by:

$$J(\theta, a, u) = \frac{\beta_1}{2}\int_\Omega |a(T) - a_d|^2 dx + \frac{\beta_2}{2}\int_0^T \int_\Omega [\theta - \theta_m]_+^2 dx ds + \frac{\beta_3}{2}\int_0^T |u - u_d|^2 ds \qquad (2.6)$$

subject to the state equations (2.1) – (2.5) in the set of admissible controls U_{ad}, (2.7)

where $U_{ad} = \{u \in U : \|u\|_{L^2(I)} \le M\}$, with $M > 0$, is the closed, bounded and convex subset of $U = L^2(I)$, denoting the maximal intensity of the laser, β_1, β_2 and β_3 being positive constants, $u_d \in L^\infty(I)$ being the desired laser intensity 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 below the melting temperature θ_m.

For theoretical, as well as computational reasons, the term $[a_{eq} - a]_+$ in (2.1) is regularized (see Figure 2) and the regularized laser surface hardening problem is now given by:

$$\min_{u_\epsilon \in U_{ad}} J(\theta_\epsilon, a_\epsilon, u_\epsilon) \text{ subject to} \tag{2.8}$$

$$\partial_t a_\epsilon = f_\epsilon(\theta_\epsilon, a_\epsilon) = \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)\mathcal{H}_\epsilon(a_{eq}(\theta_\epsilon) - a_\epsilon) \quad \text{in } Q, \tag{2.9}$$

$$a_\epsilon(0) = 0 \quad \text{in } \Omega, \tag{2.10}$$

$$\rho c_p \partial_t \theta_\epsilon - K \triangle \theta_\epsilon = -\rho L \partial_t a_\epsilon + \alpha u_\epsilon \quad \text{in } Q, \tag{2.11}$$

$$\theta_\epsilon(0) = \theta_0 \quad \text{in } \Omega, \tag{2.12}$$

$$\frac{\partial \theta_\epsilon}{\partial n} = 0 \quad \text{on } \Sigma, \tag{2.13}$$

where

$$J(\theta_\epsilon, a_\epsilon, u_\epsilon) = \frac{\beta_1}{2}\int_\Omega |a_\epsilon(T) - a_d|^2\,dx + \frac{\beta_2}{2}\int_0^T\int_\Omega [\theta_\epsilon - \theta_m]_+^2\,dx\,ds + \frac{\beta_3}{2}\int_0^T |u_\epsilon - u_d|^2\,ds,$$

and $\mathcal{H}_\epsilon \in C^{1,1}(\mathbb{R})$ is a monotone approximation of the Heaviside function satisfying $\mathcal{H}_\epsilon(s) = 0$ for $s \leq 0$.

We now make the following assumptions on the coefficients [7]:

(A1) $a_{eq}(x) \in (0,1)$ for all $x \in \mathbb{R}$ and $\|a_{eq}\|_{C^1(\mathbb{R})} \leq c_a$;

(A2) $0 < \underline{\tau} \leq \tau(x) \leq \bar{\tau}$ 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 Ω, 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 Ω.

For the sake of notational simplicity $(\theta_\epsilon, a_\epsilon, u_\epsilon)$ and f_ϵ will be replaced by (θ, a, u) and f respectively, throughout the paper.

3 Weak Formulation

Let $X = \{v \in L^2(I;V) : v_t \in L^2(I;V^*)\}$ and $Y = H^1(I;L^2(\Omega))$, where $V = H^1(\Omega)$. Together with $H = L^2(\Omega)$, the Hilbert space V and its dual V^* build a Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$. The duality pairing between V and V^* is denoted by $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{V^*\times V}$. Let $(\cdot,\cdot)$(resp. $(\cdot,\cdot)_{I,\Omega}$) and $\|\cdot\|$(resp. $\|\cdot\|_{I,\Omega}$) denote the inner product and norm in $L^2(\Omega)$(resp. $L^2(I, L^2(\Omega))$). 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 (2.9)-(2.13) takes the following form:

$$\min_{u \in U_{ad}} J(\theta, a, u) \text{ subject to} \tag{3.1}$$

$$(\partial_t a, w) = (f(\theta, a), w) \quad \forall w \in H, \text{ a.e. in } I, \tag{3.2}$$

$$a(0) = 0, \tag{3.3}$$

$$\rho c_p(\partial_t \theta, v) + K(\nabla\theta, \nabla v) = -\rho L(\partial_t a, v) + (\alpha u, v) \quad \forall v \in V, \text{ a.e. in } I, \tag{3.4}$$

$$\theta(0) = \theta_0. \tag{3.5}$$

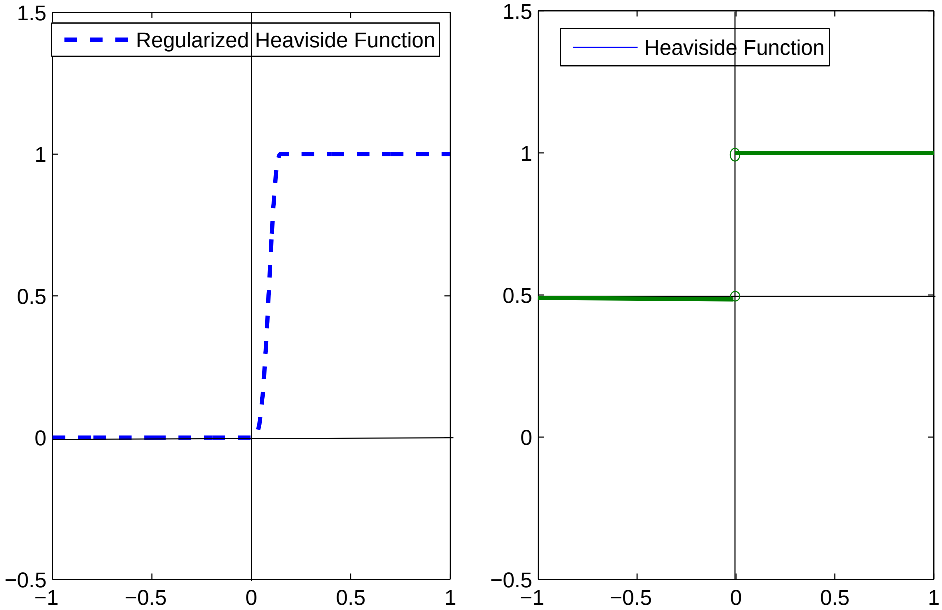


Figure 2: Heaviside ($\mathcal{H}(s)$) and Regularized Heaviside ($\mathcal{H}_\epsilon(s)$) functions

The following theorem [[8], Theorem 2.1] ensures the existence of a unique solution of the system (3.2)-(3.5).

Theorem 3.1. *Suppose that (A1)-(A6) are satisfied. Then, the system* (3.2)-(3.5) *has a unique solution*

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

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

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

Remark 3.1. *[8] Due to (A1)-(A2) and the definition of the regularized Heaviside function $\mathcal{H}_\epsilon$, there exists a constant $c_f > 0$ independent of θ and a such that*

$$\max(\|f(\theta, a)\|_{L^\infty(Q)}, \|f_a(\theta, a)\|_{L^\infty(Q)}, \|f_\theta(\theta, a)\|_{L^\infty(Q)}) \leq c_f$$

for all $(\theta, a) \in L^2(Q) \times L^\infty(Q)$.

The existence of the optimal control is guaranteed by the following theorem [[8], Theorem 2.3].

Theorem 3.2. *Suppose that (A1)-(A6) hold true. Then the optimal control problem* (3.1)-(3.5) *has at least one (global) solution.*

Let $u^* \in U_{ad}$ be a solution of (3.1)-(3.5) and (θ^*, a^*) be the solution of the corresponding state system. In the following lemma, we state the existence and uniqueness result of the corresponding adjoint system.

Lemma 3.1. *[7] Let (A1)-(A6) hold true and* $(\theta^*, a^*, u^*) \in X \times Y \times U_{ad}$ *be a solution to* (3.1)-(3.5). *Then there exists a unique solution* $(z^*, \lambda^*) \in H^{1,1} \times H^1(I, L^2(\Omega))$ *of the corresponding adjoint system defined by:*

$$\begin{aligned} -(\psi, \partial_t \lambda^*) + (\psi, f_a(\theta^*, a^*)(\rho L z^* - \lambda^*)) &= 0 \quad \forall \psi \in H, \ a.e. \ in \ I, & (3.6)\\ \lambda^*(T) &= \beta_1(a^*(T) - a_d), & (3.7)\\ -\rho c_p(\phi, \partial_t z^*) + K(\triangledown \phi, \triangledown z^*) + (\phi, f_\theta(\theta^*, a^*)(\rho L z^* - \lambda^*)) &= \beta_2(\phi, [\theta^* - \theta_m]_+) & (3.8)\\ & \quad \forall \phi \in V, \ a.e. \ in \ I, \\ z^*(T) &= 0. & (3.9)\end{aligned}$$

Moreover, z^* *satisfies the following variational inequality*

$$\left(\beta_3(u^* - u_d) + \int_\Omega \alpha z^* dx, \ p - u^* \right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \tag{3.10}$$

The existence of a unique solution to the state equation (3.2)-(3.5) ensures the existence of a control-to-state mapping $u \longmapsto (\theta, a) = (\theta(u), a(u))$ through (3.2)-(3.5). By means of this mapping, we introduce the reduced cost functional $j : U_{ad} \longrightarrow \mathbb{R}$ as

$$j(u) = J(\theta(u), a(u), u). \tag{3.11}$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u \in U_{ad}} j(u). \tag{3.12}$$

The first order necessary optimality condition for (3.12) reads as

$$j'(u^*)(p - u^*) \geq 0 \ \ \forall p \in U_{ad}, \tag{3.13}$$

where $j'(u)(p-u) = \left(\beta_3(u - u_d) + \int_\Omega \alpha z(u) dx, \ p - u \right)_{L^2(I)}$.

Remark 3.2. *The constant C will be used to denote different values at different steps of the proof throughout the paper in all results and is a generic one.*

We now discuss a regularity result for θ.

Lemma 3.2. *Under the assumptions (A1)-(A5), the solution* (θ, a) *of* (3.2)-(3.5) *satisfies:*

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

where $C > 0$ *is a constant.*

Proof. From (3.4), we have

$$\rho c_p(\partial_t \theta, v) - K(\Delta \theta, v) = -\rho L(\partial_t a, v) + (\alpha u, v).$$

Putting $v = -\Delta\theta$ and using Cauchy Schwarz inequality, we obtain

$$K\|\Delta \theta\| \leq C\Big(\|\partial_t a\| + |u| + \|\partial_t \theta\| \Big),$$

where $C = \max\{\rho c_p,\ \rho L,\ \max_Q |\alpha(x,t)|\}$.

Squaring and integrating from 0 to T, we find that

$$\|\Delta\theta\|_{I,\Omega}^2 \le C(\|\partial_t a\|_{I;\Omega}^2 + \|u\|_{L^2(I)}^2 + \|\partial_t\theta\|_{I;\Omega}^2).$$

Using $(\theta, a) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$ ([7], Theorem 2.1) we obtain the required result and this completes the rest of the proof. □

4 Semidiscrete Scheme

In this section, we discuss a semi-discrete Galerkin method with piece-wise linear polynomials for the problem (3.1)-(3.5) and establish *a priori* error estimates for the semi-discrete solution with a fixed control variable $u \in U_{ad}$. Moreover, similar analysis is also developed for the semi-discrete approximation of the adjoint problem (3.6)-(3.9).

Let $\mathcal{T}_h$ be an admissible regular triangulation of $\bar{\Omega}$ into simplexes R. Let the discretization parameter h be defined as $h = \max\limits_{R\in\mathcal{T}_h} h_R$, where h_R is the diameter of R. Further, let $V_h \subset V$ be a finite element space defined by $V_h = \{v \in C^0(\bar{\Omega}) : \quad v|_R \in P_1(R) \quad \forall R \in \mathcal{T}_h\}$ and $X_h = L^2(I, V_h)$. Here $P_1(R)$ denotes the set of all polynomials of degree ≤ 1. Then the semi-discrete formulation corresponding to the continuous problem (3.1)-(3.5) reads as

$$\min_{u\in U_{ad}} J(\theta_h, a_h, u) \text{ subject to} \tag{4.1}$$

$$(\partial_t a_h, w) = (f(\theta_h, a_h), w) \qquad \forall w \in V_h,\ \text{a.e. in } I, \tag{4.2}$$

$$a_h(0) = 0, \tag{4.3}$$

$$\rho c_p(\partial_t\theta_h, v) + K(\bigtriangledown\theta_h, \bigtriangledown v) = -\rho L(\partial_t a_h, v) + (\alpha u, v) \qquad \forall v \in V_h,\ \text{a.e. in } I, \tag{4.4}$$

$$\theta_h(0) = \theta_{h,0}, \tag{4.5}$$

where $\theta_{h,0}$ is a suitable approximation of θ_0 to be chosen later. Corresponding to the solution $\tilde{u}^* \in U_{ad}$ of (4.1)-(4.5), let (θ_h^*, a_h^*) be the solution to the state system (4.2)-(4.4). The first order optimality conditions yield the adjoint problem: Find $(z_h^*(t), \lambda_h^*(t)) \in V_h \times V_h, \quad t \in \bar{I}$ such that

$$-(\psi, \partial_t\lambda_h^*) + (\psi, f_a(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)) = 0 \quad \forall\psi \in V_h,\ \text{a.e. in } I, \tag{4.6}$$

$$\lambda_h^*(T) = \beta_1(a_h^*(T) - a_d), \tag{4.7}$$

$$-\rho c_p(\phi, \partial_t z_h^*) + K(\bigtriangledown\phi, \bigtriangledown z_h^*) + (\phi, f_\theta(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)) = \beta_2(\phi, [\theta_h^* - \theta_m]_+) \quad \forall\phi \in V_h,\ \text{a.e. in } I, \tag{4.8}$$

$$z_h^*(T) = 0. \tag{4.9}$$

Moreover, z_h^* satisfies the following variational inequality

$$\left(\beta_3(\tilde{u}^* - u_d) + \int_\Omega \alpha z_h^* dx,\ p - \tilde{u}^*\right)_{L^2(I)} \ge 0 \quad \forall p \in U_{ad}. \tag{4.10}$$

Now we consider the reduced cost functional $j_h : U_{ad} \longrightarrow \mathbb{R}$:

$$j_h(u) = J(\theta_h(u), a_h(u), u). \tag{4.11}$$

Then the semi-discrete optimal control problem can be equivalently formulated as

$$\min_{u \in U_{ad}} j_h(u). \tag{4.12}$$

The first order necessary optimality condition for (4.12) reads as

$$j_h'(\tilde{u}^*)(p - \tilde{u}^*) \geq 0 \quad \forall p \in U_{ad}, \tag{4.13}$$

where $j_h'(u)(p-u) = \left(\beta_3(u - u_d) + \int_\Omega \alpha z_h(u) dx, p - u\right)_{L^2(I)}$.

Now define the elliptic projection $\mathcal{R}_h : V \longrightarrow V_h$ by

$$K(\bigtriangledown(v - \mathcal{R}_h v), \bigtriangledown\phi) + \gamma(v - \mathcal{R}_h v, \phi) \;=\; 0 \quad \forall \phi \in V_h, \tag{4.14}$$

where γ is a positive constant.

Lemma 4.1 ([20], p. 737)**.** *For $v \in H^2(\Omega)$, there exists a positive constant C such that:*

$$\|v - \mathcal{R}_h v\| \;\leq\; Ch^2 \|v\|_{H^2(\Omega)}. \quad \square$$

We also define the L^2-projection $P_h : L^2(\Omega) \times H^2(\Omega) \longrightarrow V_h$, such that

$$(P_h v - v, w) = 0 \quad \forall w \in V_h.$$

Note that P_h satisfies the following error estimates:

$$\|v - P_h v\| \;\leq\; Ch^2 \|v\|_{H^2(\Omega)} \quad \forall v \in H^2(\Omega). \tag{4.15}$$

Theorem 4.1. *Let (θ, a) and (θ_h, a_h) be the solutions of* (3.2)-(3.5) *and* (4.2)-(4.5)*, respectively. Then, under the extra regularity assumptions that, $(\theta, a) \in L^\infty(I, H^2(\Omega)) \times L^\infty(I, H^2(\Omega))$, $\partial_t\theta \in L^2(I, H^2(\Omega))$ and $\theta_0 \in H^2(\Omega)$; for fixed $u \in U_{ad}$, there exists a positive constant C independent of h such that*

$$\begin{aligned} \|\theta(t) - \theta_h(t)\| + \|a(t) - a_h(t)\| \;\leq\; & Ch^2\Big(\|\theta_0\|_{H^2(\Omega)} + \|\theta\|_{L^\infty(I,H^2(\Omega))} + \|a\|_{L^\infty(I,H^2(\Omega))} \\ & + \|\partial_t\theta\|_{L^2(I,H^2(\Omega))}\Big) \quad \forall t \in I. \end{aligned}$$

Proof. Let $\zeta^\theta = \theta - \mathcal{R}_h\theta$ and $\eta^\theta = \mathcal{R}_h\theta - \theta_h$. Subtract (4.4) from (3.4), use (3.2), (4.2) and (4.14) to obtain

$$\rho c_p(\partial_t \eta^\theta, v) + K(\bigtriangledown\eta^\theta, \bigtriangledown v) \;=\; -\rho L(f(\theta, a) - f(\theta_h, a_h), v) - \rho c_p(\partial_t\zeta^\theta, v) + \gamma(\zeta^\theta, v),$$

where $v \in V_h$.Choose $v = \eta^\theta$. Then integrate from 0 to t, apply Cauchy Schwarz and Young's inequality to obtain

$$\|\eta^\theta(t)\|^2 \;\leq\; C\left(\|\eta^\theta(0)\|^2 + \int_0^t \left(\|f(\theta, a) - f(\theta_h, a_h)\|^2 + \|\eta^\theta\|^2 + \|\zeta^\theta\|^2 + \|\partial_t\zeta^\theta\|^2\right) ds\right). \tag{4.16}$$

By choosing $\theta_{h,0}$ as the L^2 approximation of the function $\theta_0 \in H^2(\Omega)$, we obtain

$$\|\eta^\theta(0)\|^2 \leq \|\mathcal{R}_h\theta_0 - \theta_0\|^2 + \|\theta_0 - \theta_{h,0}\|^2 \leq Ch^4\|\theta_0\|^2_{H^2(\Omega)}. \tag{4.17}$$

Since f is Lipschitz in both the arguments (see Remark 3.1), we find using (4.17) in (4.16) that

$$\begin{aligned}\|\eta^\theta(t)\|^2 &\leq C\Big(h^4\|\theta_0\|^2_{H^2(\Omega)} + \int_0^t \Big(\|\theta-\theta_h\|^2 + \|a-a_h\|^2 + \|\eta^\theta\|^2 + \|\zeta^\theta\|^2 + \|\partial_t\zeta^\theta\|^2\Big)ds\Big)\\ &\leq C\Big(h^4\|\theta_0\|^2_{H^2(\Omega)} + \int_0^t (\|\zeta^\theta\|^2 + \|\zeta^a\|^2 + \|\partial_t\zeta^\theta\|^2)ds + \int_0^t (\|\eta^\theta\|^2 + \|\eta^a\|^2)ds\Big), \quad (4.18)\end{aligned}$$

where $\zeta^a = a - P_h a$, $\eta^a = P_h a - a_h$. Now subtracting (4.2) from (3.2) for fixed $t \in I$, integrating from 0 to t, using Cauchy Schwarz, Young's inequality and the fact that $(\partial_t\zeta^a, \eta) = 0$, we obtain

$$\|\eta^a(t)\|^2 \leq C\Big(\int_0^t (\|\zeta^\theta\|^2 + \|\zeta^a\|^2)ds + \int_0^t (\|\eta^\theta\|^2 + \|\eta^a\|^2)ds\Big). \tag{4.19}$$

Adding (4.18) and (4.19), we arrive at

$$\begin{aligned}\|\eta^\theta(t)\|^2 + \|\eta^a(t)\|^2 &\leq C\Big(h^4\|\theta_0\|^2_{H^2(\Omega)} + \int_0^T (\|\zeta^\theta\|^2 + \|\zeta^a\|^2 + \|\partial_t\zeta^\theta\|^2)\, ds\Big)\\ &+ C\int_0^t (\|\eta^\theta\|^2 + \|\eta^a\|^2)\, ds.\end{aligned}$$

Using Gronwall's lemma, Lemma 4.1 and (4.15), we obtain

$$\|\eta^\theta(t)\|^2 + \|\eta^a(t)\|^2 \leq Ch^4\Big(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta\|^2_{L^2(I,H^2(\Omega))} + \|a\|^2_{L^2(I,H^2(\Omega))} + \|\partial_t\theta\|^2_{L^2(I,H^2(\Omega))}\Big).$$

A use of triangle inequality with Lemma 4.1 and (4.15) yields the required result and this completes the rest of the proof. □

Remark 4.1. *Although, the finite element space used in this article to discretize the variables θ and a is X_h, where approximation is done using continuous functions, the variable a can also be approximated using piecewise constants with appropriate changes in the proof.*

Below, we discuss error estimates for the adjoint problem.

Theorem 4.2. *Let (z^*, λ^*) and (z_h^*, λ_h^*) be the solutions of (3.6)-(3.9) and (4.6)-(4.9) corresponding to the state solutions (θ^*, a^*) and (θ_h^*, a_h^*), respectively. Then, under the extra regularity assumptions made in Theorem 4.1 and $(z^*, \lambda^*) \in L^\infty(I, H^2(\Omega)) \times L^\infty(I, H^2(\Omega))$, $\partial_t z^* \in L^2(I, H^2(\Omega))$, $a_d \in H^2(\Omega)$; for $t \in I$, there exists a positive constant C independent of h such that*

$$\begin{aligned}\|z^*(t) - z_h^*(t)\| &+ \|\lambda^*(t) - \lambda_h^*(t)\|\\ &\leq Ch^2\Big(\|\theta_0\|_{H^2(\Omega)} + \|\theta^*\|_{L^\infty(I,H^2(\Omega))} + \|a^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|_{L^2(I,H^2(\Omega))}\\ &\quad + \|a_d\|_{H^2(\Omega)} + \|z^*\|_{L^\infty(I,H^2(\Omega))} + \|\lambda^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t z^*\|_{L^2(I,H^2(\Omega))}\Big).\end{aligned}$$

Proof. Write $\lambda^* - \lambda_h^* = (\lambda^* - P_h\lambda^*) + (P_h\lambda^* - \lambda_h^*) = \zeta^\lambda + \eta^\lambda$ and $z^* - z_h^* = (z^* - \mathcal{R}_h z^*) + (\mathcal{R}_h z^* - z_h^*) = \zeta^z + \eta^z$. Subtract (4.8) from (3.8) and use (4.14) to obtain

$$\begin{aligned} -\rho c_p(\phi, \partial_t \eta^z) \quad &+ \quad K(\bigtriangledown\phi, \bigtriangledown\eta^z) = \rho c_p(\phi, \partial_t\zeta^z) + \gamma(\phi, \zeta^z) \\ &- \quad (\phi, f_\theta(\theta^*, a^*)(\rho L z^* - \lambda^*) - f_\theta(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)) + (\phi, [\theta^* - \theta_m]_+ - [\theta_h^* - \theta_m]_+) \end{aligned} \tag{4.20}$$

Choose $\phi = \eta^z$ in (4.20), integrate from t to T, apply Cauchy Schwarz inequality, Young's inequality and use Theorem 4.1, to obtain

$$\begin{aligned} \|\eta^z(t)\|^2 \quad \leq \quad & C\Bigg(h^4(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta^*\|^2_{L^\infty(I,H^2(\Omega))} + \|a^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|^2_{L^2(I,H^2(\Omega))}) \\ & + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 + \|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2) ds\Bigg). \end{aligned} \tag{4.21}$$

Subtract (4.6) from (3.6) and choose $\chi = \eta^\lambda$. Then integrate from t to T and apply Cauchy Schwarz with Young's inequality, to arrive at

$$\|\eta^\lambda(t)\|^2 \quad \leq \quad C\Bigg(\|P_h a_d - a_d\|^2 + \|P_h a^*(T) - a_h^*(T)\|^2 + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 + \|\zeta^\lambda\|^2 + \|\zeta^\lambda\|^2) ds\Bigg).$$

Using Lemma 4.1 and Theorem 4.1, we obtain

$$\begin{aligned} \|\eta^\lambda(t)\|^2 \quad \leq \quad & C\Bigg(h^4(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta^*\|^2_{L^\infty(I,H^2(\Omega))} + \|a^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|^2_{L^2(I,H^2(\Omega))} + \|a_d\|^2_{H^2(\Omega)}) \\ & + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 + \|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2) ds\Bigg). \end{aligned} \tag{4.22}$$

Adding (4.21) and (4.22), we find that

$$\begin{aligned} \|\eta^z(t)\|^2 \quad + \quad & \|\eta^\lambda(t)\|^2 \leq Ch^4\Bigg(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta^*\|^2_{L^\infty(I,H^2(\Omega))} + \|a^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|^2_{L^2(I,H^2(\Omega))} \\ & + \|a_d\|^2_{H^2(\Omega)}\Bigg) + C\int_0^T (\|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2)\, ds + C\int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2)\, ds. \end{aligned}$$

Using Gronwall's lemma, we obtain

$$\begin{aligned} \|\eta^z(t)\|^2 + \|\eta^\lambda(t)\|^2 \quad \leq \quad & Ch^4\Bigg(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta^*\|^2_{L^\infty(I,H^2(\Omega))} + \|a^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|^2_{L^2(I,H^2(\Omega))} \\ & + \|a_d\|^2_{H^2(\Omega)}\Bigg) + C\int_0^T (\|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2)\, ds. \end{aligned} \tag{4.23}$$

A use of Lemma 4.1, (4.15) and (4.23) yields the required result and this completes the rest of the proof. □

5 Completely Discrete Scheme

In this section, a temporal discretization is done using a discontinuous Galerkin finite element method with piecewise constant approximation and *a priori* error estimates are proved in Theorem 5.1 and 5.2. The control is being discretized using piecewise constants in each time interval $I_n, n =$

$1, 2, \cdots, N$. In Theorem 5.3, the convergence of discrete optimal control to an optimal control of (3.1)-(3.5) is established. In order to discretize (4.1)-(4.5) in time, we consider the following partition of I:

$$0 = t_0 < t_1 < .... < t_N = T.$$

Set $I_1 = [t_0, t_1]$ and $I_n = (t_{n-1}, t_n]$, $k_n = t_n - t_{n-1}$, for $n = 2, 3, ..., N$ and $k = \max_{1 \le n \le N} k_n$. We define

$$X_{hk}^q = \{\phi : I \to V_h;\ \ \phi|_{I_n} = \sum_{j=0}^{q-1} \psi_j t^j, \psi_j \in V_h\},\ \ q \in \mathbb{N}. \tag{5.1}$$

For a function v in X_{hk}^q, we use the following notations:

$$v_n = v(t_n),\ \ v_n^+ = \lim_{t \to t_n + 0} v(t) \text{ and } [v]_n = v_n^+ - v_n^-.$$

Then the dG(q)cG(1) discretization of (3.1)-(3.5) reads as:

$$\begin{aligned}
&\min_{u \in U_{ad}} J(\theta_{hk}, a_{hk}, u) \qquad \text{subject to} && (5.2)\\
&\sum_{n=1}^{N} (\partial_t a_{hk}, w)_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a_{hk}]_n, w_n^+) + (a_{hk,0}^+, w_0^+) = (f(\theta_{hk}, a_{hk}), w)_{I,\Omega}, && (5.3)\\
&a_{hk}(0) = 0, && (5.4)\\
&\rho c_p \sum_{n=1}^{N} (\partial_t \theta_{hk}, v)_{I_n,\Omega} + K(\nabla \theta_{hk}, \nabla v)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_{hk}]_m, v_m^+) + \rho c_p(\theta_{hk,0}^+, v_0^+)\\
&\quad = -\rho L(f(\theta_{hk}, a_{hk}, v))_{I,\Omega} + (\alpha u, v)_{I,\Omega} + \rho c_p(\theta_0, v_0^+), && (5.5)\\
&\theta_{hk}(0) = \theta_{h,0} && (5.6)
\end{aligned}$$

for all $(v, w) \in X_{hk}^q \times X_{hk}^q$.
Corresponding to the solution $\hat{u}^* \in U_{ad}$ of (5.2)-(5.6), let $(\theta_{hk}^*, a_{hk}^*)$ be the solution to the state system (5.3)-(5.6). The first order optimality conditions yield the adjoint problem:
Find $(z_{hk}^*, \lambda_{hk}^*) \in X_{hk}^q \times X_{hk}^q$ such that

$$\begin{aligned}
&-\sum_{n=1}^{N} (\psi, \partial_t \lambda_{hk}^*)_{I_n,\Omega} - \sum_{n=1}^{N-1} (\psi_n^-, [\lambda_{hk}^*]_n) - (\psi_N^-, \lambda_{hk,N}^{-,*})\\
&\quad + (\psi, f_a(\theta_{hk}^*, a_{hk}^*)(\rho L z_{hk}^* - \lambda_{hk}^*))_{I,\Omega} = -(\psi_n^-, \lambda_{hk}^*(T)), && (5.7)\\
&\lambda_{hk}^*(T) = \beta_1(a_{hk}^*(T) - a_d), && (5.8)\\
&-\rho c_p \sum_{n=1}^{N} (\phi, \partial_t z_{hk}^*)_{I_n,\Omega} + K(\nabla \phi, \nabla z_{hk}^*)_{I,\Omega} - \rho c_p \sum_{n=1}^{N-1} (\phi_m^-, [z_{hk}^*]_m) - \rho c_p(\phi_N^-, z_{hk,N}^{+,*})\\
&\quad + (\phi, f_\theta(\theta_{hk}^*, a_{hk}^*)(\rho L z_{hk}^* - \lambda_{hk}^*))_{I,\Omega} = \beta_2(\phi, [\theta_{hk}^* - \theta_m]_+)_{I,\Omega}, && (5.9)\\
&z_{hk}^*(T) = 0, && (5.10)
\end{aligned}$$

for all $(\psi, \phi) \in X_{hk}^q \times X_{hk}^q$. Moreover, z_{hk}^* satisfies the following variational inequality

$$\left(\beta_3(\hat{u}^* - u_d) + \int_\Omega \alpha z_{hk} dx,\ p - \hat{u}^*\right)_{L^2(I)} \ge 0 \quad \forall p \in U_{ad}. \tag{5.11}$$

We introduce the following space-time discrete reduced cost functional $j_{hk} : U_{ad} \longrightarrow \mathbb{R}$:

$$j_{hk}(u) = J(\theta_{hk}(u), a_{hk}(u), u). \tag{5.12}$$

Then the space-time discrete optimal control problem can be equivalently reformulated as

$$\min_{u \in U_{ad}} j_{hk}(u). \tag{5.13}$$

The first order necessary optimality condition for (5.13) reads as

$$j'_{hk}(\hat{u}^*)(p - \hat{u}^*) \geq 0 \quad \forall p \in U_{ad}. \tag{5.14}$$

We consider the case of piecewise constant approximation in time for both the state and adjoint formulation. For the case where $q = 1$ in the definition of X^q_{hk}, (5.3)-(5.6) can be rewritten as: for $n = 1, 2, \cdots, N$, find $(\theta^n_{hk}, a^n_{hk}) \in V_h \times V_h$ such that

$$\begin{aligned}
\Big(\frac{a^n_{hk} - a^{n-1}_{hk}}{k_n}, w\Big) &= \frac{1}{k_n}\Big(\int_{I_n} f(\theta^n_{hk}, a^n_{hk}) ds, w\Big), && (5.15)\\
a_{hk}(0) &= 0, && (5.16)\\
\rho c_p\Big(\frac{\theta^n_{hk} - \theta^{n-1}_{hk}}{k_n}, v\Big) + K(\nabla\theta^n_{hk}, \nabla v) &= -\rho L\Big(\frac{1}{k_n}\int_{I_n} f(\theta^n_{hk}, a^n_{hk}) ds, v\Big) \\
&\quad + \Big(\frac{1}{k_n}\int_{I_n} \alpha u ds, v\Big), && (5.17)\\
\theta_{hk}(0) &= \theta_{h,0}, && (5.18)
\end{aligned}$$

$\forall (w, v) \in V_h \times V_h$.

Before estimating the *a priori* error estimates for space-time discretization, we define the interpolant $\pi_k : C(\bar{I}, V_h) \longrightarrow X^1_{hk}$ as:

$$\pi_k v(t_n) = v(t_n) \qquad \forall n = 1, 2, \cdots, N, \tag{5.19}$$

where $C(\bar{I}, V_h)$ is the space of all continuous functions defined from $\bar{I}$ to V_h. Note that

$$\|v - \pi_k v\|_{I,\Omega} \leq Ck\|\partial_t v\|. \tag{5.20}$$

Theorem 5.1. *Let* $(\theta^m_{hk}, a^m_{hk})$ $\forall m = 1, 2, \cdots, N$ *and* (θ, a) *be the solutions of the problems* (5.15)-(5.18) *and* (3.2)-(3.5), *respectively. Then, under the extra regularity assumptions made in Theorem 4.1 and* $(\partial_{tt}\theta, \partial_{tt}a) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega))$, $\partial_t u \in L^2(I)$; *there exists a positive constant* C *independent of* h *and* k *such that*

$$\begin{aligned}
\|\theta(t_m) - \theta^m_{hk}\| &+ \|a(t_m) - a^m_{hk}\| \\
&\leq Ch^2\Big(\|\theta\|_{L^\infty(I,H^2(\Omega))} + \|a\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|_{L^2(I,H^2(\Omega))} + \|\theta_0\|_{H^2(\Omega)}\Big) \\
&\quad + Ck\Big(\|\partial_t u\|_{L^2(I)} + \|\partial_{tt}\theta\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}a\|_{L^\infty(I,L^2(\Omega))}\Big), \quad t_m \in \bar{I_m}.
\end{aligned}$$

Proof. Write $\theta(t_n) - \theta^n_{hk} = (\theta(t_n) - \mathcal{R}_h\theta(t_n)) + (\mathcal{R}_h\theta(t_n) - \theta^n_{hk}) = \zeta^{\theta,n} + \eta^{\theta,n}$ and denote $\frac{\theta^n_{hk} - \theta^{n-1}_{hk}}{k_n}$ by $\bar{\partial}\theta^n_{hk}$. Also, write $a(t_n) - a^n_{hk} = (a(t_n) - P_h a(t_n)) + (P_h a(t_n) - a^n_{hk}) = \zeta^{a,n} + \eta^{a,n}$ and denote $\frac{a^n_{hk} - a^{n-1}_{hk}}{k_n}$

by $\bar{\partial}a_{hk}^n$. Subtracting (5.17) from (3.4), we obtain at $t = t_n$

$$\rho c_p(\partial_t\theta(t_n) - \bar{\partial}\theta_{hk}^n, v) + K(\bigtriangledown(\theta(t_n) - \theta_{hk}^n), \bigtriangledown v) \;=\; -\rho L\Big(f(\theta(t_n), a(t_n)) - \frac{1}{k_n}\int_{I_n} f(\theta_{hk}^n, a_{hk}^n)ds, v\Big) + \Big(\alpha(x,t_n)u(t_n) - \frac{1}{k_n}\int_{I_n}\alpha u ds, v\Big),$$

where $v \in V_h$. Using (5.19) and (4.14), we find that

$$\begin{aligned}\rho c_p(\bar{\partial}\eta^{\theta,n}, v) \;&+\; K(\bigtriangledown(\eta^{\theta,n}, \bigtriangledown v) = -\rho c_p(\bar{\partial}\zeta^{\theta,n}, v) + \rho c_p(\bar{\partial}\theta(t_n) - \partial_t\theta(t_n), v) \\ &+\; \gamma(\zeta^{\theta,n}, v) - \rho L\Big(f(\theta(t_n), a(t_n)) - f(\theta_{hk}^n, a_{hk}^n), v\Big) \\ &+\; \Big(\frac{1}{k_n}\int_{I_n}(\pi_k(\alpha(x,t_n)u(t_n)) - \alpha u)\,ds, v\Big).\end{aligned} \tag{5.21}$$

Choose $v = \eta^{\theta,n}$ in (5.21) and Cauchy-Schwarz inequality to obtain

$$\begin{aligned}\rho c_p(\bar{\partial}\eta^{\theta,n}, \eta^{\theta,n}) \;&+\; K\|\bigtriangledown\eta^{\theta,n}\|^2 \le \rho\Big(L\|f(\theta_{hk}^n, a_{hk}^n) - f(\theta(t_n), a(t_n))\| + c_p\|\bar{\partial}\theta(t_n) - \partial_t\theta(t_n)\|\Big)\|\eta^{\theta,n}\| \\ &+\; \Big(\rho c_p\|\bar{\partial}\zeta^{\theta,n}\| + \gamma\|\zeta^{\theta,n}\| + \frac{1}{k_n^{1/2}}\|\pi_k(\alpha u) - \alpha u\|_{I_n,\Omega}\Big)\|\eta^{\theta,n}\|.\end{aligned} \tag{5.22}$$

Observe that

$$\begin{aligned}(\bar{\partial}\eta^{\theta,n}, \eta^{\theta,n}) \;&=\; \frac{1}{2k_n}\Big(\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2\Big) + \frac{1}{2k_n}\|\eta^{\theta,n} - \eta^{\theta,n-1}\|^2 \\ &\ge\; \frac{1}{2k_n}\Big(\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2\Big).\end{aligned} \tag{5.23}$$

Using (5.23) in (5.22), we find that

$$\begin{aligned}\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2 \;&\le\; Ck_n\Big(\|f(\theta_{hk}^n, a_{hk}^n) - f(\theta(t_n), a(t_n))\| + \|\bar{\partial}\theta(t_n) - \partial_t\theta(t_n)\| \\ &+\|\bar{\partial}\zeta^{\theta,n}\| + \|\zeta^{\theta,n}\|\Big)\|\eta^{\theta,n}\| + Ck_n^{1/2}\|\pi_k(\alpha u) - \alpha u\|_{I_n,\Omega}\,\|\eta^{\theta,n}\|.\end{aligned}$$

Using Young's inequality and the Remark 3.1, we obtain

$$\begin{aligned}\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2 \;&\le\; Ck_n\Big(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{\theta,n}\|^2 + \|\zeta^{a,n}\|^2 \\ &+\; \|\bar{\partial}\theta(t_n) - \partial_t\theta(t_n)\|^2 + \|\bar{\partial}\zeta^{\theta,n}\|^2 + k_n^{-1}\|u - \pi_k u\|_{I_n,\Omega}^2\Big) \\ &=\; Ck_n\Big(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{a,n}\|^2 + R_n^1\Big),\end{aligned} \tag{5.24}$$

where $R_n^1 = \|\zeta^{\theta,n}\|^2 + \|\bar{\partial}\theta(t_n) - \partial_t\theta(t_n)\|^2 + \|\bar{\partial}\zeta^{\theta,n}\|^2 + k_n^{-1}\|\alpha u - \pi_k(\alpha u)\|_{I_n,\Omega}^2$. Summing up (5.24) from $n = 1$ to m, we find that

$$\|\eta^{\theta,m}\|^2 \le \|\eta^{\theta,0}\|^2 + \sum_{n=1}^{m} Ck_n\Big(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{a,n}\|^2 + R_n^1\Big). \tag{5.25}$$

Similarly, we now consider

$$(\bar\partial\eta^{a,n}, w) \;=\; (f(\theta(t_n), a(t_n)) - f(\theta_{hk}^n, a_{hk}^n), w) - (\bar\partial a(t_n) - \partial_t a(t_n), w) - (\bar\partial\zeta^{a,n}, w),$$

where $w \in V_h$. Putting $w = \eta^{a,n}$ and proceeding as in (5.22)-(5.23) we find using Remark 3.1 that

$$\|\eta^{a,n}\|^2 - \|\eta^{a,n-1}\|^2 \le Ck_n\Big(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{\theta,n}\|^2 + R_n^2\Big), \tag{5.26}$$

where $R_n^2 = \|\zeta^{a,n}\|^2 + \|\bar\partial a(t_n) - \partial_t a(t_n)\|^2$. Summing up (5.26) from $n = 1$ to m, we arrive at

$$\|\eta^{a,m}\|^2 \le \|\eta^{a,0}\|^2 + \sum_{n=1}^{m} Ck_n\Big(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{\theta,n}\|^2 + R_n^2\Big). \tag{5.27}$$

Now adding (5.25) and (5.27) we obtain

$$\begin{aligned}\Big(\|\eta^{\theta,m}\|^2 + \|\eta^{a,m}\|^2\Big) \;&\le\; \|\eta^{\theta,0}\|^2 + \|\eta^{a,0}\|^2 + \sum_{n=1}^{m} Ck_n(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2)\\ &+\; C\sum_{n=1}^{m} k_n(R_n^1 + R_n^2).\end{aligned} \tag{5.28}$$

In order to estimate the terms in R_n^1, we use Lemma 4.1 to obtain

$$\begin{aligned}\|\bar\partial\zeta^{\theta,n}\|^2 \;&=\; \|k_n^{-1}\int_{t_{n-1}}^{t_n} \partial_t\zeta^\theta ds\|^2\\ &\le\; k_n^{-1}\int_{t_{n-1}}^{t_n} \|\partial_t\zeta^\theta\|^2\, ds \le Ck_n^{-1}h^4\int_{t_{n-1}}^{t_n} \|\partial_t\theta(s)\|^2_{H^2(\Omega)}\, ds,\end{aligned} \tag{5.29}$$

Using interpolation error, we find that

$$\|\alpha u - \pi_k(\alpha u)\|^2_{L^2(I_n)} \le Ck_n^2\int_{t_{n-1}}^{t_n} (|\partial_t u|^2 + \|\alpha_t\|^2_{L^\infty(\Omega)})\, ds. \tag{5.30}$$

A use of Taylor's expansion yields

$$\begin{aligned}\|\bar\partial\theta(t_n) - \partial_t\theta(t_n)\|^2 \;&=\; \|k_n^{-1}\int_{t_{n-1}}^{t_n} (s - t_{n-1})\partial_{tt}\theta\, ds\|^2 \le k_n^{-2}\Big(\int_{t_{n-1}}^{t_n} (s - t_{n-1})\|\partial_{tt}\theta\|\, ds\Big)^2\\ &\le\; \frac{1}{3}k_n\int_{t_{n-1}}^{t_n} \|\partial_{tt}\theta\|^2\, ds.\end{aligned} \tag{5.31}$$

A use of Lemma 4.1 with (5.20),(5.29)-(5.31) implies that

$$\begin{aligned}\sum_{n=1}^{m} k_n R_n^1 \;&\leq\; Ch^4\Big(\|\theta\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|^2_{L^2(I;H^2(\Omega))}\Big) \\ &+\; C\sum_{n=1}^{m} k_n^2\Big(|\partial_t u|^2_{L^2(I_n)} + \|\alpha_t\|^2_{L^2(I_n;L^\infty(\Omega)} + \|\partial_{tt}\theta\|^2_{L^2(I_n,L^2(\Omega))}\Big).\end{aligned} \tag{5.32}$$

Similarly, we estimate R_n^2 and write is as

$$\sum_{n=1}^{m} k_n R_n^2 \;\leq\; C\Big(h^4\|a\|^2_{L^\infty(I,H^2(\Omega))} + \sum_{n=1}^{m} k_n^2\|\partial_{tt}a\|^2_{L^2(I_n,L^2(\Omega))}\Big). \tag{5.33}$$

Note that $\|\eta^{\theta,0}\| \leq Ch^2\|\theta_0\|_{H^2(\Omega)}$. On substituting (5.32)-(5.33) in (5.28), we obtain

$$\begin{aligned}\|\eta^{\theta,m}\|^2 \;&+\; \|\eta^{a,m}\|^2 \leq Ch^4\Big(\|\theta\|^2_{L^\infty(I,H^2(\Omega))} + \|a\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|^2_{L^2(I;H^2(\Omega))}\Big) \\ &+\; C\sum_{n=1}^{m} k_n^2\Big(|\partial_t u|^2_{L^2(I_n)} + \|\alpha_t\|^2_{L^2(I_n;L^\infty(\Omega)} + \|\partial_{tt}\theta\|^2_{L^2(I_n,L^2(\Omega))} + \|\partial_{tt}a\|^2_{L^2(I_n,L^2(\Omega))}\Big). \\ &+\; \sum_{n=1}^{m-1} k_n(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2)\end{aligned} \tag{5.34}$$

Using discrete Gronwall's lemma and $k = \max_{1\leq n\leq N} k_n$, we arrive at the required result and this completes the rest of the proof. □

Similar to the error estimates for (θ, a), the following theorem yields error estimates for the adjoint variables (z, λ). The proof of the following theorem is on the same lines as Theorem 5.1 and hence is omitted.

Theorem 5.2. *Let* $(z_{hk}^{n,*}, \lambda_{hk}^{n,*})$ $\forall n = 1, 2, \cdots, N$ *and* (z^*, λ^*) *be the solutions of the adjoint problems* (5.7)-(5.10) *and* (3.6)-(3.9) *corresponding to the solutions* $(\theta_{hk}^{n,*}, a_{hk}^{n,*})$ $\forall n = 1, 2, \cdots, N$ *and* (θ^*, a^*) *of* (5.15)-(5.17) *and* (3.2)-(3.4), *respectively, with optimal control* $u^* \in U_{ad}$. *Then, under the extra regularity assumptions in Theorem 5.1 with* $(z^*, \lambda^*) \in L^\infty(I, H^2(\Omega)) \times L^\infty(I, H^2(\Omega))$, $(\partial_{tt}z, \partial_{tt}a) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega))$, $\partial_t z \in L^2(I, H^2(\Omega))$, $a_d \in H^2(\Omega)$; *there exists a positive constant* C *independent of* h *and* k, *such that*

$$\begin{aligned}\|z_{hk}^{n,*} - z^*(t_n)\| \;&+\; \|\lambda_{hk}^{n,*} - \lambda^*(t_n)\| \leq Ch^2\Big(\|\theta^*\|_{L^\infty(I,H^2(\Omega))} + \|a^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|_{L^2(I,H^2(\Omega))} \\ &+\|z^*\|_{L^\infty(I,H^2(\Omega))} + \|\lambda^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t z^*\|_{L^2(I,H^2(\Omega))} + \|\theta_0\|_{H^2(\Omega)} + \|a_d\|_{H^2(\Omega)}\Big) \\ &+Ck\Big(\|\partial_t u^*\|_{L^2(I)} + \|\partial_{tt}\theta^*\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}a\|_{L^\infty(I,L^2(\Omega))} \\ &+\|\partial_{tt}z^*\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}\lambda^*\|_{L^\infty(I,L^2(\Omega))}\Big).\end{aligned}$$

In order to completely discretize the problem (3.1)-(3.5), we choose discontinuous Galerkin piecewise constant approximation of the control variable. Let U_d be the finite dimensional subspace of U

defined by

$$U_d = \{v_d \in L^2(I) \ : \ v_d|_{I_n} = \text{constant}\} \quad \forall n = 1, 2, \cdots, N.$$

Let $U_{d,ad} = U_d \cap U_{ad}$ and $\sigma = \sigma(h, k, d)$ be the discretization parameter. The completely discretized problem reads as:

$$\min_{u_\sigma \in U_{d,ad}} J(\theta_\sigma, a_\sigma, u_\sigma) \qquad \text{subject to} \tag{5.35}$$

$$\sum_{n=1}^{N} (\partial_t a_\sigma, w)_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a_\sigma]_n, w_n^+) + (a_{\sigma,0}^+, w_0^+) = (f(\theta_\sigma, a_\sigma), w)_{I,\Omega}, \tag{5.36}$$

$$a_\sigma(0) = 0, \tag{5.37}$$

$$\begin{aligned} \rho c_p \sum_{n=1}^{N} (\partial_t \theta_\sigma, v)_{I_n,\Omega} + K(\nabla \theta_\sigma, \nabla v)_{I,\Omega} &+ \rho c_p \sum_{n=1}^{N-1} ([\theta_\sigma]_m, v_m^+) + \rho c_p (\theta_{\sigma,0}^+, v_0^+) \\ &= -\rho L(f(\theta_\sigma, a_\sigma), v)_{I,\Omega} + (\alpha u_\sigma, v)_{I,\Omega}, \\ &\quad + \rho c_p (\theta_0, v_0^+), \end{aligned} \tag{5.38}$$

$$\theta_\sigma(0) = \theta_0 \tag{5.39}$$

for all $(v, w) \in X_{hk}^q \times X_{hk}^q$.
We consider the case of piecewise constant approximation in time for the state equation (5.36)-(5.39), which can be rewritten as: for $n = 1, 2, \cdots, N$, find $(\theta_\sigma^n, a_\sigma^n) \in V_h \times V_h$ such that

$$\left(\frac{a_\sigma^n - a_\sigma^{n-1}}{k_n}, w \right) = \frac{1}{k_n} \left(\int_{I_n} f(\theta_\sigma^n, a_\sigma^n) ds, w \right), \tag{5.40}$$

$$a_\sigma(0) = 0, \tag{5.41}$$

$$\begin{aligned} \rho c_p \left(\frac{\theta_\sigma^n - \theta_\sigma^{n-1}}{k_n}, v \right) + K(\nabla \theta_\sigma^n, \nabla v) &= -\rho L \left(\frac{1}{k_n} \int_{I_n} f(\theta_\sigma^n, a_\sigma^n) ds, v \right) \\ &\quad + \left(\frac{1}{k_n} \int_{I_n} \alpha u_\sigma ds, v \right), \end{aligned} \tag{5.42}$$

$$\theta_\sigma(0) = \theta_{h,0}, \tag{5.43}$$

$\forall (w, v) \in V_h \times V_h$.

Lemma 5.1. *For a fixed control $u_\sigma \in U_{d,ad}$, the solution $(\theta_\sigma, a_\sigma) \in X_{hk}^q \times X_{hk}^q$ of (5.36)-(5.39), satisfies the following a priori bounds:*

$$\sum_{n=1}^{N} \left(\|\partial_t \theta_\sigma\|_{\Omega,I_n}^2 + \|\Delta_h \theta_\sigma\|_{\Omega,I_n}^2 \right) \le C, \qquad \sum_{n=1}^{N} \|\partial_t a_\sigma\|_{\Omega,I_n}^2 \le C. \tag{5.44}$$

Further for piecewise constant approximation, we have

$$\|\theta_\sigma^n\|^2 + \sum_{l=1}^{n} \|\nabla \theta_\sigma^l\|^2 \le C, \qquad \|a_\sigma^n\|^2 \le C \tag{5.45}$$

where $\Delta_h : V_h \times V_h$ is the discrete Laplacian defined by

$$-(\Delta_h v, w) = (\nabla v, \nabla w), \quad \forall v, w \in V_h. \tag{5.46}$$

Proof. Using (5.46) in (5.38), we have

$$\sum_{n=1}^{N}\Big(\rho c_p(\partial_t\theta_\sigma, v)_{\Omega,I_n} - (\Delta_h\theta_\sigma, v)_{\Omega,I_n} + \rho c_p([\theta_\sigma]_{n-1}, v_{n-1}^{+})\Big)$$
$$= \sum_{n=1}^{N}\Big(- \rho L(f(\theta_\sigma, a_\sigma), v)_{\Omega,I_n} + (\alpha u_\sigma, v)_{\Omega,I_n}\Big). \qquad (5.47)$$

Put $v = -\Delta_h\theta_\sigma$ in (5.47) to obtain

$$\sum_{n=1}^{N}\Big(\rho c_p(\partial_t\theta_\sigma, -\Delta_h\theta_\sigma)_{\Omega,I_n} - (\Delta_h\theta_\sigma, -\Delta_h\theta_\sigma)_{\Omega,I_n} + \rho c_p([\theta_\sigma]_{n-1}, -\Delta_h\theta_{\sigma,n-1}^{+})\Big)$$
$$= \sum_{n=1}^{N}\Big(- \rho L(f(\theta_\sigma, a_\sigma), -\Delta_h\theta_\sigma)_{\Omega,I_n} + (\alpha u_\sigma, -\Delta_h\theta_\sigma)_{\Omega,I_n}\Big). \qquad (5.48)$$

Again using (5.46) in first and third terms on the left hand side of (5.48), we obtain

$$\sum_{n=1}^{N}\Big(\rho c_p\int_{I_n}(\nabla\partial_t\theta_\sigma, \nabla\theta_\sigma)dt + \|\Delta_h\theta_\sigma\|_{\Omega,I_n}^2 + \rho c_p(\nabla[\theta_\sigma]_{n-1}, \nabla\theta_{\sigma,n-1}^{+})\Big)$$
$$= \sum_{n=1}^{N}\Big(- \rho L(f(\theta_\sigma, a_\sigma), -\Delta_h\theta_\sigma)_{\Omega,I_n} + (\alpha u_\sigma, -\Delta_h\theta_\sigma)_{\Omega,I_n}\Big). \qquad (5.49)$$

Now we find estimates for the terms in (5.49) one by one. Consider

$$\int_{I_n}(\nabla\partial_t\theta_\sigma, \nabla\theta_\sigma)dt = \int_{I_n}\frac{1}{2}\frac{d}{dt}\|\bigtriangledown\theta_\sigma\|^2 dt = \frac{1}{2}\Big(\|\bigtriangledown\theta_{\sigma,n}\|^2 - \|\bigtriangledown\theta_{\sigma,n-1}^{+}\|^2\Big) \qquad (5.50)$$

Now consider the 3^{rd} on the left side of (5.49)

$$([\nabla\theta_\sigma]_{n-1}, \nabla\theta_{\sigma,n-1}^{+}) = \frac{1}{2}\Big(\|\nabla\theta_{\sigma,n-1}^{+}\|^2 + \|[\nabla\theta_\sigma]_{n-1}\|^2 - \|\nabla\theta_{\sigma,n-1}\|^2\Big), \qquad (5.51)$$

Using (5.50), (5.51) in (5.49), Cauchy-Schwarz and Young's inequalities, we have

$$\begin{aligned}\|\nabla\theta_{\sigma,N}\|^2 - \|\nabla\theta_0\|^2 \quad &+ \quad \sum_{n=1}\|\Delta_h\theta_\sigma\|_{\Omega,I_n}^2 \\ &\leq \quad C\sum_{n=1}^{N}\Big(\|f(\theta_\sigma, a_\sigma)\|_{\Omega,I_n}^2 + \|\alpha u_\sigma\|_{\Omega,I_n}^2 + \|\Delta_h\theta_\sigma\|_{\Omega,I_n}^2\Big).\end{aligned}$$

Choosing Young's constant appropriately and using Remark 3.1, we obtain

$$\sum_{n=1}^{N}\|\Delta_h\theta_\sigma\|_{K,I_n}^2 \text{ is bounded.} \qquad (5.52)$$

Put $v = (t - t_{n-1})\partial_t\theta_\sigma$ in (5.38), use $((t - t_{n-1})\partial_t\theta_\sigma)^+_{n-1} = 0$ and (5.46) to obtain

$$\begin{aligned} \rho c_p \sum_{n=1}^{N} (\partial_t\theta_\sigma, (t - t_{n-1})\partial_t\theta_\sigma)_{\Omega,I_n} &- \sum_{n=1}^{N} \int_{I_n} (\Delta_h\theta_\sigma, (t - t_{n-1})\partial_t\theta_\sigma)_{\Omega,I_n} \\ &= \sum_{n=1}^{N} \Big(-\rho L(f(\theta_\sigma, a_\sigma), (t - t_{n-1})\partial_t\theta_\sigma)_{\Omega,I_n} \\ &\quad +(\alpha u_\sigma, (t - t_{n-1})\partial_t\theta_\sigma)_{\Omega,I_n}\Big). \end{aligned} \tag{5.53}$$

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

$$\begin{aligned} \sum_{n=1}^{N} \int_{I_n} (t - t_{n-1})\|\partial_t\theta_\sigma\|^2 dt &\le C \sum_{n=1}^{N} \Big(\|f(\theta_\sigma, a_\sigma)\|^2_{\Omega,I_n} + \|\alpha u_\sigma\|^2_{\Omega,I_n} + \|\Delta_h\theta_\sigma\|^2_{\Omega,I_n} \\ &\quad + \int_{I_n} (t - t_{n-1})\|\partial_t\theta_\sigma\|^2 dt\Big) \end{aligned}$$

Choosing Young's constant appropriately, using (5.52) and Remark 3.1, we obtain

$$\sum_{n=1}^{N} \int_{I_n} (t - t_{n-1})\|\partial_t\theta_\sigma\|^2 dt \text{ is bounded.}$$

From inverse estimate, we have

$$\sum_{n=1}^{N} \int_{I_n} \|\partial_t\theta_\sigma\|^2 dt \le C \sum_{n=1}^{N} k_n^{-1} \int_{I_n} (t - t_{n-1})\|\partial_t\theta_\sigma\|^2 dt$$

Therefore,

$$\sum_{n=1}^{N} \Big(\|\partial_t\theta_\sigma\|^2_{\Omega,I_n} + \|\Delta_h\theta_\sigma\|^2_{\Omega,I_n} \Big) \le C.$$

Similarly putting $w = (t - t_{n-1})\partial_t a_\sigma$ in (5.36) and using Inverse estimate, we obtain

$$\sum_{n=1}^{N} \|\partial_t a_\sigma\|^2_{\Omega,I_n} \le C.$$

Put $v = \theta^n_\sigma$ in (5.17) and consider

$$\rho c_p(\bar\partial\theta^n_\sigma, \theta^n_\sigma) + \mathcal{K}(\nabla\theta^n_\sigma, \nabla\theta^n_\sigma) = -\rho L(\frac{1}{k_n}\int_{I_n} f(\theta^n_\sigma, a^n_\sigma), \theta^n_\sigma) + (\frac{1}{k_n}\int_{I_n} \alpha u_\sigma dt, \theta^n_\sigma) \tag{5.54}$$

Using (5.23), Cauchy Schwarts and Young's inequality, we obtain

$$\|\theta^n_\sigma\|^2 - \|\theta^{n-1}_\sigma\| + \|\nabla\theta^n_\sigma\| \tag{5.55}$$

$$\le C\Big(\|\frac{1}{k_n}\int_{I_n} f(\theta_\sigma, a_\sigma)dt\|^2 + \|\frac{1}{k_n}\int_{I_n} \alpha u_\sigma dt\|^2 + \|\theta^n_\sigma\|^2 \Big) \tag{5.56}$$

Choosing Young's constant appropriately, Using Remark 3.1 and summing from 1 to n, we obtain

$$\|\theta_\sigma^n\|^2 + \sum_{l=1}^{n} \|\nabla \theta_\sigma^l\|^2 \leq C. \tag{5.57}$$

Similarly by putting $w = a_\sigma$ in (5.15), using Cauchy Schwartz inequality, Young's inequality and summing from 1 to n we obtain

$$\|a_\sigma^n\|^2 \leq C. \tag{5.58}$$

□

Corresponding to the solution $u_\sigma^* \in U_{ad}$ of (5.35)-(5.39), let $(\theta_\sigma^*, a_\sigma^*)$ be the solution to the state system (5.36)-(5.39). The first order optimality conditions yield the following adjoint problem: Find $(z_\sigma^*, \lambda_\sigma^*) \in X_{hk}^q \times X_{hk}^q$ such that

$$\begin{aligned}
-\sum_{n=1}^{N}(\psi, \partial_t \lambda_\sigma^*)_{I_n,\Omega} - \sum_{n=1}^{N-1}(\psi_n^-, [\lambda_\sigma^*]_n) &- (\psi_N^-, \lambda_{\sigma,N}^{-,*}) \\
+(\psi, f_a(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega} &= -(\psi_n^-, \lambda_\sigma^*(T)), & (5.59)\\
\lambda_{\sigma,N}^* &= \beta_1(a_\sigma^*(T) - a_d), & (5.60)\\
-\rho c_p \sum_{n=1}^{N}(\phi, \partial_t z_\sigma^*)_{I_n,\Omega} + K(\nabla\phi, \nabla z_\sigma^*)_{I,\Omega} &- \rho c_p \sum_{n=1}^{N-1}(\phi_n^-, [z_\sigma^*]_n) - \rho c_p(\phi_N^-, z_{\sigma,N}^{+,*}) \\
+(\phi, f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega} &= \beta_2(\phi, [\theta_\sigma^* - \theta_m]_+)_{I,\Omega}, & (5.61)\\
z_{\sigma,N}^* &= 0, & (5.62)
\end{aligned}$$

for all $(\psi, \phi) \in X_{hk}^q \times X_{hk}^q$. Moreover, z_σ^* satisfies the variational inequality,

$$\left(\beta_3(u_\sigma^* - u_d) + \int_\Omega \alpha z_\sigma^* dx, p - u_\sigma^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{d,ad}. \tag{5.63}$$

Theorem 5.3. *Let u_σ^* be the optimal control of* (5.35)-(5.39). *Then, there exists a subsequence (still denoted as $\{u_\sigma^*\}_{\sigma>0}$) $\lim_{\sigma\to 0} u_\sigma^* = u^*$ exists in $L^2(I)$ and u^* is an optimal control of* (3.1)-(3.5).

Proof: Since u_σ^* is an optimal control, we obtain

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

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

$$u_\sigma^* \longrightarrow u^* \text{ weakly in } L^2(I). \tag{5.64}$$

Since $U_{ad} \subset L^2(I)$ is a closed and convex set, we have $u^* \in U_{ad}$. Now corresponding to each u_σ^* there exists solution $(\theta_\sigma^*, a_\sigma^*)$ to (5.36)-(5.39). Thus from Lemma **??**, we have

$$\begin{aligned}
\theta_\sigma^* &\longrightarrow \theta^* \text{ weakly in } L^\infty(I, H^1(\Omega)), & (5.65)\\
\theta_\sigma^* &\longrightarrow \theta^* \text{ strongly in } C(I, L^2(\Omega)), & (5.66)\\
a_\sigma^* &\longrightarrow a^* \text{ weak}^* \text{ in } W^{1,\infty}(I, L^\infty(\Omega)), & (5.67)\\
a_\sigma^* &\longrightarrow a^* \text{ strongly in } L^\infty(I, L^2(\Omega)). & (5.68)
\end{aligned}$$

Now passing limit as $\sigma \to 0$, using (5.65)-(5.68) and Remark 3.1 in (5.36)-(5.39), we obtain that (u^*, θ^*, a^*) is an admissible solution for the optimal control problem (3.1)-(3.5). It now remains to show that (u^*, θ^*, a^*) is an optimal solution.
If possible, let $(\bar{u}^*, \bar{\theta}^*, \bar{a}^*)$ be another optimal solution of (3.1)-(3.5). Consider the auxiliary problem

$$\begin{aligned}
\sum_{n=1}^{N}\Big((\partial_t a_\sigma, w)_{\Omega,I_n} + ([a_\sigma]_{n-1}, w_{n-1}^+)\Big) &= \sum_{n=1}^{N}(f(\theta_\sigma, a_\sigma), w), && (5.69)\\
a_\sigma(0) &= 0, && (5.70)\\
\sum_{n=1}^{N}\Big(\rho c_p(\partial_t\theta_\sigma, v)_{\Omega,I_n} + K(\nabla\theta_\sigma, \nabla v)_{\Omega,I_n} + ([\theta_\sigma]_{n-1}, v_{n-1}^+)\Big) &= \sum_{n=1}^{N}\Big(-\rho L(f(\theta_\sigma, a_\sigma), v)_{\Omega,I_n}\\
&\quad +(\alpha\pi_k\bar{u}^*, v), && (5.71)\\
\theta_\sigma(0) &= \theta_0, && (5.72)
\end{aligned}$$

for all $(w, v) \in X_{hk}^q \times X_{hk}^q$. Then, there exists a solution to (5.69)-(5.72), say $(\bar{\theta}_\sigma, \bar{a}_\sigma) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$. Similar to (5.65)-(5.68), we arrive at

$$\begin{aligned}
\bar{\theta}_\sigma &\longrightarrow \bar{\theta} \text{ weakly in } L^\infty(I, H^1(\Omega)), && (5.73)\\
\bar{\theta}_\sigma &\longrightarrow \bar{\theta} \text{ strongly in } C(I, L^2(\Omega)), && (5.74)\\
\bar{a}_\sigma &\longrightarrow \bar{a} \text{ weakly in } W^{1,\infty}(I, L^\infty(\Omega)), && (5.75)\\
\bar{a}_\sigma &\longrightarrow \bar{a} \text{ strongly in } L^\infty(I, L^2(\Omega)). && (5.76)
\end{aligned}$$

Now letting $\sigma \to 0$ in (5.69)-(5.72), we obtain that $(\bar{\theta}, \bar{a})$ is a unique solution of (3.2)-(3.5) with respect to the control $\bar{u}^*$. Since the solution to (3.2)-(3.5) for a fixed control is unique, we find that $\bar{\theta} = \bar{\theta}^*$ and $\bar{a} = \bar{a}^*$.
Since u_σ^* is the optimal control for (5.35)-(5.39), we have

$$j(u_\sigma^*) \le j(\pi_k\bar{u}^*). \qquad (5.77)$$

Now letting $\sigma \to 0$ in (5.77) and using (5.64), we obtain

$$j(u^*) \le j(\bar{u}^*). \qquad (5.78)$$

Note from (5.78) that if $\bar{u}^*$ is another optimal control, then $j(\bar{u}^*)$ will be greater than or equal to $j(u^*)$ and hence, u^* is the optimal control.

Next we need to show that $\lim_{\sigma\to 0}\|u_\sigma^* - u\|_{L^2(I)} = 0$. Since $u_\sigma^* \longrightarrow u^*$ weakly in $L^2(\Omega)$, it is enough to show that $\lim_{\sigma\to 0}\|u_\sigma^*\|_{L^2(I)} = \|u^*\|_{L^2(I)}$. Using Lemma **??** and (5.64), we find that

$$\begin{aligned}
\lim_{\sigma\to 0}\frac{\beta_3}{2}\|u_\sigma^*\|_{L^2(I)}^2 &= \lim_{\sigma\to 0}\Big(J(\theta_\sigma^*, a_\sigma^*, u_\sigma^*) - \frac{\beta_1}{2}\|a_\sigma^*(T) - a_d\|^2 - \frac{\beta_2}{2}\|[\theta_\sigma^* - \theta_m]_+\|_{I,\Omega}^2\Big)\\
&= J(\theta^*, a^*, u^*) - \frac{\beta_1}{2}\|a^*(T) - a_d\|^2 - \frac{\beta_2}{2}\|[\theta^* - \theta_m]_+\|_{I,\Omega}^2\\
&= \frac{\beta_3}{2}\|u^*\|_{L^2(I)}^2,
\end{aligned}$$

that is , $\lim_{\sigma\to 0}\|u_\sigma^*\|_{L^2(I)} = \|u^*\|_{L^2(I)}$ and hence, $\lim_{\sigma\to 0}\|u_\sigma^* - u^*\| = 0$. This completes the rest of the proof. □

Remark 5.1. *In this paper, we have first fully discretized the laser surface hardening of steel problem using continuous Galerkin finite element method with piecewise linear polynomials in space, discontinuous Galerkin finite element method with constant approximation in time and control, and then nonlinear conjugate method has been used for the optimization. One can also use the strategy of optimizing the control problem first and then discretizing in space, time and control later on.*

Remark 5.2. *Since the workpiece used for the laser surface hardening is a thin sheet of steel, height of the workpiece has been ignored for the analysis and numerical implementation. By making appropriate changes in the formulation and the analysis, one can extend this to the case of a system in* $\mathbb{R}^3$.

6 Numerical Experiment

For the purpose of numerical experiment, we use cG(1) for the state and adjoint variables and dg(0) for time and control variables. We have used non-linear conjugate method [8] to evaluate the optimal control for the complete discretized problem (5.35)-(5.39).

Physical Data [8]: The parameters in the heat equation used are given by $\rho c_p = 4.91\frac{J}{cm^3K}, k = 0.64\frac{J}{cm^3K}$ and $\rho L = 627.9\frac{J}{cm^3K}$. The regularized monotone function $\mathcal{H}_\epsilon$ is chosen as

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

where $\epsilon = 0.15$. The initial temperature θ_0 and the melting temperature θ_m are chosen as 20 and 1800, respectively. Pointwise data for $aeq(\theta)$ and $\tau(\theta)$ are given by

θ	730	830	840	930
$aeq(\theta)$	0	0.91	1	1
$\tau(\theta)$	1	0.2	0.18	0.05

The shape function $\alpha(x, y, t)$ is given by $\alpha(x, y, t) = \frac{4k_1 A}{\pi D^2} exp(-\frac{2(x-vt)^2}{D^2}) exp(k_1 y)$, where $D = 0.47cm, k_1 = 60/cm, A = 0.3cm$ and $v = 1cm/s$. In the nonlinear conjugate gradient method tolerance is chosen as 10^{-7}.

Example: In the following numerical experiment we choose $\beta_1 = 7500, \beta_2 = 1000$ and $\beta_3 = 10^{-3}$. The main aim of this experiment is to achieve a constant hardening depth of 1mm , see Figure 3, with expected order of convergence $O(h^2 + k)$ for the approximation of (θ, a) and u. To apply non-linear conjugate method for the optimal control problem, we take u_0 (initial control) and u_d (desired control) as 1404.

When the finite element method is applied, the mesh used for space discretization is more refined near the area, where hardness is desired. With the initial control as u_0, we find that $\|a^0_\sigma(T) - a_d\| = 0.239547$, where a^0_σ corresponds to the austenite value for initial control u_0, which is being reduced to $\|a^{optimal}_\sigma(T) - a_d\| = 0.073632$ after applying non-linear conjugate method. Comparison of Figure 3 and Figure 4(a) shows that the goal of uniform hardening depth is nearly acheived. Also, the state constraint that $\|\theta\|_{L^\infty(Q)} < 1800$ is satisfied, since $\|\theta_\sigma\|_{L^\infty(Q)} < 1200$, see Figure 4(b). Figure 5 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. The numerical results

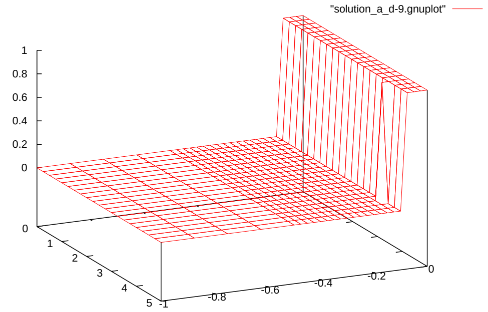
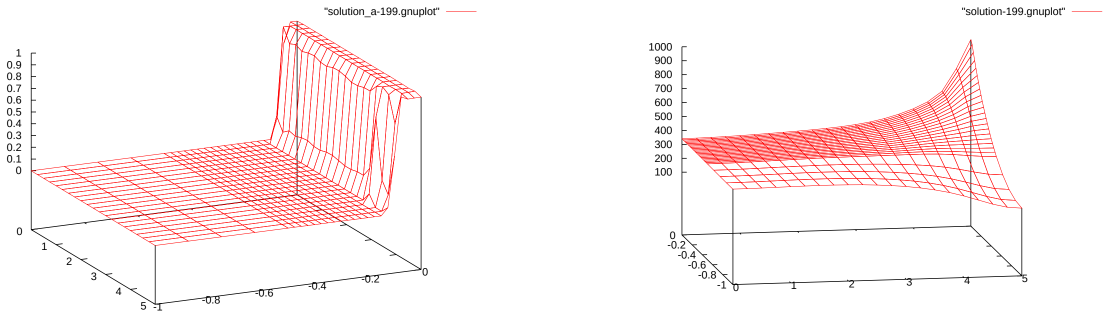


Figure 3: Goal a_d to be achieved for the volume fraction of austenite



(a) (b)

Figure 4: (a)The volume fraction of the austenite at time $t = T$ (b)The temperature at time $t = T$

confirm with those obtained in [8], though error estimates have not been developed in [8]. Figure 6 represents $\|E1\| = \|\theta - \theta_{hk}\|$ and $\|E2\| = \|a - a_{hk}\|$ as a function of the discretization step k in the log-log scale when $T = 5.25$. It is shown that the slope is approximately 2 confirming the theoretical order of convergence. Figure 7, shows $\|E1\|$ and $\|E2\|$ as a function of discretization step h in the log-log scale when $T = 5.25$. The slope is approximately 2, which justifies the theoretical order of convergence. Figure 8 represents the graph of $\|e(u)\| = \|u - u_\sigma\|$ as a function of the discretization parameter k in the log-log scale. It is shown that the slope is approximately equal to 2.

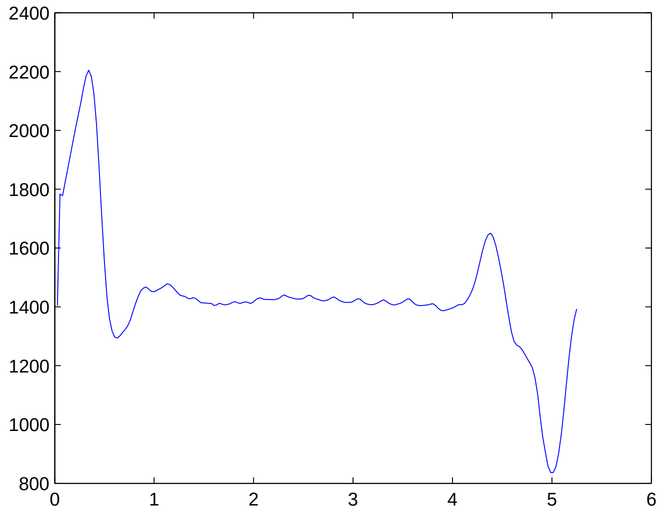


Figure 5: Laser energy

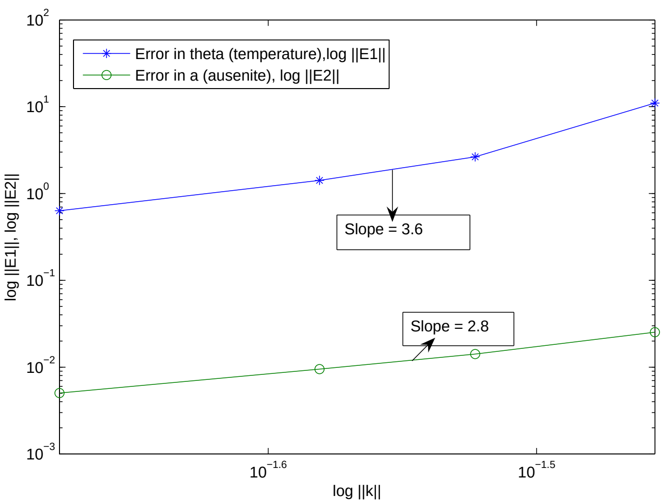


Figure 6: Refinement of the time steps for number of 525 nodes in spatial triangulation

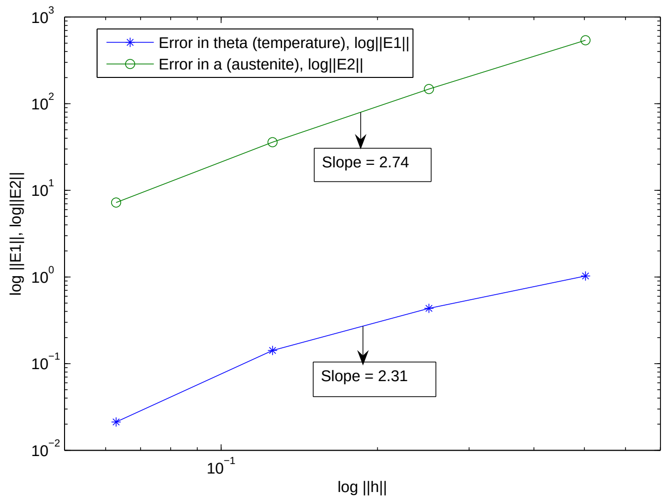


Figure 7: Refinement of spatial triangulation for 200 time steps.

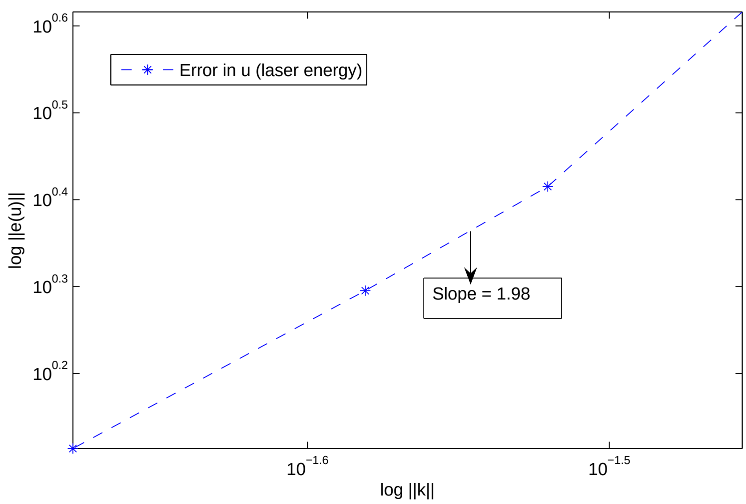


Figure 8: Evolution of control error.

References

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

[2] Leblond, J.B. and Devaux, J., A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size, Acta Metall. 32, 137(1984).

[3] Hömberg, D. and Sokolowski, J., Optimal control of laser hardening, Adv. Math. Sci. 8, 911(1998).

[4] Mazhukin, V.I. and Samarskii, A.A., Mathematical modelling in the technology of laser treatments of materials, Surveys Math. Indust. 4, 85(1994).

[5] Arnautu, V., Hömberg, D. and Sokolowski, J., Convergence results for a nonlinear parabolic control problem, Numer. Funct. Anal. Optim. 20, 805(1999).

[6] Hömberg, D. and Fuhrmann, J., Numerical simulation of surface hardening of steel, Internat. J. Numer. Meth. Heat Fluid Flow 9, 705(1999).

[7] Hömberg, D. and Volkwein, S., Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition, Math. Comput. Modelling 37, 1003(2003).

[8] Volkwein, S., Non-linear conjugate method for the optimal control of laser surface hardening, Optimization Methods and Software 19, 179(2004).

[9] Hömberg, D. and Weiss, W., PID-control of laser surface hardening of steel, IEEE Trans. Control Syst. Technol. 14, 896(2006).

[10] Thomee, V., Galerkin finite element methods for parabolic problems, Springer, 1997.

[11] Crouzeix, M., Thomee, V., Wahlbin L. B., Error estimates for spatial discrete approximation of semilinear parabolic equation with initial data of low regularity, Math. Comp. 53, 25(1989).

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

[13] Becker, R., Meidner, D. and Vexler, B., Efficient numerical solution of parabolic optimization problems by finite element methods, Optimization Methods and Software 22, 813(2007).

[14] Eriksson, K. and Johnson, C., Adaptive finite element methods for parabolic problems I: a linear model problem, SIAM J. Numer. Anal. 28, 43(1991).

[15] Eriksson, K. and Johnson, C., Adaptive finite element methods for parabolic problems II: optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$, SIAM J. Numer. Anal. 32, 706(1995).

[16] Falk, F.S., Approximation of a class of optimal control problems with order of convergence estimates, J. Math. Anal. Appl. 44, 28(1973).

[17] Meidner, D. and Vexler, B., A priori error estimates for space-time finite element discretization of parabolic optimal control problems part I: problems without control constraints, SIAM J. Control and Optimization 47, 1150(2007).

[18] Meidner, D. and Vexler, B., A priori error estimates for space-time finite element discretization of parabolic optimal control problems part I: problems with control constraints, SIAM J. Control and Optimization 47, 1301(2007).

[19] Sang-Uk Ryu, Optimal control problem governed by some parabolic equations, Nonlinear Analysis 56, 241(2004).

[20] Wheeler, M.F, A priori L^2 error estimates for Galerkin apporximation to parabolic differential equations, SIAM J. Numerical Analysis 10, 723(1973).

[21] Liu, W. and Yan, N., A posteriori error estimates for control problems governed by nonlinear elliptic equations, Applied Numerical Mathematics, 47, 173(2003).