Simultaneous Reconstruction of the Initial Temperature and Heat Radiative Coefficient

Lau Kin Wing

A thesis submitted to the Graduate School of The Chinese University of Hong Kong (Division of Mathematics)

in partial fulfillment of the requirements for the Degree of Master of Philosophy (M. Phil.)

Hong Kong July 2000

The Chinese University of Hong Kong holds the copyright of this thesis. Any person(s) intending to use a part or whole of the materials in the thesis in a proposed publication must seek copyright release from the Dean of the Graduate School.
Acknowledgement

I would like to give my deepest gratitude and thankness to my supervisor, Dr. Jun Zou for his enthusiastic and kind guidance. Although there were some unexpected troubles when I was preparing this thesis, he always encourages me and gives me continuous supervision, and finally I can overcome all the difficulties and finish this thesis. Many thanks, Dr. Zou.

Lau Kin Wing

The Chinese University of Hong Kong

Abstract

In this thesis, we investigate the simultaneous reconstruction of the initial temperature and heat radiative coefficient in a heat conduction system, when the measurement of temperature at a fixed time $\theta > 0$ and the measurement of temperature in a subregion of the physical domain are available.

For the numerical resolution of the inverse problem, the stability is the most essential issue. So we will first establish the stability for the conductive inverse problem considered in this thesis. Based on the stability results, we will then study the numerical reconstruction of the initial temperature and heat radiative coefficient. The reconstruction process is carried out in such a way that the temperature solution of the heat equation matches its fixed time observation and its subregion observation optimally in the $L^2$-norm sense, incorporated with a regularization of the $H^1$-seminorm for both the initial temperature and the radiative coefficient.

The infinite dimensional nonlinear optimization system involved will be discretized using the piecewise linear finite element method, the existence of discrete minimizers and convergence of the finite element approximation are presented. The discrete finite element problem is solved using a nonlinear gradient method, which is suggested to combine with an efficient nonlinear multigrid technique to accelerate the entire reconstruction process. Many numerical experiments are provided to demonstrate the efficiency of the proposed nonlinear multigrid gradient method.
論文簡介

本論文探討某傳熱物體在有熱源(heat source)的情況下，若已知某一時間該物體其中一部份的溫度場時，如何設計有效的數值方法求出該物體的最初溫度(initial temperature)及輻射系數(heat radiative coefficient)。

在論文中，我們會先探討上述逆問題(inverse problem)的穩定性，然後研究還原最初溫度及輻射系數的數值方法而還原的過程是以還原系統的溫度與已知之測量溫度在 L^2 范數或 H^1 半范下的誤差達到極小為標準。

本論文會以有限元方法(finite element method)，將原本是連續的最優化問題轉變為離散(discrete)問題，亦會論證最優化解的存在性及收斂性。而本論文是以非線性梯度方法(nonlinear gradient method)來求解這個離散最優化問題，以及加上多重網格(multigrid)的技巧來增加還原的效率。論文的最終部份會有實際例子以觀察多重網格方法的效果。
Contents

1 Introduction 3
  1.1 Heat conduction problem ............... 3
  1.2 Direct problem ........................ 4
  1.3 Inverse problem ........................ 4
  1.4 Difficulty of the inverse problems ...... 5
  1.5 A simple but important example for instability 5
  1.6 The purpose of this thesis ............... 7

2 Stability of the inverse problem 9
  2.1 Conditional stability results ........... 9
  2.2 Stability of the inverse problems ......... 11

3 The continuous formulation 30
  3.1 Constrained minimization problem ........ 30
  3.2 Existence of minimizers to the minimization problem 31

4 Discretization and its convergence 36
  4.1 Finite element space .................... 36
4.2 Two important discrete projection operators .... 37
4.3 Finite element problem ......................... 39
4.4 Existence of minimizers to the finite element problem 39
4.5 Discrete minimizers and global minimizers .... 42

5 Numerical algorithms .......................... 51
5.1 Gateaux derivative ............................... 51
5.2 Nonlinear single-grid gradient method ........... 53
5.3 Nonlinear multigrid gradient method ............. 55

6 Numerical experiments .......................... 60
6.1 One dimensional examples ....................... 60
6.2 Two dimensional examples ....................... 66
Chapter 1

Introduction

1.1 Heat conduction problem

Heat conduction problems are encountered in many industrial and engineering applications. Consider a body of heat conductive material $\Omega$. Let $u$ be the temperature of the material at position $x$ and time $t$, then $u$ can be modelled by the following well-known heat conduction equations:

$$u_t - \nabla \cdot (q(x) \nabla u) - p(x)u = f(x, t), \quad (x, t) \in \Omega \times (0, T), \quad (1.1)$$

with the initial temperature

$$u(x, 0) = \mu(x), \quad x \in \Omega, \quad (1.2)$$

and, for example, the Dirichlet boundary temperature

$$u(x, t) = \eta(x, t), \quad (x, t) \in \partial \Omega \times (0, T), \quad (1.3)$$

where $q(x)$ is the heat conductivity, $p(x)$ the thermal radiative coefficient of the material, and $f(x, t)$ the heat source.
1.2 Direct problem

The usual direct problem is to find the temperature distribution when the initial temperature and all the physical coefficients \( q(x), p(x) \) and the heat source \( f(x, t) \) are available. This problem is known to be well-posed: the solution exists uniquely and the solution is continuously dependent on \( u(x, 0), p(x), q(x) \) and \( f(x, t) \).

1.3 Inverse problem

In many applications, one is often interested in the inverse heat conduction problems. Three such typical inverse problems are:

1. Parameter identification problems

Suppose the heat source, boundary conditions and initial conditions are known, one wants to identify the material, i.e., to identify the heat conductivity \( q(x) \) and the thermal radiative distribution \( p(x) \) with some extra measurement of temperature \( u(x, t) \) on part of the boundary or in part of the interior of the physical domain \( \Omega \).

2. Heat source identification problems

Suppose the boundary temperature and initial conditions, the heat conductivity \( q(x) \) and the thermal radiative distribution \( p(x) \) are known, one wants to identify the heat source distribution \( f(x, t) \) with some extra measurement of temperature \( u(x, t) \) on part of the boundary or in part of the interior of the physical domain \( \Omega \).

3. Backward heat conduction problems

Suppose all the physical coefficients \( q(x), p(x) \), boundary conditions and the heat source are known, one wants to reconstruct the initial temperature \( u(x, 0) \), with the extra measurement of temperature \( u(x, t) \) at time \( t = T \).
1.4 Difficulty of the inverse problems

The inverse problems are well-known to be ill-posed and difficult to solve. By ill-posedness, here we mean that the solution may not exist, or the solutions exist but are not unique, or the solution is not continuously dependent on \( u(x, 0) \).

The major difficulty for the inverse problem lies in its instability, i.e., a small change of the observation data can cause tremendous change of the parameters, which are to be identified or to be reconstructed.

So to solve the inverse problems stably, some kind of regularization technique have to be introduced.

1.5 A simple but important example for instability

Consider the 1D parabolic equation:

\[
\frac{\partial u}{\partial t}(x, t) = \frac{\partial^2 u}{\partial x^2}(x, t), \quad x \in (0, \pi), \quad t > 0
\]

with BCs:

\[
u(0, t) = 0, \quad u(\pi, t) = 0.
\]

We are given \( u(x, 1) \); and try to find \( u(x, 0) \).

By separation of variables, we have

\[
u(x, t) = \sum_{n=1}^{\infty} e^{-n^2 t} a_n \sin nx,
\]

where \( a_n \) is given by

\[
a_n = \frac{2}{\pi} \int_{0}^{\pi} u(x, 0) \sin ny \, dy.
\]

Or write

\[
u(x, 1) = \int_{0}^{\pi} k(x, y) u(y, 0) \, dy \tag{1.4}
\]
with
\[ k(x, y) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2} \sin nx \sin ny. \]

Note that equation (1.4) is equivalent to
\[ K u(\cdot, 0) = u(\cdot, 1), \]
where one can show that \( K \) is compact and self-adjoint.

Under the condition
\[ \sum_{n=1}^{\infty} (c_n e^{n^2})^2 < \infty, \]
where \( \{c_n\} \) are the Fourier coefficients of \( u(x, 1) \), we have
\[ u(x, 0) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{n^2} c_n \sin nx \]
by singular value expansion and eigen-expansion, see [9].

So \( u(x, 0) \) exists only when \( \{c_n\} \) decays faster than \( e^{-n^2} \). This is not true even for \( u(x, 1) = x^2 \)! More importantly, we can see that small error in \( \{c_n\} \) will cause large change in \( u(x, 0) \). For example, a perturbation of \( 10^{-8} \) for \( c_5 \) will lead to an error of about 720 for \( u(x, 0) \).
1.6 The purpose of this thesis

The goal of this thesis is to investigate the possibility of the simultaneous numerical reconstruction of the initial temperature distribution $\mu$ and the heat radiative coefficient $p$ in (1.1)-(1.2). The major theoretical results and numerical analysis are taken from Yamamoto and Zou [32] but with more detailed proofs. It is well known that given only the measurement of temperature at a fixed time $\theta > 0$, the reconstruction of the initial temperature is highly ill-posed and actually impossible in most instances, let alone the case that we intend to recover both $p$ and $\mu$ here. Some extra measurement should be given in order to make the reconstruction possible. Through the stability analysis of the inverse problem to be established in Chapter 2 we find that the reconstruction of the initial status is possible with some extra observation of the temperature, say in a small subregion of the physical domain along the time direction. More precisely, Theorem 2.1 assures the Lipschitz stability in determining the coefficient $p$ but very weak stability in determining the initial status $\mu$. This leads us to assume that the following measurements are available for the heat conduction system throughout the thesis:

$$z_\theta(x) \approx u(x, \theta), \quad x \in \Omega \quad \text{and} \quad z(x, t) \approx u(x, t), \quad (x, t) \in \omega \times (0, T) \quad (1.5)$$

where $\omega$ is a subregion of $\Omega$. The distributional observation data $z_\theta$ and $z$ are possibly obtained through interpolations of the point observation values in practice.

In the literature one can find vast references on analytical and numerical methods for solving the inverse elliptic problems, see [1, 5, 9, 12, 17, 34] and the references therein. There are also many analytical and numerical investigations on the identification of the heat conductivity and diffusivity in parabolic systems, see [3, 13, 19, 22]. However to our knowledge, it seems there exists little research work which addresses the numerical reconstruction of the initial temperature and the coefficient. Surely it is not because this problem is not important. On the
contrary, the reconstruction of the initial temperature distribution and the radiative coefficient is extremely important in many practical applications. The fact may be due to the ill-posedness of the problem and the reconstruction is almost impossible in most instances, especially from the numerical point of view. However, with the guidance and encouragement of the stability analysis in Chapter 2 (see also [7] [6]), the current work will investigate such a possibility of numerical reconstruction. With the extra measurement as given in (1.5) we are indeed able to achieve certain satisfactory reconstruction results.

The reconstruction is carried out in such a way that the temperature solution of the heat conduction equation matches its subregion observation data \( z \) and the fixed time observation \( z_0 \) optimally in the \( L^2 \)-norm sense, with the help of a regularization of either \( H^1 \)-seminorm or bounded variation for both the radiative coefficient and initial temperature (Chapter 3). The continuous nonlinear optimization system and the parabolic equation involved will be discretized using the piecewise linear finite element method, the existence of discrete minimizers of the finite element system and their convergence to the global minimizers of the continuous optimization problem are proved (Chapter 4). Then the nonlinear finite element minimization problem is solved using a gradient method, which converges rather slowly in general although very stably, and is thus suggested to combine with an efficient nonlinear multigrid technique to accelerate the entire reconstruction process (Chapter 5). Numerical experiments are given to demonstrate the efficiency of the proposed nonlinear multigrid gradient method for the simultaneous reconstruction of the thermal radiative coefficient and the initial temperature status (Chapter 6).
Chapter 2

Stability of the inverse problem

2.1 Conditional stability results

In this chapter, we will review some conditional stability results for the inverse problem formulated in Chapter 1. These results are basically taken from Yamamoto and Zou [32] but with some more details. Such stability is fundamental for our subsequent numerical reconstruction of the initial temperature and the heat radiative coefficient in the heat conduction system. The stability was established in Choulli and Yamamoto [7], Choulli, Imanuvilov and Yamamoto [6] for some special heat conductive problems. Here we will prove the conditional stability for the heat conductive system (1.1)-(1.2) with the Dirichlet boundary condition (1.3), which appears to be much more technical than those previously handled cases from the numerical analysis point of view. However, the stability established here is still not perfectly adjusted to the same case as adopted in our later numerical reconstruction. The major difference lies in the considered solution classes: we adopt $H^1$-solutions in the numerical reconstruction while the stability of this chapter will mainly consider the sufficiently smooth solutions. The stability for the $H^1$-solution case is much more complicated and is still an open question.
Throughout this chapter, for $\gamma \in (0, 1)$ and $m \in \mathbb{N} \cup \{0\}$, we use

$$C^{\gamma+m}(\bar{\Omega}), \quad C^{\gamma+m,\frac{\gamma+m}{2}}(\bar{\Omega} \times [0, T]), \quad C^{\gamma+m,\frac{\gamma+m}{2}}(\partial \Omega \times [0, T])$$

to denote the usual Hölder spaces (cf. [10, 23]). We assume that $\partial \Omega$ is of class $C^{\gamma+3}$ with some $\gamma \in (0, 1)$ and the boundary function $\eta = \eta(x, t)$ in (1.3) satisfies

$$\eta \in C^{\gamma+3,\frac{\gamma+3}{2}}(\partial \Omega \times [0, T]), \quad \eta(x, t) > \eta_0 \text{ on } \partial \Omega \times [0, T] \quad (2.1)$$

with some constant $\eta_0 > 0$. For any fixed $M > 0$ and $r_0 > 0$, we set

$$\mathcal{U} = \left\{ (p, \mu) \in C^{\gamma+1}(\bar{\Omega}) \times C^{\gamma+3}(\bar{\Omega}); \mu|_{\partial \Omega} = \eta(\cdot, 0), \quad (\Delta \mu + p\mu)|_{\partial \Omega} = \eta_0(\cdot, 0), \quad \|p\|_{C^{\gamma+1}(\bar{\Omega})} \leq M, \quad \|\mu\|_{C^{\gamma+3}(\bar{\Omega})} \leq M, \quad \mu(x) \geq r_0 > 0 \text{ in } \bar{\Omega} \right\}. \quad (2.2)$$

Then by the classical parabolic theory (see [23], for example), for any $(p, \mu) \in \mathcal{U}$, there exists a unique solution $u(p, \mu) = u(p, \mu)(x, t) \in C^{\gamma+3,\frac{\gamma+3}{2}}(\bar{\Omega} \times [0, T])$ to the problem (1.1)-(1.3). Let $\omega \subset \Omega$ be any subdomain, and $\theta \in (0, T)$ be fixed. The inverse problem considered here is to determine $p = p(x)$ and $\mu = \mu(x)$ in $\Omega$ from the measurement given by $u(x, t), \ x \in \omega, \ 0 < t < T$ and $u(x, \theta), \ x \in \Omega$. This chapter aims to establish a stability estimate for the inverse parabolic problem.

We can also take the Neumann data $\frac{\partial u}{\partial n}|_{\Gamma \times (0, T)}$ as the observation data in place of $u|_{\partial \omega \times (0, T)}$, with $\Gamma \subset \partial \Omega$ being an arbitrary relatively open subset. Bukhgeim [2] and Klibanov [20] studied this case and proved the uniqueness of the inverse problem for $\theta = 0$, provided that $\Gamma$ is a sufficiently large part of $\partial \Omega$. When $\Gamma$ does not occupy a sufficient large part of $\partial \Omega$, the uniqueness is still open. On the other hand, for $\theta > 0$, Isakov [16] proved the uniqueness of the inverse problem with the above-mentioned Neumann data on an arbitrarily small part $\Gamma$ over any time interval $(0, T)$. 

10
2.2 Stability of the inverse problems

The main result of this chapter will answer the stability of the inverse problem with the Dirichlet observation data \( u_{\omega \times (0,T)} \) within the class of sufficiently smooth solutions for the case of \( \theta > 0 \). For any \((p, \mu), (q, \nu) \in \mathcal{U} \), let

\[
E((p, \mu), (q, \nu)) = \|u(p, \mu) - u(q, \nu)\|_{H^1(\omega, L^2(\omega))} + \|(u(p, \mu) - u(q, \nu))(\cdot, \theta)\|_{H^2(\Omega)}.
\]

Our main results of this chapter are stated in the following theorem:

**Theorem 2.1** There exists a constant \( C = C(M, \Omega, T, \theta, \eta, \gamma, \omega) > 0 \) such that

\[
\|p - q\|_{L^2(\Omega)} \leq C E((p, \mu), (q, \nu)), \tag{2.3}
\]

\[
\|\mu - \nu\|_{L^2(\Omega)} \leq C \left\{ \log E((p, \mu), (q, \nu)) \right\}^{-1}, \tag{2.4}
\]

for all \((p, \mu), (q, \nu) \in \mathcal{U} \).

**Remark 2.1** The determination of initial values involves the backward parabolic equation, and so one cannot expect good stability in general. However, the determination of heat radiative coefficients is more stable within the admissible set \( \mathcal{U} \) defined by (2.2). In the case where we have the Neumann data on any part \( \Gamma \) of \( \partial \Omega \), in place of \( u_{\omega \times (0,T)} \), we can prove a similar stability result, which is omitted here.

**Proof of Theorem 2.1.** We divide the proof into four steps.

**First Step.** We first establish a Carleman estimate. For simplicity, we set

\[
Q = \Omega \times (0,T), \quad Q_\omega = \omega \times (0,T), \quad \Sigma = \partial \Omega \times (0,T)
\]

and

\[
(Pu)(x, t) = u_t(x, t) - \Delta u(x, t) + \sum_{i=1}^d a_i(x, t) \frac{\partial u}{\partial x_i} + b(x, t)u, \quad x \in \Omega, \quad t \in (0,T)
\]

where the coefficients \( \{a_i\}_{i=1}^d \) and \( b \) are any functions satisfying

\[
\|b\|_{L^\infty(Q)} \leq M_1, \quad \|a_i\|_{L^\infty(Q)} \leq M_1, \quad i = 1, 2, \ldots, d. \tag{2.5}
\]
The following lemma can be found in Fursikov and Imanuvilov [11] and Imanuvilov [15]:

**Lemma 2.1** Let $\omega_0$ be an arbitrarily fixed subdomain of $\Omega$ such that $\bar{\omega}_0 \subset \omega$. Then there exists a function $\psi \in C^2(\Omega)$ such that $\psi|_{\partial \Omega} = 0$ and

$$\psi(x) > 0, \quad x \in \Omega; \quad |\nabla \psi(x)| > 0, \quad x \in \Omega \setminus \omega_0. \quad (2.6)$$

**Example of $\psi$.** Let $\Omega = \{x; |x| < 1\}$ and $\omega_0 = \{x; |x| < \frac{1}{2}\}$. Then $\psi(x) = 1 - |x|^2$ satisfies (2.6).

Set

$$\varphi(t, x) = \frac{e^{\lambda t} - e^{2\lambda t}|\psi|_{C(\Omega)}}{t(T - t)}, \quad \alpha(t, x) = \frac{e^{\lambda t} - e^{2\lambda t}|\psi|_{C(\Omega)}}{t(T - t)}.$$

The following lemma states a Carleman estimate due to Imanuvilov [11],[15], whose detailed proof for our current problem is given in the appendix.

**Lemma 2.2** There exists $\hat{\lambda} = \hat{\lambda}(M_1) > 0$ such that for any $\lambda \geq \hat{\lambda}$ we can choose $s_0(\lambda) > 0$ and a constant $C = C(\lambda, M_1) > 0$, independent of each choice $\{a_i\}_{i=1}^d$ and $b$, such that for each $s \geq s_0(\lambda)$, the following inequality holds

$$\int_Q \left\{ \frac{1}{s\varphi} \left( |z_t|^2 + \sum_{i=1}^d \left| \frac{\partial^2 z}{\partial x_i \partial x_j} \right|^2 \right) + s \varphi |\nabla z|^2 + s \varphi^3 z^2 \right\} e^{2s\alpha} dx dt \leq C \int_Q |Pz|^2 e^{2s\alpha} dx dt + C \int_{Q_\omega} s^3 \varphi^3 z^2 e^{2s\alpha} dx dt$$

for all $z \in W^{1,2}_2(Q)$ with $z|_{\partial \Omega} = 0$.

**Proof.** Our proof is based on the argument in Fursikov and Imanuvilov [11] and Imanuvilov [15]. Here we provide a more detailed proof. Since $s > 0$ can be taken to be sufficiently large we can absorb

$$\int_Q |a_i z_{x_i}|^2 e^{2s\alpha} dx dt \quad (1 \leq i \leq d), \quad \int_Q |b z|^2 e^{2s\alpha} dx dt$$

into the left hand side of the desired estimate, so it is sufficient to assume that $a_i = 0, 1 \leq i \leq d$ and $b = 0$. 

12
We set
\[ w(t, x) = e^{\alpha z(t, x)}, \quad (t, x) \in Q \] (2.7)
and introduce the operators $L_1$ and $L_2$:
\[ L_1 w = -\Delta w - \lambda^2 s^2 \varphi^2 |\nabla \psi|^2 w - s \alpha w \] (2.8)
and
\[ L_2 w = w_t + 2s \lambda \varphi \nabla \psi \cdot \nabla w + 2s \lambda^2 \varphi |\nabla \psi|^2 w. \] (2.9)
Then by direct calculations we can verify that
\[ L_1 w + L_2 w = f, \quad \text{in} \ Q \] (2.10)
where
\[ Pz = g, \quad f_z(t, x) = g e^{\alpha z} - s \lambda \varphi w \Delta \psi + s \lambda^2 \varphi |\nabla \psi|^2 w. \] (2.11)
In fact, from $z_t - \Delta z = g$, we have
\[
ge^{\alpha z} = e^{\alpha P(e^{-\alpha} w)}
= -s \alpha w + w_t - (s^2 |\nabla \alpha|^2 - s \Delta \alpha) w + 2s \nabla \alpha \cdot \nabla w - \Delta w
\]
\[
= -s \alpha w + w_t + (-s^2 \lambda^2 \varphi^2 |\nabla \psi|^2 + s \lambda (\Delta \psi) \varphi + s \lambda^2 |\nabla \psi|^2 \varphi) w
+ 2s \lambda \varphi \nabla \psi \cdot \nabla w - \Delta w,
\]
which gives (2.10). Taking the $L^2$-norms of the both hand sides of (2.10), we obtain
\[ \|f_z\|^2_{L^2(Q)} = \|L_1 w\|^2_{L^2(Q)} + \|L_2 w\|^2_{L^2(Q)} + 2(L_1 w, L_2 w)_{L^2(Q)}. \] (2.12)
We next calculate $(L_1 w, L_2 w)_{L^2(Q)}$. We have
\[
(L_1 w, L_2 w)_{L^2(Q)}
= - \int_Q (\Delta w + \lambda^2 s^2 \varphi^2 |\nabla \psi|^2 w + s \alpha w)(w_t + 2s \lambda \varphi \nabla \psi \cdot \nabla w + 2s \lambda^2 \varphi |\nabla \psi|^2 w) dx dt
\]
\[
= - \int_Q (\Delta w + \lambda^2 s^2 \varphi^2 |\nabla \psi|^2 w + s \alpha w)(w_t + 2s \lambda^2 \varphi |\nabla \psi|^2 w) dx dt
\]
\[- \int_Q (\lambda^2 s^2 \varphi^2 |\nabla \psi|^2 w + s \omega_t w) \times 2s \lambda \varphi \nabla \psi \cdot \nabla w \, dx \, dt \]
\[- \int_Q \Delta w \times 2s \lambda \varphi \nabla \psi \cdot \nabla w \, dx \, dt \]
\[\equiv A_1 + A_2 + A_3. \] (2.13)

By integration by parts and using the condition $w|_{\partial \Omega} = 0$, we have

\[A_1 = - \int_Q w_t \Delta w \, dx \, dt - \int_Q (\lambda^2 s^2 \varphi^2 |\nabla \psi|^2 + s \omega_t) w \, dx \, dt \]
\[- \int_Q 2s \lambda \varphi |\nabla \psi|^2 w \Delta w \, dx \, dt - \int_Q (\lambda^2 s^2 \varphi^2 |\nabla \psi|^2 + s \omega_t) 2s \lambda \varphi |\nabla \psi|^2 w^2 \, dx \, dt \]
\[= \int_Q \nabla w \cdot \nabla w_t \, dx \, dt - \int_Q (\lambda^2 s^2 \varphi^2 |\nabla \psi|^2 + s \omega_t) \frac{1}{2} (w^2)_t \, dx \, dt \]
\[+ \int_Q \{2s \lambda \varphi |\nabla \psi|^2 |\nabla w|^2 + \nabla (2s \lambda \varphi |\nabla \psi|^2) \cdot w \nabla w \} \, dx \, dt \]
\[- \int_Q (2s^3 \lambda^4 \varphi^3 |\nabla \psi|^4 + 2s^2 \lambda^2 \varphi \omega_t |\nabla \psi|^2) w^2 \, dx \, dt. \]

and

\[\int_Q \nabla (2s \lambda \varphi |\nabla \psi|^2) \cdot (w \nabla w) \, dx \, dt = \int_Q \nabla (s \lambda^2 \varphi |\nabla \psi|^2) \cdot \nabla (w^2) \, dx \, dt \]
\[= - \int_Q s \lambda^2 \Delta (\varphi |\nabla \psi|^2) w^2 \, dx \, dt \]

By integration by parts again and using the conditions $w(\cdot, 0) = w(\cdot, T) = 0$, we obtain

\[A_1 = \int_Q \frac{1}{2} (\lambda^2 s^2 \partial_t (\varphi^2 |\nabla \psi|^2) + s \omega_t) w^2 \, dx \, dt \]
\[+ \int_Q \{2s \lambda^2 \varphi |\nabla \psi|^2 |\nabla w|^2 - s \lambda^2 \Delta (\varphi |\nabla \psi|^2) w^2 \} \, dx \, dt \]
\[- \int_Q (2s^3 \lambda^4 \varphi^3 |\nabla \psi|^4 + 2s^2 \lambda^2 \varphi \omega_t |\nabla \psi|^2) w^2 \, dx \, dt. \] (2.14)

For the term $A_2$, we also apply integration by parts and using the condition that $w|_{\partial \Omega} = 0$ to derive

\[A_2 = - \int_Q 2s^3 \lambda^3 \varphi^3 |\nabla \psi|^2 \nabla \psi \cdot w (\nabla w) \, dx \, dt - \int_Q 2s^2 \lambda \varphi \omega_t \nabla \psi \cdot (w \nabla w) \, dx \, dt \]
\[
= - \int_Q s^3 \lambda^3 \varphi^3 |\nabla \psi|^2 \nabla \psi \cdot \nabla (w^2) \, dx \, dt - \int_Q s^2 \lambda \varphi \alpha \nabla \psi \cdot \nabla (w^2) \, dx \, dt
\]
\[
= \int_Q s^3 \lambda^3 \nabla \cdot \left( \varphi^3 |\nabla \psi|^2 \nabla \psi \right) w^2 \, dx \, dt + \int_Q s^2 \lambda \nabla \cdot \left( \varphi \nabla \psi \right) w^2 \, dx \, dt
\]
\[
= \int_Q \left\{ 3s^2 \lambda^3 \varphi^3 |\nabla \psi|^4 w^2 + s^3 \lambda^3 \varphi^3 w^2 \nabla \cdot (|\nabla \psi|^2 \nabla \psi) \right\} \, dx \, dt
\]
\[
+ \int_Q s^2 \lambda \nabla \cdot \left( \varphi \nabla \psi \right) w^2 \, dx \, dt. \tag{2.15}
\]

Finally for the term \(A_3\), we can still use integration by parts and obtain

\[
A_3 = - \sum_{i,j=1}^d \int_Q \frac{\partial^2 w}{\partial x_i^2} \left\{ 2s \lambda \varphi \frac{\partial \psi}{\partial x_j} \right\} \frac{\partial w}{\partial x_j} \, dx \, dt
\]
\[
+ \sum_{i,j=1}^d \int_Q \frac{\partial w}{\partial x_i} \frac{\partial C_j}{\partial x_i} \frac{\partial w}{\partial x_j} \, dx \, dt + \sum_{i,j=1}^d \int_Q \frac{\partial w}{\partial x_i} C_j \frac{\partial^2 w}{\partial x_i \partial x_j} \, dx \, dt
\]
\[
- \sum_{i,j=1}^d \int_{\Sigma} \frac{\partial w}{\partial x_i} \left( 2s \lambda \varphi \frac{\partial \psi}{\partial x_j} \frac{\partial w}{\partial x_j} \nu_i \right) \, d\sigma \tag{2.16}
\]

Here \(n = (\nu_1, ..., \nu_d)\) is the unit outward normal vector to \(\partial \Omega\). Set

\[
C_j = 2s \lambda \varphi \frac{\partial \psi}{\partial x_j}, \quad 1 \leq i \leq d.
\]

By \(\psi|_{\partial \Omega} = 0\), we have

\[
\frac{\partial \psi}{\partial x_j} = \frac{\partial \psi}{\partial n} \nu_j, \quad 1 \leq j \leq d.
\]

Using this we can write the last term of (2.16) as

\[
- \sum_{i,j=1}^d \int_{\Sigma} \frac{\partial w}{\partial x_i} \left( 2s \lambda \varphi \frac{\partial \psi}{\partial x_j} \frac{\partial w}{\partial x_j} \nu_i \right) \, d\sigma = - \int_{\Sigma} 2s \lambda \varphi \frac{\partial \psi}{\partial n} \left( \frac{\partial w}{\partial n} \right)^2 \, d\sigma. \tag{2.17}
\]

The second term of (2.16) can be estimated by integration by parts as follows:

\[
\sum_{i,j=1}^d \int_Q \frac{\partial w}{\partial x_i} C_j \frac{\partial^2 w}{\partial x_i \partial x_j} \, dx \, dt = \sum_{i,j=1}^d \int_Q \frac{1}{2} \frac{\partial}{\partial x_i} \left\{ \left( \frac{\partial w}{\partial x_i} \right)^2 \right\} C_j \, dx \, dt
\]
\[
- \sum_{i,j=1}^d \int_Q \frac{1}{2} \left( \frac{\partial w}{\partial x_i} \right)^2 \frac{\partial C_j}{\partial x_j} \, dx \, dt + \sum_{i,j=1}^d \int_{\Sigma} \frac{1}{2} \left( \frac{\partial w}{\partial x_i} \right)^2 C_j \nu_j \, d\sigma
\]
\[
= - \int_Q s \lambda |\nabla w|^2 \sum_{j=1}^d \frac{\partial}{\partial x_j} \left( \varphi \frac{\partial \psi}{\partial x_j} \right) \, dx \, dt + \int_{\Sigma} s \lambda \varphi \frac{\partial \psi}{\partial n} |\nabla w|^2 \, d\sigma
\]

15
\[
\begin{align*}
&= - \int_Q (s\lambda^2 \varphi |\nabla w|^2 |\nabla \psi|^2 + s\lambda \varphi (\Delta \psi) |\nabla w|^2) \, dx \, dt \\
&\quad + \int \sum_{i,j=1}^d \lambda \frac{\partial^2 \psi}{\partial x_i \partial x_j} |\nabla w|^2 \, d\sigma.
\end{align*}
\] (2.18)

By the definition of \( C_j \), we can write the first term of (2.16) as
\[
\begin{align*}
&= \sum_{i,j=1}^d \int_{Q} \frac{\partial w}{\partial x_i} \frac{\partial C_j}{\partial x_j} \, dx \, dt \\
&= \sum_{i,j=1}^d \int_{Q} 2s\lambda \left( \lambda \frac{\partial \psi}{\partial x_i} \frac{\partial \psi}{\partial x_j} + \varphi \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right) \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_j} \, dx \, dt \\
&= \sum_{i,j=1}^d \int_{Q} 2s\lambda^2 \varphi (\nabla \psi \cdot \nabla w)^2 \, dx \, dt \\
&\quad + \sum_{i,j=1}^d \int_{Q} 2s\lambda \varphi \frac{\partial^2 \psi}{\partial x_i \partial x_j} \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_j} \, dx \, dt.
\end{align*}
\] (2.19)

Since \( w_{|\partial Q} = 0 \), we have \( |\nabla w| = |\frac{\partial w}{\partial n}| \). Then summing up (2.17)-(2.19), we have
\[
\begin{align*}
A_3 &= - \int_Q (s\lambda^2 \varphi |\nabla w|^2 |\nabla \psi|^2 + s\lambda \varphi (\Delta \psi) |\nabla w|^2) \, dx \, dt \\
&\quad + \sum_{i,j=1}^d \int_{Q} 2s\lambda \varphi \frac{\partial^2 \psi}{\partial x_i \partial x_j} \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_j} \, dx \, dt - \int \sum_{i,j=1}^d s\lambda \varphi \frac{\partial \psi}{\partial n} \frac{\partial w}{\partial n} \, d\sigma. 
\end{align*}
\] (2.20)

Using the above estimates for \( A_1, A_2 \) and \( A_3 \), we obtain
\[
(L_1 w, L_2 w)_{L^2(Q)}
\]
\[
= \int_Q (s^3 \lambda^4 \varphi^3 |\nabla \psi|^4 w^2 + s\lambda^2 \varphi |\nabla \psi|^2 |\nabla w|^2) \, dx \, dt + \int Q 2s\lambda^2 \varphi (\nabla \psi \cdot \nabla w)^2 \, dx \, dt \\
&\quad - \int \sum_{i,j=1}^d s\lambda \varphi \left( \frac{\partial \psi}{\partial n} \frac{\partial w}{\partial n} \right)^2 \, d\sigma + X_1,
\] (2.21)

where
\[
\begin{align*}
X_1 &= \int_Q \left\{ \frac{1}{2} \left( \frac{\partial^2 \varphi}{\partial t^2} (\varphi^2 |\nabla \psi|^2) + s\alpha \varphi \right) w^2 - 2s\lambda^2 \alpha \varphi |\nabla \psi|^2 \frac{\partial w}{\partial \psi} \right. \\
&\quad + s^3 \lambda^2 \varphi \nabla \cdot (|\nabla \psi|^2 \nabla \psi) w^2 + s\lambda \nabla \cdot (\varphi \alpha \nabla \psi) w^2 \\
&\quad - s\lambda \varphi (\Delta \psi) |\nabla w|^2 - s\lambda \Delta (\varphi |\nabla \psi|^2) \frac{\partial w}{\partial \psi} \right. \\
&\quad + 2s\lambda \varphi \sum_{i,j=1}^d \frac{\partial^2 \psi}{\partial x_i \partial x_j} \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_j} \right\} \, dx \, dt.
\] (2.22)
Henceforth $C > 0$ and $c_0 > 0$ denote generic constants which are independent of $\lambda > 0$ and $s > 0$.

Now we estimate $|X_1|$. It is easy to verify that

$$f < \alpha t \leq C \varphi$$

which imply

$$|\alpha_t| < C \varphi, \quad |\alpha_t| < C \varphi^2, \quad |\alpha_t| < C \varphi^3,$$

Consequently,

$$|X_1| \leq C \int_\{ (\lambda^3 s^3 \varphi^3 + \lambda^4 s \varphi)w^2 + s \lambda \varphi |\nabla w|^2 \}dxdt$$

for sufficiently large $s$ and $\lambda$.

On the other hand, since $|\nabla \psi| \geq c_0$ on $\Omega \setminus \omega$ and $\frac{\partial \psi}{\partial n} < 0$ on $\partial \Omega$, using (2.25) we see from (2.21) that

$$(L_1 w, L_2 w)_{L^2(\Omega)} \geq \int_{Q \setminus Q_\omega} (s^3 \lambda^4 \varphi^3 c_0^4 w^2 + s \lambda^2 \varphi c_0^2 |\nabla w|^2)dxdt$$

Therefore we obtain

$$(L_1 w, L_2 w)_{L^2(\Omega)} + \int_{Q_\omega} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2)dxdt$$

Consequently,

$$|X_1| \leq C \int_\{ (\lambda^3 s^3 \varphi^3 + \lambda^4 s \varphi)w^2 + s \lambda \varphi |\nabla w|^2 \}dxdt$$

for sufficiently large $s$ and $\lambda$.

On the other hand, since $|\nabla \psi| \geq c_0$ on $\Omega \setminus \omega$ and $\frac{\partial \psi}{\partial n} < 0$ on $\partial \Omega$, using (2.25) we see from (2.21) that

$$(L_1 w, L_2 w)_{L^2(\Omega)} \geq \int_{Q \setminus Q_\omega} (s^3 \lambda^4 \varphi^3 c_0^4 w^2 + s \lambda^2 \varphi c_0^2 |\nabla w|^2)dxdt$$

Therefore we obtain

$$(L_1 w, L_2 w)_{L^2(\Omega)} + \int_{Q_\omega} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2)dxdt$$

Consequently,

$$|X_1| \leq C \int_\{ (\lambda^3 s^3 \varphi^3 + \lambda^4 s \varphi)w^2 + s \lambda \varphi |\nabla w|^2 \}dxdt$$

for sufficiently large $s$ and $\lambda$.

On the other hand, since $|\nabla \psi| \geq c_0$ on $\Omega \setminus \omega$ and $\frac{\partial \psi}{\partial n} < 0$ on $\partial \Omega$, using (2.25) we see from (2.21) that

$$(L_1 w, L_2 w)_{L^2(\Omega)} \geq \int_{Q \setminus Q_\omega} (s^3 \lambda^4 \varphi^3 c_0^4 w^2 + s \lambda^2 \varphi c_0^2 |\nabla w|^2)dxdt$$

Therefore we obtain

$$(L_1 w, L_2 w)_{L^2(\Omega)} + \int_{Q_\omega} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2)dxdt$$

Consequently,

$$|X_1| \leq C \int_\{ (\lambda^3 s^3 \varphi^3 + \lambda^4 s \varphi)w^2 + s \lambda \varphi |\nabla w|^2 \}dxdt$$

for sufficiently large $s$ and $\lambda$.

On the other hand, since $|\nabla \psi| \geq c_0$ on $\Omega \setminus \omega$ and $\frac{\partial \psi}{\partial n} < 0$ on $\partial \Omega$, using (2.25) we see from (2.21) that

$$(L_1 w, L_2 w)_{L^2(\Omega)} \geq \int_{Q \setminus Q_\omega} (s^3 \lambda^4 \varphi^3 c_0^4 w^2 + s \lambda^2 \varphi c_0^2 |\nabla w|^2)dxdt$$

Therefore we obtain

$$(L_1 w, L_2 w)_{L^2(\Omega)} + \int_{Q_\omega} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2)dxdt$$

Consequently,

$$|X_1| \leq C \int_\{ (\lambda^3 s^3 \varphi^3 + \lambda^4 s \varphi)w^2 + s \lambda \varphi |\nabla w|^2 \}dxdt$$

for sufficiently large $s$ and $\lambda$.
for $\lambda > 0$ and $s > 0$ sufficiently large. Hence by (2.12) we have

$$\|f_s\|^2_{L^2(Q)} \geq \frac{1}{2}\|f_s\|^2_{L^2(Q)} \geq (L_1w, L_2w)_{L^2(Q)}.$$  

This with (2.27) leads to

$$0 \leq (L_1w, L_2w)_{L^2(Q)} + \int_{Q_\omega} (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt$$

$$\leq \|f_s\|^2_{L^2(Q)} + \int_{Q_\omega} (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt. \quad (2.28)$$

Furthermore, it follows from (2.11), (2.27) and (2.28) that

$$\int_Q (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt$$

$$\leq \|f_s\|^2_{L^2(Q)} + \int_{Q_\omega} (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt$$

$$\leq C \int_Q (s^2\lambda^2\varphi^2|\Delta \psi|^2w^2 + s^3\lambda^4\varphi^3|\nabla \psi|^4w^2)dxdt + C \int_Q g^2e^{2s\alpha}dxdt$$

$$+ C \int_{Q_\omega} (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt.$$  

If we take sufficiently large $s$ and $\lambda$, then the first term on the right hand side can be absorbed into the left hand side, and so

$$\int_Q (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt$$

$$\leq C \int_Q g^2e^{2s\alpha}dxdt + C \int_{Q_\omega} (s^3\lambda^4\varphi^3w^2 + s\lambda^2\varphi|\nabla w|^2)dxdt. \quad (2.29)$$

Next we will estimate $\Delta w$ and $w_t$. First by (2.8), we have

$$\Delta w = -L_1w - \lambda^2s^2\varphi^2|\nabla \psi|^2w - s\alpha_t w.$$  

Hence using (2.23) and the fact that $1/(s\varphi) \leq C$, we can deduce

$$\int_Q \frac{1}{s\varphi}|\Delta w|^2dxdt = \int_Q \frac{1}{s\varphi}((L_1w + \lambda^2s^2\varphi^2|\nabla \psi|^2w + s\alpha_t w)^2dxdt$$

$$\leq C \int_Q (L_1w)^2dxdt + C \int_Q (s^3\lambda^4s^3\varphi^3 + s\varphi^3)w^2dxdt. \quad (2.30)$$

While for $w_t$, we have from (2.9) that

$$w_t = L_2w - 2s\lambda\varphi \nabla \psi \cdot \nabla w - 2s\lambda^2\varphi|\nabla \psi|^2w,$$
Thus
\[
\int_Q \frac{1}{s \varphi} |w_t|^2 \, dx \, dt = \int_Q \frac{1}{s \varphi} (L_2 w - 2s \lambda \varphi \nabla \psi \cdot \nabla w - 2s \lambda^2 \varphi |\nabla \psi|^2 w)^2 \, dx \, dt
\]
\[
\leq C \int_Q |L_2 w|^2 \, dx \, dt + C \int_Q (s \lambda^2 \varphi |\nabla w|^2 + s \lambda^4 \varphi w^2) \, dx \, dt.
\] (2.31)

On the other hand, by (2.28), we have
\[
|(L_1 w, L_2 w)_{L^2(Q)}| \leq \|f_3\|_{L^2(Q)}^2 + 2 \int_{Q_w} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2) \, dx \, dt.
\]
This with (2.12) gives
\[
\|L_1 w\|_{L^2(Q)}^2 + \|L_2 w\|_{L^2(Q)}^2
\leq 3 \|f_3\|_{L^2(Q)}^2 + 4 \int_{Q_w} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2) \, dx \, dt.
\] (2.32)

From its definition, \(f_3\) can be bounded by
\[
\|f_3\|_{L^2(Q)}^2 \leq C \int_Q g^2 e^{2s \alpha} \, dx \, dt
\]
\[
+ C \int_Q s^2 \lambda^2 \varphi^2 |\Delta \psi|^2 w^2 \, dx \, dt + C \int_Q s^2 \lambda^4 \varphi^2 |\nabla \psi|^2 w^2 \, dx \, dt.
\]
Substituting into (2.32) yields
\[
\|L_1 w\|_{L^2(Q)}^2 + \|L_2 w\|_{L^2(Q)}^2
\leq C \int_Q g^2 e^{2s \alpha} \, dx \, dt + C \int_{Q_w} (s^3 \lambda^4 \varphi^3 w^2 + s \lambda^2 \varphi |\nabla w|^2) \, dx \, dt
\]
\[
+ C \int_Q (s^3 \lambda^4 \varphi^3 |\Delta \psi|^2 + s \lambda^4 \varphi^2 |\nabla \psi|^2) w^2 \, dx \, dt.
\] (2.33)

Henceforth we fix a sufficiently large \(\lambda > 0\). Then for sufficiently large \(s > 0\) we obatin from (2.29)-(2.31) and (2.33) that
\[
\int_Q \frac{1}{s \varphi} (|\Delta w|^2 + |w_t|^2) \, dx \, dt
\]
\[
\leq C (\|L_1 w\|_{L^2(Q)}^2 + \|L_2 w\|_{L^2(Q)}^2) + C \int_Q (s \varphi |\nabla w|^2 + s^3 \lambda^2 \varphi w^2) \, dx \, dt
\]
\[
\leq C \int_Q g^2 e^{2s \alpha} \, dx \, dt + C \int_Q s^2 \varphi w^2 \, dx \, dt
\]
\[
+ C \int_{Q_w} (s \varphi |\nabla w|^2 + s^3 \lambda^2 \varphi w^2) \, dx \, dt.
\] (2.34)
It is easy to check that

\[ C^{-1} \varphi_0(t) \leq \varphi(x, t) \leq C \varphi_0(t), \quad (x, t) \in \overline{Q} \]

with \( \varphi_0(t) = 1/(t(T-t)) \), and

\[ \int_{\Omega} \sum_{i,j=1}^{d} \left( \frac{\partial^2 w}{\partial x_i \partial x_j} \right)^2 \, dx \leq C \int_{Q} \left( |\Delta w|^2 + |w|^2 \right) dx \]

using the uniform ellipticity of \(-\Delta\), thus we have

\[ \int Q \frac{1}{s\varphi} \sum_{i,j=1}^{d} \left( \frac{\partial^2 w}{\partial x_i \partial x_j} \right)^2 \, dx \, dt \leq C \int_{0}^{T} \frac{1}{s\varphi_0} \left( \int_{\Omega} \sum_{i,j=1}^{d} \left( \frac{\partial^2 w}{\partial x_i \partial x_j} \right)^2 \, dx \right) \, dt \]

\[ \leq C \int_{0}^{T} \frac{1}{s\varphi} \left( \int_{\Omega} \left( |\Delta w|^2 + |w|^2 \right) \, dx \right) \, dt \]

\[ = C \int_{Q} \frac{1}{s\varphi} \left( |\Delta w|^2 + |w|^2 \right) \, dx \, dt. \]

Combining with the estimate (2.34) yields

\[ \int Q \frac{1}{s\varphi} \left( \sum_{i,j=1}^{d} \left| \frac{\partial^2 w}{\partial x_i \partial x_j} \right|^2 + |w_t|^2 \right) \, dx \, dt \]

\[ \leq C \int Q g^2 e^{2\alpha} \, dx \, dt + \int Q \left( \frac{1}{s\varphi} + s^2 \varphi^2 \right) w^2 \, dx \, dt \]

\[ + C \int_{Q} (s\varphi |\nabla w|^2 + s^3 \varphi^3 w^2) \, dx \, dt. \]  

(2.35)

Summing (2.29) and (2.35), and taking \( s > 0 \) sufficiently large so that we can absorb \( \int Q \left( \frac{1}{s\varphi} + s^2 \varphi^2 \right) w^2 \, dx \, dt \) into the term \( \int Q (s\varphi)^3 w^2 \, dx \, dt \), we obtain

\[ \int Q \left\{ \frac{1}{s\varphi} \left( \sum_{i,j=1}^{d} \left| \frac{\partial^2 w}{\partial x_i \partial x_j} \right|^2 + |w_t|^2 \right) + s\varphi |\nabla w|^2 + s^3 \varphi^3 w^2 \right\} \, dx \, dt \]

\[ \leq C \int Q g^2 e^{2\alpha} \, dx \, dt + C \int_{Q} (s\varphi |\nabla w|^2 + s^3 \varphi^3 w^2) \, dx \, dt. \]  

(2.36)

We next change the function \( w \) back to \( z \). Recall that \( z = e^{-\alpha} w \), then

\[ z_{x_i} = -s\alpha x_i e^{-\alpha} w + e^{-\alpha} w_{x_i}, \quad z_t = -s\alpha e^{-\alpha} w + e^{-\alpha} w_t \]

and

\[ z_{x_ix_j} = (-s\alpha x_i x_j + s^2 \alpha x_i x_j) e^{-\alpha} w - s(\alpha x_i w_{x_j} + \alpha x_j w_{x_i}) e^{-\alpha} + e^{-\alpha} w_{x_i x_j}. \]
Noting that
\[ |\alpha_{z_i}| \leq C|\varphi|, \quad |\alpha_{z_i z_j}| \leq C|\varphi|, \quad |\alpha_t| \leq C\varphi^2 \] (2.37)
for fixed \( \lambda > 0 \), we see
\[ \begin{cases} |z_{z_i}|^2 e^{2s\alpha} \leq Cs^2 \varphi^2 w^2 + C|\nabla w|^2, \\ |z_t|^2 e^{2s\alpha} \leq Cs^2 \varphi^4 w^2 + C|w_t|^2, \\ |z_{z_i z_j}|^2 e^{2s\alpha} \leq Cs^4 \varphi^2 w^2 + Cs^2 \varphi^2 |\nabla w|^2 + C|w_{z_i z_j}|^2. \end{cases} \] (2.38)
Again by (2.37), for large \( s > 0 \), we have
\[ \begin{cases} |w|^2 = e^{2s\alpha} |z|^2, \\ |\nabla w|^2 \leq C|\nabla z|^2 e^{2s\alpha} + Cs^2 \varphi^2 |z|^2 e^{2s\alpha}, \\ |w_t|^2 \leq C|z_t|^2 e^{2s\alpha} + Cs^4 \varphi^4 |z|^2 e^{2s\alpha}, \\ |w_{z_i z_j}|^2 \leq C|z_{z_i z_j}|^2 e^{2s\alpha} + Cs^4 \varphi^2 |z|^2 e^{2s\alpha} + Cs^2 \varphi^2 |\nabla z|^2 e^{2s\alpha}. \end{cases} \] (2.39)
Applying (2.38) and (2.39) in both sides of (2.36), we obtain
\[ \int_Q \left\{ \frac{1}{s^2 \varphi} \left( \sum_{i,j=1}^d \left| \frac{\partial^2 z}{\partial x_i \partial_j} \right|^2 + |z_t|^2 \right) + s \varphi |\nabla z|^2 + s^3 \varphi^3 z^2 \right\} e^{2s\alpha} \, dx \, dt \]
\[ \leq C \int_Q s \varphi |\nabla z|^2 e^{2s\alpha} \, dx \, dt + C \int_{Q_{\omega}} (s^3 \varphi^3 z^2 + s \varphi |\nabla z|^2) e^{2s\alpha} \, dx \, dt. \] (2.40)
Finally, to conclude Lemma 2.2, we need to remove the term
\[ \int_{Q_{\omega}} s \varphi |\nabla z|^2 e^{2s\alpha} \, dx \, dt \]
from (2.40). Let us choose \( \rho \in C^\infty_0(\omega) \) such that \( \rho|_\omega = 1, 0 \leq \rho \leq 1 \) with \( \overline{\omega} \subset \omega_1 \).
We multiply the both sides of \( z_t - \Delta z - g = 0 \) with \( s \varphi z e^{2s\alpha} \) and integrate in \( x \) and \( t \) over \( Q \), then we have
\[ \left| \int_Q s \varphi z e^{2s\alpha} \rho \Delta z \, dx \, dt \right| \leq \left| \int_Q s \varphi z e^{2s\alpha} z_t \, dx \, dt \right| + \left| \int_Q s \varphi z e^{2s\alpha} g \, dx \, dt \right|. \] (2.41)
But using (2.23), integration by parts and Schwarz's inequality the right hand side of (2.41) can be estimated as follows:
\[ \left| \int_Q s \varphi z e^{2s\alpha} z_t \, dx \, dt \right| \]
\[ \begin{align*}
&= \left| \frac{1}{2} \int_{Q} \rho \varphi e^{2s \alpha} \frac{\partial}{\partial t} z^2 dxdt \right| \\
&= \left| \frac{1}{2} \int_{\Omega} s \rho \varphi e^{2s \alpha} z^2 |z|^2 dx - \frac{1}{2} \int_{Q} s \rho (\varphi e^{2s \alpha} + 2s \alpha \varphi e^{2s \alpha}) z^2 dxdt \right| \\
&\leq C \int_{Q} s^2 \varphi^3 z^2 e^{2s \alpha} dxdt \tag{2.42}
\end{align*} \]

where we have used the fact that \( \alpha(\cdot, 0) = \alpha(\cdot, T) = -\infty \), and

\[ \left| \int_{Q} s \rho \varphi z e^{2s \alpha} g dxdt \right| \leq \int_{Q} s^2 \varphi^2 z^2 e^{2s \alpha} dxdt + \int_{Q} g^2 e^{2s \alpha} dxdt. \tag{2.43} \]

On the other hand, by (2.23) and integration by parts, the left hand side of (2.41) can be bounded below as follows:

\[ \begin{align*}
\left| \int_{Q} s \rho \varphi e^{2s \alpha} \Delta z dxdt \right| &= \left| - \int_{Q} s \nabla (\varphi e^{2s \alpha} \rho) \cdot \nabla z dxdt \right| \\
&= \left| \int_{Q} s \varphi e^{2s \alpha} |\nabla z|^2 dxdt - \int_{Q} s \rho e^{2s \alpha} \varphi \nabla z \cdot z dxdt \\
- \int_{Q} s \varphi e^{2s \alpha} z \rho \cdot \nabla z dxdt - \int_{Q} s \varphi \rho \nabla (e^{2s \alpha}) \cdot \nabla z dxdt \right| \\
&\geq C \int_{Q_{\omega}} s \varphi e^{2s \alpha} |\nabla z|^2 dxdt - C \int_{Q} s \nabla |e^{2s \alpha} | \nabla z| e^{2s \alpha} dxdt \\
&- C \int_{Q_{\omega}} s^3 \varphi^3 |z| e^{2s \alpha} \nabla z |e^{2s \alpha} z dxdt \right|.
\end{align*} \]

where we have used the following bounds

\[ \left| \nabla \varphi \right| \leq C \varphi, \quad \left| \nabla (e^{2s \alpha}) \right| = 2s e^{2s \alpha} \nabla \alpha \leq C e^{2s \alpha} s \varphi \]

for fixed \( \lambda > 0 \). Consequently applying the Schwarz’s inequality and Young’s inequality: \( |ab| \leq \varepsilon a^2 + b^2/(4\varepsilon) \), the left hand side of (2.41) can be further bounded below by

\[ \begin{align*}
\left| \int_{Q} s \rho \varphi e^{2s \alpha} \Delta z dxdt \right| &\geq C \int_{Q_{\omega}} s \varphi e^{2s \alpha} |\nabla z|^2 dxdt \\
- C \int_{Q} s^2 \varphi^2 z^2 e^{2s \alpha} dxdt - C \int_{Q} |\nabla z|^2 e^{2s \alpha} dxdt \end{align*} \]
for any $\varepsilon > 0$.

Now by (2.42)-(2.44), we obtain

$$\int_{Q_s} s\varphi|\nabla z|^2 e^{2s_\alpha} dx dt \leq C(\varepsilon) \left( \int_{Q_s} s^3 \varphi^3 z^2 e^{2s_\alpha} dx dt + \|g e^{s_\alpha}\|_{L^2(Q)}^2 \right)$$

$$+ C\varepsilon \int_{Q_s} s\varphi|\nabla z|^2 e^{2s_\alpha} dx dt + C \int_{Q_s} |\nabla z|^2 e^{2s_\alpha} dx dt$$

for $s > 0$ sufficiently large. Applying (2.45) to (2.40) with $\omega = \omega_0$, we derive

$$\int_{Q_s} \left\{ \frac{1}{s\varphi} \left( \sum_{i,j=1}^{d} \left| \frac{\partial^2 z}{\partial x_i \partial x_j} \right|^2 + |z|^2 \right) + s\varphi|\nabla z|^2 + s^3 \varphi^3 z^2 \right\} e^{2s_\alpha} dx dt$$

$$\leq C \int_{Q_s} g^2 e^{2s_\alpha} dx dt + C \int_{Q_s} s^3 \varphi^3 z^2 e^{2s_\alpha} dx dt$$

$$+ C(\varepsilon) \int_{Q_s} s^3 \varphi^3 z^2 e^{2s_\alpha} dx dt + C \int_{Q_s} |\nabla z|^2 e^{2s_\alpha} dx dt$$

$$+ C\varepsilon \int_{Q_s} s\varphi|\nabla z|^2 e^{2s_\alpha} dx dt.$$
and
\[
\begin{cases}
\begin{aligned}
u_t(x, t) &= \Delta u(x, t) + p(x)u(x, t), \quad x \in \Omega, \quad -\delta < t < T \\
u(x, -\delta) &= \mu(x), \quad x \in \Omega \\
u(x, t) &= \eta(x, t), \quad x \in \partial \Omega, \quad -\delta < t < T.
\end{aligned}
\end{cases}
\tag{2.47}
\]

Then

Lemma 2.3 There exists a constant \( M_2 = M_2(\mathcal{U}, \delta, T, \Omega) > 0 \) such that
\[
\|u(p, \mu)\|_{C^{\gamma+3, \frac{\gamma+3}{2}(\Omega \times [-\delta, T])}} \leq M_2, \quad \forall (p, \mu) \in \mathcal{U}.
\]

Proof. Since \((p, \mu) \in \mathcal{U}\) implies the compatibility condition of the first order, Lemma 2.3 follows from the classical parabolic theory (cf. [10], [23]).

Third Step. We now prove (2.3). Recall that \( \theta = \frac{T}{2} \). Let
\[
\begin{align*}
w &= u(p, \mu) - u(q, \nu), \quad f = p - q, \quad R = u(q, \nu), \\
a &= u(p, \mu)(\cdot, \theta) - u(q, \nu)(\cdot, \theta).
\end{align*}
\tag{2.48}
\]

Then it is easy to see that
\[
\begin{cases}
w_t &= \Delta w + pw + f R, \quad x \in \Omega, \quad -\delta < t < T \\
w(x, \theta) &= a(x), \quad x \in \Omega \\
w &= 0 \quad \text{on} \quad \partial \Omega \times (-\delta, T).
\end{cases}
\tag{2.49}
\]

By (2.46), we can set
\[
z = \frac{w}{R}.
\]

Then, noting \( \frac{\partial R}{\partial t} = \Delta R + qR \), we obtain
\[
\begin{cases}
z_t &= \Delta z + \frac{2qR}{R} \cdot \nabla z + (p - q)z + f \quad \text{in} \quad \Omega \times (-\delta, T) \\
z(x, \theta) &= \frac{a(x)}{R(x, \theta)}, \quad x \in \Omega \\
z &= 0 \quad \text{on} \quad \partial \Omega \times (-\delta, T).
\end{cases}
\tag{2.50}
\]

Then setting
\[
y = z_t,
\tag{2.51}
\]
we have

\[
\begin{align*}
\begin{cases}
y_t = \Delta y + \frac{2VR}{R} \cdot \nabla y + (p - q)y + \frac{\partial}{\partial t} \left( \frac{2VR}{R} \right) \cdot \nabla z & \text{in } \Omega \times (-\delta, T) \\
y = 0 & \text{on } \partial \Omega \times (-\delta, T).
\end{cases}
\end{align*}
\]  
(2.52)

By Lemma 2.3, we see that \(\|\frac{2VR}{R}\|_{L^\infty(Q)}\) is uniformly bounded for \((q, \mu) \in \mathcal{U}\), and we apply Lemma 2.2 to (2.52) to obtain

\[
\begin{align*}
&\int_Q \left( \frac{1}{s^p \varphi} \left( |y_t|^2 + |\Delta y|^2 \right) + s^p \varphi |\nabla y|^2 + s^3 \varphi y^2 \right) e^{2s_\alpha} dx dt \\
&\leq C \int_Q \left| \frac{\partial}{\partial t} \left( \frac{2VR}{R} \right) \cdot \nabla z \right|^2 e^{2s_\alpha} dx dt + C \int_Q s^3 \varphi y^2 e^{2s_\alpha} dx dt
\end{align*}
\]  
(2.53)

for large \(s > 0\).

We next estimate the first term of the right hand side of (2.53). Using \(z(x, \theta) = a(x)/R(x, \theta)\), we have

\[
z(x, t) = \int_0^t y(x, \xi) d\xi + \frac{a(x)}{R(x, \theta)}.
\]  
(2.54)

then by Lemma 2.3,

\[
\begin{align*}
&\int_Q \left| \frac{\partial}{\partial t} \left( \frac{2VR}{R} \right) \cdot \nabla z \right|^2 e^{2s_\alpha} dx dt \\
&\leq C_1 \int_0^t \left| \nabla y(x, \xi) \right|^2 e^{2s_\alpha} dx dt + C_1 \int_Q \left| \nabla \left( \frac{a(x)}{R(x, \theta)} \right) \right|^2 e^{2s_\alpha} dx dt \\
&\leq C_1 \int_0^t \left| \nabla y(x, \xi) \right|^2 e^{2s_\alpha} dx dt + C_1 \left( \|ae^{s_\alpha}\|_{L^2(Q)} + \|(\nabla a)e^{s_\alpha}\|_{L^2(Q)} \right).
\end{align*}
\]

To the further estimate, we need the following result:

**Lemma 2.4** If \(\alpha \in C^1(Q)\) satisfies

\[
\left( t - \frac{T}{2} \right) \frac{\partial \alpha}{\partial t} \leq 0, \quad x \in \Omega, 0 \leq t \leq T,
\]

then for all \(u \in L^2(Q)\),

\[
\int_Q \left| \int_0^t |u(x, \xi)| d\xi \right|^2 e^{2s_\alpha} dx dt \leq C_2(T) \int_Q |u(x, t)|^2 e^{2s_\alpha} dx dt.
\]

25
Lemma 2.4 can be proved directly. This kind of results are essential for applying Carleman estimates to inverse problems (see Isakov [16], Lemma 6.4.2, p.153 or Klibanov [20], for example).

We can directly verify that \( (t - \frac{T}{2}) \frac{\partial u}{\partial t} \leq 0, \quad x \in \Omega, \quad 0 \leq t \leq T \), so the application of Lemma 2.4 yields
\[
\int_{Q} \left| \frac{\partial}{\partial t} \left( \frac{2 \nabla R}{R} \right) \cdot \nabla z \right|^2 e^{2s\alpha} \, dx \, dt \\
\leq C_2 \int_{Q} |\nabla y(x, t)|^2 e^{2s\alpha} \, dx \, dt + C_2 (\|ae^{\alpha}\|^2_{L^2(Q)} + \|(\nabla a)e^{\alpha}\|^2_{L^2(Q)})
\]
with \( C_2 > 0 \) independent of \( s > 0 \). Thus, taking large \( s > 0 \), we obtain from (2.53) that
\[
\int_{Q} \left( \frac{1}{s\varphi} (|y|^2 + |\Delta y|^2) + s\varphi|\nabla y|^2 + s^3 \varphi^3 y^2 \right) e^{2s\alpha} \, dx \, dt \\
\leq C \int_{Q} s^3 \varphi^3 y^2 e^{2s\alpha} \, dx \, dt + C (\|ae^{\alpha}\|^2_{L^2(Q)} + \|(\nabla a)e^{\alpha}\|^2_{L^2(Q)}).
\]
For such fixed \( s > 0 \), we have \( e^{2s\alpha}, s^3 \varphi^3 e^{2s\alpha} \leq C_3 \) in \( Q \) and so
\[
\int_{Q} \left( \frac{1}{s\varphi} (|y|^2 + |\Delta y|^2) + s\varphi|\nabla y|^2 + s^3 \varphi^3 y^2 \right) e^{2s\alpha} \, dx \, dt \\
\leq C \int_{Q} |z_t|^2 \, dx \, dt + C \|a\|^2_{H^1(\Omega)}.
\]
Moreover,
\[
\frac{1}{s\varphi} e^{2s\alpha}, s\varphi e^{2s\alpha}, s^3 \varphi^3 e^{2s\alpha} \geq C_4 > 0 \quad \text{in} \quad \Omega \times \left( \frac{T}{4}, \frac{3T}{4} \right)
\]
for the fixed \( s > 0 \). Hence we obtain
\[
\int_{Q} \left( |y|^2 + |\Delta y|^2 + |\nabla y|^2 + y^2 \right) \, dx \, dt \\
\leq C \int_{Q} \left\{ |u(p, \mu) - u(q, \nu)|^2 + \left| \frac{\partial}{\partial t} (u(p, \mu) - u(q, \nu)) \right|^2 \right\} \, dx \, dt \\
+ C \|a\|^2_{H^1(\Omega)}.
\]
Furthermore, by (2.50), (2.51) and (2.54) we have
\[
\frac{T}{2} f = \int_{\Omega} \left( z_t - \Delta z - \frac{2\nabla R}{R} \cdot \nabla z - (p - q)z \right) \, dt
\]
\[
\begin{align*}
&= \int_T^\frac{T}{4} \left( y - \int_\theta^t \Delta y(x, \xi) d\xi - \frac{2\nabla R}{R} \cdot \int_\theta^t \nabla y(x, \xi) d\xi - (p-q) \int_\theta^t y(x, \xi) d\xi \right) dt \\
&\quad - \frac{T}{2} \left( \Delta \left( \frac{a(x)}{R(x, \theta)} \right) + (p-q) \frac{a(x)}{R(x, \theta)} \right) - \int_T^\frac{T}{4} \frac{2\nabla R(x, t)}{R(x, t)} \cdot \nabla \left( \frac{a(x)}{R(x, \theta)} \right) dt.
\end{align*}
\]

Then using the Schwarz's inequality, Lemma 2.3 and the fact that \((p, \mu), (q, \nu) \in U\), we further deduce

\[
\frac{T}{2} \| f \|_{L^2_\omega(\Omega)}^2 \leq \int_\Omega \int_T^\frac{T}{4} \left( y - \int_\theta^t \Delta y(x, \xi) d\xi - \frac{2\nabla R}{R} \cdot \int_\theta^t \nabla y(x, \xi) d\xi \\
- (p-q) \int_\theta^t y(x, \xi) d\xi \right)^2 dx dt \\
+ C_5 (\| a \|_{L^2_\omega(\Omega)} + \| p-q \|_{L^\infty(\Omega)} \| a \|_{L^2_\omega(\Omega)}^2)
\]

\[
\leq C_5 \int_T^\frac{T}{4} \int_\Omega (y^2 + |\Delta y|^2 + |\nabla y|^2 + \| p-q \|_{L^\infty(\Omega)}^2 y^2) dx dt \\
+ C_5 (\| a \|_{L^2_\omega(\Omega)} + \| p-q \|_{L^\infty(\Omega)} \| a \|_{L^2_\omega(\Omega)}^2).
\]

Now (2.3) follows from this and (2.55).

**Fourth Step.** Finally we prove (2.4). Setting \( w_1 = w_t \) in (2.49), we have

\[
\begin{align*}
(w_1)_t &= \Delta w_1 + p(x) w_1 + f(x) R_t(x, t), \quad x \in \Omega, \quad -\delta < t < T \\
w_1(x, \theta) &= \Delta a(x) + p(x) a(x) + f(x) R(x, \theta) \equiv b(x), \quad x \in \Omega \\
w_1(x, t) &= 0, \quad x \in \partial \Omega, \quad -\delta < t < T.
\end{align*}
\]

We then decompose (2.56) as follows:

\[
\begin{align*}
v_t &= \Delta v + p(x) v + f(x) R_t(x, t), \quad x \in \Omega, \quad -\delta < t < T \\
v(x, -\delta) &= 0, \quad x \in \Omega \\
v(x, t) &= 0, \quad x \in \partial \Omega, \quad -\delta < t < T
\end{align*}
\]

and

\[
\begin{align*}
u_t &= \Delta u + p(x) u(x, t), \quad x \in \Omega, \quad -\delta < t < T \\
u(x, \theta) &= b(x) - v(x, \theta), \quad x \in \Omega \\
u(x, t) &= 0, \quad x \in \partial \Omega, \quad -\delta < t < T.
\end{align*}
\]
By the uniqueness of the backward parabolic problem (e.g. [10]), we know that if \( U \in C^2(\Omega \times [-\delta, T]) \) satisfies

\[
\begin{cases}
U_t = \Delta U + p(x)U(x,t), & x \in \Omega, \ -\delta < t < T \\
U(x,0) = 0, & x \in \Omega \\
U(x,t) = 0, & x \in \partial \Omega, \ -\delta < t < T,
\end{cases}
\]

then \( U(x,t) = 0, x \in \Omega, -\delta < t < T \). Therefore we see that

\[
w_1 = u + v \quad \text{in } Q
\]

and

\[
w_1(x,-\delta) = u(x,-\delta), \quad x \in \Omega.
\]

Since \( w_1 = w_t \), we have by Lemma 2.3 that

\[
\begin{align*}
||w_1(\cdot,-\delta)||_{L^\infty(\Omega)} &= ||w_t(\cdot,-\delta)||_{L^\infty(\Omega)} \\
& \leq C_6 ||u(p,\mu) - u(q,\nu)||_{C^{\gamma+3, \frac{\gamma+3}{2}}} \leq C_6M_2,
\end{align*}
\]

which, with (2.60), implies

\[
||u(\cdot,-\delta)||_{L^\infty(\Omega)} \leq C_6M_2.
\]

Thus we can apply the method of the logarithmic convexity (cf. Payne [26]) to obtain

\[
||u(\cdot,t)||_{L^2(\Omega)} \leq (C_6M_2)^{1-\frac{1+\gamma}{4}} ||u(\cdot,\theta)||_{L^2(\Omega)}^{\frac{1+\gamma}{4}}, \quad -\delta \leq t \leq \theta.
\]

Furthermore by the semigroup theory (e.g. Pazy [27]), we can write

\[
v(t) = v(\cdot,t) = \int_{-\delta}^{t} e^{-(t-\xi)A}pv(\xi)d\xi + \int_{-\delta}^{t} e^{-(t-\xi)A}fR_\xi(\xi)d\xi, \quad -\delta < t < \theta.
\]

Here \( A = -\Delta \) with \( \mathcal{D}(A) = H^2(\Omega) \cap H^1_0(\Omega) \). By \( ||p||_{L^\infty(\Omega)} \leq M \) and Lemma 2.3, we have

\[
||v(t)||_{L^2(\Omega)} \leq C_7 \int_{-\delta}^{t} ||v(\xi)||_{L^2(\Omega)}d\xi + C_7||f||_{L^2(\Omega)}, \quad -\delta < t < \theta.
\]
The Gronwall's inequality then yields

\[ \|v(t)\|_{L^2(\Omega)} \leq C_8 \|f\|_{L^2(\Omega)}, \quad -\delta \leq t \leq \theta. \]  

(2.62)

In view of (2.59), (2.56) and (2.58), the inequalities (2.61) and (2.62) imply

\[
\begin{align*}
\|w_t(\cdot, t)\|_{L^2(\Omega)} &\leq \|u(\cdot, t)\|_{L^2(\Omega)} + \|v(\cdot, t)\|_{L^2(\Omega)} \\
&\leq C_9 \|u(\cdot, \theta)\|_{L^2(\Omega)} + C_9 \|f\|_{L^2(\Omega)} \\
&\leq C_9 (\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})^{\frac{1}{2} + \frac{1}{8}} + C_9 \|f\|_{L^2(\Omega)}.
\end{align*}
\]

(2.63)

By (2.63) we have

\[
\|\mu - \nu\|_{L^2(\Omega)} = \|u(p, \mu)(\cdot, -\delta) - u(q, \nu)(\cdot, -\delta)\|_{L^2(\Omega)} = \|w(\cdot, -\delta)\|_{L^2(\Omega)} \\
= \left\| - \int_{-\delta}^0 w_t(\cdot, \xi) d\xi + w(\cdot, \theta) \right\|_{L^2(\Omega)} \\
\leq C_9 \int_{-\delta}^0 (\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})^{\frac{1}{2} + \frac{1}{8}} d\xi + C_9 \|f\|_{L^2(\Omega)} (\theta + \delta) + \|a\|_{L^2(\Omega)} \\
= C_9 (\theta + \delta) \left( \frac{1}{\log(\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})} \right) \left( \frac{1}{\log(\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})} \right) \\
+ C_9 \|f\|_{L^2(\Omega)} (\theta + \delta) + \|a\|_{L^2(\Omega)} \\
\leq \frac{C_{10}}{\log(\|a\|_{H^2(\Omega)} + \|f\|_{L^2(\Omega)})}.
\]

By noting that \(a = u(p, \mu)(\cdot, \theta) - u(q, \nu)(\cdot, \theta)\) and \(f = p - q\), (2.4) follows from the last estimate above and (2.3). This completes the proof of Theorem 2.1. \(\square\)
Chapter 3

The continuous formulation

3.1 Constrained minimization problem

From the stability established in Chapter 2 we know that the inverse problem of reconstructing the initial temperature and the heat radiative coefficient are conditionally stable under a very smooth class of solutions to the corresponding parabolic system and the sufficiently rich observation data, including the data \( u \in H^1(0,T; L^2(\omega)) \) and \( u(\cdot, \theta) \in H^2(\Omega) \). Clearly such a strict requirement on the observation data is not very practical. However, these stability results give us some important guidance and encouragement for the numerical reconstruction. Starting from this chapter, we will explore the possibility of the numerical reconstruction of both the initial temperature and the heat radiative coefficient. Contrary to the assumptions made in Chapter 2, the observation data and the admissible solution classes will be allowed to be much more practical here.

From now on, we assume that only the observations \( z_0(x) \) and \( z(x,t) \) given in (1.5) are available. Then we formulate the reconstruction of the initial temperature \( \mu \) and the thermal radiative coefficient \( p \) in (1.1)-(1.2) as the following constrained minimizing process:

\[
\text{minimize } J(p, \mu) = \frac{1}{2} \int_0^T \int_\omega (u(p, \mu) - z)^2 \, dx \, dt + \frac{1}{2} \int_\Omega (u(p, \mu)(x, \theta) - z_0(x))^2 \, dx
\]
\[ +\beta \int_{\Omega} |\nabla p|^2 \, dx + \gamma \int_{\Omega} |\nabla \mu|^2 \, dx \]  
subject to \( p \in K_1, \mu \in K_2 \) and \( u(\cdot, t) \equiv u(p, \mu)(\cdot, t) \in H^1(\Omega) \) satisfying

\[ u(x, 0) = \mu(x) \quad \text{in } \Omega; \quad u(x, t) = \eta(x, t) \quad \text{on } \partial \Omega \times (0, T) \quad (3.2) \]

\[ \int_{\Omega} u_\xi \phi \, dx + \int_{\Omega} \nabla u \cdot \nabla \phi \, dx = \int_{\Omega} p(x) u \phi \, dx \quad \forall \phi \in H^1_0(\Omega) \quad (3.3) \]

for a.e. \( t \in (0, T) \). Note that the system (3.2)-(3.3) is the variational formulation associated with the parabolic problem (1.1)-(1.3). In the sequel, we may denote the solution of this variational problem as \( u(p, \mu) \) or \( u(p, \mu)(x, t) \). The measurement functions \( z \) and \( z_\theta \) satisfy that \( z \in L^2(0, T; L^2(\omega)) \) and \( z_\theta \in L^2(\Omega) \), the constants \( \beta \) and \( \gamma \) are regularization parameters. We remark that the \( H^1 \)-seminorm regularization for both \( p \) and \( \mu \) used in (3.1) can be replaced by the seminorm of bounded variations when one wants to reconstruct some discontinuous thermal radiative coefficient \( p(x) \) and initial data \( \mu(x) \). With such a replacement of regularization, all the results of this paper follow, with only some natural modifications of the proofs here combined with our previous technical treatments for \( BV \)-functions as done in [5, 19]. The constrained sets \( K_1 \) and \( K_2 \) above are chosen to be as follows:

\[ K_1 = \{ p \in H^1(\Omega); \, |p(x)| \leq \alpha_1 \quad \text{a.e. in } \Omega \}, \]

\[ K_2 = \{ \mu \in H^1(\Omega); \, 0 < \mu(x) \leq \alpha_2 \quad \text{a.e. in } \Omega \}. \]

Here \( \alpha_1 \) and \( \alpha_2 \) are two positive constants.

### 3.2 Existence of minimizers to the minimization problem

To show the existence of the minimizers to the minimization problem (3.1)-(3.3), we first prove
Lemma 3.1 For any sequences \( \{p_n\} \) in \( K_1 \) which converges to some \( p^* \in K_1 \) in \( H^1(\Omega) \) and \( \{\mu_n\} \) in \( K_2 \) which converges to some \( \mu^* \in K_2 \) in \( H^1(\Omega) \) as \( n \to \infty \), we have

\[
\lim_{n \to \infty} \int_0^T \int_{\Omega} (u(p_n, \mu_n) - z)^2 dx dt = \int_0^T \int_{\Omega} (u(p^*, \mu^*) - z)^2 dx dt, \quad \text{and}
\]

\[
\lim_{n \to \infty} \int_{\Omega} (u(p_n, \mu_n)(x, \theta) - z_{\theta}(x))^2 dx = \int_{\Omega} (u(p^*, \mu^*)(x, \theta) - z_{\theta}(x))^2 dx.
\]

Proof. Take any \( p \in K_1 \) and \( \mu \in K_2 \), \( \phi = u(p, \mu) - \eta \in H_0^1(\Omega) \). Then for (3.3), integrating both sides with respect to \( t \), we have

\[
\int_0^t \int_{\Omega} u_t (u - \eta) dx dt + \int_0^t \int_{\Omega} \nabla u \cdot \nabla (u - \eta) dx dt
\]

\[
= \int_0^t \int_{\Omega} p(x)u(u - \eta) dx dt,
\]

\[
\int_0^t \int_{\Omega} u_t u \eta dx dt + \int_0^t \int_{\Omega} \nabla u \cdot \nabla u \eta dx dt
\]

\[
= \int_0^t \int_{\Omega} u \eta \eta dx dt + \int_0^t \int_{\Omega} \nabla u \cdot \nabla \eta dx dt + \int_0^t \int_{\Omega} p(x)u^2 dx dt + \int_0^t \int_{\Omega} p(x)u \eta dx dt,
\]

\[
\leq \int_0^t \int_{\Omega} (u^2 - u \eta) dt dx + \int_0^t \int_{\Omega} u(t) \eta(t) dx dt - \int_0^t \int_{\Omega} u(0) \eta(0) dx dt
\]

\[
+ \frac{1}{2} \left( \int_0^t \int_{\Omega} |\nabla u|^2 dx dt + \int_0^t \int_{\Omega} |\nabla \eta|^2 dx dt \right) + \alpha_1 \int_0^t \int_{\Omega} u^2 dx dt
\]

\[
+ \frac{\alpha_1}{2} \left( \int_0^t \int_{\Omega} u^2 dx dt + \int_0^t \int_{\Omega} \eta^2 dx dt \right).
\]

So we have

\[
\int_0^t \int_{\Omega} u^2 dx dt \leq C_1 + C_2 \int_0^t \int_{\Omega} u^2 dx dt.
\]

By Gronwall inequality, we have

\[
\int_0^t \int_{\Omega} u^2 dx dt \leq C e^{C_2 t} \leq C e^{C_2 T}
\]

which is finite for any \( t \in (0, T] \). This implies the sequence \( \{u(p_n, \mu_n)\} \) is bounded in the space \( L^2(0, T; H^1(\Omega)) \). Hence we can extract a subsequence, still denoted by
\{u(p_n, \mu_n)\} \text{ s.t. } u(p_n, \mu_n) \to u^* \text{ weakly in } L^2(0, T; H^1(\Omega)).

We next show that $u^* = u(p^*, \mu^*)$. We multiply both sides of (3.3) by a function $\eta(t) \in C^1[0, T]$ with $\eta(T) = 0$, we get

$$\int_\Omega u(p_n, \mu_n) \eta dt + \int_\Omega \nabla u(p_n, \mu_n) \cdot \nabla \eta dt = \int_\Omega p_n u(p_n, \mu_n) \eta dt.$$ 

Then integrating both sides with respect to $t$, we get

$$\int_\Omega \left( \int_0^T u(p_n, \mu_n) \eta dt \right) \phi dx + \int_\Omega \int_0^T \nabla u(p_n, \mu_n) \cdot \phi \eta dt dt = \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \int_0^T \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt,$$

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt.$$ 

Since $\int_0^T \int_\Omega (p_n - p^*) u(p_n, \mu_n) \eta \phi dx dt \to 0$ by the Cauchy-Schwarz inequality, Young's inequality and the Lebesgue dominant theorem, letting $n \to \infty$, and by $u(p_n, \mu_n) \to u^*$, we have

$$\int_\Omega \mu_n \eta(0) \phi dx + \int_\Omega \int_0^T p_n u(p_n, \mu_n) \eta \phi dx dt = \int_\Omega \int_0^T u^* \eta \phi dx dt + \int_\Omega \int_0^T \nabla u^* \cdot \nabla \eta \phi dx dt,$$

which is valid for any $\eta(t) \in C^1[0, T]$ with $\eta(T) = 0$. This implies the equation is also true for any $\eta(t) \in C_0^\infty(0, T)$. So

$$\int_\Omega \int_0^T u^* \eta \phi dx dt + \int_\Omega \int_0^T \nabla u^* \cdot \nabla \eta \phi dx dt = \int_\Omega \int_0^T p u^* \eta \phi dx dt, \quad \forall \phi \in H^1_0(\Omega).$$
and get \( u^*(0) = \mu^* \). Thus we know \( u^* = u(p^*, \mu^*) \).

Then
\[
\lim_{n \to \infty} \int_0^T \int_\Omega (u(p_n, \mu_n) - z)^2 \, dx \, dt - \int_0^T \int_\Omega (u(p^*, \mu^*) - z)^2 \, dx \, dt
\]
\[
= \lim_{n \to \infty} \int_0^T \int_\Omega (u^2(p_n, \mu_n) - u^2(p^*, \mu^*)) \, dx \, dt - 2 \lim_{n \to \infty} \int_0^T \int_\Omega z(u(p_n, \mu_n) - u(p^*, \mu^*)) \, dx \, dt
\]
\[
= 0.
\]
i.e.
\[
\lim_{n \to \infty} \int_0^T \int_\Omega (u(p_n, \mu_n) - z)^2 \, dx \, dt = \int_0^T \int_\Omega (u(p^*, \mu^*) - z)^2 \, dx \, dt.
\]

And
\[
\lim_{n \to \infty} \int_\Omega (u(p_n, \mu_n)(x, \theta) - z_\theta(x))^2 \, dx - \int_\Omega (u(p^*, \mu^*)(x, \theta) - z_\theta(x))^2 \, dx
\]
\[
= \lim_{n \to \infty} \int_\Omega (u^2(p_n, \mu_n)(x, \theta) - u^2(p^*, \mu^*)(x, \theta)) \, dx
\]
\[
- 2 \lim_{n \to \infty} \int_\Omega z_\theta(x)(u(p_n, \mu_n)(x, \theta) - u(p^*, \mu^*)(x, \theta)) \, dx
\]
\[
= 0.
\]
i.e.
\[
\lim_{n \to \infty} \int_\Omega (u(p_n, \mu_n)(x, \theta) - z_\theta(x))^2 \, dx = \int_\Omega (u(p^*, \mu^*)(x, \theta) - z_\theta(x))^2 \, dx.
\]

Then we can show

**Theorem 3.1** There exists at least one minimizer to the optimization problem (3.1)-(3.3).

**Proof.** We know that for any constant function \( p \) in \( K_1 \), \( \mu \) in \( K_2 \), \( J(p, \mu) \) is bounded. Thus \( \min J(p, \mu) \) is finite over \( K_1 \times K_2 \) and there exists a minimizing sequence \( \{p_n, \mu_n\} \) with \( p_n \in K_1 \), \( \mu_n \in K_2 \) such that
\[
\lim_{n \to \infty} J(p_n, \mu_n) = \min_{p \in K_1, \mu \in K_2} J(p, \mu),
\]

34
and there is a subsequence, still denoted as \( \{p_n, \mu_n\} \) such that \( p_n \to p^* \), \( \mu_n \to \mu^* \)
in \( L^1(\Omega) \) with \( p^* \in K_1 \) and \( \mu^* \in K_2 \).

Then by Lemma 3.1, we have

\[
J(p^*, \mu^*) \leq \lim_{n \to \infty} \frac{1}{2} \int_0^T \int_\omega (u(p_n, \mu_n) - z)^2 \, dx \, dt + \lim_{n \to \infty} \frac{1}{2} \int_\Omega \left( u(p_n, \mu_n)(x, \theta) - z_\theta(x) \right)^2 \, dx
\]
\[
+ \beta \lim_{n \to \infty} \inf \int_\Omega |\nabla p_n|^2 \, dx + \gamma \lim_{n \to \infty} \inf \int_\Omega |\nabla \mu_n|^2 \, dx
\]
\[
\leq \lim \inf_{n \to \infty} J(p_n, \mu_n)
\]
\[
= \min_{p \in K_1, \mu \in K_2} J(p, \mu)
\]

which indicates \((p^*, \mu^*)\) is a minimizer of \( J(\cdot, \cdot) \) over \( K_1 \times K_2 \). (cf. [18],[19],[21]).

\( \square \)
Chapter 4

Discretization and its convergence

4.1 Finite element space

We now propose a finite element method for solving the continuous minimization problem (3.1)-(3.3). For the purpose, we first triangulate the polyhedral domain $\Omega$ with a regular triangulation $T^h$ of simplicial elements (cf. Ciarlet [8]). Then we define the finite element space $V_h$ to be the space of all continuous piecewise linear functions over $T^h$, and $\bar{V}_h$ a subspace of $V_h$ with all functions vanishing on the boundary $\partial \Omega$. Let $N_h = \{x_i\}_{i=1}^{N_h}$ be the set of all nodal points of $T^h$, then we approximate the constrained subsets $K_1$ and $K_2$ by

$$K_{1h} = \{p_h \in V_h; \ |p_h(x_i)| \leq \alpha_1 \text{ for } x_i \in N_h\},$$

$$K_{2h} = \{\mu_h \in V_h; \ 0 < \mu_h(x_i) \leq \alpha_2 \text{ for } x_i \in N_h\}.$$

To fully discretize the system (3.1)-(3.3), we need also the time discretization. We thus divide the time interval $[0,T]$ into $M$ equally-spaced subintervals by using nodal points

$$0 = t^0 < t^1 < \cdots < t^M = T.$$
with \( t^n = n\tau, \tau = T/M \). For a continuous mapping \( u : [0, T] \rightarrow L^2(\Omega) \), we define \( u^n = u(\cdot, n\tau) \) for \( 0 \leq n \leq M \). For a given sequence \( \{u^n\}_{n=0}^M \subset L^2(\Omega) \) we define its difference quotient and the averaging \( \bar{u}^n \) of a function \( u(\cdot, t) \) as follows:

\[
\partial_\tau u^n = \frac{u^n - u^{n-1}}{\tau}, \quad \bar{u}^n = \frac{1}{\tau} \int_{t^{n-1}}^{t^n} u(\cdot, t) \, dt
\]

where for \( n = 0 \), we let \( \bar{u}^0 = u(\cdot, 0) \).

### 4.2 Two important discrete projection operators

For our later analysis, we now introduce two discrete projection operators. The first is a quasi-\( L^2 \) projection operator \( \pi_h : L^2(\Omega) \rightarrow V_h \) given by

\[
(\pi_h w, v_h)_h = (w, v_h), \quad \forall w \in L^2(\Omega), \quad v_h \in V_h
\]

where \((\cdot, \cdot)_h\) denotes the lumped mass discrete \( L^2 \)-inner product:

\[
(u_h, v_h)_h = \int_{\Omega} I_h(u_h v_h) \, dx, \quad \forall u_h, v_h \in V_h.
\]

Here \( I_h \) is the nodal value interpolation operator associated with the finite element space \( V_h \). The operator \( \pi_h \) possesses the following properties (cf. Nochetto-Verdi [25], Xu [31]):

\[
\|\pi_h w\| \leq C\|w\|, \quad \lim_{h \rightarrow 0} \|w - \pi_h w\| = 0, \quad \forall w \in L^2(\Omega); \quad (4.1)
\]

\[
\|\pi_h w\|_{H^1(\Omega)} \leq C\|w\|_{H^1(\Omega)}, \quad \lim_{h \rightarrow 0} \|w - \pi_h w\|_{H^1(\Omega)} = 0, \quad \forall w \in H^1(\Omega). \quad (4.2)
\]

where \( \| \cdot \| \) is the \( L^2 \)-norm in \( L^2(\Omega) \). Now let \( \pi_h w(x) = \sum_{i=1}^{N_h} w_i \phi_i(x) \), then it is easy to see that

\[
w_i = \frac{(w, \phi_i)}{(\phi_i, \phi_i)_h} = \frac{(w, \phi_i)}{(1, \phi_i)},
\]

therefore we have \( \pi_h w \in K_{1h} \) (resp. \( K_{2h} \)) for any \( w \in K_1 \) (resp. \( K_2 \)). This property is crucial to our later convergence analysis. It is important to note that the standard \( L^2 \) projection does not have such a nice property [4].
The second operator we need is a discrete operator $Q_h : H^1(\Omega) \to V_h$ which preserves the boundary values, namely $Q_h w = Q_h v$ on $\partial\Omega$ for any two functions $w, v \in H^1(\Omega)$ satisfying $w = v$ on $\partial\Omega$. Clearly the quasi-$L^2$ projection $\pi_h$ does not have such a property. The standard nodal value interpolation $I_h$ associated with the space $V_h$ has the aforementioned two properties but it is well-defined only for smoother functions and it does not have the stability estimates shown in (4.1)-(4.2) either. Next we introduce such a boundary value preserving operator $Q_h$, which was first proposed by Scott-Zhang [28]. For any $x_i \in N_h$ (the set of nodal points in $T^h$), choose $\tau_i$ to be a $(d - 1)$-simplex from the triangulation $T^h$ with vertices $z_l$ ($l = 1, \ldots, d$) such that $z_l = x_i$. The choice of $\tau_i$ is not unique in general, but if $x_i \in \partial\Omega$, we take $\tau_i \subset \partial\Omega$. Let $\theta_i \in P_1(\tau_i)$, the set of all linear polynomials on $\tau_i$, be the unique function satisfying

$$\int_{\tau_i} \theta_i \lambda_i dx = \delta_{ii}, \quad l = 1, \ldots, d$$

where $\lambda_i$ is the barycentric coordinate of $\tau_i$ (see Ciarlet [8]) with respect to $z_l$. Obviously,

$$\int_{\tau_i} \theta_i v dx = v(x_i) \quad \text{for any } v \in P_1(\tau_i).$$

The average nodal value interpolant $Q_h$ is then defined by

$$Q_h v(x) = \sum_{i=1}^{N_h} \phi_i(x) \int_{\tau_i} \theta_i v dx. \quad (4.3)$$

For any $v \in H^{1/2}(\partial\Omega)$, we define

$$Q_h^b v(x) = \sum_{z_l \in \partial\Omega} \phi_i(x) \int_{\tau_i} \theta_i v dx = Q_h v(x) \quad \forall x \in \partial\Omega. \quad (4.4)$$

Note that the above definitions for $Q_h$ and $Q_h^b$ do not make sense for $n = 1$. In this case, we let $Q_h$ and $Q_h^b$ to be the standard nodal value interpolant $I_h$ on $V_h$, this is well-defined since $H^1(\Omega) \subset C^0(\bar{\Omega})$ ($n = 1$).

**Lemma 4.1** The operator $Q_h$ defined by (4.3) satisfies (cf. [28])
4.3 Finite element problem

With the above notation and the discretization of the first term of (3.1) by the composite trapezoidal rule in time, we formulate the finite element approximation of the problem (3.1)-(3.3) as follows:

\[
\text{minimize } J_h(p_h, \mu_h) = \frac{T}{2} \sum_{n=0}^{M} \alpha_n \int_{\Omega} (u^n_h(p_h, \mu_h) - z^n)^2 dx + \frac{1}{2} \int_{\Omega} (u^n_{h0}(p_h, \mu_h) - z^n)^2 dx \\
+ \beta \int_{\Omega} |\nabla p_h|^2 dx + \gamma \int_{\Omega} |\nabla \mu_h|^2 dx
\]  

(4.5)

over all \(p_h \in K_{1h}\) and \(\mu_h \in K_{2h}\) with \(u^n_h \equiv u^n_h(p_h, \mu_h) \in V_h\) satisfying

\[
u^n_h = \mu_h \quad \text{and} \quad u^n_h = Q^n_h \bar{\eta} + \hat{u}^n_h,
\]

(4.6)

\[
\int_{\Omega} \partial_t u^n_h \phi_h dx + \int_{\Omega} \nabla u^n_h \cdot \nabla \phi_h dx = \int_{\Omega} p_h u^n_h \phi_h \forall \phi_h \in V_h
\]

(4.7)

for \(n = 1, 2, \cdots, M\). Here \(\alpha_0 = \alpha_M = \frac{1}{2}\) and \(\alpha_n = 1\) for all \(n \neq 0, M\) while \(\hat{u}^n_h \in V_h^0\) and \(n_0 > 0\) is an integer such that \(n_0 \tau = \theta\). Note that \(\bar{\eta}^n\) is the average of \(\eta(x, t)\) on \([t^{n-1}, t^n]\), so \(Q^n_h \bar{\eta}^n\) is easy to calculate on the boundary \(\partial \Omega\) using (4.4).

About the existence of the minimizers to the finite element problem (4.5)-(4.7), we have the following theorem, which can be proved basically along the same line as given in Keung-Zou [18, 19] combining with the stability estimates and convergence results of Lemmas 4.5 and 4.6 below.

4.4 Existence of minimizers to the finite element problem
Lemma 4.2 For any sequence \( \{p^k_h, \mu^k_h\} \) in \( K_{1h} \times K_{2h} \), which converges to some \( \{p^*_h, \mu^*_h\} \) in \( K_{1h} \times K_{2h} \) in \( H^1(\Omega) \) as \( k \) tends to \( \infty \), we have for \( n = 1, 2, \ldots, M \),
\[
u^k_h(p^k_h, \mu^k_h) \to \nu^*_h(p^*_h, \mu^*_h) \quad \text{in} \quad H^1(\Omega) \quad \text{as} \quad k \to \infty.
\]

Proof. By the definition of \( \nu^k_h(p^k_h, \mu^k_h) \) and \( \nu^*_h(p^*_h, \mu^*_h) \), we have
\[
\begin{align*}
\int_\Omega \partial_t \nu^k_h(p^k_h, \mu^k_h) \phi_h dx &+ \int_\Omega \nabla \nu^k_h(p^k_h, \mu^k_h) \cdot \nabla \phi_h dx \\
&= \int_\Omega p^k_h \nu^k_h(p^k_h, \mu^k_h) \phi_h dx, \quad \forall \phi_h \in \mathcal{V}_h \\
\int_\Omega \partial_t \nu^*_h(p^*_h, \mu^*_h) \phi_h dx &+ \int_\Omega \nabla \nu^*_h(p^*_h, \mu^*_h) \cdot \nabla \phi_h dx \\
&= \int_\Omega p^*_h \nu^*_h(p^*_h, \mu^*_h) \phi_h dx, \quad \forall \phi_h \in \mathcal{V}_h \quad (4.8)
\end{align*}
\]
Taking \( \phi_h = \tau u^*_{h0}(p^*_h, \mu^*_h) \) where \( u^*_{h0} = u^*_h \) inside \( \Omega \) and \( u^*_{h0} = 0 \) on \( \partial \Omega \). Use the fact that \( p^*_h \leq \alpha_1 \), we have
\[
\frac{1}{2} \| \nu^*_h(p^*_h, \mu^*_h) \|_{L^2(\Omega)}^2 + \frac{1}{2} \| \nu^*_{h-1}(p^*_h, \mu^*_h) \|_{L^2(\Omega)}^2 + \tau \| \nabla \nu^*_h(p^*_h, \mu^*_h) \|_{L^2(\Omega)}^2 \leq \tau \alpha_1 \| \nu^*_{h}(p^*_h, \mu^*_h) \|_{L^2(\Omega)}^2.
\]
Summing both sides over \( n = 1, 2, \ldots, k \leq M \), we have
\[
\max_{1 \leq n \leq M} \| \nu^*_h(p^*_h, \mu^*_h) \|_{L^2(\Omega)}^2 \leq C
\]
with \( C \) independent of \( h, \tau \) and \( k \).

Let \( w^*_h(k) = \nu^*_h(p^*_h, \mu^*_h) - \nu^*_h(p^*_h, \mu^*_h) \), and (4.9)- (4.8), we have
\[
\begin{align*}
\int_\Omega \partial_t w^*_h(k) \phi_h dx &+ \int_\Omega \nabla w^*_h(k) \cdot \phi_h dx \\
&= \int_\Omega p^*_h \nu^*_h(p^*_h, \mu^*_h) \phi_h dx - \int_\Omega p^*_h \nu^*_h(p^*_h, \mu^*_h) \phi_h dx \\
&= \int_\Omega p^*_h \nu^*_h(k) \phi_h dx - \int_\Omega (p^*_h - p^*_h) \nu^*_h(p^*_h, \mu^*_h) \phi_h dx, \quad \forall \phi_h \in \mathcal{V}_h \quad (4.9)
\end{align*}
\]
Then taking \( \phi_h = \tau w^*_h(k) \) gives
\[
\begin{align*}
&\frac{1}{2} \| w^*_h(k) \|_{L^2(\Omega)}^2 + \frac{1}{2} \| w^*_{h-1}(k) \|_{L^2(\Omega)}^2 + \tau \| \nabla w^*_h(k) \|_{L^2(\Omega)}^2 \\
&\leq \tau \alpha_1 \| w^*_h(k) \|_{L^2(\Omega)}^2 + \tau \max_{x \in \Omega} \| [p^*_h(x) - p^*_h(x)] \| \nu^*_h(p^*_h, \mu^*_h) \|_{L^2(\Omega)} \| w^*_h(k) \|_{L^2(\Omega)}.
\end{align*}
\]
Summing the above inequality over $n = 1, 2, \cdots, M$, and by Gronwall's inequality, we have

$$w^n_h(k) \to 0 \text{ in } H^1(\Omega) \text{ as } k \to \infty.$$ 

Thus our desire result follows. □

**Lemma 4.3** For any sequences $\{\mu^k_h\}$ in $K_{1h}$ which converges to some $\mu^*_h \in K_{1h}$ in $H^1(\Omega)$ and $\{\mu^k_h\}$ in $K_{2h}$ which converges to some $\mu^*_h \in K_{2h}$ in $H^1(\Omega)$ as $k \to \infty$, we have

$$
\lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^k_h, \mu^k_h) - z^n)^2 dx = \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^*_h, \mu^*_h) - z^n)^2 dx,
$$

and

$$
\lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^{n_0}_h(p^k_h, \mu^k_h) - z_0)^2 dx = \frac{1}{2} \int_{\Omega} (u^{n_0}_h(p^*_h, \mu^*_h) - z_0)^2 dx.
$$

**Proof.**

\[
\lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^k_h, \mu^k_h) - z^n)^2 dx = \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^*_h, \mu^*_h) - z^n)^2 dx
\]

\[
= \lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} ((u^n_h(p^k_h, \mu^k_h))^2 - (u^n_h(p^*_h, \mu^*_h))^2) dx
\]

\[
-2 \lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \int_{\Omega} \alpha_n z^n (u^n_h(p^k_h, \mu^k_h) - u^n_h(p^*_h, \mu^*_h)) dx
\]

\[
= 0.
\]

i.e.

\[
\lim_{k \to \infty} \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^k_h, \mu^k_h) - z^n)^2 dx = \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u^n_h(p^*_h, \mu^*_h) - z^n)^2 dx.
\]

And

\[
\lim_{k \to \infty} \frac{1}{2} \int_{\Omega} (u^{n_0}_h(p^k_h, \mu^k_h) - z_0)^2 dx = \frac{1}{2} \int_{\Omega} (u^{n_0}_h(p^*_h, \mu^*_h) - z_0)^2 dx
\]

\[
= \lim_{k \to \infty} \frac{1}{2} \int_{\Omega} ((u^{n_0}_h(p^k_h, \mu^k_h))^2 - (u^{n_0}_h(p^*_h, \mu^*_h))^2) dx
\]

\[
-2 \lim_{k \to \infty} \frac{1}{2} \int_{\Omega} z_0 (u^{n_0}_h(p^k_h, \mu^k_h) - u^{n_0}_h(p^*_h, \mu^*_h)) dx
\]

\[
= 0.
\]
Then we can show

**Theorem 4.1** There exists at least one minimizer to the finite element problem (4.5)-(4.7).

**Proof.** We know that for any constant function \( p_h \) in \( K_{1h} \), \( \mu_h \) in \( K_{2h} \), \( J_h(p_h, \mu_h) \) is bounded. Thus \( \min J_h(p_h, \mu_h) \) is finite over \( K_{1h} \times K_{2h} \) and there exists a minimizing sequence \( \{ p_h^k, \mu_h^k \} \) such that

\[
\lim_{k \to \infty} J_h(p_h^k, \mu_h^k) = \min_{p_h \in K_{1h}, \mu_h \in K_{2h}} J_h(p_h, \mu_h),
\]

and there is a subsequence, still denoted as \( \{ p_h^k, \mu_h^k \} \) such that \( p_h^k \to p_h^*, \mu_h^k \to \mu_h^* \) in \( L^1(\Omega) \) with \( p_h^* \in K_{1h} \) and \( \mu_h^* \in K_{2h} \).

Then by Lemma 4.3, we have

\[
J_h(p_h^*, \mu_h^*) \leq \lim \frac{1}{2} \sum_{n=1}^{M} \alpha_n \int_{\Omega} (u_h^n(p_h^k, \mu_h^k) - z^n)^2 dx + \lim \frac{1}{2} \int_{\Omega} (u_h^0(p_h^k, \mu_h^k) - z_0)^2 dx
\]

\[
+ \beta \lim \inf_{k \to \infty} \int_{\Omega} |\nabla p_h^k|^2 dx + \gamma \lim \inf_{k \to \infty} \int_{\Omega} |\nabla \mu_h^k|^2 dx
\]

\[
\leq \lim \inf_{k \to \infty} J_h(p_h^k, \mu_h^k)
\]

\[
= \min_{p_h \in K_{1h}, \mu_h \in K_{2h}} J_h(p_h, \mu_h)
\]

which indicates \( (p_h^*, \mu_h^*) \) is a minimizer of \( J_h(\cdot, \cdot) \) over \( K_{1h} \times K_{2h} \). □

### 4.5 Discrete minimizers and global minimizers

In the rest of this chapter we are going to prove the convergence of the discrete minimizers of (4.5) to the global minimizers of (3.1). For this purpose, we will first present three auxiliary lemmas.
Lemma 4.4 For any given function \( f \in H^1(0,T) \), we have
\[
E(f^2) \leq 2\tau \left( \int_0^T f'(t)^2 dt \right)^{1/2} \left( \int_0^T f^2(t) dt + \max_{t \in [0,T]} f^2(t) \right)^{1/2}.
\] (4.10)
where \( E(f^2) \) is the approximation error given by
\[
E(f^2) = \int_0^T f^2(t) dt - \frac{T}{2} \sum_{n=1}^M \left\{ f^2(t^n - 1) + f^2(t^n) \right\}.
\]

Proof. We can write
\[
E(f^2) = \frac{1}{2} \sum_{n=1}^M \int_{t^n}^{t^{n+1}} \left\{ f^2(t) - f^2(t^n) \right\} dt + \frac{1}{2} \sum_{n=1}^M \int_{t^n}^{t^{n+1}} \left\{ f^2(t) - f^2(t^n - 1) \right\} dt
\]
\[
=: E_1(f^2) + E_2(f^2).
\]
For \( E_1(f^2) \), we can estimate as follows:
\[
\left| E_1(f^2) \right| \leq \frac{1}{2} \sum_{n=1}^M \left( \int_{t^n}^{t^{n+1}} (f(t) - f(t^n))^2 dt \right)^{1/2} \left( \int_{t^n}^{t^{n+1}} (f(t) + f(t^n))^2 dt \right)^{1/2}
\]
\[
\leq \frac{1}{2} \left\{ \int_{t^n}^{t^{n+1}} (f(t) - f(t^n))^2 dt \right\}^{1/2} \left\{ \int_{t^n}^{t^{n+1}} (f(t) + f(t^n))^2 dt \right\}^{1/2}
\]
\[
\leq \frac{T}{2} \left( \int_0^T f'(t)^2 dt \right)^{1/2} \left( 2 \int_0^T f^2(t) dt + 2\tau \sum_{n=1}^M f^2(t^n) \right)^{1/2}.
\]
Similar estimate holds also for \( E_2(f^2) \):
\[
\left| E_2(f^2) \right| \leq \frac{T}{2} \left( \int_0^T f'(t)^2 dt \right)^{1/2} \left( 2 \int_0^T f^2(t) dt + 2\tau \sum_{n=1}^M f^2(t^n - 1) \right)^{1/2}.
\]
Now the estimate (4.10) follows immediately from the above estimates for \( E_1(f^2) \) and \( E_2(f^2) \). \(\square\)

From now on, we assume that the given boundary function in (1.3) satisfies
\[
\eta(x,t) \in H^1(0,T; H^{1/2}(\partial\Omega)),
\]
and then extend \( \eta \) into the domain \( \Omega \) by solving the boundary value problem at each time \( t \):
\[
\Delta \eta(x,t) = 0 \quad \text{in} \quad \Omega; \quad \eta(x,t) = \eta(x,t) \quad \text{on} \quad \partial\Omega.
\]
With this extension we have $\eta_\ast(x, t) \in H^1(0, T; H^1(\Omega))$. Note that this extension
is only used in the convergence analysis and will never be needed in the numerical
algorithm.

**Remark 4.1** Note that $\eta_\ast(x, t) = \eta(x, t)$ on $\partial\Omega$, so their averages over $[t^{n-1}, t^n]$
are also the same, i.e. $\bar{\eta}_\ast^n(x) = \bar{\eta}_\ast(t)$ on $\partial\Omega$. Using this, we can write $\eta_h\bar{\eta}_\ast^n =
\eta_h^n + \bar{\eta}_h^n$ with $\bar{\eta}_h^n \in V_h$. Then the solution $u_h^n$ in (4.6) can be decomposed as

$$u_h^n = \eta_h^n + \bar{\eta}_h^n = Q_h\bar{\eta}_\ast^n + (\bar{\eta}_h^n - \bar{\eta}_h^n). \quad (4.11)$$

This will be an important relation in our subsequent analysis.

**Lemma 4.5** Let $u_h^n(p_h, \mu_h)$ be the solution of the finite element problem (4.6)-(4.7)
corresponding to $p_h \in K_{1h}$ and $\mu_h \in K_{2h}$, then we have the following stability
estimates

$$\max_{1 \leq n \leq M} \|u_h^n\|^2_{H^1(\Omega)} + \tau \sum_{n=1}^M \left( \|\nabla u_h^n\|^2_{L^2(\Omega)} + \|\partial_t u_h^n\|^2 \right) \leq C \left( \|\mu_h\|^2_{H^1(\Omega)} + \|\eta_\ast^n\|^2_{H^1(0, T; H^1(\Omega))} \right),$$

with $C$ independent of $p_h$, $\mu_h$, $h$ and $\tau$.

**Proof.** We choose $\phi_h = \tau(\bar{\eta}_h^n - \bar{\eta}_h^n)$ in (4.7) and use (4.11) to obtain

$$(u_h^n - u_h^{n-1}, u_h^n) + \tau (\nabla u_h^n, \nabla u_h^n)$$

$$= \tau (p_h u_h^n, u_h^n) - \tau (p_h u_h^n, Q_h\bar{\eta}_\ast^n) + \tau (\partial_t u_h^n, Q_h\bar{\eta}_\ast^n) + \tau (\nabla u_h^n, \nabla Q_h\bar{\eta}_\ast^n),$$

then using the Young's inequality: $ab \leq \varepsilon a^2 + b^2/(4\varepsilon)$, $\forall a, b, \varepsilon \in R^+_1$, we derive

$$\frac{1}{2} \|u_h^n\|^2 - \frac{1}{2} \|u_h^{n-1}\|^2 + \tau \|\nabla u_h^n\|^2$$

$$\leq 2\tau\alpha_1 \|u_h^n\|^2 + \frac{1}{4}\tau \|\partial_t u_h^n\|^2 + \tau \|Q_h\bar{\eta}_\ast^n\|^2 + \frac{1}{2}\tau \|\nabla u_h^n\|^2$$

$$+ \frac{1}{4}\tau \|\nabla Q_h\bar{\eta}_\ast^n\|^2 + \frac{1}{4}\tau \|\nabla Q_h\bar{\eta}_\ast^n\|^2. \quad (4.12)$$

Next we choose $\phi_h = \tau \partial_t(\bar{\eta}_h^n - \bar{\eta}_h^n)$ in (4.7) and use (4.11) to have

$$\tau (\partial_t u_h^n, \partial_t u_h^n) + \tau (\nabla u_h^n, \partial_t \nabla u_h^n)$$

$$= \tau (p_h u_h^n, \partial_t u_h^n) - \tau (p_h u_h^n, \partial_t Q_h\bar{\eta}_\ast^n) + \tau (\partial_t u_h^n, \partial_t Q_h\bar{\eta}_\ast^n) + \tau (\nabla u_h^n, \nabla \partial_t Q_h\bar{\eta}_\ast^n).$$
Using Young's inequality gives
\[
\tau \|\partial_t u^n_h\|^2 + \frac{1}{2}(\|\nabla u^n_h\|^2 - \|\nabla u^{n-1}_h\|^2) \\
\leq 2\tau \alpha^n_2 \|u^n_h\|^2 + \frac{1}{4}\tau \|\partial_t u^n_h\|^2 + \frac{1}{2}\tau \alpha_1 \|u^n_h\|^2 + \frac{1}{2}\tau \alpha_0 \|Q_h \partial_t \tilde{\eta}^n\|^2 \\
+ 2\tau \|Q_h \partial_t \tilde{\eta}^n\|^2 + \frac{1}{2}\tau \|\nabla u^n_h\|^2 + \frac{1}{2}\tau \|\nabla Q_h \partial_t \tilde{\eta}^n\|^2.
\] (4.13)

Summing up (4.12) and (4.13) and then the resulting inequality over \(n = 1, 2, \ldots, k \leq M\), we come to
\[
\frac{1}{2}(\|u^k_h\|^2 + \|\nabla u^k_h\|^2) + \frac{1}{2}\tau \sum_{n=1}^{k} \|\partial_t u^n_h\|^2 \\
\leq \frac{1}{2}(\|\mu_h\|^2 + \|\nabla \mu_h\|^2) + 3\tau (\alpha_1 + \alpha_2) \sum_{n=1}^{k} \|u^n_h\|^2 + C \|\eta_h\|^2_{H^1(0,T;L^2(\Omega))},
\]
where we have used the stability of \(Q_h\) in Lemma 4.1 to obtain
\[
\tau \sum_{n=1}^{k} \|\nabla Q_h \partial_t \tilde{\eta}^n\|^2 \leq C \tau \sum_{n=1}^{k} \|\nabla \partial_t \tilde{\eta}^n\|^2 \leq C \|\nabla \eta_h\|^2_{H^1(0,T;L^2(\Omega))}.
\]

Lemma 4.6 For any sequence \(\{p_h, \mu_h\}\) in \(V_h \times V_h\) converging to \(\{p, \mu\}\) weakly in \(H^1(\Omega) \times H^1(\Omega)\) as \(h \to 0\), we have
\[
\int_{\Omega} \left( u^n_{h0} (p_h, \mu_h)(x) - z_\theta(x) \right)^2 \, dx \to \int_{\Omega} \left( u(p, \mu)(x, \theta) - z_\theta(x) \right)^2 \, dx,
\]
\[
\tau \sum_{n=0}^{M} \alpha_n \int_\omega (u^n(p_h, \mu_h) - z^n)^2 \, dx \to \int_0^T \int_\omega (u(p, \mu) - z)^2 \, dx \, dt
\]
when \(h \to 0\) and \(\tau \to 0\).

Proof. For \(1 \leq n \leq M\), we shall use the notation
\[
u^n = u(p, \mu)(\cdot, t^n), \quad \bar{u}^n = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} u(p, \mu)(\cdot, t) \, dt
\]
and for \(n = 0\), \(\bar{u}^0 = u^0 = \mu\).
We choose $\phi = \tau^{-1}\phi_h$ in (3.3), then integrate over $[t^{n-1}, t^n]$ to obtain
\[
\int_\Omega \partial_t u^n \phi_h dx + \int_\Omega \nabla \bar{u}^n \cdot \nabla \phi_h dx = \int_\Omega p(x) \bar{u}^n \phi_h dx \quad \forall \phi_h \in V_h. \tag{4.14}
\]
Now subtracting this equation from (4.7) yields
\[
\int_\Omega \partial_t (u_h^n - u^n) \phi_h dx + \int_\Omega \nabla (u_h^n - \bar{u}^n) \cdot \nabla \phi_h dx = \int_\Omega (p_h u_h^n - p \bar{u}^n) \phi_h dx.
\]
Set $\eta_h^n = u_h^n - Q_h \bar{u}^n$, we next show that $\eta_h^n$ converges to zero as $h, \tau \to 0$. Noting the form $u_h^n$ in (4.6) and the definition of $Q_h$, we know $u_h^n - Q_h \bar{u}^n \in V_h$. Thus we can take $\phi_h = \tau (u_h^n - Q_h \bar{u}^n)$ as a test function in the above equation and obtain
\[
\int_\Omega \eta_h^n \partial_t \eta_h^n dx + \tau \int_\Omega |\nabla \eta_h^n|^2 dx = \int_\Omega (p_h u_h^n - p \bar{u}^n) \phi_h dx.
\]
Summing up over $n = 1, 2, \cdots, k \leq M$ and using the inequality \[\frac{1}{2} a^2 - \frac{1}{2} b^2 \leq (a - b) a \quad \forall a, b > 0,\]
we derive
\[
\frac{1}{2} \|\eta_h^k\|_{L^2(\Omega)}^2 - \frac{1}{2} \|\eta_h^0\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^k \|\nabla \eta_h^n\|_{L^2(\Omega)}^2 \\
\leq \tau \alpha_1 \sum_{n=1}^k \|\eta_h^n\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^k \int_\Omega p_h (Q_h \bar{u}^n - \bar{u}^n) \eta_h^n dx + \tau \sum_{n=1}^k \int_\Omega (p_h - p) \bar{u}^n \eta_h^n dx
\]
\[
+ \tau \sum_{n=1}^k \int_\Omega \partial_t (u_h^n - \bar{u}^n) \eta_h^n dx + \tau \sum_{n=1}^k \int_\Omega \partial_t (\bar{u}^n - Q_h \bar{u}^n) \eta_h^n dx
\]
\[
+ \tau \sum_{n=1}^k \int_\Omega \nabla (\bar{u}^n - Q_h \bar{u}^n) \cdot \nabla \eta_h^n dx
\]
\[
\equiv: \sum_{m=1}^6 (I)_m. \tag{4.15}
\]
We now estimate the terms $(I)_m$ one by one. First for $(I)_2$, using the bound of $p_h$ and Cauchy-Schwarz inequality we have
\[
(I)_2 \leq \frac{1}{2} \tau \alpha_1 \sum_{n=1}^k \|\eta_h^n\|_{L^2(\Omega)}^2 + \frac{1}{2} \tau \alpha_1 \sum_{n=1}^k \|Q_h \bar{u}^n - \bar{u}^n\|_{L^2(\Omega)}^2
\]
\[
\leq \frac{1}{2} \tau \alpha_1 \sum_{n=1}^k \|\eta_h^n\|_{L^2(\Omega)}^2 + \frac{1}{2} \alpha_1 \int_0^T \|Q_h u(t) - u(t)\|_{L^2(\Omega)}^2 dt.
\]
For (I)3, we can easily derive

\[
(I)_3 \leq \tau \sum_{n=1}^{k} \|p_{n} - p\| \|\eta_{n}\| \sup_{x \in \Omega} |\tilde{u}_{n}| \\
\leq \|p_{n} - p\| \sum_{n=1}^{k} \|\eta_{n}\| \int_{\tau_{n-1}}^{\tau_{n}} \sup_{x \in \Omega} |u(x,t)| dt \\
\leq \left( \int_{0}^{T} \sup_{x \in \Omega} |u(x,t)|^{2} dx \right)^{1/2} \|p_{n} - p\| \left( \tau \sum_{n=1}^{k} \|\eta_{n}\|^{2} \right)^{1/2} \\
\leq C \tau \sum_{n=1}^{k} \|\eta_{n}\|^{2} + C\|p_{n} - p\|^{2}.
\]

To estimate (I)4, we use the formula

\[
\sum_{n=1}^{k} (a_{n} - a_{n-1}) b_{n} = a_{k} b_{k} - a_{0} b_{0} - \sum_{n=1}^{k} a_{n-1} (b_{n} - b_{n-1})
\]

(4.16)

and the stability estimates of \(u_{h}^{n}\) and \(Q_{h}\) in Lemmas 4.1 and 4.5 to obtain that

\[
(I)_4 = \int_{\Omega} (u^{k} - \tilde{u}^{k}) \eta_{h}^{k} dx - \tau \sum_{n=1}^{k} \int_{\Omega} (u^{n-1} - \tilde{u}^{n-1}) \partial_{t} \eta_{h}^{n} dx \\
\leq \sqrt{\tau} \left\{ \int_{t_{1}}^{t_{k}} \|u_{t}\|^{2} dt \right\}^{1/2} + \tau \left\{ \tau \sum_{n=1}^{k} \|\partial_{t} \eta_{h}^{n}\|^{2} \right\}^{1/2} \int_{0}^{T} \|u_{t}\|^{2} dt \right\}^{1/2} \\
\leq C \sqrt{\tau}.
\]

Using (4.16), Lemmas 4.1 and 4.5 again, the term (I)5 can be estimated as follows:

\[
(I)_5 = \int_{\Omega} (\tilde{u}^{k} - Q_{h} \tilde{u}^{k}) \eta_{h}^{k} dx - \int_{\Omega} (u^{0} - Q_{h} u^{0}) \eta_{h}^{0} dx \\
- \tau \sum_{n=1}^{k} \int_{\Omega} (\tilde{u}^{n-1} - Q_{h} \tilde{u}^{n-1}) \partial_{t} \eta_{h}^{n} dx \\
\leq C h \|\nabla \tilde{u}^{k}\| \|\eta_{h}^{k}\| + \|\mu - Q_h \mu\| \left( \|\mu_{n} - \mu\| + \|\mu - Q_{h} \mu\| \right) \\
+ \left( \tau \sum_{n=1}^{k} \|\partial_{t} \eta_{h}^{n}\|^{2} \right)^{1/2} \left( \tau \sum_{n=1}^{k} \|\tilde{u}^{n-1} - Q_{h} \tilde{u}^{n-1}\|^{2} \right)^{1/2} \\
\leq \frac{1}{4} \|\eta_{h}^{k}\|^{2} + Ch^{2} \|u\|_{L^{2}(0,T;H^{1}(\Omega))}^{2} + C \left( \|\mu - Q_{h} \mu\|^{2} + \|\mu_{n} - \mu\|^{2} \right) \\
+ C \left( \int_{0}^{T} \|u(t) - Q_{h} u(t)\|^{2} dt \right)^{1/2}.
\]
Finally for (I)₆ we have

\[(I)_6 \leq \tau \sum_{n=1}^{k} \| \nabla (u^n - Q_h u^n) \| \| \nabla \eta_h^n \| \]

\[\leq k \frac{1}{4} \tau \sum_{n=1}^{k} \| \nabla \eta_h^n \|^2 + \int_0^T \| u(t) - Q_h u(t) \|^2 dt.\]

Substituting the above estimates for (I)₂ through (I)₆ into (4.15), then using the approximation properties of Qₜ and the discrete Gronwall’s inequality we obtain the following convergence

\[\max_{1 \leq n \leq M} \| u_h^n - Q_h u^n \| \to 0 \quad \text{and} \quad \tau \sum_{n=1}^{k} \| \nabla (u_h^n - Q_h u^n) \|^2 \to 0\]

when \( h \) and \( \tau \) both tend to 0, this with the relation

\[u_h^n - \bar{u}^n = (u_h^n - Q_h \bar{u}^n) + (Q_h \bar{u}^n - \bar{u}^n)\]

implies the convergence

\[\max_{1 \leq n \leq M} \| u_h^n - \bar{u}^n \| \to 0 \quad \text{and} \quad \tau \sum_{n=1}^{k} \| \nabla (u_h^n - \bar{u}^n) \|^2 \to 0. \quad (4.17)\]

Now we are ready to show the desired convergence of Lemma 4.6. We first make the following manipulations:

\[(II) \equiv \tau \sum_{n=0}^{M} \alpha_n \int_{\omega} (u_h^n(p_h, \mu_h) - z)^2 dx - \int_0^T \int_{\omega} (u(p, \mu) - z)^2 dx dt\]

\[= \tau \sum_{n=0}^{M} \alpha_n \int_{\omega} (u_h^n - z)^2 dx - \sum_{n=1}^{M} \int_{t_{n-1}}^{t_n} \int_{\omega} (u - z)^2 dx dt\]

\[= \frac{\tau}{2} \sum_{n=1}^{M} \int_{\omega} \{(u_h^n - z)^2 + (u_h^{n-1} - z^{n-1})^2\} dx - \sum_{n=1}^{M} \int_{t_{n-1}}^{t_n} (u - z)^2 dx dt\]

\[= \frac{\tau}{2} \sum_{n=1}^{M} \left\{ \int_{\omega} \{(u_h^n - z)^2 + (u_h^{n-1} - z^{n-1})^2\} dx - \int_{\omega} \{(u^n - z)^2 + (u^{n-1} - z^{n-1})^2\} dx\right\}

+ \sum_{n=1}^{M} \int_{t_{n-1}}^{t_n} \left\{ \frac{(u^n - z)^2 + (u^{n-1} - z^{n-1})^2}{2} - (u - z)^2 \right\} dx dt\]

\[\equiv: (II)_1 + (II)_2.\]
By (4.10), \((\Pi)_2\) can be bounded by
\[
|((\Pi)_2)| \leq 2\tau \int_\omega \left( \int_0^T (u - z)^2 dt \right)^{1/2} \left( \int_0^T (u - z)^2 dt + \max_{t \in [0,T]} (u - z)^2 \right)^{1/2} dx
\]
\[
\leq 2\tau \left( \int_0^T \int_\omega (u - z)^2 dxdt \right)^{1/2} \left( \|u - z\|_{L^2(\omega)}^2 + \|u - z\|_{C(\omega)}^2 \right)^{1/2}
\]
\[
\leq C_\tau,
\]
while \((\Pi)_1\) can be estimated as follows:
\[
|((\Pi)_1)| = \left| \tau \sum_{n=0}^M \alpha_n \int_\omega \left\{ (u_h^n - z^n)^2 - (u^n - z^n)^2 \right\} dx \right|
\]
\[
\leq \tau \left( \sum_{n=0}^M \int_\omega (u_h^n - u^n)^2 dx \right)^{1/2} \left( \sum_{n=0}^M \int_\omega (u_h^n - 2z + u^n)^2 dx \right)^{1/2}
\]
\[
\leq C \max_{0 \leq n \leq M} \|u_h^n - u^n\|_{L^2(\omega)},
\]
thus \((\Pi) \to 0\) as \(h \to 0\) and \(\tau \to 0\) using (4.17). This proves the second part of Lemma 4.6 while the first part can be shown similarly. \(\Box\)

Finally we can state the following convergence theorem about the finite element problem (4.5)-(4.7):

**Theorem 4.2** Let \(\{p_h^*, \mu_h^*\}_{h>0}\) be a sequence of minimizers to the finite element problem (4.5)-(4.7). Then each subsequence of \(\{p_h^*, \mu_h^*\}_{h>0}\) has a subsequence converging in \(L^2(\Omega)\) to a minimizer of the continuous problem (3.1)-(3.3). If the minimizers of the continuous problem are unique, then the whole sequence \(\{p_h^*, \mu_h^*\}_{h>0}\) converges to the unique minimizer of (3.1)-(3.3).

**Proof.** It is not difficult to see that \(J_h(p_h^*, \mu_h^*) \leq C\) for some constant \(C\) independent of \(\tau\) and \(h\). This implies that \(\{p_h^*\}\) and \(\{\mu_h^*\}\) are both bounded in \(H^1(\Omega)\) and so there exists a subsequence of \(\{p_h^*\}\) and \(\{\mu_h^*\}\) each, still denoted as \(\{p_h^*\}\) and \(\{\mu_h^*\}\), such that
\[
p_h^* \to p^* \text{ weakly in } H^1(\Omega), \quad \mu_h^* \to \mu^* \text{ weakly in } H^1(\Omega).
\]
Then for any $p \in K_1$ and $\mu \in K_2$, we have by Lemma 4.6 and the properties of $\pi_h$ in (4.1)-(4.2) that

$$J(p^*, \mu^*) \leq \lim_{h \to 0} \frac{T}{2} \sum_{n=0}^{M} \alpha_n \int_\omega \left( u_h^n(p_h^*, \mu_h^*) - z^n \right)^2 dx + \lim_{h \to 0} \frac{1}{2} \int_\Omega \left( u_h^0(p_h^*, \mu_h^*) - z_0 \right)^2 dx$$

$$+ \beta \liminf_{h \to 0} \| \nabla p_h^* \|^2 + \gamma \liminf_{h \to 0} \| \nabla \mu_h^* \|^2$$

$$\leq \liminf_{h \to 0} J_h(p_h^*, \mu_h^*) \leq \liminf_{h \to 0} J_h(\pi_h p, \pi_h \mu)$$

$$= \frac{1}{2} \int_0^T \int_\omega (u(p, \mu) - z)^2 dx \, dt + \frac{1}{2} \int_\Omega (u(x, \theta) - z_0(x))^2 dx$$

$$+ \beta \| \nabla p \|^2 + \gamma \| \nabla \mu \|^2$$

$$= J(p, \mu),$$

that is, $\{p_h^*, \mu_h^*\}$ is a minimizer of the problem (3.1)-(3.3). The proof of the last statement is standard. \qed
Chapter 5

Numerical algorithms

5.1 Gateaux derivative

This chapter is devoted to some iterative methods for solving the discretized finite element minimization of $J_h(p_h, \mu_h)$ over $K_{1h}$ and $K_{2h}$ in (4.5)-(4.7). For the purpose we need to calculate the Gateaux derivative of $J_h(p_h, \mu_h)$ at any given direction $(q_h, \lambda_h) \in V_h \times V_h$, this derivative will be written as $J_h(p_h, \mu_h)'(q_h, \lambda_h)$. Note that the directions for $\mu_h$ are searched only in $V_h$ instead of $V_h$ because of the considered Dirichlet boundary condition. This can be easily adapted for Neumann boundary conditions.

First of all, we note that the Gateaux derivative for the discrete parabolic solution $u_h^n(p_h, \mu_h)$ at any given direction $(q_h, \lambda_h) \in V_h \times V_h$, denoted as $u_h^n(p_h, \mu_h)'(q_h, \lambda_h)$, solves the following discrete system: $U_h^n = u_h^n(p_h, \mu_h)'(q_h, \lambda_h) \in \bar{V}_h$ such that

$$
\int_{\Omega} \partial_t U_h^n \phi_h dx + \int_{\Omega} \nabla U_h^n \cdot \nabla \phi_h dx = \int_{\Omega} p_h U_h^n \phi_h dx + \int_{\Omega} q_h u_h^n(p_h, \mu_h) \phi_h dx \quad \forall \phi_h \in \bar{V}_h .
$$

Using this, we can easily obtain $J_h(p_h, \mu_h)'(q_h, \lambda_h)$ as follows:

$$
J_h(p_h, \mu_h)'(q_h, \lambda_h) = \tau \sum_{n=0}^{M} \alpha_n \int_{\omega} (u_h^n(p_h, \mu_h) - z^n) u_h^n(p_h, \mu_h)'(q_h, \lambda_h) dx
$$
\[ \begin{align*}
+ \int_{\Omega} (u_h^{n_0}(p_h, \mu_h) - z_\theta) u_h^{n_0}(p_h, \mu_h)'(q_h, \lambda_h) dx \\
+ 2\beta \int_{\Omega} \nabla p_h \cdot \nabla q_h dx + 2\gamma \int_{\Omega} \nabla \mu_h \cdot \nabla \lambda_h dx. \tag{5.2}
\end{align*} \]

From (5.2) we see that the evaluation of the Gateaux derivative of \( J_h(p_h, \mu_h) \) is very expensive. It requires solving the discrete parabolic system \( 2m \) times \( (m = \dim(V_h)) \) as the derivative \( u_h^n(p_h, \mu_h)'(q_h, \lambda_h) \) is needed for \( 2m \) directions \( \{q_h, \lambda_h\} \). Thus it would be very difficult even for some powerful supercomputers if we use (5.2) to calculate the derivative of \( J_h(p_h, \mu_h) \). But fortunately this can be avoided by means of the adjoint equation technique (cf. [24, 19]). For this purpose, we introduce a discrete sequence \( \{w_n^h\}_{n=0}^M \subset V_h \) such that \( w_0^h = 0 \) and \( w_n^h \in \mathcal{V}_h \) for \( n < M \) solves the discrete backward parabolic equation

\[ \begin{align*}
- \int_\Omega \partial_t w_n^h \phi_h dx + \int_\Omega \nabla w_n^{n-1} \cdot \nabla \phi_h dx - \int_\Omega p_h w_n^{n-1} \phi_h dx \\
= \tau \alpha_n \int_\Omega (u_h^n(p_h, \mu_h) - z^n) \phi_h dx \\
+ k_n \int_\Omega (u_h^n(p_h, \mu_h) - z_\theta) \phi_h dx \quad \forall \phi_h \in \mathcal{V}_h \tag{5.3}
\end{align*} \]

where \( k_{n_0} = 1 \) and \( k_n = 0 \) for \( n \neq n_0 \).

Using the sequence \( \{w_n^h\}_{n=0}^M \) and taking the test function \( \phi_h = u_h^n(p_h, \mu_h)'(q_h, \lambda_h) \) in (5.3), the first two terms on the right-hand side of (5.2) can be written as

\[ \begin{align*}
J_h^1 = \tau \alpha_0 \int_\Omega (\mu_h - z^0) \lambda_h dx + \tau \sum_{n=1}^M \alpha_n \int_\Omega (u_h^n(p_h, \mu_h) - z^n) U_h^n dx \\
+ \sum_{n=1}^M k_n \int_\Omega (u_h^n(p_h, \mu_h) - z_\theta) U_h^n dx \\
= \tau \alpha_0 \int_\Omega (\mu_h - z^0) \lambda_h dx + \sum_{n=1}^M \int_\Omega -\partial_r w_h^n U_h^n dx \\
+ \sum_{n=1}^M \left\{ \int_\Omega \nabla w_h^{n-1} \cdot \nabla U_h^n dx - \int_\Omega p_h w_h^{n-1} U_h^n dx \right\}.
\end{align*} \]

Using the formula (4.16) and then (5.1) we obtain

\[ \begin{align*}
J_h^1 = \tau \alpha_0 \int_\Omega (\mu_h - z^0) \lambda_h dx + \tau^{-1} \int_\Omega w_h^0 \lambda_h dx +
\end{align*} \]
\[
\begin{align*}
&+ \sum_{n=1}^{M} \left\{ \int_{\Omega} w_{h}^{n-1} \partial_{\nu} U_{h}^{n} \, dx + \int_{\Omega} \nabla w_{h}^{n-1} \cdot \nabla U_{h}^{n} \, dx - \int_{\Omega} p_{h} w_{h}^{n-1} U_{h}^{n} \, dx \right\} \\
&= \tau \alpha_{0} \int_{\omega} \left( \mu_{h} - z^{0} \right) \lambda_{h} \, dx + \tau^{-1} \int_{\Omega} w_{h}^{0} \lambda_{h} \, dx + \sum_{n=1}^{M} \int_{\Omega} q_{h} \left( p_{h}, \mu_{h} \right) w_{h}^{n-1} \, dx.
\end{align*}
\]

Plugging this in (5.2) we derive a very simple formula for the evaluation of the derivative of \( J_{h}(p_{h}, \mu_{h}) \):

\[
J_{h}(p_{h}, \mu_{h})'(q_{h}, \lambda_{h})
\]

\[
= \tau \alpha_{0} \int_{\omega} \left( \mu_{h} - z^{0} \right) \lambda_{h} \, dx + \tau^{-1} \int_{\Omega} w_{h}^{0} \lambda_{h} \, dx + \sum_{n=1}^{M} \int_{\Omega} q_{h} \left( p_{h}, \mu_{h} \right) w_{h}^{n-1} \, dx
\]

\[
+ 2\beta \int_{\Omega} \nabla p_{h} \cdot \nabla q_{h} \, dx + 2\gamma \int_{\Omega} \nabla \mu_{h} \cdot \nabla \lambda_{h} \, dx.
\]

(5.4)

### 5.2 Nonlinear single-grid gradient method

With the formula (5.4) we are now ready to present the following gradient method for solving the discrete minimization problem (4.5)-(4.7).

**Gradient method I.** Given an initial guess \((p_{h}^{(0)}, \mu_{h}^{(0)}) \in K_{1h} \times K_{2h}\).

(a) Compute \(u_{h}^{0} = \mu_{h}^{(0)} \) and \(u_{h}^{n} = u_{h}^{n}(p_{h}^{(0)}, \mu_{h}^{(0)}) = Q_{h}^{n} + \hat{u}_{h}^{n} \) by solving

\[
\left( \partial_{\nu} u_{h}^{n} \right) \left( \phi_{h} \right) + \left( \nabla u_{h}^{n} \right) \left( \nabla \phi_{h} \right) = \left( p_{h}^{n} \right) u_{h}^{n} \left( \phi_{h} \right) \quad \forall \phi_{h} \in \hat{V}_{h}.
\]

Compute \(w_{h}^{M} = 0\) and \(w_{h}^{n-1} \in \hat{V}_{h}\) for \(n = M, M - 1, \ldots, 1\) by solving

\[
-\left( \partial_{\nu} w_{h}^{n} \right) \left( \phi_{h} \right) + \left( \nabla w_{h}^{n-1} \right) \left( \nabla \phi_{h} \right) = \left( p_{h}^{(n)} \right) w_{h}^{n-1} \left( \phi_{h} \right)
\]

\[
= \tau \alpha_{n} \int_{\omega} \left( u_{h}^{n} - z^{0} \right) \phi_{h} \, dx + k_{n} \left( u_{h}^{n} - z^{0}, \phi_{h} \right) \quad \forall \phi_{h} \in \hat{V}_{h}.
\]

(b) Compute the components of \( \left( J_{h}(p_{h}^{(0)}, \mu_{h}^{(0)}) \right)' \) corresponding to all the basis functions \( \{ \phi_{m} \} \) from \( V_{h} \) and \( \{ \phi_{l} \} \) from \( \hat{V}_{h} \) using (5.4) respectively:

\[
g_{m} = \sum_{n=1}^{M} \int_{\Omega} u_{h}^{n}(p_{h}^{(0)}, \mu_{h}^{(0)}) w_{h}^{n-1} \phi_{m} \, dx + 2\beta \int_{\Omega} \nabla p_{h}^{(0)} \cdot \nabla \phi_{m} \, dx;
\]

\[
g_{l} = \tau \alpha_{0} \int_{\omega} \left( \mu_{h}^{(0)} - z^{0} \right) \phi_{l} \, dx + \tau^{-1} \int_{\omega} w_{h}^{0} \phi_{l} \, dx + 2\gamma \int_{\Omega} \nabla \mu_{h}^{(0)} \cdot \nabla \phi_{l} \, dx.
\]
Set $\tilde{g}_h = \sum_m \tilde{g}_m \phi_m$ and $g_h = \sum_i g_i \phi_i$.

(c) Find $\lambda > 0$ such that $\min_{\lambda > 0} J_h(p_h^{(0)} - \lambda \tilde{g}_h, \mu_h^{(0)} - \lambda g_h)$.

(d) Project onto the constrained sets $K_{1h}$ and $K_{2h}$:

$$
\begin{align*}
\mu_h^* &= \min \left\{ \max\{0, \mu_h^{(0)} - \lambda g_h\}, \alpha_2 \right\}, \\
p_h^* &= \min \left\{ \max\{-\alpha_1, p_h^{(0)} - \lambda \tilde{g}_h\}, \alpha_1 \right\},
\end{align*}
$$

If $|| (p_h^*, \mu_h^*) - (p_h^{(0)}, \mu_h^{(0)}) || \leq$ tolerance, stop;
otherwise set $(p_h^{(0)}, \mu_h^{(0)}) = (p_h^*, \mu_h^*), j = j + 1$, goto (a). □

For the later convenience, we denote the iterate of the above gradient method I as

$$(p_h^*, \mu_h^*) = \text{GradientI}(J_h, p_h^{(0)}, \mu_h^{(0)}, m_h)$$

i.e., $(p_h^*, \mu_h^*)$ is the approximate solution obtained using $m_h$ iterations of the gradient method I with the cost functional $J_h$ and initial guess $(p_h^{(0)}, \mu_h^{(0)})$.

Remark 5.1 There are many existing approaches for finding the step size $\lambda$ in Step (c). Here we use here the parabolic curve search method and it takes about 5 searches on the average in our numerical experiments.

Let us explain the parabolic curve search method for the 1D simple example problem $\min_{\lambda > 0} J(x - \lambda g)$

1. Let $\lambda_0^0 = 0, \lambda_1^0 = 1$, find $\lambda_2^0$ such that

   (a) If $J(x - \lambda_0^0 g) \leq J(x - \lambda_1^0 g)$, then find $\lambda_2^0 \in (0, 1)$ such that $J(x - \lambda_2^0 g) < J(x - \lambda_0^0 g)$.

   (b) If $J(x - \lambda_0^0 g) > J(x - \lambda_1^0 g)$, then find $\lambda_2^0 > 1$ such that $J(x - \lambda_2^0 g) > J(x - \lambda_1^0 g)$.

2. Find a quadratic polynomial $f(\lambda)$ such that for $i = 0, 1, 2$,

   $$f(\lambda_i^k) = J(x - \lambda_i^k g).$$

54
3. Let $\lambda^k_3$ be the minimum of $f(\lambda)$. If $|\lambda^k_3 - \lambda^{k-1}_3| < \text{tolerance}$, go to (4); otherwise, find $\lambda^{k+1}_0, \lambda^{k+1}_1, \lambda^{k+1}_2 \in \{\lambda^k_0, \lambda^k_1, \lambda^k_2, \lambda^k_3\}$ with one of $\lambda^{k+1}_i$ begin $\lambda^k_3$ such that they satisfy the conditions in (1), set $k=k+1$, go to (2).

4. The approximate solution of $\min_{\lambda>0} J(x - \lambda g)$ is $\lambda = \lambda^*_3$.

5.3 Nonlinear multigrid gradient method

Through many numerical experiments we find that the gradient method above converges in most cases globally (even with very bad initial guesses) and stably. In particular, we have observed that the first few iterations of the method often converge very fast and then the convergence slows down significantly. If one still continues with the gradient method after the first few iterations, it will take a great deal of more iterations to reach the tolerance, a few thousand iterations often needed even for one dimensional inverse parabolic problems. This is very much like the performance of the classical iterative methods for solving second order boundary value problems. For the latter, there exists a well-developed multigrid method (MGM) which can deal with such a slow-down very efficiently by making full use of the fast convergence of the first few iterations of the classical iterative methods. The MGM starts with a fine grid and iterates a few times using a classical iterative method (called a smoothing step) and then goes to a coarser grid to solve the residual equation to achieve some coarse correction for the approximate solution obtained on the fine grid, again applying the same iterative method a few iterations for the residual equation. The MGMs have been proved to be very effectively applicable to various PDE problems (direct problems), see [14, 30, 33] and the references therein. However to our knowledge, there seems still no applications of the MGM for solving inverse problems, especially for our currently considered inverse parabolic problems. The first thing one has to solve for this application is to find a way of formulating the MGM for the highly nonlinear minimization problem with constraints involving some initial-boundary
value problems. Clearly this is not straightforward.

In the rest of this chapter we are going to propose a nonlinear MGM for solving the nonlinear minimization system (4.5)-(4.7). And the numerical results will demonstrate its efficiency in solving the inverse parabolic problem considered in this paper.

Assume that we are given a nested set of shape regular triangulations \( \{ \mathcal{T}^{h_k} \}_{k=0}^{N} \), with \( \mathcal{T}^{h_{k+1}} \) being a refinement of \( \mathcal{T}^{h_k} \). And \( \{ V_{h_k} \}_{k=0}^{N} \) are the continuous piecewise linear finite element spaces defined on \( \{ \mathcal{T}^{h_k} \}_{k=0}^{N} \) satisfying

\[
V_{h_0} \subset V_{h_1} \subset \cdots \subset V_{h_N} \equiv V.
\]

Our goal is to solve the discrete minimization problem (4.5)-(4.7), which is defined on the finest space \( V^h \), by making use of the auxiliary coarser spaces \( V_{h_k} \) for \( 0 \leq k < N \).

To do so, we need to introduce some more notation. Corresponding to each coarse triangulation \( \mathcal{T}^{h_k} \), we divide the time interval \([0, T]\) into \( M_k \) equally-spaced subintervals using nodes

\[
0 = t^0_k < t^1_k < \cdots < t^{M_k}_k = T
\]

with \( t^n_k = n\tau_k \) and \( \tau_k = T/M_k \). Similarly to \( K_{1h} \) and \( K_{2h} \) we define two constrained subsets \( K_{1h_k} \) and \( K_{2h_k} \). And more, for the initialization step of the nonlinear multigrid method to be introduced below we have to solve a coarse minimization problem on each coarse space \( V_{h_k} \), which is defined as follows:

\[
\text{minimize } J_k^0(p_{h_k}, \mu_{h_k}) = \frac{\tau_k}{2} \sum_{n=0}^{M_k} \alpha_n \int_\Omega (u^n_{h_k} - z^n)^2 dx + \frac{1}{2} \int_\Omega (u^n_{h_k} - z_0)^2 dx \\
+ \beta \int_\Omega |\nabla p_{h_k}|^2 dx + \gamma \int_\Omega |\nabla \mu_{h_k}|^2 dx \quad (5.5)
\]

over all \( p_{h_k} \in K_{1h_k} \) and \( \mu_{h_k} \in K_{2h_k} \) with \( u^n_{h_k} \equiv u^n_{h_k}(p_{h_k}, \mu_{h_k}) \in V_{h_k} \) satisfying

\[
\begin{align*}
 u^0_{h_k} &= \mu_{h_k} & u^n_{h_k} &= Q_{h_k}^k \tilde{u}^n + \hat{u}^n_{h_k}, \\
 (\partial_t u^n_{h_k}, \phi_{h_k}) + (\nabla u^n_{h_k}, \nabla \phi_{h_k}) &= (p_{h_k} u^n_{h_k}, \phi_{h_k}) \quad \forall \phi_{h_k} \in V^n_{h_k} \quad (5.7)
\end{align*}
\]

56
for \( n = 1, 2, \ldots, M_k \). Here \( \alpha_0 = \alpha_{M_k} = \frac{1}{2} \) and \( \alpha_n = 1 \) for all \( n \neq 0, M_k \) while \( \tilde{w}_n^{n_k} \in V_{h_k} \) and \( n_k > 0 \) is an integer such that \( n_k \tau_k = \theta \).

Another important step of the multigrid method is the coarse grid correction, for which we define the following nonlinear functional:

\[
J_k(p_h + p_{h_k}, \mu_h + \mu_{h_k}) = \frac{\tau_k}{2} \sum_{n=0}^{M_k} \alpha_n \int_{\Omega} (u_{h_k}^n - z^n)^2 \, dx + \frac{1}{2} \int_{\Omega} (u_{h_k}^{n_k} - z_0)^2 \, dx \\
+ \beta \int_{\Omega} |\nabla (p_h + p_{h_k})|^2 \, dx + \gamma \int_{\Omega} |\nabla (\mu_h + \mu_{h_k})|^2 \, dx
\]

where \( p_{h_k} \in V_{h_k} \) and \( \mu_{h_k} \in V_{h_k} \), while \( u_{h_k}^0 = \mu_h + \mu_{h_k} \) and for \( n = 1, 2, \ldots, M_k \),

\[
u_{h_k}^n \equiv u_{h_k}^n (p_h + p_{h_k}, \mu_h + \mu_{h_k}) = Q_{h_k} \eta^n + \tilde{w}_k^{n_k} \in V_{h_k} \text{ solves}
\]

\[
(\partial_t u_{h_k}^n, \phi_{h_k}) + (\nabla u_{h_k}^n, \nabla \phi_{h_k}) = (p_h + p_{h_k}, u_{h_k}^n \phi_{h_k}) \quad \forall \phi_{h_k} \in V_{h_k} - \text{ (5.8)}
\]

Finally, corresponding to the adjoint parabolic solution \( w_{h_k}^n \) given by (5.3), we define an adjoint solution on each coarse space \( V_{h_k} \) such that \( w_{h_k}^{M_h} = 0 \) and \( w_{h_k}^{n-1} \) for \( n = M, \ldots, 1 \) solves

\[
- \int_{\Omega} \partial_t w_{h_k}^n \phi_{h_k} \, dx + \int_{\Omega} \nabla w_{h_k}^{n-1} \cdot \nabla \phi_{h_k} \, dx - \int_{\Omega} (p_h + p_{h_k}) w_{h_k}^{n-1} \phi_{h_k} \, dx \\
= \tau_k \alpha_n \int_{\Omega} (u_{h_k}^n - z^n) \phi_{h_k} \, dx + k_n \int_{\Omega} (u_{h_k}^{n_k} - z_0) \phi_{h_k} \, dx \quad \forall \phi_{h_k} \in V_{h_k} - \text{ (5.9)}
\]

where \( u_{h_k}^n \) is the solution of (5.8).

Before we state the nonlinear multigrid method, we first formulate the gradient method for solving the following minimization problem

\[
\min J_k(p_h + p_{h_k}, \mu_h + \mu_{h_k})
\]

over all \( p_{h_k} \in V_{h_k} \) and \( \mu_{h_k} \in V_{h_k} \) such that \( p_h + p_{h_k} \in K_{1h} \) and \( \mu_h + \mu_{h_k} \in K_{2h} \).

This minimization will constitute the coarse grid correction step in the multigrid method.

**Gradient method C.** Given an initial guess \( p_{h_k}^{(0)} \in K_{1h_k} \) and \( \mu_{h_k}^{(0)} \in K_{2h_k} \).
(a) Solve the parabolic and backward parabolic equations (5.8) and (5.9) for \( \{u_h^n\}_{n=0}^{M_k} \) and \( \{w_h^n\}_{n=0}^{M_k} \).

(b) Compute the components of \( J_k(p_h + p_h^{(0)}, \mu_h + \mu_h^{(0)})' \) corresponding to all the basis functions \( \{\phi_m\} \) from \( V_{h_k} \) and \( \{\phi_l\} \) from \( V_{h_k} \) respectively:

\[
\begin{align*}
\bar{g}_{h_k,m} &= \sum_{n=1}^{M_k} (u_h^n, w_h^{n-1}, \phi_m) + 2\beta (\nabla (p_h + p_h^{(0)}), \nabla \phi_m); \\
g_{h_k,l} &= \tau_k \alpha_0 \int_\omega (\mu_h + \mu_h^{(0)} - z^0) \phi_l dx + \tau_k^{-1} \int_\omega w_h^0 \phi_l dx + 2\gamma (\nabla (\mu_h + \mu_h^{(0)}), \nabla \phi_l).
\end{align*}
\]

Set \( \bar{g}_{h_k} = \sum_{m} \bar{g}_{h_k,m} \phi_m \) and \( g_{h_k} = \sum_{l} g_{h_k,l} \phi_l \).

(c) Find \( \lambda > 0 \) such that \( \min_{\lambda>0} J_k(p_h + p_h^{(0)} - \lambda \bar{g}_{h_k}, \mu_h + \mu_h^{(0)} - \lambda g_{h_k}) \).

(d) Project onto the constrained sets \( K_{1h} \) and \( K_{2h} \):

\[p^*_h = \min \left\{ \max \{-\alpha_1, p_h + p_h^{(0)} - \lambda \bar{g}_{h_k}\}, \alpha_1 \right\}; \]

\[\mu^*_h = \min \left\{ \max \{0, \mu_h + \mu_h^{(0)} - \lambda g_{h_k}\}, \alpha_2 \right\}.\]

Set \((p_h^{(0)}, \mu_h^{(0)}) := (p_h^{(0)} - \lambda \bar{g}_{h_k}, \mu_h^{(0)} - \lambda g_{h_k})\), \( j := j + 1 \), goto (a).

In what follows, we denote the iterate of the gradient method \( C \) above as

\[ (p_h^*, \mu_h^*) = \text{GradientC}(J_k, p_h, \mu_h, p_h^{(0)}, \mu_h^{(0)}, m_k) \]

i.e., \((p_h^*, \mu_h^*)\) is the coarse correction of \((p_h, \mu_h)\), obtained using \( m_k \) iterations of the gradient method \( C \) with the cost functional \( J_k \) and coarse initial guess \((p_h^{(0)}, \mu_h^{(0)})\).

With the above preparation we are now ready to formulate the nonlinear multigrid method for solving the finite element minimization problem (4.5)-(4.7).

**Nonlinear multigrid gradient method.**

Given an initial guess \((p_h^{(0)}, \mu_h^{(0)}) \in K_{1ho} \times K_{2ho} \) on the coarsest finite element space \( V_{ho} \).

I. Coarse grid initialization.
For $k = 0, 1, \ldots, N - 1$, do

If $k \neq 0$, calculate $p_{hk}^{(0)} = \Pi_{k-1}^k p_{hk-1}^*$ and $\mu_{hk}^{(0)} = \Pi_{k-1}^k \mu_{hk-1}^*$.

Compute $(p_{hk}^*, \mu_{hk}^*) = \text{GradientI}(J_k, p_{hk}^{(0)}, \mu_{hk}^{(0)}, m_k)$.

end;

Compute $p_h^{(0)} = \Pi_{N-1}^N p_{hN-1}^*$ and $\mu_h^{(0)} = \Pi_{N-1}^N \mu_{hN-1}^*$.

II. Smoothing and coarse grid correction. Set the iteration number $j = 0$.

(a) Set $p_h^{(0)} = p_h^*$ and $\mu_h^{(0)} = \mu_h^*$.

For $k = N, N - 1, \ldots, 1, 0$, do

If $k \neq N$, compute $p_h^{(0)} = p_h^*$ and $\mu_h^{(0)} = \mu_h^*$.

Compute $(p_h^*, \mu_h^*) = \text{GradientC}(J_k, p_h^{(0)}, \mu_h^{(0)}, 0, 0, n_k)$.

(b) If $\| (p_h^*, \mu_h^*) - (\overline{p}_h^{(0)}, \overline{\mu}_h^{(0)}) \| \leq \text{tolerance}$, stop;

otherwise set $p_h^{(0)} = p_h^*$ and $\mu_h^{(0)} = \mu_h^*$, $j := j + 1$, goto (a). □

Remark 5.2 The operators $\Pi_{k-1}^k$ in Step I can be any interpolation from $V_{k-1}$ onto $V_h$. We use the natural finite element injections since $V_{k-1} \subset V_h$. For the starting values $p_{hk}^{(0)}$ and $\mu_{hk}^{(0)}$ in Step II, we take the most natural zero.

Remark 5.3 The major computational costs of the nonlinear MGM are from the iterations on the finest grid, all the costs of iterations on the coarse grids are very small compared to the ones on the finest grid since the unknowns on the coarse spaces are much scaled down compared with those on the fine space. This is one of the essential ingredients of MGMs.
Chapter 6

Numerical experiments

6.1 One dimensional examples

In this chapter we show some numerical experiments on the nonlinear multi-grid gradient method (MGM) proposed in the last chapter for the simultaneous reconstruction for the initial data $\mu(x)$ and the coefficient function $p(x)$.

Firstly, we consider some 1-dimensional examples. Our test problem is

$$\frac{\partial u}{\partial t}(x,t) = \Delta u(x,t) + p(x)u(x,t) + f(x,t), \quad (x,t) \in \Omega \times (0,T)$$

$$u(x,0) = \mu(x), \quad x \in \Omega$$

$$u(x,t) = \eta(x), \quad (x,t) \in \partial \Omega \times (0,T).$$

The observed data will be taken to be

$$z_0(x) = u(x,0), \quad x \in \Omega; \quad z(x,t) = u(x,t), \quad (x,t) \in \omega \times (0,T)$$

where $\Omega = (0,1), T = 1, \theta = 1/2$ and $\omega = (1/4,3/4)$. Most parameters related in the algorithm are attached in each figure. The error shown is the relative $L^2$-norm error between the exact parameter $\{p(x), \mu(x)\}$ to be identified and the numerically reconstructed parameter $\{p_h(x), \mu_h(x)\}$. The upper bounds $\alpha_1$ and $\alpha_2$ in the two constrained sets $K_1$ and $K_2$ are both taken to be 100. We take the mesh size $h = 1/40$ in nonlinear SGM and we take 4-level nested finite element
grids with mesh sizes $h_3 = 1/40, h_2 = 1/20, h_1 = 1/10, h_0 = 1/5$, the parameters $m_k = 5$ and $n_k = 5$ in the nonlinear MGM. And for better numerical reconstruction, we have normalized the gradients $\bar{g}$ and $g$. In our implementations, we normalize the gradient $g_h$ of SGM (Gradient method I) by dividing it by $h$, but keep $\bar{g}_h$ unchanged (as it is in the SGM). Similarly, we normalize the gradient $g_{h_k}$ of MGM (Gradient method C) by dividing it by $h_k$, but keep the gradient $\bar{g}_{h_k}$ unchanged (as it is in MGM). These conditions are all the same in examples 1, 2 and 3.

In our numerical implementations, we always assume the observation data have some errors. That is, instead of using the exact data $z_0(x)$ and $z(x,t)$, we take the noised data of the following form:

$$ z_0^\delta(x) = z_0(x) + \delta \sin(3\pi x), $$
$$ z^\delta(x,t) = z(x,t) + \delta \sin(3\pi x). $$

Here $\delta$ is a noise level parameter, and we use the function $\sin(3\pi x)$ instead of the random function just for the convenience of readers' numerical verification.

**Example 1.** We take the exact solution $u(x,t)$ as

$$ u(x,t) = \exp(-3t - x^2 + x + 1/2) + \exp(t + x^2 - x - 1/2) $$

and the identifying coefficient $p(x)$ and the initial data $\mu(x)$ as

$$ p(x) = -4x^2 + 4x - 2, \quad \mu(x) = \exp(-x^2 + x + 1/2) + \exp(x^2 - x - 1/2). $$

Figures 6.1-6.2 show the exact coefficient $p(x)$ and the initial data $\mu(x)$ (the dashed lines) and the numerically reconstructed $\{p_h(x), \mu_h(x)\}$ (the solid line) when the noise level is $\delta = 1\%$. The initial guess $\{p_0^h(x), \mu_0^h(x)\}$ is taken to be the constant pair $\{-2.5, 2.0\}$ at all 6 grid points on the coarsest grid with $h_0 = 1/5$. That is, we start with only these 6 nodal point values and then iterate with the nonlinear MGM. Clearly, this initial guess is not good at all, but the
MGM converges very stably and fast (the result shown is obtained from the 10th iteration), the reconstruction of both $p$ and $\mu$ appears to be rather satisfactory with 1% noise present.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure6_1}
\caption{$p_h^{(0)} = -2.5, \beta = 5E-6, \text{iter}=10, \text{error}=2.8E-2, \delta = 1\%$.}
\end{figure}

Example 2. This example tests the effectiveness of the reconstruction of only the initial temperature $\mu(x)$, with the radiative coefficient $p(x)$ known. To construct more general testing functions $p$ and $\mu$, we add a source function $f(x,t)$ to the right-hand side of equation (6.2). Then we take the exact solution $u(x,t)$ as

$$u(x,t) = \sin \pi x(t + 1) + x^2 + 5$$

and the identifying initial data $\mu(x)$ as

$$\mu(x) = x^2 + 5 + \sin \pi x$$

with $p(x) = 3 + \cos(2\pi x)$ given. And the newly added function $f(x,t)$ is then calculated through (6.2) using $u(x,t)$ and $p(x)$.

Figure 6.3 is the exact observation $z_0 = u(x,\theta)$ given at $\theta = 1/2$, very different from the initial temperature $\mu(x)$ to be reconstructed. Figure 6.4 shows the
exact initial data $\mu(x)$ (the dashed lines) and the numerically reconstructed $\mu_h(x)$ (the solid line) obtained at the 10th iteration of the MGM with the noise level $\delta = 1\%$. The initial guess $\mu_h^0(x)$ is taken to be the constant 5.0 on the coarsest grid everywhere. We can see that the reconstructed $\mu_h$ matches the exact $\mu$ very satisfactorily with 1\% noise added in the observation data. We tried then to reduce the size of the subregion $\omega$ from $(1/4, 3/4)$ to $(3/8, 5/8)$ and then to $(7/16, 9/16)$, the reconstructions are the same accurate as the one shown in Figure 6.4 but just we need a bit more iterations: 20 and 25 respectively for the two cases. If we shifted $\omega$ to be in the left half or the right half of $\Omega = (0, 1)$, the reconstructions are still very satisfactory but not as accurate as in the previous cases. Finally we tried to switch off the observation on the subregion $\omega \times (0, 1)$ with only the measurements given by $z_\theta = u(\cdot, \theta)$, the reconstructions are always terrible even without any noise present. This is consistent with the well-known instability on the backward parabolic problem.

**Example 3.** We take the exact solution $u(x, t)$ as

$$u(x, t) = \exp(t)(2 - \sin 2\pi x)$$
Figure 6.3: The exact observation \( z_0 = u(x, 1/2) \).

Figure 6.4: \( \mu^0_\delta = 5.0, \gamma = 1.0E-7, \text{error} = 2.3E-3, \text{iter} = 10, \delta = 1\% \).
and the identifying initial data $\mu(x) = u(x,0)$ and the radiative coefficient $p(x)$ as 

$$p(x) = \frac{2 - \sin 2\pi x - 4\pi^2 \sin 2\pi x}{2 - \sin 2\pi x}.$$ 

We see that the function $p$ has a very large fluctuation between its maximum and minimum (difference $\geq 50$). Reconstruction of a function like $p$ with such a large fluctuation between its function values is very hard from the numerical point of view. Figures 6.5-6.6 show the exact coefficient $p(x)$ and the initial data $\mu(x)$ (the dashed lines) and the numerically reconstructed $\{p_h(x), \mu_h(x)\}$ (the solid line) with the noise level $\delta = 1\%$. The initial guess $\{p_h^0(x), \mu_h^0(x)\}$ is taken to be the constant pair $\{0.5, 2.0\}$ everywhere, which is certainly a very bad initial guess for the reconstruction, but the nonlinear MGM converges very stably and the reconstruction obtained using only 20 iterations (Fig. 6.5) appears to be rather satisfactory, considering such a large fluctuation of function values for $p$.

Figure 6.5: $p_h^{(0)} = 0.5$, $\beta = 5E-9$, iter=20, error=1.3E-1.
6.2 Two dimensional examples

We now consider some 2-dimensional examples. Our test problem becomes

\[
\frac{\partial u}{\partial t} (x, y, t) = \Delta u(x, y, t) + p(x, y)u(x, y, t) + f(x, y, t), \quad (x, y, t) \in \Omega \times (0, T) \\
u(x, y, 0) = \mu(x, y), \quad (x, y) \in \Omega \\
u(x, y, t) = \eta(x, y, t), \quad (x, y, t) \in \partial \Omega \times (0, T).
\]

The observed data will be taken to be

\[
z_{\theta}(x, y) = u(x, y, \theta), \quad (x, y) \in \Omega; \quad z(x, y, t) = u(x, y, t), \quad (x, y, t) \in \omega \times (0, T)
\]

where \(\Omega = (0, 1) \times (0, 1), T = 1, \theta = 1/2\) and \(\omega = (1/4, 3/4) \times (1/4, 3/4)\).

Now we assume the radiative coefficient \(p(x, y)\) is known and reconstruct the initial temperature \(\mu(x, y)\).

For nonlinear SGM, we take the mesh size \(h = 1/40\) and for nonlinear MGM, we take 3-level nested finite element grids with mesh sizes \(h_2 = 1/40, h_1 = 1/20, h_0 = 1/10\), the parameters are still \(m_k = 5\) and \(n_k = 5\). And as 1-dimensional problem, for better numerical reconstruction, we have normalized the gradients \(\bar{g}\) and \(g\).

In our implementations, we normalize the gradient \(g_h\) of SGM (Gradient method...
I) by dividing it by \( h \), but keep \( \bar{g}_h \) unchanged (as it is in the SGM). Similarly, we normalize the gradient \( g_{h_k} \) of MGM (Gradient method C) by dividing it by \( h_k \), but keep the gradient \( \bar{g}_{h_k} \) unchanged (as it is in MGM). As 1-dimensional examples, we also assume the observation data have some errors, we take the noised data of the following form:

\[
\begin{align*}
  z_\delta^\phi(x, y) &= z_\phi(x, y) + \delta \sin(3\pi(x + y)), \\
  z^\delta(x, t) &= z(x, t) + \delta \sin(3\pi(x + y)).
\end{align*}
\]

These conditions are all the same in examples 4, 5, 6 and 7.

**Example 4.** We take the exact solution \( u(x, y, t) \) as

\[
  u(x, y, t) = 4 + \sin \pi xy \sin \pi t + \cos \pi xy \cos \pi t + t^2,
\]

the radiative coefficient \( p(x, y) \) as

\[
  p(x, y) = 2 + \cos \pi xy
\]

and the identifying initial data \( \mu(x, y) = u(x, y, 0) \) is

\[
  \mu(x, y) = 4 + \cos \pi xy.
\]

Figure 6.7 shows the exact observation \( z_\theta = u(x, y, \theta) \) given at \( \theta = 1/2 \) and Figure 6.8 shows the exact initial data \( \mu(x, y) \). We can see that they are very different. We now compare the efficiency of nonlinear single-grid gradient method (SGM) and the multigrid gradient method (MGM). Figure 6.9 shows the numerically reconstructed \( \mu_h(x, y) \) by nonlinear SGM while Figure 6.10 shows the numerically reconstructed \( \mu_h(x, y) \) by nonlinear MGM, both with the noise level \( \delta = 1\% \). The initial guess \( \mu_h^{(0)}(x, y) \) is taken to be constant 5 everywhere.

We can see that for nonlinear SGM, it needs 140 iterations to have the same accuracy while only 2 iterations are required in nonlinear MGM, and the calculation of the cost functional on finest grid is greatly reduced from 1120 to 90 when nonlinear MGM is instead of nonlinear SGM.
Figure 6.7: The exact observation $z_0 = u(x, y, 1/2)$.

Figure 6.8: The exact initial data needed to be reconstructed $\mu(x, y) = u(x, y, 0)$. 
Figure 6.9: $\mu_h^{(0)} = 5.0$, $\gamma = 1.0 \times 10^{-3}$, single-grid-iter=140, error=1.30E-2, no. of cost functional calculated=1120.

Figure 6.10: $\mu_h^{(0)} = 5.0$, $\gamma = 1.0 \times 10^{-3}$, multigrid-iter=2, error=1.26E-2, no. of cost functional calculated=90.
Example 5. Here is another 2-dimensional example. We take the exact solution $u(x, y, t)$ as

$$u(x, y, t) = 4 - \sin\left(\frac{5\pi x}{4} - \frac{\pi}{8}\right) \sin\left(\frac{5\pi y}{4} - \frac{\pi}{8}\right) \cos 2\pi t + t^2,$$

the radiative coefficient $p(x, y)$ as

$$p(x, y) = 2 + \cos \pi xy$$

and the identifying initial data $\mu(x, y) = u(x, y, 0)$ is

$$\mu(x, y) = 4 - \sin\left(\frac{5\pi x}{4} - \frac{\pi}{8}\right) \sin\left(\frac{5\pi y}{4} - \frac{\pi}{8}\right).$$

Similarly as example 4, Figure 6.11 shows the exact observation $z_\theta = u(x, y, \theta)$ given at $\theta = 1/2$ and Figure 6.12 shows the exact initial data $\mu(x, y)$. Figure 6.13 shows the numerically reconstructed $\mu_h(x, y)$ by nonlinear SGM while Figure 6.14 shows the numerically reconstructed $\mu_h(x, y)$ by nonlinear MGM, both with the noise level $\delta = 1\%$. The initial guess $\mu_h^{(0)}(x, y)$ is taken to be constant 5 everywhere.

We can see that for nonlinear SGM, it needs 95 iterations to have the same accuracy while only 3 iterations are required in nonlinear MGM, and the calculation of the cost functional on finest grid is reduced from 1157 to 209 when nonlinear MGM is instead of nonlinear SGM.
Figure 6.11: The exact observation $z_0 = u(x, y, 1/2)$.

Figure 6.12: The exact initial data needed to be reconstructed $\mu(x, y) = u(x, y, 0)$.
Figure 6.13: $\mu_h^{(0)} = 5.0$, $\gamma = 5.0E-5$, single-grid-iter=95, error=1.030E-2, no. of cost functional calculated=1157.

Figure 6.14: $\mu_h^{(0)} = 5.0$, $\gamma = 5.0E-5$, multigrid-iter=3, error=1.024E-2, no. of cost functional calculated=209.
Now we need to reconstruct both the radiative coefficient \( p(x, y) \) and the initial temperature \( \mu(x, y) \).

**Example 6.** In this example, we need to reconstruct both the radiative coefficient \( p(x, y) \) and the initial temperature \( \mu(x, y) \).

For nonlinear MGM, we take 3-level nested finite element grids with mesh sizes \( h_2 = 1/40, h_1 = 1/20, h_0 = 1/10 \), the parameters are still \( m_k = 5 \) and \( n_k = 5 \). In our implementations, we normalize the gradient \( g_{h_k} \) of MGM (Gradient method C) by dividing it by \( h_k \), but keep the gradient \( g_{h_k} \) unchanged (as it is in MGM). As previous examples, we also assume the observation data have some errors, we take the noised data of the following form:

\[
\begin{align*}
    z^\delta(x, y) &= z(x, y) + \delta \sin(3\pi(x + y)), \\
    z^\delta(x, t) &= z(x, t) + \delta \sin(3\pi(x + y)).
\end{align*}
\]

We take the exact solution \( u(x, y, t) \) as

\[
u(x, y, t) = t^2 + (2 - xy(1 - x)(1 - y)) \cosh(t).
\]

The identifying radiative coefficient \( p(x, y) \) is

\[
p(x, y) = 2 - xy(1 - x)(1 - y),
\]

and the identifying initial data \( \mu(x, y) = u(x, y, 0) \) is

\[
\mu(x, y) = 2 - xy(1 - x)(1 - y).
\]

Figure 6.15 shows the exact observation \( z_\theta = u(x, y, \theta) \) given at \( \theta = 1/2 \). Figure 6.16 shows the exact radiative coefficient \( p(x, y) \) and the exact initial data \( \mu(x, y) \). Figure 6.17 shows the numerically reconstructed \( p_h(x, y) \) by nonlinear MGM and figure 6.18 shows the numerically reconstructed \( \mu_h(x, y) \) by nonlinear MGM. The noise level \( \delta = 1\% \), the initial guess \( p_h(0)(x, y) \) and \( \mu_h(0)(x, y) \) are both taken to be constant 2 everywhere.
Figure 6.15: The exact observation $z_0 = u(x, y, 1/2)$.

Figure 6.16: Both the exact radiative coefficient $p(x, y)$ and initial data $\mu(x, y)$ needed to be reconstructed.
Figure 6.17: $p_h^{(0)} = 2.0$, $\beta = 1.0 \times 10^{-4}$, multi-grid-iter=3, error=2.930E-3, no. of cost functional calculated=126.

Figure 6.18: $\mu_h^{(0)} = 2.0$, $\gamma = 1.0 \times 10^{-3}$, multi-grid-iter=3, error=6.797E-3.
Example 7. Similar to example 6, we need to identify both the radiative coefficient \( p(x, y) \) and the initial data \( \mu(x, y) \). We take the exact solution \( u(x, y, t) \) as

\[
u(x, y, t) = t^2 + (2 - \sin(\pi x)(1 - x) \sin(\pi y)(1 - y)) \cosh(t).
\]

The identifying radiative coefficient \( p(x, y) \) is

\[
p(x, y) = 4 - xy(1 - x)(1 - y),
\]

and the identifying initial data \( \mu(x, y) = u(x, y, 0) \) is

\[
\mu(x, y) = 2 - \sin(\pi x)(1 - x) \sin(\pi y)(1 - y).
\]

For nonlinear MGM, we take 3-level nested finite element grids with mesh sizes \( h_2 = 1/40, h_1 = 1/20, h_0 = 1/10 \), the parameters are still \( m_k = 5 \) and \( n_k = 5 \). In our implementations, we normalize the gradient \( g_{hk} \) of MGM (Gradient method C) by multiplying it by \( h_k \), but keep the gradient \( \bar{g}_{hk} \) unchanged (as it is in MGM). As previous examples, we also assume the observation data have some errors, we take the noised data of the following form:

\[
z^\delta(x, y) = z_0(x, y) + \delta \sin(3\pi(x + y)),
\]

\[
z^\delta(x, t) = z(x, t) + \delta \sin(3\pi(x + y)).
\]

However, we find that the previous algorithm, just use one parameter \( \lambda \), to reconstruct two parameters \( p(x, y) \) and \( \mu(x, y) \) cannot obtain good results for this example. We introduce a modified algorithm, which uses two parameters, one for calculating the step-size for \( \bar{g}_{hk} \) and the other for calculating the step-size of \( g_{hk} \), which operates alternatively as follows:

(a) Set \( \lambda_1 = \lambda_2 = 0 \).

(b) Fix \( \lambda_2 \), find \( \lambda_1 > 0 \) such that \( \min_{\lambda_1>0} J(p_h^{(0)} - \lambda_1 \bar{g}_h, \mu_h^{(0)} - \lambda_2 g_h) \).

(c) Fix \( \lambda_1 \), find \( \lambda_2 > 0 \) such that \( \min_{\lambda_2>0} J(p_h^{(0)} - \lambda_1 \bar{g}_h, \mu_h^{(0)} - \lambda_2 g_h) \). Goto (b).
After several alternative steps, we can use $\lambda_1$ as the step-size to update $p_h(x, y)$ and use $\lambda_2$ as the step-size to update $\mu_h(x, y)$. In this example, we calculate $\lambda_1$ and $\lambda_2$ alternatively twice to obtain the step-size for $\bar{g}_h$ and $\bar{g}_h$.

Now figure 6.19 shows the exact observation $z_\theta = u(x, y, \theta)$ given at $\theta = 1/2$. Figure 6.20 shows the exact radiative coefficient $p(x, y)$ and figure 6.21 shows the numerically reconstructed $p_h(x, y)$ by nonlinear MGM. Figure 6.22 shows the exact initial data $\mu(x, y)$ and figure 6.23 shows the numerically reconstructed $\mu_h(x, y)$ by nonlinear MGM. The noise level $\delta = 1\%$, the initial guess $p_h^{(0)}(x, y)$ and $\mu_h^{(0)}(x, y)$ are both taken to be constant 2 everywhere.

Figure 6.19: The exact observation $z_\theta = u(x, y, 1/2)$. 

77
Figure 6.20: The exact radiative coefficient needed to be reconstructed $p(x, y)$.

Figure 6.21: $p_h^{(0)} = 4.0$, $\beta = 1.0E-7$, multi-grid-iter=2, error=1.161E-3, no. of cost functional calculated=711.
Figure 6.22: The exact initial data needed to be reconstructed $\mu(x, y)$.

Figure 6.23: $\mu_h^{(0)} = 2.0$, $\gamma = 1.0 \times 10^{-4}$, multi-grid-iter = 2, error = $2.48 \times 10^{-2}$. 
Bibliography


