FINITE ELEMENT METHODS FOR THE OPTIMAL CONTROL PROBLEM OF LASER SURFACE HARDENING OF STEEL

Thesis
submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

by

Nupur Gupta
(04409302)

Under the Supervision of

Professor Neela Nataraj

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY
2009

Approval Sheet

The thesis entitled

"FINITE ELEMENT METHODS FOR THE OPTIMAL CONTROL PROBLEM OF LASER SURFACE HARDENING OF STEEL"

by

Nupur Gupta

is approved for the degree of

DOCTOR OF PHILOSOPHY

Examiners

____

____

Supervisors

____

____

Chairman

____

Date : ____________

Place : ____________

Dedicated to

My Parents
Shri. Suresh K. Gupta and Smt. Mithlesh Gupta

&

My Guide
Prof. Neela Nataraj

Abstract

The main focus of this thesis has been on the development of finite element methods for the optimal control problem of laser surface hardening of steel governed by a dynamical system consisting of a semi-linear parabolic equation and an ordinary differential equation. For the purpose of applying numerical methods, we use a mathematical model where there is a phase transformation from $(ferrite + pearlite)$ to *austenite*. The control function in the problem consists of a term which helps in achieving the desired volume of austenite, a penalization term, which restricts the increment of temperature beyond melting temperature of steel and the cost due to the laser energy. The dynamical system consists of an ordinary differential equation, which describes the evolution of austenite and a semi-linear parabolic equation which shows the evolution of temperature during the formation of austenite.

The ordinary differential equation consists of a non-differentiable right hand side function, which restricts the development of mathematical and numerical methods. To avoid this problem, the right hand side function has been regularized using a monotone regularized Heaviside function with regularization parameter ϵ. Then, the regularized problem has been studied in literature for the existence, uniqueness and stability of the solution. It has been shown that the solution of the regularized problem converges to the solution of original problem with the order of convergence $\mathcal{O}(\epsilon)$ and then the existence of the solution for the original problem has been established.

For the purpose of numerical approximation of the variables describing volume of austenite, temperature and laser energy, finite element methods have been used and analyzed. First of all, a continuous Galerkin method is applied for the discretization of space and a discontinuous Galerkin method is applied for the discretization of time and control variables. Optimal *a priori* error estimates of order $\mathcal{O}(h^2 + k)$, where h and k are space and time discretization parameters, respectively, are developed.

Due to irregularity of solutions of the laser surface hardening of steel problem, a non-uniformity in the triangulation of domain becomes relevant. Therefore, an *hp*-discontinuous Galerkin method has been applied for discretization of space, and a discontinuous Galerkin finite element method for time and control variables. *A priori* error estimates obtained are optimal in nature. Numerical results obtained justify the theoretical order of convergences established. Also, adaptive finite element methods for the laser surface hardening of steel problem using residual and dual weighted residual type estimators for space discretization and a discontinuous Galerkin method for time and control discretizations have been developed and used for numerical implementations.

Table of Contents

1 Introduction 6

1.1 Motivation 6

1.2 Preliminaries 9

1.3 The Laser Surface Hardening of Steel Problem 15

1.4 Literature Review 18

1.5 Organization of the Thesis 22

2 Regularization Error Estimates for the Optimal Control Problem of Laser Surface Hardening of Steel 24

2.1 Introduction 24

2.2 Existence and Uniqueness Results for the Regularized Laser Surface Hardening of Steel Problem 25

2.3 Existence and Uniqueness Results for the Laser Surface Hardening of Steel Problem 32

2.4 Summary 41

3 *A Priori* Error Estimates for the Optimal Control Problem of Laser Surface Hardening of Steel 42

3.1 Introduction 42

3.2 Weak Formulation and Semi-Discrete Scheme 43

3.3 Completely Discrete Scheme 51

3.4 Numerical Experiment 65

3.5 Summary 73

4 *A Priori* Error Estimates: A Discontinuous Galerkin Space-Time Method 74

4.1 Introduction 74

4.2 Preliminaries . . . 76
4.3 *hp*-Discontinuous Galerkin Weak Formulation . . . 81
4.4 *hp*-DGFEM-DG Space-Time-Control Discretization . . . 93
4.5 Numerical Experiment . . . 108
4.6 Summary . . . 112

5 Adaptive Finite Element Methods 113
5.1 Introduction . . . 113
5.2 Preliminaries . . . 115
5.3 Residual Method . . . 117
5.4 Dual Weighted Residual Method . . . 133
5.5 Numerical Experiments . . . 149
5.6 Summary . . . 154

6 Conclusions and Future Directions 155
6.1 Summary and Critical Assessments of the Results . . . 155
6.2 Future Extensions . . . 159

References 161

List of Tables

5.1 **Error in space for fixed time partition** 100 150

6.1 **Comparitive study between three approaches used in the thesis** . . 159

List of Figures

1.1 Possible phase transitions in steel 7

1.2 Laser hardening process 8

1.3 Regularized Heaviside($\mathcal{H}_\epsilon$) function and Heaviside($\mathcal{H}$) function . . . 17

3.1 Goal a_d to be achieved for the volume fraction of austenite 67

3.2 (a) The volume fraction of the austenite at time $t = T$ (b) The temperature at time $t = T$ 69

3.3 Laser energy 70

3.4 Refinement of the time steps for number of 525 nodes in spatial triangulation 70

3.5 Refinement of spatial triangulation for 200 time steps. 71

3.6 Evolution of control error. 71

3.7 Evolution of state error as a function of ϵ. 72

3.8 Evolution of control error as a function of ϵ. 72

4.1 An example of construction of finite elements 76

4.2 Convergence of hp-DGFEM with h-refinement: Temperature and Austenite 108

4.3 Convergence of hp-DGFEM with p-refinement. (a) Temperature (b) Austenite 109

4.4 Convergence of DGFEM with k-refinement 110

4.5 (a) Temperature (b) Austenite 111

5.1 (a) Patched mesh (b) Macrocell from four adjacent cells. 117

5.2 Initial approximate triangulation 150

5.3 Adaptive refinement using DWR type estimators 151

5.4 **Adaptive refinement using residual type estimators** 152

5.5 **Error graphs** 152

5.6 **The volume fraction of the austenite at time $t = T$ using (a) DWR estimator (b) residual estimator** 153

5.7 **The temperature at time $t = T$ using (a) DWR estimator (b) residual estimator** 153

5.8 **(a) Control after using DWR AFEM (b) Control after using residual AFEM** 154

Chapter 1

Introduction

The main focus of this thesis is to study the optimal control problem of laser surface hardening of steel, which is governed by a dynamical system consisting of a semilinear parabolic equation and an ordinary differential equation with a non-differential right hand side function. To avoid the numerical and analytical difficulties posed by the non-differentiable right hand side function, it is regularized using a monotone approximation of the Heaviside function and the regularized problem has been studied in literature. In this dissertation, first of all we establish the convergence of the solution of the regularized problem to that of the original problem and justify the existence of the solution of the original problem. We also study discretization schemes using a continuous Galerkin (cG) method for space variable and a discontinuous Galerkin (dG) method for time and control variables; using an *hp*-discontinuous Galerkin finite element method (*hp*-DGFEM) for space variable and dG method for time and control variables; Adaptive Finite Element Methods (AFEM) using residual and Dual Weighted Residual (DWR) for space discretization and a dG method for time and control variables.

1.1 Motivation

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 workpiece by rapid heating and subsequent quenching. The desired hardening results in a change of the micro structure. A few applications include cutting tools, wheels, driving axles, gears, etc. One of the methods for the rapid heating of the workpiece is to use laser beam on the top of the boundary layer. This is called *laser surface hardening of steel.* It is an optimal control problem governed by a dynamical system of differential equations, consisting of a semi-linear parabolic equation and an ordinary differential equation. The cost function in the problem consists of a term which helps in achieving the desired volume of austenite, a penalization term, which restricts the increment of temperature beyond melting temperature of steel and the cost occurred due to the laser energy.

The hardening of steel can be achieved by different surface heat treatment procedures like case hardening, induction hardening, laser hardening, etc. The reason for the possibility to change the hardness of steel by thermal treatment originate from the occurring phase transition (see Figure 1.1). At room temperature, steel in general is a mixture of ferrite, pearlite, bainite and martensite. Upon heating, these phases are transformed to austenite. The heated zone is rapidly quenched by self cooling, leading to further phase transitions and the desired hardening effect. In case hardening [59], carbon is dissolved in the surface layer of a low-carbon steel part at a temperature sufficient to render the steel austenite and then quenched to form a martensite micro structure. The induction hardening [59] relies on the transformer principle. A given current density induces eddy currents in the workpiece which lead to an increase in the temperature of the boundary layers of the workpiece. The current in then switched off and the workpiece is quenched by spray water cooling.

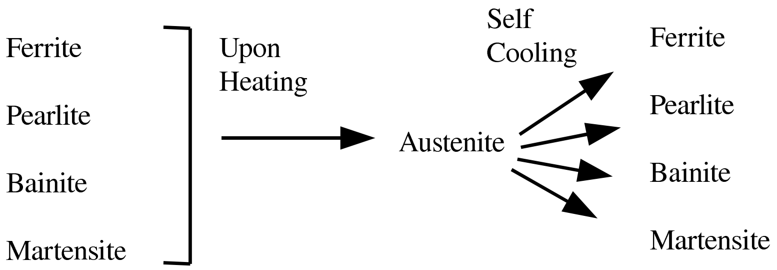


Figure 1.1: **Possible phase transitions in steel**

When the workpiece is very large or has a complicated geometry with curved edges, laser hardening becomes attractive. In this process, the laser beam moves along the surface of the workpiece, where the hardening is required, creating a heated zone around its trace (see Figure 1.2). The laser radiation is absorbed by the workpiece, leading to the rapid heating of its boundary layer. The heating process is accompanied by a phase transition in which austenite is produced. Since the penetration depth of the laser beam is very small, typically not more than 1mm, the heated zone is rapidly quenched by self cooling of the workpiece accompanied by a growth of the surface hardness. To increase the scanning width, sometimes the laser beam performs an additional oscillatory movement orthogonal to the principal moving direction. Since the velocity of the moving laser beam is tried to

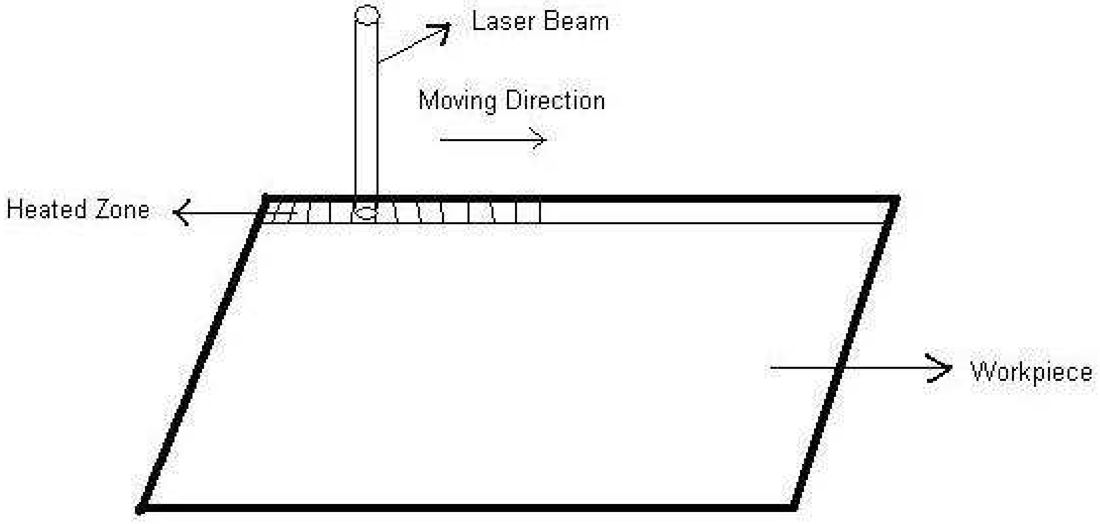


Figure 1.2: **Laser hardening process**

be kept as a constant, the most important control parameter is the laser energy. Whenever the temperature in the heated zone exceeds the melting temperature of steel, the workpiece quality is destroyed. Therefore, a precise adjustment of the laser beam is an important task, especially when the laser approaches a workpiece boundary or when there are large variations in the thickness of the workpiece.

The mathematical model for the laser surface hardening of steel has been studied in [43], [44], [46], [52], [59], [82]. In this dissertation, we follow the model described by Leblond-Devaux [52] for the case of two phase transition, that is, from *(ferrite + pearlite) to austenite.* A study of different methods of discretization and their comparison for the optimal control problem of the laser surface hardening of steel is conspicuous by its absence and hence is studied in this dissertation.

We first state the existence, uniqueness and stability results for the regularized laser surface hardening of steel problem given in the literature and then establish the convergence of the solution of the regularized problem to that of the original laser surface hardening of steel problem. Also, the existence of solution of the original problem is shown using the convergence results. The regularized problem has been discretized using a cG method for space and a dG method for time and control variables. *A priori* error estimates have been developed to show the convergence of the approximate solutions to the exact solutions and numerical results are presented to justify the theoretical results obtained.

Due to irregularity of solutions of the laser surface hardening of steel problem, a non-uniformity in the triangulation of domain becomes relevant. Therefore, Chapter 4 has been devoted to apply and analyze *hp*-DGFEM for the discretization of space, DGFEM for time and control variables. Finally, AFEM using residual type and DWR type estimators for space discretization and a dG method for time and control discretizations have been analyzed and used for numerical implementations.

1.2 Preliminaries

We need some well known results and definitions, which we state in this section.
Let Ω denote an open bounded subset of $\mathbb{R}^2$.

Definition 1.2.1. (**Functional Spaces [16]**) *For* $1 \leq p < \infty$*, let* $L^p(\Omega)$ *denote the real valued measurable functions* ϕ *on* $\Omega \subset \mathbb{R}^2$ *for which* $\displaystyle\int_\Omega |\phi|^p dx < \infty$*. The norm on* $L^p(\Omega)$ *is given by*

$$\|\phi\|_{L^p(\Omega)} = \left(\int_\Omega |\phi|^p dx \right)^{1/p} \quad \textit{for } 1 \leq p < \infty.$$

In addition, let $L^\infty(\Omega)$ *denote the real valued measurable functions which are essentially bounded in* Ω *and with norm defined by*

$$\|\phi\|_{L^\infty(\Omega)} = ess \sup_{x \in \Omega} |\phi(x)|.$$

For natural numbers $m \geq 1$ *and* $1 \leq p < \infty$*, let* $W^{m,p}(\Omega)$ *denote the Sobolev spaces, which are defined by*

$$W^{m,p}(\Omega) = \{\phi \in L^p(\Omega) : \partial^\alpha \phi \in L^p(\Omega),\ |\alpha| \leq m\},$$

and for $p = \infty$

$$W^{m,\infty}(\Omega) = \{\phi \in L^\infty(\Omega) : \partial^\alpha \phi \in L^\infty(\Omega),\ |\alpha| \leq m\},$$

where α *is multi-valued index defined as* $\alpha = (\alpha_1, \alpha_2), \alpha_i \geq 0, \alpha_i \in \mathbb{N}$ *with* $|\alpha| = \displaystyle\sum_{i=1}^{2} \alpha_i$*, endowed with norm and semi-norm as*

$$\|\phi\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha| \leq m} \|\partial^\alpha \phi\|^p_{L^p(\Omega)} \right)^{1/p} \quad \textit{for } m \geq 1,$$

and

$$|\phi|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha|=m} \|\partial^\alpha \phi\|^p_{L^p(\Omega)} \right)^{1/p} \quad \textit{for } m \geq 1, \textit{ respectively.}$$

For $p = 2$, $W^{m,2}(\Omega) = H^m(\Omega)$ *denotes the standard Hilbert Sobolev space of order* m. $H^m(\Omega)$ *is equipped with norm and semi-norm defined by*

$$\|\phi\|_{H^m(\Omega)} = \left(\sum_{|\alpha|\leq m} \|\partial^\alpha \phi\|^2_{L^2(\Omega)} \right)^{1/2} \quad \textit{for } m \geq 1.$$

and

$$|\phi|_{H^m(\Omega)} = \left(\sum_{|\alpha|=m} \|\partial^\alpha \phi\|^2_{L^2(\Omega)} \right)^{1/2} \quad \textit{for } m \geq 1, \textit{ respectively.}$$

Now, let $I = (a, b)$ *be an interval with* $-\infty < a < b < \infty$, *and let* X *be a Banach space with norm* $\|\cdot\|_X$. *For* $1 \leq p < \infty$, *we denote* $L^p(I; X)$ *the space*

$$L^p(I;X) = \{\phi : I \longrightarrow X \Big| \phi(t) \textit{ is measurable in } I \textit{ and } \int_I \|\phi(t)\|^p_X < \infty\}.$$

It is equipped with the following norm for $1 \leq p < \infty$

$$\|\phi\|_{L^p(I;X)} = \left(\int_I \|\phi(t)\|^p_X dt \right)^{1/p},$$

and for $p = \infty$

$$\|\phi\|_{L^p(I;X)} = ess \sup_{t\in I} \|\phi(t)\|_X.$$

When $-\infty < a < b < \infty$, *the space*

$$C(I;X) = \{\phi : I \longrightarrow X \Big| \phi \textit{ is continuous in } I\}$$

is a Banach space equipped with the norm, $\|\phi\|_{C(I;X)} = \max_{t\in I} \|\phi(t)\|_X$.

Definition 1.2.2. (**Lipschitz Continuous [16]**) *Let* X *and* Y *be normed spaces with norms as* $\|\cdot\|_X$ *and* $\|\cdot\|_Y$, *respectively. A function* $f : X \longrightarrow Y$ *is called Lipschitz continuous, iff there exists a smallest* $C > 0$, *called the Lipschitz constant, so that for all* $x, y \in X$, *we have*

$$\|f(x) - f(y)\|_Y \leq C\|x - y\|_X.$$

Definition 1.2.3. (**Strong Convergence** [**40**]) *A sequence* $\{x_n\} \in X$ *in a normed space* X *is said to be strongly convergent if there exists an* $x \in X$ *such that*

$$\lim_{n \to \infty} \|x_n - x\|_X = 0.$$

Definition 1.2.4. (**Lower Semi-Continuous** [**38**]) *Let* X *be a Banach Space and* $J : X \longrightarrow \mathbb{R} \bigcup \{-\infty, \infty\}$ *be an extended real valued function. Then,* J *is lower semi-continuous at* x_0 *if*

$$\lim \inf_{x \in X} J(x) \geq J(x_0).$$

Definition 1.2.5. (**Weak Convergence** [**40**]) *Let* X *be a normed space. Then* $\{x_n\} \in X$ *is said to be convergent to* $x \in X$ *weakly, iff* $f(x_n) \to f(x)$*, for all* $f \in X^*$*, where* X^* *is the dual of* X*.*

Definition 1.2.6. (**Weak-* Convergence** [**40**]) *Let* ϕ_n *be a sequence of bounded linear functions on normed space* X*, then weak-* convergence of* ϕ_n *means that there exists a function* $\phi \in X^*$ *such that*

$$\phi_n(x) \longrightarrow \phi(x) \quad \forall x \in X.$$

Lemma 1.2.1. (**Young's Inequality**) *Let* a *and* b *be two positive numbers and* $1 \leq p, q < \infty$ *be such that,* $\dfrac{1}{p} + \dfrac{1}{q} = 1$*, then the following inequality holds true:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Definition 1.2.7. (**Contraction** [**40**]) *Let* $X = (X, d)$ *be a metric space. A mapping* $T : X \to X$ *is called a contraction on* X *if there is a positive real number* $\alpha < 1$ *such that for all* $x, y \in X$

$$d(Tx, Ty) \leq \alpha d(x, y).$$

Theorem 1.2.1. (**Theorem of Carathéodory** [**87**]) *Let the function* $f : J \times K \longrightarrow \mathbb{R}$ *satisfy the Carathéodory conditions, i.e., for all* $i = 1, 2, \cdots, n$*,*

$$\begin{aligned} t &\longrightarrow f_i(t, x) \quad \text{is measurable on } J \text{ for each } x \in K; \\ x &\longrightarrow f_i(t, x) \quad \text{is continuous on } K \text{ for almost all } t \in J, \end{aligned}$$

where $J = \{t \in \mathbb{R} : |t - t_0| \leq r_0\}$ *and* $K = \{x \in \mathbb{R}^n : |x - x_0| \leq r\}$ *and there exists an integral function* $M : J \longrightarrow \mathbb{R}$ *(majorant) such that*

$$|f_i(t, x)| \leq M(t) \ \ \forall \ (t, x) \in J \times K \ \ \forall i = 1, 2, \cdots, n.$$

Then,

(a) there exists an open bounded neighbourhood U *of* t_0 *and a continuous function* $x(\cdot) : U \longrightarrow \mathbb{R}^n$ *which solves the integral equation*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds, \quad t \in U$$

(b) for almost all $t \in U$*, the derivative* $x'(t)$ *exists and*

$$x'(t) = f(t, x(t)) \ \ t \in U, \quad x(t_0) = x_0, \tag{1.2.1}$$

holds.

(c) The components $\zeta_1(\cdot), \cdots, \zeta_N(\cdot)$*, of* $x(t)$*, have generalized derivatives on* U*, and* $x(\cdot)$ *is a solution of* (1.2.1) *on* U *in the sense of generalized derivatives.*

Theorem 1.2.2. (**Banach Fixed Point Theorem** [**40**]) *Consider a nonempty metric space* (X, d)*. Suppose that* X *is complete and let* $T : X \to X$ *be a contraction on* X*. Then,* T *has precisely one fixed point.*

Lemma 1.2.2. (Hölder's inequality [**16**]**)** *Let* $1 \leq p, q < \infty$ *be such that* $1/p + 1/q = 1$*. Suppose that* $\phi \in L^p(\Omega)$ *and* $\psi \in L^q(\Omega)$*. Then*

$$\left| \int_\Omega \phi\psi \ dx \right| \leq \left(\int_\Omega |\phi|^p \ dx \right)^{1/p} \left(\int_\Omega |\psi|^q \ dx \right)^{1/q}.$$

Lemma 1.2.3. (Generalized Hölder's inequality [**16**]**)** *Let* $1 \leq p, \ q, \ r < \infty$ *be such that* $1/p + 1/q + 1/r = 1$*. Suppose that* $\phi \in L^p(\Omega)$*,* $\psi \in L^q(\Omega)$ *and* $\chi \in L^r(\Omega)$*. Then*

$$\left| \int_\Omega \phi\psi\chi \ dx \right| \leq \left(\int_\Omega |\phi|^p \ dx \right)^{1/p} \left(\int_\Omega |\psi|^q \ dx \right)^{1/q} \left(\int_\Omega |\chi|^r \ dx \right)^{1/r}.$$

Lemma 1.2.4. (Cauchy-Schwarz inequality [16]) *For $\Omega \subset \mathbb{R}^2$, suppose that $\phi, \psi \in L^2(\Omega)$. Then, $\phi\psi \in L^1(\Omega)$ and*

$$\left(\int_\Omega \phi\chi dx\right) \;\; \leq \;\; \left(\int_\Omega \phi^2 dx\right)^{1/2} \left(\int_\Omega \psi^2 dx\right)^{1/2}$$

Lemma 1.2.5. (**Gronwall's Inequality [27]**) *Let I denote an interval of the real line of the form $[a, \infty), [a, b]$ or $[a, b)$ with $a < b$. Let β and u be real valued functions defined on I that are continuous. If β is non-negative and u satisfies the integral inequality*

$$u(t) \leq \alpha + \int_a^t \beta(s)u(s)ds, \quad t \in I,$$

then

$$u(t) \leq \alpha exp\left(\int_a^t \beta(s)ds\right), \quad t \in I.$$

Definition 1.2.8. (**Directional derivative**) *Let X and Y be two normed spaces, X_0 be a non-empty open subset of X and $g : X_0 \to Y$ be a given mapping. If for two elements $x \in X_0$ and $v \in X$ the limit*

$$g'(x)(v) = \lim_{\gamma \to 0} \frac{g(x + \gamma v) - g(x)}{\gamma}$$

exists, then $g'(x)(v)$ is called the directional derivative of g at x in direction v. Moreover, if the limit

$$g''(x)(v, v) = \lim_{\gamma \to 0} \frac{g'(x + \gamma v)(v) - g'(x)(v)}{\gamma},$$

exists, then, $g''(x)(v, v)$ is called the second order directional derivative of g in the direction of v.

Definition 1.2.9. (**Gâteaux derivative**) *Let X and Y be two normed spaces, X_0 be a non-empty open subset of X. A directionally differentiable mapping $g : X_0 \to Y$ is called Gâteaux differentiable at $x \in X_0$, if the directional derivative $g'(x)$ is a continuous linear mapping from X to Y. $g'(x)$ is then called Gâteaux derivative of g at x.*

Definition 1.2.10. (**Fréchet derivative**) *Let X and Y be two normed spaces, X_0 be an open non-empty subset of X and $g : X_0 \to Y$ be a given mapping. Furthermore, let an* *element $x \in X_0$ be given. If there is a continuous linear mapping $g'(x) : X \to Y$ with the property*

$$\lim_{\|v\|_X \to 0} \frac{\|g(x+v) - g(x) - g'(x)(v)\|_Y}{\|v\|_X} = 0,$$

then $g'(x)$ is called the Fréchet derivative of g at x and g is called Fréchet differentiable at x.

Consider the control problem,

$$\min_{u \in U_{ad}} j(u), \tag{1.2.2}$$

j being the cost functional and U_{ad} being the space of admissible controls.

Definition 1.2.11. (**Local Optimal Solution**) *A control $u \in U_{ad}$ is called the local optimal solution of the optimization problem* (1.2.2) *if there exists a neighborhood $U_0 \subseteq U_{ad}$ containing u such that*

$$j(u) \leq j(p) \quad \forall p \in U_0.$$

Theorem 1.2.3. (**First Order Necessary Optimality Condition**) *Let the reduced cost functional j be Gâteaux differentiable on a convex subset $U_0 \subseteq U_{ad}$. If $u \in U_0$ is a local optimal solution of the optimization problem* (1.2.2)*, then there holds the first order necessary optimality condition*

$$j'(u)(p-u) \geq 0 \quad \forall p \in U_{ad}.$$

Definition 1.2.12. (**Uniformly (Strongly) Convex Functions** [**35**])*) Let $S \subset \mathbb{R}^2$ be a non-empty convex set. A function $f : S \longrightarrow \mathbb{R}$ is said to be uniformly convex, when, for all pairs $(x, x') \in S \times S$ and $0 < \lambda < 1$, there exists a constant $C > 0$, such that*

$$f(\lambda x + (1-\lambda)x') < \lambda f(x) + (1-\lambda) f(x') - \frac{1}{2} C \lambda (1-\lambda) \|x - x'\|^2.$$

Remark 1.2.1. *The generic constant $C > 0$ takes different values at different instances and will be used throughout this dissertation.*

1.3 The Laser Surface Hardening of Steel Problem

We present a mathematical model for the laser surface hardening of steel problem based on [52]. The laser surface hardening of steel problem is formulated in terms of a distributed optimal control problem in which the state equations are composed of a semi-linear parabolic equation and an ordinary differential equation describing the evolution of temperature and austenite, respectively.

Mathematical Modeling

Let $\Omega \subset \mathbb{R}^2$ be a convex, bounded domain with Lipschitz continuous boundary $\partial\Omega$, $Q = \Omega \times I$ and $\Sigma = \partial\Omega \times I$, where $I = (0, T)$. According to Leblond and Devaux [52], the evolution of volume fraction of the austenite $a(t)$ for given temperature evolution $\theta(t)$ is described by the initial value problem:

$$\begin{aligned} \partial_t a = f_+(\theta, a) &= \frac{1}{\tau(\theta)}[a_{eq}(\theta) - a]_+ \qquad \text{in } Q, && (1.3.1)\\ a(0) &= 0 \qquad \text{in } \Omega, && (1.3.2) \end{aligned}$$

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

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

denotes the non-negative part of $a_{eq}(\theta) - a$, that is, $[a_{eq}(\theta) - a]_+ = \dfrac{(a_{eq}(\theta) - a) + |a_{eq}(\theta) - a|}{2}$.

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

$$\begin{aligned} \rho c_p \partial_t \theta - \mathcal{K} \bigtriangleup \theta &= -\rho L a_t + \alpha u \qquad \text{in } Q, && (1.3.3)\\ \theta(0) &= \theta_0 \qquad \text{in } \Omega, && (1.3.4) \end{aligned}$$

$$\frac{\partial \theta}{\partial n} = 0 \quad \text{on } \Sigma, \tag{1.3.5}$$

where the density ρ, 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, 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:

$$J(\theta, a, u) = \frac{\beta_1}{2}\int_\Omega |a(T) - a_d|^2 dx + \frac{\beta_2}{2}\int_0^T \int_\Omega [\theta - \theta_m]_+^2 dxdt + \frac{\beta_3}{2}\int_0^T |u|^2 dt \tag{1.3.6}$$

$$\text{subject to state equations } (1.3.1) - (1.3.5) \text{ in the set of admissible controls } U_{ad}, \tag{1.3.7}$$

where $U_{ad} = \{u \in U : 0 \leq u \leq u_{max} \text{ a.e. in } I\}$; $U = L^2(I)$, $u_{max} > 0$, denoting the maximal intensity of the laser, β_1, β_2 and β_3 being positive constants, a_d being the given desired fraction of austenite. The second term in (1.3.6) 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 (1.3.1) is regularized using a monotone regularized Heaviside Function (see Figure 1.3) and the regularized laser surface hardening problem is given by:

$$\min_{u_\epsilon \in U_{ad}} J(\theta_\epsilon, a_\epsilon, u_\epsilon) \text{ subject to} \tag{1.3.8}$$

$$\partial_t a_\epsilon = f_\epsilon(\theta_\epsilon, a_\epsilon) = \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)\mathcal{H}_\epsilon(a_{eq}(\theta_\epsilon) - a_\epsilon) \quad \text{in } Q, \tag{1.3.9}$$

$$a_\epsilon(0) = 0 \quad \text{in } \Omega, \tag{1.3.10}$$

$$\rho c_p \partial_t \theta_\epsilon - \mathcal{K} \triangle \theta_\epsilon = -\rho L \partial_t a_\epsilon + \alpha u_\epsilon \quad \text{in } Q, \tag{1.3.11}$$

$$\theta_\epsilon(0) = \theta_0 \quad \text{in } \Omega, \tag{1.3.12}$$

$$\frac{\partial \theta_\epsilon}{\partial n} = 0 \quad \text{on } \Sigma, \tag{1.3.13}$$

where $J(\theta_\epsilon, a_\epsilon, u_\epsilon) = \dfrac{\beta_1}{2}\displaystyle\int_\Omega |a_\epsilon(T) - a_d|^2 dx + \dfrac{\beta_2}{2}\int_0^T \int_\Omega [\theta_\epsilon - \theta_m]_+^2 dxds + \dfrac{\beta_3}{2}\int_0^T |u_\epsilon|^2 ds$ and $\mathcal{H}_\epsilon \in C^{1,1}(\mathbb{R})$ is a monotone approximation of the Heaviside function satisfying $\mathcal{H}_\epsilon(x) = 0$ for $x \leq 0$, defined by

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

.

We now make the following assumptions [47]:

(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.

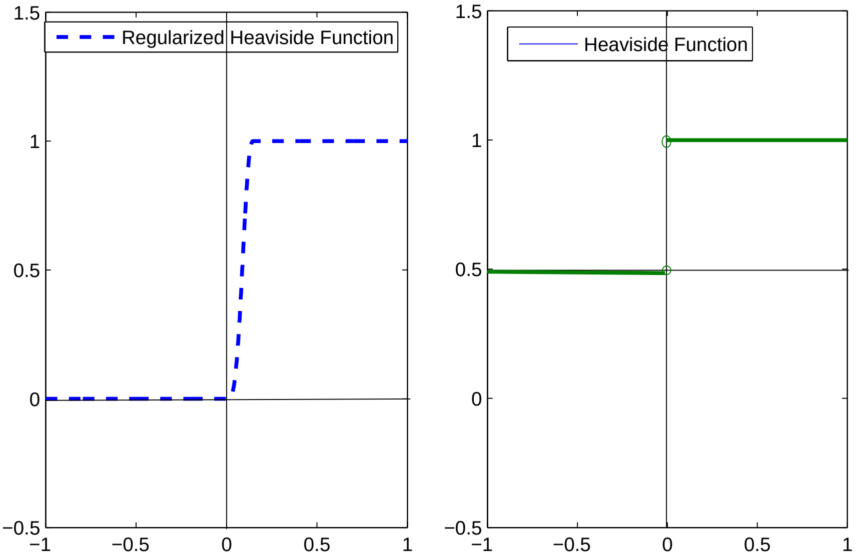


Figure 1.3: **Regularized Heaviside($\mathcal{H}_\epsilon$) function and Heaviside($\mathcal{H}$) function**

1.4 Literature Review

A lot of work has been devoted to the description of the kinetics of metallurgical transformation of steel. The reason why one can change the hardness of steel by thermal treatment is because, on heating, steel which is a mixture of ferrite, pearlite, beanite and marstentite gets transformed to austenite. During cooling, austenite is transformed back into different phases with the actual phase distribution depending on the cooling strategy. For an extensive survey on mathematical models for laser material treatments, we refer to [59]. In [2], [46], the mathematical model for the laser hardening problem is discussed and results on existence of solution, regularity and stability are derived. In [45], 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. The mathematical modeling for austenite-pearlite-martensite phase change in eutectoid carbon steel is presented in [43], [44] and [82]. The model is based on Scheil's additivity rule (see [82]) and the Koistinen-Marburger formula. Existence and uniqueness results are also established in [2], [43]-[48], [84]. In [52], first of all a model is described for case of *(ferrite + pearlite) to austenite* transformation. Later on it is generalized to the case of n phases and several possible transformations.

In [47], the optimal control problem is analyzed and the related error estimates for the state system are derived using proper orthogonal decomposition (POD) Galerkin method. For the purpose of numerical implementation, a penalized problem is also considered. In [84], 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.

During phase transition in the surface hardening of steel process, maintaining the quality of the workpiece is an important and difficult task to avoid melting of surface. In laser surface hardening of steel problem laser energy serves as a control because velocity of the laser beam is tried to be kept a constant to maintain quality of workpiece. This leads to an optimal control problem with laser energy as a control function. The cost functional in the problem is comprised of achievement of desired austenite, a penalization term, which restricts the increment of the temperature beyond melting temperature and the cost occurred due to the laser energy. In [46], the optimal control problem of laser surface hardening is given with pointwise state constraints. Using proper orthogonal decomposition, the optimal control problem is solved using gradient projection method in [47]. More details about the phase transition and simulation of surface heat treatment can be found in [45].

Control problems have been of great importance in many engineering applications. For a systematic study of optimal control problems governed by PDEs, we refer to [58] and [64]. Extensive research in the area of optimal control problems governed by PDEs are being done since the last four decades and is still going on. For optimization problems governed by linear elliptic and parabolic equations, we refer the reader to [24]-[26], [80] and [1], [12], [41], [49], [60], [76], respectively. Some research papers which discuss optimal control problems governed by non-linear elliptic and parabolic equations are [33], [34] and [2], [71], [81], respectively. For a study of applied optimal shape design for fluids, we refer to [69].

Some numerical methods have also been developed in the past for the optimal control problem of laser surface hardening of steel. In [47], the optimal control problem is analyzed and the related error estimates for the state system are derived using proper orthogonal decomposition (POD) Galerkin method. For the purpose of numerical implementation, a penalized problem is also considered. In [84], a variant of non-linear conjugate method has been applied to solve the optimal control problem numerically. For the purpose of discretization in space, finite element method has been applied with piecewise linear approximations and for the time discretization implicit Euler scheme has been used. In [78], Ricatti method has been applied for the optimization of laser surface hardening of steel with checkpointing technique. In [31], the focus is on the discretization of state, adjoint and control variables using a cG method for space discretization and a dG method for time and control discretization. *A priori* error estimate obtained for the control error in this paper are of optimal order. To solve the optimization problem, primal-dual strategy has been applied and non-linear conjugate method has been used to solve it numerically.

The original laser surface hardening of steel problem has a non differentiability in the right hand side function, which is $f_+ = \frac{1}{\tau(\theta)}[a_{eq}(\theta) - a]_+$, where τ, a_{eq} are smooth functions of θ and θ, a are the temperature and the austenite, respectively. To overcome this problem of non differentiability, regularization of f with the help of monotonic regularized Heaviside function has been done in [2], [45], [47], [48] and [84]. Then, existence, uniqueness and stability results are shown for the regularized problem. In [47], existence and uniqueness results are proved for a very specific laser surface hardening of steel problem, although existence and uniqueness for an optimal control problem governed by non linear parabolic problems are shown in [2]. Also, existence, regularity and stability results for the finite element approximation are shown, although the rate of convergence is not established. In this thesis, it is shown that, the solution of the regularized problem converges to the solution of original problem. Also the optimal control for the regularized problem converges strongly to that of the original problem in L^2 sense.

Even though the articles mentioned above discuss the existence results and different methods for the numerical approximation of the problem of laser surface hardening of steel, *a priori* error estimates have not been developed in any of the papers [2], [45], [47], [48] and [84]. It is important to estimate *a priori* error estimates for both semi-discrete and complete discrete problem of laser surface hardening of steel for the purpose of deriving the rates of convergence. For the optimal control problems governed by linear state equations, *a priori* error estimates of the finite element method were established long ago, see for example [24]. It is however, much more difficult to obtain such error estimates for control problems where the state equations are nonlinear or where there are inequality state constraints. For a class of nonlinear optimal control problems with equality constraints, *a priori* error estimates were established in [33]. In [62] and [63], *a priori* error estimates have been developed, respectively, for unconstrained and constrained optimal control problems governed by linear parabolic equation. In this thesis, *a priori* error estimates for the laser surface hardening of steel problem are developed using a cG method for space variables and a dG method for time and control variables.

Since laser surface hardening problem has a nature of irregularity near the heated zone or boundary, application of non-uniform mesh becomes important. Using non-uniform meshes can become quite expensive during the computations. Since near the boundary of the domain in this problem, a more refined mesh is needed, using DGFEM can help in obtaining a refined mesh near the boundary. Although DGFEMs for the numerical approximation of elliptic and parabolic problems were introduced in the early 1970's (see [20] and [86]), DGFEMs for non-linear problems have been studied and used only since the last few years. The motivation for the development of these methods is the flexibility of choosing local approximation spaces. These methods help in obtaining nonuniformity in the construction of meshes for the numerical approximation and are generalization of work by Nitsche [66] for treating Dirichlet boundary condition by introduction of penalty term on boundary. In 1973, Babuška [5] introduced another penalty method to impose the Dirichlet boundary condition weakly. Interior Penalty(IP) methods by Wheeler [86] and Arnold [3] 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. We refer the reader to review article [18] for various motivations in developing the dG methods over the last 30 years. For a review of work on DG methods for elliptic problems, we refer to [4], [70]. [28]-[30] discuss DG methods for quasilinear and strongly non-linear elliptic problems. In [74] and [75], a non symmertric interior penalty DGFEM is analyzed for elliptic and non-linear parabolic problems, respectively. For a detailed description of DGFEM for elliptic and parabolic problems, we refer to [73]. In [50], an *hp*-version of DGFEM has been developed for semilinear parabolic problems. In this dissertation, we develop *a priori* estimates when the space variable is discretized using an *hp*-DGFEM and the time and control variables using a DGFEM.

Adaptive finite element methods (AFEM) are one of the techniques to gain non-uniformity over the heated zone. Under this method, mesh obtained are dependent on the approximate solution and therefore it helps in refinement or coarsening of the mesh at certain areas. In AFEM, the solution is first found on a coarser mesh and then *a posteriori* error estimates are determined using the given data. Based on these estimates, the mesh is refined or coarsened further. For the purpose of deriving *a posteriori* error estimates, different methods such as, heuristic methods [83], local error estimation [83], residual [[6], [21], [22] and [83]] and Dual Weighted Residual (DWR) methods [[8], [13], [14]] have been used in literature. The use of adaptive technique based on *a posteriori* error estimation is well accepted in context of finite element discretization of partial differential equations, see Bank [9], Becker and Rannacher [8], [13], [14], Eriksson and Johnson [21], [22], Verfurth [83]. For *a posteriori* error estimates for elliptic equations using residual type estimator, see for example, [6], [9] and [83]. Estimates using DWR method are developed in [8], [13], [14] and the references cited in there. AFEM for linear parabolic problems are also studied in [21], [22] using residual type estimators and in [8] using DWR type estimators, to mention a few.

In the last few years, the application of these techniques have also been investigated for the optimization problems governed by partial differential equations. Energy type error estimation for the error in the control, state and adjoint variables are developed in Liu, Ma, Tang, Yan [53] and Liu and Yan [55], [56] in the context of optimal control problem governed by elliptic equations subject to pointwise control constraints. These techniques are also been applied to optimal control problems governed by linear parabolic differential equations, see

Liu, Ma, Tang, Yan [54], Liu, Yan [57]. In Picasso [68], an anisotropic error estimate is derived for the error due to the space discretization of an optimal control problems governed by the linear heat equation. These papers use residual type estimators.

For optimal control problems governed by elliptic equations, *a posteriori* error estimators have been developed in [10], [11], [55], [56] and [72]. However, for optimal control problems governed by the heat equation, there are only a few papers with *a posteriori* error estimators, which are based on residual method [54], [57] and [68] and DWR methods [61], [72]. DWR is useful in applications where the error bounds in global norms (energy and L^2 norm) may not be important, but error bounds on some quantity of physical interest are useful.

In Becker and Kapp [10], Becker, Kapp and Rannacher [11] and Becker and Rannacher [13],[14], a general concept for *a posteriori* estimation of the discretization error with respect to the cost functional in the context of optimal control problems is presented. The DWR approach in these papers is to obtain an estimate of the type

$$J(\theta, a, u) - J(\theta_\sigma, a_\sigma, u_\sigma) \equiv \eta_k^J + \eta_h^J + \eta_d^J,$$

where η_k^J, η_h^J and η_d^J are the errors due to the space, time and control discretizations and σ is the discretization parameter representing space, time and control discretizations and J is the cost functional. In this thesis, both residual and DWR estimators are developed for the laser surface hardening of steel problem and numerical results are presented for the methods with theoretical justifications.

1.5 Organization of the Thesis

The organization of the thesis is as follows. Chapter 1 is introductory in nature. In Chapter 2, weak formulation for the original problem (1.3.6)-(1.3.7) and regularized problem (1.3.8)-(1.3.13) are given. Then the existence and uniqueness of solution of the laser surface hardening of steel problem, for a fixed control is shown, using the results already established for the regularized laser surface hardening of steel problem in literature. We have also shown that the solutions of the regularized problem converges to that of the original problem as the regularization parameter $\epsilon \longrightarrow 0$. Numerical results justifying the theoretical order of convergence obtained, in terms of ϵ are presented at the end of the Chapter 3.

In Chapter 3, first of all a semi-discrete and a fully discrete formulation using a cG method for space discretization and a dG method for time and control discretizations for the laser surface hardening problem (1.3.8)-(1.3.13) are presented. Based on these discretizations, *a priori* error estimates have been developed, which are of optimal order, at different discretization levels. Finally, numerical results have been presented in order to justify the theoretical results obtained.

Chapter 4 is devoted to an *hp*-DGFEM, which by nature helps in obtaining local flexibility. At first a weak formulation, semi-discrete and fully discrete formulations are given with an *hp*-DGFEM method applied for space, and a dG method applied for time and control discretizations. Then, *a priori* error estimates of optimal order have been developed. Finally, numerical experiments are presented for the same.

In Chapter 5, two types of adaptive finite element methods namely residual method and DWR method have been developed. AFEM, which ensures a higher density of nodes in certain area of the computational domain, is an important tool to boost the accuracy and efficiency of the finite element discretization. Residual and DWR type error estimators depending on the error in space, time and control variables are developed and a refinement or coarsening of the computational domain based on these estimators is done for an efficient numerical implementation. Numerical results along with comparison of both the methods is also presented.

Conclusion and critical summary of the main results obtained in this thesis are discussed in Chapter 6. Also, some possible future extensions have been described.

Chapter 2

Regularization Error Estimates for the Optimal Control Problem of Laser Surface Hardening of Steel

2.1 Introduction

Consider the optimal control problem of *laser surface hardening of steel* which is governed by a dynamical system consisting of a semilinear parabolic equation and an ordinary differential equation with a non differentiable right hand side function f_+, defined as

$$\begin{aligned} &\min_{u \in U_{ad}} J(\theta, a, u) \text{ subject to} && (2.1.1)\\ \partial_t a = f_+(\theta, a) &= \frac{1}{\tau(\theta)}[a_{eq}(\theta) - a]_+ \quad \text{in } Q, && (2.1.2)\\ a(0) &= 0 \quad \text{in } \Omega, && (2.1.3)\\ \rho c_p \partial_t \theta - \mathcal{K} \triangle \theta &= -\rho L a_t + \alpha u \quad \text{in } Q, && (2.1.4)\\ \theta(0) &= \theta_0 \quad \text{in } \Omega, && (2.1.5)\\ \frac{\partial \theta}{\partial n} &= 0 \quad \text{on } \Sigma, && (2.1.6) \end{aligned}$$

with the notations as defined in Section 1.3. To avoid the numerical and analytic difficulties posed by f_+, this function is regularized using a monotone Heaviside function and the regularized problem has been studied in literature, see [2], [43]-[48], [84].

The regularized *laser surface hardening of steel problem* is given by

$$\min_{u_\epsilon \in U_{ad}} J(\theta_\epsilon, a_\epsilon, u_\epsilon) \text{ subject to} \qquad (2.1.7)$$

$$\begin{aligned}
\partial_t a_\epsilon = f_\epsilon(\theta_\epsilon, a_\epsilon) &= \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)\mathcal{H}_\epsilon(a_{eq}(\theta_\epsilon) - a_\epsilon) \quad \text{in } Q, && (2.1.8)\\
a_\epsilon(0) &= 0 \quad \text{in } \Omega, && (2.1.9)\\
\rho c_p \partial_t \theta_\epsilon - \mathcal{K} \triangle \theta_\epsilon &= -\rho L \partial_t a_\epsilon + \alpha u_\epsilon \quad \text{in } Q, && (2.1.10)\\
\theta_\epsilon(0) &= \theta_0 \quad \text{in } \Omega, && (2.1.11)\\
\frac{\partial \theta_\epsilon}{\partial n} &= 0 \quad \text{on } \Sigma. && (2.1.12)
\end{aligned}$$

where $J(\theta_\epsilon, a_\epsilon, u_\epsilon) = \dfrac{\beta_1}{2}\displaystyle\int_\Omega |a_\epsilon(T) - a_d|^2 dx + \dfrac{\beta_2}{2}\int_0^T \int_\Omega [\theta_\epsilon - \theta_m]_+^2 dxds + \dfrac{\beta_3}{2}\int_0^T |u_\epsilon|^2 ds$. In this chapter, we focus on establishing the convergence of solution of the regularized problem (2.1.7)-(2.1.12) to that of the original problem (2.1.1)-(2.1.6). The estimates, in terms of the regularization parameter ϵ, justify the existence of solution of the original problem.

This chapter is divided into three sections. In Section 2.2, the existence, uniqueness and stability results for solution of regularized *laser surface hardening of steel* problem, which are already there in the literature are presented. Section 2.3 is devoted to estimate the rate at which the solution of the regularized problem converges to that of the original problem for a fixed control u. Also, the convergence of the optimal control of the regularized problem to that of the original problem in the strong sense is proved.

2.2 Existence and Uniqueness Results for the Regularized Laser Surface Hardening of Steel Problem

In this section we present the existence results for the system (2.1.7)-(2.1.12) available in the literature [47]. For continuity of presentation, we prove the relevant results.

Lemma 2.2.1. *[47, page no. 4] With* (A1)-(A6) *holding true, we have:*
(a) if $\theta_\epsilon \in L^1(Q)$, then (2.1.8)-(2.1.9) *has a unique solution satisfying*

$$0 \leq a_\epsilon < 1 \quad a.e. \ in \ Q \tag{2.2.1}$$

and

$$\|a_\epsilon\|_{W^{1,\infty}(I,L^\infty(\Omega))} \leq M \tag{2.2.2}$$

with a constant M independent of θ_ϵ.

(b) if $\theta_{\epsilon,1}, \theta_{\epsilon,2} \in L^{2p}(Q), 1 \leq p < \infty$, and $a_{\epsilon,1}, a_{\epsilon,2}$ are the corresponding solutions to (2.1.8)-(2.1.9)*, then there exists a constant $C > 0$, such that for all $t \in I$,*

$$\|a_{\epsilon,1} - a_{\epsilon,2}\|^{2p}_{L^{2p}(\Omega)} \leq C \int_0^t \|\theta_{\epsilon,1} - \theta_{\epsilon,2}\|^{2p}_{L^{2p}(\Omega)} ds.$$

(c) if $\theta_\epsilon \in L^2(Q)$ and $\{\theta_{\epsilon,k}\}_{k\in\mathbb{N}} \subset L^2(Q)$ with $\lim\limits_{k\to 0} \|\theta_{\epsilon,k} - \theta_\epsilon\|_{L^2(Q)} = 0$, then

$$a_{\epsilon,k} \longrightarrow a_\epsilon \quad \textit{strongly in} \quad C(I, L^2(\Omega)) \bigcap H^1(I, L^2(\Omega)), \tag{2.2.3}$$

where $a_{\epsilon,k}$ and a_ϵ are the solutions to (2.1.8)-(2.1.9) *with temperatures $\theta_{\epsilon,k}$ and θ_ϵ, respectively.*

Proof: Using assumptions (A1)-(A2) and Theorem of Carathéodory, (2.1.8)-(2.1.9) has a unique solution. For (2.2.1), consider

$$\partial_t a_\epsilon \quad = \quad \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)\mathcal{H}_\epsilon(a_{eq}(\theta_\epsilon) - a_\epsilon) \leq \frac{1}{\underline{\tau}}(1 - a_\epsilon).$$

That is, we have

$$\partial_t a_\epsilon + \frac{1}{\underline{\tau}} a_\epsilon \leq \frac{1}{\underline{\tau}}. \tag{2.2.4}$$

Solving (2.2.4) with integrating factor as $e^{\frac{t}{\underline{\tau}}}$, we obtain

$$a_\epsilon(t) < 1.$$

Since $\partial_t a_\epsilon = f_\epsilon(\theta_\epsilon, a_\epsilon) \geq 0$ and $a_\epsilon(0) = 0$, we have $a_\epsilon \geq 0$. Also, using (A1)-(A2), $\|a_\epsilon\|_{L^\infty(I,L^\infty(\Omega))} \leq M, \;\; \|\partial_t a_\epsilon\|_{L^\infty(\Omega)} \leq \|\frac{1}{\tau(\theta)}(a_{eq}(\theta) - a)\|_{L^\infty(\Omega)} \leq \frac{1}{\underline{\tau}}(c_a + M)$, and hence (2.2.2) holds true.

Now we prove (b). Let $\theta_\epsilon = \theta_{\epsilon,1} - \theta_{\epsilon,2}$ and $a_\epsilon = a_{\epsilon,1} - a_{\epsilon,2}$. From (2.1.8), we have

$$\partial_t a_\epsilon = f_\epsilon(\theta_{\epsilon,1}, a_{\epsilon,1}) - f_\epsilon(\theta_{\epsilon,2}, a_{\epsilon,2}). \tag{2.2.5}$$

Multiplying (2.2.5) by $a_\epsilon^{2p-1}, \; 1 \leq p < \infty$, integrating over Q and using Lipschitz continuity of f, we obtain

$$\frac{1}{2p}\|a_\epsilon(t)\|^{2p}_{L^{2p}(\Omega)} \leq C\bigg(\int_0^t\int_\Omega\bigg(|\theta_\epsilon||a_\epsilon|^{2p-1}+|a_\epsilon|^{2p}\bigg)dxds\bigg). \tag{2.2.6}$$

For $\int_0^t\int_\Omega|\theta_\epsilon||a_\epsilon|^{2p-1}dxds$, apply Young's inequality, to obtain,

$$\frac{1}{2p}\|a_\epsilon(t)\|^{2p}_{L^{2p}(\Omega)} \leq C\bigg(\frac{1}{2p}\int_0^t\int_\Omega|\theta_\epsilon|^{2p}dxds+\bigg(1+\frac{2p-1}{2p}\bigg)\int_0^t\int_\Omega|a_\epsilon|^{2p}dxds\bigg). \tag{2.2.7}$$

Applying Gronwall's lemma, we have

$$\|a_{\epsilon,1}-a_{\epsilon,2}\|^{2p}_{L^{2p}(\Omega)} \leq C\int_0^t\|\theta_{\epsilon,1}-\theta_{\epsilon,2}\|^{2p}_{L^{2p}(\Omega)}ds,$$

which proves (b).

To prove (c), let $\{\theta_{\epsilon,k}\}\subset L^2(Q)$ with $\lim\limits_{k\to\infty}\|\theta_{\epsilon,k}-\theta_\epsilon\|_{L^2(Q)}=0$. For $p=1$ in (b), we obtain

$$\|a_{\epsilon,1}-a_{\epsilon,2}\|^2 \leq C\int_0^t\|\theta_{\epsilon,1}-\theta_{\epsilon,2}\|^2ds. \tag{2.2.8}$$

Using (2.2.8), we have

$$\|a_{\epsilon,k}-a_\epsilon\|^2 \leq C\int_0^t\|\theta_{\epsilon,k}-\theta_\epsilon\|^2ds. \tag{2.2.9}$$

In $\|\partial_t a_{\epsilon,k}-\partial_t a_\epsilon\|^2=\|f_\epsilon(\theta_{\epsilon,k},a_{\epsilon,k})-f_\epsilon(\theta_\epsilon,a_\epsilon)\|^2$, we obtain

$$\|\partial_t a_{\epsilon,k}-\partial_t a_\epsilon\|^2 \leq C\bigg(\|\theta_{\epsilon,k}-\theta_\epsilon\|^2+\|a_{\epsilon,k}-a_\epsilon\|^2\bigg). \tag{2.2.10}$$

From (2.2.9) and (2.2.10), we obtain (2.2.3), which proves (c). □

Before proving existence of solution to the problem (2.1.8)-(2.1.12), for a fixed control $u\in U_{ad}$, The next theorem ensures the existence and uniqueness of solution of the system (2.1.8)-(2.1.12) ([47], Theorem 2.1).

Theorem 2.2.1. *Suppose that* (A1)-(A6) *are satisfied. Then, for a fixed control* $u\in U_{ad}$*,* (2.1.8)-(2.1.12) *has a unique solution*

$$(\theta_\epsilon, a_\epsilon) \in H^{1,1}(Q) \times W^{1,\infty}(I; L^\infty(\Omega)),$$

where $H^{1,1} = L^2(I; H^1(\Omega)) \cap H^1(I; L^2(\Omega))$ *is a Hilbert space endowed with a common inner product.*

Proof: Let $\hat{K} = \{\hat{\theta}_\epsilon \in L^2(Q) : \|\hat{\theta}_\epsilon\| \leq \hat{M},\ \hat{\theta}_\epsilon(0) = \theta_0\}$, where $\hat{M}$ is large enough. For a fixed $\hat{\theta}_\epsilon \in \hat{K}$, let a_ϵ be solution to

$$\begin{aligned} \partial_t a_\epsilon &= f_\epsilon(\hat{\theta}_\epsilon, a_\epsilon) \quad \text{in } Q, \qquad (2.2.11) \\ a_\epsilon(0) &= 0 \quad \text{in } \Omega. \qquad (2.2.12) \end{aligned}$$

Since $\hat{\theta}_\epsilon \in L^1(Q)$, using Lemma 2.2.1(a), (2.2.11)-(2.2.12) has a unique solution satisfying $a_\epsilon \in W^{1,\infty}(I, L^\infty(\Omega))$. For this a_ϵ and fixed $u \in U_{ad}$, (2.1.10)-(2.1.12) has a unique solution θ_ϵ. Now, we prove the a priori bound for θ_ϵ. Multiply (2.1.10) by θ_ϵ and integrate over Ω to obtain

$$\frac{\rho c_p}{2}\frac{d}{dt}\|\theta_\epsilon\|^2 + \mathcal{K}\|\nabla\theta_\epsilon\|^2 = -\rho L(\partial_t a_\epsilon, \theta_\epsilon) + (\alpha u, \theta_\epsilon)$$

Integrating from 0 to t, using Cauchy-Schwarz and Young's inequalities, we obtain

$$\|\theta_\epsilon(t)\|^2 + \int_0^t \|\nabla\theta_\epsilon\|^2 ds \leq C\bigg(\|\theta_0\|^2 + \int_0^t (\|\partial_t a_\epsilon\|^2 + |u|^2) ds + \int_0^t \|\theta_\epsilon\|^2 ds\bigg). \qquad (2.2.13)$$

Using Gronwall's Lemma, we obtain

$$\|\theta_\epsilon(t)\|^2 \leq C\bigg(\|\theta_0\|^2 + \int_0^t (\|\partial_t a_\epsilon\|^2 + |u|^2) ds\bigg). \qquad (2.2.14)$$

Using (2.2.14) on the right hand side (2.2.13), we obtain

$$\|\theta_\epsilon(t)\|^2_{L^2(I,H^1(\Omega))} \leq C\bigg(\|\theta_0\|^2 + \int_0^t (\|\partial_t a_\epsilon\|^2 + |u|^2) ds\bigg). \qquad (2.2.15)$$

Now, multiply (2.1.10) with $\partial_t\theta_\epsilon$, integrate over Ω, use Cauchy-Schwarz and Young's inequalities and then integrate from 0 to t, to obtain

$$\int_0^t \|\partial_t \theta_\epsilon\|^2 ds + \|\nabla \theta_\epsilon(t)\|^2 \ \leq \ C\Big(\|\nabla\theta_0\|^2 + \int_0^t \Big(\|\partial_t a_\epsilon\|^2 + |u|^2\Big)ds\Big). \qquad (2.2.16)$$

Using (2.2.2), $u \in U_{ad}$, using (A3), (2.2.14), (2.2.15) and (2.2.16) we have the results

$$\|\theta_\epsilon\|_{L^\infty(I,H^1(\Omega))} \leq C \qquad (2.2.17)$$

and

$$\|\theta_\epsilon\|_{H^{1,1}} \leq C. \qquad (2.2.18)$$

Define an operator $F : \hat{K} \subset L^2(Q) \longrightarrow L^2(Q)$ by $F(\hat{\theta}_\epsilon) = \theta_\epsilon$. F is well-defined. From (2.2.18), we have $\theta_\epsilon \in H^{1,1}(Q)$ with $\theta_\epsilon(0) = \theta_0$ and hence $F(\hat{K}) \subset \hat{K}$, for $\hat{M}$ large enough.

Now, we want to show that F is a contraction map. Let $\hat{\theta}_{\epsilon,i} \in \hat{K}$, for $i = 1, 2$, $\theta_{\epsilon,i} = F(\hat{\theta}_{\epsilon,i})$ and $\hat{\theta}_\epsilon = \hat{\theta}_{\epsilon,1} - \hat{\theta}_{\epsilon,2}$. Then $\theta_\epsilon = \theta_{\epsilon,1} - \theta_{\epsilon,2}$ solves the system

$$\begin{aligned}
\rho c_p \partial_t \theta_\epsilon - \mathcal{K} \triangle \theta_\epsilon &= -\rho L(f(\hat{\theta}_{\epsilon,1}, a_{\epsilon,1}) - f(\hat{\theta}_{\epsilon,2}, a_{\epsilon,2})) \text{ in Q}, && (2.2.19)\\
\frac{\partial \theta_\epsilon}{\partial n} &= 0 \text{ on } \Sigma, && (2.2.20)\\
\theta_\epsilon(0) &= 0 \text{ in } \Omega. && (2.2.21)
\end{aligned}$$

Multiplying (2.2.19) with θ_ϵ, integrating over Ω, using Cauchy-Schwarz and Young's inequalities and integrating from 0 to t, we obtain

$$\|\theta_\epsilon\|^2 + \int_0^t \| \bigtriangledown \theta_\epsilon\|^2 ds \ \leq C \int_0^t \Big(\|f(\hat{\theta}_{\epsilon,1}, a_{\epsilon,1}) - f(\hat{\theta}_{\epsilon,2}, a_{\epsilon,2})\|^2 + \|\theta_\epsilon\|^2\Big) ds. \qquad (2.2.22)$$

Using Proposition 2.2.1 and Lemma 2.2.1(b), we obtain

$$\|\theta_\epsilon\|^2 + \int_0^t \| \bigtriangledown \theta_\epsilon\|^2 ds \ \leq C \int_0^t \Big(\|\hat{\theta}_\epsilon\|^2 + \|\theta_\epsilon\|^2\Big) ds.$$

Finally, use Gronwall's lemma and integrate once more from 0 to t to obtain

$$\int_0^t \|\theta_\epsilon\|^2 ds \ \leq \ tC(T) \int_0^T \|\hat{\theta}_\epsilon\|^2 ds.$$

Hence for $T^+ < T$, small enough, F is a contraction on $L^2(0, T^+; L^2(\Omega))$. Since F is also defined from $\hat{K}$ to itself, we can apply Banach fixed point theorem. This implies that F has a unique fixed point θ_ϵ, which is a local solution to (2.1.8)-(2.1.12). Using bootstrap argument and Lemma (2.2.18), we can extend the local solution θ_ϵ to the whole interval I. This completes the proof. □

Proposition 2.2.1. *Due to* (A1)-(A2) *and the definition of regularized Heaviside function, there exists a constant $c_f > 0$ independent of θ_ϵ and a_ϵ such that*

$$\max(\|f_\epsilon(\theta_\epsilon, a_\epsilon)\|_{L^\infty(Q)}, \|\frac{\partial f_\epsilon}{\partial a}(\theta_\epsilon, a_\epsilon)\|_{L^\infty(Q)}, \|\frac{\partial f_\epsilon}{\partial \theta}(\theta_\epsilon, a_\epsilon)\|_{L^\infty(Q)}) \leq c_f$$

for all $(\theta_\epsilon, a_\epsilon) \in L^2(Q) \times L^\infty(Q)$.

Proof:

$$f(\theta_\epsilon, a_\epsilon) = \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)H(a_{eq}(\theta_\epsilon) - a_\epsilon) \leq \frac{1}{\underline{\tau}} c_a$$

as from assumption (A1), $\|a_{eq}\| \leq c_a$, from assumption (A2), $\tau(\theta_\epsilon) \geq \underline{\tau}$, by definition of regularized Heaviside function $H \leq 1$ and $a_\epsilon < 1$ from Lemma 2.2.1. For the boundedness of f_{a_ϵ}, we have

$$\begin{aligned} |f_a(\theta_\epsilon, a_\epsilon)| &= |\frac{1}{\tau(\theta_\epsilon)} H(a_{eq}(\theta_\epsilon) - a_\epsilon) + \frac{1}{\tau(\theta_\epsilon)}(a_{eq}(\theta_\epsilon) - a_\epsilon)H'(a_{eq}(\theta_\epsilon) - a_\epsilon)| \\ &\leq \frac{1}{\underline{\tau}} + \frac{1}{\underline{\tau}} \epsilon C \frac{1}{\epsilon} \\ &\leq C\frac{1}{\underline{\tau}}. \end{aligned}$$

The boundedness of f_{θ_ϵ} can be proved similarly. □

For $u_\epsilon^* \in U_{ad}$, let $(\theta_\epsilon^*, a_\epsilon^*)$ be the solution of (2.1.8)-(2.1.12). The existence of a unique solution to the state equation (2.1.8)-(2.1.12) ensures the existence of a control-to-state mapping $u_\epsilon \mapsto (\theta_\epsilon, a_\epsilon) = (\theta_\epsilon(u_\epsilon), a_\epsilon(u_\epsilon))$ through (2.1.8)-(2.1.12). By means of this mapping, we introduce the reduced cost functional $j_\epsilon : U_{ad} \longrightarrow \mathbb{R}$ as

$$j_\epsilon(u_\epsilon) = J(\theta_\epsilon(u_\epsilon), a_\epsilon(u_\epsilon), u_\epsilon). \tag{2.2.23}$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u_\epsilon \in U_{ad}} j_\epsilon(u_\epsilon). \tag{2.2.24}$$

Theorem 2.2.2. *Let* (A1)-(A6) *hold true. Then there exists at least one optimal control* u_ϵ^* *to* (2.1.7)-(2.1.12).

Proof: Let $l_\epsilon = \inf\limits_{u_\epsilon \in U_{ad}} j_\epsilon(u_\epsilon)$ and $\{u_{\epsilon,n}\} \subset U_{ad}$ be a minimizing sequence such that

$$j_\epsilon(u_{\epsilon,n}) \longrightarrow l_\epsilon \text{ in } \mathbb{R}. \tag{2.2.25}$$

Since U_{ad} is bounded, the sequence $\{u_{\epsilon,n}\}$ is bounded uniformly in $L^2(I)$. Therefore, one can extract a subsequence $\{u_{\epsilon,n}\}$(say), such that

$$u_{\epsilon,n} \longrightarrow u_\epsilon^* \text{ weakly in } L^2(I).$$

Since the admissible space U_{ad} is closed and convex, we have $u_\epsilon^* \in U_{ad}$. Corresponding to each $u_{\epsilon,n}$, we have $(\theta_{\epsilon,n}, a_{\epsilon,n}) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$ satisfying (2.1.8)-(2.1.12). Also, $\theta_{\epsilon,n} \in L^\infty(I, V)$. Therefore, one can extract a subsequence $\{(\theta_{\epsilon,n}, a_{\epsilon,n})\}$ (see Corollary 4 in [77]), such that

$$\theta_{\epsilon,n} \quad \longrightarrow \quad \theta_\epsilon^* \text{ weakly in } H^{1,1}$$

$$\theta_{\epsilon,n} \quad \longrightarrow \quad \theta_\epsilon^* \text{ strongly in } C(I, L^2(\Omega))$$

From Lemma 2.2.1, we obtain

$$\begin{aligned} a_{\epsilon,n} \quad &\longrightarrow \quad a_\epsilon^* \text{ strongly in } L^\infty(I, L^2(\Omega)) \\ a_{\epsilon,n} \quad &\longrightarrow \quad a_\epsilon^* \text{ weak-* in } W^{1,\infty}(I; L^\infty(\Omega)). \end{aligned}$$

Also, $f_\epsilon(\theta_{\epsilon,n}, a_{\epsilon,n})$ converges strongly to $f_\epsilon(\theta_\epsilon, a_\epsilon)$ using Proposition (2.2.1). Letting $n \to \infty$

in

$$\begin{aligned}(\partial_t a_{\epsilon,n}, w) &= (f_\epsilon(\theta_{\epsilon,n}, a_{\epsilon,n}), w) \qquad \forall w \in V,\\ a_{\epsilon,n}(0) &= 0,\\ \rho c_p(\partial_t \theta_{\epsilon,n}, v) + \mathcal{K}(\bigtriangledown \theta_{\epsilon,n}, \bigtriangledown v) &= -\rho L(\partial_t a_{\epsilon,n}, v) + (\alpha u_{\epsilon,n}, v) \qquad \forall v \in V,\\ \theta_{\epsilon,n}(0) &= \theta_0,\end{aligned}$$

we obtain $(\theta^*_\epsilon, a^*_\epsilon)$ as unique solution of (2.1.8)-(2.1.12). This shows that $(\theta^*_\epsilon, a^*_\epsilon, u^*_\epsilon)$ is the admissible solution. Now we show that it is an optimal solution. Since j is lower semi-continuous, we have

$$j_\epsilon(u^*_\epsilon) \leq \liminf_{n\to\infty} j_\epsilon(u_{\epsilon,n})$$

and using (2.2.25), we obtain

$$j_\epsilon(u^*_\epsilon) \leq l_\epsilon$$

which implies that u^*_ϵ is the minimizer of the cost functional j_ϵ. Therefore, $(\theta^*_\epsilon, a^*_\epsilon, u^*_\epsilon)$ is an optimal solution. □

2.3 Existence and Uniqueness Results for the Laser Surface Hardening of Steel Problem

In this section, first we describe the weak formulation corresponding to (2.1.1)-(2.1.6) and (2.1.7)-(2.1.12). Let $X = \{v \in L^2(I;V) : v_t \in L^2(I;V^*)\}$, where $V = H^1(\Omega)$ and $Y = H^1(I;L^2(\Omega))$. Together with $H = L^2(\Omega)$, the Hilbert space V and its dual V^* build a Gelfand triple

$$V \hookrightarrow H \equiv H^* \hookrightarrow V^*. \tag{2.3.1}$$

In (2.3.1), V is densely embedded in H and H^* is densely embedded in V^*. Additionally, the corresponding injections are continuous. The duality pairing between V and its dual V^* is denoted by $\langle\cdot,\cdot\rangle_{V^*\times V}$. Let $i : V \to H$ be the injection of V into H. Then, its dual $i^* : H \equiv H^* \to V^*$ is an injection of H^* into V^*. Because of the definition of i^*, every element $h \in H \equiv H^*$ can be understood as linear continuous functional on V by virtue of the identity

$$\langle i^*(h), v\rangle_{V^* \times V} = (h, i(v))_H \quad \forall v \in V,$$

where $(\cdot,\cdot)_H$ is the inner product of H. Since H^* is densely embedded in V^*, every functional $\langle v^*, \cdot\rangle_{V^* \times V}$ can be uniformly approximated by inner product $(h, i(\cdot))_H$.

For $\omega \subseteq \Omega$, let $(\cdot,\cdot)_\omega$ (resp. $(\cdot,\cdot)$) and $\|\cdot\|_\omega$ (resp. $\|\cdot\|$) denote the inner product and norm in $L^2(\omega)$(resp. $L^2(\Omega)$). The inner product and norm in $L^2(\mathbb{J}), \mathbb{J} \subseteq I$ are denoted by $(\cdot,\cdot)_{L^2(\mathbb{J})}$ and $\|\cdot\|_{L^2(\mathbb{J})}$. Also, let $(\cdot,\cdot)_{\mathbb{J},\omega}$ and $\|\cdot\|_{\mathbb{J},\omega}$ denote the inner product and norm in $L^2(I, L^2(\omega))$.

Weak Formulation

The weak formulation corresponding to (2.1.1)-(2.1.6) is given by

$$\min_{u \in U_{ad}} J(\theta, a, u) \text{ subject to} \tag{2.3.2}$$

$$(\partial_t a, w) = (f_+(\theta, a), w) \quad \forall w \in H, \text{ a.e. in } I, \tag{2.3.3}$$

$$a(0) = 0, \tag{2.3.4}$$

$$\rho c_p(\partial_t \theta, v) + \mathcal{K}(\bigtriangledown\theta, \bigtriangledown v) = -\rho L(\partial_t a, v) + (\alpha u, v) \quad \forall v \in V, \text{ a.e. in } I, \tag{2.3.5}$$

$$\theta(0) = \theta_0, \tag{2.3.6}$$

where $f_+(\theta, a) = \frac{1}{\tau(\theta)}(a_{eq}(\theta) - a)\mathcal{H}(a_{eq}(\theta) - a)$. Similarly, the weak formulation corresponding to (2.1.7)-(2.1.12) is given by

$$\min_{u_\epsilon \in U_{ad}} J(\theta_\epsilon, a_\epsilon, u_\epsilon) \text{ subject to} \tag{2.3.7}$$

$$(\partial_t a_\epsilon, w) = (f_\epsilon(\theta_\epsilon, a_\epsilon), w) \quad \forall w \in H, \text{ a.e. in } I, \tag{2.3.8}$$

$$a_\epsilon(0) = 0, \tag{2.3.9}$$

$$\rho c_p(\partial_t \theta_\epsilon, v) + \mathcal{K}(\bigtriangledown\theta_\epsilon, \bigtriangledown v) = -\rho L(\partial_t a_\epsilon, v) + (\alpha u_\epsilon, v) \; \forall v \in V, \text{ a.e. in } I, \tag{2.3.10}$$

$$\theta_\epsilon(0) = \theta_0, \tag{2.3.11}$$

where $f_\epsilon(\theta, a) = \frac{1}{\tau(\theta)}(a_{eq}(\theta) - a)\mathcal{H}_\epsilon(a_{eq}(\theta) - a)$. It is already shown in Section 2.2 that (2.3.7)-(2.3.11) has a solution $u_\epsilon^* \in U_{ad}$.

The main focus of this section is to show that the optimal control of the regularized problem (2.3.7)-(2.3.11) converges to that of the original problem (2.3.2)-(2.3.6). The whole procedure of establishing the convergence is divided into a few steps. First of all, we will show that for a fixed control $u \in U_{ad}$, the system (2.3.3)-(2.3.6) has a unique solution (θ^*, a^*) and that the solution of the regularized problem (2.3.8)-(2.3.11), for the same control variable u, converges to that of the original problem (2.3.3)-(2.3.6) with order of convergence $\mathcal{O}(\epsilon)$. Using this result it will be established that the optimal control of (2.3.7)-(2.3.11) converges to that of (2.3.2)-(2.3.6).

Convergence Analysis

Theorem 2.3.1. *Let the assumptions* (A1)-(A6) *hold true. Then, for a fixed control* $u \in U_{ad}$*, there exists a unique solution* $(\theta, a) \in X \times Y$ *to* (2.3.3)-(2.3.6) *and for all* $\epsilon \in (0,1)$*,* $t \in I$*, we have*

$$\|a(t) - a_\epsilon(t)\| + \|\theta(t) - \theta_\epsilon(t)\| \leq C(\Omega, T)\epsilon, \tag{2.3.12}$$

where $C(\Omega, T)$ *is a positive constant and* $(\theta_\epsilon, a_\epsilon)$ *is the solution to the problem* (2.3.8)-(2.3.11) *for a fixed control* $u \in U_{ad}$*. Moreover,*

$$\|\theta - \theta_\epsilon\|_{L^2(I, H^1(\Omega))} \leq C(\Omega, T)\epsilon. \tag{2.3.13}$$

Proof: From Theorem 2.2.1, we have $(\theta_\epsilon, a_\epsilon) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$ and from (2.2.17) we have, $\{\theta_\epsilon\}$ is uniformly bounded in $L^\infty(I, V)$. Since V is compactly imbedded in $L^2(\Omega)$, we obtain

$$\theta_\epsilon \longrightarrow \theta \text{ strongly in } C(I, L^2(\Omega)), \tag{2.3.14}$$

$$\begin{aligned} \theta_\epsilon &\longrightarrow \theta \text{ weakly in } H^{1,1}, &(2.3.15)\\ a_\epsilon &\longrightarrow a \text{ weak}^* \text{ in } W^{1,\infty}(I, L^\infty(\Omega)). &(2.3.16)\end{aligned}$$

For $\theta \in C(I, L^2(\Omega))$ and f_+ being Lipschitz continuous, (2.3.3)-(2.3.4) has a unique solution a (say), by theorem of Carathéodary. Now subtracting (2.3.8) from (2.3.3), putting $w = a - a_\epsilon$

and using Cauchy-Schwarz and Young's inequalities, we obtain

$$\frac{d}{dt}\|a - a_\epsilon\|^2 \leq \|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_+(\theta, a)\|^2 + \|a - a_\epsilon\|^2.$$

Now integrating from 0 to t, we obtain

$$\|(a - a_\epsilon)(t)\|^2 \leq C\biggl(\int_0^t \|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_+(\theta, a)\|^2 ds + \int_0^t \|a - a_\epsilon\|^2 ds\biggr). \tag{2.3.17}$$

Note that using triangle inequality,

$$\|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_+(\theta, a)\|^2 \leq C\biggl(\|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_\epsilon(\theta, a)\|^2 + \|f_\epsilon(\theta, a) - f_+(\theta, a)\|^2\biggr). \tag{2.3.18}$$

For the first term on the right hand side of (2.3.18), use Proposition (2.2.1) to obtain

$$\|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_\epsilon(\theta, a)\|^2 \leq C\biggl(\|\theta_\epsilon - \theta\|^2 + \|a_\epsilon - a\|^2\biggr). \tag{2.3.19}$$

For the second term on the right hand side of (2.3.18), using the assumption (A2), we obtain

$$\begin{aligned}\|f_\epsilon(\theta, a) - f_+(\theta, a)\|^2 &\leq \frac{1}{\mathcal{T}}\|(a_{eq}(\theta) - a)\mathcal{H}(a_{eq}(\theta) - a) - (a_{eq}(\theta) - a)\mathcal{H}_\epsilon(a_{eq}(\theta) - a)\|^2,\\ &= \frac{1}{\mathcal{T}}\int_\Omega (a_{eq}(\theta) - a)^2(\mathcal{H}(a_{eq}(\theta) - a) - \mathcal{H}_\epsilon(a_{eq}(\theta) - a))^2 dx.\end{aligned}$$

Let $\Omega_1 = \{x \in \bar{\Omega} : a_{eq} - a \leq 0 \text{ or } a_{eq} - a \geq \epsilon\}$ and $\Omega_2 = \{x \in \bar{\Omega} : 0 < a_{eq} - a < \epsilon\}$. Since $\bar{\Omega} = \Omega_1 \bigcup \Omega_2$, we arrive at

$$\begin{aligned}\|f_\epsilon(\theta, a) - f_+(\theta, a)\|^2 &\leq \frac{1}{\mathcal{T}}\int_{\Omega_1} (a_{eq}(\theta) - a)^2(\mathcal{H}(a_{eq}(\theta) - a) - \mathcal{H}_\epsilon(a_{eq}(\theta) - a))^2 dx\\ &\quad + \frac{1}{\mathcal{T}}\int_{\Omega_2} (a_{eq}(\theta) - a)^2(\mathcal{H}(a_{eq}(\theta) - a) - \mathcal{H}_\epsilon(a_{eq}(\theta) - a))^2 dx.\end{aligned}$$

From Figure 1.3 in Chapter 1, we obtain

$$\|f_\epsilon(\theta, a) - f_+(\theta, a)\|^2 \leq \frac{1}{\mathcal{T}}\int_{\Omega_2} \epsilon^2 dx \leq C(\Omega)\epsilon^2. \tag{2.3.20}$$

Substituting (2.3.20) in (2.3.18), we obtain

$$\|f_\epsilon(\theta_\epsilon, a_\epsilon) - f_+(\theta, a)\|^2 \leq C\bigg(\|\theta_\epsilon - \theta\|^2 + \|a_\epsilon - a\|^2 + \epsilon^2\bigg). \tag{2.3.21}$$

Substituting (2.3.21) in (2.3.17), we find using Gronwall's lemma that

$$\|a - a_\epsilon\|^2 \leq C(\Omega, T)\bigg(\int_0^t \|\theta - \theta_\epsilon\|^2 ds + \epsilon^2\bigg)$$

Using (2.3.14), we arrive at

$$a_\epsilon \longrightarrow a \text{ strongly in } L^\infty(I, L^2(\Omega)). \tag{2.3.22}$$

From (2.3.21), using (2.3.14), (2.3.22), we obtain as $\epsilon \longrightarrow 0$

$$f_\epsilon(\theta_\epsilon, a_\epsilon) \longrightarrow f_+(\theta, a) \ \text{ in } L^2(\Omega) \times L^2(\Omega). \tag{2.3.23}$$

Now letting $\epsilon \to 0$ in (2.3.8)-(2.3.11) and using (2.3.14)-(2.3.16), (2.3.22), (2.3.23), we obtain the existence of solution of (2.3.3)-(2.3.6). For proving the uniqueness we proceed as follows. If possible, let (θ_1, a_1) and (θ_2, a_2) be two different solutions of (2.3.3)-(2.3.6). Therefore, from (2.3.5), we have

$$\rho c_p(\partial_t(\theta_1 - \theta_2), v) + \mathcal{K}(\nabla(\theta_1 - \theta_2), \nabla v) = -\rho L(f_+(\theta_1, a_1) - f_+(\theta_2, a_2), v). \tag{2.3.24}$$

Setting $v = \theta_1 - \theta_2$ in (2.3.24), use Young's inequality to obtain

$$\frac{d}{dt}\|\theta_1 - \theta_2\|^2 + \|\nabla(\theta_1 - \theta_2)\|^2 \leq C\bigg(\|f_+(\theta_1, a_1) - f_+(\theta_2, a_2)\|^2 + \|\theta_1 - \theta_2\|^2\bigg). \tag{2.3.25}$$

Similarly from (2.3.3), we arrive at

$$\frac{d}{dt}\|a_1 - a_2\|^2 \ \ \leq \ \ C\bigg(\|f_+(\theta_1, a_1) - f_+(\theta_2, a_2)\|^2 + \|a_1 - a_2\|^2\bigg). \tag{2.3.26}$$

Adding (2.3.25) and (2.3.26), using Lipschitz continuity of the functions a_{eq}, f_+, integrating from 0 to T and finally using Gronwall's lemma, we obtain

$$\|\theta_1 - \theta_2\|^2 + \|a_1 - a_2\|^2 \ \ \leq \ \ 0,$$

which proves uniqueness.

To prove (2.3.12), subtract (2.3.8) from (2.3.3), put $w = a - a_\epsilon$, use Cauchy-Schwarz and Young's inequalities to find that

$$\frac{d}{dt}\|a - a_\epsilon\|^2 \;\; \leq \;\; \Big(\|f_+(\theta, a) - f_\epsilon(\theta_\epsilon, a_\epsilon)\|^2 + \|a - a_\epsilon\|^2\Big). \tag{2.3.27}$$

Now integrating from 0 to t, we obtain

$$\|a(t) - a_\epsilon(t)\|^2 \;\; \leq \;\; \Big(\int_0^t \|f_+(\theta, a) - f_\epsilon(\theta_\epsilon, a_\epsilon)\|^2 ds + \int_0^t \|a - a_\epsilon\|^2 ds\Big). \tag{2.3.28}$$

Similarly, for a fixed $u \in U_{ad}$ and $u_\epsilon = u$, subtract (2.3.10) from (2.3.5), substitute $v = \theta - \theta_\epsilon$, use (2.3.3), (2.3.8) and integrate from 0 to t to arrive at

$$\begin{aligned}\|\theta(t) - \theta_\epsilon(t)\|^2 + \int_0^t \|\nabla(\theta - \theta_\epsilon)\|^2 ds \;\; &\leq \;\; C\Big(\int_0^t \|f_+(\theta, a) - f_\epsilon(\theta_\epsilon, a_\epsilon)\|^2 ds \\ &\quad + \int_0^t \|\theta - \theta_\epsilon\|^2 ds\Big).\end{aligned} \tag{2.3.29}$$

Adding (2.3.28) and (2.3.29), we find that

$$\begin{aligned}\|a(t) - a_\epsilon(t)\|^2 + \|\theta(t) - \theta_\epsilon(t)\|^2 \;\; &\leq \;\; C\Big(\int_0^t \|f_+(\theta, a) - f_\epsilon(\theta_\epsilon, a_\epsilon)\|^2 ds + \int_0^t \|a - a_\epsilon\|^2 ds \\ &\quad + \int_0^t \|\theta - \theta_\epsilon\|^2 ds\Big).\end{aligned} \tag{2.3.30}$$

Using (2.3.21), we now obtain

$$\|a(t) - a_\epsilon(t)\|^2 + \|\theta(t) - \theta_\epsilon(t)\|^2 \;\; \leq \;\; C(\Omega, T)\Big(\epsilon^2 + \int_0^t (\|\theta - \theta_\epsilon\|^2 + \|a - a_\epsilon\|^2) ds\Big).$$

Using Gronwall's lemma, we arrive at

$$\|a(t) - a_\epsilon(t)\| + \|\theta(t) - \theta_\epsilon(t)\| \;\; \leq \;\; C(\Omega, T)\epsilon. \tag{2.3.31}$$

Using (2.3.21) and (2.3.31) in the right hand side of (2.3.29), we obtain

$$\|\theta - \theta_\epsilon\|_{L^2(I, H^1(\Omega))} \;\; \leq \;\; C(\Omega, T)\epsilon. \quad \Box$$

We now show existence of solution of the optimal control problem (2.3.2)-(2.3.6). For $u^* \in U_{ad}$, let (θ^*, a^*) be the solution of (2.3.3)-(2.3.6). The existence of a unique solution to the state equations (2.3.3)-(2.3.6) ensure the existence of a control-to-state mapping $u \mapsto (\theta, a) = (\theta(u), a(u))$ through (2.3.3)-(2.3.6). By means of this mapping, we introduce the reduced cost functional $j : U_{ad} \longrightarrow \mathbb{R}$ as

$$j(u) = J(\theta(u), a(u), u). \tag{2.3.32}$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u \in U_{ad}} j(u). \tag{2.3.33}$$

Theorem 2.3.2. (2.3.2)-(2.3.6) *has at least one solution* $(\theta^*, a^*, u^*) \in X \times X \times U_{ad}$.

Proof: Let $l = \inf\limits_{u \in U_{ad}} j(u)$ and $\{u_n\} \subset U_{ad}$ be a minimizing sequence such that

$$j(u_n) \longrightarrow l \text{ in } \mathbb{R}. \tag{2.3.34}$$

Since U_{ad} is bounded, the sequence $\{u_n\}$ is bounded uniformly in $L^2(I)$. Therefore, one can extract a subsequence $\{u_n\}$(say), such that

$$u_n \longrightarrow u^* \text{ weakly in } L^2(I).$$

Since the admissible space U_{ad} is closed and convex, we have $u^* \in U_{ad}$. Corresponding to each u_n, we have $(\theta_n, a_n) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$ satisfying (2.3.3)-(2.3.6), also $\theta_n \in L^\infty(I, V)$. Therefore, we can extract a subsequence $\{(\theta_n, a_n)\}$(say), such that

$$\begin{aligned}
\theta_n &\longrightarrow \theta^* \text{ weakly in } H^{1,1} \\
\theta_n &\longrightarrow \theta^* \text{ strongly in } C(I, L^2(\Omega)) \\
a_n &\longrightarrow a^* \text{ weak}^* \text{ in } W^{1,\infty}(I, L^\infty(\Omega))
\end{aligned}$$

$$a_n \longrightarrow a^* \text{ strongly in } L^\infty(I, L^2(\Omega)).$$

Now letting $n \to \infty$ in

$$
(\partial_t a_n, w) \;=\; (f_+(\theta_n, a_n), w) \qquad \forall w \in V,
$$

$$
\begin{aligned}
a_n(0) &= 0, \\
\rho c_p(\partial_t \theta_n, v) + \mathcal{K}(\nabla \theta_n, \nabla v) &= -\rho L(\partial_t a_n, v) + (\alpha u_n, v) \qquad \forall v \in V, \\
\theta_n(0) &= \theta_0,
\end{aligned}
$$

we obtain (θ^*, a^*) as the unique solution of (2.3.3)-(2.3.6). This shows that (θ^*, a^*, u^*) is the admissible solution. Now we show that it is an optimal solution. Since j is lower semi-continuous, it follows that

$$
j(u^*) \leq \liminf_{n \to \infty} j(u_n)
$$

and using (2.3.34), we obtain

$$
j(u^*) \leq l
$$

which implies that u^* is the minimizer of the cost functional j. Therefore, (θ^*, a^*, u^*) is an optimal solution. This completes the rest of the proof. □

Convergence of the Control Function

Theorem 2.3.3. *Let u^*_ϵ be the optimal control of* (2.3.7)-(2.3.11)*, for* $0 < \epsilon < 1$*. Then,* $\lim\limits_{\epsilon \to 0} u^*_\epsilon = u^*$ *exists in $L^2(I)$ and u^* is an optimal control of* (2.3.2)-(2.3.6)*.*

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

$$
\|u^*_\epsilon\|_{L^2(I)} \leq C, \qquad 0 < \epsilon < 1,
$$

that is, $\{u^*_\epsilon\}_{0<\epsilon<1}$ is uniformly bounded in $L^2(I)$. Thus, it is possible to extract a subsequence, say $\{u^*_\epsilon\}_{0<\epsilon<1}$ in $L^2(I)$ such that

$$
u^*_\epsilon \longrightarrow u^* \;\text{ weakly in } L^2(I). \tag{2.3.35}
$$

Since $U_{ad} \subset L^2(I)$ is a closed set, we have $u^* \in U_{ad}$. Now corresponding to each u^*_ϵ there exists solution $(\theta_\epsilon^*, a_\epsilon^*)$ to (2.3.8)-(2.3.11). Also from Theorem 2.2.1, we have

$$\begin{aligned}
\theta_\epsilon^* &\longrightarrow \theta^* \text{ weakly in } H^{1,1}, && (2.3.36)\\
\theta_\epsilon^* &\longrightarrow \theta^* \text{ strongly in } C(I, L^2(\Omega)), && (2.3.37)\\
a_\epsilon^* &\longrightarrow a^* \text{ weak}^* \text{ in } W^{1,\infty}(I, L^\infty(\Omega)), && (2.3.38)\\
a_\epsilon^* &\longrightarrow a^* \text{ strongly in } L^\infty(I, L^2(\Omega)). && (2.3.39)
\end{aligned}$$

Now passing limit as $\epsilon \to 0$ and using (2.3.36)-(2.3.39) in

$$\begin{aligned}
(\partial_t a_\epsilon^*, w) &= (f_\epsilon(\theta_\epsilon^*, a_\epsilon^*), w) \quad \forall w \in H, \text{ a.e. in } I,\\
a_\epsilon^*(0) &= 0,\\
\rho c_p(\partial_t \theta_\epsilon^*, v) + \mathcal{K}(\nabla\theta_\epsilon^*, \nabla v) &= -\rho L(\partial_t a_\epsilon^*, v) + (\alpha u_\epsilon^*, v) \quad \forall v \in V, \text{ a.e. in } I,\\
\theta_\epsilon^*(0) &= \theta_0,
\end{aligned}$$

we obtain that (u^*, θ^*, a^*) is an admissible solution for the optimal control problem (2.3.2)-(2.3.6). 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 (2.3.2)-(2.3.6). Consider the auxiliary problem

$$\begin{aligned}
(\partial_t a_\epsilon, w) &= (f_\epsilon(\theta_\epsilon, a_\epsilon), w) \quad \forall w \in H, \text{ a.e. in } I, && (2.3.40)\\
a_\epsilon(0) &= 0, && (2.3.41)\\
\rho c_p(\partial_t \theta_\epsilon, v) + \mathcal{K}(\nabla\theta_\epsilon, \nabla v) &= -\rho L(\partial_t a_\epsilon, v) + (\alpha \bar{u}^*, v) \quad \forall v \in V, \text{ a.e. in } I, && (2.3.42)\\
\theta_\epsilon(0) &= \theta_0. && (2.3.43)
\end{aligned}$$

Then by Theorem 2.2.1, there exists a solution to (2.3.40)-(2.3.43), say $(\bar{\theta}_\epsilon, \bar{a}_\epsilon) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$. Similar to (2.3.36)-(2.3.39), we arrive at

$$\begin{aligned}
\bar{\theta}_\epsilon &\longrightarrow \bar{\theta} \text{ weakly in } H^{1,1}, && (2.3.44)\\
\bar{\theta}_\epsilon &\longrightarrow \bar{\theta} \text{ strongly in } C(I, L^2(\Omega)), && (2.3.45)\\
\bar{a}_\epsilon &\longrightarrow \bar{a} \text{ weakly in } W^{1,\infty}(I, L^\infty(\Omega)), && (2.3.46)\\
\bar{a}_\epsilon &\longrightarrow \bar{a} \text{ strongly in } L^\infty(I, L^2(\Omega)). && (2.3.47)
\end{aligned}$$

Now letting $\epsilon \rightarrow 0$ in (2.3.40)-(2.3.43), we obtain that $(\bar{\theta}, \bar{a})$ is a unique solution of (2.3.3)-(2.3.6) with respect to the control $\bar{u}^*$. Since the solution to (2.3.3)-(2.3.6) for a fixed control is unique, we find that $\bar{\theta} = \bar{\theta}^*$ and $\bar{a} = \bar{a}^*$.
Since u_ϵ^* is the optimal control for (2.3.7)-(2.3.11), we have

$$j(u_\epsilon^*) \leq j(\bar{u}^*). \tag{2.3.48}$$

Now letting $\epsilon \rightarrow 0$ in (2.3.48) and using (2.3.35), we obtain

$$j(u^*) \leq j(\bar{u}^*). \tag{2.3.49}$$

Hence u^* is the optimal control. Next we need to show that $\lim\limits_{\epsilon \to 0} \|u_\epsilon^* - u\|_{L^2(I)} = 0$. Since $u_\epsilon^* \longrightarrow u^*$ weakly in $L^2(\Omega)$, it is enough show that $\lim\limits_{\epsilon \to 0} \|u_\epsilon^*\|_{L^2(I)} = \|u^*\|_{L^2(I)}$.
Using Theorem 2.3.1 and (2.3.35), we find that

$$\begin{aligned}
\lim_{\epsilon \to 0} \frac{\beta_3}{2} \|u_\epsilon^*\|_{L^2(I)}^2 &= \lim_{\epsilon \to 0} \left(j(u_\epsilon^*) - \frac{\beta_1}{2} \|a_\epsilon^*(T) - a_d\|^2 - \frac{\beta_2}{2} \|[\theta_\epsilon^* - \theta_m]_+\|_{I,\Omega}^2 \right) && (2.3.50)\\
&= j(u^*) - \frac{\beta_1}{2} \|a^*(T) - a_d\|^2 - \frac{\beta_2}{2} \|[\theta^* - \theta_m]_+\|_{I,\Omega}^2 && (2.3.51)\\
&= \frac{\beta_3}{2} \|u^*\|_{L^2(I)}^2. && (2.3.52)
\end{aligned}$$

Therefore, we have $\lim\limits_{\epsilon \to 0} \|u_\epsilon^*\|_{L^2(I)} = \|u^*\|_{L^2(I)}$ and $\lim\limits_{\epsilon \to 0} \|u_\epsilon^* - u^*\| = 0$. This completes the rest of the proof. □

2.4 Summary

In this chapter, we have established the convergence of the regularized version of the laser surface hardening of steel problem to that of the original problem. In Theorem 2.3.1, it has been proved that for a fixed control $u \in U_{ad}$, the solution of the regularized problem converges to that of the solution of the original problem with rate of convergence $\mathcal{O}(\epsilon)$, ϵ being the regularization parameter. In Theorem 2.3.2, it has been shown that the original problem has at least one solution. In Theorem 2.3.3, it has been shown that the optimal control of the regularized problem also converges to that of the original problem in $L^2(I)$. Numerical experiments justifying the results of this chapter are presented in Chapter 3.

Chapter 3

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

3.1 Introduction

In this chapter, *a priori* error estimates have been developed for a finite element approximation of the optimal control problem of laser surface hardening of steel. The space discretization is done using a cG finite element method, whereas the time and control discretizations are based on a dG finite element method. *A priori* error estimates are developed for temperature, formation of austenite and control.

In literature, even though cG finite element method in space and Euler implicit method in time has been used for the discretization schemes in [2], [84], the rate of convergence of the scheme using finite element method has not been developed. In [2], an abstract control problem for a class of nonlinear parabolic equations is investigated. The surface hardening of steel problem is described as one of the applications. Also, the compactness of solution operator and the existence of the optimal control is established in [2]. Though the convergence of a finite dimensional approximation using finite element methods in space is studied, the order of convergence has not been established. In [84], focus has been given for developing the numerical optimization algorithm using nonlinear conjugate gradient method. A finite element method and an implicit Euler scheme for space and time discretization, respectively, have been used.

The organization of this chapter is as follows. Section 3.1 is introductory in nature. In Section 3.2, a weak formulation for the optimal control problem of laser surface hardening of steel is presented. Section 3.2 also describes a semi-discrete formulation with *a priori* error estimates. In Section 3.3, a complete discretization scheme with *a priori* error estimates has been developed for the state, adjoint and control variables. The numerical results which justify the theoretical results are presented in Section 3.4.

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 chapter.

3.2 Weak Formulation and Semi-Discrete Scheme

In this section, we describe a space discretization for (2.3.7)-(2.3.11) using continuous Galerkin (dG) finite element method with piecewise continuous linear approximations and develop *a priori* error estimates for the spatially discretized system of (2.3.8)-(2.3.11) (subscript ϵ being removed) and the adjoint system. (2.3.7)-(2.3.11) has atleast one global solution, which is characterized by the saddle point $(\theta^*, a^*, z^*, \lambda^*, u^*) \in X \times Y \times X \times Y \times U_{ad}$ of the Lagrangian functional

$$\begin{aligned}\mathcal{L}(\theta, a, z, \lambda, u) &= J(\theta, a, u) - \Big((\partial_t a, \lambda)_{I,\Omega} - (f(\theta, a), \lambda)_{I,\Omega}\Big) - \Big(\rho c_p(\partial_t \theta, z)_{I,\Omega} \\ &\quad + \mathcal{K}(\bigtriangledown\theta, \bigtriangledown z)_{I,\Omega} + \rho L(a_t, z)_{I,\Omega} - (\alpha u, z)_{I,\Omega}\Big),\end{aligned}$$

where $J(\theta, a, u)$ is defined by $J(\theta, a, u) = \dfrac{\beta_1}{2}\displaystyle\int_\Omega |a(T) - a_d|^2 dx + \dfrac{\beta_2}{2}\int_0^T \int_\Omega [\theta - \theta_m]_+^2 dt + \dfrac{\beta_3}{2}\int_0^T |u|^2 dt$, $X = \{v \in L^2(I; V) : v_t \in L^2(I; V^*)\}$, $V = H^1(\Omega)$, $Y = H^1(I; L^2(\Omega))$ and $U_{ad} = \{u \in L^2(I) :\ 0 \le u \le u_{max}$ a.e. in $I\}$. Also, let $H = L^2(\Omega)$.

The saddle point $(\theta^*, a^*, z^*, \lambda^*, u^*) \in X \times Y \times X \times Y \times U_{ad}$ is determined by Karush-Kuhn-Tucker(KKT) system given by:

State Equations:

$$\mathcal{L}_z(\theta^*, a^*, z^*, \lambda^*, u^*)(v) = 0 \quad \forall v \in V, \tag{3.2.1}$$

$$\mathcal{L}_\lambda(\theta^*, a^*, z^*, \lambda^*, u^*)(w) = 0 \quad \forall w \in H. \tag{3.2.2}$$

Adjoint Equations:

$$\mathcal{L}_\theta(\theta^*, a^*, z^*, \lambda^*, u^*)(\phi) \;=\; 0 \quad \forall \phi \in V, \tag{3.2.3}$$

$$\mathcal{L}_a(\theta^*, a^*, z^*, \lambda^*, u^*)(\psi) \;=\; 0 \quad \forall \psi \in H. \tag{3.2.4}$$

First Order Optimality Condition:

$$\mathcal{L}_u(\theta^*, a^*, z^*, \lambda^*, u^*)(p - u^*) \;\geq\; 0 \quad \forall p \in U_{ad}, \tag{3.2.5}$$

where $\mathcal{L}_z(\theta^*, a^*, z^*, \lambda^*, u^*)(v), \mathcal{L}_\lambda(\theta^*, a^*, z^*, \lambda^*, u^*)(w), \mathcal{L}_\theta(\theta^*, a^*, z^*, \lambda^*, u^*)(\phi)$, $\mathcal{L}_a(\theta^*, a^*, z^*, \lambda^*, u^*)(\psi)$ and $\mathcal{L}_u(\theta^*, a^*, z^*, \lambda^*, u^*)(p)$ are the directional derivatives of $\mathcal{L}(\theta^*, a^*, z^*, \lambda^*, u^*)$, with respect to $z^*, \lambda^*, \theta^*, a^*$ and u^*, in the directions of v, w, ϕ, ψ and p, respectively.

The state system (2.3.8)-(2.3.11) is obtained from (3.2.1)-(3.2.2). For the continuity of reading, we state the optimal control problem again:

$$\min_{u \in U_{ad}} J(\theta, a, u) \text{ subject to} \tag{3.2.6}$$

$$(\partial_t a, w) \;=\; (f(\theta, a), w) \qquad \forall w \in H, \text{ a.e. in } I, \tag{3.2.7}$$

$$a(0) \;=\; 0, \tag{3.2.8}$$

$$\rho c_p(\partial_t \theta, v) + \mathcal{K}(\bigtriangledown \theta, \bigtriangledown v) \;=\; -\rho L(\partial_t a, v) + (\alpha u, v) \qquad \forall v \in V, \text{ a.e. in } I, \tag{3.2.9}$$

$$\theta(0) \;=\; \theta_0. \tag{3.2.10}$$

The existence of a unique solution to the state system (3.2.7)-(3.2.10) (see Chapter 2) ensures the existence of a control-to-state mapping $u \mapsto (\theta, a) = (\theta(u), a(u))$ through (3.2.7)-(3.2.10). By means of this mapping, we introduce the reduced cost functional $j : U_{ad} \longrightarrow \mathbb{R}$ as

$$j(u) = J(\theta(u), a(u), u). \tag{3.2.11}$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u \in U_{ad}} j(u). \tag{3.2.12}$$

The adjoint system of (3.2.6)-(3.2.10) obtained from (3.2.3)-(3.2.4) is defined by:

Find $(z^*, \lambda^*) \in X \times Y$ such that

$$\begin{aligned} -(\psi, \partial_t \lambda^*) &= -(\psi, f_a(\theta^*, a^*)(\rho L z^* - \lambda^*)), && (3.2.13) \\ \lambda^*(T) &= \beta_1(a^*(T) - a_d), && (3.2.14) \\ -\rho c_p(\phi, \partial_t z^*) + \mathcal{K}(\bigtriangledown \phi, \bigtriangledown z^*) &= -(\phi, f_\theta(\theta^*, a^*)(\rho L z^* - \lambda^*)) \\ &\quad + \beta_2(\phi, [\theta^* - \theta_m]_+), && (3.2.15) \\ z^*(T) &= 0, && (3.2.16) \end{aligned}$$

$\forall(\psi, \phi) \in H \times V$. Moreover from (3.2.5), z^* satisfies the following variational inequality

$$\left(\beta_3 u^* + \int_\Omega \alpha z^* dx,\ p - u^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \qquad (3.2.17)$$

The existence and uniqueness of the solution of system (3.2.13)-(3.2.16) can be shown using similar arguments used in Chapter 2 (see Theorem 2.2.1) for the state system (3.2.7)-(3.2.10). Also, we have $(z^*, \lambda^*) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$ (see [47, Theorem 2.7]).

Semi-Discrete Scheme

Let $\mathcal{T}_h$ be an admissible regular triangulation of $\bar{\Omega}$ into quadrilaterals K, that is,

- $\bigcup\limits_{K \in \mathbb{T}_h} K = \bar{\Omega}$;
- For $K_1 \neq K_2, K_1 \bigcap K_2$ is either empty, common vertex or a common edge;
- Angles of the triangulation are bounded below by positive constant.

Remark 3.2.1. *In this thesis, we triangulate the domain Ω into rectangles throughout.*

Let the discretization parameter h be defined as $h = \max\limits_{K \in \mathcal{T}_h} h_K$, where h_K is the diameter of the quadrilateral K. Let the finite element space $V_h \subset V$ be defined as

$$V_h = \{v \in C^0(\bar{\Omega}) : \quad v(t)|_K \in Q_1(K) \ \ \forall K \in \mathcal{T}_h\}.$$

Here $Q_1(K)$ denotes the set of all polynomials of degree ≤ 1 in each variable x and y. Then the semi-discrete version corresponding to the continuous problem (3.2.6)-(3.2.10) reads as:

$$\min_{u_h \in U_{ad}} J(\theta_h, a_h, u_h) \text{ subject to} \tag{3.2.18}$$

$$(\partial_t a_h, w) = (f(\theta_h, a_h), w) \quad \forall w \in V_h, \text{ a.e. in } I, \tag{3.2.19}$$

$$a_h(0) = 0, \tag{3.2.20}$$

$$\rho c_p(\partial_t \theta_h, v) + \mathcal{K}(\nabla \theta_h, \nabla v) = -\rho L(\partial_t a_h, v) + (\alpha u_h, v) \ \forall v \in V_h, \text{ a.e. in } I, \tag{3.2.21}$$

$$\theta_h(0) = \theta_{h,0}, \tag{3.2.22}$$

where $\theta_{h,0}$ is a suitable approximation of θ_0 to be chosen later. (3.2.18)-(3.2.22) has at least one global solution, see [2], which is characterized by the saddle point $(\theta_h^*(t), a_h^*(t), z_h^*(t), \lambda_h^*(t), u_h^*(t)) \in V_h \times V_h \times V_h \times V_h \times U_{ad}$ of the Lagrangian functional defined by

$$\begin{aligned}\mathcal{L}(\theta_h, a_h, z_h, \lambda_h, u_h) = \; & J(\theta_h, a_h, u_h) - \Big((\partial_t a_h, \lambda_h)_{I,\Omega} - (f(\theta_h, a_h), \lambda_h)_{I,\Omega}\Big) - \Big(\rho c_p(\partial_t \theta_h, z_h)_{I,\Omega} \\ & + \mathcal{K}(\nabla \theta_h, \nabla z_h)_{I,\Omega} + \rho L(\partial_t a_h, z_h)_{I,\Omega} - (\alpha u_h, z_h)_{I,\Omega}\Big)\end{aligned}$$

The saddle point $(\theta_h^*(t), a_h^*(t), z_h^*(t), \lambda^*(t), u^*(t)) \in V_h \times V_h \times V_h \times V_h \times U_{ad}$ is determined by the KKT system given by:

State equations:

$$\mathcal{L}_z(\theta_h^*, a_h^*, z_h^*, \lambda_h^*, u_h^*)(v) = 0 \quad \forall v \in V_h, \tag{3.2.23}$$

$$\mathcal{L}_\lambda(\theta_h^*, a_h^*, z_h^*, \lambda_h^*, u_h^*)(w) = 0 \quad \forall w \in V_h. \tag{3.2.24}$$

Adjoint equations:

$$\mathcal{L}_\theta(\theta_h^*, a_h^*, z_h^*, \lambda_h^*, u_h^*)(\phi) = 0 \quad \forall \phi \in V_h, \tag{3.2.25}$$

$$\mathcal{L}_a(\theta_h^*, a_h^*, z_h^*, \lambda_h^*, u_h^*)(\psi) = 0 \quad \forall \psi \in V_h. \tag{3.2.26}$$

First order optimality condition:

$$\mathcal{L}_u(\theta_h^*, a_h^*, z_h^*, \lambda_h^*, u_h^*)(p - u_h^*) \geq 0 \quad \forall p \in U_{ad}. \tag{3.2.27}$$

The state system (3.2.19)-(3.2.22) is obtained from (3.2.23)-(3.2.24). The adjoint system of (3.2.18)-(3.2.22) obtained from (3.2.25)-(3.2.26) is defined by:
Find $(z_h^*(t), \lambda_h^*(t)) \in V_h \times V_h, \;\; t \in \bar{I}$ such that

$$\begin{aligned} -(\psi, \partial_t \lambda_h^*) &= -(\psi, f_a(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)), && (3.2.28)\\ \lambda_h^*(T) &= \beta_1(a_h^*(T) - a_d), && (3.2.29)\\ -\rho c_p(\phi, \partial_t z_h^*) + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z_h^*) &= -(\phi, f_\theta(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)) \\ &\quad + \beta_2(\phi, [\theta_h^* - \theta_m]_+), && (3.2.30)\\ z_h^*(T) &= 0, && (3.2.31) \end{aligned}$$

$\forall(\psi, \phi) \in V_h \times V_h$. Moreover from (3.2.27), z_h^* satisfies the following variational inequality

$$\left(\beta_3 u_h^* + \int_\Omega \alpha z_h^* dx, \; p - u_h^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \qquad (3.2.32)$$

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

$$j_h(u_h) = J(\theta_h(u_h), a_h(u_h), u_h). \qquad (3.2.33)$$

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

$$\min_{u_h \in U_{ad}} j_h(u_h). \qquad (3.2.34)$$

The first order necessary optimality condition for (3.2.34) reads as

$$j_h'(u_h^*)(p - u_h^*) \geq 0 \;\; \forall p \in U_{ad}, \qquad (3.2.35)$$

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

Below, we discuss the *a priori* error estimates for the state equations. Define the elliptic projection $\mathcal{R}_h : V \longrightarrow V_h$ by

$$\mathcal{K}(\bigtriangledown(v - \mathcal{R}_h v), \bigtriangledown\phi) + \gamma(v - \mathcal{R}_h v, \phi) \;\; = \;\; 0 \quad \forall \phi \in V_h, \qquad (3.2.36)$$

where γ is a positive constant.

Lemma 3.2.1 ([85], page no. 737). *For the interpolation operator defined by* (3.2.36) *and* $v \in H^2(\Omega)$, *we have the following error estimates:*

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

We also define the L^2-projection $P_h : H^2(\Omega) \longrightarrow V_h$ [79], such that

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

Note that P_h satisfies the following error estimates:

$$\|v - P_h v\| \;\; \leq \;\; Ch^2 \|v\|_{H^2(\Omega)} \quad \forall v \in H^2(\Omega). \tag{3.2.37}$$

Theorem 3.2.1. *Let* $(\theta(t), a(t))$ *and* $(\theta_h(t), a_h(t))$ *be the solutions of* (3.2.7)-(3.2.10) *and* (3.2.19)-(3.2.22), *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 a fixed* $u \in U_{ad}$, *there exists a positive constant* C *independent of* h *such that*

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

Proof. Let $\zeta^\theta = \theta - \mathcal{R}_h \theta$ and $\eta^\theta = \mathcal{R}_h \theta - \theta_h$. Subtract (3.2.21) from (3.2.9), use (3.2.7), (3.2.19) and (3.2.36) to obtain

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

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

$$\begin{aligned} \|\eta^\theta(t)\|^2 \;\; \leq \;\; & C \Big(\|\eta^\theta(0)\|^2 + \int_0^t \Big(\|f(\theta, a) - f(\theta_h, a_h)\|^2 + \|\partial_t \zeta^\theta\|^2 \\ & + \|\eta^\theta\|^2 + \|\zeta^\theta\|^2 \Big) ds \Big). \end{aligned} \tag{3.2.38}$$

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

3.2.1, we obtain

$$\|\eta^\theta(0)\|^2 \ \leq \ \|\mathcal{R}_h\theta_0 - \theta_0\|^2 + \|\theta_0 - \theta_{h,0}\|^2 \leq Ch^4\|\theta_0\|^2_{H^2(\Omega)}. \tag{3.2.39}$$

Using the fact that f is Lipschitz in both the arguments (see Proposition 2.2.1) and (3.2.39) in (3.2.38), we find that

$$\begin{aligned}\|\eta^\theta(t)\|^2 \ &\leq \ C\Big(h^4\|\theta_0\|^2_{H^2(\Omega)} + \int_0^t (\|\zeta^\theta\|^2 + \|\zeta^a\|^2 + \|\partial_t\zeta^\theta\|^2)ds \\ &\quad + \int_0^t (\|\eta^\theta\|^2 + \|\eta^a\|^2)ds\Big), \end{aligned} \tag{3.2.40}$$

where $\zeta^a = a - P_h a$, $\eta^a = P_h a - a_h$. Now subtracting (3.2.19) from (3.2.7) for fixed $t \in I$, integrating from 0 to t, using Cauchy-Schwarz inequality, Young's inequality and the fact that $(\partial_t\zeta^a, \eta^a) = 0$, we obtain

$$\|\eta^a(t)\|^2 \leq C\Big(\int_0^t (\|\zeta^\theta\|^2 + \|\zeta^a\|^2)ds + \int_0^t (\|\eta^\theta\|^2 + \|\eta^a\|^2)ds\Big). \tag{3.2.41}$$

Adding (3.2.40) and (3.2.41), we arrive at

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

Using Gronwall's lemma, Lemma 3.2.1 and (3.2.37), we obtain

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

Using triangle inequality to split $\theta - \theta_h$ and $a - a_h$, we obtain the required result. This completes the proof. □

Remark 3.2.2. *Although, the finite element space used in this chapter to discretize the variables θ and a is V_h, where approximation is done using continuous functions, the variable a can also be approximated using piecewise constants. Here V_h is used for both θ and a for computational ease.*

Below, we discuss the adjoint error estimates.

Theorem 3.2.2. *Let* $(z^*(t), \lambda^*(t))$ *and* $(z_h^*(t), \lambda_h^*(t))$ *be the solutions of* (3.2.13)-(3.2.16) *and* (3.2.28)-(3.2.31) *corresponding to the state solutions* (θ^*, a^*) *and* (θ_h^*, a_h^*)*, respectively. Then, under the extra regularity assumptions made in Theorem 3.2.1 and* $(z^*, \lambda^*) \in L^\infty(I, H^2(\Omega)) \times L^\infty(I, H^2(\Omega))$*,* $\partial_t z^* \in L^\infty(I, H^2(\Omega))$*,* $a_d \in H^2(\Omega)$*; there exists a positive constant* C *independent of* h *such that*

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

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 (3.2.30) from (3.2.15) to obtain

$$\begin{aligned}
&-\rho c_p(\phi, \partial_t(z^* - z_h^*)) + \mathcal{K}(\nabla\phi, \nabla(z^* - z_h^*)) \\
&+(\phi, f_\theta(\theta^*, a^*)(\rho L z^* - \lambda^*)) - (\phi, f_\theta(\theta_h^*, a_h^*)(\rho L z_h^* - \lambda_h^*)) \qquad = \beta_2(\phi, [\theta^* - \theta_m]_+ - [\theta_h^* - \theta_m]_+),
\end{aligned}$$

where $\phi \in V_h$. Using Proposition 2.2.1 and (3.2.36), we obtain

$$\begin{aligned}
-\rho c_p(\phi, \partial_t\eta^z) \ &+ \ \mathcal{K}(\nabla\phi, \nabla\eta^z) - c_f\rho L(\phi, \eta^z) \\
&\leq \ C\Big((\phi, \eta^\lambda) + (\phi, \zeta^\lambda) + (\phi, [\theta^* - \theta_m]_+ - [\theta_h^* - \theta_m]_+) + (\phi, \partial_t\zeta^z) + \gamma(\phi, \zeta^z) \Big).
\end{aligned}$$

Substitute $\phi = \eta^z$, integrate from t to T, apply Cauchy-Schwarz inequality and Young's inequality, to obtain

$$\|\eta^z(t)\|^2 \leq C\Big(\|\theta - \theta_h\|^2 + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 + \|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2) ds \Big). \quad (3.2.42)$$

Subtract (3.2.28) from (3.2.13) and choose $\chi = \eta^\lambda$. Then integrate from t to T and apply Cauchy-Schwarz inequality with Young's inequality, to obtain

$$\begin{aligned}
\|\eta^\lambda(t)\|^2 \ \leq \ & C\Big(\|P_h a_d - a_d\|^2 + \|P_h a^*(T) - a_h^*(T)\|^2 + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 \\
& + \|\zeta^z\|^2 + \|\zeta^\lambda\|^2) ds \Big).
\end{aligned}$$

Using (3.2.37), we obtain

$$\begin{aligned}\|\eta^\lambda(t)\|^2 &\leq C\Big(h^4\|a_d\|^2_{H^2(\Omega)} + \|P_h a^*(T) - a_h^*(T)\|^2 + \int_t^T (\|\eta^z\|^2 + \|\eta^\lambda\|^2 \\ &\quad + \|\zeta^z\|^2 + \|\zeta^\lambda\|^2)ds\Big). \qquad (3.2.43)\end{aligned}$$

Adding (3.2.42) and (3.2.43), we find that

$$\begin{aligned}\|\eta^z(t)\|^2 + \|\eta^\lambda(t)\|^2 &\leq C\Big(h^4\|a_d\|^2_{H^2(\Omega)} + \|\theta - \theta_h\|^2 + \|P_h a^*(T) - a_h^*(T)\|^2 + \int_0^T \Big(\|\zeta^z\|^2 \\ &\quad + \|\zeta^\lambda\|^2 + \|\partial_t \zeta^z\|^2\Big)ds + \int_t^T \Big((\|\eta^z\|^2 + \|\eta^\lambda\|^2\Big)ds\Big). \qquad (3.2.44)\end{aligned}$$

Using Gronwall's lemma and Theorem 3.2.1, we obtain

$$\begin{aligned}\|\eta^z(t)\|^2 + \|\eta^\lambda(t)\|^2 &\leq C\Big(h^4\Big(\|\theta_0\|^2_{H^2(\Omega)} + \|\theta^*\|^2_{L^\infty(I,H^2(\Omega))} + \|a^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|^2_{L^\infty(I,H^2(\Omega))} \\ &\quad + \|a_d\|^2_{H^2(\Omega)}\Big) + \int_0^T (\|\zeta^z\|^2 + \|\zeta^\lambda\|^2 + \|\partial_t\zeta^z\|^2)ds\Big). \qquad (3.2.45)\end{aligned}$$

A use of Lemma 3.2.1, (3.2.37) in (3.2.45) yields the required result. □

3.3 Completely Discrete Scheme

In this section, first of all, a temporal discretization is done using a dG finite element method with piecewise constant approximation and *a priori* error estimates are proved in Theorem 3.3.1 and 3.3.2. The control is then discretized using piecewise constants in each discrete interval $I_n, n = 1, 2, \cdots, N$. In order to discretize (3.2.18)-(3.2.22) in time, we consider the following partition of I:

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

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

$$X^q_{hk} = \{\phi : I \to V_h;\ \ \phi|_{I_n} = \sum_{j=0}^{q} \psi_j t^j, \psi_j \in V_h\},\ \ q \in \mathbb{N}. \qquad (3.3.1)$$

For a function v in X^q_{hk}, we use the following notations [79]:

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

Then a cG finite element approximation with piecewise polynomials of degree 1 and a discontinuous Galerkin (dG) finite element approximation with piecewise polynomials of degree q(denoted as cG(1)dG(q)) of (3.2.6)-(3.2.10) reads as:

$$\begin{aligned}
\min_{u_{hk} \in U_{ad}} J(\theta_{hk}, a_{hk}, u_{hk}) & \qquad \text{subject to} && (3.3.2)\\
\sum_{n=1}^{N} (\partial_t a_{hk}, w)_{I_n,\Omega} & + \sum_{n=1}^{N-1} ([a_{hk}]_n, w_n^+) + (a_{hk,0}^+, w_0^+) = (f(\theta_{hk}, a_{hk}), w)_{I,\Omega}, && (3.3.3)\\
a_{hk}(0) & = 0, && (3.3.4)\\
\rho c_p \sum_{n=1}^{N} (\partial_t \theta_{hk}, v)_{I_n,\Omega} & + \mathcal{K}(\bigtriangledown \theta_{hk}, \bigtriangledown v)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_{hk}]_n, v_n^+) + \rho c_p (\theta_{hk,0}^+, v_0^+) \\
& = -\rho L (f(\theta_{hk}, a_{hk}), v)_{I,\Omega} + (\alpha u_{hk}, v)_{I,\Omega} + \rho c_p (\theta_0, v_0^+), && (3.3.5)\\
\theta_{hk}(0) & = \theta_{h,0} && (3.3.6)
\end{aligned}$$

for all $(w, v) \in X^q_{hk} \times X^q_{hk}$. (3.3.2)-(3.3.6) has atleast one global solution, which is characterized by the saddle point $(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk}) \in X^q_{hk} \times X^q_{hk} \times X^q_{hk} \times X^q_{hk} \times U_{ad}$ of the Lagrangian functional

$$\begin{aligned}
\mathcal{L}(\theta_{hk}, a_{hk}, z_{hk}, \lambda_{hk}, u_{hk}) \; = \; & J(\theta_{hk}, a_{hk}, u_{hk}) - \Bigg(\sum_{n=1}^{N} (\partial_t a_{hk}, \lambda_{hk})_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a_{hk}]_n, \lambda_{hk,n}^+) \\
& + (a_{hk,0}^+, \lambda_{hk,0}^+) - (f(\theta_{hk}, a_{hk}), \lambda_{hk})_{I,\Omega} \Bigg) - \Bigg(\sum_{n=1}^{N} \rho c_p (\partial_t \theta_{hk}, z_{hk})_{I_n,\Omega} \\
& + \mathcal{K}(\bigtriangledown \theta_{hk}, \bigtriangledown z_{hk})_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_{hk}]_n, z_{hk,n}^+) + \rho c_p (\theta_{hk,0}^+, z_{hk,0}^+) \\
& + \rho L (f(\theta_{hk}, a_{hk}, z_{hk}))_{I,\Omega} - (\alpha u_{hk}, z_{hk})_{I,\Omega} - \rho c_p (\theta_0, z_{hk,0}^+) \Bigg).
\end{aligned}$$

The saddle point $(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk}) \in X^q_{hk} \times X^q_{hk} \times X^q_{hk} \times X^q_{hk} \times U_{ad}$ is determined by the KKT system given by:

State equations:

$$\mathcal{L}_z(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk})(v) = 0 \quad \forall v \in X^q_{hk}, \tag{3.3.7}$$

$$\mathcal{L}_\lambda(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk})(w) = 0 \quad \forall w \in X^q_{hk}. \tag{3.3.8}$$

Adjoint equations:

$$\mathcal{L}_\theta(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk})(\phi) = 0 \quad \forall \phi \in X^q_{hk}, \tag{3.3.9}$$

$$\mathcal{L}_a(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk})(\psi) = 0 \quad \forall \psi \in X^q_{hk}. \tag{3.3.10}$$

First order optimality condition:

$$\mathcal{L}_u(\theta^*_{hk}, a^*_{hk}, z^*_{hk}, \lambda^*_{hk}, u^*_{hk})(p - u^*_{hk}) \geq 0 \quad \forall p \in U_{ad}. \tag{3.3.11}$$

The adjoint system of (3.3.2)-(3.3.6) obtained from (3.3.9)-(3.3.10) is defined by:
Find $(z^*_{hk}, \lambda^*_{hk}) \in X^q_{hk} \times X^q_{hk}$ such that

$$\begin{aligned}
-\sum_{n=1}^{N}(\psi, \partial_t \lambda^*_{hk})_{I_n,\Omega} - \sum_{n=1}^{N-1}(\psi_n, [\lambda^*_{hk}]_n) &= -(\psi, f_a(\theta^*_{hk}, a^*_{hk})(\rho L z^*_{hk} - \lambda^*_{hk}))_{I,\Omega} && (3.3.12)\\
\lambda^*_{hk}(T) &= \beta_1(a^*_{hk}(T) - a_d), && (3.3.13)\\
-\rho c_p \sum_{n=1}^{N}(\phi, \partial_t z^*_{hk})_{I_n,\Omega} + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z^*_{hk})_{I,\Omega} &- \rho c_p \sum_{n=1}^{N-1}(\phi_n, [z^*_{hk}]_n) \\
= -(\phi, f_\theta(\theta^*_{hk}, a^*_{hk})(\rho L z^*_{hk} - \lambda^*_{hk}))_{I,\Omega} &+ \beta_2(\phi, [\theta^*_{hk} - \theta_m]_+)_{I,\Omega}, && (3.3.14)\\
z^*_{hk}(T) &= 0, && (3.3.15)
\end{aligned}$$

Moreover from (3.3.11), z^*_h satisfies the following variational inequality

$$\left(\beta_3 u^*_{hk} + \int_\Omega \alpha z^*_{hk} dx,\ p - u^*_{hk}\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \tag{3.3.16}$$

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

$$j_{hk}(u_{hk}) = J(\theta_{hk}(u_{hk}), a_{hk}(u_{hk}), u_{hk}). \tag{3.3.17}$$

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

$$\min_{u_{hk} \in U_{ad}} j_{hk}(u_{hk}). \tag{3.3.18}$$

The first order necessary optimality condition for (3.3.18) reads as

$$j'_{hk}(u^*_{hk})(p-u^*_{hk}) \geq 0 \quad \forall p \in U_{ad}. \tag{3.3.19}$$

We consider the case of piecewise constant approximation in time for both the state and adjoint formulation. For the case where $q=0$ in the definition of X^0_{hk}, (3.3.3)-(3.3.6) 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

$$\begin{aligned}
\left(\frac{a^n_{hk}-a^{n-1}_{hk}}{k_n}, w\right) &= \frac{1}{k_n}\left(\int_{I_n} f(\theta^n_{hk}, a^n_{hk})ds, w\right), && (3.3.20)\\
a_{hk}(0) &= 0, && (3.3.21)\\
\rho c_p\left(\frac{\theta^n_{hk}-\theta^{n-1}_{hk}}{k_n}, v\right) + \mathcal{K}(\triangledown\theta^n_{hk}, \triangledown v) &= -\rho L\left(\frac{1}{k_n}\int_{I_n} f(\theta^n_{hk}, a^n_{hk})ds, v\right) && \\
&\quad + \left(\frac{1}{k_n}\int_{I_n} \alpha u_{hk} ds, v\right), && (3.3.22)\\
\theta_{hk}(0) &= \theta_{h,0}, && (3.3.23)
\end{aligned}$$

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

Before estimating the *a priori* error estimates for space-time discretization, we define the interpolant $\pi_k : C(\bar{I}, V_h) \longrightarrow X^0_{hk}$ [79] as:

$$\pi_k v(t) = v(t_n) \qquad \text{if } t \in I_n, \quad \forall n = 1,2,\cdots,N, \tag{3.3.24}$$

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

$$\|v - \pi_k v\|_{I,\Omega} \leq Ck\|\partial_t v\|. \tag{3.3.25}$$

Theorem 3.3.1. *Let* $(\theta^n_{hk}, a^n_{hk})$ $\forall n = 1,2,\cdots,N$ *and* $(\theta(t), a(t))$ *be the solutions of the problems* (3.3.20)-(3.3.23) *and* (3.2.7)-(3.2.10), *respectively. Then, under the extra regularity assumptions made in Theorem 3.2.1 and* $(\partial_{tt}\theta, \partial_{tt}a) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega))$, $\partial_t u \in L^2(I)$, *for a fixed* $u \in U_{ad}$*; there exists a positive constant* C *independent of* h *and* k *such that, for all* $t \in \bar{I}_n$

$$\begin{aligned}
\|\theta(t_n) - \theta^n_{hk}\| + \|a(t_n) - a^n_{hk}\| \;&\leq\; C(h^2+k)\Big(\|\theta\|_{L^\infty(I,H^2(\Omega))} + \|a\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|_{L^\infty(I,H^2(\Omega))}\\
&\quad + \|\theta_0\|_{H^2(\Omega)} + \|\partial_{tt}\theta\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}a\|_{L^\infty(I,L^2(\Omega))} + \|\partial_t u\|_{L^2(I)}\Big).
\end{aligned}$$

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 $\dfrac{\theta^n_{hk}-\theta^{n-1}_{hk}}{k_n}$ by $\bar\partial\theta^n_{hk}$. Also, write $a(t_n)-a^n_{hk}=(a(t_n)-P_ha(t_n))+(P_ha(t_n)-a^n_{hk})=\zeta^{a,n}+\eta^{a,n}$ and denote $\dfrac{a^n_{hk}-a^{n-1}_{hk}}{k_n}$ by $\bar\partial a^n_{hk}$. Subtracting (3.3.22) from (3.2.9), we obtain, at $t=t_n$;

$$\begin{aligned}\rho c_p(\partial_t\theta(t_n)-\bar\partial\theta^n_{hk},v) &+ \mathcal{K}(\bigtriangledown(\theta(t_n)-\theta^n_{hk}),\bigtriangledown v)\\ &= -\rho L\Big(f(\theta(t_n),a(t_n))-\frac{1}{k_n}\int_{I_n}f(\theta^n_{hk},a^n_{hk})ds,v\Big)\\ &\quad+\Big(\alpha(x,t_n)u(t_n)-\frac{1}{k_n}\int_{I_n}\alpha u ds,v\Big),\end{aligned}$$

where $v\in V_h$. Using (3.3.24), we find that

$$\begin{aligned}\rho c_p(\partial_t\theta(t_n)-\bar\partial\theta^n_{hk},v)+\mathcal{K}(\bigtriangledown(\theta(t_n)-\theta^n_{hk}),\bigtriangledown v) &\le -\rho L\Big(f(\theta(t_n),a(t_n))-f(\theta^n_{hk},a^n_{hk}),v\Big)\\ &\quad+\max_{\Omega\times I_n}\frac{1}{k_n}|\alpha|\Big(\int_{I_n}(\pi_ku(t_n)-u)ds,v\Big).\end{aligned}$$

Using (3.2.36) and Cauchy-Schwarz inequality, we find that

$$\begin{aligned}\rho c_p(\bar\partial\eta^{\theta,n},v) &+ \mathcal{K}(\bigtriangledown\eta^{\theta,n},\bigtriangledown v)\\ &\le \rho L\|f(\theta^n_{hk},a^n_{hk})-f(\theta(t_n),a(t_n))\|\;\|v\|+\rho c_p\|\bar\partial\theta(t_n)-\partial_t\theta(t_n)\|\|v\|\\ &\quad+\rho c_p\|\bar\partial\zeta^{\theta,n}\|\|v\|+\gamma\|\zeta^{\theta,n}\|\|v\|+\max_{\Omega\times I_n}\frac{1}{k_n}|\alpha|\|\pi_ku-u\|_{L^2(I)}\|v\|.(3.3.26)\end{aligned}$$

Substituting $v=\eta^{\theta,n}$ in (3.3.26), we obtain

$$\begin{aligned}\rho c_p(\bar\partial\eta^{\theta,n},\eta^{\theta,n})+\mathcal{K}(\bigtriangledown\eta^{\theta,n},\bigtriangledown\eta^{\theta,n}) &\le \rho L\|f(\theta^n_{hk},a^n_{hk})-f(\theta(t_n),a(t_n))\|\;\|\eta^{\theta,n}\|+\gamma\|\zeta^{\theta,n}\|\|\eta^{\theta,n}\|\\ &\quad+\rho c_p\|\bar\partial\theta(t_n)-\partial_t\theta(t_n)\|\|\eta^{\theta,n}\|+\rho c_p\|\bar\partial\zeta^{\theta,n}\|\|\eta^{\theta,n}\|\\ &\quad+\max_{\Omega\times I_n}\frac{1}{k_n}|\alpha|\|\pi_ku-u\|_{L^2(I)}\|\eta^{\theta,n}\|. \qquad (3.3.27)\end{aligned}$$

Now,

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

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

$$\begin{aligned}\frac{1}{2k_n}\left(\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2\right) &= (\bar{\partial}\eta^{\theta,n},\eta^{\theta,n}) - \frac{1}{2k_n}(\eta^{\theta,n}-\eta^{\theta,n-1},\eta^{\theta,n}-\eta^{\theta,n-1})\\ &\le (\bar{\partial}\eta^{\theta,n},\eta^{\theta,n}). \qquad (3.3.28)\end{aligned}$$

Using (3.3.28) in (3.3.27), we find that

$$\begin{aligned}\|\eta^{\theta,n}\|^2 &- \|\eta^{\theta,n-1}\|^2\\ &\le C\bigg(k_n\|f(\theta^n_{hk},a^n_{hk}) - f(\theta(t_n),a(t_n))\|\ \|\eta^{\theta,n}\| + k_n\|\bar{\partial}\theta(t_n)-\partial_t\theta(t_n)\|\|\eta^{\theta,n}\|\\ &\quad + k_n\|\bar{\partial}\zeta^{\theta,n}\|\|\eta^{\theta,n}\| + k_n\|\zeta^{\theta,n}\|\|\eta^{\theta,n}\| + \max_{\Omega\times I_n}|\alpha|\|\pi_k u - u\|_{L^2(I)}\|\eta^{\theta,n}\|\bigg).\end{aligned}$$

Using Young's inequality and Proposition 2.2.1, we obtain

$$\begin{aligned}\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2 &\le C\bigg(\|\eta^{\theta,n}\|^2 + \|\zeta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{a,n}\|^2 + \|\bar{\partial}\theta(t_n)-\partial_t\theta(t_n)\|^2\\ &\quad + \|\bar{\partial}\zeta^{\theta,n}\|^2 + \|u-\pi_k u\|^2_{L^2(I_n)}\bigg)\\ &\le C\bigg(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + R^1_n\bigg), \qquad (3.3.29)\end{aligned}$$

where $R^1_n = \|\zeta^{\theta,n}\|^2 + \|\zeta^{a,n}\|^2 + \|\bar{\partial}\theta(t_n)-\partial_t\theta(t_n)\|^2 + \|\bar{\partial}\zeta^{\theta,n}\|^2 + \|u-\pi_k u\|^2_{L^2(I_n)}$. Subtracting (3.3.20) from (3.2.7), we obtain, at $t = t_n$;

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

where $w \in V_h$. Putting $w = \eta^{a,n}$, proceeding as in (3.3.27)-(3.3.28) and using Proposition 2.2.1, we find that

$$\begin{aligned}\|\eta^{a,n}\|^2 - \|\eta^{a,n-1}\|^2 &\le C\bigg(\|\eta^{\theta,n}\|^2 + \|\zeta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + \|\zeta^{a,n}\|^2\\ &\quad + \|\bar{\partial}a(t_n)-\partial_t a(t_n)\|^2\bigg). \qquad (3.3.30)\end{aligned}$$

Let $R^2_n = \|\zeta^{a,n}\|^2 + \|\zeta^{\theta,n}\|^2 + \|\bar{\partial}a(t_n)-\partial_t a(t_n)\|^2$. From (3.3.29) and (3.3.30), we obtain

$$\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 - \|\eta^{\theta,n-1}\|^2 - \|\eta^{a,n-1}\|^2 \le C\bigg(\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 + R^1_n + R^2_n\bigg)$$

Summing up from 1 to n and using the fact that $a(0) = 0$ and $a_h(0) = 0$, we arrive at

$$\|\eta^{\theta,n}\|^2 + \|\eta^{a,n}\|^2 \;\; \leq \;\; \|\eta^{\theta,0}\|^2 + Ck\sum_{m=1}^{n}(\|\eta^{\theta,m}\|^2 + \|\eta^{a,m}\|^2 + R^1_m + R^2_m).$$

From Lemma 3.2.1 and (3.2.37), we have $\|\zeta^{\theta,j}\| \leq Ch^2\|\theta\|_{H^2(\Omega)}$ and $\|\zeta^{a,j}\| \leq Ch^2\|a\|_{H^2(\Omega)}$, repectively. Also,

$$\|\bar{\partial}\zeta^{\theta,j}\| = \;\; \|k_j^{-1}\int_{t_{j-1}}^{t_j}\partial_t\zeta^{\theta,j}dt\| \leq k_j^{-1}\int_{t_{j-1}}^{t_j}\|\partial_t\zeta^{\theta,j}\|dt \leq Ch^2\|\partial_t\theta\|_{L^\infty(I,H^2(\Omega))}, \quad (3.3.31)$$

Note that,

$$\begin{aligned}\|\bar{\partial}\theta(t_j) - \partial_t\theta(t_j)\| \;\; &= \;\; \|k_j^{-1}\int_{t_{j-1}}^{t_j}(t-t_{j-1})\partial_{tt}\theta dt\| \leq k_j^{-1}\int_{t_{j-1}}^{t_j}(t-t_{j-1})\|\partial_{tt}\theta\|dt \\ &\leq \;\; Ck_j^{-1}\frac{(s-t_{j-1})^2}{2}\bigg|_{t_{j-1}}^{t_j}\|\partial_{tt}\theta\| \leq Ck\|\partial_{tt}\theta\|_{L^\infty(I,L^2(\Omega))}. \qquad (3.3.32)\end{aligned}$$

A use of Lemma 3.2.1 with (3.3.25), (3.3.31) and (3.3.32) implies that

$$R^1_n \leq C(h^4+k^2)\bigg(\|\theta\|^2_{L^\infty(I,H^2(\Omega))}+\|a\|^2_{L^\infty(I,H^2(\Omega))}+\|\partial_t\theta\|^2_{L^\infty(I,H^2(\Omega))}+\|\partial_{tt}\theta\|^2_{L^\infty(I,L^2(\Omega))}+\|\partial_t u\|^2_{L^2(I_n)}\bigg),$$

where $k = \max_{1\leq n\leq N} k_n$. Similarly, we find that

$$R^2_n \leq C(h^4+k^2)\bigg(\|a\|^2_{L^\infty(I,H^2(\Omega))} + \|\theta\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_{tt}a\|^2_{L^\infty(I,L^2(\Omega))}\bigg)$$

and $\|\eta^{\theta,0}\| \leq Ch^2\|\theta_0\|_{H^2(\Omega)}$. Hence, we obtain

$$\begin{aligned}\|\eta^{\theta,n}\| \;\; &+ \;\; \|\eta^{a,n}\| \\ &\leq \;\; C\bigg((h^2+k)\bigg(\|\theta\|_{L^\infty(I,H^2(\Omega))} + \|a\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|_{L^\infty(I,H^2(\Omega))} + \|\theta_0\|_{H^2(\Omega)} \\ &\quad + \|\partial_{tt}\theta\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}a\|_{L^\infty(I,L^2(\Omega))} + \|\partial_t u\|_{L^2(I)}\bigg) + \sum_{m=1}^{n}(\|\eta^{\theta,m}\| + \|\eta^{a,m}\|)\bigg).\end{aligned}$$

Using Gronwall's lemma, we arrive at

$$\begin{aligned}\|\eta^{\theta,n}\| + \|\eta^{a,n}\| \;\; &\leq \;\; C(h^2+k)\bigg(\|\theta\|_{L^\infty(I,H^2(\Omega))} + \|a\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta\|_{L^\infty(I,H^2(\Omega))} \\ &\quad + \|\theta_0\|_{H^2(\Omega)} + \|\partial_{tt}\theta\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}a\|_{L^\infty(I,L^2(\Omega))} + \|\partial_t u\|_{L^2(I)}\bigg).\end{aligned}$$

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 time discrete formulation for (3.3.12)-(3.3.15) for $q = 0$ in X^0_{hk} is defined as:

Find $(z^{n-1}_{hk}, \lambda^{n-1}_{hk}) \in V_h \times V_h, n = N-2, N-3, \cdots, 1$ such that

$$\begin{aligned}
-(\psi, \tilde{\partial}\lambda^{n-1}_{hk}) &= -(\psi, f_a(\theta^{n-1}_{hk}, a^{n-1}_{hk})(\rho L z^{n-1}_{hk} - \lambda^{n-1}_{hk})), && (3.3.33)\\
\lambda_{hk}(T) &= \beta_1(a_{hk}(T) - a_d), && (3.3.34)\\
-\rho c_p(\phi, \tilde{\partial} z^{n-1}_{hk}) + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z^{n-1}_{hk}) &= -(\phi, f_\theta(\theta^{n-1}_{hk}, a^{n-1}_{hk})(\rho L z^{n-1}_{hk} - \lambda^{n-1}_{hk})) && (3.3.35)\\
&\quad + \beta_2(\phi, [\theta^{n-1}_{hk} - \theta_m]_+), &&\\
z_{hk}(T) &= 0, && (3.3.36)
\end{aligned}$$

for all $(\psi, \phi) \in V_h \times V_h$ and $\tilde{\partial}\phi^{n-1} = \dfrac{\phi^n - \phi^{n-1}}{k_n}$.

Theorem 3.3.2. *Let* $(z^{n-1,*}_{hk}, \lambda^{n-1,*}_{hk})$ $\forall n = 1, 2, \cdots, N, N+1$ *and* (z^*, λ^*) *be the solutions of the adjoint problems* (3.3.33)-(3.3.36) *and* (3.2.13)-(3.2.16) *corresponding to the solutions* $(\theta^{n-1,*}_{hk}, a^{n-1,*}_{hk})$ $\forall n = 1, 2, \cdots, N, N+1$ *and* (θ^*, a^*) *of* (3.3.20)-(3.3.23) *and* (3.2.7)-(3.2.10), *respectively, with a fixed control* $u^* \in U_{ad}$. *Then, under the extra regularity assumptions in Theorem 3.3.1 and 3.2.2 with* $(\partial_{tt} z, \partial_{tt}\lambda) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega))$*; there exists a positive constant* C *independent of* h *and* k*, such that, for all* $t \in \bar{I}_n$

$$\begin{aligned}
&\|z^{n-1,*}_{hk} - z^*(t_{n-1})\| + \|\lambda^{n-1,*}_{hk} - \lambda^*(t_{n-1})\|\\
&\quad \le C(h^2 + k)\Big(\|\theta^*\|_{L^\infty(I,H^2(\Omega))} + \|a^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t\theta^*\|_{L^2(I,H^2(\Omega))} + \|\partial_{tt}\theta^*\|_{L^\infty(I,L^2(\Omega))}\\
&\qquad + \|\partial_{tt} a^*\|_{L^\infty(I,L^2(\Omega))} + \|z^*\|_{L^\infty(I,H^2(\Omega))} + \|\lambda^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t z^*\|_{L^2(I,H^2(\Omega))}\\
&\qquad + \|\partial_{tt} z^*\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}\lambda^*\|_{L^\infty(I,L^2(\Omega))} + \|\theta_0\|_{H^2(\Omega)} + \|a_d\|_{H^2(\Omega)} + \|\partial_t u^*\|_{L^2(I)}\Big).
\end{aligned}$$

Proof: Write $z^*(t_{n-1}) - z^{n-1,*}_{hk} = (z^*(t_{n-1}) - \mathcal{R}_h z^*(t_{n-1})) + (\mathcal{R}_h z^*(t_{n-1}) - z^{n-1,*}_{hk}) = \zeta^{z,n-1} + \eta^{z,n-1}$. Also, write $\lambda^*(t_{n-1}) - \lambda^{n-1,*}_{hk} = (\lambda^*(t_{n-1}) - P_h\lambda^*(t_{n-1})) + (P_h\lambda^*(t_{n-1}) - \lambda^{n-1,*}_{hk}) = \zeta^{\lambda,n-1} + \eta^{\lambda,n-1}$. Subtracting (3.3.35) from (3.2.15), we obtain, at $t = t_{n-1}$;

$$\begin{aligned}-\rho c_p(\phi, \partial_t z^*(t_{n-1}) - \tilde{\partial} z_{hk}^{n-1,*}) \;&+\; \mathcal{K}(\bigtriangledown \phi, \bigtriangledown (z^*(t_{n-1}) - z_{hk}^{n-1,*})) \\ &= -(\phi, f_\theta(\theta^*(t_{n-1}), a^*(t_{n-1}))(\rho L z^*(t_{n-1}) - \lambda^*(t_{n-1}))) \\ &\quad - (\phi, f_\theta(\theta_{hk}^{n-1,*}, a_{hk}^{n-1,*})(\rho L z_{hk}^{n-1,*} - \lambda_{hk}^{n-1,*})) \\ &\quad + \beta_2(\phi, [\theta^*(t_{n-1}) - \theta_m]_+ - [\theta_{hk}^{n-1,*} - \theta_m]_+),\end{aligned}$$

where $\phi \in V_h$. Using Proposition 2.2.1 and Cauchy-Schwarz inequality , we find that

$$\begin{aligned}-\rho c_p(\phi, \tilde{\partial}\eta^{z,n-1}) \;&+\; \mathcal{K}(\bigtriangledown\phi, \bigtriangledown\eta^{z,n-1}) \\ &\leq C\Big(\|\rho L(z^*(t_{n-1}) - z_{hk}^{n-1,*}) - (\lambda^*(t_{n-1}) - \lambda_{hk}^{n-1,*})\| \, \|\phi\| + \|\tilde{\partial}\zeta^{z,n-1}\| \, \|\phi\| \\ &\quad + \|\tilde{\partial} z^*(t_{n-1}) - \partial_t z^*(t_{n-1})\| \, \|\phi\| + \|\zeta^{z,n-1}\| \, \|\phi\| + \|\theta^*(t_{n-1}) - \theta_{hk}^{n-1,*}\| \, \|\phi\| \Big).\end{aligned}$$

Putting $\phi = \eta^{z,n-1}$ and using Young's inequality, we obtain

$$\begin{aligned}-\tilde{\partial}\|\eta^{z,n-1}\|^2 \;&\leq\; C\Big(\|\eta^{z,n-1}\|^2 + \|\zeta^{z,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2 + \|\zeta^{\lambda,n-1}\|^2 + \|\tilde{\partial}\zeta^{z,n-1}\|^2 \\ &\quad + \|\tilde{\partial} z^*(t_{n-1}) - \partial_t z^*(t_{n-1})\|^2 + \|\theta^*(t_{n-1}) - \theta_{hk}^{n-1,*}\| \Big), \\ &\leq\; C\Big(\|\eta^{z,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2 + R_{n-1}^3 \Big), \qquad (3.3.37)\end{aligned}$$

where
$R_{n-1}^3 = \|\zeta^{z,n-1}\|^2 + \|\zeta^{\lambda,n-1}\|^2 + \|\tilde{\partial} z^*(t_{n-1}) - \partial_t z^*(t_{n-1})\|^2 + \|\tilde{\partial}\zeta^{z,n-1}\|^2 + \|\theta^*(t_{n-1}) - \theta_{hk}^{n-1,*}\|^2$.
Subtracting (3.3.12) from (3.2.13), we obtain at $t = t_{n-1}$

$$\begin{aligned}-(\chi, \tilde{\partial}\eta^{\lambda,n-1}) \;&=\; -(\chi, f_a(\theta^*(t_{n-1}), a^*(t_{n-1}))(\rho L z^* - \lambda^*) - f_a(\theta_{hk}^{n-1,*}, a_{hk}^{n-1,*})(\rho L z_{hk}^{n-1,*} - \lambda_{hk}^{n-1,*})) \\ &\quad + (\chi, \tilde{\partial}\lambda^*(t_{n-1}) - \partial_t\lambda^*(t_{n-1})) + (\chi, \tilde{\partial}\zeta^{\lambda,n-1}),\end{aligned}$$

where $\chi \in V_h$. Putting $\chi = \eta^{\lambda,n-1}$ and using Proposition 2.2.1, we find that

$$\begin{aligned}-\tilde{\partial}\|\eta^{\lambda,n-1}\|^2 \;&\leq\; C\Big(\|\zeta^{z,n-1}\|^2 + \|\eta^{z,n-1}\|^2 + \|\zeta^{\lambda,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2 \\ &\quad + \|\tilde{\partial}\lambda^*(t_{n-1}) - \partial_t\lambda^*(t_{n-1})\|^2 \Big). \qquad (3.3.38)\end{aligned}$$

Let $R^4_{n-1} = \|\zeta^{\lambda,n-1}\|^2 + \|\zeta^{z,n-1}\|^2 + \|\tilde{\partial}\lambda^*(t_{n-1}) - \partial_t\lambda^*(t_{n-1})\|^2$. Adding (3.3.37) and (3.3.38), we obtain

$$-\tilde{\partial}\Big(\|\eta^{z,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2\Big) \;\; \leq \;\; C\Big(\|\eta^{z,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2 + R^3_{n-1} + R^4_{n-1}\Big).$$

Summing up from n to $N+1$ and using the fact that $z^*(T) = 0$ and $z^*_h(T) = 0$, we arrive at

$$\|\eta^{z,n-1}\|^2 + \|\eta^{\lambda,n-1}\|^2 \;\; \leq \;\; \|\eta^{\lambda,N}\|^2 + Ck\sum_{m=n-1}^{N}(\|\eta^{z,m}\|^2 + \|\eta^{\lambda,m}\|^2 + R^z_m + R^\lambda_m).$$

From Lemma 3.2.1 and (3.2.37), we have $\|\zeta^{z,j}\| \leq Ch^2\|z^*\|_{H^2(\Omega)}$ and $\|\zeta^{\lambda,j}\| \leq Ch^2\|\lambda^*\|_{H^2(\Omega)}$, $j = 1, 2, \cdot\cdot\cdot, N$, respectively. Also, using same arguments as in (3.3.31), we have

$$\|\tilde{\partial}\zeta^{z,j}\| \;\; \leq \;\; Ch^2\|\partial_t z^*\|_{L^\infty(I,H^2(\Omega))}, \tag{3.3.39}$$

and applying the same steps as in (3.3.32), we obtain

$$\|\tilde{\partial}z^*(t_{n-1}) - \partial_t z^*(t_{n-1})\| \leq Ck\|\partial_{tt}z^*\|_{L^\infty(I,L^2(\Omega))}. \tag{3.3.40}$$

A use of Lemma 3.2.1 with (3.3.25), (3.3.39) and (3.3.40) implies that

$$\begin{aligned} R^3_{n-1} \;\; \leq \;\; & C(h^4 + k^2)\Big(\|z^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\lambda^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_t z^*\|^2_{L^\infty(I,H^2(\Omega))} + \|\partial_{tt}z^*\|^2_{L^\infty(I,L^2(\Omega))}\Big) \\ & + \|\theta^*(t_{n-1}) - \theta^{n-1,*}_{hk}\|. \end{aligned}$$

Similarly, we find that $R^4_{n-1} \leq C(h^4+k^2)\Big(\|\lambda^*\|^2_{L^\infty(I,H^2(\Omega))}+\|z^*\|^2_{L^\infty(I,H^2(\Omega))}+\|\partial_{tt}\lambda^*\|^2_{L^\infty(I,L^2(\Omega))}\Big)$. Hence, we obtain

$$\begin{aligned} \|\eta^{z,n-1}\| \;\; + \;\; & \|\eta^{\lambda,n-1}\| \\ \leq \;\; & C\Big((h^2 + k)\Big(\|a_d\|_{H^2(\Omega)} + \|z^*\|_{L^\infty(I,H^2(\Omega))} + \|\lambda^*\|_{L^\infty(I,H^2(\Omega))} + \|\partial_t z^*\|_{L^\infty(I,H^2(\Omega))} \\ & + \|\partial_{tt}z^*\|_{L^\infty(I,L^2(\Omega))} + \|\partial_{tt}\lambda^*\|_{L^\infty(I,L^2(\Omega))}\Big) + \|\theta^*(t_{n-1}) - \theta^{n-1,*}_{hk}\| + \|a^*(T) - a^{N,*}_{hk}\| \\ & + \sum_{m=n-1}^{N}(\|\eta^{z,m}\| + \|\eta^{\lambda,m}\|)\Big). \end{aligned}$$

Using Gronwall's lemma and Theorem 3.3.1, we arrive at the required result. This completes the rest of the proof. □

In order to completely discretize the problem (3.2.6)-(3.2.10) we choose discontinuous 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:

$$\begin{aligned}
\min_{u_\sigma \in U_{d,ad}} \ J(\theta_\sigma, a_\sigma, u_\sigma) & \qquad \text{subject to} && (3.3.41)\\
\sum_{n=1}^{N}(\partial_t a_\sigma, w)_{I_n,\Omega} + \sum_{n=1}^{N-1}([a_\sigma]_n, w_n^+) \ &+ \ (a_{\sigma,0}^+, w_0^+) = (f(\theta_\sigma, a_\sigma), w)_{I,\Omega}, && (3.3.42)\\
a_\sigma(0) \ &= \ 0, && (3.3.43)\\
\rho c_p \sum_{n=1}^{N}(\partial_t \theta_\sigma, v)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown\theta_\sigma, \bigtriangledown v)_{I,\Omega} \ &+ \ \rho c_p \sum_{n=1}^{N-1}([\theta_\sigma]_n, v_n^+) + \rho c_p(\theta_{\sigma,0}^+, v_0^+) \\
&= \ -\rho L(f(\theta_\sigma, a_\sigma), v)_{I,\Omega} + (\alpha u_\sigma, v)_{I,\Omega}, + \rho c_p(\theta_0, v_0^+), \\
\theta_\sigma(0) \ &= \ \theta_0 && (3.3.44)
\end{aligned}$$

for all $(w, v) \in X_{hk}^q \times X_{hk}^q$.

Lemma 3.3.1. *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* (3.3.42)-(3.3.44)*, satisfies the following a priori bounds :*

$$\begin{aligned}
\sum_{n=1}^{N} \|\partial_t \theta_\sigma\|_{\Omega,I_n}^2 + \|\Delta_h \theta_\sigma\|_{\Omega,I}^2 \le C, & \qquad \sum_{n=1}^{N} \|\partial_t a_\sigma\|_{\Omega,I_n}^2 \le C,\\
\|\theta_\sigma\|^2 + \| \bigtriangledown \theta_\sigma\|_{\Omega,I}^2 \le C, & \qquad \|a_\sigma\|^2 \le C,
\end{aligned}$$

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

$$(\Delta_h v, w) = (\bigtriangledown v, \bigtriangledown w) \quad \forall v, \ w \in V_h.$$

The proof of this lemma is on the similar lines as the proof of [62, Theorem 4.6] and hence is omitted. □

The solution of (3.3.41)-(3.3.44) is characterized by the saddle point $(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*) \in X_{hk}^q \times X_{hk}^q \times X_{hk}^q \times X_{hk}^q \times U_{d,ad}$ of the Lagrangian functional given by

$$\begin{aligned}\mathcal{L}(\theta_\sigma, a_\sigma, z_\sigma, \lambda_\sigma, u_\sigma) &= J(\theta_\sigma, a_\sigma, u_\sigma) - \Bigg(\sum_{n=1}^{N} (\partial_t a_\sigma, \lambda_\sigma)_{I_n} + \sum_{n=1}^{N-1} ([a_\sigma]_n, \lambda_{\sigma,n}^+) + (a_{\sigma,0}^+, \lambda_{\sigma,0}^+) \\ &\quad - (f(\theta_\sigma, a_\sigma), \lambda_\sigma)_{I,\Omega} \Bigg) - \Bigg(\rho c_p \sum_{n=1}^{N} (\partial_t \theta_\sigma, z_\sigma)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \theta_\sigma, \bigtriangledown z_\sigma)_{I,\Omega} \\ &\quad + \rho c_p \sum_{n=1}^{N-1} ([\theta_\sigma]_n, z_{\sigma,n}^+) + \rho c_p (\theta_{\sigma,0}^+, z_{\sigma,0}^+) + \rho L (f(\theta_\sigma, a_\sigma), z_\sigma)_{I,\Omega} \\ &\quad - (\alpha u_\sigma, z_\sigma)_{I,\Omega} - \rho c_p (\theta_0, z_{\sigma,0}^+) \Bigg)\end{aligned}$$

The saddle point $(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*) \in X_{hk}^q \times X_{hk}^q \times X_{hk}^q \times X_{hk}^q \times U_{d,ad}$ is determined by the KKT system given by:

State equations:

$$\mathcal{L}_z(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(v) = 0 \quad \forall v \in X_{hk}^q, \tag{3.3.45}$$

$$\mathcal{L}_\lambda(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(w) = 0 \quad \forall w \in X_{hk}^q. \tag{3.3.46}$$

Adjoint equations:

$$\mathcal{L}_\theta(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(\phi) = 0 \quad \forall \phi \in X_{hk}^q, \tag{3.3.47}$$

$$\mathcal{L}_a(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(\psi) = 0 \quad \forall \psi \in X_{hk}^q. \tag{3.3.48}$$

First order optimality condition:

$$\mathcal{L}_u(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(p - u_\sigma^*) \geq 0 \quad \forall p \in U_{d,ad}. \tag{3.3.49}$$

The state system (3.3.41)-(3.3.44) is obtained from (3.3.45)-(3.3.46). The adjoint system of (3.3.41)-(3.3.44) obtained from (3.3.47)-(3.3.48) is defined by:

Find $(z_\sigma^*, \lambda_\sigma^*) \in X_{hk}^q \times X_{hk}^q$ such that

$$\begin{aligned}-\sum_{n=1}^{N} (\psi, \partial_t \lambda_\sigma^*)_{I_n} - \sum_{n=1}^{N-1} (\psi_n, [\lambda_\sigma^*]_n) &+ (\psi, f_a(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega} = 0 \\ \lambda_{\sigma,N}^* &= \beta_1 (a_\sigma^*(T) - a_d),\end{aligned} \tag{3.3.50}$$

$$
\begin{aligned}
-\rho c_p \sum_{n=1}^{N} (\phi, \partial_t z_\sigma^*)_{I_n} + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z_\sigma^*)_{I,\Omega} &- \rho c_p \sum_{n=1}^{N-1} (\phi_n, [z_\sigma^*]_n) + (\phi, f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega} \\
&= \beta_2(\phi, [\theta_\sigma^* - \theta_m]_+)_{I,\Omega}, \qquad (3.3.51) \\
z_{\sigma,N}^* &= 0, \qquad (3.3.52)
\end{aligned}
$$

for all $(\psi, \phi) \in X_{hk}^q \times X_{hk}^q$. Moreover from (3.3.49), z_σ^* satisfies the variational inequality,

$$
\left(\beta_3 u_\sigma^* + \int_\Omega \alpha z_\sigma^* dx, p - u_\sigma^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{d,ad}. \qquad (3.3.53)
$$

(3.3.42)-(3.3.44) forms a system of non-linear equations with Lipschitz continuous right hand side and therefore, admits a unique local solution. It ensures the existence of a control-to-state mapping $u_\sigma \mapsto (\theta_\sigma, a_\sigma) = (\theta_\sigma(u_\sigma), a_\sigma(u_\sigma))$ through (3.3.42)-(3.3.44). By means of this mapping, we introduce the reduced cost functional $j_\sigma : U_{d,ad} \longrightarrow \mathbb{R}$ as

$$
j_\sigma(u_\sigma) = J(\theta_\sigma(u_\sigma), a_\sigma(u_\sigma), u_\sigma). \qquad (3.3.54)
$$

Then the optimal control problem can be equivalently reformulated as

$$
\min_{u_\sigma \in U_{d,ad}} j_\sigma(u_\sigma). \qquad (3.3.55)
$$

Theorem 3.3.3. *Let u_σ^* be the optimal control of* (3.3.41)-(3.3.44)*. Then, $\lim_{\sigma\to 0} u_\sigma^* = u^*$ exists in $L^2(I)$ and u^* is an optimal control of* (3.2.6)-(3.2.10)*.*

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

$$
u_\sigma^* \longrightarrow u^* \text{ weakly in } L^2(I). \qquad (3.3.56)
$$

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 (3.3.42)-(3.3.44). Thus from Lemma 3.3.1, we have

$$\begin{aligned} \theta_\sigma^* &\longrightarrow \theta^* \text{ weakly in } H^{1,1}, && (3.3.57)\\ \theta_\sigma^* &\longrightarrow \theta^* \text{ strongly in } C(I, L^2(\Omega)), && (3.3.58)\\ a_\sigma^* &\longrightarrow a^* \text{ weak}^* \text{ in } W^{1,\infty}(I, L^\infty(\Omega)), && (3.3.59)\\ a_\sigma^* &\longrightarrow a^* \text{ strongly in } L^\infty(I, L^2(\Omega)). && (3.3.60) \end{aligned}$$

Now passing limit as $\sigma \to 0$, using (3.3.57)-(3.3.60) and Proposition 2.2.1 in (3.3.42)-(3.3.44), we obtain that (u^*, θ^*, a^*) is an admissible solution for the optimal control problem (3.2.6)-(3.2.10). 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 (3.2.6)-(3.2.10). Consider the auxiliary problem

$$\begin{aligned} \sum_{n=1}^{N}\Big((\partial_t a_\sigma, w)_{\Omega,I_n} + ([a_\sigma]_{n-1}, w_{n-1}^+)\Big) &= \sum_{n=1}^{N}(f(\theta_\sigma, a_\sigma), w), && (3.3.61)\\ a_\sigma(0) &= 0, && (3.3.62)\\ \sum_{n=1}^{N}\Big(\rho c_p(\partial_t\theta_\sigma, v)_{\Omega,I_n} + K(\nabla\theta_\sigma, \nabla v)_{\Omega,I_n} + ([\theta_\sigma]_{n-1}, v_{n-1}^+)\Big) &= \sum_{n=1}^{N}\Big(- \rho L(f(\theta_\sigma, a_\sigma), v)_{\Omega,I_n} \\ &\quad +(\alpha\pi_k\bar{u}^*, v), && (3.3.63)\\ \theta_\sigma(0) &= \theta_0, && (3.3.64) \end{aligned}$$

for all $(w, v) \in X_{hk}^q \times X_{hk}^q$. Then, there exists a solution to (3.3.61)-(3.3.64), say $(\bar{\theta}_\sigma, \bar{a}_\sigma) \in H^{1,1} \times W^{1,\infty}(I, L^\infty(\Omega))$. Similar to (3.3.57)-(3.3.60), we arrive at

$$\begin{aligned} \bar{\theta}_\sigma &\longrightarrow \bar{\theta} \text{ weakly in } H^{1,1}, && (3.3.65)\\ \bar{\theta}_\sigma &\longrightarrow \bar{\theta} \text{ strongly in } C(I, L^2(\Omega)), && (3.3.66)\\ \bar{a}_\sigma &\longrightarrow \bar{a} \text{ weakly in } W^{1,\infty}(I, L^\infty(\Omega)), && (3.3.67)\\ \bar{a}_\sigma &\longrightarrow \bar{a} \text{ strongly in } L^\infty(I, L^2(\Omega)). && (3.3.68) \end{aligned}$$

Now letting $\sigma \to 0$ in (3.3.61)-(3.3.64), we obtain that $(\bar{\theta}, \bar{a})$ is a unique solution of (3.2.7)-(3.2.10) with respect to the control $\bar{u}^*$. Since the solution to (2.3.3)-(2.3.6) 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 (3.3.41)-(3.3.44), we have

$$j(u_\sigma^*) \leq j(\pi_k \bar{u}^*). \tag{3.3.69}$$

Now letting
sigma → 0 in (3.3.69) and using (3.3.56), we obtain

$$j(u^*) \leq j(\bar{u}^*). \tag{3.3.70}$$

Note from (3.3.70) 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 $\lim\limits_{\sigma \to 0} \|u_\sigma^* - u\|_{L^2(I)} = 0$. Since $u_\sigma^* \longrightarrow u^*$ weakly in $L^2(\Omega)$, it is enough to show that $\lim\limits_{\sigma \to 0} \|u_\sigma^*\|_{L^2(I)} = \|u^*\|_{L^2(I)}$. Using Lemma 3.3.1 and (3.3.56), we find that

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

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

3.4 Numerical Experiment

For the purpose of numerical experiment, we use cG(1)dG(0) space-time discretization for the state and adjoint variables and dG(0) for the control variable. We have used non-linear conjugate method [84] to evaluate the optimal control for the completely discrete problem (3.3.41)-(3.3.44). The implementations in Chapters 3, 4 and 5 have been performed using the software package *deal.II* [7].

Non-Linear Conjugate Gradient Method

```
step 1: Initialize $l = 0, m = 0, r = -j'_\sigma(u_\sigma)$, $d = r, \delta_{new} = r^T r$ and $\delta_0 = \delta_{new}$.
        While($k < k_{max}$ and $\delta_{new} > \epsilon^2 \delta_0$)
step 2: Initialize $n = 0, \delta_d = d^T d, \alpha_1 = -\alpha_0$ and $\eta_{prev} = j'_\sigma(u_\sigma + \alpha_0 d)^T d$.
        Do
          (i) $\eta = j_\sigma(u_\sigma)^T d$.
          (ii) $\alpha_1 = \alpha \dfrac{\eta}{\eta_{prev} - \eta}$.
          (iii) $u_\sigma = u_\sigma + \alpha d$.
          (iii) $\eta_{prev} = \eta$ and $n = n + 1$.
        while($n < n_{max}$ and $\alpha^2 \delta_d > \epsilon^2$)
step 3: $r_{old} = r$ and $s = r_{old}^T r$.
step 4: $r = -j'_\sigma(u_\sigma)$ and $\delta_{old} = \delta_{new}$.
step 5: $\delta_{new} = r^T r$ and $\beta = \dfrac{\delta_{new} - s}{\delta_{old}}$.
step 6: $m = m + 1$
        if ($\beta < 0$)
          (a) $d = r$ and $m = 0$
        else
          (b) $d = r + \beta d$.
step 7: $l = l + 1$.
        While ends.
```

Physical Data [84]: The computational domain is chosen as $\Omega = (0, 5) \times (-1, 0)$ and T is chosen as 5.25. In (3.2.7)-(3.2.10), we consider the physical data as $\rho c_p = 4.91 \frac{J}{cm^3 K}, k = 0.64 \frac{J}{cmKs}$ and

$\rho L = 627.9 \frac{J}{cm^3}$. The regularized monotone function $\mathcal{H}_\epsilon$ is chosen as

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

The initial temperature θ_0 and the melting temperature θ_m are chosen as 20 and 1800, respectively. The pointwise data for $a_{eq}(\theta)$ and $\tau(\theta)$ are given by

θ	730	830	840	930
$a_{eq}(\theta)$	0	0.91	1	1
$\tau(\theta)$	1	0.2	0.18	0.05

We use a cubic spline interpolation to obtain approximations for the functions $a_{eq}(\theta)$ and $\tau(\theta)$. The shape function $\alpha(x, y, t)$ is given by $\alpha(x, y, t) = \frac{4k_1A}{\pi D^2}exp(-\frac{2(x-vt)^2}{D^2})exp(k_1y)$, 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}. Also, we choose $\beta_1 = 5000, \beta_2 = 1000$ and $\beta_3 = 10^{-3}$. The main aim of this experiment is to achieve a constant hardening depth of 1mm , see Figure 3.1, with expected order of convergence $\mathcal{O}(h^2 + k)$ 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) as 1404.

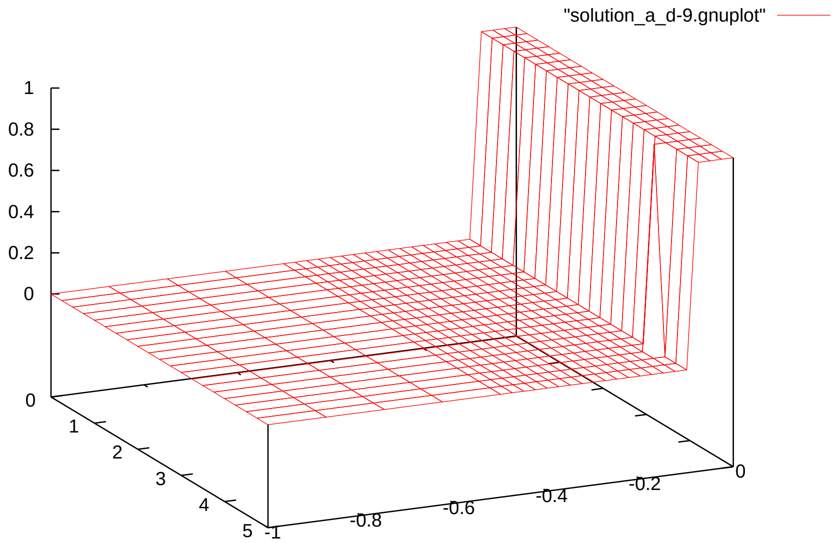


Figure 3.1: **Goal a_d to be achieved for the volume fraction of austenite**

When the finite element method is applied, the mesh used for space discretization is much 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. A comparison of Figure 3.1 and Figure 3.2(a) shows that the goal of uniform hardening depth is nearly achieved. Also, the state constraint that $\|\theta\|_{L^\infty(Q)} < 1800$ is satisfied, since $\|\theta_\sigma\|_{L^\infty(Q)} < 1200$, see Figure 3.2(b). Figure 3.3 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 confirm with those obtained in [84], though error estimates have not been developed [84]. Figure 3.4 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 3.5, 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 near 2, which confirms the convergence obtained in Theorem 3.3.3. The finite element *a priori* estimates developed in Theorem 3.3.1 yields the order of convergence

$$\|\theta_\epsilon - \theta_{\epsilon,\sigma}\| + \|a_\epsilon - a_{\epsilon,\sigma}\| = \mathcal{O}(h^2 + k),$$

where $(\theta_{\epsilon,\sigma}, a_{\epsilon,\sigma})$ is the solution to (3.2.7)-(3.2.10) obtained after a finite element discretization, h and k being the space and time discretization parameters respectively. Therefore, using Theorem 2.2.2, we have

$$\|\theta - \theta_{\epsilon,\sigma}\| + \|a - a_{\epsilon,\sigma}\| = \mathcal{O}(h^2 + k + \epsilon). \tag{3.4.1}$$

Figure 3.7 and 3.8 represents the $\|E_1\|$, $\|E_2\|$ and $\|e(u)\|$, respectively, as a function of regularization parameter ϵ in the log-log scale. For the purpose of implementation, the values of epsilon were taken as $\{0.5, 0.10, 0.15, 0.20, 0.25\}$. The numerical results obtained confirms the theoretical results obtained in Theorem 3.3.1, Theorem 3.3.3 and (3.4.1).

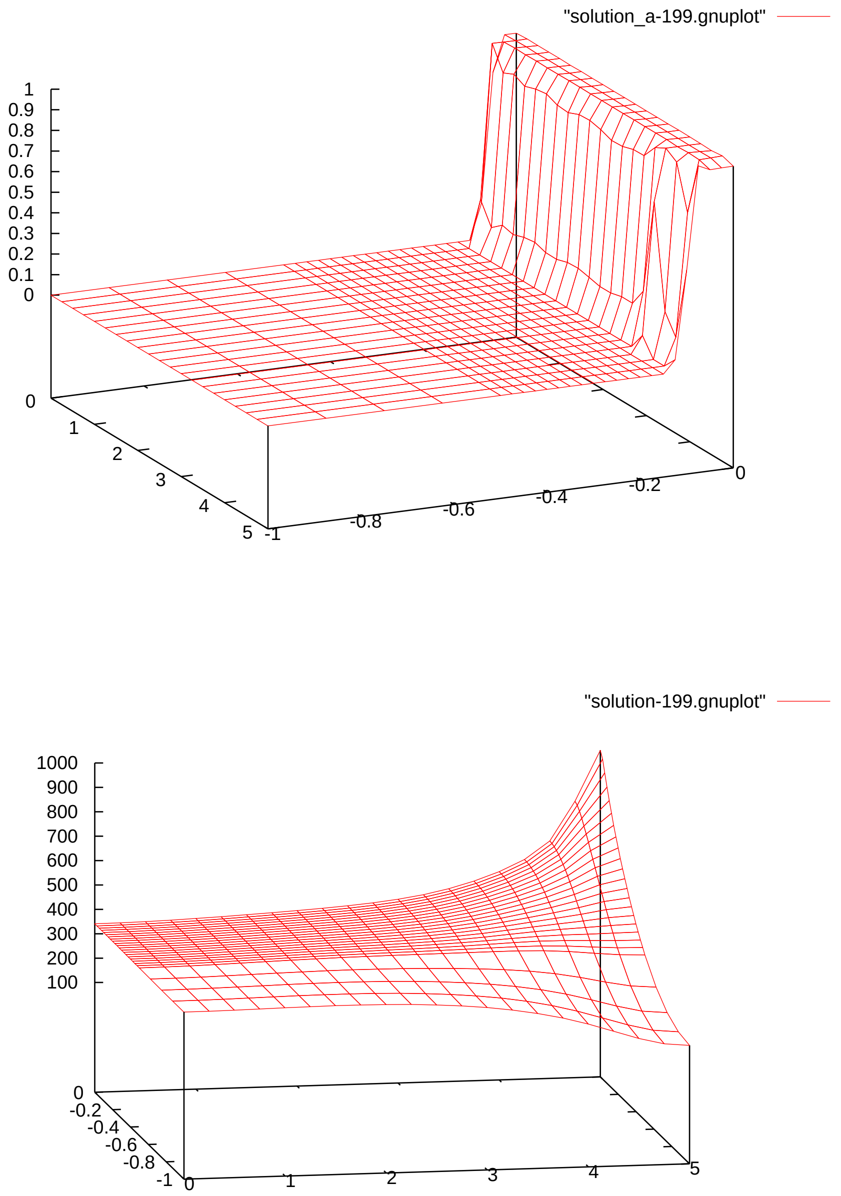


Figure 3.2: **(a) The volume fraction of the austenite at time** $t = T$ **(b) The temperature at time** $t = T$

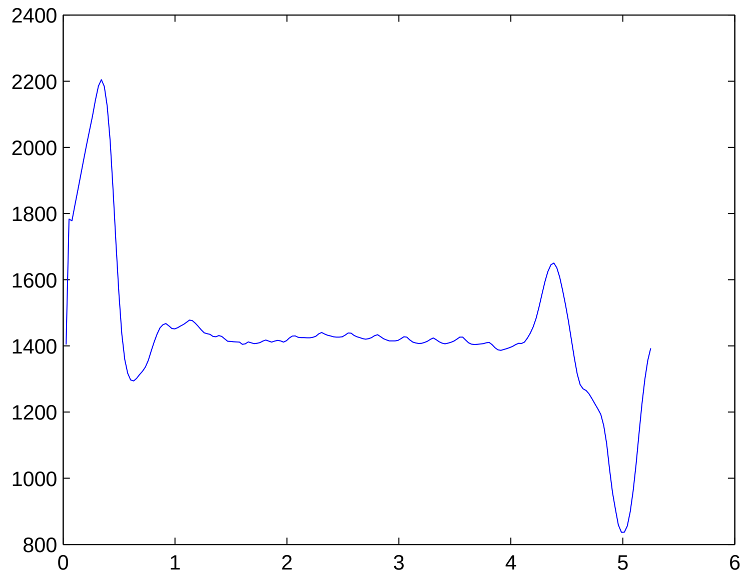


Figure 3.3: **Laser energy**

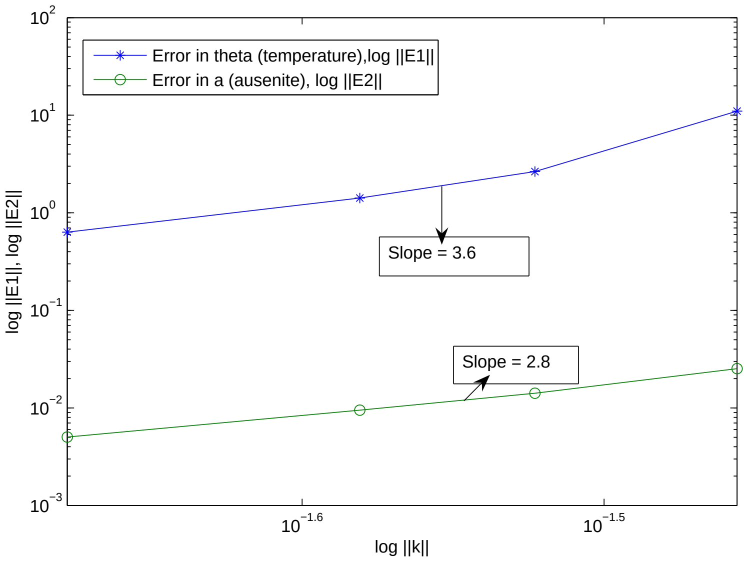


Figure 3.4: **Refinement of the time steps for number of** 525 **nodes in spatial triangulation**

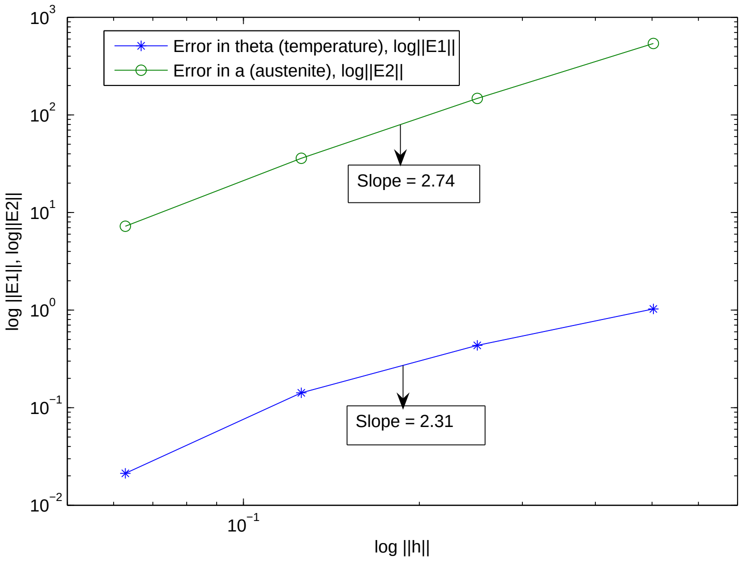


Figure 3.5: **Refinement of spatial triangulation for 200 time steps.**

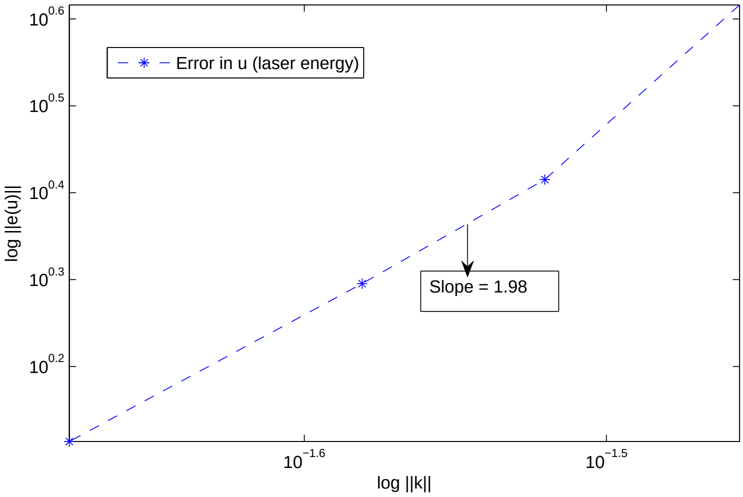


Figure 3.6: **Evolution of control error.**

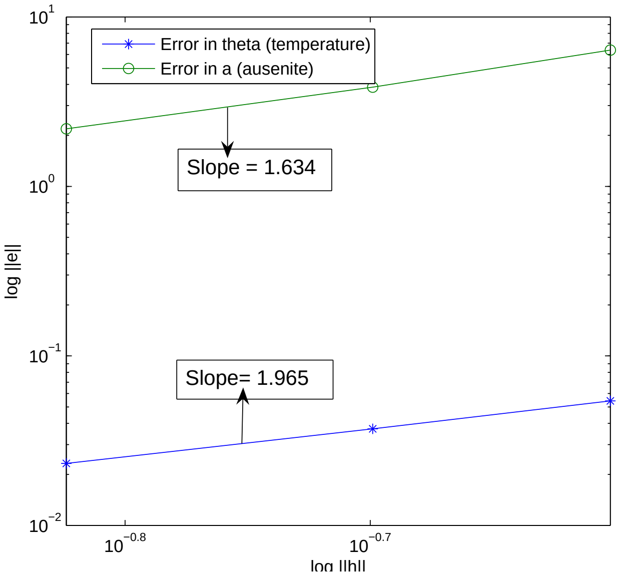


Figure 3.7: **Evolution of state error as a function of ϵ.**

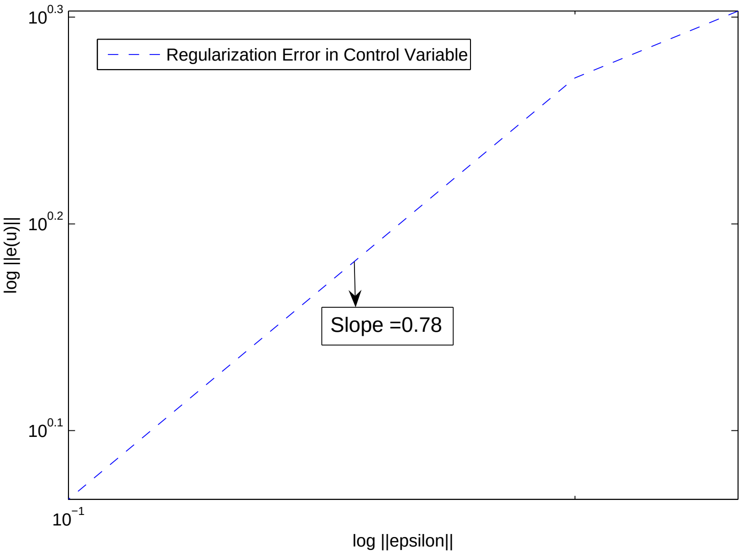


Figure 3.8: **Evolution of control error as a function of ϵ.**

3.5 Summary

This chapter discusses the convergence of a cG(1)dG(0)-dG(0) space-time-control discretization method for the laser surface hardening of steel problem. It has been shown that the approximate solution $(\theta_\sigma^*, a_\sigma^*, u_\sigma^*) \in X_{hk}^0 \times X_{hk}^0 \times U_{d,ad}$ converges to the solution of the regularized problem at the rate of $(h^2 + k)$, where h and k are the space and time discretization parameters. Also, numerical experiments attached in the last section shows that the solution of the regularized problem converges to that of the original problem atleast with the order of convergence $\mathcal{O}(h^2 + k + \epsilon)$.

Chapter 4

A Priori Error Estimates: A Discontinuous Galerkin Space-Time Method

4.1 Introduction

In this chapter, we discuss an *hp*-Discontinuous Galerkin Finite Element Method (*hp*-DGFEM) for the optimal control problem of laser surface hardening of steel. The space discretization is based on *hp*-DGFEM, time and control discretizations are based on a DGFEM. *A priori* error estimates have been developed for state, adjoint and control errors. Numerical experiment presented justifies the theoretical order of convergence obtained.

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 and are easy to implement than the continuous finite element method.

The use of DGFEM for elliptic and parabolic problems started with the work of Douglas, Dupont [20] and Wheeler [86] in the 70's. These methods are generalization of work by Nitsche [66] for treating Dirichlet boundary condition by introduction of a penalty term on boundary. In 1973, Babuška [5] introduced another penalty method to impose the Dirichlet boundary condition weakly. Interior Penalty(IP) methods by Arnold [3] and Wheeler [86] 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, Babuška and Baumann [67] proposed DGFEM for advection diffusion problems which is based on a non-symmetric formulation. This method is known to be stable when the degree of approximation is greater or equal to 2, see [67], [75]. In Houston *et al.* [36], discontinuous *hp* finite element methods are studied for diffusion reaction problems. For a review of work on DG methods for elliptic problems, we refer to [4], [70]. [28]-[30] discuss DG methods for quasilinear and strongly non-linear elliptic problems. In [74] and [75], a non symmertric interior penalty DGFEM is analyzed for elliptic and non-linear parabolic problems, respectively. For a detailed description of DGFEM for elliptic and parabolic problems, we refer to [73].

Since the laser surface hardening of steel problem is an optimal control problem, adjoint consistency becomes important. Therefore, in this chapter, a symmetric version of *hp*-DGFEM has been introduced and analyzed for the optimal control problem of laser surface hardening of steel. A similar *hp*-version of interior penalty discontinuous Galerkin method for semilinear parabolic equation with mixed Dirichlet and Neumann boundary conditions has been analyzed in [50]. 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 [74].

The outline of this chapter is as follows. This section is introductory in nature. Section 4.2 includes necessary preliminaries and a weak formulation of the regularized laser surface hardening of steel problem. In Section 4.3, 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 adjoint variables. In Section 4.4, a space-time discretization using DGFEM in time and *hp*-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 4.5, results of numerical experiment are presented.

4.2 Preliminaries

1. (**Finite Elements**) Let $\mathcal{T}_h = \{K, K \subset \Omega\}$ be a shape regular finite element subdivision of Ω 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 γh_K in its interior [17]. 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 $\{F_K\}$ be a family of invertible maps such that each F_K maps from $\hat{K}$ to K, see Figure 4.1. Let $h = \max_{K \in \mathcal{T}_h} h_K$.

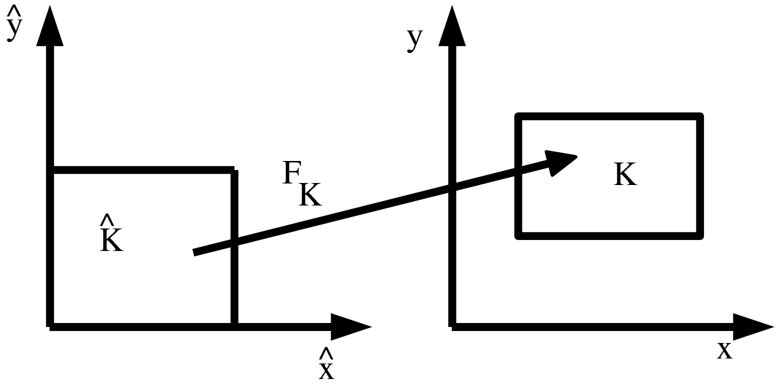


Figure 4.1: **An example of construction of finite elements**

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

$$\begin{aligned}
\mathcal{E} &= \{e : e = \partial K \cap \partial K' \text{ or } e = \partial K \cap \partial\Omega, K, K' \in \mathcal{T}_h\}, \\
\mathcal{E}_{int} &= \{e \in \mathcal{E} : e = \partial K \cap \partial K', K, K' \in \mathcal{T}_h\}, \\
\mathcal{E}_\partial &= \{e \in \mathcal{E} : e = \partial K \cap \partial\Omega, K \in \mathcal{T}_h\},
\end{aligned}$$

Note that the definition of the triangulation $\mathcal{T}_h$ admits atmost one hanging node along each side of K.

2. (**Discontinuous Spaces**) On the subdivision $\mathcal{T}_h$, we define the required broken Sobolev spaces for $s = 1, 2$ as $H^s(\Omega, \mathcal{T}_h) = \Big\{ w \in L^2(\Omega) : w|_K \in H^s(K), K \in \mathcal{T}_h \Big\}$.
The associated broken norm and semi-norm are defined by:

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

Also, let $\mathcal{U} = \{w \in H^2(\Omega, \mathcal{T}_h) : w, \bigtriangledown w.n$ are continuous along each $e \in \mathcal{E}_{int}\}$.

3. (**Discrete 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 $H^1(\Omega, \mathcal{T}_h)$,

$$S^p = \bigg\{ w \in L^2(\Omega) : w|_K o F_K \in Q_{p_K}(\hat{K}), K \in \mathcal{T}_h \bigg\},$$

   where $p = \{p_K : K \in \mathcal{T}_h\}$ and $F = \{F_K : K \in \mathcal{T}_h\}$, F_K being the affine map from $\hat{K}$ to K.

4. (**Jump and Average of a Function**) For $e_K \in \mathcal{E}_{int}$, the jump and average of $w \in H^1(\Omega, \mathcal{T}_h)$ are defined by:

$$\{w\} = \frac{1}{2}\bigg((w|_K)|_{e_K} + (w|_{K'})|_{e_K} \bigg) \quad , \quad [w] = (w|_K)|_{e_K} - (w|_{K'})|_{e_K}.$$

   The jump and average on $e_K \in \mathcal{E}_\partial$ are defined as

$$\{w\} = (w|_K)|_{e_K}, \qquad\qquad [w] = (w|_K)|_{e_K}, \quad \text{respectively.}$$

5. (**Broken Energy Norm**) We define the broken energy norm for $w \in H^1(\Omega, \mathcal{T}_h)$ as

$$|||w||| = \bigg(\sum_{K \in \mathcal{T}_h} \|w\|^2_{H^1(K)} + \mathcal{J}^\gamma(w, w) \bigg)^{1/2},$$

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

Assumptions on the mesh and degree of approximation

Assumption (P):

- The finite element subdivision $\mathcal{T}_h$ satisfies the *bounded local variation* condition in the sense that if $|\partial K \cap \partial K'| > 0$, for any K and $K' \in \mathcal{T}_h$, then there exists a constant κ independent of h_K, $h_{K'}$ such that

$$\frac{h_K}{h_{K'}} \leq \kappa.$$

In particular, this implies that for any element K, the number of neighboring elements $K' \in \mathcal{T}_h$ with $|\partial K \cap \partial K'| > 0$ is bounded by N_κ uniformly, for some positive integer N_κ.

- The discontinuous finite element space $S^{\mathbf{p}}$ satisfies the following *bounded local variation*: If $|\partial K \cap \partial K'| > 0$, for any K and $K' \in \mathcal{T}_h$, then there exists a constant $\varrho > 0$ independent of p_K and $p_{K'}$ such that

$$\frac{p_K}{p_{K'}} \leq \varrho \; .$$

Here, $|\cdot|$ denotes the one dimensional Euclidean measure.

We now present some examples (see [17]) which satisfy the assumption (**P**).
(i) **Regular subdivision** is a subdivision of Ω into shape regular elements $K \in \mathcal{T}_h$ such that for any two elements K and K', the common portion $\partial K \cap \partial K'$ is either empty or a vertex of K or an entire edge e of K, that is, $e = \partial K \cap \partial K'$ and there is no other element $K_1 \in \mathcal{T}_h$ such that $|e \cap \partial K_1| > 0$.
(ii) **1-irregular subdivision** is a shape regular subdivision $\{K\}_{K\in\mathcal{T}_h}$ of Ω is such that for any side of an element K, there can be at most one hanging node.
From the assumptions (**P**) and shape regularity, it is easy to see that if $e_K \subset \partial K$ then there exist constants $c_1(\kappa)$, $c_2(\kappa)$, $c_3(\varrho)$ and $c_4(\varrho)$ which are independent of h and p such that

$$c_1(\kappa)h_K \leq |e_K| \leq c_2(\kappa)h_K, \qquad c_3(\varrho)p_K \leq p_{e_K} \leq c_4(\varrho)p_K, \tag{4.2.1}$$

where p_{e_K} is the degree of the polynomial used for the approximation of the unknown variables over the edge e_K.

Approximation Properties of Finite Element Spaces

Lemma 4.2.1. *[50]: Let $w|_K \in H^{s'}(K), s' \geq 0$. Then there exists a sequence $z^{h_K}_{p_k} \in Q_{p_K}(K)$, $p_K = 1, 2 \cdots$, such that for $0 \leq l \leq s'$,*

$$\begin{aligned} \|w - z^{h_K}_{p_k}\|_{H^l(K)} &\leq C\frac{h_K^{s-l}}{p_K^{s'-l}}\|w\|_{H^{s'}(K)} \qquad \forall K \in \mathcal{T}_h, \\ \|w - z^{h_K}_{p_k}\|_{L^2(e)} &\leq C\frac{h_K^{s-\frac{1}{2}}}{p_K^{s'-\frac{1}{2}}}\|w\|_{H^{s'}(K)} \quad \forall e \in \mathcal{E}_{int}, \end{aligned}$$

and

$$\| \bigtriangledown (w - z_{p_k}^{h_K})\|_{L^2(e)} \leq C \frac{h_K^{s-\frac{3}{2}}}{p_K^{s'-\frac{3}{2}}} \|w\|_{H^{s'}(K)} \quad \forall e \in \mathcal{E}_{int},$$

where $1 \leq s \leq \min(p_K + 1, s')$, $p_K \geq 1$, *and* C *is a constant independent of* w, h_K, *and* p_K, *but dependent on* s'.

Remark 4.2.1. *Given* $w \in H^2(\Omega, \mathcal{T}_h)$, *define the interpolant* $I_h w = \hat{w} \in S^p$ *by*

$$I_h w|_K = \hat{w}|_K = z_{p_k}^{h_K}(w|_K) \quad \forall K \in \mathcal{T}_h. \tag{4.2.2}$$

Trace Inequalities

Lemma 4.2.2. *[70, Appendix A.2] Let* $w \in H^{j+1}(K), K \in \mathcal{T}_h$. *Then, there exists a constant* $C > 0$ *such that*

$$\|w\|_{H^j(e_K)}^2 \leq C \left(\frac{1}{h_K} \|w\|_{H^j(K)}^2 + \|w\|_{H^j(K)} \| \bigtriangledown^{(j+1)} w\|_{L^2(K)} \right),$$

where $j = 0, 1$.

Lemma 4.2.3. *[75, Lemma 2.1] Let* $v_h \in Q_{p_K}(K)$. *Then, there exists a constant* $C > 0$ *such that*

$$\|\nabla^l v_h\|_{e_k} \leq C p_K h_K^{-1/2} \|\nabla^l v_h\|_K, \quad l = 0, 1. \tag{4.2.3}$$

Inverse Inequalities

Below, we state without proof a lemma on inverse inequalities. For a proof, we refer to [51, page no. 6], [15, Theorem 6.1].

Lemma 4.2.4. *Let* $v_h \in Q_{p_K}(K)$. *Then, for* $r \geq 2$, *there exists a constant* $C > 0$ *such that*

$$\|v_h\|_{L^r(K)} \leq C p_K^{1-2/r} h_K^{(2/r-1)} \|v_h\|_K, \tag{4.2.4}$$

$$|v_h|_{H^l(K)} \leq C p_K^2 h_K^{-1} |v_h|_{H^{l-1}(K)}, \quad l \geq 1 \tag{4.2.5}$$

$$\|v_h\|_{L^r(e_K)} \leq C p_K^{1-2/r} |e_K|^{(1/r-1/2)} \|v_h\|_{e_K}, \tag{4.2.6}$$

where $e_k \subset \partial K$ *is an edge.*

Now consider the regularized laser surface hardening of steel problem:

$$\min_{u\in U_{ad}} J(\theta, a, u) \text{ subject to} \tag{4.2.7}$$

$$\begin{aligned}
\partial_t a &= f(\theta, a) \quad \text{ in } Q, && (4.2.8)\\
a(0) &= 0 \quad \text{ in } \Omega, && (4.2.9)\\
\rho c_p \partial_t \theta - \mathcal{K} \triangle \theta &= -\rho L \partial_t a + \alpha u \quad \text{ in } Q, && (4.2.10)\\
\theta(0) &= \theta_0 \quad \text{ in } \Omega, && (4.2.11)\\
\frac{\partial \theta}{\partial n} &= 0 \quad \text{ on } \Sigma. && (4.2.12)
\end{aligned}$$

The weak formulation corresponding to (4.2.8)-(4.2.12), for a fixed $u \in U_{ad}$, reads as:
Find $(\theta, a) \in X \times Y$ such that

$$\begin{aligned}
(\partial_t a, w) &= (f(\theta, a), w) \quad \forall w \in H, && (4.2.13)\\
a(0) &= 0, && (4.2.14)\\
\rho c_p(\partial_t \theta, v) + \mathcal{K}(\bigtriangledown \theta, \bigtriangledown v) &= -\rho L(\partial_t a, v) + (\alpha u, v) \quad \forall v \in V, && (4.2.15)\\
\theta(0) &= \theta_0, && (4.2.16)
\end{aligned}$$

where $X = \{v \in L^2(I; V) : v_t \in L^2(I; V^*)\}$, $V = H^1(\Omega)$, $Y = H^1(I; L^2(\Omega))$, $H = L^2(\Omega)$ and $U_{ad} = \{u \in L^2(I) : \ 0 \le u \le u_{max} \text{ a.e. in } I\}$.
Therefore, the weak formulation for the optimal control problem can be stated as

$$\min_{u\in U_{ad}} J(\theta, a, u) \quad \text{subject to the constraints (4.2.13)-(4.2.16).} \tag{4.2.17}$$

The adjoint system of (4.2.17) is defined by (for explanations see Chapter 3):
Find $(z^*, \lambda^*) \in X \times Y$ such that

$$\begin{aligned}
-(\chi, \partial_t \lambda^*) = -(\chi, f_a(\theta^*, a^*) g(z^*, \lambda^*)), && (4.2.18)\\
\lambda^*(T) = \beta_1(a^*(T) - a_d), && (4.2.19)\\
-\rho c_p(\phi, \partial_t z^*) + \mathcal{K}(\bigtriangledown \phi, \bigtriangledown z^*) + (\phi, f_\theta(\theta^*, a^*) g(z^*, \lambda^*)) \ = \beta_2(\phi, [\theta^* - \theta_m]_+), && (4.2.20)\\
z^*(T) = 0, && (4.2.21)
\end{aligned}$$

for all $(\chi, \phi) \in H \times V$ and $g(z, \lambda) = \rho L z - \lambda$. Moreover, z^* satifies the following variational inequality

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

One can easily check that g is a Lipschitz continuous function.

4.3 *hp*-Discontinuous Galerkin Weak Formulation

Space Discretization

The *hp*-DGFEM formulation corresponding to (4.2.13)-(4.2.16) can be stated as:
Find $(\theta_h(t), a_h(t)) \in S^p \times S^p$, a.e. in $\bar{I}$ such that

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h} (\partial_t a_h, w)_K &= \sum_{K\in\mathcal{T}_h} (f(\theta_h, a_h), w)_K \quad \forall w \in S^p, && (4.3.1)\\
a_h(0) &= 0, && (4.3.2)\\
\rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t \theta_h, v)_K + B(\theta_h, v) &= -\rho L \sum_{K\in\mathcal{T}_h} (\partial_t a_h, v)_K + \sum_{K\in\mathcal{T}_h} (\alpha u_h, v)_K \ \forall v \in S^p, && (4.3.3)\\
\theta_h(0) = \theta_0, & && (4.3.4)
\end{aligned}$$

where $B(\theta, v) = \mathcal{K} \sum_{K\in\mathcal{T}_h} (\bigtriangledown\theta, \bigtriangledown v)_K - \mathcal{K} \sum_{e\in\mathcal{E}_{int}} \int_e \{\bigtriangledown\theta.n\}[v]de - \mathcal{K} \sum_{e\in\mathcal{E}_{int}} \int_e \{\bigtriangledown v.n\}[\theta]de + \mathcal{J}^\gamma(\theta, v)$
and $\mathcal{J}^\gamma(\theta, v) = \sum_{e\in\mathcal{E}_{int}} \frac{\gamma}{|e|} \int_e [\theta][v]de$, $\gamma > 0$ is the penalty parameter to be chosen later.

Remark 4.3.1. *Note that the bilinear form $B(\cdot,\cdot)$ is symmetric. Therefore,* (4.3.1)-(4.3.4) *corresponds to the SIPG formulation for the regularized laser surface hardening of steel problem.*

Let $\{\phi_1, \phi_2, \cdots, \phi_M\}$ 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 $v = \phi_j, w = \phi_j, j = 1, 2, \cdots, M$ in (4.3.1)-(4.3.4), we obtain

$$\begin{aligned}
\mathbf{A}\ \partial_t \bar{\mathbf{a}} &= \bar{\mathbf{F}}(\bar{\theta}, \bar{\mathbf{a}}), && (4.3.5)\\
\bar{\mathbf{a}}(0) &= 0, && (4.3.6)\\
\rho c_p \mathbf{A}\ \partial_t \bar{\theta} + \mathbf{B}\ \bar{\theta} &= -\rho L \bar{\mathbf{F}}(\bar{\theta}, \bar{\mathbf{a}}) + u_h(t)\bar{\alpha}, && (4.3.7)\\
\bar{\theta}(0) &= \theta_0, && (4.3.8)
\end{aligned}$$

where $\quad \bar{\mathbf{a}} = \Big(a_i(t)\Big)_{1\le i\le M}, \ \bar{\theta} = \Big(\theta_i(t)\Big)_{1\le i\le M}, \ \mathbf{A} = \Big(\sum_{K\in\mathcal{T}_h}(\phi_i,\phi_j)_K\Big)_{1\le i,j\le M},$

$$\mathbf{B} = \Big(B(\phi_i,\phi_j)\Big)_{1\le i,j\le M}, \bar{\mathbf{F}}(\bar{\theta},\bar{\mathbf{a}}) = \Big(\sum_{K\in\mathcal{T}_h}(f(\sum_{i=1}^{M}\theta_i(t)\phi_i, \sum_{i=1}^{M}a_i(t)\phi_i),\phi_j)_K\Big)_{1\le j\le M},$$

$$\bar{\alpha} = \Big((\alpha(t),\phi_j)_K\Big)_{1\le j\le M}. \qquad (4.3.9)$$

(4.3.5)-(4.3.8) is a system of ordinary differential equations in independent variable t, with Lipschitz continuous right hand side $(\bar{\theta},\bar{\mathbf{a}})$ and hence admits a unique solution in the neighbourhood of $t = 0$.

The *hp*-DGFEM scheme corresponding to the optimal control problem is

$$\min_{u\in U_{ad}} J(\theta_h, a_h, u_h) \quad \text{subject to the constraints (4.3.1)-(4.3.4).} \qquad (4.3.10)$$

The adjoint system of (4.3.10) determined from the KKT system (as developed in Chapter 3) is defined by:

Find $(z_h^*(t), \lambda_h^*(t)) \in S^p \times S^p, \quad$ a.e. in $\bar{I}$ such that

$$\begin{aligned}
-\sum_{K\in\mathcal{T}_h}(\chi,\partial_t\lambda_h^*)_K &= -\sum_{K\in\mathcal{T}_h}(\chi, f_a(\theta_h^*,a_h^*)g(z_h^*,\lambda_h^*))_K, && (4.3.11)\\
\lambda_h^*(T) &= \beta_1(a_h^*(T)-a_d), && (4.3.12)\\
-\rho c_p\sum_{K\in\mathcal{T}_h}(\phi,\partial_t z_h^*)_K + B(\phi,z_h^*) &= -\sum_{K\in\mathcal{T}_h}\Big(\phi,(f_\theta(\theta_h^*,a_h^*)g(z_h^*,\lambda_h^*))_K \\
&\quad + \beta_2(\phi,[\theta_h^*-\theta_m]_+)_K\Big), && (4.3.13)\\
z_h^*(T) &= 0, && (4.3.14)
\end{aligned}$$

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

$$\Big(\beta_3 u_h^* + \int_\Omega \alpha z_h^* dx,\ p-u_h^*\Big)_{L^2(I)} \ge 0 \quad \forall p\in U_{ad}.$$

Continuous Time *A Priori* Error Estimates

For estimating the *a priori* error estimates for the *hp*-DGFEM formulation of the laser surface hardening of steel problem, we would like to define the broken projector

$\Pi : H^2(\Omega, \mathcal{T}_h) \longrightarrow S^p$ satisfying;

$$B(\Pi v - v, w) + \nu(\Pi v - v, w) = 0 \;\; \forall w \in S^p, \tag{4.3.15}$$

where $\nu > 0$ is a constant. We now proceed to show that Π is well-defined.

Lemma 4.3.1. *There exists a constant $C > 0$, independent of h such that*

$$|B_\nu(v, w)| \leq C|||v|||\;|||w||| \quad \forall v, w \in H^2(\Omega, \mathcal{T}_h),$$

where $B_\nu(v, w) = B(v, w) + \nu(v, w) \;\; \forall v, w \in H^2(\Omega, \mathcal{T}_h)$.

Proof: For $v, w \in H^2(\Omega, \mathcal{T}_h)$, we have

$$\begin{aligned} |B_\nu(v, w)| \;\; &\leq \;\; \mathcal{K}\Big(|\sum_{K\in\mathcal{T}_h}(v, w)_{H^1(K)}| + |\sum_{e\in\mathcal{E}_{int}}(\{\triangledown v.n\}, [w])_e| + |\sum_{e\in\mathcal{E}_{int}}(\{\triangledown w.n\}, [v])_e|\Big) \\ &+ |\mathcal{J}^\gamma(v, w)| + \nu\sum_{K\in\mathcal{T}_h}|(v, w)_K| = I_1 + I_2 + I_3 + I_4 + I_5. \end{aligned}$$

We need to obtain bounds for I_1, I_2, I_3, I_4 and I_5. Clearly, we have

$$I_1 \;\; \leq \;\; \mathcal{K}\sum_{K\in\mathcal{T}_h}\|v\|_{H^1(K)}\|w\|_{H^1(K)} \leq \mathcal{K}|||v||||||w|||.$$

Now from Lemma 4.2.3 for $l = 1$ and using (4.2.1), we have
$\| \triangledown v.n\|^2_{e_K} \leq Cp^2_{e_K}|e_K|^{-1}\| \triangledown v\|^2_K$. Hence,

$$\begin{aligned} I_2 \;\; &\leq \;\; \mathcal{K}\sum_{e\in\mathcal{E}_{int}}|(\{\triangledown v.n\}, [w])_e| \leq \mathcal{K}\Big(\sum_{e\in\mathcal{E}_{int}}\|\sqrt{\frac{\gamma}{|e|}}[w]\|^2_e\Big)^{1/2}\Big(\sum_{K\in\mathcal{T}_h}\|\sqrt{\frac{|e|}{\gamma}}\{\triangledown v.n\}\|^2_e\Big)^{1/2} \\ &\leq \;\; C|||v|||\;|||w|||. \end{aligned}$$

Similarly $I_3 \leq C|||v||||||w|||$. Using definition of $||| \cdot |||$, we can easily obtain that $I_4 \leq C\mathcal{K}|||v||||||w|||$ and $I_5 \leq C\mathcal{K}|||v||||||w|||$. Using bounds for I_1, I_2, I_3, I_4 and I_5, we obtain the required result. □

Lemma 4.3.2. *For a sufficiently large penalty parameter γ, there exists $C > 0$ such that*

$$B_\nu(w, w) \geq C|||w|||^2 \;\; \forall w \in S^p.$$

Proof: For $w \in S^p$, we have

$$\begin{aligned}
B_\nu(w,w) &= \mathcal{K}\sum_{K\in\mathcal{T}_h} \| \bigtriangledown w\|_K^2 - 2\mathcal{K}\sum_{e\in\mathcal{E}_{int}} (\{\bigtriangledown w.n\},[w])_e + \sum_{e\in\mathcal{E}_{int}} \int_e \frac{\gamma}{|e|}[w]^2 de + (\nu+\mathcal{K})\sum_{K\in\mathcal{T}_h} \|w\|_K^2 \\
&= \mathcal{K}\sum_{K\in\mathcal{T}_h} \| \bigtriangledown w\|_K^2 + \sum_{e\in\mathcal{E}_{int}} \int_e \Big(\frac{\gamma}{|e|}[w]^2 - 2\mathcal{K}\{\bigtriangledown w.n\}\ [w]\Big) de + (\nu+\mathcal{K})\sum_{K\in\mathcal{T}_K} \|w\|_K^2. \qquad (4.3.16)
\end{aligned}$$

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

$$-2\mathcal{K}\sum_{e\in\mathcal{E}_{int}} \int_e |\{\nabla w.n\}|\ |[w]| de \ \geq \ -2\sum_{e\in\mathcal{E}_{int}} \Big(\mathcal{K}\int_e \delta |e|\ |\nabla w|^2 de + \mathcal{K}\int_e \frac{\delta^{-1}}{|e|}[w]^2 de\Big),$$

where $\delta > 0$ is positive constant which will be suitably chosen. Using Lemma 4.2.3, (4.2.1) and the fact that the summation over $e \in \mathcal{E}_{int}$ may count an element at the most 8 times, if we allow one hanging node at each interface, we have

$$-2\mathcal{K}\sum_{e\in\mathcal{E}_{int}} \int_e |\{\nabla w.n\}|\ |[w]| de \ \geq \ -\Big(16\mathcal{K}\sum_{K\in\mathcal{T}_h} \int_K \delta C p_e^2 |\nabla w|^2 dx + 2\mathcal{K}\sum_{e\in\mathcal{E}_{int}} \int_e \frac{\delta^{-1}}{|e|}[w]^2 de\Big).$$

Choose $\delta = \frac{1}{32}(Cp_e^2)^{-1}$ to obtain

$$-2\mathcal{K}\sum_{e\in\mathcal{E}_{int}} \int_e |\{\nabla w.n\}|\ |[w]| de \ \geq \ -\Big(\frac{\mathcal{K}}{2}\sum_{K\in\mathcal{T}_h} \int_K |\nabla w|^2 dx + 64\mathcal{K}\sum_{e\in\mathcal{E}_{int}} \int_e \frac{Cp_e^2}{|e|}[w]^2 de\Big).$$

Using the above expression in (4.3.16), we obtain

$$B_\nu(w,w) \ \geq \ \frac{\mathcal{K}}{2}\sum_{K\in\mathcal{T}_h} \| \bigtriangledown w\|_K^2 + \sum_{e\in\mathcal{E}_{int}} \int_e \frac{\gamma - 64\mathcal{K}Cp_e^2}{|e|}[w]^2 de + (\nu+\mathcal{K})\sum_{K\in\mathcal{T}_h} \|w\|_K^2.$$

Now choose $\gamma = 2\mathcal{K}\delta^{-1}$ to obtain the required result. □

Using Lemma 4.3.1 and 4.3.2, Πv is well defined for $v \in H^2(\Omega, \mathcal{T}_h)$. Now, we establish an estimate for $\|v - \Pi v\|$, the proof of which are in the similar lines as in [50].

Lemma 4.3.3. *Let Πv be the projection of $v \in H^2(\Omega, \mathcal{T}_h)$ onto S^p defined by* (4.3.15), *then the following error estimate holds true:*

$$\|v - \Pi v\|^2 \ \leq \ C\Big(\max_{K\in\mathcal{T}_h} \frac{h_K^2}{p_K}\Big) \sum_{K\in\mathcal{T}_h} \frac{h_K^{2s-2}}{p_K^{2s'-3}} \|v\|_{H^{s'}(K)}^2,$$

where $s = \min(p_K + 1, s'), s' \geq 2,\ p_K \geq 2.$

Proof: Choose $v - \Pi v = (v - \hat{v}) + (\hat{v} - \Pi v) = \zeta + \eta$, where $\hat{v} \in S^p$ is the interpolant of v defined in Remark 4.2.2. From Lemma 4.3.2 and (4.3.15), we have

$$|||\eta|||^2 \;\leq\; CB_\nu(\eta,\eta) \leq C\Big(|B_\nu(\eta+\zeta,\eta)| + |B_\nu(\zeta,\eta)|\Big) = C|B_\nu(\zeta,\eta)|.$$

From Lemma 4.3.1, we have

$$|||\eta|||^2 \;\leq\; C|||\zeta|||\ |||\eta|||.$$

Therefore, we obtain

$$\begin{aligned} |||\eta|||^2 \leq C|||\zeta|||^2 &= C\Big(\sum_{K\in\mathcal{T}_h} \|\zeta\|^2_{H^1(K)} + \sum_{e\in\mathcal{E}_{int}} \int_e \frac{\gamma}{|e|}[\zeta]^2 de\Big), \\ &= C\Big(\sum_{K\in\mathcal{T}_h} \|\zeta\|^2_{H^1(K)} + \sum_{e\in\mathcal{E}_{int}} \frac{\gamma}{|e|}\|\zeta\|^2_e\Big). \end{aligned} \tag{4.3.17}$$

From Lemma 4.2.1, we have

$$\|\zeta\|^2_{H^1(K)} \leq C\frac{h_K^{2s-2}}{p_K^{2s'-2}}\|v\|^2_{H^{s'}(K)}, \qquad \|\zeta\|^2_e \leq C\frac{h_K^{2s-1}}{p_K^{2s'-1}}\|v\|^2_{H^{s'}(K)}, s' \geq 2. \tag{4.3.18}$$

Substituting (4.3.18) in (4.3.17), choosing $\gamma = 2\mathcal{K}\delta^{-1}$ as in Lemma 4.3.2 and using the local bounded variation property (4.2.1), we obtain

$$|||\eta|||^2 \;\leq\; C\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big). \tag{4.3.19}$$

We have, $\displaystyle\sum_{K\in\mathcal{T}_h} \|v - \Pi v\|^2_{H^1(K)} \leq C\sum_{K\in\mathcal{T}_h}\Big(\|\eta\|^2_{H^1(K)} + \|\zeta\|^2_{H^1(K)}\Big).$
Using $\displaystyle\sum_{K\in\mathcal{T}_h} \|\eta\|^2_{H^1(K)} \leq |||\eta|||^2$, Lemma 4.2.1 and (4.3.19), we obtain

$$\sum_{K\in\mathcal{T}_h} \|v - \Pi v\|^2_{H^1(K)} \;\leq\; C\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big). \tag{4.3.20}$$

From the definition of broken norm, we have

$$|||v - \Pi v|||^2 \;=\; \sum_{K\in\mathcal{T}_h} \|v - \Pi v\|^2_{H^1(K)} + \sum_{e\in\mathcal{E}_{int}} \frac{\gamma}{|e|}\|[v - \Pi v]\|^2_e. \tag{4.3.21}$$

Using (4.3.18), (4.3.19) and (4.2.1) in (4.3.21), we obtain

$$|||v - \Pi v|||^2 \;\leq\; C\sum_{K\in\mathcal{T}_h}\left(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\right). \tag{4.3.22}$$

Now we proceed to estimate $\|v - \Pi v\|$.

$$\|v - \Pi v\| = \sup_{g\in L^2(\Omega), g\neq 0} \frac{(v - \Pi v, g)}{\|g\|}.$$

Let $w \in H^2(\Omega)$ be the solution of

$$-\mathcal{K}\Delta w + \nu w \;=\; g \quad \text{in } \Omega, \tag{4.3.23}$$

$$\frac{\partial w}{\partial n} \;=\; 0 \quad \text{on } \partial\Omega, \tag{4.3.24}$$

satisfying,

$$\|w\|_{H^2(\Omega)} \leq C\|g\|. \tag{4.3.25}$$

and where ν as defined in (4.3.15). Then a discontinuous weak formulation of (4.3.23)-(4.3.24) is given by: find $w \in \mathcal{U}$ such that

$$B_\nu(w, \phi) \;=\; (g, \phi) \quad \forall\phi \in H^2(\Omega, \mathcal{T}_h). \tag{4.3.26}$$

Also, $B_\nu(v - \Pi v, \hat{w}) = 0$, where $\hat{w}$ is the interpolant of w defined in Remark 4.2.2.

$$\text{Now, } (v - \Pi v, g) \;=\; B_\nu(v - \Pi v, w) = B_\nu(v - \Pi v, w - \hat{w}). \tag{4.3.27}$$

Therefore, from Lemma 4.3.1 and definition of $|||\cdot|||$, we obtain $(v - \Pi v, g) \leq C|||v -$

$\Pi v|||\ |||w - \hat{w}|||$. From (4.3.22), we obtain

$$(v - \Pi v, g) \;\le\; C\Big[\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big) \times \Big(\sum_{K\in\mathcal{T}_h}\|w-\hat{w}\|^2_{H^1(K)} + \mathcal{J}^{\gamma}(w-\hat{w}, w-\hat{w})\Big)\Big]^{1/2}. \tag{4.3.28}$$

Using Lemma 4.2.1 for $s' = 2$, we obtain

$$\sum_{K\in\mathcal{T}_h}\|w-\hat{w}\|^2_{H^1(K)} + \mathcal{J}^{\gamma}(w-\hat{w}, w-\hat{w}) \;\le\; C\sum_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\|w\|^2_{H^2(K)}. \tag{4.3.29}$$

Using (4.3.29) in (4.3.28), we obtain

$$\begin{aligned}(v - \Pi v, g) \;&\le\; C\Big[\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big) \times \Big(\sum_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\|w\|^2_{H^2(K)}\Big)\Big]^{1/2},\\ &\le\; C\Big[\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big) \times \Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\Big]^{1/2}\|w\|_{H^2(\Omega)}.\end{aligned} \tag{4.3.30}$$

Using (4.3.30) in (4.3.25), we obtain

$$(v - \Pi v, g) \;\le\; C\Big(\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big) \times \Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\Big)^{1/2}\|g\|.$$

Hence, we have the required result

$$\|v - \Pi v\| = \sup_{g\in L^2(\Omega:g\neq 0)}\frac{(v-\Pi v, g)}{\|g\|} \le C\Big(\Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\sum_{K\in\mathcal{T}_h}\Big(\frac{h_K^{2s-2}}{p_K^{2s'-3}}\|v\|^2_{H^{s'}(K)}\Big)\Big)^{1/2}.$$

This completes the proof. □

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 Remark 4.2.2. Using the triangle inequality, we have

$$\|\theta - \theta_h\| \;\le\; \|\eta^\theta\| + \|\zeta^\theta\|, \quad \|a - a_h\| \le \|\eta^a\| + \|\zeta^a\|.$$

In the next theorem, we develop an *a priori* error estimate for $\|\theta(t) - \theta_h(t)\|$ and

$\|a(t) - a_h(t)\|$, for a fixed $u \in U_{ad}$ and $t \in \bar{I}$.

Theorem 4.3.1. *Let $(\theta(t), a(t))$ and $(\theta_h(t), a_h(t))$ be the solutions of* (4.2.8)-(4.2.12) *and* (4.3.1)-(4.3.4), *respectively, for a fixed $u \in U_{ad}$. Then,*

$$\begin{aligned}
&\|\theta(t) - \theta_h(t)\|^2 + \|a(t) - a_h(t)\|^2 \\
&\quad \leq C\Big(\max_{K\in\mathcal{T}_h} \frac{h_K^2}{p_K}\Big) \sum_{K\in\mathcal{T}_h} \frac{h_K^{2s-2}}{p_K^{2s'-3}} \Big(\|\theta_0\|^2_{H^{s'}(K)} + \|\theta\|^2_{L^2(I,H^{s'}(K))} + \|\partial_t\theta\|^2_{L^2(I,H^{s'}(K))} \\
&\qquad + \|a\|^2_{L^2(I,H^{s'}(K))} + \|\partial_t a\|^2_{L^2(I,H^{s'}(K))} + \|\theta\|^2_{L^\infty(I,H^{s'}(K))} + \|a\|^2_{L^\infty(I,H^{s'}(K))}\Big), \ t \in \bar{I},
\end{aligned}$$

where $C > 0$ is independent of p_K, h_K and (θ, a), also $s = \min(p_K + 1, s')$ and $s', p_K \geq 2$.

Proof: A solution $(\theta, a) \in X \times Y$ of (4.2.13)-(4.2.16), under the regularity assumption that $\theta(t) \in \mathcal{U}, t \in \bar{I}$, satisfies the equation

$$\rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t\theta, v)_K + B_\nu(\theta, v) = -\rho L \sum_{K\in\mathcal{T}_h} (f(\theta, a), v)_K + \sum_{K\in\mathcal{T}_h} (\alpha u, v)_K + \nu \sum_{K\in\mathcal{T}_h} (\theta, v)_K (4.3.31)$$

Subtracting (4.3.3) from (4.3.31) and using $B_\nu(\zeta^\theta, v) = 0 \ \ \forall v \in S^p$ (see (4.3.15)), we obtain

$$\begin{aligned}
\rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t\eta^\theta, v)_K + B_\nu(\eta^\theta, v) &= -\rho L \sum_{K\in\mathcal{T}_h} (f(\theta, a) - f(\theta_h, a_h), v)_K - \rho c_p \sum_{K\in\mathcal{T}_h} \Big((\partial_t\zeta^\theta, v)_K \\
&\quad + \nu(\eta^\theta + \zeta^\theta, v)_K\Big).
\end{aligned}$$

Choose $v = \eta^\theta$, use Lemma 4.3.2 and integrate from 0 to t to obtain

$$\begin{aligned}
\frac{1}{2}\|\eta^\theta(t)\|^2 + \int_0^t |||\eta^\theta|||^2 ds &\leq C\Big(\|\eta^\theta(0)\|^2 + \sum_{K\in\mathcal{T}_h} \int_0^t |(f(\theta, a) - f(\theta_h, a_h), \eta^\theta)_K| ds \\
&\quad + \sum_{K\in\mathcal{T}_h} \int_0^t |(\partial_t\zeta^\theta, \eta^\theta)_K| ds + \sum_{K\in\mathcal{T}_h} \int_0^t |(\zeta^\theta, \eta^\theta)_K| ds \\
&\quad + \sum_{K\in\mathcal{T}_h} \int_0^t \|\eta^\theta\|_K^2 ds\Big) \\
&= C\|\eta^\theta(0)\|^2 + I_1 + I_2 + I_3 + \int_0^t \|\eta^\theta\|^2 ds, \text{ say.} \qquad (4.3.32)
\end{aligned}$$

Now we estimate I_1, I_2 and I_3 in the right hand side of (4.3.32). Using Cauchy-Schwarz inequality, Young's inequality and Proposition 2.2.1, we obtain

$$I_1 = \sum_{K\in\mathcal{T}_h}\int_0^t |(f(\theta,a)-f(\theta_h,a_h),\eta^\theta)_K|ds \le C\int_0^t \Big(\|\eta^\theta\|^2+\|\zeta^\theta\|^2+\|\eta^a\|^2+\|\zeta^a\|^2\Big)ds \quad (4.3.33)$$

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

$$\begin{aligned} I_2 &\le \sum_{K\in\mathcal{T}_h}\int_0^t |(\partial_t\zeta^\theta,\eta^\theta)_K|ds \le C\int_0^t \Big(\|\partial_t\zeta^\theta\|^2+\|\eta^\theta\|^2\Big)ds. && (4.3.34)\\ I_3 &\le \sum_{K\in\mathcal{T}_h}\int_0^t |(\zeta^\theta,\eta^\theta)_K|ds \le C\int_0^t \Big(\|\zeta^\theta\|^2+\|\eta^\theta\|^2\Big)ds. && (4.3.35)\end{aligned}$$

Using (4.3.33)-(4.3.35) in (4.3.32), we obtain

$$\begin{aligned} \frac{1}{2}\|\eta^\theta(t)\|^2+\int_0^t |||\eta^\theta|||^2 ds &\le C\Big(\|\eta^\theta(0)\|^2+\int_0^t\Big(\|\zeta^\theta\|^2+\|\zeta^a\|^2+\|\partial_t\zeta^\theta\|^2\Big)ds\\ &+\int_0^t\Big(\|\eta^\theta\|^2+\|\eta^a\|^2\Big)ds\Big).\\ \text{That is,}\qquad \frac{1}{2}\|\eta^\theta(t)\|^2 &\le C\Big(\|\eta^\theta(0)\|^2+\int_0^t\Big(\|\zeta^\theta\|^2+\|\zeta^a\|^2+\|\partial_t\zeta^\theta\|^2\Big)ds\\ &+\int_0^t\Big(\|\eta^\theta\|^2+\|\eta^a\|^2\Big)ds\Big). && (4.3.36)\end{aligned}$$

Now subtracting (4.3.1) from (4.2.13), we obtain

$$\sum_{K\in\mathcal{T}_h}(\partial_t(a-a_h),w)_K = \sum_{K\in\mathcal{T}_h}(f(\theta,a)-f(\theta_h,a_h),w)_K \quad \forall w\in S^p.$$

Using $a-a_h=\eta^a+\zeta^a$ and substituting $w=\eta^a$, we obtain

$$\sum_{K\in\mathcal{T}_h}(\partial_t\eta^a,\eta^a)_K = \sum_{K\in\mathcal{T}_h}\Big((f(\theta,a)-f(\theta_h,a_h),\eta^a)_K-(\partial_t\zeta^a,\eta^a)_K\Big).$$

Now integrating from 0 to t, using Cauchy-Schwarz inequality, Young's inequality and Proposition 2.2.1, we obtain

$$\|\eta^a(t)\|^2 \le C\int_0^t\Big(\|\eta^\theta\|^2+\|\zeta^\theta\|^2+\|\eta^a\|^2+\|\zeta^a\|^2+\|\partial_t\zeta^a\|^2\Big)ds. \quad (4.3.37)$$

Adding (4.3.36) and (4.3.37), we obtain

$$\begin{aligned}\|\eta^\theta(t)\|^2 + \|\eta^a(t)\|^2 \;&\leq\; C\Big(\|\eta^\theta(0)\|^2 + \int_0^t \Big(\|\zeta^\theta\|^2 + \|\zeta^a\|^2 + \|\partial_t\zeta^\theta\|^2 + \|\partial_t\zeta^a\|^2\Big)ds \\ &\quad + \int_0^t \Big(\|\eta^\theta\|^2 + \|\eta^a\|^2\Big)ds\Big).\end{aligned}$$

Use Gronwall's Lemma to obtain

$$\|\eta^\theta(t)\|^2 + \|\eta^a(t)\|^2 \;\leq\; C\Big(\|\eta^\theta(0)\|^2 + \int_0^T \Big(\|\zeta^\theta\|^2 + \|\zeta^a\|^2 + \|\partial_t\zeta^\theta\|^2 + \|\partial_t\zeta^a\|^2\Big)ds\Big).$$

From Lemma 4.2.1 and Lemma 4.3.3, we have

$$\begin{aligned}\|\eta^\theta(t)\|^2 + \|\eta^a(t)\|^2 \;&\leq\; C\Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\sum_{K\in\mathcal{T}_h}\frac{h_K^{2s-2}}{p_K^{2s'_k-3}}\Big(\|\theta_0\|^2_{H^{s'}(K)} + \|\theta\|^2_{L^2(I,H^{s'}(K))} \\ &\quad + \|\partial_t\theta\|^2_{L^2(I,H^{s'}(K))} + \|a\|^2_{L^2(I,H^{s'}(K))} + \|\partial_t a\|^2_{L^2(I,H^{s'}(K))}\Big).\end{aligned}$$

Using triangle inequality we obtain the required estimate. This completes the proof. □

Next we develop error estimates for the system (4.2.18)-(4.2.21), which is the adjoint system corresponding to (4.2.13)-(4.2.16). Denote $z - z_h = \eta^z + \zeta^z$ and $\lambda - \lambda_h = \eta^\lambda + \zeta^\lambda$, where $\eta^z = \Pi z - z_h,\ \zeta^z = z - \Pi z,\ \eta^\lambda = \hat{\lambda} - \lambda_h,\ \zeta^\lambda = \lambda - \hat{\lambda}$ and $\hat{\lambda}$ is the interpolant of λ as defined in Remark 4.2.2. We denote (z_h^*, λ_h^*) as (z_h, λ_h) for notational convenience.

Theorem 4.3.2. *Let* $(z(t), \lambda(t))$ *and* $(z_h(t), \lambda_h(t))$ *be the solutions for* (4.2.18)-(4.2.21) *and* (4.3.11)-(4.3.14)*, respectively. Then, there exists a positive constant* C *such that*

$$\begin{aligned}&\|z(t) - z_h(t)\|^2 + \|\lambda(t) - \lambda_h(t)\|^2 \\ &\quad\leq\; C\Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\sum_{K\in\mathcal{T}_h}\frac{h_K^{2s-2}}{p_K^{2s'-3}}\Big(\|a_d\|^2_{H^2(K)} + \|\theta_0\|^2_{H^{s'}(K)} + \|\theta\|^2_{L^2(I,H^{s'}(K))} \\ &\quad\quad + \|\partial_t\theta\|^2_{L^2(I,H^{s'}(K))} + \|a\|^2_{L^2(I,H^{s'}(K))} + \|\partial_t a\|^2_{L^2(I,H^{s'}(K))} + \|z\|^2_{L^2(I,H^{s'}(K))} \\ &\quad\quad + \|\partial_t z\|^2_{L^2(I,H^{s'}(K))} + \|\lambda\|^2_{L^2(I,H^{s'}(K))} + \|\partial_t\lambda\|^2_{L^2(I,H^{s'}(K))} + \|\theta\|^2_{L^\infty(I,H^{s'}(K))} \\ &\quad\quad + \|z\|^2_{L^\infty(I,H^{s'}(K))} + \|a\|^2_{L^\infty(I,H^{s'}(K))} + \|\lambda\|^2_{L^\infty(I,H^{s'}(K))}\Big),\quad t\in\bar{I},\end{aligned}$$

where C *is independent of* p_K, h_K*,* (θ, a) *and* (z, λ)*, also* $s = \min(p_K + 1, s')$ *and* $s', p_K \geq 2$.

Proof: A solution $(z, \lambda) \in X \times Y$ of (4.2.18)-(4.2.21), under the regularity assumption that $z(t) \in \mathcal{U},\ t \in \bar{I}$, satisfies the equation

$$-\rho c_p \sum_{K\in\mathcal{T}_h} (\phi, \partial_t z)_K + B(\phi, z) = -\sum_{K\in\mathcal{T}_h} (f_\theta(\theta, a)g(z,\lambda), \phi)_K + \ \beta_2 \sum_{K\in\mathcal{T}_h} ([\theta - \theta_m]_+, \phi)_K. \quad (4.3.38)$$

Subtracting (4.3.13) from (4.3.38) and using $B_\nu(\zeta^z, \phi) = 0 \ \ \forall \phi \in S^p$, we obtain

$$\begin{aligned} -\rho c_p \sum_{K\in\mathcal{T}_h} (\phi, \partial_t \eta^z)_K + B_\nu(\phi, \eta^z) \ &= \ -\sum_{K\in\mathcal{T}_h} (\phi, f_\theta(\theta, a)g(z,\lambda) - f_\theta(\theta_h, a_h)g(z_h, \lambda_h))_K \\ &\quad + \beta_2 \sum_{K\in\mathcal{T}_h} (\phi, [\theta - \theta_m]_+ - [\theta_h - \theta_m]_+)_K \\ &\quad + \rho c_p \sum_{K\in\mathcal{T}_h} (\phi, \partial_t \zeta^z)_K + \nu(\phi, \eta^z + \zeta^z). \end{aligned}$$

Choose $\phi = \eta^z$, use Lemma 4.3.2, $\eta^z(T) = 0$ and integrate from t to T to obtain

$$\begin{aligned} \frac{1}{2}\|\eta^z(t)\|^2 + \int_t^T \||\eta^z\||^2 ds \ &\leq \ C\Bigg|\Bigg(-\sum_{K\in\mathcal{T}_h} \int_t^T (\eta^z, f_\theta(\theta, a)g(z,\lambda) - f_\theta(\theta_h, a_h)g(z_h, \lambda_h))_K ds \\ &\quad + \sum_{K\in\mathcal{T}_h} \int_t^T (\eta^z, [\theta - \theta_m]_+ - [\theta_h - \theta_m]_+) ds + \sum_{K\in\mathcal{T}_h} \int_t^T (\eta^z, \partial_t \zeta^z)_K ds \\ &\quad + \sum_{K\in\mathcal{T}_h} \int_t^T (\eta^z, \zeta^z)_K ds + \sum_{K\in\mathcal{T}_h} \int_t^T \|\eta^z\|^2 ds \Bigg)\Bigg| \\ &\leq \ C\Bigg(I_1 + I_2 + I_3 + I_4 + \int_t^T \|\eta^z\|^2 ds \Bigg), \text{ say.} \end{aligned} \quad (4.3.39)$$

Using Cauchy-Schwarz inequality, Young's inequality, Proposition 2.2.1 and Lipschitz continuity of g, we obtain

$$\begin{aligned} I_1 \ &= \ \sum_{K\in\mathcal{T}_h} \int_t^T |(f_\theta(\theta, a)g(z,\lambda) - f_\theta(\theta_h, a_h)g(z_h, \lambda_h), \eta^z)_K| ds \\ &\leq \ C \int_t^T \Big(\|\eta^z\|^2 + \|\zeta^z\|^2 + \|\eta^\lambda\|^2 + \|\zeta^\lambda\|^2 \Big) ds. \end{aligned} \quad (4.3.40)$$

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

$$I_2 \ \leq \ \sum_{K\in\mathcal{T}_h} \int_t^T |(\eta^z, \partial_t \zeta^z)_K| ds \leq C \sum_{K\in\mathcal{T}_h} \int_t^T \Big(\|\partial_t \zeta^z\|_K^2 + \|\eta^z\|_K^2 \Big) ds \quad (4.3.41)$$

$$\begin{aligned} I_3 &\leq \sum_{K\in\mathcal{T}_h}\int_t^T |(\eta^z, [\theta-\theta_m]_+ - [\theta_h-\theta_m]_+)_K|ds \\ &\leq C\sum_{K\in\mathcal{T}_h}\int_t^T \Big(\|\theta-\theta_h\|_K^2 + \|\eta^z\|_K^2\Big)ds \qquad (4.3.42) \\ I_4 &\leq \sum_{K\in\mathcal{T}_h}\int_t^T |(\eta^z,\zeta^z)_K|ds \leq C\sum_{K\in\mathcal{T}_h}\int_t^T \Big(\|\zeta^z\|_K^2 + \|\eta^z\|_K^2\Big)ds. \qquad (4.3.43) \end{aligned}$$

Using (4.3.40)-(4.3.43) in (4.3.39), we obtain

$$\begin{aligned} \frac{1}{2}\|\eta^z(t)\|^2 + \int_t^T |||\eta^z|||^2 ds &\leq C\sum_{K\in\mathcal{T}_h}\bigg(\int_t^T \Big(\|\zeta^z\|_K^2 + \|\zeta^\lambda\|_K^2 + \|\partial_t\zeta^z\|_K^2 + \|\theta-\theta_h\|_K^2\Big)ds \\ &\quad + \int_t^T \Big(\|\eta^z\|_K^2 + \|\eta^\lambda\|_K^2\Big)ds\bigg). \end{aligned}$$

That is,

$$\begin{aligned} \|\eta^z(t)\|^2 &\leq C\sum_{K\in\mathcal{T}_h}\bigg(\int_t^T \Big(\|\zeta^z\|_K^2 + \|\zeta^\lambda\|_K^2 + \|\partial_t\zeta^z\|_K^2 + \|\theta-\theta_h\|_K^2\Big)ds \\ &\quad + \int_t^T \Big(\|\eta^z\|_K^2 + \|\eta^\lambda\|_K^2\Big)ds\bigg). \qquad (4.3.44) \end{aligned}$$

Now subtracting (4.3.11) from (4.2.18), we obtain

$$-\sum_{K\in\mathcal{T}_h}(\chi, \partial_t(\lambda-\lambda_h))_K = -\sum_{K\in\mathcal{T}_h}(\chi, f_a(\theta,a)g(z,\lambda) - f_a(\theta_h,a_h)g(z_h,\lambda_h))_K.$$

Using $\lambda-\lambda_h = \eta^\lambda + \zeta^\lambda$ and substituting $\chi = \eta^\lambda$, we obtain

$$-\sum_{K\in\mathcal{T}_h}(\eta^\lambda, \partial_t\eta^\lambda)_K = \sum_{K\in\mathcal{T}_h}\Big(-(\eta^\lambda, f_a(\theta,a)g(z,\lambda) - f_a(\theta_h,a_h)g(z_h,\lambda_h))_K + (\eta^\lambda, \partial_t\zeta^\lambda)_K\Big).$$

Now integrating from t to T, using $\lambda(T) = \beta_1(a(T)-a_d)$, Cauchy-Schwarz inequality, Young's inequality and Proposition 2.2.1, we obtain

$$\begin{aligned} \sum_{K\in\mathcal{T}_h}\|\eta^\lambda(t)\|_K^2 &\leq C\sum_{K\in\mathcal{T}_h}\bigg(\|\hat{a}(T)-a(T)\|_K^2 + \|a(T)-a_h(T)\|_K^2 + \|\hat{a}_d - a_d\|_K^2 \\ &\quad + \int_t^T \Big(\|\eta^z\|_K^2 + \|\zeta^z\|_K^2 + \|\eta^\lambda\|_K^2 + \|\zeta^\lambda\|_K^2 + \|\partial_t\zeta^\lambda\|_K^2\Big)ds\bigg)(4.3.45) \end{aligned}$$

Adding (4.3.44) and (4.3.45), we obtain

$$\sum_{K\in\mathcal{T}_h}\Big(\|\eta^z(t)\|_K^2 + \|\eta^\lambda(t)\|_K^2\Big) \leq C\sum_{K\in\mathcal{T}_h}\Big(\|\hat{a}(T)-a(T)\|_K^2 + \|a(T)-a_h(T)\|_K^2 + \|\hat{a}_d - a_d\|_K^2$$

$$+ \int_t^T \Big(\|\theta - \theta_h\|_K^2 + \|\zeta^z\|_K^2 + \|\zeta^\lambda\|_K^2 + \|\partial_t \zeta^z\|_K^2 + \|\partial_t \zeta^\lambda\|_K^2 \Big) ds + \int_t^T \Big(\|\eta^z\|_K^2 + \|\eta^\lambda\|_K^2 \Big) ds \Big).$$

Use Gronwall's Lemma to obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \Big(\|\eta^z(t)\|_K^2 + \|\eta^\lambda(t)\|_K^2 \Big) &\le C \sum_{K\in\mathcal{T}_h} \Big(\|\hat{a}(T) - a(T)\|_K^2 + \|a(T) - a_h(T)\|_K^2 + \|\hat{a}_d - a_d\|_K^2 \\ &\quad + \int_0^T \Big(\|\theta - \theta_h\|_K^2 + \|\zeta^z\|_K^2 + \|\zeta^\lambda\|_K^2 + \|\partial_t \zeta^z\|_K^2 + \|\partial_t \zeta^\lambda\|_K^2 \Big) ds \Big).\end{aligned}$$

From Lemma 4.2.1, Lemma 4.3.3 and Theorem 4.3.1, we have

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \Big(\|\eta^z(t)\|_K^2 &+ \|\eta^\lambda(t)\|_K^2 \Big) \\ &\le C \Big(\max_{K\in\mathcal{T}_h} \frac{h_K^2}{p_K} \Big) \sum_{K\in\mathcal{T}_h} \frac{h_K^{2s-2}}{p_K^{2s'-3}} \Big(\|a_d\|_{H^2(K)}^2 + \|\theta_0\|_{H^{s'}(K)}^2 + \|\theta\|_{L^2(I,H^{s'}(K))}^2 \\ &\quad + \|\partial_t \theta\|_{L^2(I,H^{s'}(K))}^2 + \|a\|_{L^2(I,H^{s'}(K))}^2 + \|\partial_t a\|_{L^2(I,H^{s'}(K))}^2 + \|z\|_{L^2(I,H^{s'}(K))}^2 \\ &\quad + \|\partial_t z\|_{L^2(I,H^{s'}(K))}^2 + \|\lambda\|_{L^2(I,H^{s'}(K))}^2 + \|\partial_t \lambda\|_{L^2(I,H^{s'}(K))}^2 + \|\theta\|_{L^\infty(I,H^{s'}(K))}^2 \\ &\quad + \|a\|_{L^\infty(I,H^{s'}(K))}^2 \Big).\end{aligned}$$

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

4.4 *hp*-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 Theorem 4.4.1 and 4.4.2. The control is then discretized using piecewise constants in each discrete interval $I_n, n = 1, 2, \cdots, N$ and show the convergence of u_σ^* to u^* in $L^2(I)$ in Theorem 4.4.3. In order to discretize (4.3.1)-(4.3.4) in time, we consider the following partition of I:

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

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

$$X_{hk}^q = \{\phi : I \rightarrow S^p;\ \phi|_{I_n} = \sum_{j=0}^{q} \psi_j t^j, \psi_j \in S^p\},\ \ q \in \mathbb{N}. \tag{4.4.1}$$

For $q = 0$, the space time *hp*-DGFEM scheme corresponding to (4.3.1)-(4.3.4) reads as;
Find $(\theta_{hk}^n, a_{hk}^n) \in S^p \times S^p, n = 1, 2, \cdots, N$ such that

$$\begin{aligned}
\sum_{K \in \mathcal{T}_h} (\bar{\partial} a_{hk}^n, w)_K &= \sum_{K \in \mathcal{T}_h} (f(\theta_{hk}^n, a_{hk}^n), w)_K, && (4.4.2)\\
a_{hk}(0) &= 0, && (4.4.3)\\
\rho c_p \sum_{K \in \mathcal{T}_h} (\bar{\partial} \theta_{hk}^n, v)_K + B(\theta_{hk}^n, v) &= -\rho L \sum_{K \in \mathcal{T}_h} (f(\theta_{hk}^n, a_{hk}^n), v)_K \\
&\quad + \sum_{K \in \mathcal{T}_h} \left(\frac{1}{k_n} \int_{I_n} \alpha u_{hk}(t) ds, v \right)_K, && (4.4.4)\\
\theta_{hk}(0) &= \theta_0, && (4.4.5)
\end{aligned}$$

$\forall (w, v) \in S^p \times S^p$, where $\bar{\partial}\phi^n = \dfrac{\phi^n - \phi^{n-1}}{k_n} \forall \phi \in S^p$. Expanding $a_{hk}^n = \sum_{i=1}^{M} a_i^n \phi_i$ and $\theta_{hk}^n = \sum_{i=1}^{M} \theta_i^n \phi_i$, where $\{\phi_i\}_{i=1}^N$ is the basis for S^p, we obtain the system

$$\begin{aligned}
\mathbf{A}\ \bar{\mathbf{a}}^{\mathbf{n}} &= k_n \bar{\mathbf{F}}(\bar{\theta}^{\mathbf{n}}, \bar{\mathbf{a}}^{\mathbf{n}}) + \mathbf{A}\ \bar{\mathbf{a}}^{\mathbf{n-1}}, && (4.4.6)\\
\bar{\mathbf{a}}^0 &= 0, && (4.4.7)\\
(\rho c_p \mathbf{A} + k_n \mathbf{B})\ \bar{\theta}^{\mathbf{n}} &= -k_n \rho L \bar{\mathbf{F}}(\bar{\theta}^{\mathbf{n}}, \bar{\mathbf{a}}^{\mathbf{n}}) + k_n u_{hk}(t_n) \bar{\alpha}^{\mathbf{n}} + \mathbf{A}\ \bar{\theta}^{\mathbf{n-1}}, && (4.4.8)\\
\bar{\theta}(0) &= \theta_0, && (4.4.9)
\end{aligned}$$

where $\mathbf{A}, \bar{\mathbf{a}}, \bar{\mathbf{F}}, \mathbf{B}, \bar{\theta}$ and $\bar{\alpha}$ are defined in (4.3.9). (4.4.6)-(4.4.9) forms a system of non-linear equations with Lipschitz continuous right hand side and hence admits a unique local solution in the neighbourhood of $t = 0$.

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

$$\min_{u_{hk} \in U_{ad}} J(\theta_{hk}, a_{hk}, u_{hk}) \quad \text{subject to the constraints (4.4.2)-(4.4.5).} \tag{4.4.10}$$

The adjoint system of (4.4.10) determined by the KKT system is defined by:

find $(z_{hk}^{n,*}, \lambda_{hk}^{n,*}) \in S^p \times S^p$ such that

$$-\sum_{K\in\mathcal{T}_h} (\chi, \tilde{\partial}\lambda_{hk}^{n-1,*})_K \;=\; -\sum_{K\in\mathcal{T}_h} (\chi, f_a(\theta_{hk}^{n-1,*}, a_{hk}^{n-1,*})g(z_{hk}^{n-1,*}, \lambda_{hk}^{n-1,*}))_K, \qquad (4.4.11)$$

$$\begin{aligned}
\lambda_{hk}^{*}(T) &= \beta_1(a_{hk}^{*}(T) - a_d), && (4.4.12)\\
-\rho c_p \sum_{K\in\mathcal{T}_h} (\phi, \tilde{\partial} z_{hk}^{n-1,*})_K + B(\phi, z_{hk}^{n-1,*}) &= -\sum_{K\in\mathcal{T}_h} \Big((\phi, f_\theta(\theta_{hk}^{n-1,*}, a_{hk}^{n-1,*})g(z_{hk}^{n-1,*}, \lambda_{hk}^{n-1,*}))_K \\
&\quad + \beta_2(\phi, [\theta_{hk}^{n-1,*} - \theta_m]_+)_K\Big), && (4.4.13)\\
z_{hk}^{*}(T) &= 0, && (4.4.14)
\end{aligned}$$

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

Discrete Time *A Priori* Error Estimates

Before estimating the *a priori* error estimates for space-time discretization, we define the interpolant $\pi_k : C(\bar{I}, S^p) \longrightarrow X_{hk}^1$, $\pi_k v|_{I_n} \in \mathcal{P}_0(I_n, S^p)$, (see [79]) as:

$$\pi_k v(t) = v(t_n) \qquad \forall t \in \bar{I}_n, n = 1, 2, \cdots, N, \qquad (4.4.15)$$

where $\mathcal{P}_0(I_n, S^p)$ 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

$$\|v - \pi_k v\|_{I_n,K} \le C k_n \|\partial_t v\|_{I_n,K}. \qquad (4.4.16)$$

Theorem 4.4.1. *Let* $(\theta(t), a(t))$ *and* $(\theta_{hk}^n, a_{hk}^n)$, $n = 1, 2, \cdots, N$ *be the solutions of* (4.2.8)-(4.2.12) *and* (4.4.2)-(4.4.5), *respectively. Then,*

$$\begin{aligned}
\|\theta(t_n) \;-\; \theta_{hk}^n\|^2 &+ \|a(t_n) - a_{hk}^n\|^2 \\
&\le C \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h} \left(\left(\max_{K\in\mathcal{T}_h} \frac{h_K^2}{p_K}\right) \frac{h_K^{2s-2}}{p_K^{2s'-3}} + k_n^2\right) \Big(\|\theta_0\|_{H^{s'}(K)}^2 + \|\theta\|_{L^\infty(I_n;H^{s'}(K))}^2 + \|\partial_t\theta\|_{L^\infty(I_n;H^{s'}(K))}^2 \\
&\quad + \|a\|_{L^\infty(I_n;H^{s'}(K))}^2 + \|\partial_t a\|_{L^\infty(I_n;H^{s'}(K))}^2 + \|\partial_{tt}\theta\|_{L^\infty(I;L^2(K))}^2 + \|\partial_{tt} a\|_{L^\infty(I_n,L^2(K))}^2 \\
&\quad + \|\partial_t u\|_{L^2(I_n)}^2\Big), \forall t \in \bar{I}_n.
\end{aligned}$$

where $C > 0$ is independent of p_K, h_K and (θ, a), also $s = \min(p_K + 1, s')$ and $s', p_K \geq 2$.

Proof: Subtracting (4.4.4) from (4.3.31) at $t = t_n$, we obtain

$$\begin{aligned}
\rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t\theta(t_n) - \bar{\partial}\theta^n_{hk}, v)_K \quad &+ \quad B_\nu(\theta(t_n) - \theta^n_{hk}, v) \\
&= \quad -\rho L \sum_{K\in\mathcal{T}_h} \Big(f(\theta(t_n), a(t_n)) - f(\theta^n_{hk}, a^n_{hk})ds, v\Big)_K \\
&\quad\quad + \sum_{K\in\mathcal{T}_h} \Big(\alpha(x, t_n)u(t_n) - \frac{1}{k_n}\int_{I_n} \alpha u ds, v\Big)_K \\
&\quad\quad + \nu \sum_{K\in\mathcal{T}_h} (\theta(t_n) - \theta^n_{hk}), v)_K,
\end{aligned}$$

where $v \in S^p$. Using (4.4.15), we find that

$$\begin{aligned}
&\rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t\theta(t_n) - \bar{\partial}\theta^n_{hk}, v)_K \quad + \quad B_\nu(\theta(t_n) - \theta^n_{hk}, v) \\
&\quad\quad \leq - \sum_{K\in\mathcal{T}_h} \rho L\Big(f(\theta(t_n), a(t_n)) - f(\theta^n_{hk}, a^n_{hk}), v\Big)_K + \sum_{K\in\mathcal{T}_h} \max_{K\times I_n} \frac{1}{k_n}|\alpha|\Big(\int_{I_n}(\pi_k u(t_n) - u)ds, v\Big)_K \\
&\quad\quad + \nu \sum_{K\in\mathcal{T}_h} (\theta(t_n) - \theta^n_{hk}), v)_K.
\end{aligned}$$

Write $\theta(t_n) - \theta^n_{hk} = (\theta(t_n) - \Pi\theta(t_n)) + (\Pi\theta(t_n) - \theta^n_{hk}) = \zeta^{\theta,n} + \eta^{\theta,n}$, $a(t_n) - a^n_{hk} = (a(t_n) - \hat{a}(t_n)) + (\hat{a}(t_n) - a^n_{hk}) = \zeta^{a,n} + \eta^{a,n}$ and using $B_\nu(\zeta, v) = 0$, we obtain

$$\begin{aligned}
&\rho c_p \sum_{K\in\mathcal{T}_h} (\bar{\partial}\eta^{\theta,n}, v)_K \quad + \quad B_\nu(\eta^{\theta,n}, v) \\
&\leq -\rho L \sum_{K\in\mathcal{T}_h} (f(\theta(t_n), a(t_n)) - f(\theta^n_{hk}, a^n_{hk}), v)_K + \frac{\max\limits_{\Omega\times I_n}|\alpha|}{k_n} \sum_{K\in\mathcal{T}_h} (\int_{I_n}(\pi_k u - u)ds, v)_K \\
&- \rho c_p \sum_{K\in\mathcal{T}_h} (\partial_t\theta(t_n) - \bar{\partial}\theta(t_n), v)_K - \rho c_p \sum_{K\in\mathcal{T}_h} (\bar{\partial}\zeta^{\theta,n}, v)_K + \nu \sum_{K\in\mathcal{T}_h} (\eta^{\theta,n} + \zeta^{\theta,n}, v)_K. \quad (4.4.17)
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{1}{2k_n}\Big(\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2\Big) \quad &= \quad \frac{1}{2k_n}\Big((\eta^{\theta,n}, \eta^{\theta,n}) - (\eta^{\theta,n-1}, \eta^{\theta,n-1})\Big) \\
&= \quad \frac{1}{2k_n}\Big((\eta^{\theta,n} - \eta^{\theta,n-1}, \eta^{\theta,n}) - (\eta^{\theta,n-1}, \eta^{\theta,n-1} - \eta^{\theta,n})\Big)
\end{aligned}$$

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

$$\begin{aligned}\frac{1}{2k_n}\left(\|\eta^{\theta,n}\|^2-\|\eta^{\theta,n-1}\|^2\right) &= (\bar\partial\eta^{\theta,n},\eta^{\theta,n})-\frac{1}{2k_n}(\eta^{\theta,n}-\eta^{\theta,n-1},\eta^{\theta,n}-\eta^{\theta,n-1})\\ &\le (\bar\partial\eta^{\theta,n},\eta^{\theta,n}). \qquad (4.4.18)\end{aligned}$$

Substituting $v=\eta^{\theta,n}$ and using (4.4.18) in (4.4.17) , we obtain

$$\begin{aligned}\|\eta^{\theta,n}\|^2 &- \|\eta^{\theta,n-1}\|^2\\ &\le C\Bigg(\sum_{K\in\mathcal{T}_h}|(f(\theta(t_n),a(t_n))-f(\theta^n_{hk},a^n_{hk}),\eta^{\theta,n})_K|+\sum_{K\in\mathcal{T}_h}|(\int_{I_n}(\pi_k u-u)ds,\eta^{\theta,n})_K|\\ &\quad+\sum_{K\in\mathcal{T}_h}|(\partial_t\theta(t_n)-\bar\partial\theta(t_n),\eta^{\theta,n})_K|+\sum_{K\in\mathcal{T}_h}|(\bar\partial\zeta^{\theta,n},\eta^{\theta,n})_K|+\sum_{K\in\mathcal{T}_h}|(\zeta^{\theta,n},\eta^{\theta,n})_K|\\ &\quad+\sum_{K\in\mathcal{T}_h}\|\eta^{\theta,n}\|^2_K\Bigg)\\ &= C\Bigg(I_1+I_2+I_3+I_4+I_5+\|\eta^{\theta,n}\|^2\Bigg),\quad \text{say} \qquad (4.4.19)\end{aligned}$$

From Cauchy-Schwarz inequality, Young's inequality and Proposition 2.2.1, we have

$$\begin{aligned}I_1 &\le \rho L\sum_{K\in\mathcal{T}_h}\|f(\theta(t_n),a(t_n))-f(\theta^n_{hk},a^n_{hk})\|_K\|\eta^{\theta,n}\|_K\\ &\le C\sum_{K\in\mathcal{T}_h}\Bigg(\|\zeta^{\theta,n}\|^2_K+\|\zeta^{a,n}\|^2_K+\|\eta^{a,n}\|^2_K+\|\eta^{\theta,n}\|^2_K\Bigg). \qquad (4.4.20)\end{aligned}$$

For I_2, use Cauchy-Schwarz inequality, Young's inequality and (4.4.16) to obtain

$$\begin{aligned}I_2\le\sum_{K\in\mathcal{T}_h}\|\pi_k u-u\|_{L^2(I_n)}\|\eta^{\theta,n}\|_K &\le C\Bigg(\|\pi_k u-u\|^2_{L^2(I_n)}+\sum_{K\in\mathcal{T}_h}\|\eta^{\theta,n}\|^2_K\Bigg),\\ &\le C\Bigg(k_n^2\|\partial_t u\|^2_{L^2(I_n)}+\sum_{K\in\mathcal{T}_h}\|\eta^{\theta,n}\|^2_K\Bigg). \qquad (4.4.21)\end{aligned}$$

Now consider I_3. Using Cauchy-Schwarz inequality and Young's inequality, we obtain

$$I_3 \le C\sum_{K\in\mathcal{T}_h}\Bigg(\|\bar\partial\theta(t_n)-\partial_t\theta(t_n)\|^2_K+\|\eta^{\theta,n}\|^2_K\Bigg). \qquad (4.4.22)$$

For the first term on the right hand side of (4.4.22), we have

$$\begin{aligned}\|\bar{\partial}\theta(t_n) - \partial_t\theta(t_n)\|_K &= \|k_n^{-1}\int_{t_{n-1}}^{t_n}(t-t_{n-1})\partial_{tt}\theta\ dt\|_K \leq k_n^{-1}\int_{t_{n-1}}^{t_n}(t-t_{n-1})\|\partial_{tt}\theta\|_K\ dt \\ &\leq Ck_n^{-1}\frac{(t-t_{n-1})^2}{2}\Big|_{t_{n-1}}^{t_n}\|\partial_{tt}\theta\|_{L^\infty(I_n,L^2(K))} \leq Ck_n\|\partial_{tt}\theta\|_{L^\infty(I_n,L^2(K))}.\end{aligned}$$

Therefore, we have

$$I_3 \leq C\sum_{K\in\mathcal{T}_h}\Big(k_n^2\|\partial_{tt}\theta\|^2_{L^\infty(I_n,L^2(K))} + \|\eta^{\theta,n}\|_K^2\Big). \tag{4.4.23}$$

Also for I_4, using Cauchy-Schwarz inequality and Young's inequality, we have

$$I_4 \leq C\sum_{K\in\mathcal{T}_h}\Big(\|\bar{\partial}\zeta^{\theta,n}\|_K^2 + \|\eta^{\theta,n}\|_K^2\Big). \tag{4.4.24}$$

$$\text{Also,}\quad \|\bar{\partial}\zeta^{\theta,n}\|_K = \|k_n^{-1}\int_{t_{n-1}}^{t_n}\partial_t\zeta^\theta\ dt\|_K \leq k_n^{-1}\int_{t_{n-1}}^{t_n}\|\partial_t\zeta^\theta\|_K dt \leq \|\partial_t\zeta^{\theta,n}\|_{L^\infty(I_n;L^2(K))}.$$

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

$$I_5 = \sum_{K\in\mathcal{T}_h}|(\zeta^{\theta,n},\eta^{\theta,n})_K| \leq C\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{\theta,n}\|_K^2 + \|\zeta^{\theta,n}\|_K^2\Big). \tag{4.4.25}$$

Using (4.4.20)-(4.4.25) in (4.4.19), we have

$$\begin{aligned}\|\eta^{\theta,n}\|^2 - \|\eta^{\theta,n-1}\|^2 \leq\ & C\sum_{K\in\mathcal{T}_h}\Big(\|\zeta^{\theta,n}\|_K^2 + \|\zeta^{a,n}\|_K^2 + k_n^2\|\partial_{tt}\theta\|^2_{L^\infty(I_n,L^2(K))} \\ & + \|\partial_t\zeta^\theta\|^2_{L^\infty(I_n,L^2(K))} + \|\eta^{\theta,n}\|_K^2 + \|\eta^{a,n}\|_K^2 \\ & + k_n^2\|\partial_t u\|^2_{L^2(I_n)}\Big).\end{aligned} \tag{4.4.26}$$

Subtracting (4.2.8) from (4.4.2), we obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}(\bar{\partial}\eta^{a,n},w)_K =\ & \sum_{K\in\mathcal{T}_h}(f(\theta(t_n),a(t_n)) - f(\theta^n_{hk},a^n_{hk}),w)_K - \sum_{K\in\mathcal{T}_h}(\bar{\partial}a(t_n) - \partial_t a(t_n),w)_K \\ & - \sum_{K\in\mathcal{T}_h}(\bar{\partial}\zeta^{a,n},w)_K,\end{aligned}$$

where $w \in S^p$. Substituting $w = \eta^{a,n}$, proceeding as in (4.4.17)-(4.4.18) and using Proposition 2.2.1, we obtain

$$\begin{aligned}\|\eta^{a,n}\|^2 &- \|\eta^{a,n-1}\|^2 \\ &\leq C\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{\theta,n}\|_K^2+\|\eta^{a,n}\|_K^2+\|\zeta^{\theta,n}\|_K^2+\|\zeta^{a,n}\|_K^2+\|\bar{\partial}a(t_n)-\partial_t a(t_n)\|_K^2+\|\partial_t\zeta^{a,n}\|_K^2\Big) \\ &\leq C\sum_{K\in\mathcal{T}_h}\Big(\|\zeta^{\theta,n}\|_K^2+\|\zeta^{a,n}\|_K^2+k_n^2\|\partial_{tt}a\|^2_{L^\infty(I_n,L^2(K))}+\|\partial_t\zeta^a\|^2_{L^\infty(I_n,L^2(K))}+\|\eta^{\theta,n}\|_K^2 \\ &\quad+\|\eta^{a,n}\|_K^2\Big). \qquad (4.4.27)\end{aligned}$$

Adding (4.4.26) and (4.4.27), we obtain

$$\begin{aligned}\|\eta^{a,n}\|^2+\|\eta^{\theta,n}\|^2 &- \|\eta^{a,n-1}\|^2-\|\eta^{\theta,n-1}\|^2 \\ &\leq C\sum_{K\in\mathcal{T}_h}\Big(\|\zeta^{\theta,n}\|_K^2+\|\zeta^{a,n}\|_K^2+k_n^2\|\partial_{tt}\theta\|^2_{L^\infty(I_n,L^2(K))}+k_n^2\|\partial_{tt}a\|^2_{L^\infty(I_n,L^2(K))} \\ &\quad+\|\partial_t\zeta^\theta\|^2_{L^\infty(I_n,L^2(K))}+\|\partial_t\zeta^a\|^2_{L^\infty(I_n,L^2(K))}+\|\eta^{\theta,n}\|_K^2+\|\eta^{a,n}\|_K^2 \\ &\quad+k_n^2\|\partial_t u\|^2_{L^2(I_n)}\Big).\end{aligned}$$

Summing from 1 to n and using the fact that $\theta(0)=\theta_0$ and $a(0)=0$, we obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{a,n}\|_K^2+\|\eta^{\theta,n}\|_K^2\Big) &\leq C\Big(\sum_{K\in\mathcal{T}_h}\|\eta^{\theta,0}\|_K^2+\sum_{m=1}^n\sum_{K\in\mathcal{T}_h}\Big(\|\zeta^{\theta,m}\|_K^2+\|\zeta^{a,m}\|_K^2 \\ &\quad+\|\partial_t\zeta^\theta\|^2_{L^\infty(I_m,L^2(K))}+\|\partial_t\zeta^a\|^2_{L^\infty(I_m,L^2(K))} \\ &\quad+k_m^2\|\partial_{tt}\theta\|^2_{L^\infty(I_m,L^2(K))}+k_n^2\|\partial_{tt}a\|^2_{L^\infty(I_m,L^2(K))}+k_m^2\|\partial_t u\|^2_{L^2(I_m)}\Big) \\ &\quad+\sum_{m=1}^n\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{\theta,m}\|_K^2+\|\eta^{a,m}\|_K^2\Big)\Big)\end{aligned}$$

Now using Gronwall's Lemma, Lemma 4.2.1 and Lemma 4.3.3, we obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{a,n}\|_K^2 &+ \|\eta^{\theta,n}\|_K^2\Big) \\ &\leq C\sum_{n=1}^N\sum_{K\in\mathcal{T}_h}\Big(\Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\frac{h_K^{2s-2}}{p_K^{2s'-3}}+k_n^2\Big)\Big(\|\theta_0\|^2_{H^{s'}(K)}+\|\theta\|^2_{L^\infty(I_n;H^{s'}(K))} \\ &\quad+\|\partial_t\theta\|^2_{L^\infty(I_n,H^{s'}(K))}+\|a\|^2_{L^\infty(I_n;H^{s'}(K))}+\|\partial_t a\|^2_{L^\infty(I_n,H^{s'}(K))}+\|\partial_{tt}\theta\|^2_{L^\infty(I_n,L^2(K))} \\ &\quad+\|\partial_{tt}a\|^2_{L^\infty(I_n,L^2(K))}+\|\partial_t u\|^2_{L^2(I_n)}\Big).\end{aligned}$$

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

Next we show the discrete time error for the adjoint equation (4.2.18)-(4.2.21).

Theorem 4.4.2. *let* $(z(t),\lambda(t))$ *and* $(z_{hk}^n,\lambda_{hk}^n)$, $n=1,2,\cdots,N$ *be the solutions for* (4.2.18)-(4.2.21) *and* (4.4.11)-(4.4.14)*, respectively. Then,*

$$\begin{aligned}\|z(t_{n-1}) &- z_{hk}^{n-1}\|^2+\|\lambda(t_{n-1})-\lambda_{hk}^{n-1}\|^2\\ &\leq C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\bigg(\Big(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\Big)\frac{h_K^{2s-2}}{p_K^{2s'-3}}+k_n^2\bigg)\bigg(\|\theta_0\|_{H^{s'}(K)}^2+\|a_d\|_{H^{s'}(K)}^2+\|\theta\|_{L^\infty(I_n,H^{s'}(K))}^2\\ &+\|\partial_t\theta\|_{L^\infty(I_n,H^{s'}(K))}^2+\|a\|_{L^\infty(I_n,H^{s'}(K))}^2+\|\partial_t a\|_{L^\infty(I_n,H^{s'}(K))}^2+\|\partial_{tt}\theta\|_{L^\infty(I_n,L^2(K))}^2\\ &+\|\partial_{tt}a\|_{L^\infty(I_n,L^2(K))}^2+\|z\|_{L^\infty(I_n,H^{s'}(K))}^2+\|\partial_t z\|_{L^\infty(I_n,H^{s'}(K))}^2+\|\lambda\|_{L^\infty(I_n,H^{s'}(K))}^2\\ &+\|\partial_t\lambda\|_{L^\infty(I_n,H^{s'}(K))}^2+\|\partial_{tt}z\|_{L^\infty(I_n,L^2(K))}^2+\|\partial_{tt}\lambda\|_{L^\infty(I_n,L^2(K))}^2+\|\partial_t u\|_{L^2(I_n)}^2\bigg),\\ &\forall t\in\bar{I}_n,\end{aligned}$$

where $C>0$ *is independent of* p_K,h_K, (θ,a) *and* (z,λ)*, also* $s=\min(p_K+1,s')$ *and* $s',p_K\geq 2$.

Proof: Subtracting (4.4.13) from (4.3.38) at $t=t_{n-1}$ and splitting $z(t_{n-1})-z_{hk}^{n-1}=\eta^{z,n-1}+\zeta^{z,n-1}$, we obtain

$$\begin{aligned}&-\rho c_p\sum_{K\in\mathcal{T}_h}(\phi,\tilde{\partial}\eta^{z,n-1})_K+B_\nu(\phi,\eta^{z,n-1})\\ &=\ -\sum_{K\in\mathcal{T}_h}(\phi,f_\theta(\theta(t_{n-1}),a(t_{n-1}))g(z(t_{n-1}),\lambda(t_{n-1}))-f_\theta(\theta_{hk}^{n-1},a_{hk}^{n-1})g(z_{hk}^{n-1},\lambda_{hk}^{n-1}))_K\\ &+\beta_2\sum_{K\in\mathcal{T}_h}(\phi,[\theta(t_{n-1})-\theta_m]_+-[\theta_{hk}^{n-1}-\theta_m]_+)_K+\rho c_p\sum_{K\in\mathcal{T}_h}(\phi,\partial_t z(t_{n-1})-\tilde{\partial}z(t_{n-1}))_K\\ &+\rho c_p\sum_{K\in\mathcal{T}_h}\Big((\phi,\tilde{\partial}\zeta^{z,n-1})_K+\nu(\phi,z(t_{n-1})-z_{hk}^{n-1})_K\Big),\end{aligned}$$

Substitute $\phi=\eta^{z,n-1}$ and Lemma 4.3.2 to obtain

$$\begin{aligned}&-\ \tilde{\partial}\|\eta^{z,n-1}\|^2\\ &\leq\ C\bigg(\sum_{K\in\mathcal{T}_h}|(\eta^{z,n-1},f_\theta(\theta(t_{n-1}),a(t_{n-1}))g(z(t_{n-1}),\lambda(t_{n-1}))-f_\theta(\theta_{hk}^{n-1},a_{hk}^{n-1})g(z_{hk}^{n-1},\lambda_{hk}^{n-1}))_K|\\ &+\beta_2\sum_{K\in\mathcal{T}_h}|(\eta^{z,n-1},[\theta(t_{n-1})-\theta_m]_+-[\theta_{hk}^{n-1}-\theta_m]_+)_K|+\sum_{K\in\mathcal{T}_h}|(\eta^{z,n-1},\partial_t z(t_{n-1})-\tilde{\partial}z(t_{n-1}))_K|\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{K\in\mathcal{T}_h} |(\eta^{z,n-1}, \tilde{\partial}\zeta^{z,n-1})_K| + |\Big(\sum_{K\in\mathcal{T}_h} (\eta^{z,n-1}, \zeta^{z,n-1})_K + \sum_{K\in\mathcal{T}_h} \|\eta^{z,n-1}\|_K^2 \Big)| \Big), \\
&= I_1 + I_2 + I_3 + I_4 + I_5 + \|\eta^{z,n-1}\|^2, \text{ say.} \qquad (4.4.28)
\end{aligned}
$$

From Cauchy-Schwarz inequality, Young's inequality, Proposition 2.2.1 and Lipschitz continuity of $g(\cdot,\cdot)$, we obtain

$$
\begin{aligned}
I_1 &\leq \sum_{K\in\mathcal{T}_h} \|f_\theta(\theta(t_{n-1}), a(t_{n-1}))g(z(t_{n-1}), \lambda(t_{n-1})) - f_\theta(\theta_{hk}^{n-1}, a_{hk}^{n-1})g(z_{hk}^{n-1}, \lambda_{hk}^{n-1})\|_K \|\eta^{z,n-1}\|_K, \\
&\leq C \sum_{K\in\mathcal{T}_h} \Big(\|\zeta^{z,n-1}\|_K^2 + \|\zeta^{\lambda,n-1}|_K + \|\eta^{\lambda,n-1}\|_K^2 + \|\eta^{z,n-1}\|_K^2 \Big). \qquad (4.4.29)
\end{aligned}
$$

For I_2 and I_3, use Cauchy Schwarz and Young's inequality to obtain

$$
\begin{aligned}
I_2 &\leq \sum_{K\in\mathcal{T}_h} \|[\theta(t_{n-1}) - \theta_m]_+ - [\theta_{hk}^{n-1} - \theta_m]_+\|_K \|\eta^{z,n-1}\|_K \\
&\leq C \sum_{K\in\mathcal{T}_h} \Big(\|\theta(t_{n-1}) - \theta_{hk}^{n-1}\|_K^2 + \|\eta^{z,n-1}\|_K^2 \Big). \qquad (4.4.30) \\
I_3 &\leq C \sum_{K\in\mathcal{T}_h} \Big(\|\tilde{\partial} z(t_{n-1}) - \partial_t z(t_{n-1})\|_K^2 + \|\eta^{z,n-1}\|_K^2 \Big). \qquad (4.4.31)
\end{aligned}
$$

For the first term on the right hand side of (4.4.31), we have

$$
\begin{aligned}
\sum_{K\in\mathcal{T}_h} \|\tilde{\partial} z(t_{n-1}) &- \partial_t z(t_{n-1})\|_K \\
&= \sum_{K\in\mathcal{T}_h} \|k_n^{-1} \int_{t_{n-1}}^{t_n} (t_n - t)\partial_{tt} z \, dt\|_K \leq k_n^{-1} \int_{t_{n-1}}^{t_n} (t_n - t) \sum_{K\in\mathcal{T}_h} \|\partial_{tt} z\|_K \, dt \\
&\leq Ck_n^{-1} \frac{(t_n - t)^2}{2} \Big|_{t_{n-1}}^{t_n} \sum_{K\in\mathcal{T}_h} \|\partial_{tt} z\|_{L^\infty(I_n, L^2(K))} \\
&\leq Ck_n \sum_{K\in\mathcal{T}_h} \|\partial_{tt} z\|_{L^\infty(I_n, L^2(K))}.
\end{aligned}
$$

Therefore, we have

$$
I_3 \leq C \sum_{K\in\mathcal{T}_h} \Big(k_n^2 \|\partial_{tt} z\|_{L^\infty(I_n, L^2(K))}^2 + \|\eta^{z,n-1}\|_K^2 \Big). \qquad (4.4.32)
$$

Also for I_4 and I_5, using Cauchy-Schwarz inequality and Young's inequality, we have

$$
\begin{aligned}
I_4 &\leq C \sum_{K \in \mathcal{T}_h} \Big(\|\tilde{\partial} \zeta^{z,n-1}\|_K^2 + \|\eta^{z,n-1}\|_K^2 \Big). && (4.4.33) \\
I_5 &= \sum_{K \in \mathcal{T}_h} (\eta^{z,n-1}, \zeta^{z,n-1})_K \leq C \sum_{K \in \mathcal{T}_h} \Big(\|\eta^{z,n-1}\|_K^2 + \|\zeta^{z,n-1}\|_K^2 \Big). && (4.4.34)
\end{aligned}
$$

Also,

$$
\sum_{K \in \mathcal{T}_h} \|\tilde{\partial} \zeta^{z,n-1}\|_K = \sum_{K \in \mathcal{T}_h} \|k_n^{-1} \int_{t_{n-1}}^{t_n} \partial_t \zeta^z dt\|_K \leq k_n^{-1} \sum_{K \in \mathcal{T}_h} \int_{t_{n-1}}^{t_n} \|\partial_t \zeta^z\|_{L^\infty(I_n, L^2(K))} dt.
$$

Using (4.4.29)-(4.4.34) in (4.4.28), we have

$$
\begin{aligned}
-\tilde{\partial} \|\eta^{z,n-1}\|^2 \leq\ & C \sum_{K \in \mathcal{T}_h} \Big(\|\zeta^{z,n-1}\|_K^2 + \|\zeta^{\lambda,n-1}\|_K^2 + \|\theta(t_{n-1}) - \theta_{hk}^{n-1}\|_K^2 + k_n^2 \|\partial_{tt} z\|_{L^\infty(I_n, L^2(K))}^2 \\
& + \|\partial_t \zeta^z\|_{L^\infty(I_n, L^2(K))}^2 + \|\eta^{z,n-1}\|_K^2 + \|\eta^{\lambda,n-1}\|_K^2 \Big). \qquad (4.4.35)
\end{aligned}
$$

Subtracting (4.4.11) from (4.2.18), using $\lambda(t_{n-1}) - \lambda_{hk}^{n-1} = \eta^{\lambda,n-1} + \zeta^{\lambda,n-1}$ and substituting $\chi = \eta^{\lambda,n-1}$, proceeding as in (4.4.29)-(4.4.34), we obtain

$$
\begin{aligned}
-\tilde{\partial} \|\eta^{\lambda,n-1}\|^2 \leq\ & C \sum_{K \in \mathcal{T}_h} \Big(\|\zeta^{z,n-1}\|_K^2 + \|\zeta^{\lambda,n}\|_K^2 + k_n^2 \|\partial_{tt} \lambda\|_{L^\infty(I_n, L^2(K))}^2 \\
& + \|\partial_t \zeta^\lambda\|_{L^\infty(I_n, L^2(K))}^2 + \|\eta^{z,n-1}\|_K^2 + \|\eta^{\lambda,n-1}\|_K^2 \Big). \qquad (4.4.36)
\end{aligned}
$$

Adding (4.4.35) and (4.4.36), we obtain

$$
\begin{aligned}
-\tilde{\partial} \|\eta^{\lambda,n-1}\|^2 -\ & \tilde{\partial} \|\eta^{z,n-1}\|^2 \\
\leq\ & C \sum_{K \in \mathcal{T}_h} \Big(\|\zeta^{z,n-1}\|_K^2 + \|\zeta^{\lambda,n-1}\|_K^2 + \|\theta(t_{n-1}) - \theta_{hk}^{n-1}\|_K^2 + k_n^2 \|\partial_{tt} z\|_{L^\infty(I_n, L^2(K))}^2 \\
& + k_n^2 \|\partial_{tt} \lambda\|_{L^\infty(I_n, L^2(K))}^2 + \|\partial_t \zeta^z\|_{L^\infty(I_n, L^2(K))}^2 + \|\partial_t \zeta^\lambda\|_{L^\infty(I_n, L^2(K))}^2 \\
& + \|\eta^{z,n-1}\|_K^2 + \|\eta^{\lambda,n-1}\|_K^2 \Big).
\end{aligned}
$$

Summing from n to $N+1$, using the fact that $\lambda(T) = \beta_2(a(T) - a_d), z(T) = 0$ and $\lambda_{hk}(T) = \beta_2(a_{hk}(T) - a_d), z_{hk}(T) = 0$, we obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{\lambda,n-1}\|_K^2 &+ \|\eta^{z,n-1}\|_K^2\Big)\\ &\le C\bigg(\|a_{hk}^N-\hat{a}(t_N)\|_K^2+\|a_d(t_N)-\hat{a}_d(t_N)\|_K^2+\sum_{m=n}^N\bigg(k_n\sum_{K\in\mathcal{T}_h}\bigg(\|\zeta^{z,m}\|_K^2\\ &+\|\zeta^{\lambda,m}\|_K^2+\|\partial_t\zeta^z\|^2_{L^\infty(I_m,L^2(K))}+\|\partial_t\zeta^\lambda\|^2_{L^\infty(I_m,L^2(K))}\\ &+k_m^2\|\theta(t_m)-\theta_{hk}^m\|_K+k_m^2\|\partial_{tt}z\|^2_{L^\infty(I_m,L^2(K))}+k_m^2\|\partial_{tt}\lambda\|^2_{L^\infty(I_m,L^2(K))}\bigg)\bigg)\\ &+\sum_{m=n}^N\sum_{K\in\mathcal{T}_h}\Big(\|\eta^{z,m}\|_K^2+\|\eta^{\lambda,m}\|_K^2\Big)\bigg).\end{aligned}$$

Now using Gronwall's Lemma, Lemma 4.2.1, Lemma 4.3.3, Theorem 4.4.1 and triangle inequality to split $z(t_{n-1})-z_{hk}^{n-1}$ and $\lambda(t_{n-1})-\lambda_{hk}^{n-1}$, we obtain the required the result. □

Complete Discretization

Now, we will discretize the control by using a DGFEM. In order to completely discretize the problem (4.2.17) we choose a discontinuous Galerkin piecewise constant approximation of the control variable. Let U_d be the finite dimensional subspace of U defined by

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

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: find $(\theta_\sigma,a_\sigma)\in X_{hk}^q\times X_{hk}^q$ such that

$$\begin{aligned}\sum_{n=1}^N\Big((\partial_t a_\sigma,w)_{\Omega,I_n}+(<a_\sigma>_{n-1},w_{n-1}^+)\Big) &= \sum_{n=1}^N(f(\theta_\sigma,a_\sigma),w)_{\Omega,I_n} &(4.4.37)\\ a_\sigma(0) &= 0 &(4.4.38)\\ \sum_{n=1}^N\Big(\rho c_p(\partial_t\theta_\sigma,v)_{\Omega,I_n}+\int_{I_n}B(\theta_\sigma,v)dt+\rho c_p(<\theta_\sigma>_{n-1},v_{n-1}^+)\Big) &= \sum_{n=1}^N\Big(-\rho L(f(\theta_\sigma,a_\sigma),v)_{\Omega,I_n}\\ &\quad+(\alpha u_\sigma,v)_{\Omega,I_n}\Big) &(4.4.39)\\ \theta_\sigma(0) &= \theta_0. &(4.4.40)\end{aligned}$$

for all $(w,v)\in X_{hk}^q\times X_{hk}^q$. Next we show a stability estimate for θ_σ and a_σ.

Lemma 4.4.1. *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* (4.4.37)-(4.4.40)*, satisfies the following a priori bounds:*

$$\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \left(\|\partial_t \theta_\sigma\|_{K,I_n}^2 + \|\Delta_h \theta_\sigma\|_{K,I_n}^2 \right) \leq C, \quad \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \|\partial_t a_\sigma\|_{K,I_n}^2 \leq C, \tag{4.4.41}$$

where $\Delta_h : S^p \times S^p$ is the discrete Laplacian defined by

$$- \sum_{K \in \mathcal{T}_h} (\Delta_h v, w)_K = B(v,w), \quad \forall v, w \in S^p. \tag{4.4.42}$$

Proof: Using (4.4.42) in (4.4.39), we have

$$\begin{aligned} \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(\rho c_p (\partial_t \theta_\sigma, v)_{K,I_n} - (\Delta_h \theta_\sigma, v)_{K,I_n} + \rho c_p (< \theta_\sigma >_{n-1}, v_{n-1}^{+})_K \Big) \\ = \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(- \rho L (f(\theta_\sigma, a_\sigma), v)_{K,I_n} + (\alpha u_\sigma, v)_{K,I_n} \Big). \end{aligned} \tag{4.4.43}$$

Put $v = -\Delta_h \theta_\sigma$ in (4.4.43) to obtain

$$\begin{aligned} \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(\rho c_p (\partial_t \theta_\sigma, -\Delta_h \theta_\sigma)_{K,I_n} - (\Delta_h \theta_\sigma, -\Delta_h \theta_\sigma)_{K,I_n} + \rho c_p (< \theta_\sigma >_{n-1}, -\Delta_h \theta_{\sigma,n-1}^{+})_K \Big) \\ = \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(- \rho L (f(\theta_\sigma, a_\sigma), -\Delta_h \theta_\sigma)_{K,I_n} + (\alpha u_\sigma, -\Delta_h \theta_\sigma)_{K,I_n} \Big). \end{aligned} \tag{4.4.44}$$

Again using (4.4.42) in first and third terms on the left hand side of (4.4.44), we obtain

$$\begin{aligned} \sum_{n=1}^{N} \Big(\rho c_p \int_{I_n} B(\partial_t \theta_\sigma, \theta_\sigma) dt + \sum_{K \in \mathcal{T}_h} \|\Delta_h \theta_\sigma\|_{K,I_n}^2 + \rho c_p B(< \theta_\sigma >_{n-1}, \theta_{\sigma,n-1}^{+}) \Big) \\ = \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(- \rho L (f(\theta_\sigma, a_\sigma), -\Delta_h \theta_\sigma)_{K,I_n} + (\alpha u_\sigma, -\Delta_h \theta_\sigma)_{K,I_n} \Big). \end{aligned} \tag{4.4.45}$$

Now we find estimates for the terms in (4.4.45) one by one. Consider

$$\begin{aligned}\int_{I_n} B(\partial_t\theta_\sigma,\theta_\sigma)dt &= \mathcal{K}\sum_{K\in\mathcal{T}_h}\int_{I_n}(\partial_t\bigtriangledown\theta_\sigma,\bigtriangledown\theta_\sigma)_K dt-\mathcal{K}\sum_{e\in\mathcal{E}_{int}}\int_{I_n}(\{\bigtriangledown(\partial_t\theta_\sigma).\mathbf{n}\},[\theta_\sigma])_e dt\\ &- \mathcal{K}\sum_{e\in\mathcal{E}_{int}}\int_{I_n}(\{\bigtriangledown\theta_\sigma.\mathbf{n}\},[\partial_t\theta_\sigma])_e dt+\sum_{e\in\mathcal{E}_{int}}\frac{\gamma}{|e|}\int_{I_n}([\partial_t\theta_\sigma],[\theta_\sigma])_e dt,\\ &= \sum_{K\in\mathcal{T}_h}\mathcal{K}I_1-\sum_{e\in\mathcal{E}_{int}}\mathcal{K}\Big(I_2+I_3\Big)-\sum_{e\in\mathcal{E}_{int}}I_4,\qquad \text{say.}\end{aligned}\qquad(4.4.46)$$

For I_1, we have

$$I_1=\int_{I_n}(\partial_t\bigtriangledown\theta_\sigma,\bigtriangledown\theta_\sigma)_K dt=\int_{I_n}\frac{1}{2}\frac{d}{dt}\|\bigtriangledown\theta_\sigma\|_K^2 dt=\frac{1}{2}\Big(\|\bigtriangledown\theta_{\sigma,n}\|_K^2-\|\bigtriangledown\theta_{\sigma,n-1}^+\|_K^2\Big)\qquad(4.4.47)$$

Using by parts intergration for I_2, we have

$$\begin{aligned}I_2=\int_{I_n}(\partial_t\{\bigtriangledown\theta_\sigma.\mathbf{n}\},[\theta_\sigma])_e dt &= (\{\bigtriangledown\theta_\sigma.\mathbf{n}\},[\theta_\sigma])_e\Big|_{I_n}-\int_{I_n}(\{\bigtriangledown\theta_\sigma.\mathbf{n}\},[\partial_t\theta_\sigma])_e dt\\ &= (\{\bigtriangledown\theta_\sigma.\mathbf{n}\},[\theta_\sigma])_e\Big|_{I_n}-I_3.\end{aligned}\qquad(4.4.48)$$

For I_4, we have

$$I_4=\int_{I_n}(\partial_t[\theta_\sigma],[\theta_\sigma])_e dt = \int_{I_n}\frac{1}{2}\frac{d}{dt}\|[\theta_\sigma]\|_e^2 dt=\frac{1}{2}\|[\theta_\sigma]\|_e^2\Big|_{I_n}.\qquad(4.4.49)$$

Using (4.4.47)-(4.4.49) in (4.4.46), we obtain

$$\begin{aligned}\int_{I_n} B(\partial_t\theta_\sigma,\theta_\sigma)dt &= \frac{1}{2}\sum_{K\in\mathcal{T}_h}\mathcal{K}\Big(\|\bigtriangledown\theta_{\sigma,n}\|_K^2-\|\bigtriangledown\theta_{\sigma,n-1}^+\|_K^2\Big)-\sum_{e\in\mathcal{E}_{int}}\mathcal{K}(\{\bigtriangledown\theta_\sigma\},[\theta_\sigma])_e\Big|_{I_n}\\ &+\frac{1}{2}\sum_{e\in\mathcal{E}_{int}}\frac{\gamma}{|e|}\|[\theta_\sigma]\|_e^2\Big|_{I_n}.\end{aligned}\qquad(4.4.50)$$

Using the definition of $B(\cdot,\cdot)$ in the third term on the left hand side of the (4.4.45), we obtain

$$\begin{aligned}B(<\theta_\sigma>_{n-1},\theta_{\sigma,n-1}^+) &= \sum_{K\in\mathcal{T}_h}\mathcal{K}(<\bigtriangledown\theta>_{n-1},\bigtriangledown\theta_{\sigma,n-1}^+)_K-\sum_{e\in\mathcal{E}_{int}}\Big(\mathcal{K}(\{\bigtriangledown<\theta_\sigma>_{n-1}.\mathbf{n}\},[\theta_{\sigma,n-1}^+])_e\\ &+\mathcal{K}(\{\bigtriangledown\theta_{\sigma,n-1}^+.\mathbf{n}\},[<\theta_\sigma>_{n-1}])_e-\frac{\gamma}{|e|}([<\theta_\sigma>_{n-1}],[\theta_{\sigma,n-1}^+])_e\Big).\end{aligned}\qquad(4.4.51)$$

Using

$$(<\bigtriangledown\theta_\sigma>_{n-1},\bigtriangledown\theta^+_{\sigma,n-1})_K=\frac{1}{2}\Big(\|\bigtriangledown\theta^+_{\sigma,n-1}\|^2_K+\|<\bigtriangledown\theta_\sigma>_{n-1}\|^2_K-\|\bigtriangledown\theta_{\sigma,n-1}\|^2_K\Big),\tag{4.4.52}$$

in (4.4.51), we have

$$\begin{aligned}B(<\theta_\sigma>_{n-1},\theta^+_{\sigma,n-1}) &= \sum_{K\in\mathcal{T}_h}\frac{\mathcal{K}}{2}\Big(\|\bigtriangledown\theta^+_{\sigma,n-1}\|^2_K+\|<\bigtriangledown\theta_\sigma>_{n-1}\|^2_K-\|\bigtriangledown\theta_{\sigma,n-1}\|^2_K\Big)\\ &- \sum_{e\in\mathcal{E}_{int}}\Big(\mathcal{K}(\{\bigtriangledown<\theta_\sigma>_{n-1}.\mathbf{n}\},[\theta^+_{\sigma,n-1}])_e+\mathcal{K}(\{\bigtriangledown\theta^+_{\sigma,n-1}.\mathbf{n}\},[<\theta_\sigma>_{n-1}])_e\\ &\quad -\frac{\gamma}{|e|}([<\theta_\sigma>_{n-1}],[\theta^+_{\sigma,n-1}])_e\Big).\end{aligned}\tag{4.4.53}$$

Using (4.4.50), (4.4.53) in (4.4.45), Cauchy-Schwarz and Young's inequalities, we have

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\Big(\|\bigtriangledown\theta_{\sigma,N}\|^2_K-\|\bigtriangledown\theta_0\|^2_K\Big) &+ \sum_{n=1}\sum_{K\in\mathcal{T}_h}\|\Delta_h\theta_\sigma\|^2_{K,I_n}\\ &\leq C\sum_{n=1}^{N}\Big(\sum_{K\in\mathcal{T}_h}\Big(\|f(\theta_\sigma,a_\sigma)\|^2_{K,I_n}+\|\alpha u_\sigma\|^2_{K,I_n}+\|\Delta_h\theta_\sigma\|^2_{K,I_n}\\ &\quad+\sum_{e\in\mathcal{E}_{int}}\Big(\|\bigtriangledown\theta_{\sigma,n}.\mathbf{n}\|^2_e+\|\bigtriangledown\theta^+_{\sigma,n-1}.\mathbf{n}\|^2_e+\|\theta_{\sigma,n}\|^2_e+\|\theta_{\sigma,n-1}\|^2_e\\ &\quad+\|\theta^+_{\sigma,n-1}\|^2_e\Big)\Big).\end{aligned}$$

Choosing Young's constant appropriately, using Remark 1.2 and $\theta_\sigma\in L^2(I,H^1(\Omega,\mathcal{T}_h))$, we obtain

$$\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|\Delta_h\theta_\sigma\|^2_{K,I_n}\text{ is bounded.}\tag{4.4.54}$$

Put $v=(t-t_{n-1})\partial_t\theta_\sigma$ in (4.4.39), use $((t-t_{n-1})\partial_t\theta_\sigma)^+_{n-1}=0$ and (4.4.42) to obtain

$$\begin{aligned}\rho c_p\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}(\partial_t\theta_\sigma,(t-t_{n-1})\partial_t\theta_\sigma)_{K,I_n} &- \sum_{n=1}^{N}\int_{I_n}(\Delta_h\theta_\sigma,(t-t_{n-1})\partial_t\theta_\sigma)_{K,I_n}\\ &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(-\rho L(f(\theta_\sigma,a_\sigma),(t-t_{n-1})\partial_t\theta_\sigma)_{K,I_n}\\ &\quad+(\alpha u_\sigma,(t-t_{n-1})\partial_t\theta_\sigma)_{K,I_n}\Big).\end{aligned}\tag{4.4.55}$$

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

$$\begin{aligned}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(t-t_{n-1})\|\partial_t\theta_\sigma\|_K^2dt \;\; &\leq \;\; C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Bigg(\|f(\theta_\sigma,a_\sigma)\|_{K,I_n}^2+\|\alpha u_\sigma\|_{K,I_n}^2+\|\Delta_h\theta_\sigma\|_{K,I_n}^2\\ &\qquad +\int_{I_n}(t-t_{n-1})\|\partial_t\theta_\sigma\|_K^2dt\Bigg)\end{aligned}$$

Choosing Young's constant appropriately, using (4.4.54) and Proposition 2.2.1, we obtain

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

From inverse estimate, we have

$$\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\|\partial_t\theta_\sigma\|_K^2dt\leq C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}k_n^{-1}\int_{I_n}(t-t_{n-1})\|\partial_t\theta_\sigma\|_K^2dt$$

Therefore,

$$\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\left(\|\partial_t\theta_\sigma\|_{K,I_n}^2+\|\Delta_h\theta_\sigma\|_{K,I_n}^2\right)\leq C.$$

Similarly putting $w=(t-t_{n-1})\partial_t a_\sigma$ in (4.4.37) and using inverse estimate, we obtain

$$\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|\partial_t a_\sigma\|_{K,I_n}^2\leq C. \quad \square$$

The discrete time DGFEM scheme for the optimal control problem is

$$\min_{u_\sigma\in U_{d,ad}} J(\theta_\sigma,a_\sigma,u_\sigma) \quad \text{subject to the constraints (4.4.37)-(4.4.40),} \tag{4.4.56}$$

where $(\theta_\sigma(t),a_\sigma(t),u_\sigma(t))=(\theta_\sigma^n,a_\sigma^n,u_\sigma^n),\;\;$ for $t\in I_n$.

Theorem 4.4.3. *Let u_σ^* be the optimal control of* (4.4.56). *Then,* $\lim_{\sigma\to 0}u_\sigma^*=u^*$ *exists in* $L^2(I)$ *and u^* is an optimal control of* (4.2.17).

The proof of this theorem is not given as it can be obtained in similar lines of proof of the Theorem 3.3.3 in Chapter 3. □

4.5 Numerical Experiment

In this section, we observe the performance of hp-DGFEM for the laser surface hardening of steel problem. For numerical experiments we consider the problem given by (4.2.7)-(4.2.12) with the given data as prescribed in the numerical experiment section of Chapter 3. We investigate the convergence of hp-DGFEM on a sequence of uniform meshes for each of degree of approximation $p = 1$ and 2. Similarly, convergence has been established by enriching the polynomial degree p for a fixed mesh.

For the purpose of computation penalty parameter is taken as $\sigma = 10$. In Figure 4.2, we plot the L^2-norm of the error against the mesh function h for polynomial degree $p = 1, 2$. Here, we observe that $\|\theta - \theta_\sigma\|$ and $\|a - a_\sigma\|$ converges to zero at the rate of $\mathcal{O}(h^p)$ as the mesh is refined. These experiments illustrates the theoretical results obtained in Theorem 4.3.1. In Figure 4.3, we present the convergence of the error in L^2-norm as the degree of polynomials increases on the fixed mesh. Figure 4.4(a) shows the convergence of k-refinement with piecewise constant approximation for temperature $'\theta'$ and austenite $'a'$. We plot that in L^2-norm the error against the time mesh function k. In Figure 4.4(b), we show the the error of control function u in L^2-norm against the time mesh function k. Figure 4.5(a) and 4.5(b) shows the temperature and austenite graph at the final time T after using hp-DGFEM for the discretization in space.

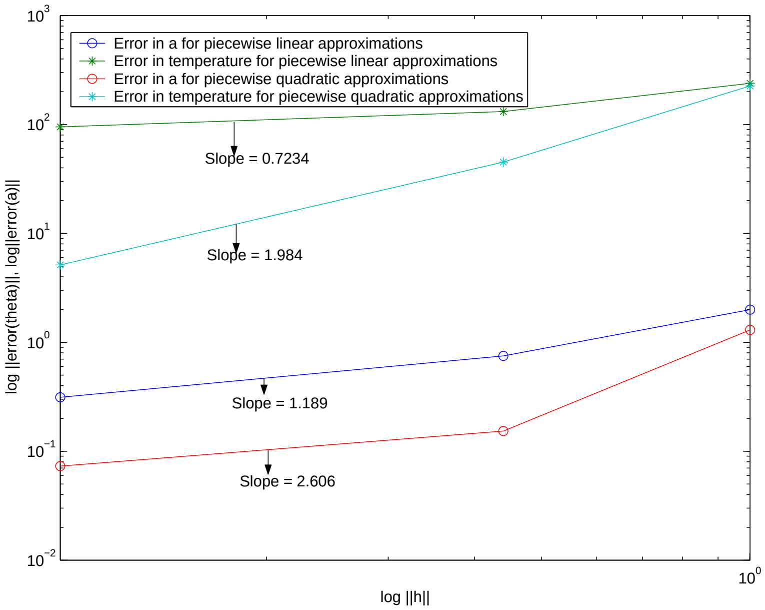
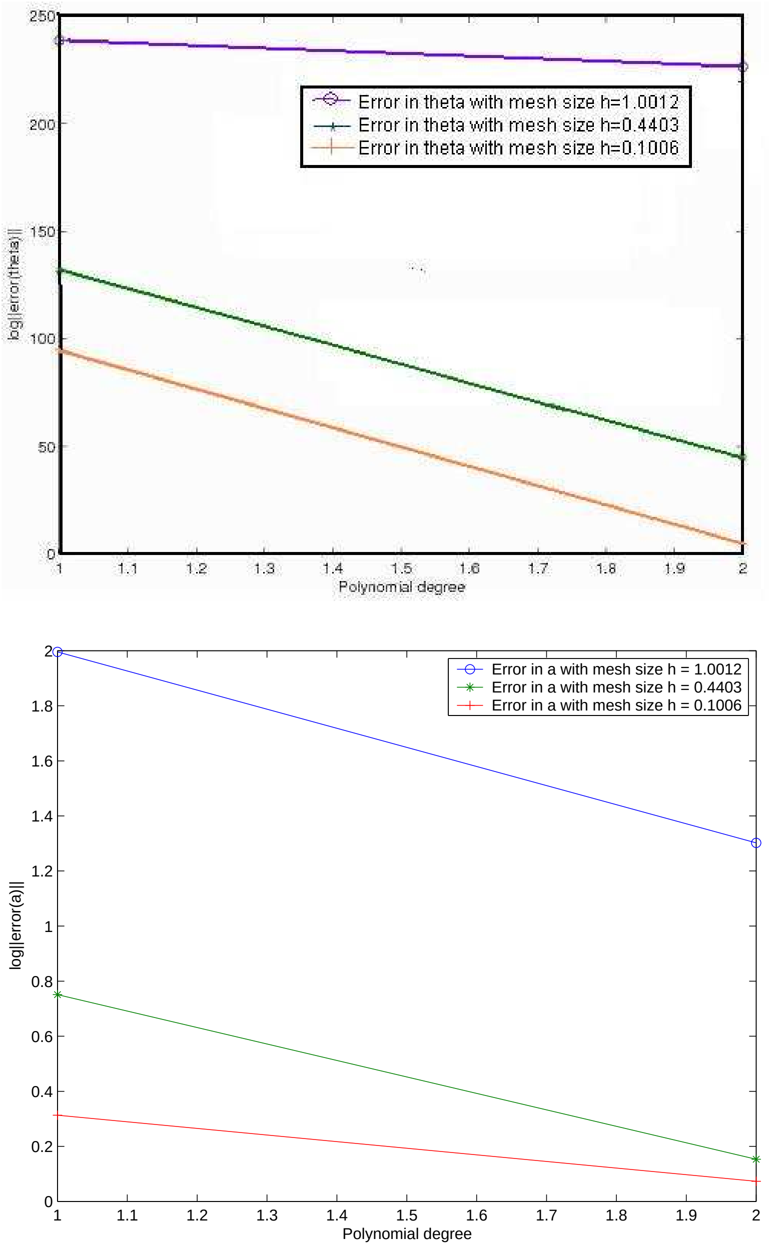


Figure 4.2: **Convergence of hp-DGFEM with h-refinement: Temperature and Austenite**



(a)

(b)

Figure 4.3: **Convergence of *hp*-DGFEM with *p*-refinement. (a) Temperature (b) Austenite**

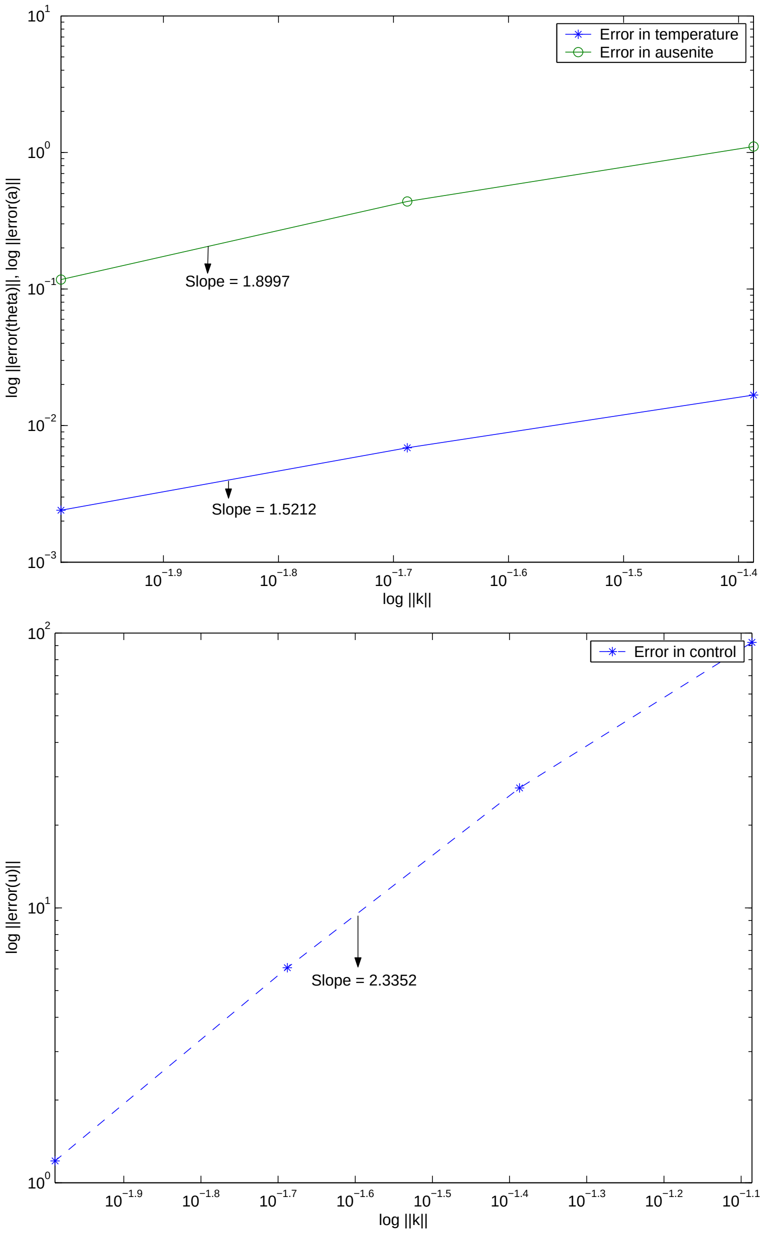


Figure 4.4: **Convergence of DGFEM with k-refinement**

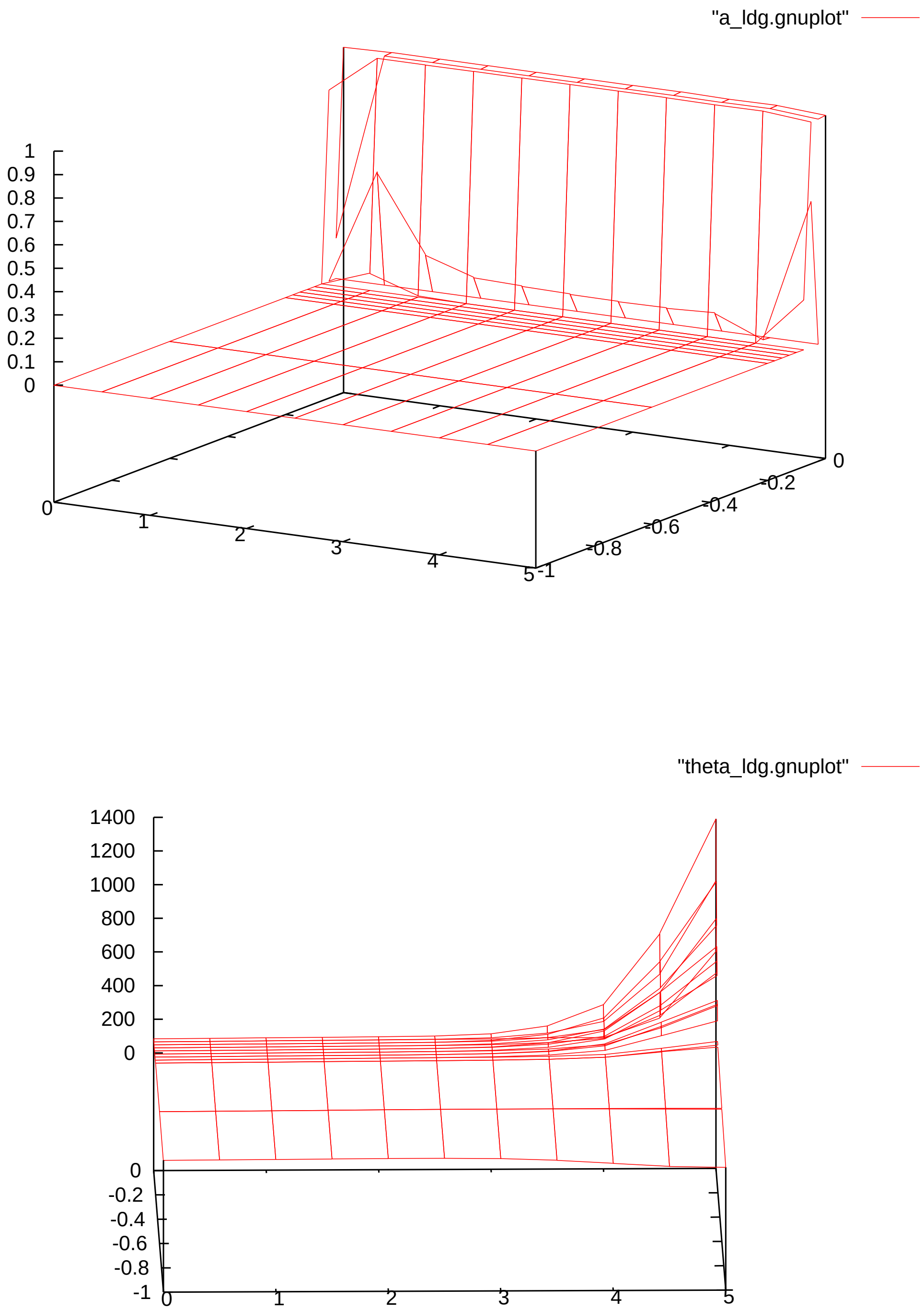


Figure 4.5: **(a) Temperature (b) Austenite**

4.6 Summary

A use of *hp*-DGFEM for space discretization, dG(0) for time and control discretizations has been done for the laser surface hardening of steel problem. In Theorems 4.4.1 and 4.4.2, it has been proved that the approximate solution converges to the solution of the regularized problem at the rate $\mathcal{O}\Bigg(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Bigg(\bigg(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\bigg)\frac{h_K^{2s-2}}{p_K^{2s'-3}}+k_n^2\Bigg)\Bigg)$. It has been observed through the numerical experiments that the rate of convergence is optimal.

Chapter 5

Adaptive Finite Element Methods

5.1 Introduction

Adaptive Finite Element Methods (AFEM) are amongst one of the important means to boost the accuracy and efficiency of the finite element discretization. It ensures higher density of nodes in certain areas of computational domain, where the solution is more difficult to approximate. Estimates obtained are called *a posteriori* error estimates as they depend on the approximate solution and data given, and the refinement of meshes is done based on the estimate of the discretization error. *A posteriori* error estimation for finite element methods for two point elliptic boundary value problems began with the pioneering work of Babuška and Rheinboldt [6]. The use of adaptive technique based on *a posteriori* error estimation is well accepted in the context of finite element discretization of partial differential equations, see Bank [9], Becker and Rannacher [8], [13], [14], Eriksson and Johnson [21], [22], Verfurth [83]. For *a posteriori* error estimates for elliptic equations using residual type estimator, see [6], [9] and [83]. Estimates using Dual Weighted Residual (DWR) method are developed in [8], [13], [14] and the references cited in there. AFEM for linear parabolic problems are also studied in [21], [22] using residual type estimators and in [8] using DWR type estimators, to mention a few.

Energy type error estimation for the error in the control, state and adjoint variables using residual method are developed in Liu and Yan [53], [55] and [56] in the context of distributed optimal control problems governed by elliptic equation subject to pointwise control constraints. These techniques are also been applied to optimal control problem governed by linear parabolic differential equations, see Liu, Ma, Tang, Yan [54] and Liu, Yan [57].

DWR method ([[8], [10], [11], [13], [14], [61] and [72]]) is a refined approach than the residual based adaptive strategy in the sense it helps in providing optimal meshes. Residual based estimator are estimated in L^2 or energy based norms involving local residuals of the computed solution, whereas DWR method is useful in estimating the error bounds not only in energy, L^2 norm but also on some quantity of physical interest, like, point value error, point value derivative error, mean normal flux etc. (see [13], [14] and [72]).

In this chapter we will discuss two types of AFEM namely, a residual based AFEM and a DWR type AFEM for the optimal control problem of laser surface hardening of steel. *A posteriori* error estimates are developed to keep the temperature under control near the heated zone. In the earlier chapters on the same problem, for the implementation purpose non-uniform meshes have been used to apply Galerkin approximation. Even though non-uniform meshes are helpful in giving the results near to desired observation, they can be quite expensive. Triangulations used in Chapter 3 and 4 are more refined near the heated zone and coarse far from the operational area but the mesh used for the approximation, chosen a priori, is independent of the approximate solution of the problem. In this chapter, error estimates have been developed using the discrete solution of the problem to help the refinement of the triangulation near the heating zone.

First of all, residual type estimators, which are based on error in L^2-norms are developed. Then, a DWR method which is based on duality argument has been applied to develop an *a posteriori* error estimate of the form:

$$|J(\theta^*, a^*, u^*) - J(\theta^*_\sigma, a^*_\sigma, u^*_\sigma)| \leq \eta_h + \eta_k + \eta_d,$$

where η_h is the space discretization error, η_k is the time discretization error and η_d is the error due to the discretization of control variable. Here $h > 0$, $k > 0$ and $d > 0$ respectively, are space, time and control discretization parameters.

The outline for this chapter is as follows. Section 5.1 is introductory in nature. In Section 5.2, a detailed description of residual method to compute *a posteriori* error estimates for the laser surface hardening of steel problem is given. In Section 5.3, *a posteriori* error estimates using DWR method are developed. Section 5.4 is devoted to numerical implementation, where results obtained using the both the methods are presented and compared.

5.2 Preliminaries

As in Chapters 3 and 4, for the sake of notational simplicity $(\theta_\epsilon, a_\epsilon, u_\epsilon)$ and f_ϵ in the regularized form will be replaced by (θ, a, u) and f respectively, throughout the chapter.

- For the sake of continuity of reading, we state the weak formulation of the regularized version of laser surface hardening of steel problem given by:

$$\min_{u\in U_{ad}} J(\theta,a,u) \text{ subject to} \tag{5.2.1}$$

$$\begin{aligned}(\partial_t a, w) &= (f(\theta,a),w) \quad \forall w\in H, \text{ a.e. in } I, &(5.2.2)\\ a(0) &= 0, &(5.2.3)\\ \rho c_p(\partial_t\theta, v)+\mathcal{K}(\bigtriangledown\theta,\bigtriangledown v) &= -\rho L(\partial_t a, v)+(\alpha u, v) \quad \forall v\in V, \text{ a.e. in } I, &(5.2.4)\\ \theta(0) &= \theta_0. &(5.2.5)\end{aligned}$$

where $J(\theta,a,u)=\dfrac{\beta_1}{2}\displaystyle\int_\Omega |a(T)-a_d|^2dx+\dfrac{\beta_2}{2}\int_0^T\int_\Omega[\theta-\theta_m]_+^2dxds+\dfrac{\beta_3}{2}\int_0^T|u|^2ds$ with other notations as defined in Chapter 2. The existence of a unique solution to the state equations (5.2.2)-(5.2.5) (see Chapter 2) ensures the existence of a control-to-state mapping $u\mapsto(\theta,a)=(\theta(u),a(u))$ through (5.2.2)-(5.2.5). By means of this mapping, we introduce the reduced cost functional $j: U_{ad}\longrightarrow\mathbb{R}$ as

$$j(u)=J(\theta(u),a(u),u). \tag{5.2.6}$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u\in U_{ad}} j(u). \tag{5.2.7}$$

Also $j(\cdot)$ satisfies,

$$j'(u^*)(p-u^*) \;\geq\; 0 \quad \forall p\in U_{ad}, \tag{5.2.8}$$

where $u^*\in U_{ad}$ is the optimal control of (5.2.7) and

$$j'(u)(p-u) \;=\; \left(\beta_3 u+\int_\Omega \alpha z dx,\ p-u\right)_{L^2(I)}. \tag{5.2.9}$$

- **Average interpolation Operator [37]**: Let $\pi_h : V \longrightarrow V_h$ be the *average interpolation* operator satisfying the following error estimates: for $v \in H^1(\Omega)$

$$\|v - \pi_h v\|_{H^l(K)} \leq C \sum_{K \bigcap K' \neq 0} h_K^{m-l} |v|_{H^m(K)},\ v \in H^1(K),\ l = 0, 1,\ l \leq m \leq 2. \quad (5.2.10)$$

- **Space-time interpolation operator [37]**: Let $\phi_I \in X_{hk}^q$ be the interpolant of ϕ such that

$$\phi_I|_{\Omega \times I_n} = \pi_{h,n} \pi_n \phi \quad n = 1, 2, \cdot \cdot \cdot, N, \quad (5.2.11)$$

where $\pi_{h,n}$ is the average interpolation operator satisfying (5.2.10) and $\pi_n : L^2(I_n) \longrightarrow P_0(I_n)$ is the L^2-projection operator, where $P_0(I_n)$ is the space of constant polynomials in I_n defined in the variable t. Also, we have

$$\|\phi - \pi_n \phi\|_{I_n,\Omega} \leq C k_n \|\partial_t \phi\|_{I_n,\Omega}. \quad (5.2.12)$$

- **Trace Inequality [37]**: For $v \in H^1(\Omega), 1 \leq q \leq \infty$,

$$\|v\|_{L^2(\partial K)} \leq C \bigg(h_K^{\frac{1}{q}} \|v\|_K + h_K^{1-\frac{1}{q}} |v|_{H^1(K)} \bigg). \quad (5.2.13)$$

- **Patch-wise interpolation** [8], [12]: Let $I_h : X_{kh}^q \rightarrow X_{k\ 2h}^q$ be the piecewise biquadratic spatial interpolant which is constructed with the help of the patch structure of the underlying mesh (see Figure 5.1) by conforming four adjacent cells to a macrocell on which the biquadratic interpolation defined. Also, let $I_k : X_{kh}^0 \rightarrow X_{kh}^1$ be the piecewise linear interpolation of the piecewise constant functions in time variable.

- **Hanging nodes continuity** [8]: To facilitate the refinement and coarsening procedure, use of hanging nodes becomes important. In order to use the conformal finite element method, global continuity is preserved by eliminating the unknowns at the hanging nodes by interpolating between neighbouring regular nodes. The interpolation is based on the transition of the cells with width h to $h/2$, as shown in Figure 5.1.

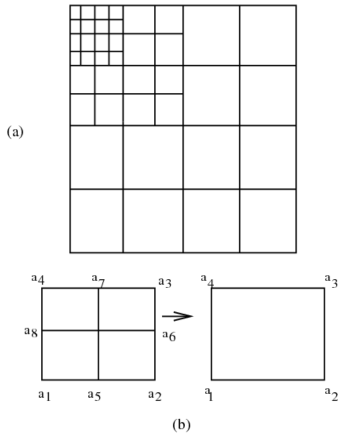


Figure 5.1: **(a) Patched mesh (b) Macrocell from four adjacent cells.**

5.3 Residual Method

Now we state the completely discrete problem (3.3.41)-(3.3.44) for the continuity of reading. The discretized problem reads as:

$$\begin{aligned}
\min_{u_\sigma \in U_{d,ad}} J(\theta_\sigma, a_\sigma, u_\sigma) & \qquad \text{subject to} && (5.3.1)\\
\sum_{n=1}^{N} (\partial_t a_\sigma, w)_{I_n,\Omega} &+ \sum_{n=1}^{N-1} ([a_\sigma]_n, w_n^+) + (a_{\sigma,0}^+, w_0^+) = (f(\theta_\sigma, a_\sigma), w)_{I,\Omega}, && (5.3.2)\\
a_\sigma(0) &= 0, && (5.3.3)\\
\rho c_p \sum_{n=1}^{N} (\partial_t \theta_\sigma, v)_{I_n,\Omega} &+ \mathcal{K}(\bigtriangledown \theta_\sigma, \bigtriangledown v)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_\sigma]_n, v_n^+) + \rho c_p(\theta_{\sigma,0}^+, v_0^+) && \\
&= -\rho L(f(\theta_\sigma, a_\sigma), v)_{I,\Omega} + (\alpha u_\sigma, v)_{I,\Omega} + \rho c_p(\theta_0, v_0^+), && (5.3.4)\\
\theta_\sigma(0) &= \theta_0, && (5.3.5)
\end{aligned}$$

for all $(w, v) \in X_{hk}^q \times X_{hk}^q$, where X_{hk}^q is the complete discrete space, defined by

$$X_{hk}^q = \{\phi : I \to V_h;\ \phi|_{I_n} = \sum_{j=0}^{q} \psi_j t^j, \psi_j \in V_h\},\ \ q \in \mathbb{N}. \qquad (5.3.6)$$

The existence of a unique solution to the state equation (5.3.2)-(5.3.5) (discussed in Chapter 3), ensures the existence of a control-to-state mapping $u_\sigma \mapsto (\theta_\sigma, a_\sigma) = (\theta_\sigma(u_\sigma), a_\sigma(u_\sigma))$ through (5.3.2)-(5.3.5). By means of this mapping, we introduce the reduced cost functional $j_\sigma : U_{d,ad} \longrightarrow \mathbb{R}$ as

$$j_\sigma(u_\sigma) = J(\theta_\sigma(u_\sigma), a_\sigma(u_\sigma), u_\sigma). \qquad (5.3.7)$$

Then the optimal control problem can be equivalently reformulated as

$$\min_{u_\sigma \in U_{d,ad}} j_\sigma(u_\sigma). \tag{5.3.8}$$

Also $j_\sigma(\cdot)$ satisfies,

$$j'_\sigma(u^*_\sigma)(p - u^*_\sigma) \ \geq \ 0 \quad \forall p \in U_{d,ad}, \tag{5.3.9}$$

where $u^*_\sigma \in U_{ad}$ is the optimal control of (5.3.8) and

$$j'_\sigma(u_\sigma)(p - u_\sigma) \ = \ \left(\beta_3 u_\sigma + \int_\Omega \alpha z_\sigma dx, \ p - u_\sigma\right)_{L^2(I)}. \tag{5.3.10}$$

The corresponding adjoint system is given by: Find $(z^*_\sigma, \lambda^*_\sigma) \in X^q_{kh} \times X^q_{kh}$ such that

$$\begin{aligned}
-\sum_{n=1}^{N}(\psi, \partial_t \lambda^*_\sigma)_{I_n,\Omega} - \sum_{n=1}^{N-1}(\psi_n, [\lambda^*_\sigma]_n) \ &= \ -(\psi, f_a(\theta^*_\sigma, a^*_\sigma)(\rho L z^*_\sigma - \lambda^*_\sigma))_{I,\Omega}, && (5.3.11)\\
\lambda^*_\sigma(T) \ &= \ \beta_1(a^*_\sigma(T) - a_d), && (5.3.12)\\
-\rho c_p \sum_{n=1}^{N}(\phi, \partial_t z^*_\sigma)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z^*_\sigma)_{I,\Omega} \ &- \ \rho c_p \sum_{n=1}^{N-1}(\phi_n, [z^*_\sigma]_n) = -(\phi, f_\theta(\theta^*_\sigma, a^*_\sigma)(\rho L z^*_\sigma - \lambda^*_\sigma))_{I,\Omega} \\
&\quad + \beta_2(\phi, [\theta^*_\sigma - \theta_m]_+)_{I,\Omega}, && (5.3.13)\\
z^*_\sigma(T) \ &= \ 0, && (5.3.14)
\end{aligned}$$

for all $(\phi, \psi) \in X^q_{hk} \times X^q_{hk}$.

Now we define the following auxiliary problem, which will help us in estimating the errors. Let $(\theta^{u_\sigma}, a^{u_\sigma}) \in X \times Y$ be the solution of the state system with control u chosen as u^*_σ in the right hand side of (5.2.4), that is, $(\theta^{u_\sigma}, a^{u_\sigma}) \in X \times Y$ is the solution of

$$\begin{aligned}
(\partial_t a^{u_\sigma}, w) \ &= \ (f(\theta^{u_\sigma}, a^{u_\sigma}), w) \quad \forall w \in H, \ \text{a.e. in } I, && (5.3.15)\\
a^{u_\sigma}(0) \ &= \ 0, && (5.3.16)\\
\rho c_p(\partial_t \theta^{u_\sigma}, v) + \mathcal{K}(\bigtriangledown\theta^{u_\sigma}, \bigtriangledown v) \ &= \ -\rho L(\partial_t a^{u_\sigma}, v) + (\alpha u^*_\sigma, v) \ \ \forall v \in V, \ \text{a.e. in } I, && (5.3.17)\\
\theta^{u_\sigma}(0) \ &= \ \theta_0, && (5.3.18)
\end{aligned}$$

where $X = \{v \in L^2(I; V) : v_t \in L^2(I; V^*)\}$, $V = H^1(\Omega)$, $Y = H^1(I; L^2(\Omega))$, $H = L^2(\Omega)$. Let $(z^{u_\sigma}, \lambda^{u_\sigma}) \in X \times Y$ denote the solution of the corresponding adjoint system defined by:

$$\begin{aligned}
-(\psi, \partial_t \lambda^{u_\sigma}) &= -(\psi, f_a(\theta^{u_\sigma}, a^{u_\sigma})(\rho L z^{u_\sigma} - \lambda^{u_\sigma})) && (5.3.19)\\
\lambda^{u_\sigma}(T) &= \beta_1(a^{u_\sigma}(T) - a_d), && (5.3.20)\\
-\rho c_p(\phi, \partial_t z^{u_\sigma}) + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z^{u_\sigma}) &= -(\phi, f_\theta(\theta^{u_\sigma}, a^{u_\sigma})(\rho L z^{u_\sigma} - \lambda^{u_\sigma}))\\
&\quad + \beta_2(\phi, [\theta^{u_\sigma} - \theta_m]_+) && (5.3.21)\\
z^*(T) &= 0, && (5.3.22)
\end{aligned}$$

for all $(\phi, \chi) \in V \times H$ and a.e. in I.

We now proceed to develop the *a posteriori* error estimates based on residual type estimators. In Lemmas 5.3.1, 5.3.2 and 5.3.3, we develop local estimators for control, adjoint and state errors, respectively. In Theorem 5.3.1 we present the *a posteriori* error estimator for control, state and adjoint variables.

Lemma 5.3.1. *Let* $(\theta^*, a^*, u^*) \in X \times Y \times U_{ad}$ *and* $(\theta^*_\sigma, a^*_\sigma, u^*_\sigma) \in X^q_{hk} \times X^q_{hk} \times U_{d,ad}$ *be the solutions of* (5.2.1)-(5.2.5) *and* (5.3.1)-(5.3.5), *respectively. Then we have*

$$\|u^* - u^*_\sigma\|^2_{L^2(I)} \;\le\; C(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} \eta^2_{1,n,K} + \max_{t\in I_n} \|\alpha\|^2_K \|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega}),$$

where $\eta^2_{1,n,K} = \|u^*_\sigma\|^2_{L^2(I_n)} + \max_{I_n} \|\alpha\|^2_K \|z^*_\sigma\|^2_{I_n,K}$, *and* $(z^{u_\sigma}, \lambda^{u_\sigma})$ *is the solution of the adjoint problem* (5.3.19)-(5.3.22).

Proof: From the first order optimality condition (5.2.8), we have

$$j'(u^*)(u^* - u^*_\sigma) \le 0. \tag{5.3.23}$$

We have

$$C\|u^* - u^*_\sigma\|^2_{L^2(I)} \;\le\; j'(u^*)(u^* - u^*_\sigma) - j'(u^*_\sigma)(u^* - u^*_\sigma). \tag{5.3.24}$$

Using (5.3.23) in (5.3.24), we obtain

$$C\|u^* - u^*_\sigma\|^2_{L^2(I)} \;\le\; -j'(u^*_\sigma)(u^* - u^*_\sigma). \tag{5.3.25}$$

Adding and subtracting the term $j'_\sigma(u^*_\sigma)(u^* - u^*_\sigma)$ in the right hand side of (5.3.25), using the definitions of $j'_\sigma(u^*_\sigma)(u^* - u^*_\sigma)$ and $j'(u^*_\sigma)(u^* - u^*_\sigma)$, we obtain

$$\begin{aligned} C\|u^* - u^*_\sigma\|^2_{L^2(I)} &\leq \sum_{n=1}^{N}\Big((\beta_3 u^*_\sigma + \int_\Omega \alpha z^*_\sigma dx, u^*_\sigma - u^*)_{L^2(I_n)} + (\int_\Omega \alpha(z^*_\sigma - z^{u_\sigma})dx, u^* - u^*_\sigma)_{L^2(I_n)}\Big) \\ &= J_1 + J_2, \quad \text{say.} \end{aligned} \tag{5.3.26}$$

Consider

$$\begin{aligned} J_1 &= \sum_{n=1}^{N}(\beta_3 u^*_\sigma + \int_\Omega \alpha z^*_\sigma dx, u^*_\sigma - u^*)_{L^2(I_n)} = \sum_{n=1}^{N}\Big((\beta_3 u^*_\sigma, u^*_\sigma - u^*)_{L^2(I_n)} \\ &\quad + (\int_\Omega \alpha z^*_\sigma dx, u^*_\sigma - u^*)_{L^2(I_n)}\Big). \end{aligned}$$

Using Cauchy-Schwarz inequality, we obtain

$$J_1 \leq C\sum_{n=1}^{N}\Big(\|u^*_\sigma\|_{L^2(I_n)}\|u^* - u^*_\sigma\|_{L^2(I_n)} + \sum_{K\in\mathcal{T}_h}\max_{t\in I_n}\|\alpha\|_K\|z^*_\sigma\|_{I_n,K}\|u^* - u^*_\sigma\|_{L^2(I_n)}\Big). \tag{5.3.27}$$

Now consider

$$\begin{aligned} J_2 &= \sum_{n=1}^{N}\Big((\int_\Omega \alpha(z^*_\sigma - z^{u_\sigma})dx, u^* - u^*_\sigma)_{L^2(I_n)}\Big) \\ &\leq C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(\max_{t\in I_n}\|\alpha\|_K\ \|z^*_\sigma - z^{u_\sigma}\|_{I_n,K}\|u^* - u^*_\sigma\|_{L^2(I_n)}\Big). \end{aligned}$$

Using estimates (5.3.27), (5.3.28) in (5.3.26) and Young's inequality, we obtain

$$\begin{aligned} \|u^* - u^*_\sigma\|^2_{L^2(I)} &\leq C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(\|u^*_\sigma\|^2_{L^2(I_n)} + \max_{t\in I_n}\|\alpha\|^2_K\ \|z^*_\sigma\|^2_{I_n,K} + \max_{t\in I_n}\|\alpha\|^2_K\ \|z^*_\sigma - z^{u_\sigma}\|^2_{I_n,K} \\ &\quad + \mu\|u^* - u^*_\sigma\|^2_{L^2(I_n)}\Big). \end{aligned}$$

Choosing the constants in the Young's inequality appropriately, we have the final result

$$\|u^* - u^*_\sigma\|^2_{L^2(I)} \leq C(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta^2_{1,n,K} + \max_{t\in I_n}\|\alpha\|^2_K\|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega}).$$

where

$$\eta^2_{1,n,K} = \|u^*_\sigma\|^2_{L^2(I_n)} + \max_{I_n} \|\alpha\|^2_K \ \|z^*_\sigma\|^2_{I_n,K}. \quad \square$$

Lemma 5.3.2. *Let (θ^*, a^*), $(\theta^*_\sigma, a^*_\sigma)$ and $(\theta^{u_\sigma}, a^{u_\sigma})$ be respectively the solutions of* (5.2.2)-(5.2.5), (5.3.2)-(5.3.5) *and* (5.3.15)-(5.3.18) *with (z^*, λ^*), $(z^*_\sigma, \lambda^*_\sigma)$ and $(z^{u_\sigma}, a^{u_\sigma})$ as the corresponding adjoint solutions. Then,*

$$\|z^{u_\sigma} - z^*_\sigma\|^2 + \|\lambda^{u_\sigma} - \lambda^*_\sigma\|^2 \ \leq \ C\Bigg(\sum_{j=2}^{10} \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \eta^2_{j,n,K} + \|\theta^{u_\sigma} - \theta^*_\sigma\|^2 \Bigg),$$

where

$$\begin{aligned}
\eta^2_{2,n,K} &= h^4_K \|r_z(x,t)\|^2_{I_n,K}, & (5.3.28)\\
r_z(x,t) &= \partial_t z^*_\sigma + \beta_2[\theta^*_\sigma - \theta_m]_+ + \mathcal{K}\Delta z^*_\sigma + \rho c_p \frac{[z^*_\sigma]_{n-1}}{k_n} - f_\theta(\theta^*_\sigma, a^*_\sigma)(\rho L z^*_\sigma - \lambda^*_\sigma) & (5.3.29)\\
\eta^2_{3,n,K} &= k^2_n(\|[\theta^*_\sigma - \theta_m]_+\|^2_{I_n,K} + \|\Delta z^*_\sigma\|^2_{I_n,K} + \|\rho L z^*_\sigma - \lambda^*_\sigma\|^2_{I_n,K}), & (5.3.30)\\
\eta^2_{4,n,K} &= h^3_K \|\mathcal{K}[\bigtriangledown z^*_\sigma].\eta\|^2_{L^2(I_n, L^2(\partial K))}, \ \eta^2_{5,n,K} = \|z^*_\sigma\|^2_{I_n,K}, & (5.3.31)\\
\eta^2_{6,n,K} &= k_n \|[z^*_\sigma]_{n-1}\|^2_{I_n,K}, & (5.3.32)\\
\eta^2_{7,n,K} &= k^2_n \|\rho L z^*_\sigma - \lambda^*_\sigma\|^2_{I_n,K}, \ \eta^2_{8,n,K} = \|\lambda^*_\sigma\|^2_{I_n,K}, \eta^2_{9,n,K} = \|z^*_\sigma\|^2_{I_n,K}, & (5.3.33)\\
\eta^2_{10,n,K} &= k_n \|[\lambda^*_\sigma]_{n-1}\|^2_K. & (5.3.34)
\end{aligned}$$

Proof: Consider the auxiliary problem defined by:
For given $g \in L^2(I, L^2(\Omega))$, find ϕ such that

$$\begin{aligned}
\rho c_p \partial_t \phi - \mathcal{K}\Delta\phi + \rho L f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\phi &= g \quad \text{in } Q, & (5.3.35)\\
\frac{\partial \phi}{\partial \eta}|_{\partial\Omega} &= 0 \quad \text{in } I, & (5.3.36)\\
\phi(0) &= 0 \quad \text{in } \Omega, & (5.3.37)
\end{aligned}$$

where $\frac{\partial \phi}{\partial \eta}|_{\partial\Omega}$ denotes the outward normal derivative to $\partial\Omega$. Then the solution to (5.3.35)-(5.3.37) satisfies [23]:

$$\begin{aligned}
\|\phi\|_{L^\infty(I;L^2(\Omega))} &\leq C\|g\|_{I,\Omega}, \quad \|\phi\|_{L^2(I;H^1(\Omega))} \leq C\|g\|_{I,\Omega}, & (5.3.38)\\
\|\phi\|_{L^2(I;H^2(\Omega))} &\leq C\|g\|_{I,\Omega}, \quad \|\partial_t \phi\|_{I,\Omega} \leq C\|g\|_{I,\Omega}. & (5.3.39)
\end{aligned}$$

Substitute $g = z_\sigma^* - z^{u_\sigma}$ in (5.3.35) and consider

$$\begin{aligned}
\|z_\sigma^* - z^{u_\sigma}\|_{I,\Omega}^2 &= \int_0^T (z_\sigma^* - z^{u_\sigma}, \rho c_p \partial_t \phi - \mathcal{K}\Delta\phi + \rho L f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\phi)\, dt \\
&= \int_0^T \Big(- \rho c_p(\partial_t(z_\sigma^* - z^{u_\sigma}), \phi) + \mathcal{K}(\bigtriangledown(z_\sigma^* - z^{u_\sigma}), \bigtriangledown\phi) \\
&\quad + (\rho L f_\theta(\theta^{u_\sigma}, a^{u_\sigma})(z_\sigma^* - z^{u_\sigma}), \phi)\Big) dt - \rho c_p \sum_{n=1}^{N}([z_\sigma^*]_{n-1}, \phi_{n-1}^+) \qquad (5.3.40)
\end{aligned}$$

Adding and subtracting the terms $(\beta_2[\theta_\sigma^* - \theta_m]_+, \phi)$, $(f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*), \phi)$, $(f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\lambda^{u_\sigma}, \phi)$, $\rho c_p(\frac{[z_\sigma^*]_{n-1}}{k_n}, \phi)$ on the right hand side of (5.3.40) and using (5.3.11)-(5.3.14), we obtain

$$\begin{aligned}
\|z_\sigma^* - z^{u_\sigma}\|_{I,\Omega}^2 &= \sum_{n=1}^{N} \Big[\int_{I_n} \Big(- \rho c_p(\partial_t z_\sigma^*, \phi) - (\beta_2[\theta_\sigma^* - \theta_m]_+, \phi) + \mathcal{K}(\bigtriangledown z_\sigma^*, \bigtriangledown\phi) \\
&\quad + (f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*), \phi) + \beta_2([\theta_\sigma^* - \theta_m]_+ - [\theta^{u_\sigma} - \theta_m]_+, \phi) \\
&\quad - (\rho c_p \frac{[z_\sigma^*]_{n-1}}{k_n}, \phi) + \rho c_p(\frac{[z_\sigma^*]_{n-1}}{k_n}, \phi - \phi_{n-1}^+) + (f_\theta(\theta_\sigma^*, a_\sigma^*)\lambda_\sigma^* - f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\lambda^{u_\sigma}, \phi) \\
&\quad + (\rho L(f_\theta(\theta^{u_\sigma}, a^{u_\sigma}) - f_\theta(\theta_\sigma^*, a_\sigma^*))z_\sigma^*, \phi)\Big) dt \Big]
\end{aligned}$$

First adding (5.3.13) to the right hand side of the above equation after replacing ϕ by the interpolant ϕ_I and then, adding and subtracting $(\rho c_p \frac{[z_\sigma^*]_{n-1}}{k_n}, \phi_I)$, we have

$$\begin{aligned}
\|z_\sigma^* - z^{u_\sigma}\|_{I,\Omega}^2 &= \sum_{n=1}^{N} \Big[\int_{I_n} \Big(- \rho c_p(\partial_t z_\sigma^*, \phi - \phi_I) - (\beta_2[\theta_\sigma^* - \theta_m]_+, \phi - \phi_I) + \mathcal{K}((\bigtriangledown z_\sigma^*), \bigtriangledown(\phi - \phi_I)) \\
&\quad + \beta_2([\theta_\sigma^* - \theta_m]_+ - [\theta^{u_\sigma} - \theta_m]_+, \phi) + (f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*), \phi - \phi_I) \\
&\quad - (\rho c_p \frac{[z_\sigma^*]_{n-1}}{k_n}, \phi - \phi_I) + (f_\theta(\theta_\sigma^*, a_\sigma^*)\lambda_\sigma^* - f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\lambda^{u_\sigma}, \phi) + (\rho L(f_\theta(\theta^{u_\sigma}, a^{u_\sigma}) \\
&\quad - f_\theta(\theta_\sigma^*, a_\sigma^*))z_\sigma^*, \phi) + \rho c_p(\frac{[z_\sigma^*]_{n-1}}{k_n}, (\phi_I)_{n-1}^+ - \phi_I + \phi - \phi_{n-1}^+)\Big) dt \Big]
\end{aligned}$$

Integrating by parts the 3^{rd} term on the right hand side, we obtain

$$\begin{aligned}
\|z_\sigma^* &- z^{u_\sigma}\|_{I,\Omega}^2 \\
&= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big[\int_{I_n} \Big(- \rho c_p(\partial_t z_\sigma^*, \phi - \phi_I)_K - (\beta_2[\theta_\sigma^* - \theta_m]_+, \phi - \phi_I)_K \\
&\quad - \mathcal{K}(\Delta z_\sigma^*, \phi - \phi_I)_K + \mathcal{K}([\bigtriangledown z_\sigma^*].\eta, \phi - \phi_I)_{L^2(\partial K)} + \beta_2([\theta_\sigma^* - \theta_m]_+ - [\theta^{u_\sigma} - \theta_m]_+, \phi)_K \\
&\quad + (f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*), \phi - \phi_I)_K + (\rho L(f_\theta(\theta^{u_\sigma}, a^{u_\sigma}) - f_\theta(\theta_\sigma^*, a_\sigma^*))z_\sigma^*, \phi)_K
\end{aligned}$$

$$
\begin{aligned}
&+ (f_\theta(\theta^*_\sigma, a^*_\sigma)\lambda^*_\sigma - f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\lambda^{u_\sigma}, \phi)_K - (\rho c_p \frac{[z^*_\sigma]_{n-1}}{k_n}, \phi - \phi_I)_K \\
&+ \rho c_p (\frac{[z^*_\sigma]_{n-1}}{k_n}, (\phi_I)^+_{n-1} - \phi_I + \phi - \phi^+_{n-1})_K \Big) dt \Big] \\
= \;& J_1 + J_2 + J_3 + J_4 + J_5 + J_6, \text{ say,} \qquad (5.3.41)
\end{aligned}
$$

$$
\begin{aligned}
\text{where} \quad J_1 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(- (r_z(x,t), \phi - \phi_I)_K \Big) dt, \\
r_z(x,t) &= \partial_t z^*_\sigma + \beta_2[\theta^*_\sigma - \theta_m]_+ + \mathcal{K}\Delta z^*_\sigma + \rho c_p \frac{[z^*_\sigma]_{n-1}}{k_n} - f_\theta(\theta^*_\sigma, a^*_\sigma)(\rho L z^*_\sigma - \lambda^*_\sigma) \\
J_2 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(\mathcal{K}([\bigtriangledown z^*_\sigma].\eta, \phi - \phi_I)_{L^2(\partial K)} \Big) dt, \\
J_3 &= \beta_2 \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(([\theta^*_\sigma - \theta_m]_+ - [\theta^{u_\sigma} - \theta_m]_+, \phi)_K \Big) dt \\
J_4 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big((f_\theta(\theta^*_\sigma, a^*_\sigma)\lambda^*_\sigma - f_\theta(\theta^{u_\sigma}, a^{u_\sigma})\lambda^{u_\sigma}, \phi)_K \Big) dt, \\
J_5 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big((\rho L (f_\theta(\theta^{u_\sigma}, a^{u_\sigma}) - f_\theta(\theta^*_\sigma, a^*_\sigma)) z^*_\sigma, \phi)_K \Big) dt, \\
J_6 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(\rho c_p (\frac{[z^*_\sigma]_{n-1}}{k_n}, (\phi_I)^+_{n-1} - \phi_I + \phi - \phi^+_{n-1})_K \Big) dt.
\end{aligned}
$$

Use Cauchy-Schwarz's inequality, (5.2.10), (5.2.11) and (5.2.12) to obtain

$$
\begin{aligned}
J_1 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(- (r_z(x,t), \phi - \pi_n\phi + \pi_n\phi - \pi_{h,n}\pi_n\phi)_K \Big) dt \qquad (5.3.42) \\
&\leq C \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(h_K^2 \|r_z(x,t)\|_{I_n,K} \|\phi\|_{L^2(I_n;H^2(K))} + k_n (\|[\theta^*_\sigma - \theta_m]_+\|_{I_n,K} + \|\Delta z^*_\sigma\|_{I_n,K} \\
&\quad + \|\rho L z^*_\sigma - \lambda^*_\sigma\|_{I_n,K}) \|\partial_t \phi\|_{I_n,K} \Big)
\end{aligned}
$$

Using Cauchy-Schwarz's inequality, (5.2.10) and (5.3.43), we obtain

$$
\begin{aligned}
J_2 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \int_{I_n} \Big(\mathcal{K}([\bigtriangledown z^*_\sigma].\eta, \phi - \phi_I)_{L^2(\partial K)} \Big) dt \\
&\leq C \Big(\sum_{n=1}^{N} \sum_{T \in \mathcal{T}_h} h_K^{\frac{3}{2}} \|\mathcal{K}[\bigtriangledown z^*_\sigma].\eta\|_{L^2(I_n, L^2(\partial K))} \|\phi\|_{L^2(I, H^2(\Omega))} \Big)
\end{aligned}
$$

Consider,

$$J_3 \;=\; \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}([\theta^*_\sigma-\theta_m]_+-[\theta^{u_\sigma}-\theta_m]_+,\phi)_K dt \le C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\|\theta^*_\sigma-\theta^{u_\sigma}\|_K\|\phi\|_K dt$$

Using Remark (2.2.1), Cauchy-Schwarz's inequality, and (5.3.38), we obtain

$$J_4 = \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\Big((f_\theta(\theta^*_\sigma,a^*_\sigma)\lambda^*_\sigma - f_\theta(\theta^{u_\sigma},a^{u_\sigma})\lambda^{u_\sigma},\phi)_K\Big)dt \le C\|\lambda^*_\sigma-\lambda^{u_\sigma}\|_{I,\Omega}\|\phi\|_{I,\Omega}.$$

Repeating similar calculations as for the term J_4, we obtain

$$\begin{aligned} J_5 \;&=\; \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(\rho L(f_\theta(\theta^{u_\sigma},a^{u_\sigma})-f_\theta(\theta^*_\sigma,a^*_\sigma))z^*_\sigma,\phi)_K dt \\ &\le\; C\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|z^*_\sigma\|_{I_n,K}\Big)\|\phi\|_{I,\Omega} \end{aligned}$$

We have,

$$\begin{aligned} J_6 \;&=\; \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\Big(\rho c_p(\frac{[z^*_\sigma]_{n-1}}{k_n},(\phi_I)^+_{n-1}-\phi_I+\phi-\phi^+_{n-1})_K\Big)dt \\ &\le\; C\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}k_n^{1/2}\|[z^*_\sigma]_{n-1}\|_{I_n,K}\Big(\|\partial_t\phi_I\|_{I,\Omega}+\|\partial_t\phi\|_{I,\Omega}\Big)\Big) \end{aligned}$$

Using (5.3.38)-(5.3.39) with $g = z^{u_\sigma}-z^*_\sigma$, we have

$$\|\phi\|_{L^2(I,H^2(\Omega))} \le C\|z^{u_\sigma}-z^*_\sigma\|_{I,\Omega} \;\text{ and }\; \|\partial_t\phi\|_{I,\Omega} \le C\|z^{u_\sigma}-z^*_\sigma\|_{I,\Omega} \tag{5.3.43}$$

Using estimates for J_1 to J_6 in (5.3.41), (5.3.43) and Young's inequality with the Young's constant chosen appropriately, we obtain

$$\begin{aligned} \|z^*_\sigma-z^{u_\sigma}\|^2_{I,\Omega} \;&\le\; C\Big(\sum_{i=2}^{6}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta^2_{i,n,K}+\|\theta^*_\sigma-\theta^{u_\sigma}\|^2_{I,\Omega}+\frac{\mu_1}{2}\|\lambda^*_\sigma-\lambda^{u_\sigma}\|^2_{I,\Omega} \\ &\quad+\frac{\mu}{2}\|z^*_\sigma-z^{u_\sigma}\|^2_{I,\Omega}\Big), \end{aligned} \tag{5.3.44}$$

where $\eta^2_{i,n,K}, 2 \le i \le 6$ are defined by

$$\begin{aligned}
\eta_2^2 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(h_K^4\|r_z(x,t)\|^2_{I_n,K} + k_n^2(\|[\theta^*_\sigma - \theta_m]_+\|^2_{I_n,K} + \|\Delta z^*_\sigma\|^2_{I_n,K} \\
&\quad +\|\rho L(z^*_\sigma - \lambda^*_\sigma)\|^2_{I_n,K})\Big), \\
\eta_3^2 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} h_K^3\|\mathcal{K}[\bigtriangledown z^*_\sigma].\eta\|^2_{L^2(I_n,L^2(\partial K))}, \\
\eta_4^2 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|z^*_\sigma\|^2_{I_n,K}, \eta_5^2 = \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} k_n\|[z^*_\sigma]_{n-1}\|^2_{I_n,K},
\end{aligned}$$

To estimate $\|\lambda^{u_\sigma} - \lambda^*_\sigma\|$, we proceed as follows.

Consider the auxiliary problem: for $G \in L^2(I, L^2(\Omega))$, find $\psi \in H^1(I, L^2(\Omega))$ such that

$$\partial_t\psi - f_a(\theta^{u_\sigma}, a^{u_\sigma})\psi = G \quad \text{in } Q, \tag{5.3.45}$$

$$\psi(0) = 0 \quad \text{in } \Omega. \tag{5.3.46}$$

(5.3.45)-(5.3.46) has a unique solution and we have [23]:

$$\|\psi\|_{L^\infty(I;L^2(\Omega))} \le C\|G\|_{I,\Omega}, \tag{5.3.47}$$

$$\|\partial_t\psi\|_{I,\Omega} \le C\|G\|_{I,\Omega}. \tag{5.3.48}$$

Let $G = \lambda^*_\sigma - \lambda^{u_\sigma}$ in (5.3.45) to obtain

$$\|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega} = \int_0^T(\lambda^*_\sigma - \lambda^{u_\sigma}, \partial_t\psi - f_a(\theta^{u_\sigma}, a^{u_\sigma})\psi)dt$$

$$\begin{aligned}
= \sum_{n=1}^{N}\Big[\int_{I_n}\Big(&-(\partial_t(\lambda^*_\sigma - \lambda^{u_\sigma}),\psi) - (f_a(\theta^*_\sigma, a^*_\sigma)(\lambda^*_\sigma - \lambda^{u_\sigma}),\psi)\Big)dt \\
&-([\lambda^*_\sigma]_{n-1}, \psi^+_{n-1})\Big].
\end{aligned}$$

Adding (5.3.11) to the right hand side of the above equation after replacing ψ by $\pi_n\psi$, we obtain

$$
\begin{aligned}
\|\lambda_\sigma^* - \lambda^{u_\sigma}\|_{I,\Omega}^2 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Bigg[\int_{I_n}\Bigg((r_\lambda(x,t),\psi-\pi_n\psi)_K - ((f_a(\theta^{u_\sigma},a^{u_\sigma}) - f_a(\theta_\sigma^*,a_\sigma^*))\lambda_\sigma^*,\psi)_K \\
&\quad - (\rho L(f_a(\theta_\sigma^*,a_\sigma^*) - f_a(\theta^{u_\sigma},a^{u_\sigma}))z_\sigma^*,\psi)_K - (\rho L f_a(\theta^{u_\sigma},a^{u_\sigma})(z_\sigma^* - z^{u_\sigma}),\psi)_K \\
&\quad + (\frac{[\lambda_\sigma^*]_{n-1}}{k_n},(\pi_n\psi)_{n-1}^+ + \psi - \pi_n\psi - \psi_{n-1}^+)_K\Bigg)dt\Bigg] = \sum_{i=1}^{5} J_i, \text{ say} \qquad (5.3.49)
\end{aligned}
$$

where $r_\lambda(x,t) = -\partial_t\lambda_\sigma^* + f_a(\theta_\sigma^*,a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*) - \dfrac{[\lambda_\sigma^*]_{n-1}}{k_n}$ and

$$
\begin{aligned}
J_1 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(r_\lambda(x,t),\psi-\pi_n\psi)_K\ dt, \\
J_2 &= -\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}((f_a(\theta^{u_\sigma},a^{u_\sigma}) - f_a(\theta_\sigma^*,a_\sigma^*))\lambda_\sigma^*,\psi)_K\ dt \\
J_3 &= -\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(\rho L(f_a(\theta_\sigma^*,a_\sigma^*) - f_a(\theta^{u_\sigma},a^{u_\sigma}))z_\sigma^*,\psi)_K\ dt \\
J_4 &= -\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(\rho L f_a(\theta^{u_\sigma},a^{u_\sigma})(z_\sigma^* - z^{u_\sigma}),\psi)_K\ dt \\
J_5 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(\frac{[\lambda_\sigma^*]_{n-1}}{k_n},(\pi_n\psi)_{n-1}^+ - \psi - \pi_n\psi - \psi_{n-1}^+)_K\ dt.
\end{aligned}
$$

Using Remark (2.2.1), Cauchy Schwarz inequality, (5.2.12) and (5.3.47), we obtain

$$
\begin{aligned}
J_1 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(r_\lambda(x,t),\psi-\pi_n\psi)_K\ dt \le C\Bigg(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} k_n\|\rho L z_\sigma^* - \lambda_\sigma^*\|_{I_n,K}\|\psi\|_{I,\Omega}^2\Bigg) \\
J_2 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}((f_a(\theta^{u_\sigma},a^{u_\sigma}) - f_a(\theta_\sigma^*,a_\sigma^*))\lambda_\sigma^*,\psi)_K dt \le C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|\lambda_\sigma^*\|_{I_n,K}\|\psi\|_{I,\Omega}^2.
\end{aligned}
$$

Similarly,

$$J_3 \le C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|z_\sigma^*\|_{I_n,K}\|\psi\|_{I,\Omega}.$$

and

$$\begin{aligned} J_4 &\leq C\|z^*_\sigma - z^{u_\sigma}\|_{I,\Omega}\|\psi\|^2_{L^2(I,H^2(\Omega))}. \\ J_5 &\leq C\bigg(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} k_n^{\frac{1}{2}}\|[\lambda^*_\sigma]_{n-1}\|_K\|\psi\|_{I,\Omega}\bigg). \end{aligned}$$

Using estimates for J_1 to J_5 in (5.3.49) and Young's inequality with the Young's contant chosen appropriately, we obtain

$$\|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega} \leq C\bigg(\sum_{i=7}^{10}\eta^2_{i,n,K} + frac\mu 2\|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega} + \frac{\mu_1}{2}\|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega}\bigg),(5.3.50)$$

where $\eta_{i,n,K}, i = 7, 8, 9, 10$, are defined by

$$\begin{aligned} \eta^2_7 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|\lambda^*_\sigma\|^2_{I_n,K}, \\ \eta^2_8 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\|z^*_\sigma\|^2_{I_n,K}, \\ \eta^2_9 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}k_n\|[\lambda^*_\sigma]_{n-1}\|^2_K. \end{aligned}$$

Adding (5.3.44) and (5.3.50), we obtain

$$\begin{aligned} \|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega} + \|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega} &\leq C\bigg(\sum_{i=2}^{10}\eta^2_{i,n,K} + \mu\|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega} + \mu_1\|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega} \\ &\quad +\|\theta^*_\sigma - \theta^{u_\sigma}\|^2_{I,\Omega}\bigg). \end{aligned}$$

We has choose Young's constant such that $C\mu < 1$ and $C\mu_1 < 1$, we obtain

$$\|z^*_\sigma - z^{u_\sigma}\|^2_{I,\Omega} + \|\lambda^*_\sigma - \lambda^{u_\sigma}\|^2_{I,\Omega} \leq C\bigg(\sum_{i=2}^{10}\eta^2_{i,n,K} + \|\theta^*_\sigma - \theta^{u_\sigma}\|^2_{I,\Omega}\bigg). \qquad (5.3.51)$$

This completes the proof. □

Lemma 5.3.3. *Let (θ^*, a^*), $(\theta^*_\sigma, a^*_\sigma)$ and $(\theta^{u_\sigma}, a^{u_\sigma})$ be respectively the solutions of* (5.2.2)-

(5.2.5), (5.3.2)-(5.3.5) and (5.3.15)-(5.3.18). Then,

$$\|\theta^{u_\sigma} - \theta_\sigma^*\|^2 + \|a^{u_\sigma} - a_\sigma^*\|^2 \leq C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\bigg(\sum_{j=11}^{14}\eta_{j,n,K}^2 + \eta_{a,n,K}^2\bigg),$$

where

$$\begin{aligned}
\eta_{11,n,K}^2 =& h_K^4\|r_\theta(x,t)\|_{I_n,K}^2, && (5.3.52)\\
r_\theta(x,t) &= \rho c_p\partial_t\theta_\sigma^* - \alpha u_\sigma^* + \rho Lf(\theta_\sigma^*, a_\sigma^*) - \mathcal{K}\Delta\theta_\sigma^* + \rho c_p\frac{[\theta_\sigma^*]_n}{k_n}, && (5.3.53)\\
\eta_{12,n,K}^2 &= k_n^2\|\rho Lf(\theta_\sigma^*, a_\sigma^*) - \mathcal{K}\Delta\theta_\sigma^*\|_{I_n,K}^2, && (5.3.54)\\
\eta_{13,n,K}^2 &= h_K^3\|\mathcal{K}[\bigtriangledown\theta_\sigma^*].\eta\|_{L^2(I_n,L^2(\partial K))}^2, \quad \eta_{14,n,K}^2 = k_n\|[\theta_\sigma^*]_n\|_K^2, && (5.3.55)\\
\eta_{a,n,K}^2 &= k_n^2\|f(\theta_\sigma^*, a_\sigma^*)\|_K^2 + k_n\|[a_\sigma^*]_n^-\|_K^2. && (5.3.56)
\end{aligned}$$

Proof: Consider the problem: find $v \in H^1(\Omega)$ such that

$$\begin{aligned}
-\rho c_p\partial_t v - \mathcal{K}\Delta v - \rho LFv &= g_1 \quad \text{in } Q, && (5.3.57)\\
\frac{\partial v}{\partial\eta}\bigg|_{\partial\Omega} &= 0 \quad \text{in } I && (5.3.58)\\
v(T) &= 0 \quad \text{in } \Omega, && (5.3.59)
\end{aligned}$$

where

$$F = \begin{cases} -\dfrac{f(\theta^{u_\sigma}, a^{u_\sigma}) - f(\theta_\sigma^*, a_\sigma^*)}{\theta_\sigma^* - \theta^{u_\sigma}} & \text{whenever } \theta_\sigma^* \neq \theta^{u_\sigma}\\ f_\theta(\theta_\sigma^*, a_\sigma^*) & \theta_\sigma^* = \theta^{u_\sigma}.\end{cases}$$

Moreover, we have [23]:

$$\begin{aligned}
\|v\|_{L^\infty(I;L^2(\Omega))} &\leq C\|g_1\|_{I,\Omega}, \quad \|v\|_{L^2(I;H^1(\Omega))} \leq C\|g_1\|_{I,\Omega}, && (5.3.60)\\
\|v\|_{L^1(I;H^2(\Omega))} &\leq C\|g_1\|_{I,\Omega}, \quad \|\partial_t v\|_{I,\Omega} \leq C\|g_1\|_{I,\Omega}. && (5.3.61)
\end{aligned}$$

Put $g_1 = \theta_\sigma^* - \theta^{u_\sigma}$ in (5.3.57) and consider

$$\begin{aligned}
\|\theta_\sigma^* - \theta^{u_\sigma}\|_{I,\Omega}^2 &= \int_0^T(\theta_\sigma^* - \theta^{u_\sigma}, -\rho c_p\partial_t v - \mathcal{K}\Delta v + \rho LFv)dt\\
&= \sum_{n=1}^{N}\int_{I_n}\bigg((\rho c_p\partial_t(\theta_\sigma^* - \theta^{u_\sigma}), v) + \mathcal{K}(\bigtriangledown(\theta_\sigma^* - \theta^{u_\sigma}), \bigtriangledown v)\\
&\quad - (\rho L(f(\theta^{u_\sigma}, a^{u_\sigma}) - f(\theta_\sigma^*, a_\sigma^*)), v) + \rho c_p(\frac{[\theta]_n}{k_n}, v_n^-)\bigg)dt.
\end{aligned}$$

Adding (5.3.4) to the right hand side of the above equation after replacing v by the interpolant v_I, we obtain

$$\begin{aligned}\|\theta^*_\sigma - \theta^{u_\sigma}\|^2_{I,\Omega} &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\Big((\rho c_p \partial_t \theta^*_\sigma - \alpha u^*_\sigma + \rho L f(\theta^*_\sigma, a^*_\sigma) - \mathcal{K}\Delta\theta^*_\sigma, v - v_I)_K \\ &\quad + \mathcal{K}(\bigtriangledown\theta^*_\sigma.\eta, v)_{L^2(\partial K)} + \rho c_p(\frac{[\theta^*_\sigma]_n}{k_n}, v_n - (v_I)_n)_K\Big)dt\end{aligned}$$

Letting $r_\theta(x,t) = \rho c_p \partial_t\theta^*_\sigma - \alpha u^*_\sigma + \rho L f(\theta^*_\sigma, a^*_\sigma) - \mathcal{K}\Delta\theta^*_\sigma + \rho c_p \dfrac{[\theta^*_\sigma]_n}{k_n}$ and adding, subtracting $(\rho c_p \frac{[\theta^*_\sigma]_n}{k_n}, v - v_I)$ to the right hand side of the above equation, we obtain

$$\begin{aligned}\|\theta^*_\sigma - \theta^{u_\sigma}\|^2_{I,\Omega} &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\Big((r_\theta(x,t), v - v_I)_K + \mathcal{K}(\bigtriangledown\theta^*_\sigma.\eta, v)_{L^2(\partial K)} \\ &\quad - \rho c_p(\frac{[\theta^*_\sigma]_n}{k_n}, (v_I)_n + v - v_I - v_n)_K\Big)dt \\ &= \sum_{i=1}^{3} J_3, \text{ say} \qquad (5.3.62)\end{aligned}$$

where

$$\begin{aligned}J_1 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(r_\theta(x,t), v - v_I)_K dt, \quad J_2 = \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\mathcal{K}(\bigtriangledown\theta^*_\sigma.\eta, v)_{L^2(\partial K)}dt, \\ J_3 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}\rho c_p(\frac{[\theta^*_\sigma]_n}{k_n}, (v_I)_n - v + v_I - v_n)_K dt.\end{aligned}$$

Use Cauchy-Schwarz inequality, (5.2.10), (5.2.11) and (5.3.60)-(5.3.61) to obtain,

$$\begin{aligned}J_1 &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\int_{I_n}(r_\theta(x,t), v - \pi_n v + \pi_n v - \pi_{h,n}v)_K dt, \\ &\leq C\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(h^2_K\|r_\theta(x,t)\|^2_{I_n,K} + k_n\|\rho L f(\theta^*_\sigma, a^*_\sigma) - \mathcal{K}\Delta\theta^*_\sigma\|^2_{I_n,K}\Big)\|\theta^*_\sigma - \theta^{u_\sigma}\|_{I,\Omega}\Big)\end{aligned}$$

Repeating the same steps used in the calculation of the term J_1, we obtain

$$J_2 \;\leq\; C\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Big(h_K^{\frac{3}{2}}\|\mathcal{K}[\bigtriangledown\theta_\sigma^*].\eta\|_{L^2(I_n,L^2(\partial K))}\Big)\|\theta_\sigma^*-\theta^{u_\sigma}\|_{I,\Omega}\Big)$$

Also,

$$J_3 \;\leq\; C\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h} k_n^{\frac{1}{2}}\|[\theta_\sigma^*]_n\|_K\|\theta_\sigma^*-\theta^{u_\sigma}\|_{I,\Omega}\Big)$$

Using the estimates for J_1 to J_3 in (5.3.62) and Young's inequality with Young's constant chosen appropriately, we obtain

$$\|\theta_\sigma^*-\theta^{u_\sigma}\|_{I,\Omega}^2 \;\leq\; C\Big(\sum_{i=11}^{14}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{i,n,K}^2+\mu_3\|\theta_\sigma^*-\theta^{u_\sigma}\|_{I,\Omega}^2\Big), \qquad (5.3.63)$$

where $\eta_i, i=11,\cdots,14$ are defined by (5.3.52)-(5.3.55). Choosing Young's constant in (5.3.63) such that $C\mu_3<1$, we have

$$\|\theta_\sigma^*-\theta^{u_\sigma}\|_{I,\Omega}^2 \;\leq\; C\Big(\sum_{i=11}^{14}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{i,n,K}^2\Big), \qquad (5.3.64)$$

Now we proceed to estimate $\|a^{u_\sigma}-a_\sigma^*\|$.

Consider the problem: given $g\in L^2(\Omega)$, find w such that

$$\begin{aligned} -\partial_t w &= F_1 w + g_2 \quad \text{in } Q, && (5.3.65)\\ w(T) &= 0, && (5.3.66)\end{aligned}$$

where

$$F_1=\begin{cases}\dfrac{f(\theta_\sigma^*,a_\sigma^*)-f(\theta^{u_\sigma},a^{u_\sigma})}{a_\sigma^*-a^{u_\sigma}} & \text{whenever } a^{u_\sigma}\neq a_\sigma^*\\ f_a(\theta_\sigma^*,a_\sigma^*) & a^{u_\sigma}=a_\sigma^*.\end{cases}$$

Moreover, we have [23]:

$$\begin{aligned} \|w\|_{L^\infty(I;L^2(\Omega))} &\leq C\|g_2\|_{I,\Omega}, && (5.3.67)\\ \|\partial_t w\|_{I,\Omega} &\leq C\|g_2\|_{I,\Omega}. && (5.3.68)\end{aligned}$$

Substitute $g_2 = a^*_\sigma - a^{u_\sigma}$ in (5.3.65), use Cauchy-Schwarz's inequality, Young's inequality with Young's constant chosen appropriately, (5.2.12) and (5.3.67)-(5.3.68) to obtain

$$\begin{aligned}
\|a^*_\sigma - a^{u_\sigma}\|^2_{I,\Omega} &= \int_0^T (a^*_\sigma - a^{u_\sigma}, -\partial_t w - F_1 w)dt \\
&= \sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \int_{I_n} \Big((\partial_t((a^*_\sigma - a^{u_\sigma}), w)_K - (F_1(a^*_\sigma - a^{u_\sigma}), w)_K + (\frac{[a^*_\sigma]_n}{k_n}, w_n) \Big) dt \\
&= \sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \int_{I_n} \Big((\partial_t a^*_\sigma - f(\theta^*_\sigma, a^*_\sigma) + \frac{[a^*_\sigma]_n}{k_n}, w - \pi_n w)_K \\
&\quad + (\frac{[a^*_\sigma]_n}{k_n}, (\pi_n w)_n - w + \pi_n w - w_n) \Big) dt \\
&\leq C \sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \Big(k_n^2 \|f(\theta^*_\sigma, a^*_\sigma)\| + k_n \|[a^*_\sigma]^-_n\|^2_K \Big) + \mu_4 \|a^{u_\sigma} - a^*_\sigma\|^2_{I,\Omega}, \\
&\leq C \Big(\sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \eta^2_{a,n,K} + \mu_4 \|a^{u_\sigma} - a^*_\sigma\|^2_{I,\Omega} \Big).
\end{aligned}$$

Choose Young's constant such that $C\mu_4 < 1$ to obtain

$$\|a^{u_\sigma} - a^*_\sigma\|^2 \leq C \sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \eta^2_{a,n,K}. \tag{5.3.69}$$

Adding (5.3.64) and (5.3.69), we obtain the required result. This completes the proof. □

Theorem 5.3.1. *Let* (θ^*, a^*, u^*), $(\theta^*_\sigma, a^*_\sigma, u^*_\sigma)$ *and* $(\theta^{u_\sigma}, a^{u_\sigma})$ *be the solutions to* (5.2.1)-(5.2.5), (5.4.31)-(5.4.35) *and* (5.3.15)-(5.3.18) *with* (z, λ), $(z_\sigma, \lambda_\sigma)$ *and* $(z^{u_\sigma}, \lambda^{u_\sigma})$ *as the corresponding adjoint solutions. Then, we have*

$$\begin{aligned}
\|u^* - u^*_\sigma\|^2_{L^2(I)} + \|\theta^* - \theta^*_\sigma\|^2_{I,\Omega} + \|z^* - z^*_\sigma\|^2_{I,\Omega} &+ \|a^* - a^*_\sigma\|^2_{I,\Omega} + \|\lambda^* - \lambda^*_\sigma\|_{I,\Omega} \\
&\leq C \sum_{n=1}^N \sum_{K\in\mathcal{T}_h} \Big(\max_{t\in I_n} \|\alpha\|_K \sum_{j=1}^{14} \eta^2_{j,n,K} + \eta^2_{a,n,K} \Big),
\end{aligned}$$

where $\eta_{j,n,K}$*'s and* $\eta_{a,n,K}$ *are as defined in Lemma 5.3.1, 5.3.2 and 5.3.3.*

Proof: From triangle inequality, we have

$$\|\theta^* - \theta^*_\sigma\|^2_{I,\Omega} \leq 2\|\theta^* - \theta^{u_\sigma}\|^2_{I,\Omega} + 2\|\theta^{u_\sigma} - \theta^*_\sigma\|^2_{I,\Omega}.$$

Subtracting (5.3.17) from (5.2.4) for $(\theta, a, u) = (\theta^*, a^*, u^*)$ ((θ^*, a^*, u^*) satisfies (2.1.6)), we obtain

$$\rho c_p(\partial_t(\theta^* - \theta^{u_\sigma}), v) + \mathcal{K}(\bigtriangledown(\theta^* - \theta^{u_\sigma}), \bigtriangledown v) \;\; = \;\; -\rho L(f(\theta^*, a^*) - f(\theta^{u^*_\sigma}, a^{u_\sigma}), v) + (\alpha(u^* - u^*_\sigma), v)$$

Putting $v = \theta - \theta^{u_\sigma}$, integrating from 0 to t, using Remark (2.2.1), we obtain

$$\|\theta^*(t) - \theta^{u_\sigma}(t)\|^2_\Omega \leq C\bigg(\|\theta^* - \theta^{u_\sigma}\|^2_{I,\Omega} + \|a^* - a^{u_\sigma}\|^2_{I,\Omega} + \|u^* - u^*_\sigma\|^2_{L^2(I)}\bigg). \qquad (5.3.70)$$

Using part (b) of Lemma 2.2.1, we have

$$\|a^* - a^{u_\sigma}\|^2_{I,\Omega} \;\; \leq \;\; C\|\theta^* - \theta^{u_\sigma}\|^2_{I,\Omega}. \qquad (5.3.71)$$

Using (5.3.71) in (5.3.70) and Gronwall's Lemma, we obtain

$$\|\theta^* - \theta^{u_\sigma}\|^2_{I,\Omega} \;\; \leq \;\; C\|u^* - u^*_\sigma\|^2_{L^2(I)}.$$

Use same arguments used to prove part (b) in Lemma 2.2.1, that is, subtract (3.2.13) from (5.3.44), integrating from t to T and using Gronwall's lemma to obtain

$$\|\lambda^* - \lambda^{u_\sigma}\|^2_{I,\Omega} \leq C\|z^* - z^{u_\sigma}\|^2_{I,\Omega}. \qquad (5.3.72)$$

Similarly subtracting (3.2.15) from (5.3.21), integrating from t to T, using (5.3.72) and Gronwall's Lemma, we have

$$\|z^* - z^{u_\sigma}\|^2_{I,\Omega} \;\; \leq \;\; C\|\theta^* - \theta^{u_\sigma}\|^2_{I,\Omega}.$$

Using all the above estimates, that is, (5.3.70)-(5.3.73), Lemma 5.3.1, 5.3.2 and 5.3.3, we obtain the required result. This completes the proof. □

Remark 5.3.1. *The* a posteriori *error estimates obtained in Theorem 5.3.1 can be divided into errors due to space, time and control discretiztion, that is,*

$$\begin{aligned}\|u^* - u^*_\sigma\|^2_{L^2(I)} + \|\theta^* - \theta^*_\sigma\|^2_{I,\Omega} + \|z^* - z^*_\sigma\|^2_{I,\Omega} \;\; &+ \;\; \|a^* - a^*_\sigma\|^2_{I,\Omega} + \|\lambda^* - \lambda^*_\sigma\|_{I,\Omega} \\ &\leq \;\; \eta_h + \eta_k + \eta_d,\end{aligned}$$

where η_h, η_k and η_d are the errors occurred due to space, time and control discretizations and are given by

$$\begin{aligned}\eta_k &= C\Big(\sum_{i=3,6,7,10,12,14}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{i,n,K}^2+\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{a,n,K}^2\Big),\\ \eta_h &= C\sum_{i=2,4,5,8,9,11,13}\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{i,n,K}^2,\\ \eta_d &= C\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\eta_{1,n,K}^2.\end{aligned}$$

5.4 Dual Weighted Residual Method

In Section 5.2, a residual type estimator has been developed for the purpose of deriving *a posteriori* error estimators. To apply residual method, (5.2.1)-(5.2.5) has been discretized first in space, then in time and control for the ease of computation. In this section, DWR type estimators have been developed and discretization of the system has been done first in time, then in space and control.

We recall the following from Chapter 3 for the continuity of reading:

The system (5.2.1) - (5.2.5) has atleast one global solution, which is characterized by the saddle point $(\theta^*, a^*, z^*, \lambda^*, u^*) \in X \times Y \times X \times Y \times U_{ad}$ of the Lagrangian functional defined by

$$\begin{aligned}\tilde{\mathcal{L}}(\theta, a, z, \lambda, u) &= J(\theta, a, u) - \Big((\partial_t a, \lambda)_{I,\Omega} - (f(\theta, a), \lambda)_{I,\Omega}\Big) - \Big(\rho c_p(\partial_t \theta, z)_{I,\Omega}\\ &\quad + \mathcal{K}(\bigtriangledown\theta, \bigtriangledown z)_{I,\Omega} + \rho L(a_t, z)_{I,\Omega} - (\alpha u, z)_{I,\Omega}\Big)\end{aligned}$$

The adjoint system of (5.2.1)-(5.2.5) obtained from KKT conditions is defined by:

Find $(z^*, \lambda^*) \in X \times Y$ such that

$$\begin{aligned}-(\psi, \partial_t\lambda^*) &= -(\psi, f_a(\theta^*, a^*)(\rho L z^* - \lambda^*)), && (5.4.1)\\ \lambda^*(T) &= \beta_1(a^*(T) - a_d), && (5.4.2)\\ -\rho c_p(\phi, \partial_t z^*) + \mathcal{K}(\bigtriangledown\phi, \bigtriangledown z^*) &= -(\phi, f_\theta(\theta^*, a^*)(\rho L z^* - \lambda^*)) + \beta_2(\phi, [\theta^* - \theta_m]_+), && (5.4.3)\\ z^*(T) &= 0, && (5.4.4)\end{aligned}$$

for all $(\psi, \phi) \in H \times V$. Moreover, z^* satisfies the following variational inequality

$$\left(\beta_3 u^* + \int_\Omega \alpha z^* dx,\ p - u^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \tag{5.4.5}$$

Discretizations

In this section, a temporal discretization is done using a discontinuous Galerkin finite element method with piecewise constant approximation and then a space discretization is done using continuous Galerkin finite element method using piecewise linear polynomials, that is a space-time discretization is done using dG(0)cG(1). The control is being discretized using piecewise constants in each discrete interval $I_n, n = 1, 2, \cdots, N$.

Time Discretization

In order to discretize (5.2.1)-(5.2.5) in time, we consider the following partition of I:

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

Set $I_1 = [t_0, t_1],\ I_n = (t_{n-1}, t_n]$, $k_n = t_n - t_{n-1}$, for $n = 2, ..., N$ and $k = \max\limits_{1\leq n\leq N} k_n$. We define the spaces

$$X_k^q \;=\; \{\phi : I \to H^1(\Omega);\ \ \phi|_{I_n} = \sum_{j=0}^q \psi_j t^j, \psi_j \in H^1(\Omega)\},\ \ q \in \mathbb{N}, \tag{5.4.6}$$

$$Y_k^q \;=\; \{\phi : I \to L^2(\Omega);\ \ \phi|_{I_n} = \sum_{j=0}^q \psi_j t^j, \psi_j \in L^2(\Omega)\},\ \ q \in \mathbb{N}. \tag{5.4.7}$$

For a function v in X_k^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) discretization of (5.2.1)-(5.2.5) reads as:

$$\min_{u_k \in U_{ad}} J(\theta_k, a_k, u_k) \qquad \text{subject to} \tag{5.4.8}$$

$$\sum_{n=1}^N (\partial_t a_k, w)_{I_n,\Omega} \;+\; \sum_{n=1}^{N-1} ([a_k]_n, w_n^+) + (a_{k,0}^+, w_0^+) = (f(\theta_k, a_k), w)_{I,\Omega}, \tag{5.4.9}$$

$$a_k(0) \;=\; 0, \tag{5.4.10}$$

$$
\begin{aligned}
\rho c_p \sum_{n=1}^{N} (\partial_t \theta_k, v)_{I_n,\Omega} &+ \mathcal{K}(\bigtriangledown \theta_k, \bigtriangledown v)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_k]_n, v_n^+) + \rho c_p (\theta_{k,0}^+, v_0^+) \\
&= -\rho L (f(\theta_k, a_k), v)_{I,\Omega} + (\alpha u_k, v)_{I,\Omega} + \rho c_p (\theta_0, v_0^+), \qquad (5.4.11) \\
\theta_k(0) &= \theta_{h,0} \qquad (5.4.12)
\end{aligned}
$$

for all $(v, w) \in X_k^q \times X_k^q$.

The solution of (5.4.8)-(5.4.12) is characterized by the saddle point $(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u) \in X_k^q \times Y_k^q \times X_k^q \times Y_k^q \times U_{ad}$ of the Lagrangian functional given by

$$
\begin{aligned}
\mathcal{L}(\theta_k, a_k, z_k, \lambda_k, u_k) &= J(\theta_k, a_k, u_k) - \Bigg(\sum_{n=1}^{N} (\partial_t a_k, \lambda_k)_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a_k]_n, \lambda_{k,n}^+) \\
&+ (a_{k,0}^+, \lambda_{k,0}^+) - (f(\theta_k, a_k), \lambda_k)_{I,\Omega} \Bigg) - \Bigg(\sum_{n=1}^{N} \rho c_p (\partial_t \theta_k, z_k)_{I_n,\Omega} \\
&+ \mathcal{K}(\bigtriangledown \theta_k, \bigtriangledown z_k)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta_k]_n, z_{k,n}^+) + \rho c_p (\theta_{k,0}^+, z_{k,0}^+) \\
&+ \rho L (f(\theta_k, a_k), z_k)_{I,\Omega} - (\alpha u_k, z_k)_{I,\Omega} - \rho c_p (\theta_0, z_{k,0}^+) \Bigg).
\end{aligned}
$$

The adjoint system of (5.4.8)-(5.4.12) obtained from KKT conditions is defined by: Find $(z_k^*, \lambda_k^*) \in X_k^q \times Y_k^q$ such that

$$
\begin{aligned}
-\sum_{n=1}^{N} (\psi, \partial_t \lambda_k^*)_{I_n,\Omega} - \sum_{n=1}^{N-1} (\psi_n, [\lambda_k^*]_n) &= -(\psi, f_a(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k^*))_{I,\Omega}, \qquad (5.4.13) \\
\lambda_k^*(T) &= \beta_1 (a_k^*(T) - a_d), \qquad (5.4.14) \\
-\rho c_p \sum_{n=1}^{N} (\phi, \partial_t z_k^*)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \phi, \bigtriangledown z_k^*)_{I,\Omega} &- \rho c_p \sum_{n=1}^{N-1} (\phi_n, [z_k^*]_n) = -(\phi, f_\theta(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k^*))_{I,\Omega} \\
&+ \beta_2 (\phi, [\theta_k^* - \theta_m]_+)_{I,\Omega}, \qquad (5.4.15) \\
z_k^*(T) &= 0, \qquad (5.4.16)
\end{aligned}
$$

for all $(\psi, \phi) \in X_k^q \times Y_k^q$. Moreover, z_k^* satisfies the following variational inequality

$$
\left(\beta_3 u_k^* + \int_\Omega \alpha z_k dx, \ p - u_k^* \right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \qquad (5.4.17)
$$

Space Discretization

We describe a space discretization for (5.4.8)-(5.4.12) using a continuous Galerkin finite element method with piecewise linear approximations. Let $\mathcal{T}_h$ be an admissible regular triangulation of $\bar{\Omega}$ into quadrilaterals K as defined in Chapter 1. Let the discretization parameter h be defined as $h = \max_{K\in\mathcal{T}_h} h_K$, where h_K is the diameter of the quadrilateral K. Let the finite element space $V_h \subset V$ be defined as $V_h = \{v \in C^0(\bar{\Omega}) : \ v(t)|_K \in Q_1(K) \ \ \forall K \in \mathcal{T}_h\}$ and

$$X_{kh}^q = \{\phi : I \to V_h; \ \ \phi|_{I_n} = \sum_{j=0}^{q} \psi_j t^j, \psi_j \in V_h\}, \ \ q \in \mathbb{N}. \tag{5.4.18}$$

Here $Q_1(K)$ denotes the set of all polynomials of degree ≤ 1 in each variable x and y. Then the dG(q)cG(1) discretization of (5.4.8)-(5.4.12) reads as:

$$\begin{aligned}
\min_{u_{kh}\in U_{ad}} J(\theta_{kh}, a_{kh}, u_{kh}) \qquad & \text{subject to} & (5.4.19)\\
\sum_{n=1}^{N}(\partial_t a_{kh}, w)_{I_n,\Omega} \ \ + \ \ & \sum_{n=1}^{N-1}([a_{kh}]_n, w_n^+) + (a_{kh,0}^+, w_0^+) = (f(\theta_{kh}, a_{kh}), w)_{I,\Omega}, & (5.4.20)\\
a_{kh}(0) \ \ = \ \ & 0, & (5.4.21)
\end{aligned}$$

$$\begin{aligned}
\rho c_p \sum_{n=1}^{N}(\partial_t \theta_{kh}, v)_{I_n,\Omega} \ \ + \ \ & \mathcal{K}(\bigtriangledown\theta_{kh}, \bigtriangledown v)_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1}([\theta_{kh}]_n, v_n^+) + \rho c_p(\theta_{kh,0}^+, v_0^+)\\
= \ \ & -\rho L(f(\theta_{kh}, a_{kh}), v)_{I,\Omega} + (\alpha u_{kh}, v)_{I,\Omega} + \rho c_p(\theta_0, v_0^+), & (5.4.22)\\
\theta_{kh}(0) \ \ = \ \ & \theta_{h,0}, & (5.4.23)
\end{aligned}$$

for all $(v, w) \in X_{kh}^q \times X_{kh}^q$.

The solution of the (5.4.19)-(5.4.23) is characterized by the saddle point $(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) \in X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times U_{ad}$ of the Lagrangian functional given by

$$\begin{aligned}
\mathcal{L}(\theta_{kh}, a_{kh}, z_{kh}, \lambda_{kh}, u_{kh}) \ \ = \ \ & J(\theta_{kh}, a_{kh}, u_{kh}) - \Bigg(\sum_{n=1}^{N}(\partial_t a_{kh}, \lambda_{kh})_{I_n,\Omega} + \sum_{n=1}^{N-1}([a_{kh}]_n, \lambda_{kh,n}^+)\\
& + (a_{kh,0}^+, \lambda_{kh,0}^+) - (f(\theta_{kh}, a_{kh}), \lambda_{kh})_{I,\Omega}\Bigg) - \Bigg(\sum_{n=1}^{N}\rho c_p(\partial_t\theta_{kh}, z_{kh})_{I_n,\Omega}\\
& + \mathcal{K}(\bigtriangledown\theta_{kh}, \bigtriangledown z_{kh})_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1}([\theta_{kh}]_n, z_{kh,n}^+) + \rho c_p(\theta_{kh,0}^+, z_{kh,0}^+)\\
& + \rho L(f(\theta_{kh}, a_{kh}, z_{kh}))_{I,\Omega} - (\alpha u_{kh}, z_{kh})_{I,\Omega} - \rho c_p(\theta_0, z_{kh,0}^+)\Bigg)(5.4.24)
\end{aligned}$$

The adjoint system of (5.4.19)-(5.4.23) obtained using KKT conditions is defined by: Find $(z_{kh}^*, \lambda_{kh}^*) \in X_{kh}^q \times X_{kh}^q$ such that

$$\begin{aligned} -\sum_{n=1}^{N}(\psi, \partial_t \lambda_{kh}^*)_{I_n,\Omega} - \sum_{n=1}^{N-1}(\psi_n, [\lambda_{kh}^*]_n) &= -(\psi, f_a(\theta_{kh}^*, a_{kh}^*)(\rho L z_{kh}^* - \lambda_{kh}^*))_{I,\Omega}, && (5.4.25) \\ \lambda_{kh}^*(T) &= \beta_1(a_{kh}^*(T) - a_d), && (5.4.26) \end{aligned}$$

$$\begin{aligned} -\rho c_p \sum_{n=1}^{N}(\phi, \partial_t z_{kh}^*)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \phi, \bigtriangledown z_{kh}^*)_{I,\Omega} & \\ -\rho c_p \sum_{n=1}^{N-1}(\phi_n, [z_{kh}^*]_n) &= -(\phi, f_\theta(\theta_{kh}^*, a_{kh}^*)(\rho L z_{kh}^* - \lambda_{kh}^*))_{I,\Omega} && (5.4.27) \\ & \quad + \beta_2(\phi, [\theta_{kh}^* - \theta_m]_+)_{I,\Omega}, && (5.4.28) \\ z_{kh}^*(T) &= 0, && (5.4.29) \end{aligned}$$

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

$$\left(\beta_3 u_{kh}^* + \int_\Omega \alpha z_{kh}^* dx,\ p - u_{kh}^*\right)_{L^2(I)} \geq 0 \quad \forall p \in U_{ad}. \qquad (5.4.30)$$

Complete discretization

In order to completely discretize the problem (5.2.1)-(5.2.5) we choose discontinuous Galerkin piecewise constant approximation of the control variable. Let U_d be the finite dimensional subspace of U defined by

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

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:

$$\begin{aligned} \min_{u_\sigma \in U_{d,ad}} J(\theta_\sigma, a_\sigma, u_\sigma) & \qquad \text{subject to} && (5.4.31) \\ \sum_{n=1}^{N}(\partial_t a_\sigma, w)_{I_n,\Omega} + \sum_{n=1}^{N-1}([a_\sigma]_n, w_n^+) &+ (a_{\sigma,0}^+, w_0^+) = (f(\theta_\sigma, a_\sigma), w)_{I,\Omega}, && (5.4.32) \\ a_\sigma(0) &= 0, && (5.4.33) \end{aligned}$$

$$
\begin{aligned}
\rho c_p \sum_{n=1}^{N} (\partial_t \theta_\sigma, v)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \theta_\sigma, \bigtriangledown v)_{I,\Omega} &+ \rho c_p \sum_{n=1}^{N-1} ([\theta_\sigma]_n, v_n^+) + \rho c_p(\theta_{\sigma,0}^+, v_0^+) \\
&= -\rho L(f(\theta_\sigma, a_\sigma), v)_{I,\Omega} + (\alpha u_\sigma, v)_{I,\Omega}, \\
&\quad + \rho c_p(\theta_0, v_0^+), \qquad (5.4.34) \\
\theta_\sigma(0) &= \theta_0, \qquad (5.4.35)
\end{aligned}
$$

for all $(v, w) \in X_{kh}^q \times X_{kh}^q$.

The solution of the (5.4.31)-(5.4.35) is characterized by the saddle point $(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*) \in X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times U_{d,ad}$ of the Lagrangian functional given by

$$
\begin{aligned}
\mathcal{L}(\theta_\sigma, a_\sigma, z_\sigma, \lambda_\sigma, u_\sigma) &= J(\theta_\sigma, a_\sigma, u_\sigma) - \Bigg(\sum_{n=1}^{N} (\partial_t a_\sigma, \lambda_\sigma)_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a_\sigma]_n, \lambda_{\sigma,n}^+) + (a_{\sigma,0}^+, \lambda_{\sigma,0}^+) \\
&- (f(\theta_\sigma, a_\sigma), \lambda_\sigma)_{I,\Omega} \Bigg) - \Bigg(\rho c_p \sum_{n=1}^{N} (\partial_t \theta_\sigma, z_\sigma)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \theta_\sigma, \bigtriangledown z_\sigma)_{I,\Omega} \\
&+ \rho c_p \sum_{n=1}^{N-1} ([\theta_\sigma]_n, z_{\sigma,n}^+) + \rho c_p(\theta_{\sigma,0}^+, z_{\sigma,0}^+) + \rho L(f(\theta_\sigma, a_\sigma), z_\sigma)_{I,\Omega} - (\alpha u_\sigma, z_\sigma)_{I,\Omega} \\
&- \rho c_p(\theta_0, z_{\sigma,0}^+) \Bigg)
\end{aligned}
$$

The adjoint system of (5.4.31)-(5.4.35) obtained from KKT conditions is defined by:
Find $(z_\sigma^*, \lambda_\sigma^*) \in X_{kh}^q \times X_{kh}^q$ such that

$$
\begin{aligned}
-\sum_{n=1}^{N} (\psi, \partial_t \lambda_\sigma^*)_{I_n,\Omega} - \sum_{n=1}^{N-1} (\psi_n, [\lambda_\sigma^*]_n) &= -(\psi, f_a(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega}, \qquad (5.4.36) \\
\lambda_{\sigma,N}^* &= \beta_1(a_\sigma^*(T) - a_d), \qquad (5.4.37) \\
-\rho c_p \sum_{n=1}^{N} (\phi, \partial_t z_\sigma^*)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown \phi, \bigtriangledown z_\sigma^*)_{I,\Omega} &- \rho c_p \sum_{n=1}^{N-1} (\phi_n, [z_\sigma^*]_n) = -(\phi, f_\theta(\theta_\sigma^*, a_\sigma^*)(\rho L z_\sigma^* - \lambda_\sigma^*))_{I,\Omega} \\
&+ \beta_2(\phi, [\theta_\sigma^* - \theta_m]_+)_{I,\Omega}, \qquad (5.4.38) \\
z_{\sigma,N}^* &= 0, \qquad (5.4.39)
\end{aligned}
$$

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

$$
\left(\beta_3 u_\sigma^* + \int_\Omega \alpha z_\sigma^* dx, p - u_\sigma^* \right)_{L^2(I)} \geq 0 \quad \forall p \in U_{d,ad}. \qquad (5.4.40)
$$

In the next section, *a posteriori* error estimates have been calculated using DWR method.

To arrive at an estimate for $|J(\theta^*,a^*,u^*) - J(\theta^*_\sigma,a^*_\sigma,u^*_\sigma)|$, we follow the approach in [61]. We will split the discretization error occurred by different discretization such as follows:

$$\begin{aligned} J(\theta^*_\sigma,a^*_\sigma,u^*_\sigma) &- J(\theta^*,a^*,u^*) \\ &= \Big(J(\theta^*_\sigma,a^*_\sigma,u^*_\sigma) - J(\theta^*_{kh},a^*_{kh},u^*_{kh})\Big) + \Big(J(\theta^*_{kh},a^*_{kh},u^*_{kh}) - J(\theta^*_k,a^*_k,u^*_k)\Big) \\ &\quad + \Big(J(\theta^*_k,a^*_k,u^*_k) - J(\theta^*,a^*,u^*)\Big), \qquad (5.4.41) \end{aligned}$$

where (θ^*,a^*,u^*), (θ^*_k,a^*_k,u^*_k), $(\theta^*_{kh},a^*_{kh},u^*_{kh})$ and $(\theta^*_\sigma,a^*_\sigma,u^*_\sigma)$ are solutions of (5.2.1)-(5.2.5), (5.4.8)-(5.4.12), (5.4.19)-(5.4.23) and (5.4.31)-(5.4.35), respectively. In Lemma 5.4.1, we first estimate the terms on the right hand side of (5.4.41) and then Theorem 5.4.1 we present the *a posteriori* error in terms of local estimators.

Remark 5.4.1. *Since the solution of the problem* (5.2.1)-(5.2.5) *will also be the stationary point for the Lagrangian* $\mathcal{L}$, *under the regularity assumption that* $(\theta,a) \in H^1(I,H^2(\Omega)) \times H^1(I,H^2(\Omega))$ *and* $(z,\lambda) \in H^1(I,H^2(\Omega)) \times H^1(I,H^2(\Omega))$, *we have*

$$\mathcal{L}(\theta,a,z,\lambda,u) = \tilde{\mathcal{L}}(\theta,a,z,\lambda,u).$$

Remark 5.4.2. *The Lagrangian functional* $\mathcal{L}$ *is two times differentiable.*

Lemma 5.4.1. *The Lagrangian functional* $\mathcal{L}(\cdot,\cdot,\cdot,\cdot,\cdot)$ *has stationary points* $(\theta^*,a^*,z^*,\lambda^*,u^*) \in X\times Y\times X\times Y\times U_{ad}$, $(\theta^*_k,a^*_k,z^*_k,\lambda^*_k,u^*_k) \in X^q_k\times Y^q_k\times X^q_k\times Y^q_k\times U_{ad}$, $(\theta^*_{kh},a^*_{kh},z^*_{kh},\lambda^*_{kh},u^*_{kh}) \in X^q_{kh}\times X^q_{kh}\times X^q_{kh}\times X^q_{kh}\times U_{ad}$ *and* $(\theta^*_\sigma,a^*_\sigma,z^*_\sigma,\lambda^*_\sigma,u^*_\sigma) \in X^q_{kh}\times X^q_{kh}\times X^q_{kh}\times X^q_{kh}\times U_{d,ad}$ *on different level of discretization, that is,*

$$\begin{aligned} &\forall(\theta,a,z,\lambda,u) \in X\times Y\times X\times Y\times U_{ad}, \\ &\qquad \mathcal{L}'(\theta^*,a^*,z^*,\lambda^*,u^*)(\theta,a,z,\lambda,u) = 0, \qquad (5.4.42) \\ &\forall(\theta_k,a_k,z_k,\lambda_k,u_k) \in X^q_k\times Y^q_k\times X^q_k\times Y^q_k\times U_{ad}, \\ &\qquad \mathcal{L}'(\theta^*_k,a^*_k,z^*_k,\lambda^*_k,u^*_k)(\theta_k,a_k,z_k,\lambda_k,u_k) = 0, \qquad (5.4.43) \end{aligned}$$

$$\forall(\theta_{kh}, a_{kh}, z_{kh}, \lambda_{kh}, u_{kh}) \in X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times U_{ad},$$

$$\mathcal{L}'(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*)(\theta_{kh}, a_{kh}, z_{kh}, \lambda_{kh}, u_{kh}) = 0, \tag{5.4.44}$$

$$\forall(\theta_\sigma, a_\sigma, z_\sigma, \lambda_\sigma, u_\sigma) \in X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times X_{kh}^q \times U_{d,ad},$$

$$\mathcal{L}'(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(\theta_\sigma - \theta_\sigma^*, a_\sigma - a_\sigma^*, z_\sigma - z_\sigma^*, \lambda_\sigma - \lambda_\sigma^*, u_\sigma - u_\sigma^*) \geq 0. \tag{5.4.45}$$

Then, the following error representation holds true:

$$\begin{aligned} J(\theta_k^*, a_k^*, u_k^*) &- J(\theta^*, a^*, u^*) && (5.4.46)\\ &\leq |\mathcal{L}'(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*)(\theta^* - \theta_k, a^* - a_k, z^* - z_k, \lambda^* - \lambda_k, u^* - u_k)| + |R_k|, \\ J(\theta_{kh}^*, a_{kh}^*, u_{kh}^*) &- J(\theta_k^*, a_k^*, u_k^*) && (5.4.47)\\ &\leq |\mathcal{L}'(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*)(\theta_k^* - \theta_{kh}, a_k^* - a_{kh}, z_k^* - z_{kh}, \lambda_k^* - \lambda_{kh}, u_k^* - u_{kh})| + |R_h|, \\ J(\theta_\sigma^*, a_\sigma^*, u_\sigma^*) &- J(\theta_{kh}^*, a_{kh}^*, u_{kh}^*) && (5.4.48)\\ &\leq |\mathcal{L}'(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(\theta_{kh}^* - \theta_\sigma, a_{kh}^* - a_\sigma, z_{kh}^* - z_\sigma, \lambda_{kh}^* - \lambda_\sigma, u_{kh}^* - u_\sigma)| + |R_d|, \end{aligned}$$

where the remainders R_k, R_h and R_d are quadratic in $(e_k^\theta, e_k^a, e_k^z, e_k^\lambda, e_k^u)$, $(e_{kh}^\theta, e_{kh}^a, e_{kh}^z, e_{kh}^\lambda, e_{kh}^u)$ and $(e_\sigma^\theta, e_\sigma^a, e_\sigma^z, e_\sigma^\lambda, e_\sigma^u)$ respectively, are defined by,

$$\begin{aligned} R_k &= \frac{1}{2}\mathcal{L}''(\theta_k^* + se_k^\theta, a_k^* + se_k^a, z_k^* + se_k^z, \lambda_k^* + se_k^\lambda, u_k^* + se_k^u)((e_k^\theta, e_k^a, e_k^z, e_k^\lambda, e_k^u), \\ &\quad (e_k^\theta, e_k^a, e_k^z, e_k^\lambda, e_k^u)), && (5.4.49)\\ R_h &= \frac{1}{2}\mathcal{L}''(\theta_{kh}^* + se_{kh}^\theta, a_{kh}^* + se_{kh}^a, z_{kh}^* + se_{kh}^z, \lambda_{kh}^* + se_{kh}^\lambda, u_{kh}^* + se_{kh}^u)((e_{kh}^\theta, e_{kh}^a, e_{kh}^z, e_{kh}^\lambda, e_{kh}^u), \\ &\quad (e_{kh}^\theta, e_{kh}^a, e_{kh}^z, e_{kh}^\lambda, e_{kh}^u)), && (5.4.50)\\ R_d &= \frac{1}{2}\mathcal{L}''(\theta_\sigma^* + se_\sigma^\theta, a_\sigma^* + se_\sigma^a, z_\sigma^* + se_\sigma^z, \lambda_\sigma^* + se_\sigma^\lambda, u_\sigma^* + se_\sigma^u)((e_\sigma^\theta, e_\sigma^a, e_\sigma^z, e_\sigma^\lambda, e_\sigma^u), \\ &\quad (e_\sigma^\theta, e_\sigma^a, e_\sigma^z, e_\sigma^\lambda, e_\sigma^u)), && (5.4.51) \end{aligned}$$

and $(e_k^\theta = \theta^* - \theta_k^*,\ e_k^a = a^* - a_k^*,\ e_k^z = z^* - z_k^*,\ e_k^\lambda = \lambda^* - \lambda_k^*,\ e_k^u = u^* - u_k^*)$, $(e_{kh}^\theta = \theta_k^* - \theta_{kh}^*,\ e_{kh}^a = a_k^* - a_{kh}^*,\ e_{kh}^z = z_k^* - z_{kh}^*,\ e_{kh}^\lambda = \lambda_k^* - \lambda_{kh}^*,\ e_{kh}^u = u_k^* - u_{kh}^*)$ and $(e_\sigma^\theta = \theta_{kh}^* - \theta_\sigma^*,\ e_\sigma^a = a_{kh}^* - a_\sigma^*, e_\sigma^z = z_{kh}^* - z_\sigma^*, e_\sigma^\lambda = \lambda_{kh}^* - \lambda_\sigma^*,\ e_\sigma^u = u_{kh}^* - u_\sigma^*)$.

Proof: Using Taylor series expansion, we have

$$\begin{aligned} \mathcal{L}(\theta^*, a^*, z^*, \lambda^*, u^*) &- \mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) \\ &= \mathcal{L}'(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*)(e_k^\theta, e_k^a, e_k^z, e_k^\lambda, e_k^u) + R_k, \end{aligned}$$

$$\begin{aligned}
\mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) &- \mathcal{L}(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) \\
&= \mathcal{L}'(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*)(e_{kh}^\theta, e_{kh}^a, e_{kh}^z, e_{kh}^\lambda, e_{kh}^u) + R_h, \\
\mathcal{L}(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*) &- \mathcal{L}(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) \\
&= \mathcal{L}'(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(-e_{kh}^\theta, -e_{kh}^a, -e_{kh}^z, -e_{kh}^\lambda, -e_{kh}^u) - R_d,
\end{aligned}$$

where R_k, R_h and R_d are defined by (5.4.49)-(5.4.51). Using (5.4.42)-(5.4.45) in above expressions, after replacing

$$\begin{aligned}
\theta^* - \theta_k^* &= (\theta^* - \theta_k) + (\theta_k - \theta_k^*), a^* - a_k^* = (a^* - a_k) + (a_k - a_k^*), \\
z^* - z_k^* &= (z^* - z_k) + (z_k - z_k^*), \lambda^* - \lambda_k^* = (\lambda^* - \lambda_k) + (\lambda_k - \lambda_k^*), \\
u^* - u_k^* &= (u^* - u_k) + (u_k - u_k^*), \theta_k^* - \theta_{kh}^* = (\theta_k^* - \theta_{kh}) + (\theta_{kh} - \theta_{kh}^*), \\
a_k^* - a_{kh}^* &= (a_k^* - a_{kh}) + (a_{kh} - a_{kh}^*), z_k^* - z_{kh}^* = (z_k^* - z_{kh}) + (z_{kh} - z_{kh}^*), \\
\lambda_k^* - \lambda_{kh}^* &= (\lambda_k^* - \lambda_{kh}) + (\lambda_{kh} - \lambda_{kh}^*), u_k^* - u_{kh}^* = (u_k^* - u_{kh}) + (u_{kh} - u_{kh}^*), \\
\theta_{kh}^* - \theta_\sigma^* &= (\theta_{kh}^* - \theta_\sigma) + (\theta_\sigma - \theta_\sigma^*), a_{kh}^* - a_\sigma^* = (a_{kh}^* - a_\sigma) + (a_\sigma - a_\sigma^*), \\
z_{kh}^* - z_\sigma^* &= (z_{kh}^* - z_\sigma) + (z_\sigma - z_\sigma^*), \lambda_{kh}^* - \lambda_\sigma^* = (\lambda_{kh}^* - \lambda_\sigma) + (\lambda_\sigma - \lambda_\sigma^*), \\
u_{kh}^* - u_\sigma^* &= (u_{kh}^* - u_\sigma) + (u_\sigma - u_\sigma^*),
\end{aligned}$$

respectively, we obtain

$$\begin{aligned}
\mathcal{L}(\theta^*, a^*, z^*, \lambda^*, u^*) &- \mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) \\
&= \mathcal{L}'(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*)(\theta^* - \theta_k, a^* - a_k, z^* - z_k, \lambda^* - \lambda_k, u^* - u_k) + R_k, \\
\mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) &- \mathcal{L}(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) \\
&= \mathcal{L}'(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*)(\theta_k^* - \theta_{kh}, a_k^* - a_{kh}, z_k^* - z_{kh}, \lambda_k^* - \lambda_{kh}, u_k^* - u_{kh}) + R_h, \\
\mathcal{L}(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*) &- \mathcal{L}(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) \\
&\leq \mathcal{L}'(\theta_\sigma^*, a_\sigma^*, z_\sigma^*, \lambda_\sigma^*, u_\sigma^*)(\theta_\sigma - \theta_{kh}^*, a_\sigma - a_{kh}^*, z_\sigma - z_{kh}^*, \lambda_\sigma - \lambda_{kh}^*, u_\sigma - u_{kh}^*) - R_d.
\end{aligned}$$

Since, all the solution pairs are optimal solution of the optimization problem at different levels of discretization, we obtain

$$\begin{aligned}
\mathcal{L}(\theta^*, a^*, z^*, \lambda^*, u^*) - \mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) &= J(\theta^*, a^*, u^*) - J(\theta_k^*, a_k^*, u_k^*), \\
\mathcal{L}(\theta_k^*, a_k^*, z_k^*, \lambda_k^*, u_k^*) - \mathcal{L}(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*, u_{kh}^*) &= J(\theta_k^*, a_k^*, u_k^*) - J(\theta_{kh}^*, a_{kh}^*, u_{kh}^*),
\end{aligned}$$

$$\mathcal{L}(\theta^*_{kh}, a^*_{kh}, z^*_{kh}, \lambda^*_{kh}, u^*_{kh}) - \mathcal{L}(\theta^*_\sigma, a^*_\sigma, z^*_\sigma, \lambda^*_\sigma, u^*_\sigma) \;=\; J(\theta^*_{kh}, a^*_{kh}, u^*_{kh}) - J(\theta^*_\sigma, a^*_\sigma, u^*_\sigma).$$

Therefore, we have the required result. This completes the proof. □

Define the residuals for different level of discretizations as :

$$\begin{aligned}
\rho^\theta(\theta, a, u)(\cdot) &= \mathcal{L}_z(\theta, a, z, \lambda, u)(\cdot), \quad \rho^z(\theta, a, z, \lambda)(\cdot) = \mathcal{L}_\theta(\theta, a, z, \lambda, u)(\cdot),\\
\rho^a(\theta, a)(\cdot) &= \mathcal{L}_\lambda(\theta, a, z, \lambda, u)(\cdot), \quad \rho^\lambda(\theta, a, z, \lambda)(\cdot) = \mathcal{L}_a(\theta, a, z, \lambda, u)(\cdot),\\
\rho^u(z, u)(\cdot) &= \mathcal{L}_u(\theta, a, z, \lambda, u)(\cdot),
\end{aligned}$$

where

$$\begin{aligned}
\rho^\theta(\theta, a, u)(\cdot) &= \rho c_p \sum_{n=1}^{N} (\partial_t\theta, \cdot)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown\theta, \bigtriangledown(\cdot))_{I,\Omega} + \rho c_p \sum_{n=1}^{N-1} ([\theta]_n, (\cdot)_n^+)\\
&\quad + \rho c_p(\theta_0^+, (\cdot)_0^+) + \rho L(f(\theta, a), \cdot)_{I,\Omega} - (\alpha u, \cdot)_{I,\Omega} - \rho c_p(\theta_0, (\cdot)_0^+),\\
\rho^a(\theta, a, u)(\cdot) &= \sum_{n=1}^{N} (\partial_t a, \cdot)_{I_n,\Omega} + \sum_{n=1}^{N-1} ([a]_n, (\cdot)_n^+) + (a_0^+, (\cdot)_0^+) - (f(\theta, a), \cdot)_{I,\Omega},\\
\rho^z(\theta, a, z, \lambda)(\cdot) &= -\rho c_p \sum_{n=1}^{N} (\cdot, \partial_t z)_{I_n,\Omega} + \mathcal{K}(\bigtriangledown(\cdot), \bigtriangledown z)_{I,\Omega} - \rho c_p \sum_{n=1}^{N-1} ((\cdot)_n, [z]_n)\\
&\quad + (\cdot, f_\theta(\theta, a)(\rho L z - \lambda))_{I,\Omega} - \beta_2(\cdot, [\theta - \theta_m]_+)_{I,\Omega},\\
\rho^\lambda(\theta, a, z, \lambda)(\cdot) &= -\sum_{n=1}^{N} (\cdot, \partial_t\lambda)_{I_n,\Omega} - \sum_{n=1}^{N-1} ((\cdot)_n, [\lambda]_n) + ((\cdot), f_a(\theta, a)(\rho L z - \lambda))_{I,\Omega},
\end{aligned}$$

and

$$\rho^u(z, u)(\cdot) \;=\; \left(\beta_3 u + \int_\Omega \alpha z dx, \;\; \cdot\right)_{L^2(I)}.$$

Theorem 5.4.1. *Let* $(\theta^*, a^*, u^*), (\theta^*_k, a^*_k, u^*_k), (\theta^*_{kh}, a^*_{kh}, u^*_{kh})$ *and* $(\theta^*_\sigma, a^*_\sigma, u^*_\sigma)$ *be the solutions of*(5.2.1)-(5.2.5), (5.4.8)-(5.4.12), (5.4.19)-(5.4.23) *and* (5.4.31)-(5.4.35), *respectively, with adjoint solutions as* $(z^*, \lambda^*), (z^*_k, \lambda^*_k)$, $(z^*_{kh}, \lambda^*_{kh})$ *and* $(z^*_\sigma, \lambda^*_\sigma)$. *Then, the following error estimates holds true:*

$$J(\theta^*_\sigma, a^*_\sigma, u_\sigma) - J(\theta^*, a^*, u^*) \;\le\; C\Bigg(\sum_{n=1}^{N} \Bigg(\sum_{K\in\mathcal{T}_h} \Bigg(\sum_{i=1}^{9} \rho^n_{i,K}\omega^n_{i,K} \Bigg) + \sum_{i=10}^{17} \rho^n_i \omega^n_i \Bigg) + R_k + R_h + R_d.$$

where

$$
\begin{aligned}
\rho_{1,K}^n &= \|\mathcal{K}\Delta\theta_{kh}^* - \rho c_p \partial_t \theta_{kh}^* - \rho L f(\theta_{kh}^*, a_{kh}^*)\|_{I_n,K} + h_K^{\frac{-1}{2}} \frac{\mathcal{K}}{2} \|\partial_\eta \theta_{kh}^*\|_{I_n,\partial K} + |K|^{\frac{1}{2}} \max_{I_n \times K} |\alpha| \|u_{kh}^*\|_{I_n}, \\
\omega_{1,K}^n &= \|z_k^* - I_h z_{kh}^*\|_{I_n,K} + h_K^{\frac{1}{2}} \|z_k^* - I_h z_{kh}^*\|_{I_n,\partial K}, \\
\rho_{2,K}^n &= \rho c_p k^{\frac{-1}{2}} \|[\theta_{kh}^*]_{n-1}\|_K, \quad \omega_{2,K}^n = \|z_k^* - I_h z_{kh}^*\|_{I_n,K} + k^{\frac{1}{2}} \|(z_k^* - I_h z_{kh}^*)_{n-1}^+\|_K, \\
\rho_{3,K}^n &= \|\mathcal{K}\Delta z_{kh}^* + \rho c_p \partial_t z_{kh}^* - f_\theta(\theta_{kh}^*, a_{kh}^*)(\rho L z_{kh}^* - \lambda_{kh}^*) + \beta_2 [\theta_{kh}^* - \theta_m]_+\|_{I_n,K} \\
&\quad + h_K^{\frac{-1}{2}} \frac{\mathcal{K}}{2} \|\partial_\eta z_{kh}^*\|_{I_n,\partial K}, \quad \omega_{3,K}^n = \|\theta_k^* - I_h \theta_{kh}^*\|_{I_n,K} + h_K^{\frac{1}{2}} \|\theta_k^* - I_h \theta_{kh}^*\|_{I_n,\partial K} \\
\rho_{4,K}^n &= \rho c_p k^{\frac{-1}{2}} \|[z_{kh}^*]_{n-1}\|, \quad \omega_{4,K}^n = \|\theta_k^* - I_h \theta_{kh}^*\|_{I_n,K} + k^{\frac{1}{2}} \|(\theta_k^* - I_h \theta_{kh}^*)_{n-1}^+\|_K, \\
\rho_{5,K}^n &= \|f(\theta_{kh}^*, a_{kh}^*) - \partial_t a_{kh}^*\|_{I_n,K}, \quad \omega_{5,K}^n = \|\lambda_k^* - I_h \lambda_{kh}^*\|_{I_n,K}, \\
\rho_{6,K}^n &= k^{-\frac{1}{2}} \|[a_{kh}^*]_n\|_K, \quad \omega_{6,K}^n = \|\lambda_k^* - I_h \lambda_{kh}^*\|_{I_n,K} + k^{\frac{1}{2}} \|(\lambda_k^* - I_h \lambda_{kh}^*)_{n-1}^+\|_K, \\
\rho_{7,K}^n &= \|\partial_t \lambda_{kh}^* - f_a(\theta_{kh}^*, a_{kh}^*)(\rho L z_{kh}^* - \lambda_{kh}^*)\|_{I_n,K}, \quad \omega_{7,K}^n = \|a_k^* - I_h a_{kh}^*\|_{I_n,K} \\
\rho_{8,K}^n &= k^{-\frac{1}{2}} (\|[\lambda_{kh}^*]_{n-1}\|_K), \quad \omega_{8,K}^n = (\|a_k^* - I_h a_{kh}^*\|_{I_n,K} + k^{\frac{1}{2}} \|(a_k^* - I_h a_{kh}^*)_{n-1}^+\|_K), \\
\rho_9^n &= \|\mathcal{K}\Delta\theta_k^* - \rho c_p \partial_t \theta_k^* - \rho L f(\theta_k^*, a_k^*)\|_{I_n} + |\Omega|^{\frac{1}{2}} \max_{I_n,\Omega} |\alpha| \|u_k^*\|_{I_n}, \quad \omega_9^n = \|z^* - I_k z_k^*\|_{I_n,\Omega}, \\
\rho_{10}^n &= \rho c_p k^{\frac{-1}{2}} \|[\theta_k^*]_{n-1}\|, \quad \omega_{10}^n = \|z^* - I_k z_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}} \|(z^* - I_k z_k^*)_{n-1}^+\|, \\
\rho_{11}^n &= \|\mathcal{K}\Delta z_k^* + \rho c_p \partial_t z_k^* - f_\theta(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k^*) + \beta_2 [\theta_k^* - \theta_m]_+\|_{I_n,\Omega} \\
\omega_{11}^n &= \|\theta^* - I_k \theta_k^*\|_{I_n,\Omega} \\
\rho_{12}^n &= \rho c_p k^{\frac{-1}{2}} \|[z_k^*]_{n-1}\|, \quad \omega_{12}^n = \|\theta^* - I_k \theta_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}} \|(\theta^* - I_k \theta_k^*)_{n-1}^+\|, \\
\rho_{13}^n &= \|f(\theta_k^*, a_k^*) - \partial_t a_k^*\|_{I_n,\Omega}, \quad \omega_{13}^n = \|\lambda^* - I_k \lambda_k^*\|_{I_n,\Omega} \\
\rho_{14}^n &= k^{-\frac{1}{2}} \|[a_k^*]_n\|, \quad \omega_{14}^n = \|\lambda^* - I_k \lambda_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}} \|(\lambda^* - I_k \lambda_k^*)_{n-1}^+\|, \\
\rho_{15}^n &= \|\partial_t \lambda_k^* - f_a(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k)\|_{I_n,\Omega}, \quad \omega_{15}^n = \|a^* - I_k a_k^*\|_{I_n,\Omega} \\
\rho_{16}^n &= k^{-\frac{1}{2}} \|[\lambda_k^*]_{n-1}\|, \quad \omega_{16}^n = (\|a^* - I_k a_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}} \|(a^* - I_h a_k^*)_{n-1}^+\|), \\
\rho_{17}^n &= \|\beta_3 u_\sigma + \int_\Omega \alpha z_\sigma^* dx\|_{L^2(I_n)}, \quad \omega_{17}^n = \|u_{kh}^* - I_k u_\sigma^*\|_{L^2(I_n)},
\end{aligned}
$$

where interpolation operators I_h and I_k are defined in the preliminaries in the beginning of the chapter.

Proof: Using (5.4.46)-(5.4.48) one can rewrite estimate for $J(\theta_\sigma^*, a_\sigma^*, u_\sigma^*) - J(\theta^*, a^*, u^*)$ as

$$
J(\theta_\sigma^*, a_\sigma^*, u_\sigma^*) - J(\theta^*, a^*, u^*) \quad \leq \quad \rho^\theta(\theta_k^*, a_k^*, u_k^*)(z^* - z_k) + \rho^a(\theta_k^*, a_k^*)(\lambda^* - \lambda_k)
$$

$$
\begin{aligned}
&+ \rho^z(\theta_k^*, a_k^*, z_k^*, \lambda_k^*)(\theta^* - \theta_k) + \rho^\lambda(\theta_k^*, a_k^*, z_k^*, \lambda_k^*)(a^* - a_k) \\
&+ \rho^u(z_\sigma^*, u_\sigma^*)(u_{kh}^* - u_\sigma) + \rho^\theta(\theta_{kh}^*, a_{kh}^*, u_{kh}^*)(z_k^* - z_{kh}) \\
&+ \rho^a(\theta_{kh}^*, a_{kh}^*)(\lambda_k^* - \lambda_{kh}) + \rho^z(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*)(\theta_k^* - \theta_{kh}) \\
&+ \rho^\lambda(\theta_{kh}^*, a_{kh}^*, z_{kh}^*, \lambda_{kh}^*)(a_k^* - a_{kh}) + R_k + R_h + R_d, \\
= \;& \sum_{i=1}^{5} I_i + \sum_{j=1}^{4} J_j + R_k + R_h + R_d. \qquad (5.4.52)
\end{aligned}
$$

where R_k, R_h and R_d are defined in Lemma 5.4.1 by (5.4.49)-(5.4.51).
For $\psi \in X_{kh}^q$, consider

$$
\begin{aligned}
J_1 &= |\rho^\theta(\theta_{kh}^*, a_{kh}^*, u_{kh}^*)(z_k^* - \psi)| \\
&= | - \rho c_p \sum_{n=1}^{N} (\partial_t \theta_{kh}^*, z_k^* - \psi)_{I_n,\Omega} - \mathcal{K}(\bigtriangledown \theta_{kh}^*, \bigtriangledown(z_k^* - \psi))_{I,\Omega} - \rho c_p \sum_{n=1}^{N} ([\theta_{kh}^*]_{n-1}, z_k^* - \psi_{n-1}^+) \\
&\quad - \rho L(f(\theta_{kh}^*, a_{kh}^*), z_k^* - \psi)_{I,\Omega} + (\alpha u_{kh}^*, z_k^* - \psi)_{I,\Omega}|.
\end{aligned}
$$

Applying integration by parts for the second term, we obtain

$$
\begin{aligned}
J_1 &= |\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(- \rho c_p \int_{I_n} \int_K \partial_t \theta_{kh}^* (z_k^* - \psi) dx dt + \mathcal{K} \int_{I_n} \int_K \Delta \theta_{kh}^* (z_k^* - \psi) dx dt \\
&\quad - \frac{\mathcal{K}}{2} \int_{I_n} \int_{\partial K} \partial_\eta \theta_{kh}^* (z_k^* - \psi) ds dt - \rho c_p \int_K [\theta_{kh}^*]_{n-1} (z_k^* - \psi)_{n-1}^+ dx dt \\
&\quad - \int_{I_n} \int_K \rho L f(\theta_{kh}^*, a_{kh}^*)(z_k^* - \psi) dx dt + \int_{I_n} \int_K \alpha u_{kh}^* (z_k^* - \psi) dx dt \Big)| \\
&= |\sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(\int_{I_n} \int_K (\mathcal{K} \Delta \theta_{kh}^* - \rho c_p \partial_t \theta_{kh}^* - \rho L f(\theta_{kh}^*, a_{kh}^*))(z_k^* - \psi) dx dt \\
&\quad - \rho c_p \int_K [\theta_{kh}^*]_{n-1} (z_k^* - \psi)_{n-1}^+ dx dt - \frac{\mathcal{K}}{2} \int_{I_n} \int_{\partial K} \partial_\eta \theta_{kh}^* (z_k^* - \psi) ds dt \\
&\quad + \int_{I_n} (u_{kh}^* \int_K \alpha (z_k^* - \psi) dx) dt \Big)| \\
&\leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_h} \Big(\int_{I_n} \| \mathcal{K} \Delta \theta_{kh}^* - \rho c_p \partial_t \theta_{kh}^* - \rho L f(\theta_{kh}^*, a_{kh}^*) \|_K \| z_k^* - \psi \|_K dt \\
&\quad + \frac{\mathcal{K}}{2} \int_{I_n} \| \partial_\eta \theta_{kh}^* \|_{\partial K} \| z_k^* - \psi \|_{\partial K} dt + \rho c_p \| [\theta_{kh}^*]_{n-1} \|_K \| (z_k^* - \psi)_{n-1}^+ \|_K \\
&\quad + |K|^{\frac{1}{2}} \max_{I_n \times K} |\alpha| \int_{I_n} |u_{kh}^*| \| z_k^* - \psi \|_K dt \Big)
\end{aligned}
$$

Substituting $\psi = I_h z_{kh}^*$, we obtain

$$\begin{aligned}
J_1 \;\leq\; & \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Bigg(\int_{I_n}(\|\mathcal{K}\Delta\theta^*_{kh}-\rho c_p\partial_t\theta^*_{kh}-\rho Lf(\theta^*_{kh},a^*_{kh})\|_K+h_K^{\frac{-1}{2}}\frac{\mathcal{K}}{2}\|\partial_\eta\theta^*_{kh}\|_{\partial K}\\
&+|K|^{\frac{1}{2}}\max_{I_n\times K}|\alpha|\;|u^*_{kh}|)(\|z^*_k-I_hz^*_{kh}\|_K+h_K^{\frac{1}{2}}\|z^*_k-I_hz^*_{kh}\|_{\partial K})dt\\
&+\rho c_pk^{\frac{-1}{2}}\|[\theta^*_{kh}]_{n-1}\|_K(\|z^*_k-I_hz^*_{kh}\|_{I_n,K}+k^{\frac{1}{2}}\|(z^*_k-I_hz^*_{kh})^+_{n-1}\|_K)\Bigg)\\
=\; & \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}(\rho^n_{1,K}\omega^n_{1,K}+\rho^n_{2,K}\omega^n_{2,K})
\end{aligned}$$

where

$$\begin{aligned}
\rho^n_{1,K} &= \|\mathcal{K}\Delta\theta^*_{kh}-\rho c_p\partial_t\theta^*_{kh}-\rho Lf(\theta^*_{kh},a^*_{kh})\|_{I_n,K}+h_K^{\frac{-1}{2}}\frac{\mathcal{K}}{2}\|\partial_\eta\theta^*_{kh}\|_{I_n,\partial K}+|K|^{\frac{1}{2}}\max_{I_n\times K}|\alpha|\|u^*_{kh}\|_{I_n},\\
\omega^n_{1,K} &= \|z^*_k-I_hz^*_{kh}\|_{I_n,K}+h_K^{\frac{1}{2}}\|z^*_k-I_hz^*_{kh}\|_{I_n,\partial K},\\
\rho^n_{2,K} &= \rho c_pk^{\frac{-1}{2}}\|[\theta^*_{kh}]_{n-1}\|_K,\quad \omega^n_{2,K}=\|z^*_k-I_hz_{kh}\|_{I_n,K}+k^{\frac{1}{2}}\|(z^*_k-I_hz^*_{kh})^+_{n-1}\|_{I_n,K}.
\end{aligned}$$

Now consider,

$$\begin{aligned}
J_2 \;=\; & \rho^z(\theta^*_{kh},a^*_{kh},z^*_{kh},\lambda^*_{kh})(\theta^*_k-v)=\Bigg(\rho c_p\sum_{n=1}(\theta^*_k-v,\partial_tz^*_{kh})_{I_n,K}-\mathcal{K}(\bigtriangledown(\theta^*_k-v),\bigtriangledown z^*_{kh})_{I,\Omega}\\
&+\rho c_p\sum_{n=1}^{N}((\theta^*_k-v)_n,[z^*_{kh}]_n)_{I_n,K}-(\theta^*_k-v,f_\theta(\theta^*_{kh},a^*_{kh})(\rho Lz^*_{kh}-\lambda^*_{kh}))_{I,\Omega}\\
&+\beta_2((\theta^*_k-v),[\theta^*_{kh}-\theta_m]_+)_{I,\Omega}\Bigg).
\end{aligned}$$

Integrating by parts the second term on the right hand side, applying Cauchy-Schwarz inequality, Young's inequality and replacing v by $I_hz^*_{kh}$, we have

$$\begin{aligned}
J_2 \;\leq\; & \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\Bigg(\int_{I_n}(\|\mathcal{K}\Delta z^*_{kh}+\rho c_p\partial_tz^*_{kh}-f_\theta(\theta^*_{kh},a^*_{kh})(\rho Lz^*_{kh}-\lambda^*_{kh})+\beta_2[\theta^*_{kh}-\theta_m]_+\|_K\\
&+h_K^{\frac{-1}{2}}\frac{\mathcal{K}}{2}\|\partial_\eta z^*_{kh}\|_{\partial K})(\|\theta^*_k-I_h\theta^*_{kh}\|_K+h_K^{\frac{1}{2}}\|\theta^*_k-I_h\theta^*_{kh}\|_{\partial K})dt\\
&+\rho c_pk^{\frac{-1}{2}}\|[z^*_{kh}]_{n-1}\|((\|\theta^*_k-I_h\theta^*_{kh}\|_{I_n,K}+k^{\frac{1}{2}}\|(\theta^*_k-I_h\theta^*_{kh})^+_{n-1}\|_K)\Bigg)\\
=\; & \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}(\rho^n_{3,K}\omega^n_{3,K}+\rho^n_{4,K}\omega^n_{4,K})
\end{aligned}$$

where

$$\begin{aligned}
\rho^n_{3,K} &= \|\mathcal{K}\Delta z^*_{kh} + \rho c_p \partial_t z^*_{kh} - f_\theta(\theta^*_{kh}, a^*_{kh})(\rho L z^*_{kh} - \lambda^*_{kh}) + \beta_2[\theta^*_{kh} - \theta_m]_+\|_{I_n,K} \\
&\quad + h_K^{\frac{-1}{2}} \frac{\mathcal{K}}{2} \|\partial_\eta z^*_{kh}\|_{I_n,\partial K} \\
\omega^n_{3,K} &= \|\theta^*_k - I_h\theta^*_{kh}\|_{I_n,K} + h_K^{\frac{1}{2}} \|\theta^*_k - I_h\theta^*_{kh}\|_{I_n,\partial K} \\
\rho^n_{4,K} &= \rho c_p k^{\frac{-1}{2}} \|[z^*_{kh}]_{n-1}\|, \quad \omega^n_{4,K} = \|(\theta^*_k - I_h\theta^*_{kh})^+_{n-1}\|_{I_n,K} + k^{\frac{1}{2}}\|\theta^*_k - I_h\theta^*_{kh}\|_K.
\end{aligned}$$

Let

$$\begin{aligned}
J_3 &= \rho^a(\theta^*_{kh}, a^*_{kh})(\lambda^*_k - \phi) \\
&= \Bigg(-\sum_{n=1}^{N} (\partial_t a^*_{kh}, \lambda^*_k - \phi)_{I_n,\Omega} - \sum_{n=1}^{N} ([a^*_{kh}]_{n-1}, (\lambda^*_k - \phi)^+_{n-1}) + (f(\theta^*_{kh}, a^*_{kh}), \lambda^*_k - \phi)_{I,\Omega} \Bigg)
\end{aligned}$$

Applying Cauchy-Schwarz inequality, Young's inequality and replacing ϕ by $I_h\lambda^*_{kh}$, we obtain

$$\begin{aligned}
J_3 &\leq \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h} \Bigg(\|f(\theta^*_{kh}, a^*_{kh}) - \partial_t a^*_{kh}\|_{I_n,K} \|\lambda^*_k - I_h\lambda^*_{kh}\|_{I_n,K} + k^{-\frac{1}{2}} \|[a^*_{kh}]_{n-1}\|_K (\|\lambda^*_k - I_h\lambda^*_{kh}\|_{I_n,K} \\
&\quad + k^{\frac{1}{2}} \|(\lambda^*_k - I_h\lambda^*_{kh})^+_{n-1}\|_K) \Bigg) = \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h} (\rho^n_{5,K}\omega^n_{5,K} + \rho^n_{6,K}\omega^n_{6,K}),
\end{aligned}$$

where

$$\begin{aligned}
\rho^n_{5,K} &= \|f(\theta^*_{kh}, a^*_{kh}) - \partial_t a^*_{kh}\|_{I_n,K}, \quad \omega^n_{5,K} = \|\lambda^*_k - I_h\lambda^*_{kh}\|_{I_n,K}, \\
\rho^n_{6,K} &= k^{-\frac{1}{2}} \|[a^*_{kh}]_{n-1}\|_K, \quad \omega^n_{6,K} = \|\lambda^*_k - I_h\lambda^*_{kh}\|_{I_n,K} + k^{\frac{1}{2}} \|(\lambda^*_k - I_h\lambda^*_{kh})^+_{n-1}\|_K.
\end{aligned}$$

Consider

$$\begin{aligned}
J_4 &= \rho^\lambda(\lambda^*_{kh})(a^*_k - w) = \Bigg(\sum_{n=1}^{N} (\partial_t \lambda^*_{kh}, a^*_k - w)_{I_n,\Omega} + \sum_{n=1}^{N} ([\lambda^*_{kh}]_{n-1}, (a^*_k - w)^+_{n-1}) \\
&\quad - (f_a(\theta^*_{kh}, a^*_{kh})(\rho L z^*_{kh} - \lambda^*_{kh}), a^*_k - w)_{I,\Omega} \Bigg).
\end{aligned}$$

Applying Cauchy-Schwarz inequality, Young's inequality and replacing w by $I_h a^*_{kh}$, we obtain

$$\begin{aligned}
J_4 &\leq \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h} \Bigg(\|\partial_t \lambda^*_{kh} - f_a(\theta^*_{kh}, a^*_{kh})(\rho L z^*_{kh} - \lambda^*_{kh})\|_{I_n,K} \|a^*_k - I_h a^*_{kh}\|_{I_n,K} + k^{-\frac{1}{2}} \|[\lambda^*_{kh}]_{n-1}\|_K \\
&\quad (\|a^*_k - I_h a^*_{kh}\|_{I_n,K} + k^{\frac{1}{2}} \|(a^*_k - I_h a^*_{kh})^+_{n-1}\|_K) \Bigg) = \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h} (\rho^n_{7,K}\omega^n_{7,K} + \rho^n_{8,K}\omega^n_{8,K}),
\end{aligned}$$

where

$$\begin{aligned}
\rho_{7,K}^n &= \|\partial_t \lambda_{kh}^* - f_a(\theta_{kh}^*, a_{kh}^*)(\rho L z_{kh}^* - \lambda_{kh}^*)\|_{I_n,K}, \quad \omega_{7,K}^n = \|a_k^* - I_h a_{kh}^*\|_{I_n,K} \\
\rho_{8,K}^n &= k^{-\frac{1}{2}}\|[\lambda_{kh}^*]\|_K, \quad \omega_{8,K}^n = \|a_k^* - I_h a_{kh}^*\|_{I_n,K} + k^{\frac{1}{2}}\|(a_k^* - I_h a_{kh}^*)_{n-1}^+\|_K.
\end{aligned}$$

We proceed in a similar manner for the time discretization error estimator as for space-time discretization. We use Cauchy-Schwarz inequality, Young's inequality to complete the estimation for the time discretization. Replace ψ by $I_k z_k^*$ and consider

$$I^1 = |\rho^\theta(\theta_k^*, a_k^*, u_k^*)(z_k^* - \psi)| = \sum_{n=1}^N (\rho_9^n \omega_9^n + \rho_{10}^n \omega_{10}^n)$$

where

$$\begin{aligned}
\rho_9^n &= \|\mathcal{K}\Delta\theta_k^* - \rho c_p \partial_t \theta_k^* - \rho L f(\theta_k^*, a_k^*)\|_{I_n,\Omega} + |\Omega|^{\frac{1}{2}} \max_{I_n,\Omega} |\alpha| \|u_k^*\|_{I_n}, \quad \omega_9^n = \|z^* - I_k z_k^*\|_{I_n,\Omega}, \\
\rho_{10}^n &= \rho c_p k^{\frac{-1}{2}} \|[\theta_{kh}^*]_{n-1}\|, \quad \omega_{10}^n = \|z^* - I_k z_k\|_{I_n,,\Omega} + k^{\frac{1}{2}}\|(z^* - I_k z_k^*)_{n-1}^+\|_{I_n}.
\end{aligned}$$

Also,

$$I^2 = \rho^z(\theta_k^*, a_k^*, z_k^*, \lambda_k^*)(\theta^* - I_k \theta_k^*) = \sum_{n=1}^N (\rho_{11}^n \omega_{11}^n + \rho_{12}^n \omega_{12}^n),$$

where

$$\begin{aligned}
\rho_{11}^n &= \|\mathcal{K}\Delta z_k^* + \rho c_p \partial_t z_k^* - f_\theta(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k^*) - \beta_2[\theta_{kh}^* - \theta_m]_+\|_{I_n,\Omega}, \quad \omega_{11}^n = \|\theta^* - I_k \theta_k^*\|_{I_n,\Omega} \\
\rho_{12}^n &= \rho c_p k^{\frac{-1}{2}} \|[z_k^*]_{n-1}\|, \text{ and } \omega_{12}^n = \|\theta^* - I_k \theta_k^*\|_{I_n} + k^{\frac{1}{2}}\|(\theta^* - I_k \theta_k^*)_{n-1}^+\|,
\end{aligned}$$

and

$$I^3 = \rho^a(\theta^*, a_k^*)(\lambda^* - I_k \lambda_k^*) = \sum_{n=1}^N (\rho_{13}^n \omega_{13}^n + \rho_{14}^n \omega_{14}^n)$$

where

$$\begin{aligned}
\rho_{13}^n &= \|f(\theta_k^*, a_k^*) - \partial_t a_k^*\|_{I_n,\Omega}, \quad \omega_{13}^n = \|\lambda^* - I_k \lambda_k^*\|_{I_n,\Omega}, \\
\rho_{14}^n &= k^{-\frac{1}{2}}\|[a_k^*]_{n-1}\|, \quad \omega_{14}^n = \|\lambda^* - I_k \lambda_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}}\|(\lambda^* - I_k \lambda_k^*)_{n-1}\|.
\end{aligned}$$

$$I^4 = \rho^\lambda(\lambda_k^*)(a^* - I_k a_k^*) = \sum_{n=1}^N (\rho_{15}^n \omega_{15}^n + \rho_{16}^n \omega_{16}^n),$$

where

$$\begin{aligned}\rho_{15}^n &= \|\partial_t\lambda_k^* - f_a(\theta_k^*, a_k^*)(\rho L z_k^* - \lambda_k^*)\|_{I_n,\Omega}, \quad \omega_{15}^n = \|a^* - I_k a_k^*\|_{I_n,\Omega} \\ \rho_{16}^n &= k^{-\frac{1}{2}}\|[\lambda_k^*]_{n-1}\|, \quad \omega_{16}^n = \|a^* - I_k a_k^*\|_{I_n,\Omega} + k^{\frac{1}{2}}\|(a^* - I_k a_k^*)_{n-1}^+\|.\end{aligned}$$

The control error is given by

$$I^5 = \rho^u(u_\sigma^*)(u_{kh}^* - I_k u_\sigma^*) \leq \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\rho_{17}^n\omega_{17}^n,$$

where

$$\rho_{17}^n = \|\beta_3 u_\sigma + \int_\Omega \alpha z_\sigma^*\|_{L^2(I_n)} \text{ and } \omega_{17}^n = \|u_{kh}^* - I_k u_\sigma^*\|_{L^2(I_n)}.$$

Using J_1 to J_4 and I^1 to I^5 in (5.4.52), we finally obtain the *a posteriori* error estimate, which is given by

$$J(\theta_\sigma^*, a_\sigma^*, u_\sigma) - J(\theta^*, a^*, u^*) \leq C\left(\sum_{n=1}^{N}\left(\sum_{K\in\mathcal{T}_h}\left(\sum_{i=1}^{9}\rho_{i,K}^n\omega_{i,K}^n\right) + \sum_{i=10}^{17}\rho_i^n\omega_i^n\right) + R_k + R_h + R_d.$$

This completes the proof. □

Remark 5.4.3. *In DWR method, the space, time and control error estimators are given by*

$$\begin{aligned}\eta_k &= \sum_{n=1}^{N}\sum_{i=10}^{16}\rho_i^n\omega_i^n + R_k, \\ \eta_h &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\sum_{i=1}^{9}\rho_{i,K}^n\omega_{i,K}^n + R_h, \\ \eta_d &= \sum_{n=1}^{N}\rho_{17}^n\omega_{17}^n + R_d.\end{aligned}$$

Remark 5.4.4. *For the computational purpose all the solutions at different discretization levels are replaced by the solutions at complete discretization level.*

Remark 5.4.5. *Remainder terms R_k, R_h and R_d defined in* (5.4.49)-(5.4.51) *are of order $\mathcal{O}(k^2)$, $\mathcal{O}(h_K^2)$ and $\mathcal{O}(k^2)$, respectively, and therefore, are bounded.*

In the next section, the AFEM algorithm using residual and DWR methods are presented and compared.

5.5 Numerical Experiments

In order to use the error estimate obtained in Theorem 5.3.1 and Theorem 5.4.1 for the adaptive refinement, we use the following algorithm.

Adaptive Finite Element Algorithm

1. Find approximate solution $(\theta_\sigma^*, a_\sigma^*, u_\sigma^*)$ for the problem (5.4.31) -(5.4.35) (resp. (5.3.1)-(5.3.5)).

2. Using the given data and approximate solution find error estimate η_h, given in Theorem 5.3.1 (resp. Theorem 5.4.1), for the purpose of flagging those elements in the triangulation which are to be adapted.

3. Adapt the flagged in the state dependent triangulation using fixed fraction strategy

$$\beta \frac{k}{T}\frac{TOL}{2N_n} \leq \eta_h \leq \frac{k}{T}\frac{TOL}{2N_n} \quad (\beta = \frac{1}{4}),$$

   where N_n is the total number of unknowns in the space direction and TOL is the tolerance, which is taken as 10^{-5} in the numerical experiments.

4. If the error η_h less than the given tolerance, then stop else go to step 1 and repeat these steps with new refined grids.

Remark 5.5.1. *For the numerical experiments, piecewise continuous linear polynomial and piecewise constants have been used for the space and time discretization, repectively. Therefore, we have*

$$\Delta\theta_\sigma^* = 0, \quad \Delta z_\sigma^* = 0, \quad \partial_t\theta_\sigma^* = 0, \quad \partial_t a_\sigma^* = 0, \quad \partial_t z_\sigma^* = 0 \quad \textit{and} \quad \partial_t\lambda_\sigma^* = 0.$$

For numerical experiments we consider the laser surface hardening of steel problem given by (5.2.1)-(5.2.5) with the given data as prescribed in the numerical experiment section of Chapter 3. To start with the adaptivity procedure first the problem is solved on the initial triangulation given by Figure 5.2. Table 5.1 shows the error occurred due to space mesh refinement, where η_h is defined in Remark 5.3.1 and 5.4.3 using residual and DWR method, respectively.

Figure 5.3 and 5.4 shows the development of meshes over adaptive loop. Error estimator

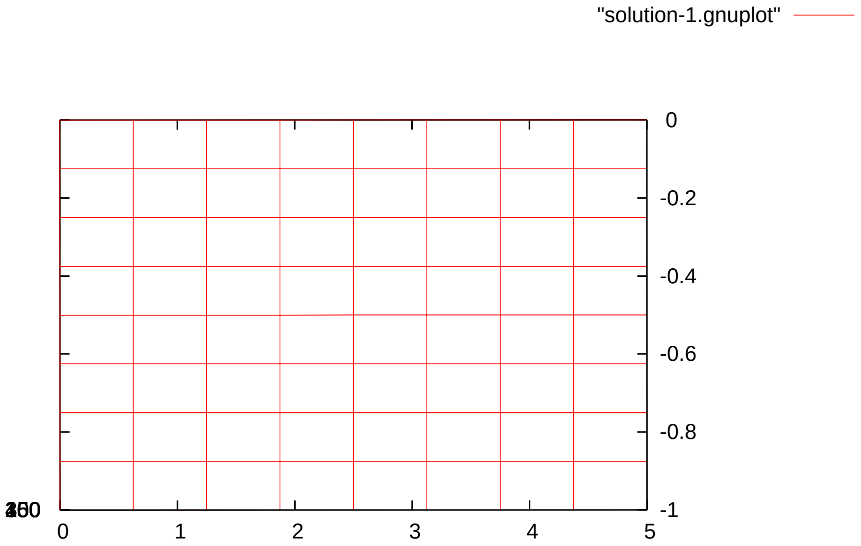
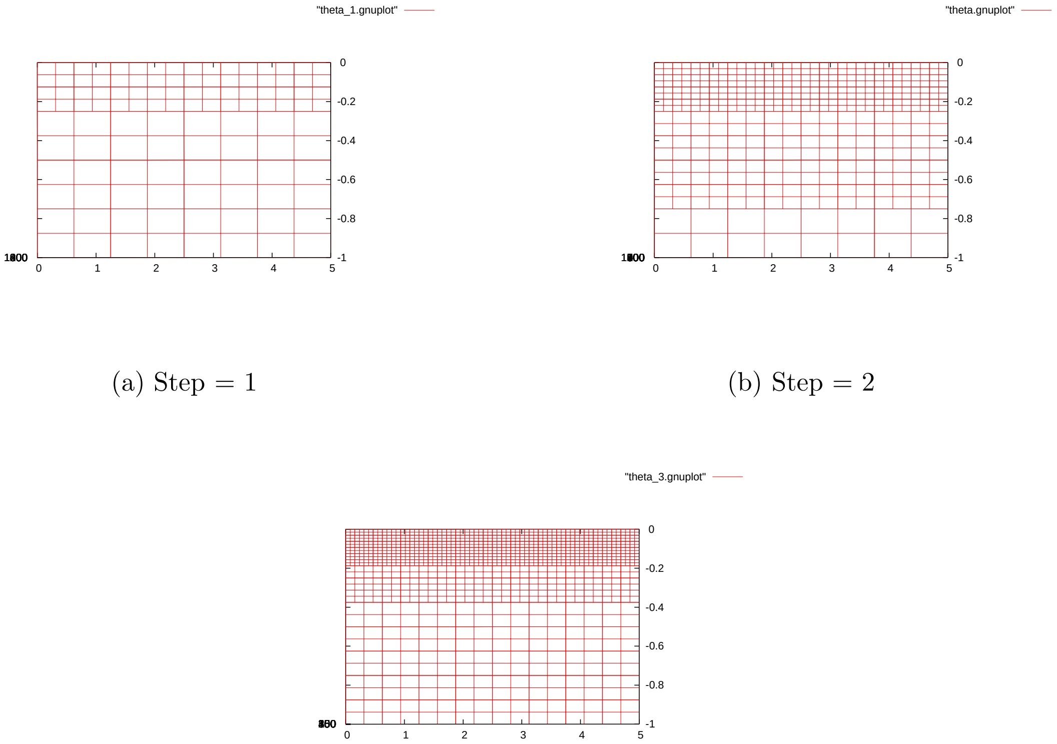


Figure 5.2: **Initial approximate triangulation**

N_n	η_h/J (DWR estimator)	η_h/J (Residual estimator)
81	0.000102	0.00022
143	0.000085	0.00019
463	0.000013	0.00007

Table 5.1: **Error in space for fixed time partition** 100



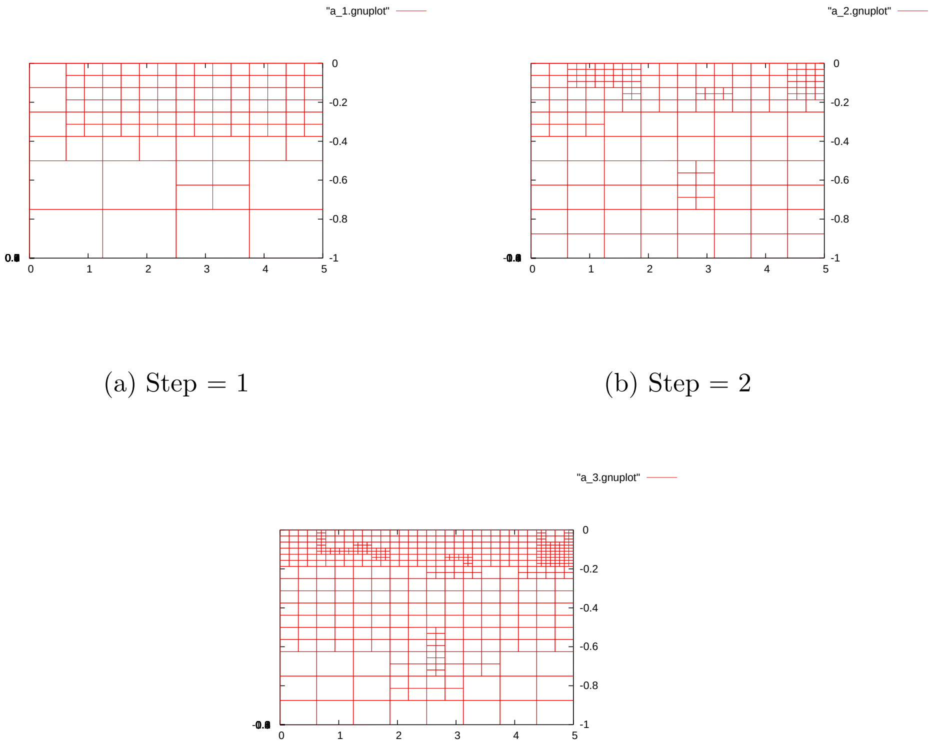
(a) Step = 1

(b) Step = 2

(c) Step = 3

Figure 5.3: **Adaptive refinement using DWR type estimators**

used for refinement strategy used in Figure 5.3 is DWR method and for that in Figure 5.4 is Residual method. It depicts that the triangulation gets more and more refined near the zone of heating, which is the boundary area. Comparison between Figure 5.3 and 5.4 shows that DWR type estimators provide better refinement strategy than residual type error estimator. Even though the refinement using both the methods can be seen near the boundary, Figure 5.4 shows extra refinement of the triangulation far from the boundary area. Figure 5.5 shows that increment in the mesh size causes the decrease in the error occurred due to the adaptive refinement. From Figure 5.5 also one can draw the conclusion that error due to the use of DWR type error estimator decreases at a faster rate than due to the residual type error estimator. Figure 5.6(a) depicts the austenite value at the final step on the final adaptive mesh using DWR type error estimates and Figure 5.6(b) the austenite value on final adaptive mesh using residual type estimator. Similarly Figure 5.7 shows temperature θ on the final mesh. Figure 5.8 shows the control at the final time $T = 5.25$.



(a) Step = 1

(b) Step = 2

(c) Step = 3

Figure 5.4: **Adaptive refinement using residual type estimators**

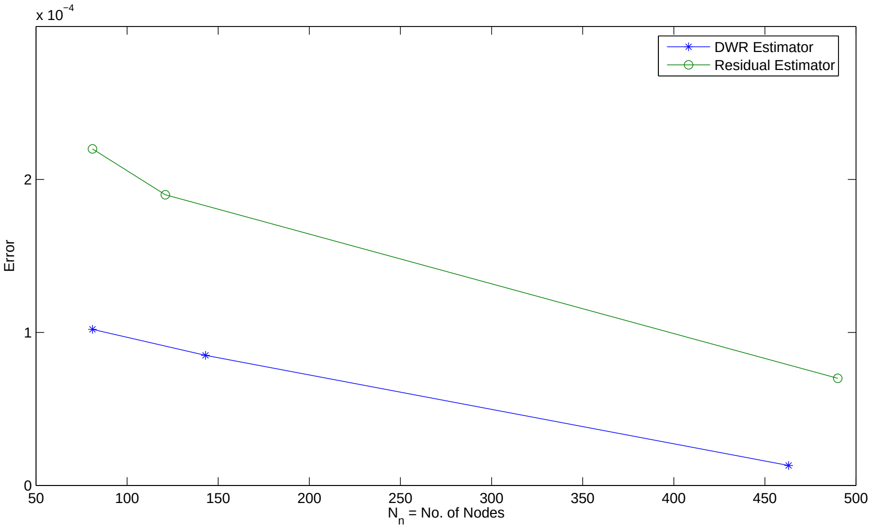


Figure 5.5: **Error graphs**

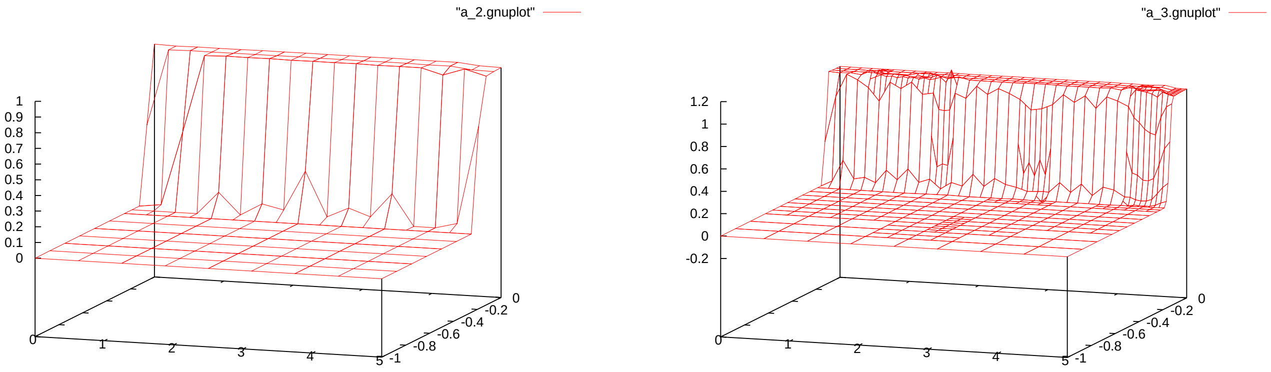


Figure 5.6: **The volume fraction of the austenite at time $t = T$ using (a) DWR estimator (b) residual estimator**

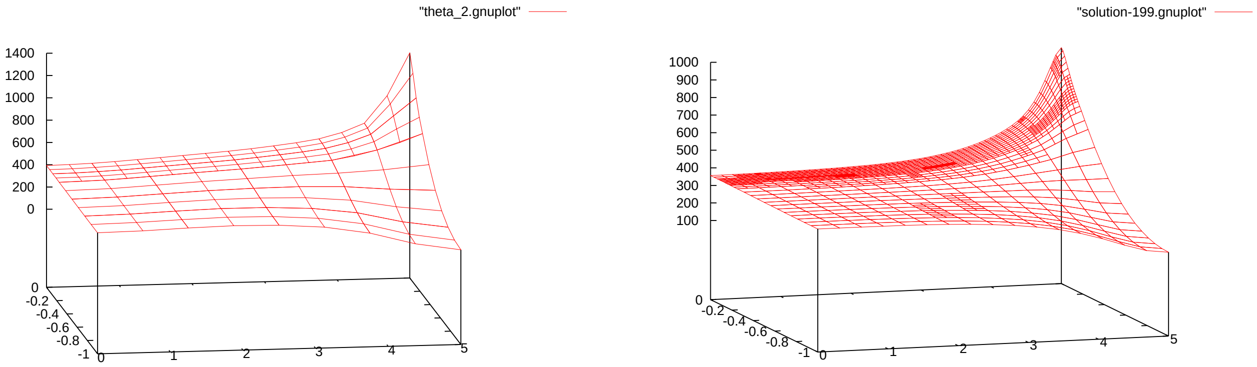


Figure 5.7: **The temperature at time $t = T$ using (a) DWR estimator (b) residual estimator**

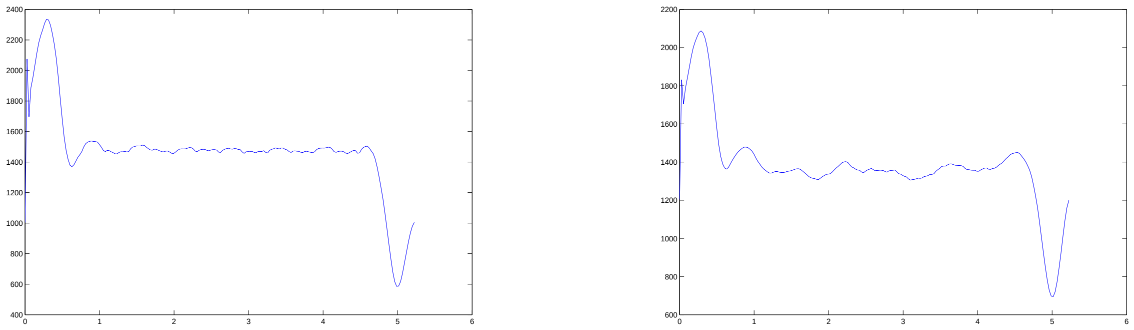


Figure 5.8: **(a) Control after using DWR AFEM (b) Control after using residual AFEM**

5.6 Summary

Since, laser surface hardening of steel problem has a nature of irregularity near the boundary due to use of laser energy, mesh obtained using AFEM is more refined near the boundary and coarse elsewhere. Adaptive finite element methods has helped in obtaining the mesh which depends on approximate solution and data. It has been shown that the mesh obtained using residual and DWR type *a posteriori* error estimates has helped in getting a approximate solution to the laser surface hardening of steel problem. Also, it has been observed through numerical experiments that the mesh obtained using DWR method is better than the one obtained using residual type error estimates.

Chapter 6

Conclusions and Future Directions

In this concluding chapter, we give highlights of the main results obtained in each chapter of the thesis. Further, we discuss the future possible extensions in the same direction.

6.1 Summary and Critical Assessments of the Results

In this thesis, numerical methods have been developed for the laser surface hardening of steel problem. Laser surface hardening of steel problem has been modeled as an optimal control problem governed by a semi-linear parabolic equation and an ordinary differential equation. Since, originally the laser surface hardening problem posseses non-differentiable load function, it has been regularized using a monotone regularized Heaviside function.

We have derived estimates for error due to regularization and have discussed different methods of discretization, a continuous Galerkin finite element method for space discretization and a discontinuous Galerkin finite element method for time and control discretization; an *hp*-discontinuous Galerkin finite element method for space discretization and a discontinuous Galerkin finite element method for time and control discretization; adaptive finite element method using residual and dual weighted residual type *a posteriori* error estimators; for the regularized laser surface hardening of steel problem. For adaptive finite element method, a continuous Galerkin finite element method for space discretization and a discontinuous Galerkin finite element method for time and control discretizations has been used.

All the numerical methods developed and analyzed in the literature are for the regularized problem. So, it became interesting to know whether the solution of the regularized problem converges to the solution of original problem. In Chapter 2, it has been established that the solution of the state system of the regularized problem converges to that of original problem for a fixed control $u \in U_{ad}$ with order of convergence $\mathcal{O}(\epsilon)$, that is,

$$\|\theta - \theta_\epsilon\| + \|a - a_\epsilon\| \leq C(\theta, a)\epsilon.$$

Also, existence and uniqueness of solution of the state system has been established for a fixed control $u \in U_{ad}$. Further, it has been shown that the optimal control of the regularized problem converges strongly in $L^2(I)$ to the optimal control of original problem.

In Chapter 3, discretization using a continuous Galerkin finite element method for space and a discontinuous Galerkin method for time and control has been developed. First, it has been shown that the solution of the semi-discrete state system converges to the weak solution, for a fixed control $u \in U_{ad}$, with order of convergence $\mathcal{O}(h^2)$, that is,

$$\|\theta_\epsilon - \theta_{\epsilon,h}\| + \|a_\epsilon - a_{\epsilon,h}\| \leq C(\theta_\epsilon, a_\epsilon)h^2.$$

It has then been shown that the solution of the space-time discrete state system converges with order of convergence $\mathcal{O}(h^2 + k)$, that is,

$$\|\theta_\epsilon(t_n) - \theta^n_{\epsilon,hk}\| + \|a_\epsilon(t_n) - a^n_{\epsilon,hk}\| \leq C(\theta_\epsilon, a_\epsilon, u_\epsilon)\Big(h^2 + k\Big).$$

The error estimates obtained at different level of discretizations are optimal in nature. Further, it has been shown that the optimal control obtained at fully discrete level converges with the order of convergence $\mathcal{O}(k)$, that is,

$$\|u_{\epsilon,\sigma} - u_\epsilon\|_{L^2(I)} \leq C(\theta_\epsilon, a_\epsilon, z_\epsilon, \lambda_\epsilon, u_\epsilon)k.$$

It has been assumed that the partition used for the control variable is same as that of the time variable. Combining the results from Chapter 2 and 3, it has been proved that state solutions of the regularized problem converges to that of the original problem with the order of convergence $\mathcal{O}(h^2 + k + \epsilon)$, that is,

$$\|\theta(t_n) - \theta^n_{\epsilon,hk}\| + \|a(t_n) - a^n_{\epsilon,hk}\| \leq C(\theta, a, u)\Big(h^2 + k + \epsilon\Big).$$

Also, the convergence of optimal control of the completely discrete system to that of the original problem has been shown in L^2-norm. In the last section of Chapter 3, numerical results has been attached for the justification of the theoretical results obtained in Chapter 2 and Chapter 3.

For the numerical implementation in Chapter 3, the mesh used was non-uniform in nature. Since laser surface hardening of steel problem has a irregularity near the boundary it became important to use a highly refined mesh near the boundary and a coarse mesh else where. Also, using non-uniform triangulation is expensive with same degree of piecewise polynomial approximations in each element of the triangulation. *hp*-DGFEM helps in choosing piecewise polynomial approximation with different degrees in each element and also permits non-uniform grids with hanging nodes. This method is easier to implement than the conformal finite element method used in Chapter 3, as an assembly of piecewise discontinuous polynomials in different elements is easier than compared to the global assembly of continuous polynomials. It has been established through Theorem 4.4.1, 4.4.2 and numerical experiments that solution of the completely discrete problem converges with the order $\mathcal{O}\bigg(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\bigg(\bigg(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\bigg)\frac{h_K^{2s-2}}{p_K^{2s'-2}}+k_n^2\bigg)\bigg)$, that is,

$$\begin{aligned}\|\theta_\epsilon-\theta_{\epsilon,\sigma}\|^2_{L^\infty(I,L^2(\Omega))} &+ \|a_\epsilon-a_{\epsilon,\sigma}\|^2_{L^\infty(I,L^2(\Omega))}+\|u_\epsilon-u_{\epsilon,\sigma}\|^2_{L^2(I)}\\ &\le C(\theta_\epsilon,a_\epsilon,z_\epsilon,\lambda_\epsilon,u_\epsilon)\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h}\bigg(\bigg(\max_{K\in\mathcal{T}_h}\frac{h_K^2}{p_K}\bigg)\frac{h_K^{2s-2}}{p_K^{2s'-2}}+k_n^2\bigg),\end{aligned}$$

where h_K is the diameter of $K\in\mathcal{T}_h$ and p_K, is the degree of piecewise polynomial used in each element $K\in\mathcal{T}_h$, k_n being the length of the interval I_n. Numerical results presented at the end of Chapter 4 justifies the theoretical results obtained.

Even though *hp*-DGFEM developed in Chapter 4 is easier to implement than the continuous Galerkin method developed in Chapter 3 both the methods fail in obtaining the right amount of non-uniformity of triangulation over the computational domain. Since triangulation of the domain does not directly depend on the solution of the problem, non-uniformity of triangulation used may solve the issue of expensiveness to only some extent. To resolve this problem, adaptive finite element methods can been used. Adaptive finite element methods developed in Chapter 5 has provided a efficient way to choose a triangulation which depends on the approximate solutions obtained. To use adaptive finite element method, *a posteriori* error estimates have been developed, which depends on the approximate solution obtained on the initial mesh and data given. *A posteriori* error estimators provides a way to choose a triangulation which is refined near area where the solution is not regular and coarse where it is not. Therefore, AFEM developed in Chapter 5 ensures higher density of nodes in certain areas of the computational domain.

To find the *a posteriori* error estimates, residual and dual weighted residual type error estimators are developed. Though an extensive study of residual and DWR methods are available for elliptic/parabolic problems, there is not much literature in which a use of these methods for the nonlinear problem of the laser surface hardening of steel has been studied. Residual type error estimators provides a bound in terms of global norms, for example, energy norm and L^2 norm, whereas DWR type error estimates provides a bound on the cost functional under consideration. Since laser surface hardening of steel problem is an optimal control problem, numerical results illustrates that the DWR method does provide better results for the laser surface hardening of steel problem.

Some points of comparison of three approaches used in the thesis for the laser surface hardening of steel problem

- In CG method, the degree of piecewise polynomials used is same in all the elements of the triangulation while in *hp*-DGFEM this condition is relaxed.
- A discretization using *hp*-DGFEM is memory intensive compared to discretization using cG FEM.
- The assembly of local matrices to corresponding global matrices is much easier in DGFEM compared to cG FEM.
- Optimal order estimates are obtained for both the cases.
- AFEM has the advantage over classical cG FEM and DGFEM that the triangulation used need not be chosen a priori; it can be adapted based on the *a posteriori* error estimates, which depend on the approximate solution and data.
- AFEM using cG method demands global continuity, which involve interpolation expenses.
- It has been observed that AFEM using DWR estimates gives more accurate results than the other three methods used in this thesis.

Table 6.1: **Comparitive study between three approaches used in the thesis**

	cG Method	dG Method	Adaptive Method
Degrees of polynomial:	Same	Different	Same
Memory:	Less Expensive	Memory intensive	Expensive
Assembly:	Global	Local	Global
Solution dependence of mesh used :	*A priori*	*A priori*	Dependent
Result:	Optimal	Optimal	Better than the rest

6.2 Future Extensions

In Chapter 5, *a posteriori* error estimates have been developed and used for adaptive finite element method with a continuous Galerkin discretization in space and a discontinuous Galerkin method for discretization in time and control variable. An extension can be done to develop error estimators for the discontinuous Galerkin method used for space discretization. Also, being an optimal control problem governed by a nonlinear system, the numerical method used becomes expensive and hence domain decomposition methods with parallel implementation for the laser surface hardening of steel problem can be a interesting extension.

In this thesis, the domain of computation under consideration is an open bounded subset of $\mathcal{R}^2$. Since, the domain is a thin sheet of steel, height of the sheet was ignored for the purpose of analysis. This work can be extended easily for the three dimensional problem with appropriate changes in the analysis done for the two dimensional problem. Similar results for the convergence for finite element method and adaptive finite element can also be done.

Another possible extension of the work presented in this thesis is as follows. Throughout in this thesis, we have considered change only in phase transition for austenite due to the heating. To estimate or develop similar kind of results for surface hardening of problem describing changes in martensite and austenite can be a possible extension of this work. In this case, the model under consideration, for $\Omega \subset \mathcal{R}^3$ with smooth boundary $\partial\Omega$ reads as:

$$\min_{u \in U_{ad}} J(u) \quad \left(= \frac{\beta_1}{2}\|a(T) - a_d\|^2 + \frac{\beta_2}{2}\int_0^T \|[\theta - \theta_m]_+\|^2 dt + \frac{\beta_3}{2}\|m(T) - m_d\|^2 + \frac{\beta_4}{2}\int_0^T |u|^2 dt \right),$$

subject to

$$\begin{aligned}
\partial_t a &= F_1(\theta, a) \quad \text{in} \quad \Omega \times I, \\
\partial_t m &= F_2(\theta, a, m) \quad \text{in} \quad \Omega \times I, \\
a(0) = 0, &\qquad m(0) = 0 \quad \text{in} \ \ \Omega, \\
\rho c_p \partial_t \theta - \mathcal{K}\Delta\theta &= -\rho L_1 F_1(\theta, a) + \rho L_2 F_2(\theta, a, m) + \alpha u \quad \text{in} \quad \Omega \times I,
\end{aligned}$$

$$\begin{aligned}
\mathcal{K}\frac{\partial\theta}{\partial n} + c\theta &= 0 \qquad \text{on} \quad \partial\Omega \times I, \\
\theta(0) &= \theta_0 \quad \text{in} \quad \Omega,
\end{aligned}$$

where, in the cost functional $J(\cdot)$, $\beta_i,\ i = 1, 2, 3, 4$ are given weights, m_d and a_d are desired volume fractions for martensite and austenite. Also,

$$\begin{aligned}
F_1(\theta, a) &= \frac{1}{\tau_1(\theta)}(a_{eq}(\theta) - a)\mathcal{H}(a_{eq}(\theta) - a), \\
F_2(\theta, a, m) &= \frac{1}{\tau_2(\theta)}(a.m_{eq}(\theta) - m)\mathcal{H}(a.m_{eq}(\theta) - m)\mathcal{H}(M_s - \theta),
\end{aligned}$$

where a and m are the volume fractions of occurring phases, $\tau_1, \tau_2, a_{eq}(\theta)$, $m_{eq}(\theta)$ are positive, Lipschitz continuous functions and M_s is a threshold temperature. The coefficients appearing in heat conduction equation are all positive constants.

Another possible extension can be to replace positive constants $c_p, \mathcal{K}, L_1, L_2$ and α with Lipschitz functions in θ. In this case, the heat conduction equation is replaced by

$$\rho c_p(\theta)\partial_t\theta - div(\mathcal{K}\nabla\theta) = -\rho L_1(\theta) F_1(\theta, a) + \rho L_2(\theta) F_2(\theta, a, m) + \alpha(\theta) u \quad \text{in} \quad \Omega \times I.$$

A study of mathematical analysis, taking into account the effect of regularization, and various numerical methods for this problem will be an interesting and challenging future problem.

References

[1] Alt, W. and Mackenroth, U., *Convergence of finite element approximations to state constrained convex parabolic boundary control problems*, SIAM J. Control Optim., 27, pp 718-736, 1989 .

[2] Arnăutu, V., Hömberg, D. and Sokołowski, J., *Convergence results for a nonlinear parabolic control problem*, Numer. Funct. Anal. Optim., 20, pp. 805-824, 1999.

[3] Arnold, D.N., *An interior penalty method for discontinuous elements*, SIAM J. Numer. Anal., pp.742-760, 1982.

[4] Arnold, D.N., Brezzi, F., Cockburn, B. and Marini, L.D., *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39, pp. 1749-1779, 2002.

[5] Babuška, I., *The finite element method with penalty*, Math. Comput., 27, pp. 221-228, 1973.

[6] Babuška, I., and Rheinboldt, W.C., *A posteriori error estimates for finite element methods*, Int. J. Numer. Methods Eng, 12, pp. 1597-1615, 1978.

[7] Bangerth, W., Hartmann, R. and Kanschat, G., *deal.II–A general-purpose object-oriented finite element library*, ACM Transactions on Mathematical Software, 33, Article 24, 2007.

[8] Bangerth, W. and Rannacher, R., *Adaptive finite element methods for differential equations*, Lectures in Mathematics, ETH Zurich, Birkhäuser Verlag, Basel 2003.

[9] Bank, R.E., and Weiser, A., *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44, pp. 283-301, 1985.

[10] Becker, R. and Kapp, H., *Optimization in PDE models with adaptive finite element discretization*, Report 98-20(SFB 359), University of Heidelberg, Heidelberg, Germany, 1998.

[11] Becker, R., Kapp, H. and Rannacher, R.,*Adaptive finite element methods for optimal control of partial differential equations: basic concepts*, SIAM Journal on Control Optim., 39, pp. 113-132, 2000.

[12] Becker, R., Meidner, D. and Vexler, B., *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optimization Methods and Software, 22, pp. 813-833, 2007.

[13] Becker, R. and Rannacher, R.,*A feed-back approach to error control in finite element methods: basic analysis and examples*, East-West J. Numer. Anal., 4, pp. 237-264, 1996.

[14] Becker, R. and Rannacher, R., *An optimal control approach to error estimation and mesh adaptation in finite element methods*, Acta Numerica 2000 (A. Iserles, ed.), Cambridge University Press, pp. 1-102, 2001.

[15] Bernardi, C., Dauge, M. and Maday, Y., *Polynomials in the Sobolev world*, Preprint of the Laboratoire Jacques-Louis Lions, No. R03038, 2003.

[16] Brenner, S. and Scott, R., *The mathematical theory of finite element methods*, 3^{rd} edition, Springer, 2007.

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

[18] Cockburn, B., Karniadakis, G.E. and Shu, C.W., *The development of discontinuous Galerkin methods, in Discontinuous Galerkin Methods, Theory, Computation and Applications*, Cockburn, B., Karniadakis, G.E. and Shu, C.W., eds, Lecture Notes in Comput. Sci. Engrg., Springer-Verlag, 11, pp. 3-50, 2000.

[19] Crouzeix, M., Thomée, V. and Wahlbin L. B., *Error estimates for spatially discrete approximation of semilinear parabolic equations with initial data of low regularity*, Math. Comp. 53, pp. 24-41, 1989.

[20] Douglas,Jr. J., and Dupont, T., *Interior penalty procedures for elliptic and parabolic Galerkin methods*, Computing Methods in Applied Sciences, Lecture Notes in Phys. 58, Springer-Verlag, Berlin, pp. 207-216, 1976.

[21] Eriksson, K. and Johnson, C.,*Adaptive finite element methods for parabolic problems I: a linear model problem*, SIAM J. Numer. Anal., 28, pp. 43-77, 1991.

[22] 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, pp. 706-740, 1995.

[23] Evans, L.C., *Partial differential equations*, AMS, 1998.

[24] Falk, F.S., *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44, pp. 28-47, 1973.

[25] French, D. and King, J.T., *Approximation of an elliptic control problem by the finite element method*, Numer. Funct. Anal. Optim., 12, pp. 299-314, 1991.

[26] Geveci, T., *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, RAIRO Anal. Numer., 13, pp. 313-328, 1979.

[27] Gronwall, T.H., *Note on the derivative with respect to a parameter of the solutions of a system of differential equations*, Ann. of Math, 20, pp. 292-296, 1919.

[28] Gudi, T., Nataraj, N. and Pani, A.K., *hp-discontinuous Galerkin methods for strongly nonlinear elliptic boundary value problems*, Numer. Math., 109, pp. 233-268, 2008.

[29] Gudi, T., Nataraj, N. and Pani, A.K., *An hp-local discontinuous Galerkin method for some quasilinear elliptic boundary value problems of nonmonotone type*, Math. Comp., 77, pp. 731-756, 2008.

[30] Gudi, T., Nataraj, N. and Pani, A.K., *Discontinuous Galerkin methods for quasi-linear elliptic problems of nonmonotone type*, SIAM J. Numer. Anal., 45, pp. 163-192, 2007.

[31] Gupta, N., Nataraj, N. and Pani, A.K., *A priori error estimates for optimal control of laser surface hardening of steel*, Under review.

[32] Gupta, N., Nataraj, N. and Pani, A.K., *An optimal control problem of laser surface hardening of steel*, Under review.

[33] Gunzburger, M.D. and Hou, L., *Finite dimensional approximation of a class of constrained nonlinear optimal control problems*, SIAM J. Control Optim., 34, pp. 1001-1043, 1996.

[34] Gunzburger, M.D., Hou, L. and Svobodny, T., *Analysis and finite element approximation of optimal control problems for stationary Navier-Stokes equations with Dirichlet controls*, RAIRO Model. Math. Anal. Numer., 25, pp. 711-748, 1991.

[35] Hiriat-Urruty, J-B., *Fundamentals of convex functions*, Springer-Verlag, Berlin, 2001.

[36] Houston, P., Schwab, C., and Süli, E., Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems, SIAM J. Numer. Anal., 39, pp. 2133-2163, 2002.

[37] Houston, P. and Süli, E., *A posteriori error analysis for linear convection-diffusion problems under weak mesh regularity assumptions*, Report 97/03, Oxford University Computing Laboratory, Oxford, UK, 1997.

[38] Hyers, D.H., Isac, G. and Rassias, T.M., *Topics in non-linear analysis and applications*, World-Scientific, 1997.

[39] Kesavan, S., *Topics in functional analysis and applications*, Wiley-Eastern Ltd., 1989.

[40] Kreyszig, E., *Introductory functional analysis with applications*, John-Wiley & Sons, Singapore, 1989.

[41] Knowles, G., *Finite element approximation of parabolic time optimal control problems*, SIAM J. Control Optim., 20, pp. 414-427, 1982.

[42] Kunisch, K., *Sequential and parallel splitting methods for bilinear control problems in Hilbert spaces*, SIAM J. Numer. Anal. 34, pp. 91-118, 1997.

[43] Hömberg, D., *A mathematical model for the phase transitions in eutectoid carbon steel*, IMA Journal of Applied Mathematics, 54, pp. 31-57, 1995.

[44] Hömberg, D., *Irreversible phase transitions in steel*, Math. Methods Appl. Sci., 20, pp. 59-77, 1997.

[45] Hömberg, D. and Fuhrmann, J., *Numerical simulation of surface hardening of steel*, Int. J. Numer. Meth. Heat Fluid Flow, 9, pp. 705-724, 1999.

[46] Hömberg, D. and Sokolowski, J., *Optimal control of laser hardening*, Adv. Math. Sci., 8, pp. 911-928, 1998.

[47] Hömberg, D. and Volkwein, S., *Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition*, Math. Comput. Modelling, 37, pp. 1003-1028, 2003.

[48] Hömberg, D. and Weiss, W., *PID-control of laser surface hardening of steel*, IEEE Trans. Control Syst. Technol., 14, pp. 896-904, 2006.

[49] Lasiecka, I., *Ritz-Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions*, SIAM J. Control Optim., 22, pp. 477-499, 1984.

[50] Lasis, A. and Süli, E., *hp-version discontinuous Galerkin finite element methods for semilinear parabolic problems*, Report no. 03/11, Oxford university computing laboratory, 2003.

[51] Lasis, A. and Süli, E., *Poincaré-type inequalities for broken Sobolev spaces*, Isaac Newton Institute for Mathematical Sciences, Preprint No. NI03067-CPD, 2003.

[52] 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, pp. 137-146, 1984.

[53] Li, R., Liu, W. B., Ma, H. P. and Tang, T.,*Adaptive finite element approximation for distributed elliptic optimal control problems*, SIAM J. on Control Optim., 41, pp. 1321-1349, 2002.

[54] Li, R., Liu, W. B., Ma, H. P. and Tang, T.,*A posteriori error estimates and discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*, SIAM J. Numer. Anal., 42, pp. 1032-1061, 2004.

[55] Liu, W.B. and Yan, N., *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15, pp. 285-309, 2001.

[56] Liu W. B. and Yan, N.,*A posteriori error estimates for convex boundary control problems*, SIAM J. Numer. Anal., 39, pp. 100-127, 2001.

[57] Liu, W. B. and Yan, N., *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., 93, pp. 497-521, 2003.

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

[59] Mazhukin, V.I. and Samarskii, A.A., *Mathematical modelling in the technology of laser treatments of materials*, Surveys Math. Indust., 4, pp. 85-149, 1994.

[60] McKnight, R.S. and Bosarge, Jr. W.S., *The Ritz-Galerkin procedure for parabolic control problems*, SIAM J. Control Optim., 11, pp. 510-524, 1973.

[61] Meidner, D. and Vexler, B., *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. on Control Optim., 46, pp. 116-142, 2007.

[62] Meidner, D. and Vexler, B., *A priori error estimates for the space-time finite element discretization of the parabolic optimal control problems part I: problems without constraints*, SIAM J. on Control Optim., 47, pp. 1150- 1177, 2008.

[63] Meidner, D. and Vexler, B., *A priori error estimates for space-time finite element discretization of parabolic optimal control problems part II: problems with control constraints*, SIAM J. Control Optim., 47, 1301-1329, 2008.

[64] Neittaanmaki, P and Tiba, D., *Optimal control of nonlinear parabolic systems*, Marcel Dekker, Inc., New York, 1994.

[65] Ohm, M.R., Lee, H.Y. and Shin, J.Y., *Error estimates for discontinuous Galerkin method for nonlinear parabolic equations*, J. Math. Anal. Appl. 315, 132, 2006.

[66] Nitsche, J.A., Uber ein Variationprinzip zur L Ísung Dirichlet-Problemen bei Verwendung von Teilr Íumen, die keinen Randbedingungen unteworfen sind, Abh. Math. Sem., Univ. Hamburg, 36 , pp. 9-15, 1971.

[67] Oden, J.T., Babuška, I. and Baumann, C.E., *A discontinuous hp finite element method for diffusion problems*, J. Comput. Phys., 146, pp. 491-519, 1998.

[68] Picasso, M., *Anisotropic a posteriori error estimates for optimal control problem governed by the heat equation*, Int. J. Numer. Methods Partial Differential Equations, 22, pp. 1314-1336, 2006.

[69] Pironneau, O., *Optimal shape design for elliptic systems*, Springer, Berlin, 1984.

[70] Prudhomme, S., Pascal, F. and Oden, J.T., *Review of error estimation for discontinuous Galerkin method*, TICAM-report 00-27, October 17, 2000.

[71] Pundir, A.K., Joshi, M.C. and Pani, A.K., *On approximation theorems for controllability of non-linear parabolic problems*, IMA J. Math. Control Inform., 24, pp. 115-136, 2007.

[72] Rannacher, R., *Adaptive finite element methods for PDE-constrained optimal control problems*, Reactive Flows, Diffusion and Transport, Abschlussband SFB 359, Universitt Heidelberg, Springer, Heidelberg, 2005.

[73] Rivière, B., *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*, SIAM, 2008.

[74] Rivière, B. and Wheeler, M.F., *A discontinuous Galerkin method applied to nonlinear parabolic equations*, The Center for Substance Modeling, TICAM, The University of Texas, Austin TX 78712, USA.

[75] Rivière, B., Wheeler, M.F. and Girault, V., *A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems*, SIAM J. Numer. Anal., 39, pp. 902-931, 2001.

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

[77] Simon, J., *Compact sets in the space $L_p([0T] : B)$*. Annali di Matematica pura ed applicata (IV), CXLVI, pp. 65-96, 1987.

[78] Sternberg, J. and Griewank, A., *Reduction of storage requirement by checkpointing for time-dependent optimal control problems in ODEs*, Automatic Differentiation: Applications, Theory, and Implementations, Springer, Berlin, 2006.

[79] Thomée, V., *Galerkin finite element methods for parabolic problems*, Springer, 1997.

[80] Tiba, D. and Tröltzsch, F., *Error estimates for the discretization of state constrained convex control problems*, Numerical Funct. Anal. Optim., 17, pp. 1005-1028, 1996.

[81] Tröltzsch, F., *Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems-Strong convergence of optimal control*, Appl. Math. Optim., 29, pp. 309-329, 1994.

[82] Verdi, C., Visintin, A., *A mathematical model of the austenite-pearlite transformation in plain steel based on the Scheil's additivity rule*, Acta Metall., 35, pp. 2711-2717, 1987.

[83] Verfürth, R., *A review of a posteriori error estimation and adaptive mesh refinement techniques,* Wiley-Teubner, New York, Stuttgart, 1996.

[84] Volkwein, S., *Non-linear conjugate method for the optimal control of laser surface hardening*, Optimization Methods and Software, 19, pp. 179-199, 2004.

[85] Wheeler, M.F, *A priori L^2 error estimates for Galerkin approximations to parabolic differential equations*, SIAM J. Numerical Analysis, 10, pp. 723-759, 1973.

[86] Wheeler, M.F., *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15, pp. 152-161, 1978.

[87] Zeidler, E., *Nonlinear functional analysis and its applications*, Springer-verlag, New York, 1990.

Acknowledgements

It would not have been possible to write my dissertation without the help and support of lots of people around me. Even though I would like to acknowledge all of them, few of them are mentioned here.

Above all, I would like to thank my parents Shri Suresh Kumar Gupta and Smt. Mithlesh Gupta for their constant support and patience in pursuing my goals of higher studies in the field of Mathematics. Without their help and support, it would not have been possible to come to Indian Institute of Technology, Bombay and work as a research scholar in the Department of Mathematics.

I would not have imagined to complete my thesis without the help, support and patience of my thesis supervisor Prof. Neela Nataraj. She has been my guide on professional or personal level throughout my stay in IIT-Bombay. Putting together all my work without her advice, knowledge and lectures in subjects related to the work presented in this thesis would have been the hardest goal to achieve. I owe her a lot of gratitude for showing me the way of research in applied mathematics. She has been my force to understand the fundamentals and advances of applied mathematics. I can never thank her enough for the time she has spend over helping me to get a mature look over my work.

I am deeply grateful to Prof. Amiya Kumar Pani for his suggestions and advice in my research work. Lectures, workshops and conferences organized by him has helped me a lot in completing my thesis work. I would like to thank him for coming and listening to my presentations. His constant pressure to present work every week has forced me to work hard and be upto date to the new research fields in the area of applied mathematics.

I would like to express my gratitude towards Prof. Andreas Griewank for inviting me to Humboldt University, Berlin during December, 2005 to December, 2006 and giving me the platform to learn the software, DEAL (Differential Equations Algebraic Library) and checkpointing techniques, which I have used for all my numerical experiments. I would also like to thank him to support my visit to University of Heidelberg to meet Prof. Rolf Rannacher and have healthy discussion on DWR methods.

I am greatly thankful to my research progress committee members, Prof. A.K. Pani and Prof. Suresh Kumar for their suggestions and comments. I would also like to thank Department of Mathematics, IIT-Bombay for providing me the academic atmosphere to do research and implement my thesis problem. I would like to acknowledge University Grants

Commission for their fellowship support for 5 years.

Many thanks to my sister Kamakshi and my brother-in-law Amit for their encouragement, love and moral support throughout my research. With their support it had became easy and comfortable to pursue my goals. I would like to specially mention my sister for being my first best teacher of mathematics. She was my inspiration to join M.Sc. (Mathematics) in IIT-Delhi, where I started my career in applied mathematics.

My special thanks to my dear friend Varsha, who has helped me morally throughout my stay in IIT-Bombay. Without her support and push I would not have completed my thesis in time. Special thanks to my friends from the department, Sushmita, Sajid, Deepjyoti, to mention a few, for their help in my personal and academic problems. Last but not the least to thank my hostel inmates and friends, Smriti, Saagrika and Shruti to make friendly and homely environment in the hostel.

Finally, I would like to thank all, whose direct and indirect support helped me in completing my thesis in time.

IIT Bombay **Nupur Gupta**