Finite element methods for parameter identification problem of linear and nonlinear steady-state diffusion equations

Edgardo Ramirez

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Tao Lin, Chair
John Burns
Robert Rogers
David Russell
Shu-Ming Sun

November 19, 1997
Blacksburg, Virginia

Keywords: Parameter identification, Elliptic PDEs, Finite Element Methods,
Optimization Algorithms

Finite element methods for parameter identification problem of linear and nonlinear steady-state diffusion equations

by
Edgardo Ramirez
Committee Chair: Tao Lin
Department of Mathematics

(ABSTRACT)

We study a parameter identification problem for the steady state diffusion equations. In this thesis, we transform this identification problem into a minimization problem by considering an appropriate cost functional and propose a finite element method for the identification of the parameter for the linear and nonlinear partial differential equation. The cost functional involves the classical output least square term, a term approximating the derivative of the piezometric head $u(x)$, an equation error term plus some regularization terms, which happen to be a norm or a semi-norm of the variables in the cost functional in an appropriate Sobolev space. The existence and uniqueness of the minimizer for the cost functional is proved. Error estimates in a weighted H^{-1}-norm, L^2-norm and L^1-norm for the numerical solution are derived. Numerical examples will be given to show features of this numerical method.

Acknowledgments

I would like to take this opportunity to express my appreciation to a number of loving people who gave me support throughout my Ph.D. program.

To begin with, I am endlessly grateful to my advisor Dr. Lin who taught me, among other things, how to code finite element routines in one and two dimension and provided me with an exciting and wonderful dissertation topic. In addition, my special thanks goes to the Interdisciplinary Center for Applied Mathematics (ICAM) and all its wonderful people, Ms. Melissa Chase, Dr. Herdman, Dr. Burns and Dr. Cliff who allowed me to work at ICAM and use all the facilities for over three years. Finally I want to thank Dr. D. Russell for letting me stay at MIPAC for one and a half years. I want to thank especially each member of my committee for reading the manuscript and providing me with constructive suggestions for its improvement.

I am grateful to my family for their emotional support and especially to my dear mother, Elida Gomez, who supported me financially throughout my carrer in the Mathematical Sciences. Without her help, I would have gotten nowhere.

Finally, I want to thank to the Mathematics Department at Virginia Tech for giving me the great opportunity to pursue my Ph.D. degree overseas. My stay in the States has been a fruitful experience in my private and professional life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Statement of the parameter identification problem

We are concerned with the identification of the coefficient function for the linear elliptic partial differential equation defined below on the basis of an observed source term and state function,

$$-\nabla \cdot (a(x)\nabla u(x)) = f(x), \text{ for } x \in \Omega \subset \mathbb{R}^n, \tag{1.1.1}$$

with boundary conditions given by,

$$u\mid_\Gamma = 0, \text{or } \frac{\partial u}{\partial \eta}\mid_\Gamma = 0, \text{ where } \Gamma = \partial\Omega, \tag{1.1.2}$$

or

$$u\mid_{\Gamma_0} = \frac{\partial u}{\partial \eta}\mid_{\Gamma_1} = 0, \tag{1.1.3}$$

where Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary $\partial\Omega = \bar{\Gamma}_0 \bigcup \bar{\Gamma}_1$, Γ_0 and Γ_1 are open disjoint subsets of $\partial\Omega$ and η is the unit outer normal to $\partial\Omega$. We also consider the parameter identification problem for the nonlinear elliptic partial differential equation which is defined by,

$$-\nabla \cdot (a(u(x))\nabla u(x)) = f(x), \text{ for } x \in \Omega \subset \mathbb{R}^n \tag{1.1.4}$$

with similar boundary conditions. The partial differential equation (1.1.1) corresponds to the steady state case of the parabolic partial differential equation below,

$$b\frac{\partial u}{\partial t} = \nabla \cdot (a\nabla u) + f, \ (x, t) \in \Omega \times [0, \tau] \tag{1.1.5}$$

supplemented with appropriate initial and boundary conditions. The partial differential equation (1.1.5) happens to be a basic model equation in many applications such as, oil reservoir and groundwater flow simulation. The quantity u represents pressure or "piezometric head", f is a source term, and a and b are coefficients which are often referred to as the "transmissivity" and "storage coefficient", respectively. These coefficients are commonly taken as functions of the space coordinates. A detailed discussion of the various terms and a derivation of equation (1.1.5) can be found in [8].

We say that a problem is well-posed (in the sense of Hadamard) if

1. there exists a solution,
2. the solution is unique,
3. the solution depends continuously on the data.

If any one of these conditions does not hold, a problem is said to be ill-posed. Of course, the meaning of the term *continuity with respect to the data* has to be made more precise by a choice of norms for each problem at hand. To see that in general the parameter identification problem is not well-posed, we rewrite equation (1.1.1) as a first order hyperbolic partial differential equation in the unknown $a(x)$, which reduces to

$$\nabla u \cdot \nabla a + a\Delta u = -f. \tag{1.1.6}$$

As we can see, the problem of identifying $a(x)$ from equation (1.1.6) becomes ill-posed when $\nabla u = 0$ on some open set, since equation (1.1.6) provides no information about $a(x)$ on this set and consequently we have no uniqueness of the transmissivity (or diffusion) coefficient $a(x)$.

Now if $a(x)$ is prescribed along that portion of the boundary where $\frac{\partial u}{\partial \eta} < 0$, the so-called inflow boundary, then in the article [35], it was proved that equation (1.1.1) can be solved for $a(x)$ uniquely under the condition

$$\inf_{\Omega} \max\{|\nabla u|, \Delta u\} > 0 \tag{1.1.7}$$

We adopt the view that a good approximation to the transmissivity coefficient $a = a(x)$ is the one which yields a good solution to the elliptic forward problem. More precisely, let u and v be two solutions of the following boundary value problem corresponding to differing coefficients α and β:

$$\begin{aligned} &-\nabla \cdot (\alpha \nabla u) = -\nabla \cdot (\beta \nabla v) \text{ in } \Omega \\ &u = v \text{ on } \Gamma \text{ or } \alpha\frac{\partial u}{\partial \eta} = \beta\frac{\partial v}{\partial \eta} \text{ on } \Gamma, \end{aligned} \tag{1.1.8}$$

where η is the unit outer normal to Γ. Then, if we assume $\beta > 0$, an integration by parts against $u - v$ gives us

$$\|\nabla(u-v)\|_0 \leq \frac{\|\nabla u\|_\infty}{\inf_\Omega \beta}\|(\alpha - \beta)\|_0. \tag{1.1.9}$$

In view of (1.1.9), we shall think of a "good" reconstructed transmissivity as one that is close to $a(x)$ in the L^2 norm.

If the goal were an accurate reconstruction of $a(x)$ in the L^∞ norm, then the problem would be ill-posed in the sense of Hadamard. This is due to the smoothing property of elliptic operators, on account of which rapid oscillations in the transmissivity coefficient are suppressed in the forward solution. The following one-dimensional example, which was taken from [26], demonstrates this phenomenon: let $a(x) = \frac{1}{2}$, $u(x) = x^2$, and consider $a_N(x) = (2 + \cos Nx)^{-1}$, with the corresponding noisy observation $z_N(x) = x^2 + (\frac{x}{N}) \sin Nx + (\frac{1}{N^2}) \cos Nx (= u + \delta_N)$. It is easy to see that for every N

$$(au')' = (a_N z_N')' \text{ on } (0, \pi)$$

$$(au')_{(0)} = (a_N z_N')_{(0)}, (au')_{(\pi)} = (a_N z_N')_{(\pi)}$$

$$\text{and } \|z_N - u\|_{L^\infty} \leq \frac{c}{N} \to 0, \text{ while } \|a_N - a\|_{L^\infty} = \frac{1}{2} \text{ for every N.}$$

In summary, $z_N \to u$, yet $a_N \not\to a$ and a is therefore not a continuous function of the data with respect to the L^∞ norm. In a more general way, the ill-posedness of the identification problem can be shown by arguments from the homogenization theory [9].

On the other hand, one can show that the identification problem is well-posed with respect to the H^{-1}-norm. We have taken the following example in two-dimension from [26]. Indeed, we can show (with some reasonable hypotheses) that if v solves

$$\begin{aligned} -\nabla \cdot (a \nabla u) &= -\nabla \cdot (b \nabla v) \text{ in } \Omega, \\ a\frac{\partial u}{\partial \eta} &= b\frac{\partial v}{\partial \eta} \text{ on } \partial\Omega, \end{aligned} \tag{1.1.10}$$

then

$$\|a - b\|_{H^{-1}} \leq \|u - v\|_1, \tag{1.1.11}$$

where H^{-1} is the dual of $H^1(\Omega)$. Assume that $u \in C^2(\bar{\Omega})$ and $\|\nabla u\| \neq 0$ in $\bar{\Omega}$. Lemma 5 in [26], shows that for each $\phi \in H^1(\Omega)$, there is a ψ_ϕ with

$$\begin{aligned} \nabla u \cdot \nabla \psi_\phi &= \phi, \\ \|\psi_\phi\|_1 &\leq C\|\phi\|_1, \end{aligned} \tag{1.1.12}$$

where C is independent of ϕ. Using (1.1.10) and integrating by parts against ψ_ϕ gives

$$\int_\Omega (a - b) \nabla u \cdot \nabla \psi_\phi = \int_\Omega b \nabla (v - u) \cdot \nabla \psi_\phi.$$

Using (1.1.12), this gives

$$\begin{array}{rcl} |\int_\Omega (a-b)\phi\,| & = & |\int_\Omega b\nabla(u-v)\cdot\nabla\psi_\phi\,| \\ & \leq & \|b\|_\infty \|\nabla(u-v)\|_0 \|\nabla\psi_\phi\|_0 \\ & \leq & C\|u-v\|_1\|\phi\|_1 \end{array} \tag{1.1.13}$$

since ϕ was an arbitrary H^1 function, this gives (1.1.11). Unfortunately, this way to get well-posedness is not practicable since a good approximation in H^{-1} of the diffusion coefficient is physically useless. For our identification problem we are interested in special identification methods that are well-posed with respect to the L^2 norm.

Since the list of all the contributions that have recently been made to parameter identification problems and which are predominantly mathematical or numerical in nature is huge, we prefer to explain with some detail those that are closely related to our method. There are different methods which have been proposed in the literature for parameter identification problems such as least-squares methods [12], [20], Tikhonov regularization methods [6], [19], [27], [41], equation error methods [17], [26] or quasi-inversion methods [3], [28], a singular perturbation technique [3], the augmented Lagrangian technique and the large time asymptotics of an associated dynamical system [23]. We will give a more detailed description of these approaches in the next section.

The thesis is arranged as follows. In Chapter 2 we present the model boundary value problem and formulate the parameter identification (PI) problem as a minimization problem. The existence and uniqueness of the minimizer for the cost functional is proved. The finite element method to be used for the identification of the parameter is presented, along with error estimations of the approximations realized in a weighted H^{-1} and L^2 norm for a two point boundary value problem with Dirichlet and Dirichlet Neumann boundary conditions. We also present some numerical examples which corroborate our theoretical results, along with interesting features when the state has singularities. Chapter 3 is devoted to the PI problem for the nonlinear steady state diffusion equation. The algorithm to recover the nonlinear coefficient is formulated. The finite element method for the reconstruction of the nonlinear coefficient is presented. Error estimates in the weighted H^{-1} and L^2 norm are derived. Very interesting features for the identification of the nonlinear parameter are displayed. Chapter 4 is devoted to the PI problem in two dimension. The minimization scheme is stated. A weighted error estimate in the L^1 norm is obtained for the model equation with homogeneous Dirichlet boundary conditions. A weighted error estimate in the H^{-1} and L^2 norm are derived for the model equation with Dirichlet Neumann boundary conditions.

1.2 REVIEW OF METHODS

In this section we discuss some of the methods that have been used for solving parameter identification problems.

A. Finite difference scheme

The author in the article [36] proposes a finite difference approach for the identification of the coefficient in equation (1.1.1) under condition (1.1.7) as long as $a(x)$ is prescribed along the "inflow" portion of the boundary Γ of Ω (essentially that portion of Γ where the outer normal derivative of u is negative). For this scheme, equation (1.1.1) is viewed as a first order hyperbolic partial differential equation in the unknown transmissivity $a(x)$, which reduces to

$$\nabla u \cdot \nabla a + a\Delta = -f,\ x \in \Omega \subset I\!R^n, \tag{1.2.14}$$

This scheme has two characteristics, first it is explicit in the direction of increasing grid values of u (the characteristics of (1.2.14) are curves of steepest ascent in u), and secondly it is self-starting in the vicinity of a relative minimum of u (where (1.2.14) is degenerate and $a(x)$ is given by $-f/\Delta u$ if $\Delta u \neq 0$). First we describe the numerical method in [36] on the unit square $(0,1) \times (0,1)$.

We define a uniform grid as follows,

$$(x_i, y_j) = (ih, jh), 0 \leq i, j \leq n+1, h = \frac{1}{n+1}.$$

denoting by Ω_h the set of interior grid points,

$$\Omega_h = \{(x_i, y_j) | 1 \leq i, j \leq n\},$$

We also have to define a discrete inflow boundary Γ_1^h. A grid point in Γ is in Γ_1^h if its nearest neighboring grid point in Ω_h has a higher u value; e.g., $(x_i, y_0) \in \Gamma_1^h$ for $i \in \{1, \cdots, n\}$ if $u(x_i, y_1) > u(x_i, y_0)$. We shall denote grid values of $a(x,y)$, $u(x,y)$ and $f(x,y)$ by a_{ij}, u_{ij} and f_{ij}, respectively.

We approximate the differential equation (1.2.14) by $\mathcal{L}^h(a_{ij}; u_{ij}) = -f_{ij}$, $1 \leq i, j \leq n$, where

$$\mathcal{L}^h(a_{ij}; u_{ij}) \equiv \frac{a_{ij} - a_{kj}}{h} \cdot \frac{u_{ij} - u_{kj}}{h} + \frac{a_{ij} - a_{il}}{h} \cdot \frac{u_{ij} - u_{il}}{h} + a_{ij} H u_{ij}, \tag{1.2.15}$$

with

$$H u_{ij} = \frac{u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{ij}}{h^2}, \tag{1.2.16}$$

where

$$k \text{ is the first index of the minimum of } \{u_{i-1,j}, u_{ij}, u_{i+1,j},\}$$
$$l \text{ is the second index of the minimum of } \{u_{i,j-1}.u_{ij}, u_{i,j+1}.\}$$

Solving the equation $\mathcal{L}^h(a_{ij}; u_{ij}) = -f_{ij}$ for a_{ij}, we obtain

$$a_{ij} = \frac{a_{kj}(\frac{u_{ij}-u_{kj}}{h}) + a_{il}(\frac{u_{ij}-u_{il}}{h}) - hf_{ij}}{\frac{u_{ij}-u_{kj}}{h} + \frac{u_{ij}-u_{il}}{h} + hHu_{ij}}, \tag{1.2.17}$$

Now we want to make three points:

(i) This difference scheme is explicit in the direction of increasing u_{ij} since a_{ij} is given in terms of a_{kj} and a_{il}, where u_{kj} and u_{il} are $\leq u_{ij}$. Thus, the discrete solution is developed in a manner consistent with the characteristics of the continuous problem, which are curves of steepest ascent in u.

(ii) If u_{i^*,j^*} is a relative minimum grid value of u, the associated a_{ij} is given by

$$a_{i^*,j^*} = -\frac{f_{i^*,j^*}}{Hu_{i^*,j^*}}.$$

This corresponds to the case of degeneracy in the continuous problem, $\nabla a \cdot \nabla u + a\Delta u = -f$, where if $\nabla u(P) = 0$ and $\Delta u(P) \neq 0$, then

$$a(P) = -\frac{f(P)}{\Delta u(P)}$$

(iii) Initial data for $\{a_{ij}\}$ is required along Γ_1^h.

A condition equivalent to (1.1.7) is that the domain Ω can be divided into subregions Ω_1 and Ω_2 in which $|\nabla u|$ and Δu, respectively, are uniformly positive:

$$\begin{aligned} &\Omega = \Omega_1 \bigcup \Omega_2, \\ &|\nabla u| \geq k_1 > 0 \text{ in } \Omega_1, \\ &\Delta u \geq k_2 > 0 \text{ in } \Omega_2, \end{aligned} \tag{1.2.18}$$

To establish a uniqueness and convergence result, for the difference scheme

$$\mathcal{L}^h(a_{ij}; u_{ij}) = -f_{ij},\ 1 \leq i,j \leq n, a_{ij} \text{ given on } \Gamma_1^h, \tag{1.2.19}$$

for the unit square under the assumption that (1.2.18) holds, we have to postulate a discrete analog of (1.2.18). We assume that the interior grid points Ω^h can be divided into two sets Ω_1^h and Ω_2^h such that

$$\begin{aligned} &\Omega^h = \Omega_1^h \bigcup \Omega_2^h, \\ &\max\{\frac{u_{ij}-u_{kj}}{h}, \frac{u_{ij}-u_{il}}{h}\} \geq k_1^h > 0, \ (x_i, y_j) \in \Omega_1^h, \\ &Hu_{ij} \geq k_2^h > 0, \ (x_i, y_j) \in \Omega_2^h. \end{aligned} \tag{1.2.20}$$

It can be shown that if condition (1.2.20) holds and $hq_1^h < 1$, then the discrete problem (1.2.19) has a unique solution $\{a_{ij}\}$ satisfying,

$$\|a_{ij}\|_\infty \leq C_1^h(u)\left[\max\{\max_{\Gamma_1^h}|a_{ij}|, \frac{\|f\|_\infty}{k_2^h}\} + \frac{[u]\|f\|_\infty}{(k_1^h)^2}\right], \tag{1.2.21}$$

where

$$\rho = 1/(1-hq_1^h), \quad C_1^h(u) \equiv \max\{1, \rho\}^{[u]/hk_1^h}, \quad [u] \equiv \sup_\Omega u - \inf_\Omega u.$$

Now we are in condition to formulate the convergence result. It can be shown that if $u \in C^2(\overline{\Omega})$ satisfies the condition (1.2.18), then for h sufficiently small, the discrete problem (1.2.19) has a unique solution $\{a_{ij}\}$ assuming prescribed values on Γ_1^h. It satisfies the bound

$$\|a_{ij}\|_\infty \leq (C_1(u))^2\left[\max\{\max_{\Gamma_1^h}|a_{ij}|, \frac{\|f\|_\infty}{k_2}\} + \frac{2[u]}{k_1}\|f\|_\infty\right](1+o(h)). \tag{1.2.22}$$

where

$$C_1(u) \equiv \max\{1, \exp(\tfrac{q_1[u]}{k_1})\}.$$

Moreover, if $u \in C^3(\overline{\Omega})$ and the solution $a(x,y)$ of the continuous problem (1.2.14) is in $C^2(\overline{\Omega})$, then

$$\max_{0\leq i,j\leq n+1}|a_{ij} - a(x_i,y_j)| = O(h) \text{ as } h \to 0$$

assuming $a_{ij} = a(x_i, y_j)$ on Γ_1^h. The author in [36], was able to extend the applicability of this difference scheme to irregular domains and to problems in which the basic condition (1.1.7) does not hold but ∇u and Δu do not simultaneously vanish anywhere in Ω. A direct numerical method to solve a parameter identification problem within the finite difference approach can be found in [30].

B. Output Least-Squares Minimization

To the best of our knowledge, it was first applied to the boundary value problem,

$$\begin{gathered} -\nabla\cdot(a(x)\nabla u(x)) = f(x) \text{ for } x \in \Omega \subset I\!R^n, \\ a(x)\frac{\partial u}{\partial \eta} = g \text{ on } \partial\Omega, \end{gathered} \tag{1.2.23}$$

in the article [21]. The least-squares approach says that if $u(b)$ is the solution of (1.2.23), with the coefficient a replaced with b, then b is a good approximation of a if the difference of a measurement z of u and $u(b)$, is small in $L^2(\Omega)$. For practical purposes, we need to define finite dimensional spaces to implement this approach. Let $\{\Delta_h\}$, $0 < h < 1$, be a triangulation of $\bar{\Omega}$ with triangles T of diameter less than or equal to h. Given an L^2 measurement z of u, select finite dimensional subspaces A_h and V_h. To each coefficient $a_h \in A_h$, we associate a $u(a_h) \in V_h$, where $u(a_h)$ solves (1.2.23) in a Galerkin approximation:

$$\int_\Omega a_h \nabla u_h(a_h) \cdot \nabla v_h dx = \int_\Omega f v_h dx + \int_\Gamma g v_h ds \tag{1.2.24}$$

for all $v_h \in V_h$, and

$$\int_\Omega u_h(a_h) dx = \int_\Omega z dx \tag{1.2.25}$$

To state this numerical approach, we need to define the following cost functional:

$$\mathcal{J}(b_h) = \|u_h(b_h) - z\|_0^2,$$

where $b_h \in A_h$. The least-squares approach to the approximate determination of a is to solve the following problem,

(P_h) Find $a_h \in A_h$ such that

$$\mathcal{J}(a_h) = \inf_{b_h \in K_h} \mathcal{J}(b_h)$$

where

$$K_h = \{b \in A_h | 0 < c_0 \le b \le c_1\}$$

is the set for admissible parameters, c_0, $c_1 \in \mathbb{R}^+$ are given constants.

We would like to present an interesting error estimate for the approximation scheme (P_h) found in [20]. To this end, we need to formulate some hypothesis,

(A1) There exists a constant unit vector $\vec{\nu}$ and a constant $\sigma > 0$ such that $\nabla u \cdot \vec{\nu} \ge \sigma > 0$ $\forall x \in \Omega$.

(A2) $u \in W^{r+3,\infty}(\Omega)$ and $\Gamma_1 = \{x \in \Gamma : \dfrac{\partial u}{\partial \eta} > 0\} \subset C^{r+2}$, where $r \ge 1$.

(A3) Suppose that $a \in H^{r+1}(\Omega)$, $u \in H^{r+2}(\Omega)$, with $A_h = S_h^r$ and $V_h = S_h^{r+1}$, where

$$S_h^r = \{v \in C^0(\bar{\Omega}) : v\,|_T \in P_r \forall T \in \Delta_h\},$$

where P_r is the space of polynomials of degree less than or equal to r in the variables x_1 and x_2.

(A4) The observation error is of the form

$$\|u - z\|_0 \le \epsilon.$$

Then, for all h sufficiently small, we have

$$\|a - a_h\|_0 \le C[h^r + h^{-2}\epsilon],$$

where a_h is any solution of problem (P_h) and C is a constant independent of h and ϵ. Another Numerical method related to the output least squares method to solve an inverse problem for a time dependent initial boundary value problem can be found in [32].

C. Equation error method

For this method we also replace u by z in (1.1.1). With z and f given, we consider the affine mapping $\psi(b) = \nabla \cdot (b\nabla z) + f$ and solve $\psi(b) = 0$ for the "true" coefficient $a = a(x)$ by means of

$$\min_{b \in \mathcal{Q}_{ad}} \|\nabla \cdot (b\nabla z) + f\|_H^2, \tag{1.2.26}$$

where $\mathcal{Q}_{ad}$ is the set of admissible parameters and H is a Hilbert space chosen appropriately. Differently from the least squares method, (1.2.26) is convex in a and existence of the unique minimizer follows. The major drawback of the equation error method seems to be the necessity of differentiating the data twice. Boundary conditions on z or b can easily be incorporated into (1.2.26).

Under an identifiability assumption the equation error method is realized with $H = L^2(\Omega)$ in [39]. A multigrid algorithm is devised to solve the linear matrix equation which arises from discretization of (1.2.26) and application of a necessary optimality condition.

An alternative approach can be based on the weak formulation of (1.2.26). In the case of homogeneous Dirichlet boundary conditions $z\mid_{\partial\Omega}= 0$ it is given by

$$\min_b \|\nabla \cdot (b\nabla z) + f\|_{H^{-1}}^2, \tag{1.2.27}$$

Observe that (1.2.27) is equivalent to

$$\min_b \|\Delta^{-1}(\nabla \cdot (b\nabla z) + f)\|_{H_0^1}^2,$$

where Δ denotes the Laplacian from H_0^1 to H^{-1} with homogeneous Dirichlet boundary conditions. Since

$$\|\Delta^{-1}(\nabla \cdot (b\nabla z) + f)\|_{H_0^1}^2 = \sup_{\phi \in H_0^1} [\langle -b\nabla z, \nabla\phi\rangle + \langle f, \phi\rangle],$$

it is evident that the data are only differentiated once in the weak formulation of the equation error method (1.2.27) as opposed to two differentiations which are required in (1.2.26) with $H = L^2(\Omega)$. The analogue of the weak formulation with the Dirichlet boundary condition replaced by the assumption of the availability of flux boundary data $b\frac{\partial z}{\partial \eta} = g$ on $\partial\Omega$ and its numerical treatment for smooth as well as for discontinuous coefficients b is given in [1]. Here η denotes the unit outward normal derivative to $\partial\Omega$.

D. Variational approach

Another numerical scheme for the reconstruction of the coefficient $a(x)$ in (1.2.23) is a variational method. This approach was first developed in the article [26] and is motivated by the simple observation that for any positive weights γ_1 and γ_2,

$$\|\sigma - a \cdot \nabla u\|_{L^2}^2 + \gamma_1 \|\mathrm{div}\sigma + f\|_{L^2}^2 + \gamma_2 \|\sigma \cdot \eta - g\|_{L^2(\Gamma)}^2 \geq 0, \tag{1.2.28}$$

for any choice of $a(x)$ and any vector field σ, the minimum being achieved only when $\sigma = a\nabla u$ with $a(x)$ a solution of (1.2.23). This variational method for reconstructing the unknown coefficient involves minimizing (1.2.28) numerically over suitable finite-dimensional spaces of coefficients and vector fields, replacing u, f, and g in (1.2.28) by their measurements interpolated on a triangulation of Ω, with the well sites as the vertices of the triangles.

Let $\{\Delta_h\}$, $0 < h < 1$, be a family of regular, quasi-uniform triangulations of Ω, a domain with a Lipschitz continuous boundary. Given an L^2 measurement u^m, f^m, g^m of u, f and g, select finite dimensional subspaces A_h, K_h for the coefficient and the vector field variables. The variational method to identify the unknown coefficient a in (1.2.23), involves minimizing the cost functional

$$\mathcal{J}(a, \sigma) = \|\sigma - a \cdot \nabla u^m\|_{L^2}^2 + \gamma_1 \|\mathrm{div}\sigma + f^m\|_{L^2}^2 + \gamma_2 \|\sigma \cdot \eta - g^m\|_{L^2(\Gamma)}^2 \geq 0, \tag{1.2.29}$$

over the finite-dimensional spaces of coefficients A_h and vector fields K_h by using the measured data u^m, f^m, and g^m. The advantage of this method is that we are dealing with a quadratic minimization problem which is extremely easy to implement. The disadvantage of this method is the large number of variables it uses, for example if σ and a are piecewise linear on a triangulation with N^2 nodes, then the functional to be minimized depends on $3N^2$ variables. Another disadvantage is that we have to differentiate the data u^m. The weights γ_1 and γ_2 are chosen so that each term of the sum in (1.2.29) has the same magnitude after the finite dimensional spaces have been fixed.

We can also implement other variational methods by defining variations of the cost functional (1.2.29). For instance, one might consider using $\sigma \cdot \eta = g^m$ and $\mathrm{div}\sigma = -f^m$ as constraints and minimizing $\|\sigma - a\nabla u^m\|_{L^2}^2$ or perhaps $\|\sigma - a\nabla u^m\|_{L^2}^2 + \epsilon\|a\|_{H^1}^2$ for some small positive ϵ. Another interesting possibility includes minimizing

$$J(a,\sigma) = \frac{1}{2}\int_{\Omega} |a^{-\frac{1}{2}}\sigma - a^{\frac{1}{2}}\nabla u|^2 dx \tag{1.2.30}$$

with respect to $a(x) \geq 0$ (by hand), then with respect to σ satisfying $\mathrm{div}\sigma = -f^m$ and $\sigma \cdot \eta = g^m$ (numerically). Minimization over $a(x)$ leads to the choice

$$a(x) = \frac{|\sigma|}{|\nabla u|},$$

which on substitution into the cost functional (1.2.30) gives the convex functional

$$E(\sigma) = \int_{\Omega} (|\sigma||\nabla u| - \langle \sigma, \nabla u \rangle) dx$$

In order to state the main results in [26], we need to make some assumptions.

(A1) Let a and u satisfy equation (1.1.1) with boundary condition

$$a\frac{\partial u}{\partial \eta} = g, \text{ on } \partial\Omega, \tag{1.2.31}$$

and let them have the following regularities: $a \in H^2(\Omega)$, $u \in H^3(\Omega)$, $\Delta u \in C^0(\bar{\Omega})$.

(A2) Set $Q_h^{(k)} = \{w \in C^0(\bar{\Omega}) : w\ |_T \in P_k, \forall T \in \Delta_h\}$, where P_k is the set of polynomials of degree less than or equal to k,

$$A_h = \{w \in Q_h^{(1)} : 0 < \alpha \leq w \leq \beta\},$$
$$K_h = Q_h^{(1)} \times Q_h^{(1)}.$$

where α and β are positive constants.

(A3) Let u^m, f^m, and g^m be measurements of u, f and g corresponding to equation (1.1.1) with boundary condition (1.2.31), where $\|u - u^m\|_1 < \epsilon$, $\|f - f^m\|_0 < \lambda_1$, and $\|g - g^m\|_{L^2(\partial\Omega)} < \lambda_2$.

(A4) Assume that $u^m \in Q_h^{(k)}$ for some fixed k.

(A5) Let $\sigma_* \in K_h$, $a_* \in A_h$ be such that,

$$\mathcal{J}(\sigma_*, a_*) = \min_{\sigma \in K_h, a \in A_h} \mathcal{J}(\sigma, a),$$

where $\mathcal{J}$ is a cost functional appropriately defined.

(A6) Assume that: $u \in C^2(\bar{\Omega})$ and $|\nabla u| \neq 0$ on $\bar{\Omega}$. Then for each $\psi \in H^1(\Omega)$, there exists a $v_\psi \in H^1(\Omega)$ satisfying

$$\nabla u \cdot \nabla v_\psi = \psi \text{ in } \Omega$$

and the estimate

$$\|v_\psi\|_1 \leq C\|\psi\|_1,$$

where the constant C is independent of ψ.

If (A3-5) and $\inf_\Omega \max\{|\nabla u|, \Delta u\} > 0$ hold, then

$$\|a_* - a\|_0 \leq C\{h + \epsilon h^{-1} + \lambda_1 + h^{-1/2}\lambda_2\}.$$

If (A3) and (A5) hold, then

$$\int_\Omega |a_* - a||\nabla u|^2 \leq C\{h + \epsilon h^{-1} + \lambda_1 + h^{-1/2}\lambda_2\}.$$

If (A3), (A5) and (A6) hold, then

$$\|a_* - a\|_0 \leq C(h + \epsilon + \lambda_1 + h^{-1/2}\lambda_2)^{1/2}.$$

All the constants C in these error estimates are independent of h, ϵ, λ_1 and λ_2.

F. Modified equation error and least-squares method

This approach was introduced in [24]. The method in this article considers the identification of the coefficient function in the homogeneous elliptic boundary value problem

$$\begin{aligned} -\nabla \cdot (b(x)\nabla u(x)) &= f(x) \text{ in } \Omega, \\ u|_{\Gamma_0} = \frac{\partial u}{\partial \eta}|_{\Gamma_1} &= 0, \end{aligned} \tag{1.2.32}$$

where Ω is a bounded domain in $I\!R^n$, with smooth boundary $\partial\Omega = \bar{\Gamma_0} \bigcup \bar{\Gamma_1}$, Γ_0 and Γ_1 are open disjoint subsets of $\partial\Omega$. The main idea of this method is to include an extra term to the least-squares cost functional which takes into account the underlying equation (1.2.32), times a weight chosen according to the finite dimensional spaces, to balance the different amount of differentiation on both terms. This approach combines the output least least squares method with the equation error method to transform the identification problem into a minimization problem.

Let $\{\mathcal{T}_h\}$ be a family of triangulation of Ω of mesh $0 < h < 1$. If the boundary of Ω is curved, we use triangles with one edge replaced by the curved segment of the boundary. It is assumed that the family $\{\mathcal{T}_h\}$ is regular and quasi-uniform. Given an L^2 measurement z of u, select finite dimensional subspaces B_h, and U_h for the coefficient and the state, respectively. To each coefficient $b_h \in B_h$, we associate a $u_h(b_h) \in U_h$, where $u_h(b_h)$ solves (1.2.32) in a Galerkin approximation:

$$\int_\Omega b_h \nabla u_h(b_h) \cdot \nabla v_h dx = \int_\Omega f v_h dx \tag{1.2.33}$$

for all $v_h \in U_h \subset \tilde{H}^1 = \{v \in H^1 | v\ |_{\Gamma_0} = 0\}$. The cost functional is defined through the following formula:

$$\mathcal{J}(b_h) = \|u_h(b_h) - z\|_0^2 + \gamma \|\nabla \cdot (b_h \nabla u_h(b_h)) + f\|_0^2 \tag{1.2.34}$$

Here $u_h(b_h)$ is the solution of equation (1.2.33), which corresponds to a given parameter $b_h \in B_h$. The advantage of this method over others found in the literature is that, it does not substitute the observation z by u directly into the operator $\nabla \cdot (b_h \nabla u)$, which is a cause of a huge error in the numerical implementation. To find an approximation of b , we select that $b_h^* \in M_h$ which minimizes (1.2.34) over M_h, where

$$M_h = \{b \in B_h | 0 < \lambda_1 \leq b \leq \lambda_2, \text{ in } \Omega\}$$

is the set for admissible parameters.

The error estimates in [25] were derived under the following assumptions,

(A1) Let $z(x)$ be a distributed L^2 observation of the state u with an observation error

$$\|u - z\|_0 \leq \epsilon.$$

(A2) The functions in (1.2.32) have the following regularity,

$$u(x) \in \tilde{H}^1 \bigcap H^{r+2} \bigcap W^{2,\infty}, b(x) \in H^{r+1} \bigcap W^{1,\infty} \text{ and } f \in H^r,$$

where $\tilde{H}^1 = \{v \in H^1 : v\ |_{\Gamma_0} = 0\}$ and $r \geq \frac{n}{2}$.

(A3) We take

$$U_h = S_{h,2}^{r+1,0}, \\ B_h = S_{h,1}^{r},$$

where

$$S_{h,l}^r = \{v \in C^{l-1}(\Omega) : v \mid_T \in P_r \; \forall T \in \mathcal{T}_h\}.$$

Here, P_r is the space of polynomials of degree less than or equal to r. We denote by $S_{h,l}^{r,0}$, the subspace of $S_{h,l}^r$ of functions which vanish on $\Gamma_0 \subset \partial\Omega$.

For h small enough, the minimizer b_h^* and the original parameter b satisfy

$$\int_\Omega |b_h^* - b||\nabla u|^2 \, dx \leq C(h^r + h^{-2}\epsilon).$$

We can get better estimates for the case $n = 1$ if we change the cost functional (1.2.34) to

$$\mathcal{J}(b_h) = \|u_h(b_h) - z\|_0^2 + \tilde{\gamma}\|(b_h u_h'(b_h))' + f\|_{-1}^2 \tag{1.2.35}$$

Moreover, we have to take at least one of the boundary conditions to be Neumann. In (1.2.35), ' denotes differentiation with respect to x and the norm in the second term is realized in the dual space $\tilde{H}^{-1}$ of the test function space $\tilde{H}^1$.

Under these new assumptions, for h small enough, we have the error estimate

$$\|(b_h^* - b)u'\|_0 \leq C(h^{r+1} + h^{-1}\epsilon),$$

which holds if b_h^* is the minimizer of (1.2.35).

G. Output-Least-Squares Method to identify a nonlinear parameter

The authors in [40] proposed an interesting approach to identify a nonlinear parameter through the identification of a linear one. The method consists in identifying the nonlinear parameter in the nonlinear elliptic system:

$$\begin{cases} -\nabla \cdot (a(u)\nabla u) &= f \text{ in } \Omega \\ a(u)\dfrac{\partial u}{\partial \eta} &= g \text{ on } \partial\Omega \end{cases} \tag{1.2.36}$$

where Ω is an open bounded domain in $I\!R^2$ with a Lipschitz boundary, the state u and the parameter $a(u)$ are scalar functions. In this approach, the nonlinear parameter is regarded as a linear spatially varying parameter defined by:

$$b(x_1, x_2) = a(u(x_1, x_2)), \tag{1.2.37}$$

and the scheme uses the finite element method to identify it. Under this assumption, the nonlinear system (1.2.36) becomes a linear one which can be rewritten as follows:

$$\begin{cases} -\nabla \cdot (b\nabla u) &= f \text{ in } \Omega \\ b\dfrac{\partial u}{\partial \eta} &= g \text{ on } \partial\Omega. \end{cases} \tag{1.2.38}$$

In order to implement this method numerically, let $\{\Delta_h\}$ be a triangulation of Ω of mesh h. Given an L^2 measurement z of u, select finite dimensional spaces S_h and T_h of coefficients and state variable. Since the least squares method is used to identify the linear parameter, the approximation of the true linear coefficient is defined to be the minimizer of the following problem:

$$\begin{cases} \text{Find } b_h(\cdot,\cdot) \in M_h \text{ such that} \\ \mathcal{J}_h(b_h) = \min_{\tilde{b}_h \in M_h} \mathcal{J}_h(\tilde{b}_h), \end{cases} \tag{1.2.39}$$

where

$$\mathcal{J}(\tilde{b}_h) = \|u_h(\tilde{b}_h) - z\|_0^2, \tag{1.2.40}$$

$$M_h = \{b \in T_h | 0 < c_0 \leq b \leq c_1\}$$

(with c_0, c_1 given bounds for the transmissivity), and $u_h(\tilde{b}_h) \in S^h$ is the solution of

$$\int_\Omega \tilde{b}_h \nabla u_h(\tilde{b}_h) \nabla v_h \, dx = \int_\Omega f v_h \, dx + \int_{\partial\Omega} g v_h \, ds \text{ for all } v_h \in S^h \tag{1.2.41}$$

along with the compatibility condition

$$\int_\Omega u_h(\tilde{b}_h) \, dx = \int_\Omega z \, dx. \tag{1.2.42}$$

In order to prevent the linear parameter $\tilde{b}_h$ from being a multi-valued function when it is regarded as a function of u, the following constraint will be imposed

$$\frac{\partial \tilde{b}_h}{\partial x_1} \frac{\partial u}{\partial x_2} = \frac{\partial \tilde{b}_h}{\partial x_2} \frac{\partial u}{\partial x_1}. \tag{1.2.43}$$

This constraint will be taken as a penalty term into the cost functional (1.2.40). To this end, the cost functional that is minimized is given by,

$$\mathcal{J}_{h,\mu}(\tilde{b}_h) = \int_\Omega |u_h(\tilde{b}_h) - z|^2 \, dx + \mu \int_\Omega |u_{x_1} \tilde{b}_{hx_2} - u_{x_2} \tilde{b}_{hx_1}|^2 \, dx.$$

where μ is the penalty term. Finally, after the identification of the linear parameter from (1.2.39), the nonlinear parameter $a(s)$ can be recovered from (1.2.37), since $b(x)$ and $u(x)$ are known.

1.3 Some related methods

In this section we discuss some of the methods (other than numerical methods) that have been used for solving parameter identification problems.

A. Integrating along characteristics

In fact it is easily seen that (1.1.1) can be interpreted as a first order hyperbolic partial differential equation for the unknown coefficient a. In this case, the equation reduces to

$$\mathcal{L}(a,u) = \nabla a \cdot \nabla u + a\Delta u \quad = \quad -f, \text{ in } \Omega \subset R^2, \tag{1.3.44}$$

We found some interesting results in [35] for the identification of the diffusion coefficient in (1.3.44). It can be shown that under the assumption (1.1.7) or its equivalent form (1.2.18), the equation (1.3.44) has a unique solution a for any f, subject to specification of appropriate initial data, and that a depends continuously on f, ∇u, and Δu. In order to state the existence, uniqueness, and continuous dependence results for the diffusion coefficient a, we need to make some assumptions,

(A1) $u \in C^2(\bar{\Omega})$, $f \in L^\infty(\Omega)$.

(A2) By a "solution" to the hyperbolic equation (1.3.44), we mean a function $a \in L^\infty(\Omega)$ which is continuous and differentiable along the characteristic curves of (1.3.44), and which satisfies the ordinary differential equation to which (1.3.44) reduces along such curves. Now the main result in [35] says that

If condition (1.2.18) holds, then for any f, $\mathcal{L}(a;u) = -f$ has a unique solution $a = a(x)$ assuming prescribed values along the inflow boundary Γ_1, and

$$\|a\|_\infty \quad \leq \quad C_1(u)\left[\max\{\sup_{\Gamma_1}|a|, \frac{\|f\|_\infty}{k_2}\} + \frac{[u]}{k_1^2}\|f\|_\infty\right], \tag{1.3.45}$$

$$\tag{1.3.46}$$

where

$$[u] \quad = \quad \sup_{\Omega} u - \inf_{\Omega} u \tag{1.3.47}$$

$$q_1 \quad = \quad \sup_{\Omega_1}\{-\frac{\Delta u}{|\nabla u|}\} \tag{1.3.48}$$

$$C_1(u) \quad = \quad \max\{1, \exp(\frac{q_1[u]}{k_1})\} \tag{1.3.49}$$

Now, suppose that $\mathcal{L}(b; v) = -g$, where b is the diffusion coefficient produced by a perturbed solution $v \approx u$ and forcing function $g \approx f$. Applying the previous result to the identity

$$\mathcal{L}(b - a; u) = -\mathcal{L}(b; v - u) + (f - g),$$

we obtain the following continuous dependence

$$\|a - b\|_\infty \quad \leq \quad C_1(u)[\max\{\sup_{\Gamma_1} |a - b|, \frac{C_2}{k_2}\} + \frac{[u]}{k_1^2} C_2], \tag{1.3.50}$$

where

$$C_2 = [\|\nabla b\|_\infty \|\nabla(v - u)\|_\infty + \|b\|_\infty \|\Delta(v - u)\|_\infty + \|g - f\|_\infty].$$

B. Singular Perturbation

In the article [3], a singular perturbation technique is proposed to determine the spatially varying coefficient in the special case $f = 0$ in the partial differential equation (1.1.1), when the boundary condition is

$$u = g, \text{ on } \partial\Omega,$$

where Ω is a simply connected, C^2-smooth, bounded domain in $I\!R^2$ and g is a smooth function which is precisely known. Moreover it is assumed that a satisfies the ellipticity condition

$$0 < \lambda_1 \leq a(x) \leq \lambda_2, \; x \in \Omega,$$

along with the following regularity hypothesis

$$|a'(x)| \leq E, \; x \in \Omega,$$

where λ_1, λ_2, E are fixed positive numbers. It was also proved that, if g has a finite number, N, of relative maxima and minima on $\partial\Omega$, then the gradient of u vanishes only at a finite number of interior points, and only with a finite multiplicity. Moreover, the number of interior critical points and their multiplicities are controlled in terms of N. This algorithm consists of an approximation procedure. It can be shown that as $\epsilon \to 0$, the solution a_ϵ of the elliptic boundary value problem

$$\varepsilon \Delta a_\varepsilon + \nabla \cdot (a_\varepsilon \nabla u) \quad = \quad 0 \text{ in } \Omega \tag{1.3.51}$$

$$a_\varepsilon \quad = \quad a \text{ on } \Gamma \tag{1.3.52}$$

converges to a in $L^p_{loc}(\Omega)$ for every $p < \infty$. Thus an approximate identification is performed solving the problem (1.3.51), (1.3.52) with a suitably chosen value of ϵ. It is worth mentioning that under very smooth hypothesis for a, a_ϵ, u, g and Ω, the following estimate holds

$$\int_\Omega |a_\varepsilon - a||\nabla u|^2 dx \leq C\varepsilon^{\frac{1}{2}}$$

C. Long-time behavior of an associated dynamical system

The authors in [23] have proposed a new and ingenious technique to reconstruct coefficients in elliptic equations. An algorithm is developed to identify the unknown coefficients without a minimization technique. This method is based on the construction of certain time-dependent problems which contain the original equation as asymptotic steady state. The specific problem is formulated as follows;

(P) Given an open and bounded set $\Omega \subset I\!R^n$, $n \geq 1$, $u^* \in H^1_0(\Omega)$ and $f^* \in H^{-1}(\Omega)$, find a pointwise symmetric matrix function $A^* \in L^\infty(\Omega)$ such that

$$-\nabla \cdot (A^*(x)\nabla u^*) = f^*. \tag{1.3.53}$$

The main idea of this method is to regard (1.3.53) as an asymptotic (for $t \to \infty$) steady state of the following system of parabolic equations;

$$\begin{aligned} \dot{u} - \nabla \cdot (A(t)\nabla u(t)) &= f^*, \text{ a.e. on } (0,T), && (1.3.54)\\ u(0) &= u^0 \in H^1_0(\Omega)\ , \\ \dot{A}(t) &= (\nabla u(t) \otimes \nabla(u(t) - u^*)), \text{ a.e. on } (0,T), && (1.3.55)\\ A(0) &= A^0 \in L^\infty(\Omega), \text{ symmetric }. \end{aligned}$$

Here for u, $w \in I\!R^n$ the $(n \times n)$ -matrix $u \otimes v$ is defined by

$$(u \otimes v)_{i,j} = \frac{1}{2}(u_i w_j + u_j w_i), i, j = 1, \cdots, n \tag{1.3.56}$$

In this method, equation (1.3.54) is replaced by a regularizing equation which will yield an a-priori estimate that will enable us to get some convergence results.

For $\epsilon > 0$, we consider the following dynamical system

$$\begin{aligned} -\epsilon\frac{\partial}{\partial t}\Delta u - \nabla \cdot (A(t)\nabla u(t)) &= f^*, \text{ a.e. on } (0,T), && (1.3.57)\\ u(0) &= u^0 \in H^1_0(\Omega)\ , \\ \dot{A}(t) &= (\nabla u(t) \otimes \nabla(u(t) - u^*)), \text{ a.e. on } (0,T), && (1.3.58)\\ A(0) &= A^0 \in L^\infty(\Omega), \text{ symmetric }. \end{aligned}$$

Clearly (1.3.57), (1.3.58) have (1.3.53) as a steady state. After considering the weak formulation of (1.3.57), (1.3.58) and under the hypothesis that (P) has at least one positive definite solution $A^* \in L^\infty(\Omega)$, it can be proved that for each sequence $t_n \to \infty$, there exists a subsequence $\{t_{n_k}\}$ such that $A(t_{n_k}) \to A_\infty$ weakly in $L^2(\Omega)$ where A_∞ satisfies (1.3.53). The key tool in this result is the a-priori estimate:

$$\sup_{t\geq 0}\{\|\nabla u(t) - \nabla u^*\|_0^2 + \|A(t) - A^*\|_0^2\} + \textstyle\int_0^\infty \|\nabla u(t) - \nabla u^*\|_0^2 \, dt \leq C < \infty, \quad (1.3.59)$$

where $C = C(u^0, A^0, A^*)$. For practical purposes it is necessary to replace (1.3.57), (1.3.58) by a finite-dimensional scheme. To this end, a semidiscrete Galerkin approximation is proposed. Under more assumptions, it can be shown that the a-priori estimation (1.3.59) holds in the finite dimensional case. This estimate is used again to show that $A_\infty^n = \lim_{t\to\infty} A^n(t)$, with $A_\infty^n \in L^\infty(\Omega)$, and $A_\infty^n \to A_\infty$ weakly in $L^2(\Omega)$ where A_∞ satisfies (1.3.53), $A^n(t)$ is the n-dimensional Galerkin solution of (1.3.57), (1.3.58). It is worth noticing that this method gives a matrix coefficient, not a scalar one. Besides, in this context the solution of (1.3.53) is not unique. This method picks a particular solution, but it does not say which one.

Chapter 2

Parameter identification of a two point boundary value problem

2.1 Introduction

In this chapter we are concerned with the identification of the unknown coefficient $a(x)$ for the elliptic boundary value problem with homogeneous Dirichlet and Dirichlet-Newmann boundary conditions defined below. We present a least-squares mixed finite element method to identify the coefficient $a(x)$. We describe the problem here, and in later sections we specify more precisely our assumptions.

Let Ω be the open interval $(0,1)$ in $I\!R$. We consider a two point boundary value problem with a differential equation

$$\mathcal{L}(a)u = -(a(x)u'(x))' = f(x) \text{ in } \Omega, \tag{2.1.1}$$

with boundary condition

$$u(0) = u(1) = 0, \tag{2.1.2}$$

or

$$u(0) = u'(1) = 0, \tag{2.1.3}$$

where $f \in L^2(\Omega)$, and $\lambda \leq a(x) \leq \Lambda$ for some positive constants λ and Λ. The parameter identification (PI) problem is to recover a parameter $a = a(x)$ within some suitable set, from an L^2 observation $z(x)$ of $u(x)$ such that $a(x)$ and $u(x)$ satisfy (2.1.1) along with the boundary condition given by either (2.1.2) or (2.1.3).

Equation (2.1.1) corresponds to a steady state solution of the parabolic partial differential equation below,

$$\frac{\partial u}{\partial t} = (au')' + f. \tag{2.1.4}$$

This equation is important in hydrology, where it describes the flow of water through an aquifer. In equation (2.1.4), u is the pressure or piezometric head and a is the transmissivity, the function of proportionality in Darcy's law which measures the ability of the water to move in the aquifer. The quantity f represents a source term. A detailed discussion of equation (2.1.4) can be found in [8]. Because f represents a source term, an analysis of equation (2.1.1) when f is a delta distribution is very important and such a case was considered in [2]. In this chapter, we focus on the parameter identification problem for the linear elliptic partial differential equation (2.1.1).

To describe the method, we let $H^s(0,1)$ be the standard Sobolev space. We use $\|\cdot\|_s$ to denote its norm, and use $|\cdot|_s$ to denote its usual seminorm. The inner product of $H^s(0,1)$ will be denoted by $(\cdot,\cdot)_s$ with $(\cdot,\cdot) = (\cdot,\cdot)_0$ for the inner product in $L^2(0,1)$. We set

$$\tilde{H}^1(0,1) = \begin{cases} H_0^1(0,1), & \text{if (2.1.2) holds,} \\ \{v \in H^1(0,1) : v(0) = 0\}, & \text{if (2.1.3) holds.} \end{cases}$$

As a matter of fact, our PI problem has two unknowns: $a(x)$ and $u(x)$. Following the standard set up in the mixed finite element method, we introduce the flux $\sigma(x) = a(x)u'(x)$ as an additional unknown. These unknowns satisfy the equivalent first order system:

$$\begin{cases} -\sigma' &= f, \\ au' &= \sigma. \end{cases} \tag{2.1.5}$$

We then try to find an approximation to the solution of the PI problem by minimizing the following cost functional,

$$\mathcal{J}(a,u,\sigma) = \begin{cases} \mathcal{J}_1(a,u,\sigma), & \text{if (2.1.2) holds,} \\ \mathcal{J}_2(a,u,\sigma), & \text{if (2.1.3) holds,} \end{cases} \tag{2.1.6}$$

over the space

$$\mathcal{H} = H^1(0,1) \times \tilde{H}^1(0,1) \times H^1(0,1), \tag{2.1.7}$$

where

$$\begin{aligned} \mathcal{J}_1(a,u,\sigma) &= \|u-z\|_0^2 + \alpha\|\sigma - au'\|_0^2 + \beta\|\sigma' + f\|_0^2 + \\ &\quad \gamma\|a\|_{r_a}^2 + \rho\|u\|_{r_u}^2 + \delta\|\sigma\|_{r_\sigma}^2, \\ \mathcal{J}_2(a,u,\sigma) &= \mathcal{J}_1(a,u,\sigma) + \xi(\sigma(1))^2 + \theta(a(1)u'(1) - \sigma(1))^2, \end{aligned} \tag{2.1.8}$$

where the integers r_a, r_u, and r_σ depend on the Sobolev spaces in which the regularization takes place. Usually, they take the following values

$$r_a = r_u = r_\sigma = 1.$$

The nonnegative numbers $\alpha, \beta, \gamma, \rho, \delta, \xi, \theta$ can be considered as regularization parameters whose choices will be given later according to the error estimates. It is not necessary to use ρ, δ for both the computation and the error estimation. However, as we will see in the next section, choosing appropriate values for these terms may give the cost functional certain preferable features.

The first term in $\mathcal{J}(a, u, \sigma)$ represents the least-squares fit to data, the second and third terms are the residuals (equation error) of the involved differential equation in the mixed formulation, the rest of the terms are added for the regularization purpose. Note that if the data is perfect, i.e., there exist $u(x)$ and $a(x)$ satisfying (2.1.1) with boundary condition (2.1.2) or (2.1.3) such that $z(x) = u(x)$, then $(a, u, \sigma)^T$ with $\sigma = au'$ can make zero the first three terms (and the last two terms if the boundary condition (2.1.3) is required) of $\mathcal{J}(a, u, \sigma)$, and makes $\mathcal{J}(a, u, \sigma)$ small if the regularization parameters are small. Therefore, the solution to the PI problem seems to be in a neighborhood of a minimizer of $\mathcal{J}(a, u, \sigma)$. In other words, minimizing the cost functional here may give an approximation to a solution of the PI problem.

Our method is related to those in [25], [26], it is a combination of the least-squares method and the equation error method, but is formulated in a least-squares mixed approach. It is interesting to notice that a similar least-squares mixed/FOSLS (first order system least-squares) formulation has been used to develop finite element methods for linear and nonlinear forward problems in [10], [44], [45], [11], [16].

2.1.1 Existence and uniqueness issues for the minimizer in 1d

In this section, we intend to prove existence and uniqueness of the minimizer for the continuous case of our PI problem.

Let $\mathcal{H}$ be the following triple product

$$\mathcal{H} = H^1(0,1) \times H^1_0(0,1) \times H^1(0,1) \tag{2.1.9}$$

and let us assume that $\mathcal{H}$ is equipped with the usual norm of the product space, i.e., for any $\mathbf{w} = (a, u, \sigma)^T \in \mathcal{H}$, we have

$$\|\mathbf{w}\|_{\mathcal{H}} = \sqrt{\|a\|_1^2 + \|u\|_1^2 + \|\sigma\|_1^2}. \tag{2.1.10}$$

We first notice that $\mathcal{J}(\mathbf{w})$ is weakly lower semicontinuous on $\mathcal{H}$ and weakly coercive. Then the standard existence theory in [46] leads to the following results.

Theorem 2.1.1 *The cost functional (2.1.6) must have a minimizer in $\mathcal{H}$.*

The minimizers of (2.1.6) have interesting features stated in the next lemmas which eventually lead to the uniqueness of the minimizer under certain conditions.

Lemma 2.1.1 *Assume the data function $z(x)$ is such that there exist $\tilde{a}(x)$ and $\tilde{u}(x)$ satisfying (2.1.1) and boundary condition (2.1.2) or (2.1.3), $z = \tilde{u}$, and $\tilde{\mathbf{w}} = (\tilde{a}, \tilde{u}, \tilde{\sigma})^T \in \mathcal{H}$ with $\tilde{\sigma} = \tilde{a}\tilde{u}'$. In addition, assume*

$$\min\{\gamma, \rho, \delta\} > 0.$$

Then all the minimizers of $\mathcal{J}(\mathbf{w})$ in $\mathcal{H}$ belong to the set,

$$\mathcal{P} = \{\mathbf{w} \in \mathcal{H} : \|\mathbf{w}\|_{\mathcal{H}} \leq \sqrt{\frac{max\{\gamma,\rho,\delta\}}{min\{\gamma,\rho,\delta\}}}\|\tilde{\mathbf{w}}\|_{\mathcal{H}}\}$$

Proof. Let $\mathbf{w}_* = (a_*, u_*, \sigma_*)^T$ be a minimizer. Then

$$\mathcal{J}(\mathbf{w}_*) \leq \mathcal{J}(\tilde{\mathbf{w}}) = \gamma\|\tilde{a}\|_1^2 + \rho\|\tilde{u}\|_1^2 + \delta\|\tilde{\sigma}\|_1^2.$$

On the other hand,

$$\begin{aligned} \min\{\gamma,\rho,\delta\}\|a_*\|_1^2 &\leq \mathcal{J}(\mathbf{w}_*), \\ \min\{\gamma,\rho,\delta\}\|u_*\|_1^2 &\leq \mathcal{J}(\mathbf{w}_*), \\ \min\{\gamma,\rho,\delta\}\|\sigma_*\|_1^2 &\leq \mathcal{J}(\mathbf{w}_*). \end{aligned}$$

Combining the above four inequalities, we get that

$$\min\{\gamma,\rho,\delta\}\|\mathbf{w}_*\|_{\mathcal{H}}^2 \leq \max\{\gamma,\rho,\delta\}\|\tilde{\mathbf{w}}\|_{\mathcal{H}}^2$$

which implies that $\mathbf{w}_* \in \mathcal{P}$. The following lemma gives a necessary condition satisfied by a minimizer.

Lemma 2.1.2 *Let $\mathbf{w}_* = (a_*, u_*, \sigma_*)^T$ be a minimizer of $\mathcal{J}(a, u, \sigma)$, then*

$$\mathcal{A}(\mathbf{w}_*, \mathbf{v}) = \mathcal{B}(\mathbf{w}_*, \mathbf{v})$$

for any $\mathbf{v} = (\tilde{a}, \tilde{u}, \tilde{\sigma})^T \in \mathcal{H}$, where for any $\mathbf{w}$, $\mathbf{v} \in \mathcal{H}$,

$$\begin{aligned} \mathcal{A}(\mathbf{w},\mathbf{v}) &= \gamma[(a',\tilde{a}') + (a,\tilde{a})] + \rho(u',\tilde{u}') + (\rho+1)(u,\tilde{u}) + \\ &\quad (\delta+\beta)(\sigma',\tilde{\sigma}') + (\delta+\alpha)(\sigma,\tilde{\sigma}), \\ \mathcal{B}(\mathbf{w},\mathbf{v}) &= \begin{cases} \mathcal{B}^1(\mathbf{w},\mathbf{v}), & \text{for boundary condition (2.1.2)}, \\ \mathcal{B}^2(\mathbf{w},\mathbf{v}), & \text{for boundary condition (2.1.3)}, \end{cases} \\ \mathcal{B}^1(\mathbf{w},\mathbf{v}) &= -\alpha(u'(au'-\sigma),\tilde{a}) - \alpha(a(au'-\sigma),\tilde{u}') + \\ &\quad \alpha(au',\tilde{\sigma}) + (z,\tilde{u}) - \beta(f,\tilde{\sigma}'), \\ \mathcal{B}^2(\mathbf{w},\mathbf{v}) &= \mathcal{B}^1(\mathbf{w},\mathbf{v}) - \xi\sigma(1)\tilde{\sigma}(1) \\ &\quad -\theta(a(1)u'(1) - \sigma(1))(u'(1)\tilde{a}(1) + a(1)\tilde{u}'(1) - \tilde{\sigma}(1)). \end{aligned} \tag{2.1.11}$$

Proof. The above forms are Euler's equations in its weak form satisfied by the minimizer.

The functional $\mathcal{A}(\mathbf{w},\mathbf{v})$ satisfies the following

Lemma 2.1.3 *$\mathcal{A}(\mathbf{w},\mathbf{v})$ is a bilinear form on $\mathcal{H}\times\mathcal{H}$ satisfying the following estimates:*

$$\begin{aligned} \mid \mathcal{A}(\mathbf{w},\mathbf{v}) \mid &\leq \max\{\gamma,\rho+1,\delta+\beta,\delta+\alpha\}\|\mathbf{w}\|_{\mathcal{H}}\|\mathbf{v}\|_{\mathcal{H}}, \\ \mathcal{A}(\mathbf{w},\mathbf{w}) &\geq \min\{\gamma,\rho,\delta+\beta,\delta+\alpha\}\|\mathbf{w}\|^2_{\mathcal{H}}. \end{aligned} \tag{2.1.12}$$

The form $\mathcal{B}(\mathbf{w},\mathbf{v})$ satisfies the following

Lemma 2.1.4 *For each $\mathbf{w}\in\mathcal{H}$, $\mathcal{B}(\mathbf{w},\mathbf{v})$ is a continuous linear functional in $\mathbf{v}$.*

Proof. The linearity is easily seen. Now let $\{\mathbf{v}_n\}_{n=1}^{\infty} = \{(\tilde{a}_n,\tilde{u}_n,\tilde{\sigma}_n)\}$ be a sequence in $\mathcal{H}$ converging to 0. Then we have $\tilde{a}_n\to 0$, $\tilde{u}_n\to 0$, $\tilde{\sigma}_n\to 0$ in the H^1 norm. By Sobolev's embedding Theorem in one dimension, we have the following estimates

$$\begin{aligned} \mid (u'(au'-\sigma),\tilde{a}_n) \mid &\leq \mid (a(u')^2,\tilde{a}_n) \mid + \mid (u'\sigma,\tilde{a}_n) \mid \\ &\leq \|a\|_\infty\|u\|_1^2\|\tilde{a}_n\|_\infty + \|u\|_1\|\sigma\|_1\|\tilde{a}_n\|_\infty \\ &\leq (\|a\|_1\|u\|_1^2 + \|u\|_1\|\sigma\|_1)\|\tilde{a}_n\|_1 \\ \mid (a(au'-\sigma),\tilde{u}_n') \mid &\leq \mid (a^2u',\tilde{u}_n') \mid + \mid (a\sigma,\tilde{u}_n') \mid \\ &\leq \|a^2\|_\infty\|u\|_1\|\tilde{u}_n\|_1 + \|a\|_\infty\|\sigma\|_1\|\tilde{u}_n\|_1 \\ &\leq (\|a\|_1^2\|u\|_1 + \|a\|_1\|\sigma\|_1)\|\tilde{u}_n\|_1 \\ \mid (au',\tilde{\sigma}_n) \mid &\leq \|a\|_\infty\|u\|_1\|\tilde{\sigma}_n\|_1 \leq \|a\|_1\|u\|_1\|\tilde{\sigma}_n\|_1 \end{aligned}$$

All three terms approach zero as $\mathbf{v}_n\to 0$. Hence $\mathcal{B}(\mathbf{w},\mathbf{v})$ is a continuous linear functional in $\mathbf{v}$.

From Lemma 2.1.3 and Riesz Representation Theorem, there exists a bounded linear operator $\tilde{\mathcal{A}}:\mathcal{H}\to\mathcal{H}^*$ such that

$$\mathcal{A}(\mathbf{w},\mathbf{v}) = (\tilde{\mathcal{A}}\mathbf{w},\mathbf{v})_{\mathcal{H}}$$

Moreover, $\tilde{\mathcal{A}}^{-1}$ exists and $\tilde{\mathcal{A}}^{-1}\in\mathcal{L}(\mathcal{H}^*,\mathcal{H})$ with

$$\|\tilde{\mathcal{A}}^{-1}\| \leq \frac{1}{\min\{\gamma,\rho,\delta+\beta,\delta+\alpha\}}.$$

From Lemma 2.1.4, there exists a nonlinear operator $\tilde{\mathcal{B}}:\mathcal{H}\to\mathcal{H}^*$ such that

$$\mathcal{B}(\mathbf{w},\mathbf{v}) = (\tilde{\mathcal{B}}\mathbf{w},\mathbf{v})$$

By Lemma 2.1.2, a minimizer $\mathbf{w}_* = (a_*, u_*, \sigma_*)^T$ of $\mathcal{J}(a, u, \sigma)$ is a solution of the following operator equation:

$$\tilde{\mathcal{A}}\mathbf{w} = \tilde{\mathcal{B}}\mathbf{w}, \tag{2.1.13}$$

i.e., $\mathbf{w}_*$ should be a fixed point of the nonlinear operator $\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}$.

Lemma 2.1.5 *Let us assume that (2.1.2) is used in the boundary value problem, and*

$$\sqrt{\frac{\max\{\gamma, \rho, \delta\}}{\min\{\gamma, \rho, \delta\}}} \leq C < \infty$$

then the operator $\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}$ is a contraction on $\mathcal{P}$ if α is chosen small enough.

Proof. For any $\mathbf{v} = (a, u, \sigma)^T \in \mathcal{H}$ and $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{P}$ with

$$\mathbf{w}_1 = \begin{pmatrix} a_1 \\ u_1 \\ \sigma_1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} a_2 \\ u_2 \\ \sigma_2 \end{pmatrix}.$$

By Sobolev's embedding Theorem in one dimension, we have

$$\begin{aligned}
&| \, (u_1'(a_1u_1' - \sigma_1), a) - (u_2'(a_2u_2' - \sigma_2), a) \, | \\
&\leq | \, (a_1(u_1')^2 - a_2(u_2')^2, a) \, | + | \, (u_1'\sigma_1 - u_2'\sigma_2, a) \, | \\
&\leq | \, ((a_1 - a_2)(u_1')^2, a) \, | + | \, (a_2(u_1' + u_2')(u_1' - u_2'), a) \, | + \\
&| \, (u_1'(\sigma_1 - \sigma_2), a) \, | + | \, ((u_1' - u_2')\sigma_2, a) \, | \\
&\leq \|u_1\|_1^2\|a_1 - a_2\|_\infty\|a\|_\infty + \|a_2\|_\infty\|u_1 + u_2\|_1\|u_1 - u_2\|_1\|a\|_\infty + \\
&\|u_1\|_1\|\sigma_1 - \sigma_2\|_1\|a\|_\infty + \|\sigma_2\|_1\|u_1 - u_2\|_1\|a\|_\infty \\
&\leq \|u_1\|_1^2\|a_1 - a_2\|_1\|a\|_1 + \|a_2\|_1\|u_1 + u_2\|_1\|u_1 - u_2\|_1\|a\|_1 + \\
&\|u_1\|_1\|\sigma_1 - \sigma_2\|_1\|a\|_1 + \|\sigma_2\|_1\|u_1 - u_2\|_1\|a\|_1 \\
&\leq (\|u_1\|_1^2\|a\|_1 + \|a_2\|_1\|u_1 + u_2\|_1\|a\|_1 + \|u_1\|_1\|a\|_1 + \|\sigma_2\|_1\|a\|_1)\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H} \\
&\leq (\|u_1\|_1^2\|a\|_1 + \|a_2\|_1(\|u_1\|_1 + \|u_2\|_1)\|a\|_1 + \|u_1\|_1\|a\|_1 + \|\sigma_2\|_1\|a\|_1)\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H} \\
&\leq C\|a\|_1\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H}
\end{aligned}$$

$$\begin{aligned}
& \quad | \, (a_1(a_1u_1' - \sigma_1), u') - (a_2(a_2u_2' - \sigma_2), u') \, | \\
\leq & \quad | \, (a_1^2u_1' - a_2^2u_2', u') \, | + | \, (a_1\sigma_1 - a_2\sigma_2, u') \, | \\
\leq & \quad C\|u\|_1\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H} \\
& \quad | \, (a_1u_1', \sigma) - (a_2u_2', \sigma) \, | \\
\leq & \quad C\|\sigma\|_1\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H}
\end{aligned}$$

All the C's in these estimates depend only on the diameter of $\mathcal{P}$ and lead to the following inequality

$$\|\tilde{\mathcal{B}}\mathbf{w}_1 - \tilde{\mathcal{B}}\mathbf{w}_2\|_{\mathcal{H}^*} \leq \alpha L_\mathcal{P}\|\mathbf{w}_1 - \mathbf{w}_2\|_\mathcal{H},$$

Here $L_{\mathcal{P}}$ is a constant depending only on the diameter of $\mathcal{P}$. Hence we obtain

$$\begin{aligned}\|\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}\mathbf{w}_1 - \tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}\mathbf{w}_2\|_{\mathcal{H}} &\leq \|\tilde{\mathcal{A}}^{-1}\|\|\tilde{\mathcal{B}}\mathbf{w}_1 - \tilde{\mathcal{B}}\mathbf{w}_2\|_{\mathcal{H}^*} \\ &\leq \|\tilde{\mathcal{A}}^{-1}\|\alpha L_{\mathcal{P}}\|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathcal{H}}\end{aligned}$$

and this inequality implies that $\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}$ is a contraction mapping if α is chosen small enough.

Similar results can be withdrawn for the boundary value problem (2.1.1), (2.1.3).

Lemma 2.1.6 *Assume that (2.1.3) is used in the boundary value problem, and*

$$\sqrt{\frac{max\{\gamma,\rho,\delta\}}{min\{\gamma,\rho,\delta\}}} \leq C < \infty$$

then the operator $\tilde{\mathcal{A}}^{-1}\tilde{\mathcal{B}}$ *is a contraction on* $\mathcal{P}$ *if* α, ξ, *and* θ *are chosen small enough.*

Finally, we have the following uniqueness result for the minimizer.

Theorem 2.1.2 *If the conditions of Lemma 2.1.1 and 2.1.5 or 2.1.6 are satisfied, then* $\mathcal{J}(a, u, \sigma)$ *has a unique minimizer in* $\mathcal{H}$.

2.2 Finite Element Method

In this section, we formulate a finite element discretization of the cost functional and discuss the related minimization problem. To avoid repeating ourselves twice, we present computation formulas only for the PI problem with boundary condition (2.1.2). Minor modifications lead to formulas for boundary condition (2.1.3). In our estimations, we regard C as a generic constant, which may vary in different contexts, but is always independent of h, unless specified otherwise. For the sake of completeness, we first give some definitions in $I\!R^n$.

In order to define the finite element spaces, let $\{\Delta_h\}$, with $h \in (0, 1)$, be a family of regular, quasi-uniform partitions of $\Omega \subset I\!R^n$. That is, for each triangle $T \in \{\Delta_h\}$,

$$\frac{h_T}{\rho_T} \leq \sigma \text{ and } \frac{h}{h_T} \leq \nu,$$

where h_T =diameter of $T \leq h$, and ρ_T = sup {diam(S); S is a ball contained in T}. For fixed integers $r \geq 1$, $l \geq 0$, we let

$$S^r_{h,l} = \{v \in C^{l-1}(\bar{\Omega}) : v|_T \in P_r \text{ for all } T \in \Delta_h\}, \tag{2.2.14}$$

be a Lagrange type finite element space, where P_r is the space of polynomials of degree less than or equal to r and $C^{-1}(\Omega)$ is interpreted as $L^2(\Omega)$. By $S^{r,0}_{h,l}$, we denote the subspace of functions of $S^r_{h,l}$ which vanish at $x=0$ and $x=1$. By the results in [14], we know that for all $v \in W^{m,p}(\Omega)$, there is (an interpolant) $v_h \in S^r_{h,l}$ such that

$$\|v - v_h\|_{k,p} \leq Ch^{m-k}\|v\|_{m,p} \text{ for } 0 \leq k \leq l,\ k \leq m \leq r+1,\ 1 \leq p \leq \infty\ . \tag{2.2.15}$$

Also, these spaces satisfy the following inverse inequalities,

$$\|v_h\|_{1,p} \leq Ch^{-1}\|v_h\|_{0,p} \text{ for all } v_h \in S^r_{h,l}, 1 \leq p \leq \infty \tag{2.2.16}$$

and

$$\|v_h\|_\infty \leq Ch^{-\frac{n}{2}}\|v_h\|_0 \text{ for all } v_h \in S^r_{h,l}. \tag{2.2.17}$$

We denote by $\tilde{H}^{-1} = \tilde{H}^{-1}(\Omega)$ the dual space $(\tilde{H}^1(\Omega))^*$ equipped with the norm

$$\|v\|_{-1} = \sup_{\psi \in \tilde{H}^1(\Omega)} \frac{|(v,\psi)|}{\|\psi\|_1}. \tag{2.2.18}$$

Let us now introduce those finite dimensional spaces needed to define the computational procedure. We need altogether three different discretization spaces, namely, U_h for the forward solution u, A_h for the parameter a, and finally Σ_h which will be used to discretize the flux variable σ. We assume that these finite element spaces are,

$$\begin{array}{rcll} A_h & = & S^{d_a}_{h,l}, & \text{here } l = 0,1, \\ U_h & = & S^{d_u}_{h,1} \cap \tilde{H}^1(0,1), & \\ \Sigma_h & = & S^{d_\sigma}_{h,1}, & \end{array} \tag{2.2.19}$$

where d_a, d_u, d_σ are integers that we will fix later.

Now we are ready to define the numerical method to solve the PI problem:
(P_h) seek $a_h^* \in A_h$, $u_h^* \in U_h$, $\sigma_h^* \in \Sigma_h$, such that

$$\mathcal{J}_h(a_h^*, u_h^*, \sigma_h^*) = \min_{a \in A_h, u \in U_h, \sigma \in \Sigma_h} \mathcal{J}_h(a,u,\sigma) \tag{2.2.20}$$

where $\mathcal{J}_h$ is defined by,

$$\mathcal{J}_h(\mathbf{w}_h) = \begin{cases} \mathcal{J}_{h,1}(\mathbf{w}_h), & \text{if (2.1.2) holds,} \\ \mathcal{J}_{h,2}(\mathbf{w}_h), & \text{if (2.1.3) holds,} \end{cases} \tag{2.2.21}$$

where

$$\begin{array}{rcl} \mathcal{J}_{h,1}(\mathbf{w}_h) & = & \|u_h - z\|_0^2 + \alpha\|\sigma_h - a_h u_h'\|_0^2 + \beta\|\sigma_h' + f\|_0^2 + \\ & & \gamma\|a_h\|_{r_a}^2 + \rho\|u_h\|_{r_u}^2 + \delta\|\sigma_h\|_{r_\sigma}^2, \\ \mathcal{J}_{h,2}(\mathbf{w}_h) & = & \mathcal{J}_{h,1}(\mathbf{w}_h) + \xi(\sigma_h(1))^2 + \theta(a_h(1)u_h'(1) - \sigma_h(1))^2. \end{array} \tag{2.2.22}$$

where $\mathbf{w}_h = (a_h, u_h, \sigma_h) \in A_h \times U_h \times \Sigma_h$. As we can see, the cost functional (2.2.21) consists of several terms, the first term represents the usual output least squares formulation with L^2-observation, the second term is the approximation of au' with the flux variable σ, while the third term takes into account the actual state equation (2.1.1) where we have replaced au' by the flux variable σ. The remaining terms are regularization terms. The advantage of this variational approach is that we allow all variables (a, u, and σ) to vary in the corresponding discrete subspace without differentiating any kind of error corrupted data z. The weights γ, ρ, and δ will be chosen so that all terms of $\mathcal{J}_h(\mathbf{w}_h)$ are balanced.

2.3 Some error estimates in weighted norms

In this section, we establish some error estimates for an approximate solution to the PI problem generated by the finite element method in the previous section. The estimates and the arguments are influenced by [20], [25], [26] and [29], but our presentation here leads to important applications such as the choice of the regularization parameters and the choice of the finite element spaces for the optimal convergence rates. In all the discussion below, C denotes a generic constant which may take a different value at different places and is independent of h unless specified otherwise. Let I_a^h, I_u^h, and I_σ^h be the interpolation operators in the finite element spaces A_h, U_h, and Σ_h, respectively. The regularization spaces are chosen so that

$$1 \leq r_a \leq d_a + 1, 1 \leq r_u \leq d_u + 1, 1 \leq r_\sigma \leq d_\sigma + 1.$$

Moreover, there exists a constant C such that the interpolation operators satisfy,

$$\|I_a^h a\|_{r_a} \leq C\|a\|_{r_a+1}, \|I_u^h u\|_{r_u} \leq C\|u\|_{r_u+1}, \|I_\sigma^h \sigma\|_{r_\sigma} \leq C\|\sigma\|_{r_\sigma+1},$$

for any $a \in H^{r_a+1}(0,1), u \in H^{r_u+1}(0,1), \sigma \in H^{r_\sigma+1}(0,1)$. We first derive an estimate for a minimizer of $\mathcal{J}_h(\mathbf{w})$.

Lemma 2.3.1 *Let us assume the following hypothesis,*

* *Let $z(x)$ be a distributed L^2 observation of the state $\tilde{u}$ with an observation error of the form*

$$\|\tilde{u} - z\|_0 \leq \epsilon.$$

* *Let $\tilde{a}$, $\tilde{u}$ satisfy the differential equation (2.1.1) together with the boundary condition either (2.1.2) or (2.1.3). Moreover, $\tilde{a}$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}\tilde{u}'$, have the following regularities:*

$$\tilde{a} \in H^{d_a+1}(0,1), \tilde{u} \in H^{d_u+1}(0,1), \tilde{\sigma} \in H^{d_\sigma+1}(0,1).$$

* *The finite element space U_h is such that*

$$\|(I_u^h\tilde{u})'\|_\infty \le C,$$

for h small enough.

* *The regularization parameters γ, ρ, and δ are chosen such that*

$$\gamma+\rho+\delta \sim C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\}.$$

with a constant C independent of h.

* *If the boundary condition (2.1.3) is used in (2.1.1), we assume*

$$\xi = 1, \theta = h^2. \tag{2.3.23}$$

Let $\mathbf{w}_h^* = (a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h(a,u,\sigma)$ defined by (2.2.21) in $A_h \times U_h \times \Sigma_h$, then we have the following estimate for the cost functional $\mathcal{J}_h$,

$$\mathcal{J}_h(\mathbf{w}_h^*) \le C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\}, \tag{2.3.24}$$

here C is independent of ϵ and h. Moreover, we get the following bounds,

$$\begin{aligned}
\|\sigma_h^* - a_h^*(u_h^*)'\|_0 &\le \tfrac{C}{\sqrt{\alpha}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\},\\
\|(\sigma_h^*)' + f\|_0 &\le \tfrac{C}{\sqrt{\beta}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}.
\end{aligned}$$

Proof. We treat the case in which the boundary condition is (2.1.2). Similar arguments can be applied for boundary condition (2.1.3). Let $(a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h$, which will be denoted by $\mathbf{w}_h^*$, and let $I_a^h\tilde{a}$, $I_u^h\tilde{u}$, and $I_\sigma^h\tilde{\sigma}$ be the interpolates of $\tilde{a}$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}\tilde{u}'$ on Δ_h, respectively. Now

$$\mathcal{J}_h(\mathbf{w}_h^*) \le \mathcal{J}_h(I_a^h\tilde{a}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}),$$

and then

$$\begin{aligned}
\mathcal{J}_h(I_a^h\tilde{a}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}) &= \|I_u^h\tilde{u} - z\|_0^2 + \alpha\|I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}(I_u^h\tilde{u})'\|_0^2 + \beta\|(I_\sigma^h\tilde{\sigma})' + f\|_0^2 +\\
&\quad \gamma\|I_a^h\tilde{a}\|_{r_a}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2.
\end{aligned}$$

The first three terms of the equality above have the following bounds,

$$\begin{aligned}
\|I_u^h\tilde{u} - z\|_0 &\le \|I_u^h\tilde{u} - \tilde{u}\|_0 + \|\tilde{u} - z\|_0\\
&\le Ch^{d_u+1}\|\tilde{u}\|_{d_u+1} + \epsilon,\\
\|I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}(I_u^h\tilde{u})'\|_0 &\le \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_0 + \|\tilde{a}\tilde{u}' - \tilde{a}(I_u^h\tilde{u})'\|_0 + \|\tilde{a}(I_u^h\tilde{u})' - I_a^h\tilde{a}(I_u^h\tilde{u})'\|_0\\
&\le Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + \|\tilde{a}\|_\infty\|\tilde{u}' - (I_u^h\tilde{u})'\|_0 + \|(I_u^h\tilde{u})'\|_\infty\|\tilde{a} - I_a^h\tilde{a}\|_0\\
&\le Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + Ch^{d_u}\|\tilde{u}\|_{d_u+1} + Ch^{d_a+1}\|\tilde{a}\|_{d_a+1},\\
\|(I_\sigma^h\tilde{\sigma})' + f\|_0 &\le \|(I_\sigma^h\tilde{\sigma})' - \tilde{\sigma}'\|_0\\
&\le Ch^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1}.
\end{aligned}$$

Then, so far

$$\begin{array}{rcl}\mathcal{J}_h(\mathbf{w}_h^*) & \leq & C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_a+1)}) + \beta h^{2d_\sigma}\} + \\ & & \gamma\|I_a^h\tilde{a}\|_{r_a}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2,\end{array}$$

since

$$\begin{array}{rcl}\|I_a^h\tilde{a}\|_{r_a} & \leq & C\|\tilde{a}\|_{r_a+1}, \\ \|I_u^h\tilde{u}\|_{r_u} & \leq & C\|\tilde{u}\|_{r_u+1}, \\ \|I_\sigma^h\tilde{\sigma}\|_{r_\sigma} & \leq & C\|\tilde{\sigma}\|_{r_\sigma+1},\end{array}$$

and choosing γ, ρ, and δ as in the hypothesis, we get the main assertion of the lemma:

$$\mathcal{J}_h(\mathbf{w}_h^*) \quad \leq \quad C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_a+1)}) + \beta h^{2d_\sigma}\}.$$

We can get the next two bounds from the definition of $\mathcal{J}_h(a, u, \sigma)$ and the bound above,

$$\begin{array}{rcl}\|\sigma_h^* - a_h^*(u_h^*)'\|_0 & \leq & \frac{C}{\sqrt{\alpha}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\} \\ \|(\sigma_h^*)' + f\|_0 & \leq & \frac{C}{\sqrt{\beta}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}\end{array}$$

As a consequence of this lemma, it can be shown that under certain conditions, the minimizer of $\mathcal{J}_h(\mathbf{w}_h)$ is uniformly bounded with respect to h. The result is stated in the following

Corollary 2.3.1 *If all the conditions of Lemma 2.3.1 hold, and*

$$\begin{array}{rcccl}0 & < & \gamma & = & C\{(\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\}, \\ 0 & < & \rho & = & C\{(\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\}, \\ 0 & < & \delta & = & C\{(\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\},\end{array}$$

then there exists a constant C such that,

$$\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C, \ \textit{for all } h.$$

Proof. In this proof, C denotes a generic constant. Let $\mathbf{w}_h^*$ be a minimizer of $\mathcal{J}_h$. Then according to the definition of $\mathcal{J}_h$ and the bound (2.3.24), we have that

$$\begin{array}{rcccl}\gamma\|a_h^*\|_{r_a}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C, \\ \rho\|u_h^*\|_{r_u}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C, \\ \delta\|\sigma_h^*\|_{r_\sigma}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C.\end{array}$$

Since γ, ρ, and δ are positive and finite, then

$$\begin{array}{rcccl}\|a_h^*\|_{r_a}^2 & \leq & \frac{1}{\sqrt{\gamma}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C, \\ \|u_h^*\|_{r_u}^2 & \leq & \frac{1}{\sqrt{\rho}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C, \\ \|\sigma_h^*\|_{r_\sigma}^2 & \leq & \frac{1}{\sqrt{\delta}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C.\end{array}$$

Therefore $\|\mathbf{w}_h^*\|_{\mathcal{H}} \le C$, for all h.

The following theorem establishes our basic error estimates.

Theorem 2.3.1 *Suppose that all the conditions in Lemma 2.3.1 and Corollary 2.3.1 hold. Let β be chosen such that*

$$\frac{C}{\sqrt{\beta}} + 1 \le \frac{C'}{\sqrt{\beta}} \tag{2.3.25}$$

for some constants C and C' independent of h. Then the minimizer $\mathbf{w}_h^* = (a_h^*, u_h^*, \sigma_h^*)^T$ of $\mathcal{J}_h(\mathbf{w}_h)$ satisfies the following weighted H^{-1}-norm estimation,

$$\|((\tilde{a} - a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} \le C(\tfrac{1}{\sqrt{\alpha}} + \tfrac{1}{\sqrt{\beta}} + \tfrac{1}{h})(\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta} h^{d_\sigma})$$

Proof. The boundary condition satisfied by $\tilde{u}$ and integration by parts give us the following estimation,

$$\begin{aligned} \|((\tilde{a} - a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} &= \sup_{v \in \tilde{H}^1(0,1)} \frac{|(((\tilde{a} - a_h^*)\tilde{u}')', v)|}{\|v\|_1} \\ &= \sup_{v \in \tilde{H}^1(0,1)} \frac{|((\tilde{a} - a_h^*)\tilde{u}', v')|}{\|v\|_1} \end{aligned} \tag{2.3.26}$$

We now estimate the term $\int_0^1 (\tilde{a} - a_h^*)\tilde{u}'v'\,dx$ when $v \in \tilde{H}^1(0,1)$. By using the variational form of equation (2.1.1) with the boundary condition (2.1.2), adding and subtracting various terms and integrating by parts, we get

$$\begin{aligned} ((\tilde{a} - a_h^*)\tilde{u}', v') &= (\tilde{a}\tilde{u}' - a_h^*(u_h^*)', v') + (a_h^*(u_h^* - \tilde{u})', v') \\ &= (f + (\sigma_h^*)', v) + (\sigma_h^* - a_h^*(u_h^*)', v') + (a_h^*(u_h^* - \tilde{u})', v'), \end{aligned}$$

By the same reasons as above, but with boundary condition (2.1.3), we get

$$\begin{aligned} ((\tilde{a} - a_h^*)\tilde{u}', v') &= (\tilde{a}\tilde{u}' - a_h^*(u_h^*)', v') + (a_h^*(u_h^* - \tilde{u})', v') \\ &= (f + (\sigma_h^*)', v) - \sigma_h^*(1)v(1) + (\sigma_h^* - a_h^*(u_h^*)', v') + \\ &\quad (a_h^*(u_h^* - \tilde{u})', v'), \end{aligned}$$

By Lemma 2.3.1 and (2.3.25), we have

$$\left.\begin{array}{r} |\ (f + (\sigma_h^*)', v)\ | \\ |\ (f + (\sigma_h^*)', v)\ | + |\ \sigma_h^*(1)v(1)\ | \end{array}\right\} \le \frac{C}{\sqrt{\beta}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}\|v\|_1 \tag{2.3.27}$$

$$|\ (\sigma_h^* - a_h^*(u_h^*)', v')\ | \le \frac{C}{\sqrt{\alpha}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}\|v'\|_0 \tag{2.3.28}$$

By Lemma 2.3.1 and Corollary 2.3.1 and the inverse inequality (2.2.16), we get

$$\begin{aligned} |\,(a_h^*(u_h^*-\tilde{u})',v')\,| &\le \|a_h^*\|_\infty \|(u_h^*-\tilde{u})'\|_0\|v'\|_0 \\ &\le C(\|(u_h^*-I_u^h\tilde{u})'\|_0+\|(I_u^h\tilde{u}-\tilde{u})'\|_0)\|v\|_1 \\ &\le C(\frac{1}{h}\|u_h^*-I_u^h\tilde{u}\|_0+h^{d_u}\|\tilde{u}\|_{d_u+1})\|v\|_1 \\ &\le C(\frac{1}{h}(\|u_h^*-z\|_0+\|z-\tilde{u}\|_0+\|\tilde{u}-I_u^h\tilde{u}\|_0)+h^{d_u}\|\tilde{u}\|_{d_u+1})\|v\|_1 \\ &\le C\{\frac{1}{h}(\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_a+1}+h^{d_u}+h^{d_\sigma+1})+\sqrt{\beta}h^{d_\sigma})\}\|v\|_1. \end{aligned}$$

Finally, we get the assertion of the theorem by combining the above three estimations,

$$\|((\tilde{a}-a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} \le C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{h})(\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma})$$

The estimate in Theorem 2.3.1 becomes more meaningful for the boundary condition (2.1.3), as is indicated by the following result.

Corollary 2.3.2 *Suppose all the conditions in Theorem 2.3.1 hold and that the boundary condition in the PI problem is (2.1.3), then a minimizer $\mathbf{w}_h^*$ of $\mathcal{J}_h(a,u,\sigma)$ satisfies the following weighted L^2-norm estimation,*

$$\|(\tilde{a}-a_h^*)\tilde{u}'\|_0 \le C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{h})(\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma})$$

Proof. Let ψ be the solution of the following initial value problem,

$$\begin{cases} \psi'(x) = (\tilde{a}(x)-a_h^*(x))\tilde{u}'(x), \ x\in(0,1), \\ \psi(0) = 0. \end{cases}$$

The solution to the above initial value problem is $\psi(x)=\int_0^x(\tilde{a}(t)-a_h^*(t))\tilde{u}'(t)\,dt$, so $\psi\in\tilde{H}^1(0,1)$ and by (2.3.26) and theorem 2.3.1, we have

$$\begin{aligned} ((\tilde{a}-a_h^*)\tilde{u}',\psi') &\le \|((\tilde{a}-a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|\psi\|_1, \\ ((\tilde{a}-a_h^*)\tilde{u}',\psi') &\le \|((\tilde{a}-a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|\psi'\|_0 \text{ by Poincare's inequality,} \\ ((\tilde{a}-a_h^*)\tilde{u}',\psi') &\le \|((\tilde{a}-a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|(\tilde{a}-a_h^*)\tilde{u}'\|_0. \end{aligned}$$

It follows that

$$\|(\tilde{a}-a_h^*)\tilde{u}'\|_0 \le \|((\tilde{a}-a_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}$$

and the assertion of the Corollary follows straightforward from Theorem 2.3.1.

Corollary 2.3.3 *Assume that the boundary condition (2.1.2) holds in the PI problem. Then under the conditions of Theorem 2.3.1, we have the following weighted estimation in the H^{-1}-norm,*

$$\|((\tilde{a}-a_h^*)\tilde{u}')'\|_{-1} \le C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{h})(\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma})$$

Proof. The result follows from theorem 2.3.1, since in this case $\tilde{H}^1(0,1) = H_0^1(0,1)$, and $(\tilde{H}^1(0,1))^* = H^{-1}(0,1)$.

Unfortunately, the error estimation in the weighted norm above is not practicable since a good approximation in H^{-1} of the coefficient $a(x)$ is physically useless. That is why we pursue a weighted L^1 norm error estimate for the PI problem with Dirichlet boundary condition in the next section.

2.4 A weighted-L^1 error estimation

In this section, we want to get a weighted L^1-norm estimation for the error $\phi = a_h^* - \tilde{a}$. Here $\tilde{a}$ is the solution to the equation (2.1.1), and a_h^* is the first component of the minimizer of the cost functional (2.1.6) with homogeneous Dirichlet boundary condition. Then we have the following

Theorem 2.4.1

$$\int_0^1 |a_h^* - a||u'|^2\,dx \le C\{\delta + c_0 + c_1(\delta^{-1} + 1)\}$$

for any $\delta > 0$*, where*

$$c_0 = \|f + (\sigma_h^*)'\|_0$$

and

$$c_1 = \|\sigma_h^* - a_h^*(u_h^*)'\|_0 + |a_h^*|_{\infty,(0,1)}\|(u_h^* - u)'\|_0$$

Proof. Remember that the variational form of equation (2.1.1) reads as,

$$\int_0^1 au'v'\,dx = \int_0^1 fv\,dx \text{ for all } v \in H_0^1(0,1) \tag{2.4.29}$$

Set $\phi = a_* - a$. Now, we estimate the term

$$|\int_0^1 \phi u'v'\,dx\,|,$$

where v is a test function in $H_0^1(0,1)$. To this end, we add and subtract suitable terms and integrate by parts, to get

$$\begin{aligned} -\textstyle\int_0^1 \phi u'v'\,dx &= \textstyle\int_0^1 (f + \sigma_*')v\,dx + \int_0^1 (\sigma_* - a_*u_*')v'\,dx + \\ &\quad \textstyle\int_0^1 a_*(u_* - u)'v'\,dx \end{aligned} \tag{2.4.30}$$

Now taking the absolute value of $-\int_0^1 \phi u'v'\,dx$, we get the following estimation:

$$\begin{array}{rcl} \mid \int_0^1 \phi u'v'\,dx \mid & \leq & \|f+\sigma_*'\|_0\|v\|_0 + \|\sigma_* - a_*u_*'\|_0\|v'\|_0 + \\ & & |a_*|_{\infty,(0,1)}\|(u_*-u)'\|_0\|v'\|_0 \end{array} \tag{2.4.31}$$

which leads to the following inequality,

$$\mid \int_0^1 \phi u'v'\,dx \mid \leq c_0\|v\|_0 + c_1\|v'\|_0 \tag{2.4.32}$$

for all $v \in H_0^1(0,1)$, where $\phi = a_* - a$, and

$$c_0 = \|f+\sigma_*'\|_0$$

and

$$c_1 = \|\sigma_* - a_*u_*'\|_0 + |a_*|_{\infty,(0,1)}\|(u_*-u)'\|_0$$

We shall rewrite the left side of equation 2.4.32 as a sum of several terms. First, we approximate ϕ by a sequence of smooth functions $\psi_r \in C^\infty([0,1])$ with

$$\|\psi_r - \phi\|_1 \to 0 \text{ as } r \to \infty. \tag{2.4.33}$$

Next, for $\delta > 0$ and r fixed, we consider the following test functions,

$$v_1 = \delta^{-1}(\psi_r^+ \wedge \delta)u,\ v_2 = \delta^{-1}((-\psi_r)^+ \wedge \delta)u, \tag{2.4.34}$$

where $\psi_r^+ = \max\{\psi_r, 0\}$, $\psi_r^+ \wedge \delta = \min\{\psi_r^+, \delta\}$. Since $\|\psi_r\|_1 \leq C$ as a consequence that $\phi \in H^1(0,1)$ and the convergence in $H^1(0,1)$, we have

$$\begin{array}{rcl} \|v_i\|_0 & \leq & \|u\|_0, \\ \|v_i'\|_0 & \leq & C(\delta^{-1}+1), \end{array} \tag{2.4.35}$$

with C depending on u and for $i = 1,2$. We claim that

$$\begin{array}{rcl} \int_{\psi_r \geq \delta} \psi_r|u'|^2 & \leq & C\delta + \mid \int_0^1 \psi_r u'v_1' \mid \\ -\int_{\psi_r \leq -\delta} \psi_r|u'|^2 & \leq & C\delta + \mid \int_0^1 \psi_r u'v_2' \mid, \end{array} \tag{2.4.36}$$

We will only prove the first inequality, since the second one is parallel in every respect. The definitions (2.4.34) give

$$\begin{array}{rcl} \int_0^1 \psi_r u'v_1' & = & \int_{\psi_r \geq \delta} \psi_r|u'|^2 + \delta^{-1}\int_{0<\psi_r<\delta} \psi_r^2|u'|^2 + \\ & & \delta^{-1}\int_{0<\psi_r<\delta} \psi_r\psi_r'uu'. \end{array} \tag{2.4.37}$$

The second term on the right satisfies

$$\delta^{-1}\int_{0<\psi_r<\delta} \psi_r^2(u')^2 \leq C\delta. \tag{2.4.38}$$

In estimating the third term, we shall suppose (without loss of generality) that 0 and δ are regular values of ψ_r. Then

$$\delta^{-1} \int_{0<\psi_r<\delta} \psi_r \psi_r' u u' \, dx = (4\delta)^{-1} \int_{0<\psi_r<\delta} (\psi_r^2)'(u^2)' \, dx \tag{2.4.39}$$

It is very easy to see that, the set $\mathcal{O} = \{x \in (0,1) : 0 < \psi_r < \delta\}$, can be written as a countable union of disjoint open intervals, i.e., $\mathcal{O} = \bigcup I_i$, where each $I_i = (a_i, b_i)$.

$$\begin{aligned} (4\delta)^{-1} \int_{\{0<\psi_r<\delta\} \bigcap I_i} (\psi_r^2)'(u^2)' \, dx &= (4\delta)^{-1} \{ \psi_r^2 (u^2)' \,|_{a_i}^{b_i} - \\ &\quad \int_{\{0<\psi_r<\delta\} \bigcap I_i} \psi_r^2 (u^2)'' \, dx \}. \end{aligned} \tag{2.4.40}$$

Now, the factor ψ_r^2 in the integrated terms may be zero or δ^2. So in either case,

$$| \ \psi_r^2 (u^2)' \ |_{a_i}^{b_i} | \leq C\delta^2,$$

We proceed to estimate the integral on the right hand side. Clearly

$$\int_{\{0<\psi_r<\delta\} \bigcap I_i} \psi_r^2 (u^2)'' \, dx \leq C\delta^2,$$

Combining all these results with (2.4.39) gives,

$$-\delta^{-1} \int_{0<\psi_r<\delta} \psi_r \psi_r' u u' \, dx \leq C\delta,$$

Putting all these inequalities together, we get;

$$\int_{\psi_r \geq \delta} \psi_r |u|^2 \, dx \leq C\delta + \left| \int_0^1 \psi_r u' v_1' \, dx \right|,$$

Assembling (2.4.36) and (2.4.37) with the obvious result,

$$\int_{|\psi_r|<\delta} |\psi_r| |u'|^2 \leq C\delta,$$

We conclude that

$$\int_0^1 |\psi_r| |u'|^2 \, dx \leq C\{ \delta + \left| \int_0^1 \psi_r u' v_1' \, dx \right| + \left| \int_0^1 \psi_r u' v_2' \, dx \right| \}$$

By taking the limit of the expression above when $r \to \infty$, we get:

$$\int_0^1 |\phi| |u'|^2 \, dx \leq C\{ \delta + \left| \int_0^1 \phi u' v_1' \, dx \right| + \left| \int_0^1 \phi u' v_2' \, dx \right| \}$$

Applying (2.4.32) with $v = v_1$ and $v = v_2$, and using (2.4.35), we conclude that

$$\int_0^1 |\phi| |u'|^2 \, dx \leq C\{ \delta + c_0 + c_1(\delta^{-1} + 1) \}$$

This concludes the discussion of this Theorem.

2.5 Applications of the error estimates

One of the important applications of these error estimates is the suggestion for the choice of the regularization parameters. Lemma 2.3.1 implies that the regularization parameters γ, ρ, and δ are decided by α and β. The error estimates in Theorem 2.3.1 and its corollaries indicate that we should pick α and β so that all the terms in

$$\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_a+1} + h^{d_u} + h^{d_\sigma+1}) + \sqrt{\beta}h^{d_\sigma}$$

are not greater than h^{d_u+1}, since this is the leading term when the observation $z(x)$ equals $u = u(x)$. On the other hand, the smaller the α and β, the weaker the regularization effect. Therefore for the finite element method here, it seems to be natural to choose the regularization parameters according to the following rules:
1. Choose the largest possible α and β such that

$$\begin{aligned}
\sqrt{\alpha}h^{d_a+1} &\leq Ch^{d_u+1}, \\
\sqrt{\alpha}h^{d_u} &\leq Ch^{d_u+1}, \\
\sqrt{\alpha}h^{d_\sigma+1} &\leq Ch^{d_u+1}, \\
\sqrt{\beta}h^{d_\sigma} &\leq Ch^{d_u+1}.
\end{aligned}$$

2. Then use corollary 2.3.1 to decide regularization parameters γ, ρ, and δ.
3. Use (2.3.23) to choose ξ and θ if boundary condition (2.1.3) holds.

Note that the parameters chosen according to these rules will make all the terms in $\mathcal{J}_h(\mathbf{w}_h)$ to be the same power of h, i.e., all the terms in the cost functional are balanced in the involved finite element spaces. The authors in [44], [45] also used similar balancing ideas in the FOSLS method for forward problems, but balance in their work was realized through the cost functional in the continuous case.

The error estimates also suggest the choice of the finite element spaces. Most of the finite element spaces in one dimension can satisfy the conditions required by Lemma 2.3.1. However, the spaces used for a_h, u_h, and σ_h should be selected in a collective way so that the rate of convergence in the approximation to the PI problem can be optimal from the point of view of the finite element space used for a_h. This can be seen in the next corollary.

Corollary 2.5.1 *Assume that all the conditions required by theorem 2.3.1 are satisfied, and the finite element spaces are chosen such that,*

$$r_a = 1, r_u = 1, r_\sigma = 1,$$
$$d_a = 1, l = 1, d_u = 2, d_\sigma = 2.$$

Then choosing the regularization parameters according to the above rules, we have the following estimations,

$$\begin{aligned} \|(\tilde{a} - a_h^*)\tilde{u}'\|_0 &\leq C(\frac{\epsilon}{h} + h^2), \text{ for the boundary condition (2.1.3)} , \\ \|((\tilde{a} - a_h^*)\tilde{u}')'\|_{-1} &\leq C(\frac{\epsilon}{h} + h^2), \text{ for the boundary condition (2.1.2)} , \end{aligned} \tag{2.5.41}$$

Proof. It is well known, see [37], that the involved finite element spaces satisfy the conditions of Lemma 2.3.1. According to the approximation capability of these finite element spaces and the rules above, we should choose

$$\alpha = Ch^2, \beta = Ch^2,$$

and other parameters accordingly. Then the two estimates follow from Corollary 2.3.2 and 2.3.3, respectively.

For the finite element spaces configured in this corollary, the choice of the regularization parameters such that $\alpha = Ch^2$, $\beta = Ch^2$ seems to be the optimal since the error bound in the approximation for the PI problem has reached the best possible approximation capability of the finite element space $A_h = S^1_{h,0}$,

$$\|(\tilde{a} - a_h^*)\tilde{u}'\|_0 \leq Ch^2,$$

in a weighted L^2-norm sense when the data is perfect ($\epsilon = 0$). On the other hand, if we use $A_h = S^2_{h,0}$ for a_h, but use the same finite element spaces for u_h and σ_h, then the same argument as that in the previous corollary leads to the following choice of the regularization parameters,

$$\alpha = Ch^2, \beta = Ch^2,$$

and consequently, we should have

$$\|(\tilde{a} - a_h^*)\tilde{u}'\|_0 \leq Ch^2,$$

when the data is perfect ($\epsilon = 0$). In this case, the error estimate is not the optimal in a weighted L^2-norm sense for the related finite element space $A_h = S^2_{h,0}$.

2.6 Some implementation issues

The evaluation of $\mathcal{J}_h(\mathbf{w}_h)$ can be done by a standard numerical quadrature formula together with the interpolation over the involved finite element spaces. Here we present an alternative approach which can reduce the time of evaluating $\mathcal{J}_h(\mathbf{w}_h)$ during a minimization iteration. Let $\{\phi_i(x)\}_{i=1,2,\ldots,N_a}$, $\{\psi_i(x)\}_{i=1,2,\ldots,N_u}$, and $\{\eta_i(x)\}_{i=1,2,\ldots,N_\sigma}$ be the basses for the finite element spaces A_h, U_h, and Σ_h, respectively, and let

$$\begin{aligned} A_{\phi lm} &= ((\phi_i^{(l)}, \phi_j^{(m)}))_{i,j=1,\ldots,N_a} \\ A_{\psi lm} &= ((\psi_i^{(l)}, \psi_j^{(m)}))_{i,j=1,\ldots,N_u} \\ A_{\eta lm} &= ((\eta_i^{(l)}, \eta_j^{(m)}))_{i,j=1,\ldots,N_\sigma} \end{aligned}$$

with $l, m = 0, 1$. The matrices for the regularization terms are formed according to the regularization Sobolev spaces we use. For instance, if the $H^1(0,1)$ norm is used for the regularization of a_h, then $r_a = 1$, and we set

$$A_{r_a} = A_{\phi 00} + A_{\phi 11},$$

Similar definitions hold for A_{r_u} and A_{r_σ}. Then the cost functional (2.1.8) becomes,

$$\begin{aligned} \mathcal{J}_h(\mathbf{w}_h) = \; & \vec{u}_h^T A_{\psi 00} \vec{u}_h - 2\vec{u}_h^T \cdot \vec{z}_h + \|z\|_0^2 + \alpha \textstyle\int_0^1 (a_h u_h' - \sigma_h)^2 \, dx + \\ & \beta(\vec{\sigma}_h^T A_{\eta 11} \vec{\sigma} + 2\vec{\sigma}_h^T \cdot \vec{F}_h + \|f\|_0^2) + \gamma \vec{a}_h^T A_{r_a} \vec{a}_h + \rho \vec{u}_h^T A_{r_u} \vec{u}_h + \delta \vec{\sigma}_h^T A_{r_\sigma} \vec{\sigma}, \end{aligned} \tag{2.6.42}$$

where $\mathbf{w}_h = (a_h, u_h, \sigma_h)$ and,

$$a_h = \textstyle\sum_{i=1}^{N_a} a_h^i \phi_i(x), \quad u_h = \sum_{i=1}^{N_u} u_h^i \psi_i(x), \quad \sigma_h = \sum_{i=1}^{N_\sigma} \sigma_h^i \eta_i(x),$$

$$\vec{a}_h = \begin{pmatrix} a_h^1 \\ a_h^2 \\ \vdots \\ a_h^{N_a} \end{pmatrix}, \vec{u}_h = \begin{pmatrix} u_h^1 \\ u_h^2 \\ \vdots \\ u_h^{N_u} \end{pmatrix}, \vec{\sigma}_h = \begin{pmatrix} \sigma_h^1 \\ \sigma_h^2 \\ \vdots \\ \sigma_h^{N_\sigma} \end{pmatrix},$$

and

$$\vec{z}_h = \begin{pmatrix} (z, \psi_1) \\ (z, \psi_2) \\ \vdots \\ (z, \psi_{N_u}) \end{pmatrix}, \vec{F}_h = \begin{pmatrix} (f, \eta_1') \\ (f, \eta_2') \\ \vdots \\ (f, \eta_{N_\sigma}') \end{pmatrix},$$

We also need to compute the gradient of $\mathcal{J}_h(\mathbf{w}_h)$ if a gradient type (such as the steepest descent) algorithm is used to find its minimizer. The simple structure of the cost functional leads to the following formula for the gradient of $\mathcal{J}_h(\mathbf{w}_h)$:

$$\nabla \mathcal{J}_h(\mathbf{w}_h) = \begin{pmatrix} \nabla_{\vec{a}_h} \mathcal{J}_h(\mathbf{w}_h) \\ \nabla_{\vec{u}_h} \mathcal{J}_h(\mathbf{w}_h) \\ \nabla_{\vec{\sigma}_h} \mathcal{J}_h(\mathbf{w}_h) \end{pmatrix},$$

with

$$\nabla_{\vec{a}_h}\mathcal{J}_h(\mathbf{w}_h) = 2\alpha \begin{pmatrix} (u_h'(a_h u_h' - \sigma_h), \phi_1) \\ (u_h'(a_h u_h' - \sigma_h), \phi_2) \\ \vdots \\ (u_h'(a_h u_h' - \sigma_h), \phi_{N_a}) \end{pmatrix} + 2\gamma A_{r_a}\vec{a}_h,$$

$$\nabla_{\vec{u}_h}\mathcal{J}_h(\mathbf{w}_h) = 2 \begin{pmatrix} (u_h - z, \psi_1) \\ (u_h - z, \psi_2) \\ \vdots \\ (u_h - z, \psi_{N_u}) \end{pmatrix} + 2\alpha \begin{pmatrix} (a_h(a_h u_h' - \sigma_h), \psi_1') \\ (a_h(a_h u_h' - \sigma_h), \psi_2') \\ \vdots \\ (a_h(a_h u_h' - \sigma_h), \psi_{N_u}') \end{pmatrix} + 2\rho A_{r_u}\vec{u}_h,$$

$$\nabla_{\vec{\sigma}_h}\mathcal{J}_h(\mathbf{w}_h) = -2\alpha \begin{pmatrix} (a_h u_h' - \sigma_h, \eta_1) \\ (a_h u_h' - \sigma_h, \eta_2) \\ \vdots \\ (a_h u_h' - \sigma_h, \eta_{N_\sigma}) \end{pmatrix} + 2\beta \begin{pmatrix} (\sigma_h' + f, \eta_1') \\ (\sigma_h' + f, \eta_2') \\ \vdots \\ (\sigma_h' + f, \eta_{N_\sigma}') \end{pmatrix} + 2\delta A_{r_\sigma}\vec{\sigma}_h,$$

Notice that all the matrices mentioned above can be assembled by a standard finite element program at the same time in a minimization procedure. In each iteration, $\mathcal{J}_h(\mathbf{w}_h)$ can be formed by a numerical integration and several matrix(sparse)-vector multiplications. The gradient $\nabla\mathcal{J}_h(\mathbf{w}_h)$ can be formed by calling the load vector assembling routines in a standard finite element program only four times plus several matrix(sparse)-vector multiplications. Therefore , the method here is rather less time consuming even though for the same discretization parameter h the cost functional have three independent variables compared to only one in the output least squares method. Moreover, the method has a convergence feature stated in the following

Theorem 2.6.1 *Lte $\{\mathbf{w}_h^i\}_1^\infty$ be a sequence generated by applying the steepest descent method to $\mathcal{J}_h(\mathbf{w}_h)$. Then $\{\mathbf{w}_h^i\}_1^\infty$ must have an accumulation point, and every accumulation point $\mathbf{w}_h^*$ of $\{\mathbf{w}_h^i\}_1^\infty$ satisfies*

$$\nabla\mathcal{J}_h(\mathbf{w}_h^*) = 0.$$

Proof. This result is the consequence of the fact that $\mathcal{J}_h(\mathbf{w}_h)$ is a C^1 functional and the level set

$$\{\mathbf{w}_h | \mathcal{J}_h(\mathbf{w}_h) \leq \mathcal{J}_h(\mathbf{w}_h^1)\}$$

is bounded.

If the Hessian matrix of the cost functional is available, an algorithm faster than the steepest descent method might be used to compute a minimizer. While it is rather difficult to form the Hessian for the cost functional in the output least squares method, we can show that the cost functional here has the following simple formula for its Hessian matrix;

$$\nabla^2 \mathcal{J}_h(a,u,\sigma) = \begin{pmatrix} \nabla^2_{aa}\mathcal{J}_h & \nabla^2_{au}\mathcal{J}_h & \nabla^2_{a\sigma}\mathcal{J}_h \\ \nabla^2_{ua}\mathcal{J}_h & \nabla^2_{uu}\mathcal{J}_h & \nabla^2_{u\sigma}\mathcal{J}_h \\ \nabla^2_{\sigma a}\mathcal{J}_h & \nabla^2_{\sigma u}\mathcal{J}_h & \nabla^2_{\sigma\sigma}\mathcal{J}_h \end{pmatrix}$$

where we have the following symmetry for the Hessian matrix;

$$\nabla^2_{ua}\mathcal{J}_h = (\nabla^2_{au}\mathcal{J}_h)^T, \nabla^2_{\sigma a}\mathcal{J}_h = (\nabla^2_{a\sigma}\mathcal{J}_h)^T, \nabla^2_{\sigma u}\mathcal{J}_h = (\nabla^2_{u\sigma}\mathcal{J}_h)^T,$$

and then, the Hessian of $\mathcal{J}_h$ reduces to

$$\nabla^2 \mathcal{J}_h(a,u,\sigma) = \begin{pmatrix} \nabla^2_{aa}\mathcal{J}_h & (\nabla^2_{ua}\mathcal{J}_h)^T & (\nabla^2_{\sigma a}\mathcal{J}_h)^T \\ \nabla^2_{ua}\mathcal{J}_h & \nabla^2_{uu}\mathcal{J}_h & (\nabla^2_{\sigma u}\mathcal{J}_h)^T \\ \nabla^2_{\sigma a}\mathcal{J}_h & \nabla^2_{\sigma u}\mathcal{J}_h & \nabla^2_{\sigma\sigma}\mathcal{J}_h \end{pmatrix}$$

each submatrix above is defined by

$$\begin{aligned}
\nabla^2_{aa}\mathcal{J}_h &= 2\alpha(((u'_h)^2\phi_k,\phi_l))_{k=1,l=1}^{N_a,N_a} + 2\gamma A_{r_a}, \\
\nabla^2_{au}\mathcal{J}_h &= 2\alpha(((2a_h u'_h - \sigma_h)\phi_k,\psi'_l))_{k=1,l=1}^{N_a,N_u} \\
\nabla^2_{a\sigma}\mathcal{J}_h &= -2\alpha((u'_h\phi_k,\eta_l))_{k=1,l=1}^{N_a,N_\sigma} \\
\nabla^2_{uu}\mathcal{J}_h &= 2\alpha((a_h^2\psi'_k,\psi'_l))_{k=1,l=1}^{N_u,N_u} + 2A_{\psi 00} + 2\rho A_{r_u}, \\
\nabla^2_{u\sigma}\mathcal{J}_h &= -2\alpha((a_h\psi'_h,\eta_l))_{k=1,l=1}^{N_u,N_\sigma} \\
\nabla^2_{\sigma\sigma}\mathcal{J}_h &= 2\alpha A_{\eta 00} + 2\beta A_{\eta 11} + 2\delta A_{r_\sigma}
\end{aligned}$$

Note that the Hessian $\nabla^2\mathcal{J}_h(a,u,\sigma)$ is a sparse matrix which can be formed efficiently by any standard finite element program. In fact, all the matrices involved in the Hessian can be formed by the matrix assembling subroutines in a standard finite element program and the rest of the matrices can be prepared before the minimization iteration is carried out. Due to this reason, a minimization algorithm utilizing the Hessian seems to be a preferable choice. We do not use Newton's method in this implementation since its local convergence property requires a good initial guess for the minimizer. Also, the straightforward Newton's method may fail if the Hessian turns out to be singular or indefinite at one of the iteration points. There have been algorithms developed to overcome various shortcomings of the straightforward Newton's methods, (see [7], [18], and [33]). The following modified Newton's algorithm is one of them:

Initialization Step
Choose a scalar $\epsilon > 0$ as a tolerance, and supposed that $\mathbf{w}_h^k$ is given. To generate $\mathbf{w}_h^{k+1}$ go

to the main step.

Main Step
1. Compute $\nabla J_h(\mathbf{w}_h^k)$ and $\nabla^2\mathcal{J}_h(\mathbf{w}_h^k)$. If $\nabla\mathcal{J}_h(\mathbf{w}_h^k) = 0$, then stop and output $\mathbf{w}_h^k$ as the solution. Otherwise, go to step 2.
2. If $\nabla^2\mathcal{J}_h(\mathbf{w}_h^k)$ is singular, then let

$$\mathbf{w}_h^{k+1} = \mathbf{w}_h^k - \tilde{r}\nabla\mathcal{J}_h(\mathbf{w}_h^k),$$

where $\tilde{r}$ minimizes the function $\phi(s) = \mathcal{J}_h(\mathbf{w}_h^k - s\nabla\mathcal{J}_h(\mathbf{w}_h^k))$, for all $s \geq 0$. Otherwise, go to step 3.
3. Let

$$\tilde{\mathbf{p}} = -(\nabla^2\mathcal{J}_h(\mathbf{w}_h^k))^{-1}\nabla\mathcal{J}_h(\mathbf{w}_h^k),$$

and use one of the following three ways to form $\mathbf{w}_h^{k+1}$:
(a). If $|\tilde{\mathbf{p}}^t\nabla\mathcal{J}_h(\mathbf{w}_h^k)| \leq \epsilon\|\tilde{\mathbf{p}}\|\nabla\mathcal{J}_h(\mathbf{w}_h^k)\|$, where ϵ is a constant chosen according to the accuracy requirement, then let

$$\mathbf{w}_h^{k+1} = \mathbf{w}_h^k - \tilde{r}\nabla\mathcal{J}_h(\mathbf{w}_h^k),$$

where $\tilde{r}$ minimizes the function $\phi(s) = \mathcal{J}_h(\mathbf{w}_h^k - s\nabla\mathcal{J}_h(\mathbf{w}_h^k))$, for all $s \geq 0$.
(b). If $\tilde{\mathbf{p}}^t\nabla\mathcal{J}_h(\mathbf{w}_h^k) > \epsilon\|\tilde{\mathbf{p}}\|\nabla\mathcal{J}_h(\mathbf{w}_h^k)\|$, then let

$$\mathbf{w}_h^{k+1} = \mathbf{w}_h^k - \tilde{r}\tilde{\mathbf{p}},$$

where $\tilde{r}$ minimizes the function $\phi(s) = \mathcal{J}_h(\mathbf{w}_h^k - s\tilde{p})$, for all $s \geq 0$.
(c). If $\tilde{\mathbf{p}}^t\nabla\mathcal{J}_h(\mathbf{w}_h^k) < -\epsilon\|\tilde{\mathbf{p}}\|\nabla\mathcal{J}_h(\mathbf{w}_h^k)\|$, then let

$$\mathbf{w}_h^{k+1} = \mathbf{w}_h^k + \tilde{r}\tilde{\mathbf{p}},$$

where $\tilde{r}$ minimizes the function $\phi(s) = \mathcal{J}_h(\mathbf{w}_h^k + s\tilde{\mathbf{p}})$, for all $s \geq 0$.

The convergence feature of the above algorithm is given in the following

Theorem 2.6.2 *Let $\{\mathbf{w}_h^i\}_1^\infty$ be a sequence generated by applying the above modified Newton algorithm to $\mathcal{J}_h(\mathbf{w}_h)$. Then $\{\mathbf{w}_h^i\}_1^\infty$ must have an accumulation point, and every accumulation point $\mathbf{w}_h^*$ of $\{\mathbf{w}_h^i\}_1^\infty$ satisfies*

$$\nabla\mathcal{J}_h(\mathbf{w}_h^*) = 0.$$

Proof The result follows from Zangwill's convergence theorem (see [7], [43]) plus the fact that the level set

$$\{\mathbf{w}_h|\mathcal{J}_h(\mathbf{w}_h) \leq \mathcal{J}_h(\mathbf{w}_h^1)\}$$

is bounded.

Like many nonlinear problems [4], [5], [34], the discretized cost functional here seems to have a mesh independent property, i.e., the number of iterations for Newton's method to converge within a given tolerance is nearly independent of the mesh size. A multi-level Newton iteration can be derived to take advantage of this property so that a good approximation can be obtained quickly from a rather crude initial guess. The basic idea here is to find a good initial guess for a fine grid computation by interpolating a minimizer generated in a coarse grid to the fine grid.

2.7 Numerical examples

We present some numerical examples which display interesting features of the finite element method in this chapter. To describe the accuracy of a finite element solution $a_h(x)$, we let $\tilde{a}(x)$ and $\tilde{u}(x)$ be the functions satisfying the differential equation (2.1.1) and the boundary condition (2.1.2) or (2.1.3) and let

$$E^0(a_h) = (\int_0^1 (a_h(x) - \tilde{a}(x))^2 (\tilde{u}'(x))^2 \, dx)^{1/2}, \tag{2.7.43}$$

$$E^*(a_h) = \|p'\|_0, \tag{2.7.44}$$

where the function $p(x)$ is the solution of

$$(p', v') = ((\tilde{a} - a_h)\tilde{u}', v'), \text{ for all } v \in \tilde{H}^1(0,1). \tag{2.7.45}$$

Now we prove that $\|p'\|_0$ is equivalent to $\|((\tilde{a} - a_h)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}$. There exists a one to one correspondence between $\tilde{H}^1(0,1)$ and $L^2(\Omega)$

$$v \in \tilde{H}^1(0,1) \to v' = \phi \in L^2(\Omega)$$

since for each ϕ in $L^2(\Omega)$ there exists a function $v \in \tilde{H}^1(0,1)$ defined by,

$$v(x) = \int_0^x \phi(s) \, ds$$

and by Poincare's inequality we get

$$\|v\|_1 \cong \|\phi\|_0, \tag{2.7.46}$$

dividing both sides of (2.7.45) by $\|v'\|_0$ and taking the supremum over all $v \in \tilde{H}^1(0,1)$ and by the equivalence (2.7.46) we have

$$\sup_{0 \neq \phi \in L^2} \frac{|(p', \phi)|}{\|\phi\|_0} \cong \sup_{0 \neq v \in \tilde{H}^1(0,1)} \frac{|(p', v')|}{\|v'\|_0} =$$
$$\sup_{0 \neq v \in \tilde{H}^1(0,1)} \frac{|((\tilde{a} - a_h)\tilde{u}', v')|}{\|v'\|_0} \cong \sup_{0 \neq v \in \tilde{H}^1(0,1)} \frac{|(((\tilde{a} - a_h)\tilde{u}')', v)|}{\|v'\|_0}$$

The left most side is $\|p'\|_0$ and the right most side is $\|((\tilde{a} - a_h)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}$ and the equivalence between the two norms is proved.

Obviously, $\tilde{a}(x)$ can be considered as the exact solution of the PI problem, $E^0(a_h)$ and $E^*(a_h)$ can be considered as a weighted $L^2(0,1)$ and $H^{-1}(0,1)$ norms of the error in a finite element solution a_h, respectively. All the numerical results presented in this section were carried out with the finite element spaces such that

$$d_a = 1, l = 1, d_u = 2, d_\sigma = 2,$$

and they were obtained by a two-level or three-level Newton method with initial guesses generated by perturbing the exact solutions with a random variable of magnitude 5.

Let us not forget that in all our calculations, we are calculating not only the coefficient function $a(x)$, which is one of our unknowns, but also the state function $u(x)$, as well as the flux $\sigma(x)$. The weighted L^2 and H^{-1}-error for various values of h are shown in the data tables below along with some pictures.

Example 1: The first problem consists in identifying $a(x) = 1 + \frac{1}{2}\cos(\pi x)$ from the data function $z(x) = \sin(\pi x) + \pi x$ in the boundary value problem,

$$\begin{aligned} -(au')' &= f \\ u(0) = u'(1) &= 0 \end{aligned} \tag{2.7.47}$$

We assume that the data is perfect, i.e., $z(x) = u(x)$ and $f(x) = \frac{1}{2}\pi^2(3\sin(\pi x) + \sin(2\pi x))$, so that $u(x)$ is the exact solution of (2.7.47) with the given $a(x)$. The errors in the weighted L^2 norm (2.7.43) for the finite element solutions with various h are given in Table 2.1.

To see the rate of convergence, we assume that the error obeys

$$E^0(a_h) = Ch^r.$$

Then the data in Table 2.1 satisfies

$$E^0(a_h) = Ch^{2.00131551208585},$$

which corroborates the result in Corollary 2.3.2.

In all the numerical examples below, we consider the PI problem for the partial differential equation (2.1.1) with boundary condition (2.1.2),

Example 2: In this example, we identify $a(x) = 1 + \frac{1}{2}\sin(2\pi x)$ from the data function $z(x) = \cos(4\pi x) - 1$ in the boundary value problem (2.1.1) with boundary condition (2.1.2). We also assume that the data is perfect and $f(x) = 2\pi^2(8\cos(4\pi x) - \sin(2\pi x) + 3\sin(6\pi x))$.

Table 2.1			
h	weighted L^2-error	h	weighted L^2-error
1/30	$1.7662646 * 10^{-3}$	1/110	$1.1686027 * 10^{-4}$
1/40	$9.2132820 * 10^{-4}$	1/120	$9.8173945 * 10^{-5}$
1/50	$5.7633162 * 10^{-4}$	1/130	$8.3642902 * 10^{-5}$
1/60	$3.9649014 * 10^{-4}$	1/140	$7.2121541 * 10^{-5}$
1/70	$2.8998400 * 10^{-4}$	1/150	$6.2827951 * 10^{-5}$
1/80	$2.2148973 * 10^{-4}$	1/160	$5.5225451 * 10^{-5}$
1/90	$1.7477196 * 10^{-4}$	1/170	$4.8921647 * 10^{-5}$
1/100	$1.4145962 * 10^{-4}$	1/180	$4.3644275 * 10^{-5}$

Table 2.1: Errors in finite element solutions with various h for the PI problem with boundary condition (2.1.3)

The errors of the finite element solutions in both the weighted L^2 and H^{-1} norms are listed in Table 2.2 and 2.3, respectively. The data listed in Table 2.2 indicates that the actual error in this example obeys

$$E^0(a_h) = Ch^{2.09368629170294},$$

This implies that the method can produce an approximation for the PI problem with boundary condition (2.1.2) within the expected accuracy in the weighted L^2 norm even though no theoretical error estimate has been established yet. The data listed in Table 2.3

Table 2.2			
h	weighted L^2-error	h	weighted L^2-error
1/30	$7.4469581 * 10^{-2}$	1/110	$1.0693435 * 10^{-3}$
1/40	$2.0250231 * 10^{-2}$	1/120	$8.8690251 * 10^{-4}$
1/50	$8.0845654 * 10^{-3}$	1/130	$7.5256289 * 10^{-4}$
1/60	$4.5401557 * 10^{-3}$	1/140	$6.4417956 * 10^{-4}$
1/70	$3.0140973 * 10^{-3}$	1/150	$5.5994340 * 10^{-4}$
1/80	$2.1546992 * 10^{-3}$	1/160	$4.8990444 * 10^{-4}$
1/90	$1.6563471 * 10^{-3}$	1/170	$4.3340711 * 10^{-4}$
1/100	$1.3063933 * 10^{-3}$	1/180	$3.8541150 * 10^{-4}$

Table 2.2: Errors in finite element solutions with various h for the PI problem with boundary condition (2.1.2)

indicates that the actual error obeys

$$E^*(a_h) = Ch^{2.62812273008935},$$

which is within the prediction of Corollary 2.3.3.

Table 2.3			
h	H^{-1}-error	h	H^{-1}-error
1/30	$3.9895322 * 10^{-4}$	1/110	$4.4503722 * 10^{-6}$
1/40	$9.8899125 * 10^{-5}$	1/120	$2.5717919 * 10^{-6}$
1/50	$5.2377479 * 10^{-5}$	1/130	$2.9168802 * 10^{-6}$
1/60	$2.0951958 * 10^{-5}$	1/140	$1.7530075 * 10^{-6}$
1/70	$1.6400521 * 10^{-5}$	1/150	$2.0694999 * 10^{-6}$
1/80	$8.0347477 * 10^{-6}$	1/160	$1.2793015 * 10^{-6}$
1/90	$7.6810592 * 10^{-6}$	1/170	$1.5503353 * 10^{-6}$
1/100	$4.1775909 * 10^{-6}$	1/180	$9.7356926 * 10^{-7}$

Table 2.3: Errors in finite element solutions with various h for the PI problem with boundary condition (2.1.3)

Notice that the first derivative of $u(x) = \cos(4\pi x) - 1$ vanishes at the points in the set $S = \{0, 1/4, 1/2, 3/4, 1\}$. Even though the finite element solution converges in a weighted norm, the approximation is less accurate at these singular points. However, the smaller the h, the better the accuracy in the finite element solution even at these singular points, as indicated by the plots in Figure 2.1.

Example 3: In this example, we identify the discontinuous coefficient function,

$$a(x) = \begin{cases} 1, & \text{for } 0 \le x \le \frac{1}{2} \\ 1.5, & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

from the data function

$$z(x) = \begin{cases} xe^{-2x}, & \text{for } 0 \le x \le \frac{1}{2} \\ (1-x)e^{(2x-2)}, & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

in the boundary value problem (2.1.1), (2.1.2). We assume again that the data is perfect and

$$f(x) = \begin{cases} 4(1-x)e^{(-2x)}, & \text{for } 0 \le x \le \frac{1}{2} \\ 6xe^{2(x-1)}, & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

is chosen so that $u(x) = z(x)$ is the exact solution of (2.1.1), (2.1.2) with $a(x)$ defined above. Notice that the exact solution $a(x)$ of the PI problem has a jump discontinuity at $x = 1/2$

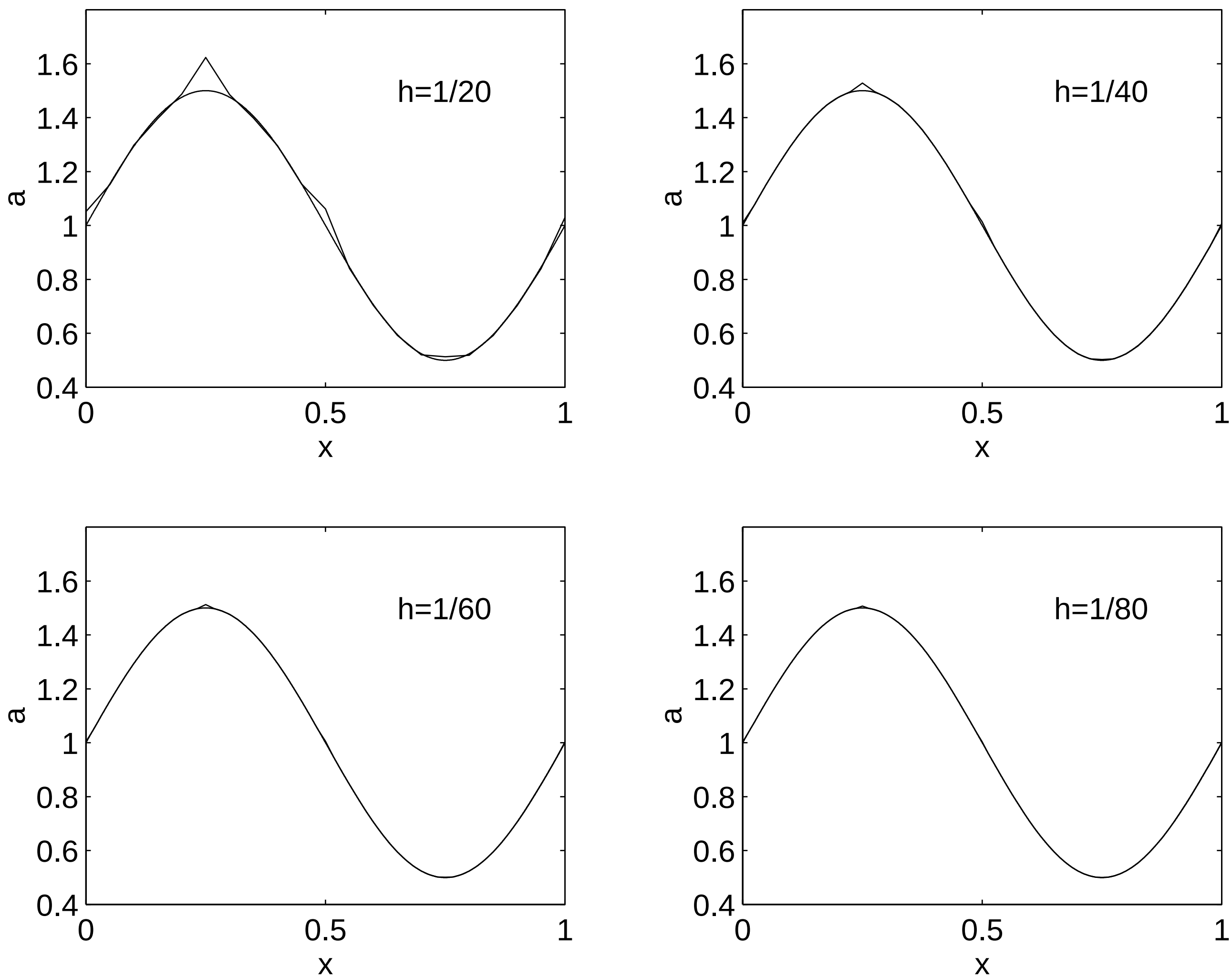

Figure 2.1: The approximation is less accurate around the singularity points of the state u, but the accuracy improves as h becomes smaller

which coincides with the only one singularity of the state $u(x)$. In this experiment, we have chosen $a(x)$ and $u(x)$ such that both functions satisfy the necessary condition below,

$$a(1/2-)u'(1/2-) = a(1/2+)u'(1/2+),$$

As we can see in Figure 2.2, the approximation $a_h(x)$ becomes better as h tends to 0, which means that this singularity of $u(x)$, right at the jump discontinuity, does not spoil the identification of $a(x)$. The approximation $a_h(x)$ gets even steeper at $x = 1/2$ as h gets smaller trying to match with the jump. In this example we get a very good qualitative behavior of the true coefficient function $a(x)$.

Example 4: In this example, we identify the discontinuous coefficient function,

$$a(x) = \begin{cases} 1, & \text{for } 0 \le x \le \frac{1}{2} \\ 1.5, & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

for the boundary value proble (2.1.1), (2.1.2) from the data function

$$z(x) = \begin{cases} x\sin(b_1 x + c_1), & \text{for } 0 \le x \le \frac{1}{2} \\ (1-x)\sin(b_2 x + c_2), & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

where b_1, b_2, c_1 and c_2, are defined as follows, $b_1 = 2\pi$, $b_2 = 4/3(5/2+\pi)$, $c_1 = -3/4\pi$ and $c_2 = -5/3 - 5/12\pi$. We assume that the data is perfect, i.e., $z(x) = u(x)$ and

$$f(x) = \begin{cases} -(2b_1\cos(b_1 x + c_1) - b_1^2 x\sin(b_1 x + c_1)), & \text{for } 0 \le x \le \frac{1}{2} \\ 3b_2\cos(b_2 x + c_2) + 3/2b_2^2(1-x)\sin(b_2 x + c_2) & \text{for } \frac{1}{2} < x \le 1 \end{cases}$$

is chosen so that $u(x)$ is the exact solution to the boundary value problem with the given $a(x)$.

In this experiment, $a(x)$ and $u(x)$ are chosen such that both functions satisfy the necessary condition

$$a(1/2-)u'(1/2-) = a(1/2+)u'(1/2+).$$

As we can see in Figure 2.3, the approximation $a_h(x)$ has a jump at the discontinuity point, as before. The smaller the h, the better the approximation. The numerical results have an oscillation around the discontinuity point of $a(x)$, but the oscillation region seems to shrink as h gets smaller, though the error in the approximation (in the supremum norm) does not reduce with h. We observed again that at the singularities of the state $u = u(x)$, which are $x = 0.22$, $x = 0.57$, $x = 0.9$, the approximation $a_h(x)$ has a bump that reduces in width but does not disappear completely as h approaches 0.

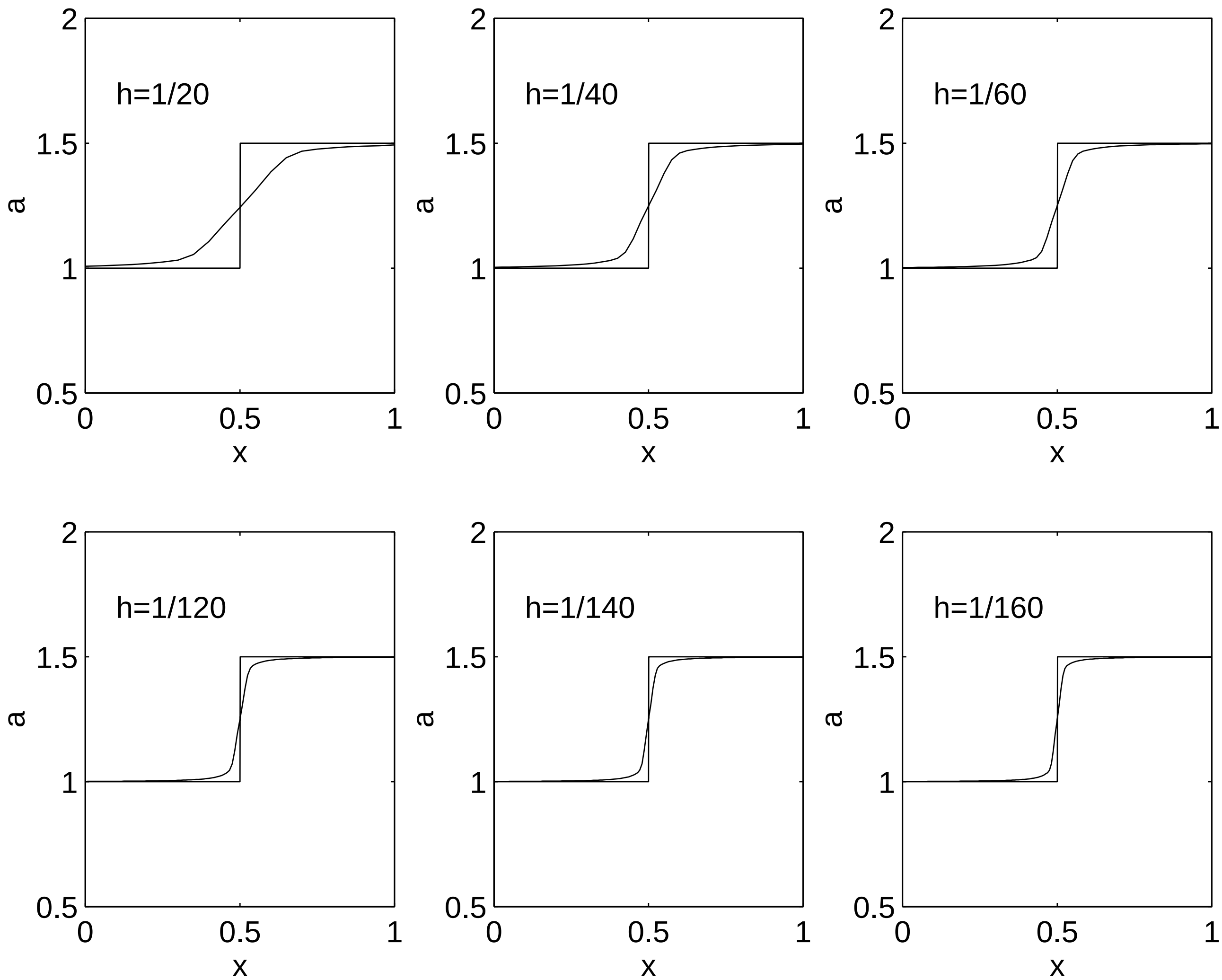

Figure 2.2: Numerical solutions compared against the exact solution with a jump discontinuity

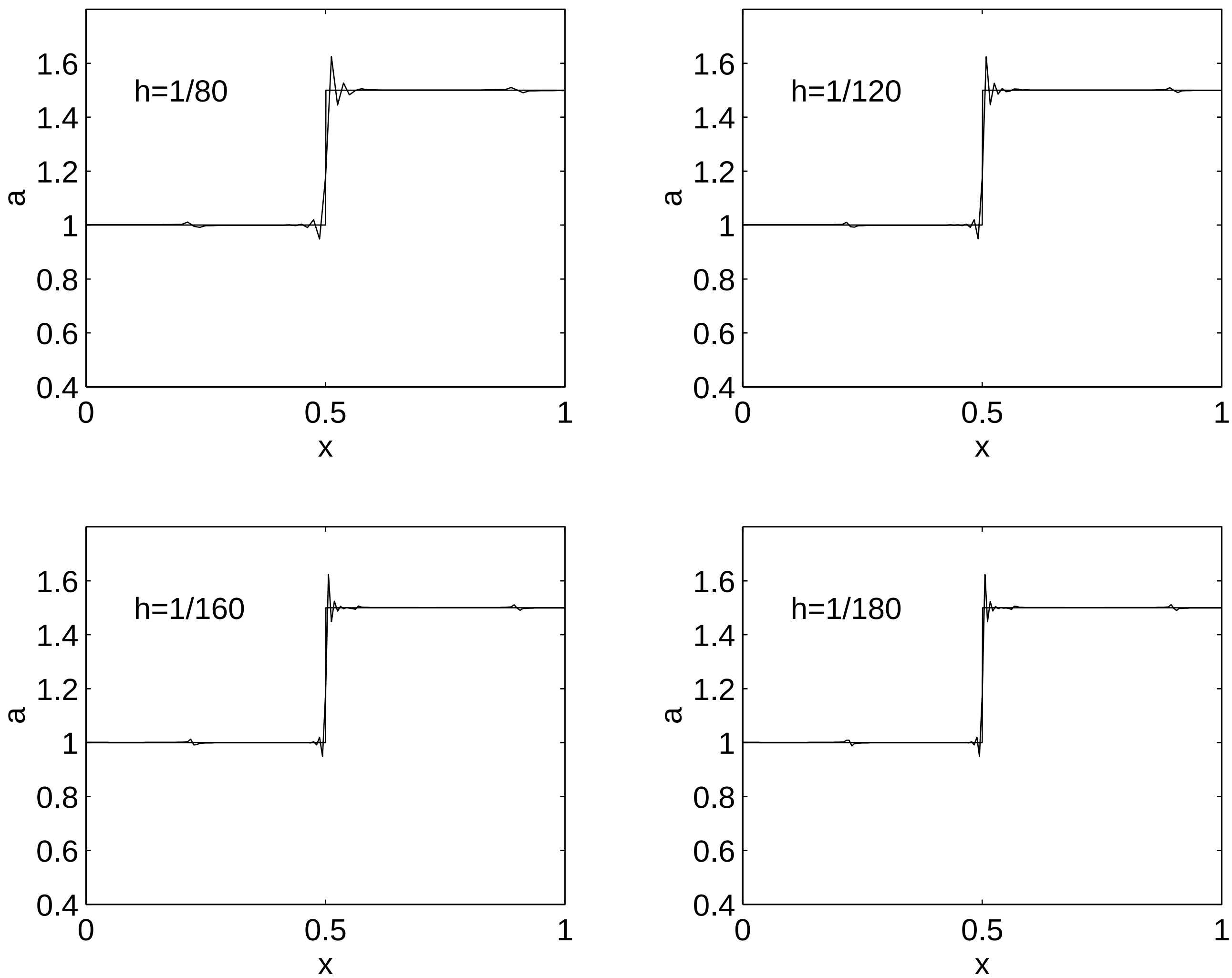

Figure 2.3: Numerical solutions plotted against the exact solution. There is an oscillation around the jump discontinuity whose width shrinks with h. The approximations have a bump at the singularities of the state that reduces in width but do not disappear as h approaches 0.

Example 5: In this example, we identify the coefficient

$$a(x) = -x(x-1) + 1,$$

for the boundary value problem (2.1.1), (2.1.2) from the data function

$$z(x) = \cos(4\pi x) - 1,$$

We assume again that the data is perfect ($\epsilon = 0$) and

$$f(x) = 16\pi^2(1 + x - x^2)\cos(4\pi x) + 4\pi(1-2x)\sin(4\pi x))$$

is chosen so that $u(x) = z(x)$ is the exact solution of (2.1.1), (2.1.2) with the given $a(x)$. In this numerical example we want to focus in the features of the numerical approximation when the state $u(x)$ has many singularities. The singularities of the state are at $x = 0, 1/4, 1/2, 3/4, 1$. Even though the state has more singularities in this example than in the previous ones, the approximation $a_h(x)$ has still a bump at the singularity points, which shrinks in size but does not vanish as h gets smaller. We have displayed our results in Figure 2.4.

The next example was taken from the joint paper [31].

Example 6: The error estimates in the previous section indicates that the error bounds depend on the error in the data linearly for a fixed h. To observe this, we carried out computations for the PI problem for the boundary value problem (2.1.1), (2.1.2) with the following data

$$z(x) = \sin(\pi x) + mag \times rand(x),$$

where mag is the magnitude of the error in the data, $rand(x)$ is a random function with range between -0.5 and 0.5. The function $f(x)$ was chosen so that $u(x) = \sin(\pi x)$ is the exact solution of the boundary value problem with

$$a(x) = 1 + \frac{1}{2}\cos(x).$$

The weighted L^2 norm errors are listed in Table 2.4 which indicates that for this example we have

$$E^0(a_h) = 3.2369(mag)^{1.0177},$$

which implies that the accuracy in the numerical solutions depends linearly on the magnitude of the error in the data function.

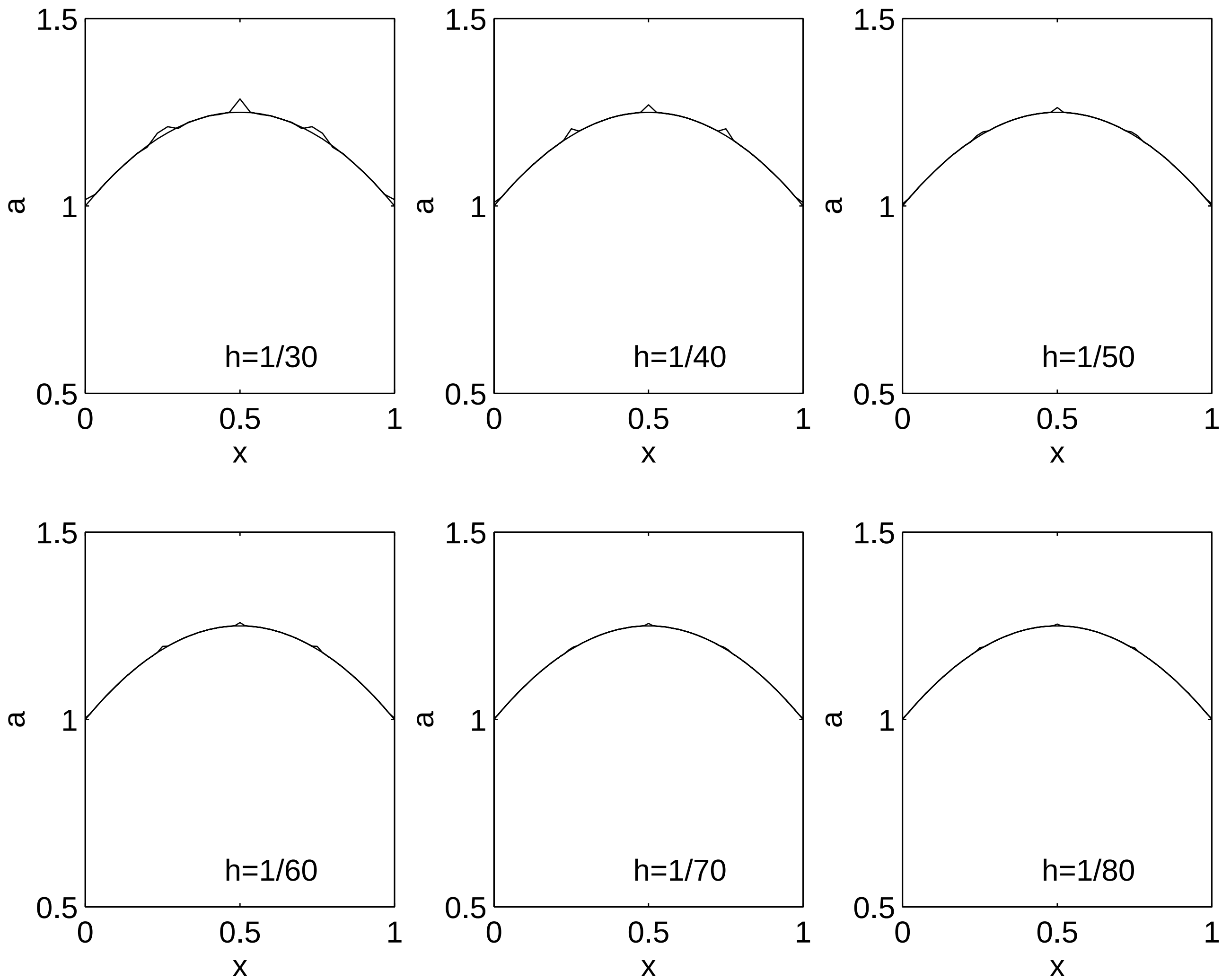

Figure 2.4: Numerical approximation plotted against the exact solution. The finite element solution has a bump at all singularities, as in the previous cases, which reduce in size but do not disappear

Table 2.4	
magnitude	weighted L^2-error
0.01	0.23991473506168
0.005	0.10400176567014
0.0025	0.0587257149189
0.00125	0.03403396156303
0.000625	0.01232753115644

Table 2.4: The error in the numerical solution in weighted L^2 norm become smaller as the data becomes more accurate.

Chapter 3

Parameter identification of a nonlinear two point boundary value problem

3.1 Introduction

In this chapter we are concerned with the identification of the unknown nonlinear coefficient $a(u)$ for the nonlinear elliptic boundary value problem with homogeneous Dirichlet and Dirichlet-Neumann boundary conditions defined below. We present a least-squares mixed finite element method to identify the nonlinear coefficient $a(u)$. We describe the problem here, and in later sections we specify more precisely our assumptions.

Let Ω be the open interval $(0, 1)$ in $I\!R$. We consider the two point boundary value problem with a nonlinear differential equation

$$\mathcal{L}(a)u = -(a(u(x))u'(x))' = f(x) \text{ in } \Omega, \tag{3.1.1}$$

with boundary condition

$$u(0) = u(1) = 0, \tag{3.1.2}$$

or

$$u(0) = u'(1) = 0, \tag{3.1.3}$$

where $f \in L^2(\Omega)$, and $\lambda \leq a(s) \leq \Lambda$ for some positive constants λ and Λ and $s \in I\!R$. The parameter identification (PI) problem is to recover the parameter $a = a(s), s \in I\!R$ within some suitable set of admissible elements, from an L^2 observation $z(x)$ of $u(x)$ such that

$a(u)$ and $u(x)$ satisfy (3.1.1) along with the boundary condition given by either (3.1.2) or (3.1.3).

Equation (3.1.1) corresponds to a steady state solution of the nonlinear parabolic partial differential equation below,

$$\begin{array}{c} \frac{\partial u}{\partial t} - (a(u)u')' = f, \text{ in } \Omega,\, t > 0, \\ \text{plus appropriate boundary conditions} \end{array} \tag{3.1.4}$$

Typically (3.1.4) models heat conduction with a temperature dependent heat conductivity given by the function $a = a(s)$. To describe the method, we let $H^s(0,1)$ be the standard Sobolev space. We use the standard notation to denote its norm and seminorm, as well as the inner product in $L^2(0,1)$. We set

$$\tilde{H}^1(0,1) = \begin{cases} H_0^1(0,1), \text{ if (3.1.2) holds,} \\ \{v \in H^1(0,1) : v(0) = 0\}, \text{ if (3.1.3) holds.} \end{cases}$$

As a matter of fact, our PI problem has two unknowns: $a(s)$ and $u(x)$. In our scheme, we regard the nonlinear parameter $a(u)$ as a spatially varying parameter, by denoting $b(x) = a(u(x))$, and carry out all our numerical implementation in terms of this "linear parameter". Following the standard set up in the mixed finite element method, we introduce the flux variable $\sigma(x) = b(x)u'(x)$ as an additional unknown. These unknowns satisfy the equivalent first order system:

$$\begin{cases} -\sigma' &= f, \\ bu' &= \sigma. \end{cases} \tag{3.1.5}$$

We then try to find an approximation to the solution of the PI problem by minimizing the following cost functional,

$$\mathcal{J}(b,u,\sigma) = \begin{cases} \mathcal{J}_1(b,u,\sigma), \text{ if (3.1.2) holds,} \\ \mathcal{J}_2(b,u,\sigma), \text{ if (3.1.3) holds,} \end{cases} \tag{3.1.6}$$

over the space

$$\mathcal{H} = H^1(0,1) \times \tilde{H}^1(0,1) \times H^1(0,1), \tag{3.1.7}$$

where

$$\begin{aligned} \mathcal{J}_1(b,u,\sigma) &= \|u - z\|_0^2 + \alpha\|\sigma - bu'\|_0^2 + \beta\|\sigma' + f\|_0^2 + \\ &\quad \gamma\|b\|_{r_b}^2 + \rho\|u\|_{r_u}^2 + \delta\|\sigma\|_{r_\sigma}^2, \\ \mathcal{J}_2(b,u,\sigma) &= \mathcal{J}_1(b,u,\sigma) + \xi(\sigma(1))^2 + \theta(b(1)u'(1) - \sigma(1))^2, \end{aligned} \tag{3.1.8}$$

where the integers r_b, r_u, and r_σ depend on the Sobolev spaces in which the regularization takes place. Usually, they take the following values

$$r_b = r_u = r_\sigma = 1.$$

The choice for the regularization parameters $\alpha, \beta, \gamma, \rho, \delta, \xi, \theta$ will be given according to the error estimates. The meaning of each term in the cost functional (3.1.6) is the same as that for Chapter2. Note that if the data is perfect, i.e., there exist $u(x)$ and $a(s)$ satisfying (3.1.1) with boundary condition (3.1.2) or (3.1.3) such that $z(x) = u(x)$, then $(a, u, \sigma)^T$ with $\sigma = a(u)u'$ can make zero the first three terms (and the last two terms if the boundary condition (3.1.3) is required) of $\mathcal{J}(b, u, \sigma)$, and makes $\mathcal{J}(b, u, \sigma)$ small if the regularization parameters are small. Therefore, as it happened in Chapter 2, minimizing the cost functional may give an approximation to a solution of the PI problem.

Our method for this nonlinear case is a combination of the least-squares method and the equation error method, but is formulated in a least-squares mixed approach.

This Chapter is organized as follows: we formulate the identification problem as an optimization problem by discretizing the cost functional, which is minimized in the computational procedure over a discrete space. This is followed by error estimates in the approximation. In the very last section, we present some computational results, which are calculated with the proposed scheme.

3.2 Finite element method

In this section, we formulate a finite element discretization of the cost functional (3.1.6) and discuss the related minimization problem. In our estimations, we regard C as a generic constant, which may vary in different contexts, but is always independent of h, unless specified otherwise. In order to define the finite element spaces, let $\{\Delta_h\}$, with $h \in (0,1)$, be a family of regular, quasi-uniform partitions of $\Omega \subset I\!R^1$. For fixed integers $r \geq 1$, $l \geq 0$, we let $S^r_{h,l}$ be the space defined in (2.2.14), $C^{-1}(\Omega)$ again is interpreted as $L^2(\Omega)$. By $S^{r,0}_{h,l}$, we denote the subspace of functions of $S^r_{h,l}$ which vanish at $x = 0$ and $x = 1$. The bounds (2.2.15), (2.2.16), (2.2.17) hold. $\tilde{H}^{-1}$ stands for the dual space $(\tilde{H}^1(\Omega))^*$ equipped with the norm (2.2.18).

Three different discretization spaces are needed for the computational procedure: U_h for the solution u, B_h for the parameter b and Σ_h for the flux variable σ, we assume that these spaces are:

$$\begin{aligned} B_h &= S^{d_b}_{h,l}, && \text{here } l = 0, 1, \\ U_h &= S^{d_u}_{h,1} \cap \tilde{H}^1(0,1), \\ \Sigma_h &= S^{d_\sigma}_{h,1}, \end{aligned} \tag{3.2.9}$$

where d_b, d_u, d_σ are integers that we will fix later.

Now we are ready to define the numerical method to solve the PI problem: (P_h) seek $b_h^* \in B_h$, $u_h^* \in U_h$, $\sigma_h^* \in \Sigma_h$, such that

$$\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) = \min_{b\in B_h, u\in U_h, \sigma\in\Sigma_h} \mathcal{J}_h(b, u, \sigma) \tag{3.2.10}$$

where $\mathcal{J}_h$ is defined by,

$$\mathcal{J}_h(\mathbf{w}_h) = \begin{cases} \mathcal{J}_{h,1}(\mathbf{w}_h), & \text{if (3.1.2) holds,} \\ \mathcal{J}_{h,2}(\mathbf{w}_h), & \text{if (3.1.3) holds,} \end{cases} \tag{3.2.11}$$

where

$$\begin{aligned} \mathcal{J}_{h,1}(\mathbf{w}_h) &= \|u_h - z\|_0^2 + \alpha\|\sigma_h - b_h u_h'\|_0^2 + \beta\|\sigma_h' + f\|_0^2 + \\ &\quad \gamma\|b_h\|_{r_b}^2 + \rho\|u_h\|_{r_u}^2 + \delta\|\sigma_h\|_{r_\sigma}^2, \\ \mathcal{J}_{h,2}(\mathbf{w}_h) &= \mathcal{J}_{h,1}(\mathbf{w}_h) + \xi(\sigma_h(1))^2 + \theta(b_h(1)u_h'(1) - \sigma_h(1))^2. \end{aligned} \tag{3.2.12}$$

where $\mathbf{w}_h = (b_h, u_h, \sigma_h) \in B_h \times U_h \times \Sigma_h$. The advantage of this variational approach is that we allow all variables (b, u, and σ) to vary in the corresponding discrete subspace without differentiating any kind of error corrupted data z. The weights γ, ρ, and δ will be chosen so that all terms of $\mathcal{J}_h(\mathbf{w}_h)$ are balanced. Another advantage of our numerical approach to recover the nonlinear coefficient is that the same code can be used for both linear and nonlinear cases with no modifications at all.

The next algorithm shows how to construct the nonlinear coefficient $\tilde{a}(s)$, $s \in [\min_\Omega \tilde{u}, \max_\Omega \tilde{u}]$ from $b_h^*(x)$ and $u_h^*(x)$ in the case that $d_a = 1$, $d_u = 2$, and the Dirichlet boundary condition is used in the boundary value problem (3.1.1).

(*Step* 1) After the minimization procedure, we get the following functions,

$$b_h^*(x) = \sum_{i=1}^{N_a} b_i^h \phi_i(x),$$

$$u_h^*(x) = \sum_{i=1}^{N_u} u_i^h \psi_i(x),$$

where $\{\phi_i(x)\}_{i=1}^{N_a}$ are the basis functions for the linear finite element space B_h, and $\{\psi_i(x)\}_{i=1}^{N_u}$ are the basis functions for the quadratic finite element space U_h.

(*Step* 2) Define $\bar{u} = (\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_m)^T$ by arranging $\{0, u_1^h, u_3^h, \cdots, u_{N_u}^h, 0\}$ in an ascending order, repeated entries counted only once.

(*Step* 3) Let $\{\bar{\phi}_i(s)\}_{i=1}^m$ be the basis functions for the linear finite element space defined over the interval

$$[\min_{i=1,\cdots,N_u} u_i^h, \max_{i=1,\cdots,N_u} u_i^h]$$

with the partition $\{\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_m\}$.

(*Step* 4) Form the vector $\bar{b} = (\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_m)^T$ by applying the same arrangement to $(b_1^h, b_2^h, \cdots, b_{N_a}^h)$ as that applied to $\{0, u_1^h, u_3^h, \cdots, u_{N_u}^h, 0\}$ to define $\bar{u}$

(*Step* 5) Define the finite element approximation to $\tilde{a}(\tilde{u})$ as follows,

$$a_h^*(s) = \sum_{i=1}^{m} \bar{b}_i \bar{\phi}_i(s), \;\; s \in [\min_{i=1,\cdots,N_u} u_i^h, \max_{i=1,\cdots,N_u} u_i^h]$$

In the next section, we derive some error estimates which demonstrate that $b_h^*(x)$ is a reasonable approximation to $\tilde{a}(\tilde{u}(x))$ as a function of x.

3.3 Some error estimations for the nonlinear parameter

In this section, we establish some error estimates for an approximate solution to the PI problem generated by the finite element method presented in the previous section. This approach leads to the best choice of the regularization parameters and the choice of the finite element spaces for the optimal convergence rates. Let I_b^h, I_u^h and I_σ^h be the interpolation operators in the finite element spaces B_h, U_h and Σ_h, respectively. The regularization finite element spaces are chosen so that

$$1 \le r_b \le d_b + 1, 1 \le r_u \le d_u + 1, 1 \le r_\sigma \le d_\sigma + 1.$$

Moreover, there exists a constant C such that the interpolation operators satisfy,

$$\|I_b^h b\|_{r_b} \le C\|b\|_{r_b+1}, \|I_u^h u\|_{r_u} \le C\|u\|_{r_u+1}, \|I_\sigma^h \sigma\|_{r_\sigma} \le C\|\sigma\|_{r_\sigma+1},$$

for any $b \in H^{r_b+1}(0,1), u \in H^{r_u+1}(0,1), \sigma \in H^{r_\sigma+1}(0,1)$. We first derive an estimate for a minimizer of $\mathcal{J}_h$.

Lemma 3.3.1 *Let us assume the following hypothesis,*

* *Let $z(x)$ be a distributed L^2 observation of the state $\tilde{u}$ with an observation error of the form*

$$\|\tilde{u} - z\|_0 \le \epsilon.$$

* *Assume that* $\tilde{b} = \tilde{a}(\tilde{u})$ *and* $\tilde{u}$ *satisfy the differential equation (3.1.1) along with the boundary condition either (3.1.2) or (3.1.3). Moreover,* $\tilde{b}$, $\tilde{u}$ *and* $\tilde{\sigma} = \tilde{b}\tilde{u}'$ *have the following regularities:*

$$\tilde{b} \in H^{d_b+1}(0,1), \tilde{u} \in \begin{cases} H^{d_u+1}(0,1), \ \text{if (3.1.2) holds} \\ H^{d_u+1}(0,1) \bigcap W^{d_u+1,\infty}(0,1), \ \text{if (3.1.3) holds} \end{cases}, \tilde{\sigma} \in H^{d_\sigma+1}(0,1),$$

* *The finite element space* U_h *is such that*

$$\|(I_u^h\tilde{u})'\|_\infty \leq C,$$

for h *small enough.*

* *The regularization parameters* γ, ρ *and* δ *are chosen such that*

$$\gamma, \rho, \ \text{and } \delta \sim C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\}$$

with a constant C *independent of* h*. The term containing* θ *above is dropped in the choice of the regularization parameters and in all the estimations after now if the boundary condition (3.1.2) holds.*

Let $\mathbf{w}_h^* = (b_h^*, u_h^*, \sigma_h^*)$ *be a minimizer of* $\mathcal{J}_h(b_h, u_h, \sigma_h)$ *defined by (3.2.11) in* $B_h \times U_h \times \Sigma_h$*, then we have the following estimate for the cost functional*

$$\mathcal{J}_h(\mathbf{w}_h^*) \leq C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\}$$

here C *is independent of* ϵ *and* h*. Moreover, we get the following bounds,*

$$\begin{aligned} \|\sigma_h^* - b_h^*(u_h^*)'\|_0 &\leq \tfrac{C}{\sqrt{\alpha}}\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\} \\ \|(\sigma_h^*)' + f\|_0 &\leq \tfrac{C}{\sqrt{\beta}}\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\} \end{aligned}$$

Proof. We treat the case in which the boundary condition (3.1.3) holds. Similar estimations can be obtained for boundary condition (3.1.2) by dropping all the terms containing θ. Let $(b_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h$, which will be denoted by $\mathbf{w}_h^*$, and let $I_b^h\tilde{b}$, $I_u^h\tilde{u}$ and $I_\sigma^h\tilde{\sigma}$ be the interpolates of $\tilde{b} = \tilde{a}(\tilde{u})$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}(\tilde{u})\tilde{u}'$ in B_h, U_h, and Σ_h, respectively. Now

$$\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) \leq \mathcal{J}_h(I_b^h\tilde{b}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}),$$

and

$$\begin{aligned} \mathcal{J}_h(I_b^h\tilde{b}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}) = \ & \|I_u^h\tilde{u} - z\|_0^2 + \alpha\|I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}(I_u^h\tilde{u})'\|_0^2 + \beta\|(I_\sigma^h\tilde{\sigma})' + f\|_0^2 + \\ & \gamma\|I_b^h\tilde{b}\|_{r_b}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2 + \theta(I_b^h\tilde{b}(1)(I_u^h\tilde{u})'(1) - I_\sigma^h\tilde{\sigma}(1))^2. \end{aligned}$$

The first three terms on the right hand side of the inequality above have the following estimates,

$$
\begin{aligned}
\|I_u^h\tilde{u} - z\|_0 &\le \|I_u^h\tilde{u} - \tilde{u}\|_0 + \|\tilde{u} - z\|_0 \\
&\le Ch^{d_u+1}\|\tilde{u}\|_{d_u+1} + \epsilon, \\
\|I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}(I_u^h\tilde{u})'\|_0 &\le \|I_\sigma^h\tilde{\sigma} - \tilde{b}(\tilde{u})'\|_0 + \|\tilde{b}(\tilde{u} - I_u^h\tilde{u})'\|_0 + \|(\tilde{b} - I_b^h\tilde{b})(I_u^h\tilde{u})'\|_0 \\
&\le \|I_\sigma^h\tilde{\sigma} - \tilde{b}(\tilde{u})'\|_0 + \|\tilde{b}\|_\infty\|(\tilde{u} - I_u^h\tilde{u})'\|_0 + \|\tilde{b} - I_b^h\tilde{b}\|_0\|(I_u^h\tilde{u})'\|_\infty \\
&\le C\{h^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + h^{d_u}\|\tilde{u}\|_{d_u+1} + h^{d_b+1}\|\tilde{b}\|_{d_b+1}\},
\end{aligned}
$$

$$
\begin{aligned}
\|(I_\sigma^h\tilde{\sigma})' + f\|_0 &\le \|(I_\sigma^h\tilde{\sigma} - \tilde{a}(\tilde{u})\tilde{u}')'\|_0 \\
&\le Ch^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1}.
\end{aligned}
$$

For the last term, since Lagrange type finite elements are used here,

$$
\begin{aligned}
\mid I_b^h\tilde{b}(1)(I_u^h\tilde{u})'(1) - I_\sigma^h\tilde{\sigma}(1) \mid &= \mid \tilde{a}(\tilde{u}(1))(I_u^h\tilde{u})'(1) - \tilde{\sigma}(1) \mid \\
&= \mid \tilde{a}(\tilde{u}(1))(I_u^h\tilde{u})'(1) - \tilde{a}(\tilde{u}(1))\tilde{u}'(1) \mid \\
&= C \mid (I_u^h\tilde{u})'(1) - \tilde{u}'(1) \mid \\
&\le C\|(I_u^h\tilde{u} - \tilde{u})'\|_\infty \\
&\le Ch^{d_u}\|\tilde{u}\|_{d_u+1,\infty},
\end{aligned}
$$

Then, so far

$$
\begin{aligned}
\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) &\le C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma}\} + \\
&\quad \gamma\|I_b^h\tilde{b}\|_{r_b}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2 + C\theta h^{2d_u},
\end{aligned}
$$

since

$$
\|I_b^h\tilde{b}\|_{r_b}, \|I_u^h\tilde{u}\|_{r_u}, \|I_\sigma^h\tilde{\sigma}\|_{r_\sigma} \le \text{ constant}
$$

and choosing γ, ρ and δ as in the hypothesis, we get the main assertion of the lemma,

$$
\mathcal{J}_h(\mathbf{w}_h^*) \le C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\}
$$

We can get the next two bounds from the definition of $\mathcal{J}_h(b_h, u_h, \sigma_h)$ and the bound above,

$$
\begin{aligned}
\|\sigma_h^* - b_h^*(u_h^*)'\|_0 &\le \tfrac{C}{\sqrt{\alpha}}\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\} \\
\|(\sigma_h^*)' + f\|_0 &\le \tfrac{C}{\sqrt{\beta}}\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\}
\end{aligned}
$$

As a consequence of this lemma, the minimizer of $\mathcal{J}_h(\mathbf{w}_h)$ is uniformly bounded with respect to h. This result is stated in the following

Corollary 3.3.1 *If all the conditions of Lemma 3.3.1 hold, and*

$$
\begin{aligned}
0 &< \gamma = C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\}, \\
0 &< \rho = C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\}, \\
0 &< \delta = C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \theta h^{2d_u}\},
\end{aligned}
$$

then there exists a constant C such that

$$\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C, \ \textit{for all } h.$$

Proof. This proof is identical to that in Corollary 2.3.1.

Theorem 3.3.1 *Suppose that all the conditions in Lemma 3.3.1 and Corollary 3.3.1 hold. If $(b_h^*, u_h^*, \sigma_h^*)$ is a minimizer of $\mathcal{J}_h(b_h, u_h, \sigma_h)$, then we get the following estimation in the weighted H^{-1}-norm,*

$$\begin{aligned} \|((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} \ &\leq \ C(\tfrac{1}{\sqrt{\alpha}} + \tfrac{1}{\sqrt{\beta}} + \tfrac{1}{\sqrt{\xi}} + \tfrac{1}{h})\{\epsilon + h^{(d_u+1)} + \\ &\sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\}. \end{aligned}$$

where C is independent of ϵ and h. The term containing $\sqrt{\theta}$ is dropped if the boundary condition (3.1.2) holds.

Proof. Due to the boundary condition satisfied by $\tilde{u}$ and integration by parts the following holds,

$$\begin{aligned} \|((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} \ &= \ \sup_{v\in\tilde{H}^1(0,1)} \frac{|(((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}')', v)|}{\|v\|_1} \\ &= \ \sup_{v\in\tilde{H}^1(0,1)} \frac{|((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}', v')|}{\|v\|_1} \end{aligned}$$

We now estimate the term $((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}', v')$ when $v \in \tilde{H}^1(0,1)$. By using the weak form of equation (3.1.1) with the boundary condition (3.1.2), adding and subtracting various terms and integrating by parts, we get

$$((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}', v') \ = \ (f + (\sigma_h^*)', v) + (\sigma_h^* - b_h^*(u_h^*)', v') + (b_h^*(u_h^* - \tilde{u})', v').$$

By the same reasons as above, but with boundary condition (3.1.3), we get

$$\begin{aligned} ((\tilde{a}(\tilde{u}) - b_h^*)\tilde{u}', v') \ &= \ (f + (\sigma_h^*)', v) - \sigma_h^*(1)v(1) + (\sigma_h^* - b_h^*(u_h^*)', v') + \\ &(b_h^*(u_h^* - \tilde{u})', v'). \end{aligned}$$

We proceed to estimate each term in the equality above. By Lemma 3.3.1

$$\begin{aligned} \left.\begin{array}{r} |\ (f + (\sigma_h^*)', v)\ | \\ |\ (f + (\sigma_h^*)', v)\ | + |\ \sigma_h^*(1)v(1)\ | \end{array}\right\} \ &\leq \ C(\tfrac{1}{\sqrt{\beta}} + \tfrac{1}{\sqrt{\xi}})\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \\ &\sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\}\|v\|_1, \\ |\ (\sigma_h^* - b_h^*(u_h^*)', v')\ | \ &\leq \ \tfrac{C}{\sqrt{\alpha}}\{\epsilon + h^{(d_u+1)} + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \\ &\sqrt{\beta}h^{d_\sigma} + \sqrt{\theta}h^{d_u}\}\|v\|_1. \end{aligned}$$

By Lemma 3.3.1 and the inverse inequality (2.2.16) we get

$$\begin{aligned}
|\,(b_h^*(u_h^*-\tilde{u})', v')\,| &\leq C\|(u_h^*-\tilde{u})'\|_0\|v\|_1 \\
&\leq C(\|(u_h^*-I_u^h\tilde{u})'\|_0+\|(I_u^h\tilde{u}-\tilde{u})'\|_0)\|v\|_1 \\
&\leq C(\frac{1}{h}\|u_h^*-I_u^h\tilde{u}\|_0+h^{d_u}\|\tilde{u}\|_{d_u+1})\|v\|_1 \\
&\leq C(\frac{1}{h}(\|u_h^*-z\|_0+\|z-\tilde{u}\|_0+\|\tilde{u}-I_u^h\tilde{u}\|_0)+h^{d_u}\|\tilde{u}\|_{d_u+1})\|v\|_1 \\
&\leq C\frac{1}{h}(\epsilon+h^{(d_u+1)}+\sqrt{\alpha}(h^{(d_\sigma+1)}+h^{d_u}+h^{(d_b+1)})+\sqrt{\beta}h^{d_\sigma}+\sqrt{\theta}h^{d_u})\|v\|_1.
\end{aligned}$$

Finally, we get the assertion of the Theorem by combining the above three estimations,

$$\begin{aligned}
\|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*} \leq\ & C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{\sqrt{\xi}}+\tfrac{1}{h})(\epsilon+h^{(d_u+1)}+\sqrt{\alpha}(h^{(d_\sigma+1)}+h^{d_u}+h^{(d_b+1)})+ \\
& \sqrt{\beta}h^{d_\sigma}+\sqrt{\theta}h^{d_u}).
\end{aligned}$$

Corollary 3.3.2 *Suppose all the conditions in Theorem 3.3.1 are satisfied and that the boundary condition in the PI problem is (3.1.3), then a minimizer* $\mathbf{w}_h^*=(b_h^*,u_h^*,\sigma_h^*)$ *of* $\mathcal{J}_h(b_h,u_h,\sigma_h)$ *satisfies the following weighted* L^2 *norm error estimation,*

$$\begin{aligned}
\|(\tilde{a}(\tilde{u})-b_h^*)\tilde{u}'\|_0 \leq\ & C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{\sqrt{\xi}}+\tfrac{1}{h})(\epsilon+h^{(d_u+1)}+\sqrt{\alpha}(h^{(d_\sigma+1)}+h^{d_u}+h^{(d_b+1)})+ \\
& \sqrt{\beta}h^{d_\sigma}+\sqrt{\theta}h^{d_u}).
\end{aligned}$$

Proof. Let χ be the solution of the following initial value problem,

$$\begin{cases}
\chi'(x) = (\tilde{a}(\tilde{u})-b_h^*)\tilde{u}'(x), \ x\in\Omega=(0,1) \\
\chi(0) = 0
\end{cases}$$

The solution to the above initial value problem is $\chi(x)=\int_0^x(\tilde{a}(\tilde{u})-b_h^*(\tau))\tilde{u}'(\tau)\,d\tau$, so $\chi\in\tilde{H}^1(\Omega)$ and by Theorem 3.3.1, we have

$$\begin{aligned}
((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}',\chi') &\leq \|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|\chi\|_1, \\
((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}',\chi') &\leq \|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|\chi'\|_0 \text{ by Poincare's inequality,} \\
((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}',\chi') &\leq \|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*}\|(\tilde{a}(\tilde{u})-b_h^*)\tilde{u}'\|_0.
\end{aligned}$$

It follows that

$$\|(\tilde{a}(\tilde{u})-b_h^*)\tilde{u}'\|_0\leq\|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{(\tilde{H}^1(0,1))^*},$$

and the assertion of the Corollary follows straightforward from Theorem 3.3.1.

Corollary 3.3.3 *Assume that the boundary condition (3.1.2) holds in the PI problem. Then under the conditions of Theorem 3.3.1, we have the following weighted error estimation in the* H^{-1} *norm,*

$$\begin{aligned}
\|((\tilde{a}(\tilde{u})-b_h^*)\tilde{u}')'\|_{-1} \leq\ & C(\tfrac{1}{\sqrt{\alpha}}+\tfrac{1}{\sqrt{\beta}}+\tfrac{1}{\sqrt{\xi}}+\tfrac{1}{h})\{\epsilon+h^{(d_u+1)}+ \\
& \sqrt{\alpha}(h^{(d_\sigma+1)}+h^{d_u}+h^{(d_b+1)})+\sqrt{\beta}h^{d_\sigma}\}.
\end{aligned}$$

Proof. The result follows from Theorem 3.3.1, since in this case $\tilde{H}^1(0,1) = H_0^1(0,1)$, and $(\tilde{H}^1(0,1))^* = H^{-1}(0,1)$.

3.4 Numerical examples

We present some numerical examples which display interesting features of the finite element method in this chapter. To describe the accuracy of a finite element solution $b_h = a_h(u_h)$, we let $\tilde{a}(s)$ and $\tilde{u}(x)$ be the functions satisfying the differential equation (3.1.1) along with the boundary condition (3.1.2) or (3.1.3). Let

$$E^0(a_h) = (\int_0^1 (\tilde{a}(\tilde{u}) - a_h(u_h))^2(\tilde{u}'(x))^2\, dx)^{1/2}, \tag{3.4.13}$$

Obviously, $\tilde{a}(s)$ can be considered as the exact solution of the PI problem, $E^0(a_h)$ can be considered as a weighted L^2 norm of the error in a finite element solution $b_h = a_h(u_h)$. All the numerical results presented in this section were carried out with the finite element spaces such that

$$d_b = 1, l = 1, d_u = 2, d_\sigma = 2,$$

and they were obtained by a two-level or three-level Newton method with initial guesses generated by perturbing the exact solutions with a random variable of magnitude 5. Let us not forget that in all our calculations, we are calculating not only the coefficient function $a(s)$, which is one of our unknowns, but also the state function $u(x)$, as well as the flux function $\sigma(x)$. The weighted L^2 errors for various values of h are shown in the data tables along with some pictures.

Example 1: In this last example, we identify the nonlinear coefficient $a(s) = e^{-s}$, from the data function $z(x) = \sin(\pi x) + \pi x$ for the boundary value problem (3.1.1), (3.1.3). We assume that the data is perfect, i.e., $z(x) = u(x)$ and $f(x)$ is chosen so that $u(x)$ is the exact solution of the two point boundary value problem with the given $a(s)$. The errors in the weighted L^2 norm for the finite element solutions $b_h = a_h(u_h)$ with various h are given in Table 3.1. To see the rate of convergence, we assume that the error obeys

$$E^0(a_h) = Ch^r.$$

Then the data in Table 3.1 satisfies

$$E^0(a_h) = Ch^{2.00042895934973}.$$

This power of h agrees with the theoretical rate of convergence given by Corollary 3.3.2 which should be 2 with no observation error.

The same rate of convergence will be obtained for the next PI problem with ho-

Table 3.1			
h	weighted L^2-error	h	weighted L^2-error
1/30	$6.4394614 * 10^{-3}$	1/110	$4.7778042 * 10^{-4}$
1/40	$3.6177598 * 10^{-3}$	1/120	$4.0145181 * 10^{-4}$
1/50	$2.3141633 * 10^{-3}$	1/130	$3.4205410 * 10^{-4}$
1/60	$1.6066215 * 10^{-3}$	1/140	$2.9492553 * 10^{-4}$
1/70	$1.1801810 * 10^{-3}$	1/150	$2.5690657 * 10^{-4}$
1/80	$9.0347604 * 10^{-4}$	1/160	$2.2579088 * 10^{-4}$
1/90	$7.1380052 * 10^{-4}$	1/170	$2.0000498 * 10^{-4}$
1/100	$5.7814271 * 10^{-4}$	1/180	$1.7839625 * 10^{-4}$

Table 3.1: Errors in finite element solutions with various h for the PI problem with boundary condition (3.1.3)

mogeneous Dirichlet boundary conditions by using the same finite element spaces in the approximation. A rigorous proof for the weighted L^2 norm error estimation for the next PI problem still remains an open issue that needs further investigation.

Example 2: The first problem consists in identifying $a(s) = s^4 - s^2 + \frac{1}{2}$ from the data function $z(x) = \sin(2\pi x)$ in the boundary value problem (3.1.1), (3.1.2). We assume that the data is perfect, i.e., $z(x) = u(x)$ and $f(x)$ is chosen so that $u(x)$ is the exact solution to the boundary value problem with the given $a(x)$. The errors in the weighted L^2 norm for the finite element solutions with various h are given in Table 3.2. To see the rate of convergence, we assume that the error obeys

$$E^0(a_h) = Ch^r.$$

Then the data in Table 3.2 satisfies

$$E^0(a_h) = Ch^{1.99373239302467},$$

This example strongly suggests that the inequality in Corollary 3.3.2 should also hold for boundary condition (3.1.2) as was the case for the PI problem with boundary condition (3.1.3). We displayed the identification of a_h^* in the figure 3.1.

In the next example, we want to explore the features of our method as to identify a discontinuous coefficient.

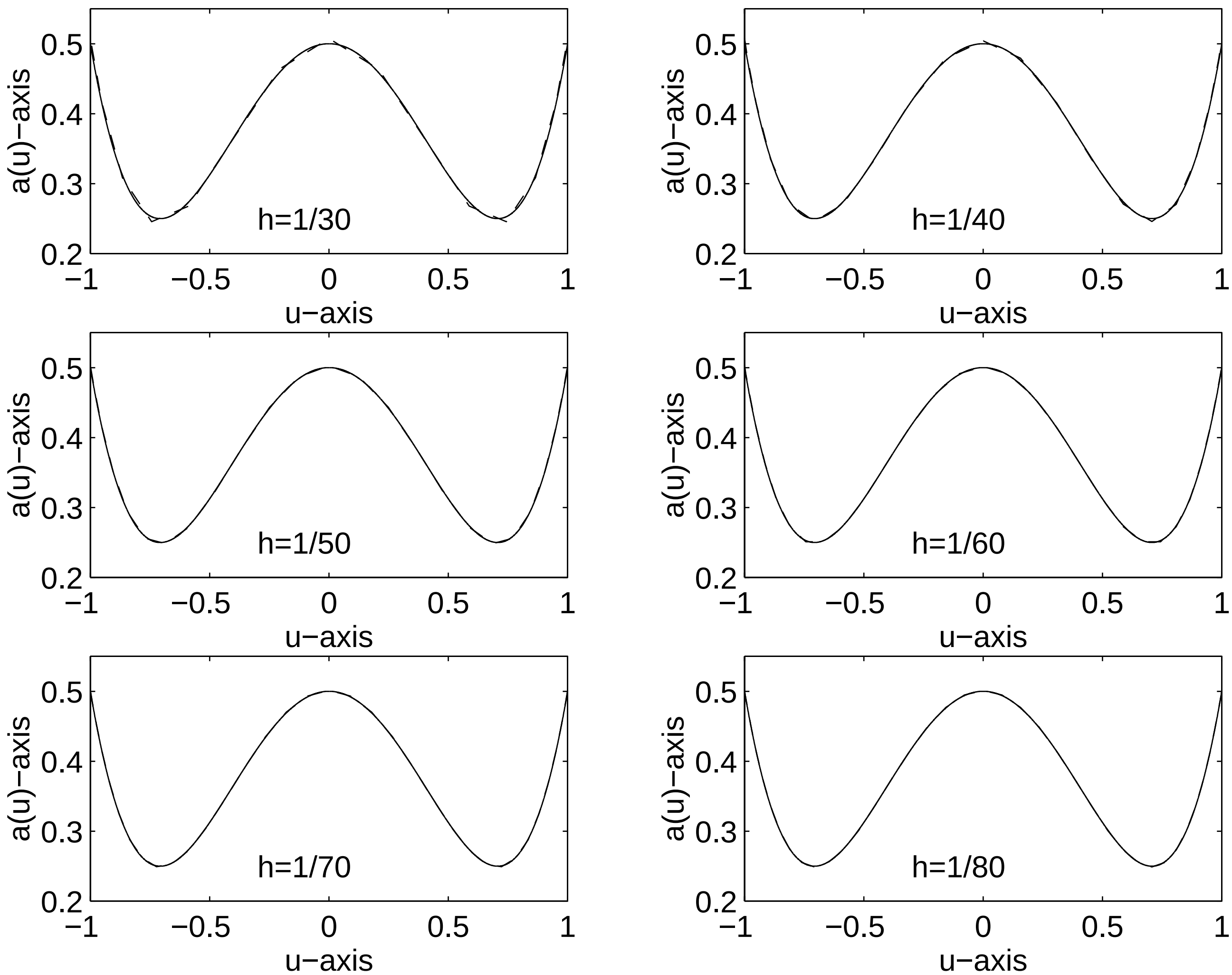

Figure 3.1: finite element approximation of $\tilde{a}(\tilde{u})$ as a function of $\tilde{u}$ for various h

Table 3.2			
h	weighted L^2-error	h	weighted L^2-error
1/30	$2.1774801 * 10^{-2}$	1/110	$1.7014844 * 10^{-3}$
1/40	$1.2536272 * 10^{-2}$	1/120	$1.4306409 * 10^{-3}$
1/50	$8.1097540 * 10^{-3}$	1/130	$1.2196122 * 10^{-3}$
1/60	$5.6652307 * 10^{-3}$	1/140	$1.0520232 * 10^{-3}$
1/70	$4.1769626 * 10^{-3}$	1/150	$9.1672019 * 10^{-4}$
1/80	$3.2054390 * 10^{-3}$	1/160	$8.0592304 * 10^{-4}$
1/90	$2.5366946 * 10^{-3}$	1/170	$7.1405110 * 10^{-4}$
1/100	$2.0570738 * 10^{-3}$	1/180	$6.3703234 * 10^{-4}$

Table 3.2: Errors in finite element solutions with various h for the PI problem with boundary condition (3.1.2)

Example 3: Finally, we identify the piecewise constant coefficient

$$a(s) = \begin{cases} 1.2 & \text{for } u_{min} \leq s \leq \frac{3}{8} \\ 1.5 & \text{for } \frac{3}{8} < s \leq u_{max} \end{cases} \tag{3.4.14}$$

from the data

$$z(x) = \begin{cases} \frac{3}{2}x & \text{for } 0 \leq x \leq \frac{1}{4} \\ -\frac{68}{30}x^2 + \frac{7}{3}x - \frac{2}{30} & \text{for } \frac{1}{4} < x \leq 1 \end{cases} \tag{3.4.15}$$

for the boundary value problem (3.1.1), (3.1.2). We assume that the data is perfect and $f(x) = -(a(u)u')'$. We carried out this example so that the nonlinear constraint,

$$a(u_-)u'_- = a(u_+)u'_+$$

holds. Some pictures are depicted below in figure 3.2. As we can see, our method does not provide us with a good approximation to the nonlinear coefficient in this case.

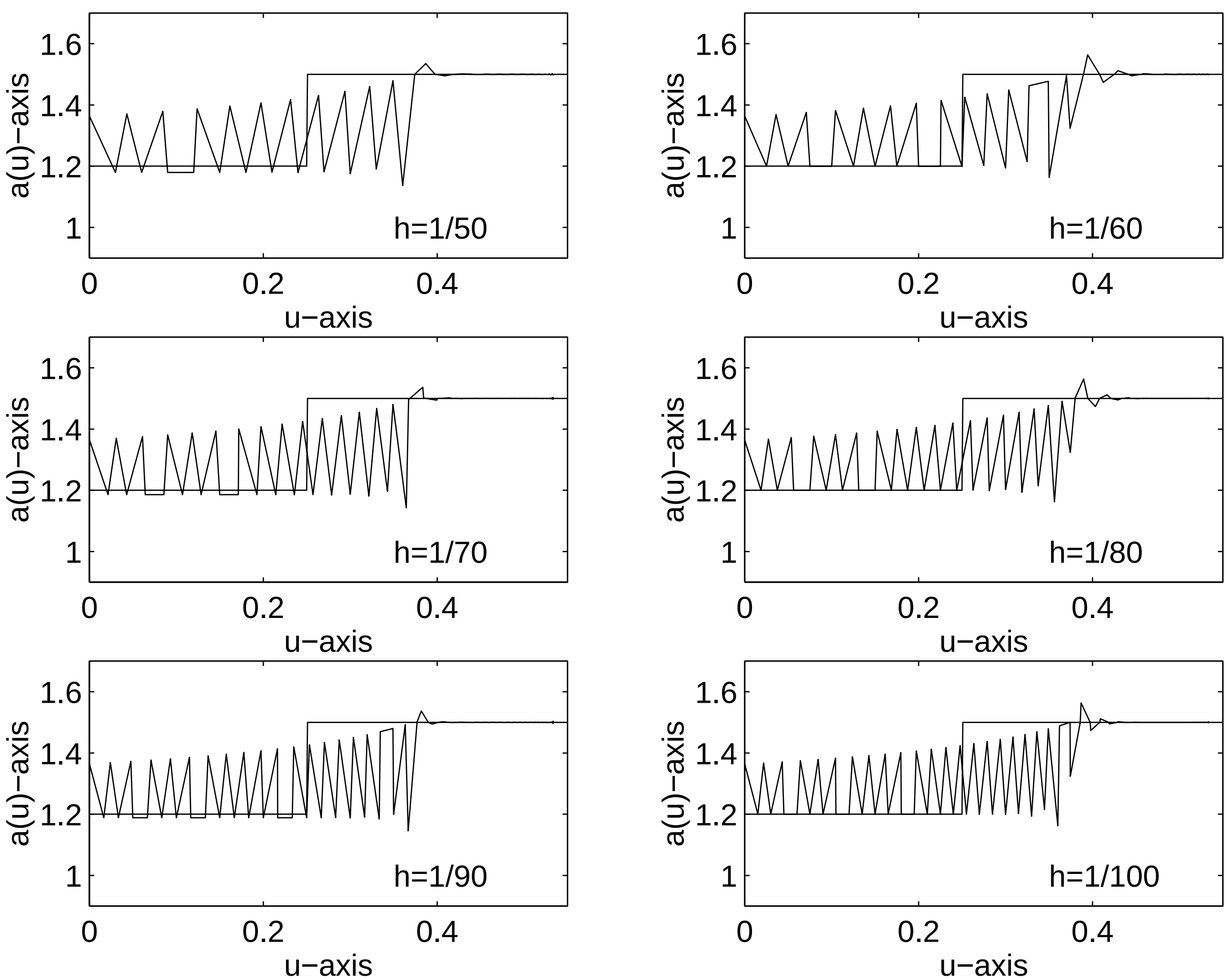

Figure 3.2: finite element approximation of $\tilde{a}(\tilde{u})$ for various h

Chapter 4

Parameter identification of two dimensional elliptic boundary value problem

4.1 Introduction

In this chapter we are concerned with the identification of the unknown coefficient $a(x)$ for a two dimensional elliptic boundary value problem. We present and analyze a finite element method based on the mixed least-squares idea.

Let Ω be a bounded polygonal domain in $\mathbb{R}^n$ with a Lipschitz continuous boundary Γ. We consider a boundary value problem (BVP) with a linear partial differential equation

$$-\nabla \cdot (a(x)\nabla u(x)) = f(x) \text{ in } \Omega \subset \mathbb{R}^n, \tag{4.1.1}$$

or nonlinear partial differential equation

$$-\nabla \cdot (a(u(x))\nabla u(x)) = f(x) \text{ in } \Omega \subset \mathbb{R}^n, \tag{4.1.2}$$

with the Dirichlet boundary condition

$$u \mid_\Gamma = 0, \tag{4.1.3}$$

or Neumann boundary condition

$$\frac{\partial u}{\partial \eta}|_\Gamma = 0, \tag{4.1.4}$$

where $f \in L^2(\Omega)$, and $\lambda \leq a(x) \leq \Lambda$ for some positive constants λ and Λ. The parameter identification (PI) problem is to recover the parameter $a = a(x)$ within some suitable set, from an L^2 observation $z(x)$ of $u(x)$ such that $a(x)$ and $u(x)$ satisfy (4.1.1) along with the boundary condition given by (4.1.3).

The differential equations here correspond to a steady state solution of the parabolic partial differential equation below,

$$-\nabla \cdot (a(u(x,t))\nabla u(x,t)) = f(x,t) + \frac{\partial u(x,t)}{\partial t}, \ (x,t) \in \Omega \times [0,T], \tag{4.1.5}$$

supplemented with appropriate initial and boundary conditions. This differential equation models the flow of oil in a reservoir, $a(x)$ is the transmissivity of the porous medium, the dependent variable u represents pressure and f accounts for the withdrawal or injection of fluid in the reservoir.

To describe the method, we let $H^s(\Omega)$ be the standard Sobolev space. We use $\|\cdot\|_s$ to denote its norm, and use $|\cdot|_s$ to denote its usual seminorm. The inner product of $H^s(\Omega)$ will be denoted by $(\cdot,\cdot)_s$ with $(\cdot,\cdot) = (\cdot,\cdot)_0$ for the inner product in $L^2(\Omega)$. Another space we will need throughout this chapter is the Sobolev space $W^{m,\infty}(\Omega)$, for m a positive integer, equipped with the norm

$$\|v\|_{m,\infty} = \max_{|\alpha|\leq m} \|\frac{\partial^\alpha}{\partial x^\alpha} v(x)\|_\infty$$

4.2 The linear BVP with Dirichlet boundary condition

In this section, we consider the PI problem for BVP formed by the linear partial differential equation (4.1.1) and the Dirichlet boundary condition (4.1.3).

As before, our PI problem has two unknowns: $a(x)$ and $u(x)$. Following the standard set up in the mixed formulation, we introduce the flux $\sigma(x) = a(x)\nabla u(x)$ as an additional unknown. These unknowns satisfy the equivalent first order system:

$$\left\{ \begin{array}{rcl} -\nabla \cdot \sigma & = & f, \\ a\nabla u & = & \sigma. \end{array} \right. \tag{4.2.6}$$

We then try to find an approximation to the solution of the PI problem by minimizing the following cost functional,

$$\begin{aligned} \mathcal{J}(a,u,\sigma) \ = \ & \|u - z\|_0^2 + \alpha\|\sigma - a\nabla u\|_0^2 + \beta\|\text{div}\sigma + f\|_0^2 + \\ & \gamma\|a\|_{r_a}^2 + \rho\|u\|_{r_u}^2 + \delta\|\sigma\|_{r_\sigma}^2, \end{aligned} \tag{4.2.7}$$

over the space $H^1(\Omega) \times (H_0^1(\Omega) \cap W^{1,\infty}(\Omega)) \times (H^1(\Omega) \times (H^1(\Omega))$, where the integers r_a, r_u and r_σ depend on the Sobolev spaces in which the regularization takes place. The nonnegative numbers α, β, γ, ρ and δ can be considered as regularization parameters. The meaning of each term in (4.2.7) is the same as that for the previous two chapters.

4.2.1 Finite element method

In this section, we formulate a finite element discretization of the cost functional (4.2.7) and discuss the related minimization problem. We regard C as a generic constant, which may vary in different contexts, but is always independent of h, unless specified otherwise.

In order to define the finite element spaces, let $\{\Delta_h\}$, with $h \in (0,1)$, be a family of regular, quasi-uniform partitions of $\Omega \subset I\!R^n$. For fixed integers $r \geq 1$, $l \geq 0$, we let

$$S^r_{h,l} = \{v \in C^{l-1}(\bar{\Omega}) : v|_T \in P_r \text{ for all } T \in \Delta_h\}, \tag{4.2.8}$$

be a Lagrange type finite element space, where P_r is the space of polynomials of degree less than or equal to r and $C^{-1}(\Omega)$ is interpreted as $L^2(\Omega)$. By $S^{r,0}_{h,l}$, we denote the subspace of functions of $S^r_{h,l}$ which vanish on $\partial\Omega$. By the results in [14] and [38], we know that for all $v \in W^{m,p}(\Omega)$, there is (an interpolant) $v_h \in S^r_{h,l}$ such that all three inequalities (2.2.15), (2.2.16), (2.2.17) hold.

We denote by $H^{-1}= H^{-1}(\Omega)$ the dual space $(H_0^1(\Omega))^*$ equipped with the norm

$$\|v\|_{-1} = \sup_{0 \neq \psi \in H_0^1(\Omega)} \frac{|(v,\psi)|}{\|\psi\|_1}. \tag{4.2.9}$$

Let us now introduce those finite dimensional spaces needed to define the computational procedure. We need altogether three different discretization spaces, namely, U_h for the solution u, A_h for the parameter a, and finally Σ_h, which will be used to discretize the flux variable σ. We assume that these finite element spaces are,

$$\begin{aligned} A_h &= S^{d_a}_{h,l}, \\ U_h &= S^{d_u}_{h,1}, \\ \Sigma_h &= S^{d_\sigma}_{h,1} \times S^{d_\sigma}_{h,1}, \end{aligned} \tag{4.2.10}$$

where d_a, d_u, d_σ are integers.

Now we are ready to define the numerical method to solve the PI problem:
(P_h) seek $a_h^* \in A_h$, $u_h^* \in U_h$, $\sigma_h^* \in \Sigma_h$, such that

$$\mathcal{J}_h(a_h^*, u_h^*, \sigma_h^*) = \min_{a \in A_h, u \in U_h, \sigma \in \Sigma_h \times \Sigma_h} \mathcal{J}_h(a,u,\sigma), \tag{4.2.11}$$

where $\mathcal{J}_h$ is defined by,

$$\begin{aligned}\mathcal{J}_h(\mathbf{w}_h) &= \|u_h - z\|_0^2 + \alpha\|\sigma_h - a_h\nabla u_h\|_0^2 + \beta\|\text{div}\sigma_h + f\|_0^2 + \\ &\quad \gamma\|a_h\|_{r_a}^2 + \rho\|u_h\|_{r_u}^2 + \delta\|\sigma_h\|_{r_\sigma}^2,\end{aligned} \tag{4.2.12}$$

for $\mathbf{w}_h = (a_h, u_h, \sigma_h) \in A_h \times U_h \times \Sigma_h$. As we can see, the cost functional (4.2.12) consists of several terms, the first term represents the usual output least squares fit of the data measured in $L^2(\Omega)$, the second and third terms are the equation error residuals of the involved differential equation in mixed formulation. The remaining terms are for regularization purposes. The advantage of this variational approach is that we allow all variables $(a, u,$ and $\sigma)$ to vary in the corresponding discrete subspace without differentiating any kind of error corrupted data z. The weights γ, ρ, and δ will be chosen so that all terms of $\mathcal{J}_h(\mathbf{w}_h)$ are balanced.

4.2.2 Some error estimates in weighted norms

In this section, we establish some error estimates for an approximate solution to the PI problem generated by the finite element method in the previous section. The estimates and the arguments are influenced by [20], [25], [26]. In all the discussion below, C denotes a generic constant which may take a different value at different places and is independent of h unless specified otherwise. Let I_a^h, I_u^h, and I_σ^h be the interpolation operators in the finite element spaces A_h, U_h, and Σ_h, respectively. The regularization spaces are chosen so that

$$1 \leq r_a \leq d_a + 1, 1 \leq r_u \leq d_u + 1, 1 \leq r_\sigma \leq d_\sigma + 1.$$

Moreover, there exists a constant C such that the interpolation operators satisfy,

$$\|I_a^h a\|_{r_a} \leq C\|a\|_{r_a+1}, \|I_u^h u\|_{r_u} \leq C\|u\|_{r_u+1}, \|I_\sigma^h \sigma\|_{r_\sigma} \leq C\|\sigma\|_{r_\sigma+1},$$

for any $a \in H^{r_a+1}(\Omega), u \in H^{r_u+1}(\Omega), \sigma \in (H^{r_\sigma+1}(\Omega) \times H^{r_\sigma+1}(\Omega))$. We first derive an estimate for a minimizer of $\mathcal{J}_h(\mathbf{w})$.

Lemma 4.2.1 *Let us assume the following hypothesis,*

* *Let $z(x)$ be a distributed L^2 observation of the state $\tilde{u}$ with an observation error of the form*

$$\|z - \tilde{u}\|_0 \leq \epsilon.$$

* *Let $\tilde{a}$, $\tilde{u}$ satisfy the differential equation (4.1.1) together with the boundary condition (4.1.3). Moreover, $\tilde{a}$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}\nabla\tilde{u}$, have the following regularities:*

$$\tilde{a} \in H^{d_a+1}(\Omega), \tilde{u} \in H^{d_u+1}(\Omega), \tilde{\sigma} \in H^{d_\sigma+1}(\Omega) \times H^{d_\sigma+1}(\Omega).$$

* *The finite element space U_h is such that*

$$\|\nabla(I_u^h\tilde{u})\|_\infty \leq C,$$

for h small enough.

* *The regularization parameters γ, ρ, and δ are chosen such that*

$$\gamma+\rho+\delta \sim C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\},$$

with a constant C independent of h.

Let $\mathbf{w}_h = (a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h(a,u,\sigma)$ defined by (4.2.12) in $A_h \times U_h \times \Sigma_h$, then we have the following estimate for the cost functional $\mathcal{J}_h$,

$$\mathcal{J}_h(\mathbf{w}_h^*) \leq C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma}\}, \tag{4.2.13}$$

here C is independent of ϵ and h. Moreover, we get the following bounds,

$$\begin{aligned}
\|\nabla(u_h^* - \tilde{u})\|_0 &\leq \frac{C}{h}\{h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma}\}\\
\|\sigma_h^* - a_h^*\nabla u_h^*\|_0 &\leq \frac{C}{\sqrt{\alpha}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}\\
\|\mathrm{div}\sigma_h^* + f\|_0 &\leq \frac{C}{\sqrt{\beta}}\{\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_u} + h^{d_a+1}) + \sqrt{\beta}h^{d_\sigma}\}
\end{aligned}$$

Proof. Let $(a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h$, which will be denoted by $\mathbf{w}_h^*$, and let $I_a^h\tilde{a}$, $I_u^h\tilde{u}$, and $I_\sigma^h\tilde{\sigma}$ be the interpolants of $\tilde{a}$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}\nabla\tilde{u}$ in A_h, U_h, and Σ_h, respectively. Now

$$\mathcal{J}_h(\mathbf{w}_h^*) \leq \mathcal{J}_h(I_a^h\tilde{a}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}),$$

and

$$\begin{aligned}
\mathcal{J}_h(I_a^h\tilde{a}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}) &= \|I_u^h\tilde{u} - z\|_0^2 + \alpha\|I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\|_0^2 + \beta\|\mathrm{div}I_\sigma^h\tilde{\sigma} + f\|_0^2 +\\
&\quad \gamma\|I_a^h\tilde{a}\|_{r_a}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2.
\end{aligned}$$

The first three terms of the equality above have the following bounds,

$$\begin{aligned}
\|I_u^h\tilde{u} - z\|_0 &\leq \|I_u^h\tilde{u} - \tilde{u}\|_0 + \|\tilde{u} - z\|_0\\
&\leq Ch^{d_u+1}\|\tilde{u}\|_{d_u+1} + \epsilon,\\
\|I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\|_0 &\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_0 + \|\tilde{a}\nabla\tilde{u} - \tilde{a}\nabla(I_u^h\tilde{u})\|_0 + \|\tilde{a}\nabla(I_u^h\tilde{u}) - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\|_0\\
&\leq Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + \|\tilde{a}\|_\infty\|\nabla(\tilde{u} - I_u^h\tilde{u})\|_0 + \|\nabla(I_u^h\tilde{u})\|_\infty\|\tilde{a} - I_a^h\tilde{a}\|_0\\
&\leq Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + Ch^{d_u}\|\tilde{u}\|_{d_u+1} + Ch^{d_a+1}\|\tilde{a}\|_{d_a+1},\\
\|\mathrm{div}I_\sigma^h\tilde{\sigma} + f\|_0 &\leq \|\mathrm{div}(I_\sigma^h\tilde{\sigma} - \tilde{\sigma})\|_0\\
&\leq Ch^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1},
\end{aligned}$$

Then, so far

$$\begin{aligned}\mathcal{J}_h(\mathbf{w}_h^*) &\leq C\{h^{2(d_u+1)}+\epsilon^2+\alpha(h^{2(d_\sigma+1)}+h^{2d_u}+h^{2(d_a+1)})+\beta h^{2d_\sigma}\}+\\ &\quad \gamma\|I_a^h\tilde{a}\|_{r_a}^2+\rho\|I_u^h\tilde{u}\|_{r_u}^2+\delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2,\end{aligned}$$

since

$$\begin{aligned}\|I_a^h\tilde{a}\|_{r_a} &\leq C\|\tilde{a}\|_{r_a+1},\\ \|I_u^h\tilde{u}\|_{r_u} &\leq C\|\tilde{u}\|_{r_u+1},\\ \|I_\sigma^h\tilde{\sigma}\|_{r_\sigma} &\leq C\|\tilde{\sigma}\|_{r_\sigma+1},\end{aligned}$$

and choosing γ, ρ, and δ as in the hypothesis, we get the main assertion of the lemma:

$$\mathcal{J}_h(\mathbf{w}_h^*) \leq C\{h^{2(d_u+1)}+\epsilon^2+\alpha(h^{2(d_\sigma+1)}+h^{2d_u}+h^{2(d_a+1)})+\beta h^{2d_\sigma}\}.$$

We bound $\|\nabla(u_h^*-\tilde{u})\|_0$ by applying the inverse inequality (2.2.16) to it and the bound above for $\mathcal{J}_h^{1/2}(\mathbf{w}_h^*)$.

$$\begin{aligned}\|\nabla(u_h^*-\tilde{u})\|_0 &\leq \|\nabla(u_h^*-I_u^h\tilde{u})\|_0+\|\nabla(I_u^h\tilde{u}-\tilde{u})\|_0\\ &\leq C\{\frac{1}{h}\|u_h^*-I_u^h\tilde{u}\|_0+h^{d_u}\|\tilde{u}\|_{d_u+1}\}\\ &\leq C\{\frac{1}{h}(\|u_h^*-z\|_0+\|z-\tilde{u}\|_0+\|\tilde{u}-I_u^h\tilde{u}\|_0)+h^{d_u}\}\\ &\leq C\{\frac{1}{h}(\|u_h^*-z\|_0+\epsilon+h^{d_u+1})+h^{d_u}\}\\ &\leq C\{\frac{1}{h}\mathcal{J}_h^{1/2}(b_h^*,u_h^*,\sigma_h^*)+\frac{\epsilon}{h}+h^{d_u}\}\\ &\leq \frac{C}{h}\{\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma}\}.\end{aligned}$$

We can get the next two bounds from the definition of $\mathcal{J}_h(a,u,\sigma)$ and the bound above for $\mathcal{J}_h(\mathbf{w}_h^*)$

$$\begin{aligned}\|\sigma_h^*-a_h^*\nabla u_h^*\|_0 &\leq \tfrac{C}{\sqrt{\alpha}}\{\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma}\},\\ \|\mathrm{div}\sigma_h^*+f\|_0 &\leq \tfrac{C}{\sqrt{\beta}}\{\epsilon+h^{d_u+1}+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_u}+h^{d_a+1})+\sqrt{\beta}h^{d_\sigma}\}.\end{aligned}$$

As a consequence of this lemma, it can be shown that under certain conditions, the minimizer of $\mathcal{J}_h(\mathbf{w}_h)$ is uniformly bounded with respect to h. The result is stated in the following

Corollary 4.2.1 *If all the conditions of Lemma 4.2.1 hold, and*

$$\begin{aligned}0 &< \gamma = C\{\epsilon^2+h^{2(d_u+1)}+\alpha(h^{2(d_a+1)}+h^{2d_u}+h^{2(d_\sigma+1)})+\beta h^{2d_\sigma}\},\\ 0 &< \rho = C\{\epsilon^2+h^{2(d_u+1)}+\alpha(h^{2(d_a+1)}+h^{2d_u}+h^{2(d_\sigma+1)})+\beta h^{2d_\sigma}\},\\ 0 &< \delta = C\{\epsilon^2+h^{2(d_u+1)}+\alpha(h^{2(d_a+1)}+h^{2d_u}+h^{2(d_\sigma+1)})+\beta h^{2d_\sigma}\},\end{aligned}$$

then there exists a constant C such that,

$$\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C, \text{ for all } h. \tag{4.2.14}$$

Proof. In this proof, C denotes a generic constant. Let $\mathbf{w}_h^*$ be a minimizer of $\mathcal{J}_h$. Then according to the definition of $\mathcal{J}_h$ and the bound (4.2.13), we have that

$$\begin{array}{rcccl} \gamma\|a_h^*\|_{r_a}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C, \\ \rho\|u_h^*\|_{r_u}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C, \\ \delta\|\sigma_h^*\|_{r_\sigma}^2 & \leq & \mathcal{J}_h(\mathbf{w}_h^*) & \leq & C. \end{array}$$

Since γ, ρ, and δ are positive and finite, then

$$\begin{array}{rcccl} \|a_h^*\|_{r_a}^2 & \leq & \frac{1}{\sqrt{\gamma}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C, \\ \|u_h^*\|_{r_u}^2 & \leq & \frac{1}{\sqrt{\rho}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C, \\ \|\sigma_h^*\|_{r_\sigma}^2 & \leq & \frac{1}{\sqrt{\delta}}\mathcal{J}_h^{1/2}(\mathbf{w}_h^*) & \leq & C. \end{array}$$

Therefore $\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C$, for all h.

Corollary 4.2.2 *Assume the hypothesis of Corollary 4.2.1 and take H^2 to be the Sobolev space for the regularization of a_h. If $A_h = S_{h,l}^{d_a}$ with $l \geq 2$, then*

$$\|a_h^*\|_\infty < \infty.$$

Proof. The proof follows from the fact that $A_h \subset C^1(\bar{\Omega}) \bigcap H^2(\Omega)$, the boundedness of $\mathbf{w}_h^*$ in the norm of $\mathcal{H}$ in the Corollary 4.2.1, and by the Sobolev's embedding Theorem in two dimension under the hypothesis we have assumed for Ω.

The proof of the following result is motivated by the work in [26] and [3].

Theorem 4.2.1 *Assume all the hypothesis of Corollary 4.2.2. Let $\tilde{a}(x)$, $\tilde{u}(x)$ satisfy equation (4.1.1), where both have the same regularity as in the Lemma 4.2.1, in addition to this, $\tilde{u}(x)$ has to be in $W^{1,\infty}(\Omega)$. Let $\mathbf{w}_h^* = (a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of the cost functional (4.2.12), then we have the following error estimation in the weighted L^1 norm,*

$$\begin{aligned} \int_\Omega |a_h^* - \tilde{a}||\nabla\tilde{u}|^2\,dx \;\leq\; & C\{\delta + \frac{1}{\sqrt{\beta}}(h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma}) + \\ & (\delta^{-1}+1)(\frac{1}{\sqrt{\alpha}} + \frac{1}{h})(h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma})\} \end{aligned}$$

where C is independent of ϵ and h, and δ is just any positive parameter.

Proof. Recall that the weak form of equation (4.1.1) reads as,

$$\int_\Omega \tilde{a}\nabla\tilde{u}\cdot\nabla v\,dx = \int_\Omega fv\,dx \text{ for all } v\in H_0^1(\Omega) \tag{4.2.15}$$

Set $\phi = a_h^* - \tilde{a}$. Now, we estimate the term

$$\mid \int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\,dx \mid,$$

where v is a test function in $H_0^1(\Omega)$. To this end, we add and subtract suitable terms and integrate by parts, to get

$$\begin{aligned} -\textstyle\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\,dx &= \textstyle\int_\Omega (f+\mathrm{div}\sigma_h^*)v\,dx + \int_\Omega(\sigma_h^* - a_h^*\nabla u_h^*)\cdot\nabla v\,dx + \\ &\ \textstyle\int_\Omega a_h^*\nabla(u_h^* - \tilde{u})\cdot\nabla v\,dx, \end{aligned} \tag{4.2.16}$$

Now taking the absolute value of $-\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\,dx$, we get the following estimation,

$$\begin{aligned} \mid \textstyle\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\,dx \mid &\leq \|\mathrm{div}\sigma_h^* + f\|_0\|v\|_0 + \|\sigma_h^* - a_h^*\nabla u_h^*\|_0\|\nabla v\|_0 + \\ &\ \|a_h^*\|_{\infty,\Omega}\|\nabla(u_h^* - \tilde{u})\|_0\|\nabla v\|_0, \end{aligned} \tag{4.2.17}$$

which leads to the following inequality,

$$\mid \int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\,dx \mid\leq c_0\|v\|_0 + c_1\|\nabla v\|_0, \tag{4.2.18}$$

for all $v\in H_0^1(\Omega)$, where $\phi = a_h^* - a$, and

$$c_0 = \|\mathrm{div}\sigma_h^* + f\|_0, \tag{4.2.19}$$

and

$$c_1 = \|\sigma_h^* - a_h^*\nabla u_h^*\|_0 + \|a_h^*\|_{\infty,\Omega}\|\nabla(u_h^* - \tilde{u})\|_0. \tag{4.2.20}$$

Next, we rewrite the left side of the inequality (4.2.18) as a sum of several terms. First, we approximate ϕ by a sequence of smooth functions $\psi_r \in C^\infty(\bar{\Omega})$ with

$$\|\psi_r - \phi\|_1 \to 0 \text{ as } r\to\infty. \tag{4.2.21}$$

Next, for $\delta > 0$ and r fixed, we consider the following test functions,

$$v_1 = \delta^{-1}(\psi_r^+\wedge\delta)\tilde{u},\ v_2 = \delta^{-1}((-\psi_r)^+\wedge\delta)\tilde{u}, \tag{4.2.22}$$

where $\psi_r^+ = \max\{\psi_r, 0\}$, $\psi_r^+\wedge\delta = \min\{\psi_r^+,\delta\}$. Since $\|\psi_r\|_1 \leq C$ as a consequence of (4.2.14) and (4.2.21), we have

$$\begin{aligned} \|v_i\|_0 &\leq \|\tilde{u}\|_0, \\ \|\nabla v_i\|_0 &\leq C(\delta^{-1}+1), \end{aligned} \tag{4.2.23}$$

for $i = 1, 2$. We claim that

$$\int_{\psi_r \geq \delta} \psi_r |\nabla \tilde{u}|^2 \, dx \leq C\delta + \left| \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_1 \, dx \right| \tag{4.2.24}$$

$$-\int_{\psi_r \leq -\delta} \psi_r |\nabla \tilde{u}|^2 \, dx \leq C\delta + \left| \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_2 \, dx \right| . \tag{4.2.25}$$

We will prove the first inequality, since the proof for the second one is similar. The definitions (4.2.22) give

$$\begin{aligned} \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_1 \, dx &= \int_{\psi_r \geq \delta} \psi_r |\nabla \tilde{u}|^2 \, dx + \delta^{-1} \int_{0<\psi_r<\delta} \psi_r^2 |\nabla \tilde{u}|^2 \, dx + \\ & \quad \delta^{-1} \int_{0<\psi_r<\delta} \psi_r \nabla \psi_r \cdot \tilde{u} \nabla \tilde{u} \, dx. \end{aligned} \tag{4.2.26}$$

The second term on the right satisfies

$$\delta^{-1} \int_{0<\psi_r<\delta} \psi_r^2 |\nabla \tilde{u}|^2 \leq C\delta. \tag{4.2.27}$$

In estimating the third term, we shall suppose (without loss of generality) that 0 and δ are regular values of ψ_r. Then

$$\begin{aligned} \delta^{-1} \int_{0<\psi_r<\delta} \psi_r \nabla \psi_r \cdot \tilde{u} \nabla \tilde{u} \, dx &= (4\delta)^{-1} \int_{0<\psi_r<\delta} \nabla(\psi_r^2) \cdot \nabla(\tilde{u}^2) \, dx \\ &= (4\delta)^{-1} \{ \int_{\partial\{0<\psi_r<\delta\}} \psi_r^2 \frac{\partial(\tilde{u}^2)}{\partial \eta} \, ds - \int_{0<\psi_r<\delta} \psi_r^2 \Delta(\tilde{u}^2) \, dx \}. \end{aligned} \tag{4.2.28}$$

where η stands for the unit outer normal Now, the boundary $\partial\{0 < \psi_r < \delta\}$ can be written as follows,

$$\partial\{0 < \psi_r < \delta\} = \Gamma_1 \bigcup \Gamma_2 \bigcup \Gamma_3,$$

where

$$\Gamma_1 = \{\psi_r = 0\} \bigcap \Omega, \Gamma_2 = \{\psi_r = \delta\} \bigcap \Omega, \Gamma_3 = \{0 \leq \psi_r \leq \delta\} \bigcap \partial\Omega.$$

Now, we proceed to estimate each boundary term. Clearly

$$\int_{\Gamma_1} \psi_r^2 \frac{\partial \tilde{u}^2}{\partial \eta} \, ds = 0.$$

As for Γ_2,

$$\begin{aligned} \int_{\Gamma_2} \psi_r^2 \frac{\partial \tilde{u}^2}{\partial \eta} \, ds &= \delta^2 \int_{\partial\{\psi_r<\delta\} \bigcap \Omega} \frac{\partial \tilde{u}^2}{\partial \eta} \, ds, \\ &= \delta^2 \int_{\psi_r<\delta} \Delta \tilde{u}^2 \, dx - \delta^2 \int_{\partial\{\psi_r<\delta\} \bigcap \partial\Omega} \frac{\partial \tilde{u}^2}{\partial \eta} \, ds. \end{aligned} \tag{4.2.29}$$

Hence, we have the following estimations,

$$| \int_{\Gamma_2} \psi_r^2 \frac{\partial \tilde{u}^2}{\partial \eta} \, ds | \leq C\delta^2, \tag{4.2.30}$$

using that $\partial \tilde{u}^2 / \partial \eta$ is bounded on $\partial \Omega$. Similarly, we get

$$\begin{aligned} | \int_{\Gamma_3} \psi_r^2 \frac{\partial \tilde{u}^2}{\partial \eta} \, ds | &\leq C\delta^2, \\ | \int_{0<\psi_r<\delta} \psi_r^2 \Delta(\tilde{u}^2) \, dx | &\leq C\delta^2. \end{aligned} \tag{4.2.31}$$

Combining all these results with (4.2.28) gives:

$$-\delta^{-1} \int_{0<\psi_r<\delta} \psi_r \nabla \psi_r \cdot \tilde{u} \nabla \tilde{u} \, dx \leq C\delta. \tag{4.2.32}$$

Putting (4.2.27), (4.2.32) together, we get:

$$\int_{\psi_r \geq \delta} \psi_r |\nabla \tilde{u}|^2 \, dx \leq C\delta + | \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_1 \, dx |$$

Assembling (4.2.24) and (4.2.25) with the obvious result:

$$\int_{|\psi_r|<\delta} |\psi_r| |\nabla \tilde{u}|^2 \, dx \leq C\delta,$$

we conclude that

$$\int_\Omega |\psi_r| |\nabla \tilde{u}|^2 \, dx \leq C\{\delta + | \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_1 \, dx | + | \int_\Omega \psi_r \nabla \tilde{u} \cdot \nabla v_2 \, dx |\}.$$

By taking the limit of the expression above when $r \to \infty$, we get:

$$\int_\Omega |\phi| |\nabla \tilde{u}|^2 \, dx \leq C\{\delta + | \int_\Omega \phi \nabla \tilde{u} \cdot \nabla v_1 \, dx | + | \int_\Omega \phi \nabla \tilde{u} \cdot \nabla v_2 \, dx |\}.$$

Applying (4.2.18) with $v = v_1$ and $v = v_2$, and using (4.2.23), we conclude that

$$\int_\Omega |\phi| |\nabla \tilde{u}|^2 \, dx \leq C\{\delta + c_0 + c_1(\delta^{-1} + 1)\},$$

which concludes the proof of the Theorem.

We can also derive an error estimate in the H^{-1} norm for a_h^*. According to the definition, we have

$$\begin{aligned} \|\mathrm{div}((\tilde{a} - a_h^*)\nabla u)\|_{-1} &= \sup_{v \in H_0^1(\Omega)} \frac{|(\mathrm{div}((\tilde{a} - a_h^*)\nabla \tilde{u}), v)|}{\|v\|_1}, \\ &= \sup_{v \in H_0^1(\Omega)} \frac{1}{\|v\|_1} | ((\tilde{a} - a_h^*)\nabla \tilde{u}, \nabla v) |, \\ &\leq \sup_{v \in H_0^1(\Omega)} \frac{1}{\|v\|_1} (c_0 \|v\|_0 + c_1 \|\nabla v\|_0), \\ &\leq c_0 + c_1, \end{aligned}$$

where c_0 and c_1 are given by (4.2.19) and (4.2.20), respectively. The following theorem follows from the previous theorem.

Theorem 4.2.2 *Under the same assumptions as in Lemma (4.2.1), we have the following weighted error estimate in the H^{-1} norm,*

$$\|\mathrm{div}((\tilde{a}-a_h^*)\nabla\tilde{u})\|_{-1} \leq C(\frac{1}{\sqrt{\alpha}}+\frac{1}{\sqrt{\beta}}+\frac{1}{h})(h^{d_u+1}+\epsilon+\sqrt{\alpha}(h^{d_\sigma+1}+h^{d_a+1}+h^{d_u})+\sqrt{\beta}h^{d_\sigma}).$$

4.3 The linear BVP with a mixed boundary condition

In this section, we are concerned with the identification of the unknown coefficient $a(x)$ for the elliptic boundary value problem with mixed boundary conditions

$$\left\{\begin{array}{rcl} -\mathrm{div}(a(x)\nabla u(x)) & = & f(x), \quad \text{in } \Omega, \\ u\mid_{\Gamma_0} & = & 0, \\ a(x)\dfrac{\partial u}{\partial \eta}\mid_{\Gamma_1} & = & g, \end{array}\right. \tag{4.3.33}$$

where $f \in L^2(\Omega)$, and $\lambda \leq a(x) \leq \Lambda$ for some positive constants λ and Λ. Here $\partial\Omega = \bar{\Gamma}_0 \bigcup \bar{\Gamma}_1$, where Γ_0 and Γ_1 are open disjoint subsets of $\partial\Omega$ with positive measure each. The PI problem is to recover a parameter $a = a(x)$ within some suitable set of admissible elements, from an L^2 observation $z(x)$ of $u(x)$ such that $a(x)$ and $u(x)$ satisfy (4.3.33).

To describe the method, we let

$$\tilde{H}^1(\Omega) = \{v \in H^1(\Omega) : v\mid_{\Gamma_0} = 0\}.$$

be the standard Sobolev space. We use the same notation for its norm and seminorm. Since we are going to use the mixed least-squares ideas again, we introduce the flux variable $\sigma(x) = a(x)\nabla u(x)$ as an additional unknown, and reduce the given linear differential equation in (4.3.33) to the equivalent linear system in (4.2.6).

As before, we try to find an approximation to the solution of the PI problem by minimizing the following cost functional,

$$\begin{array}{rcl} \mathcal{J}(a,u,\sigma) & = & \|u-z\|_0^2+\alpha\|\sigma-a\nabla u\|_0^2+\beta\|\mathrm{div}\sigma+f\|_0^2+ \\ & & \gamma\|a\|_{r_a}^2+\rho\|u\|_{r_u}^2+\delta\|\sigma\|_{r_\sigma}^2+ \\ & & \xi\|g-a\nabla u\cdot\eta\|_{L^2(\Gamma_1)}^2+\mu\|(\sigma-a\nabla u)\cdot\eta\|_{L^2(\Gamma_1)}^2, \end{array} \tag{4.3.34}$$

over the space

$$\mathcal{H} = H^1(\Omega) \times (\tilde{H}^1(\Omega) \cap W^{1,\infty}(\Omega)) \times (H^1(\Omega) \times H^1(\Omega)), \tag{4.3.35}$$

where the integers r_a, r_u, and r_σ depend on the Sobolev spaces in which the regularization takes place. The nonnegative numbers $\alpha, \beta, \gamma, \rho, \delta, \xi, \mu$ are the regularization parameters as in the previous Chapters. The meaning of the first six terms in (4.3.34) is the same as in the previous chapters, but now the boundary conditions are incorporated in the last two terms in (4.3.34). Note that if the data is perfect, i.e., there exist $u(x)$ and $a(x)$ satisfying (4.3.33) such that $z(x) = u(x)$, then $(a, u, \sigma)^T$ with $\sigma = a\nabla u$ can make zero the first three and the last two terms of (4.3.34), and makes $\mathcal{J}(a, u, \sigma)$ small if the regularization parameters are small. In other words, minimizing the cost functional may give an approximation to a solution of the PI problem.

4.3.1 An estimate for error $|v - \Pi_T v|_{m,q,\partial\Omega}$

At this point, we intend to give an estimation of the interpolation error $|v - \Pi_T v|_{m,q,\partial\Omega}$ for an affine family of finite elements. Since we are going to work with a triangle $\hat{T}$, it is easy to show that $\hat{T}$ is a domain which has a Lipschitz continuous boundary $\partial\hat{T}$. For domains $\Omega \subset I\!R^n$ with Lipschitz continuous boundaries, there is a continuous embedding (see [22], page 151)

$$W^{k,p}(\Omega) \subset C^m_B(\Omega) \text{ for } 0 \leq m < k - \frac{n}{p}$$

where $C^m_B(\Omega) = \{u \in C^m(\Omega) : D^\alpha u \in L^\infty(\Omega) \text{ for } |\alpha| \leq m\}$. In our situation, the following inclusion holds:

$$W^{k+1,p}(\hat{T}) \subset C^0_B(\hat{T})$$

provided that $0 \leq m = 0 < k + 1 - \frac{2}{p}$, which is always true with $p = 2$ in $I\!R^2$. The continuous embedding

$$W^{k+1,p}(\hat{T}) \subset W^{m,q}(\hat{T})$$

holds with $q = \infty$, $m = 0, 1$, $k > 1$, and $p = 2$ in $I\!R^2$. The double inclusion

$$P_k(\hat{T}) \subset \hat{P} \subset W^{m,q}(\hat{T})$$

is always true for $m = 0, 1$, $q = \infty$ in $I\!R^2$. Then by Theorem 16.1 on page 126 in [15], there exists a constant $C(\hat{T}, \hat{P}, \hat{\Sigma})$ such that, for all affine equivalent finite elements (T, P, Σ) and all functions $v \in W^{k+1,p}(T)$,

$$\|v - \Pi_T v\|_{m,q,T} \leq C\{\text{meas}(T)\}^{1/q-1/p} \frac{h_T^{k+1}}{\rho_T^m} \|v\|_{k+1,p,T},$$

where $\Pi_T v$ denotes the P_T-interpolant of the function v, and

$$\begin{aligned} \mathrm{meas(T)} &= \text{ dx-measure of T}, \\ \mathrm{h_T} &= \mathrm{diam(T)}, \\ \rho_{\mathrm{T}} &= \sup\{\ \mathrm{diam(S); S \text{ is a ball contained in T}}\ \}. \end{aligned} \tag{4.3.36}$$

If we are working with a regular family of finite elements, then the above inequality reduces to

$$\|v - \Pi_T v\|_{m,q,T} \le C\{\mathrm{meas}(T)\}^{1/q-1/p} h_T^{k+1-m} \|v\|_{k+1,p,T}.$$

For the particular case $n = 2$, $p = 2$, $q = \infty$, we are left with

$$\|v - \Pi_T v\|_{m,\infty,T} \le C h^{k-m} \|v\|_{k+1,2,T},$$

Since we are working with a continuous finite element, we can extend the inequality above all the way up to the boundary,

$$\|v - \Pi_h v\|_{m,\infty,\bar{\Omega}} \le C h^{k-m} \|v\|_{k+1,2,\Omega},$$

In particular, over the boundary $\Gamma = \partial\Omega$ we have,

$$\|v - \Pi_h v\|_{m,\infty,\Gamma} \le C h^{k-m} \|v\|_{k+1},$$

Therefore, for $m = 0, 1$, the inequality

$$\|v - \Pi_h v\|_{m,2,\Gamma} \le C h^{k-m} \|v\|_{k+1},$$

holds for all $v \in H^{k+1}(\Omega)$.

4.3.2 Finite Element Method

In this section, we formulate a finite element discretization of the cost functional and discuss the related minimization problem. In order to define the finite element spaces, let $\{\Delta_h\}$, with $h \in (0, 1)$, be a family of regular, quasi-uniform partitions of $\Omega \subset I\!R^n$. We consider the finite element spaces $S^r_{h,l}$ as in the previous section. By $S^{r,0}_{h,l}$, we denote the subspace of functions of $S^r_{h,l}$ which vanish on Γ_0. For all $v \in W^{m,p}(\Omega)$, there is (an interpolant) $v_h \in S^r_{h,l}$ such that the inequalities (2.2.15), (2.2.16) and (2.2.17) hold. We denote by $\tilde{H}^{-1} = \tilde{H}^{-1}(\Omega)$ the dual space $(\tilde{H}^1(\Omega))^*$ equipped with the norm

$$\|v\|_{-1} = \sup_{0 \ne \psi \in \tilde{H}^1(\Omega)} \frac{|(v, \psi)|}{\|\psi\|_1}. \tag{4.3.37}$$

Let us now introduce those finite dimensional spaces needed to define the computational procedure. We need three different discretization spaces, namely, U_h for the solution u, A_h for the parameter a, and finally Σ_h which will be used to discretize the flux variable σ. We assume that these finite element spaces are,

$$\begin{array}{rcl} A_h & = & S_{h,l}^{d_a}, \\ U_h & = & S_{h,1}^{d_u} \cap \tilde{H}^1(\Omega), \\ \Sigma_h & = & S_{h,1}^{d_\sigma} \times S_{h,1}^{d_\sigma}, \end{array} \tag{4.3.38}$$

where d_a, d_u, d_σ are integers.

Now we are ready to define the numerical method to solve the PI problem:
(P_h) seek $a_h^* \in A_h$, $u_h^* \in U_h$, $\sigma_h^* \in \Sigma_h$, such that

$$\mathcal{J}_h(a_h^*, u_h^*, \sigma_h^*) = \min_{a \in A_h, u \in U_h, \sigma \in \Sigma_h} \mathcal{J}_h(a, u, \sigma), \tag{4.3.39}$$

where $\mathcal{J}_h$ is defined by,

$$\begin{array}{rcl} \mathcal{J}_h(\mathbf{w}_h) & = & \|u_h - z\|_0^2 + \alpha\|\sigma_h - a_h \nabla u_h\|_0^2 + \beta\|\mathrm{div}\sigma_h + f\|_0^2 + \\ & & \gamma\|a_h\|_{r_a}^2 + \rho\|u_h\|_{r_u}^2 + \delta\|\sigma_h\|_{r_\sigma}^2 + \\ & & \xi\|g - a_h \nabla u_h \cdot \eta\|_{L^2(\Gamma_1)}^2 + \mu\|(\sigma_h - a_h \nabla u_h) \cdot \eta\|_{L^2(\Gamma_1)}^2, \end{array} \tag{4.3.40}$$

where $\mathbf{w}_h = (a_h, u_h, \sigma_h) \in A_h \times U_h \times \Sigma_h$.

4.3.3 Some error estimates in weighted norms

In this section, we establish some error estimates for an approximate solution to the PI problem generated by the finite element method in the previous section. In all the discussion below, C denotes a generic constant which may take a different value at different places and is independent of h unless specified otherwise. Let I_a^h, I_u^h, and I_σ^h be the interpolation operators in the finite element spaces A_h, U_h, and Σ_h, respectively. The regularization spaces are chosen so that

$$1 \leq r_a \leq d_a + 1, 1 \leq r_u \leq d_u + 1, 1 \leq r_\sigma \leq d_\sigma + 1.$$

Moreover, there exists a constant C such that the interpolation operators satisfy,

$$\|I_a^h a\|_{r_a} \leq C\|a\|_{r_a+1}, \|I_u^h u\|_{r_u} \leq C\|u\|_{r_u+1}, \|I_\sigma^h \sigma\|_{r_\sigma} \leq C\|\sigma\|_{r_\sigma+1},$$

for any $a \in H^{r_a+1}(\Omega), u \in H^{r_u+1}(\Omega), \sigma \in (H^{r_\sigma+1}(\Omega) \times H^{r_\sigma+1}(\Omega))$. We first derive an estimate for a minimizer of $\mathcal{J}_h(\mathbf{w}_h)$.

Lemma 4.3.1 *Let us assume the following hypothesis,*

- $*$ *Let $z(x)$ be a distributed L^2 observation of the state $\tilde{u}$ with an observation error of the form*
$$\|\tilde{u} - z\|_0 \leq \epsilon.$$
- $*$ *Let $\tilde{a}$, $\tilde{u}$ satisfy the boundary value problem (4.3.33) which have the following regularities:*
$$\tilde{a} \in H^{d_a+1}(\Omega), \tilde{u} \in H^{d_u+1}(\Omega) \bigcap W^{1,\infty}(\bar{\Omega}), \tilde{\sigma} \in H^{d_\sigma+1}(\Omega) \times H^{d_\sigma+1}(\Omega).$$
- $*$ *The finite element space U_h is such that*
$$\|\nabla(I_u^h \tilde{u})\|_\infty \leq C,$$
for h small enough.
- $*$ *The regularization parameters γ, ρ, and δ are chosen such that*
$$\begin{aligned} \gamma + \rho + \delta \quad \sim \quad & C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma} + \\ & \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\}. \end{aligned} \tag{4.3.41}$$
with C independent of h.

Let $\mathbf{w}_h^* = (a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h(a, u, \sigma)$ defined by (4.3.40) in $A_h \times U_h \times \Sigma_h$, then we have the following estimate for the cost functional $\mathcal{J}_h$,

$$\begin{aligned} \mathcal{J}_h(\mathbf{w}_h^*) \quad \leq \quad & C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2(d_a+1)} + h^{2d_u}) + \beta h^{2d_\sigma} + \\ & \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\}, \end{aligned} \tag{4.3.42}$$

here C is independent of ϵ and h.

Proof. Let $(a_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h$, which will be denoted by $\mathbf{w}_h^*$, and let $I_a^h \tilde{a}$, $I_u^h \tilde{u}$, and $I_\sigma^h \tilde{\sigma}$ be the interpolates of $\tilde{a}$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}\nabla\tilde{u}$ in A_h, U_h, and Σ_h, respectively. Now

$$\mathcal{J}_h(\mathbf{w}_h^*) \leq \mathcal{J}_h(I_a^h \tilde{a}, I_u^h \tilde{u}, I_\sigma^h \tilde{\sigma}),$$

and

$$\begin{aligned} \mathcal{J}_h(I_a^h \tilde{a}, I_u^h \tilde{u}, I_\sigma^h \tilde{\sigma}) \quad = \quad & \|I_u^h \tilde{u} - z\|_0^2 + \alpha\|I_\sigma^h \tilde{\sigma} - I_a^h \tilde{a}\nabla(I_u^h \tilde{u})\|_0^2 + \beta\|\mathrm{div} I_\sigma^h \tilde{\sigma} + f\|_0^2 + \\ & \gamma\|I_a^h \tilde{a}\|_{r_a}^2 + \rho\|I_u^h \tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h \tilde{\sigma}\|_{r_\sigma}^2 + \\ & \xi\|g - I_a^h \tilde{a}\nabla(I_u^h \tilde{u}) \cdot \eta\|_{L^2(\Gamma_1)}^2 + \mu\|(I_\sigma^h \tilde{\sigma} - I_a^h \tilde{a}\nabla(I_u^h \tilde{u})) \cdot \eta\|_{L^2(\Gamma_1)}^2. \end{aligned}$$

The first three terms of the equality above have the following bounds,

$$\begin{aligned}
\|I_u^h\tilde{u} - z\|_0 &\leq \|I_u^h\tilde{u} - \tilde{u}\|_0 + \|\tilde{u} - z\|_0 \\
&\leq Ch^{d_u+1}\|\tilde{u}\|_{d_u+1} + \epsilon, \\
\|I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\|_0 &\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_0 + \|\tilde{a}\nabla\tilde{u} - \tilde{a}\nabla(I_u^h\tilde{u})\|_0 + \|\tilde{a}\nabla(I_u^h\tilde{u}) - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\|_0 \\
&\leq Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + \|\tilde{a}\|_\infty\|\nabla(\tilde{u} - I_u^h\tilde{u})\|_0 + \|\nabla(I_u^h\tilde{u})\|_\infty\|\tilde{a} - I_a^h\tilde{a}\|_0 \\
&\leq Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + Ch^{d_u}\|\tilde{u}\|_{d_u+1} + Ch^{d_a+1}\|\tilde{a}\|_{d_a+1}, \\
\|\mathrm{div} I_\sigma^h\tilde{\sigma} + f\|_0 &= \|\mathrm{div}(I_\sigma^h\tilde{\sigma} - \tilde{\sigma})\|_0 \\
&\leq Ch^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1},
\end{aligned}$$

For the next two bounds, we use the results from section (4.3.1),

$$\begin{aligned}
\|g - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma_1)} &= \|(\tilde{a}\nabla\tilde{u}\cdot\eta - I_a^h\tilde{a}\nabla\tilde{u}\cdot\eta) + (I_a^h\tilde{a}\nabla\tilde{u}\cdot\eta - I_a^h\tilde{a}\nabla(I_u^h\tilde{u})\cdot\eta)\|_{L^2(\Gamma_1)} \\
&\leq \|(\tilde{a} - I_a^h\tilde{a})\frac{\partial\tilde{u}}{\partial\eta}\|_{L^2(\Gamma_1)} + \|I_a^h\tilde{a}(\nabla(\tilde{u} - I_u^h\tilde{u})\cdot\eta)\|_{L^2(\Gamma_1)} \\
&\leq C\{\|\tilde{a} - I_a^h\tilde{a}\|_{L^2(\Gamma_1)}\|\frac{\partial\tilde{u}}{\partial\eta}\|_{\infty,\Gamma} + \|\nabla(\tilde{u} - I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma_1)}\|I_a^h\tilde{a}\|_\infty\} \\
&\leq C\{\|\tilde{a} - I_a^h\tilde{a}\|_{L^2(\Gamma)} + \|\nabla(\tilde{u} - I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma)}\} \\
&\leq C\{h^{d_a}\|\tilde{a}\|_{d_a+1} + h^{d_u-1}\|\tilde{u}\|_{d_u+1}\} \\
&\leq C\{h^{d_a} + h^{d_u-1}\},
\end{aligned}$$

and

$$\begin{aligned}
\|(I_\sigma^h\tilde{\sigma} - I_a^h\tilde{a}\nabla(I_u^h\tilde{u}))\cdot\eta\|_{L^2(\Gamma_1)} &\leq \|(I_\sigma^h\tilde{\sigma} - \tilde{\sigma})\cdot\eta\|_{L^2(\Gamma_1)} + \|(\tilde{a} - I_a^h\tilde{a})\nabla\tilde{u}\cdot\eta\|_{L^2(\Gamma_1)} + \\
&\quad \|I_a^h\tilde{a}\nabla(\tilde{u} - I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma_1)} \\
&\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_{L^2(\Gamma_1)} + \|\tilde{a} - I_a^h\tilde{a}\|_{L^2(\Gamma_1)}\|\frac{\partial\tilde{u}}{\partial\eta}\|_{\infty,\Gamma} + \\
&\quad \|\nabla(\tilde{u} - I_u^h\tilde{u})\|_{L^2(\Gamma_1)}\|I_a^h\tilde{a}\|_\infty \\
&\leq C\{\|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_{L^2(\Gamma)} + \|\tilde{a} - I_a^h\tilde{a}\|_{L^2(\Gamma)} + \|\nabla(\tilde{u} - I_u^h\tilde{u})\|_{L^2(\Gamma)}\} \\
&\leq C\{h^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1} + h^{d_a}\|\tilde{a}\|_{d_a+1} + h^{d_u-1}\|\tilde{u}\|_{d_u+1}\} \\
&\leq C\{h^{d_\sigma} + h^{d_a} + h^{d_u-1}\}
\end{aligned}$$

in the above two estimations on the boundary Γ, we have used that $\partial\tilde{u}/\partial\eta$ and $I_a^h\tilde{a}$ are bounded on $\partial\Omega$, where η is the outer unit normal to Γ. Then,

$$\begin{aligned}
\mathcal{J}_h(a_h^*, u_h^*, \sigma_h^*) &\leq C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2(d_a+1)} + h^{2d_u}) + \beta h^{2d_\sigma} + \\
&\quad \gamma\|I_a^h\tilde{a}\|_{r_a}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2 + \\
&\quad \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\}.
\end{aligned}$$

Since

$$\begin{aligned}
\|I_a^h\tilde{a}\|_{r_a} &\leq C\|\tilde{a}\|_{r_a+1}, \\
\|I_u^h\tilde{u}\|_{r_u} &\leq C\|\tilde{u}\|_{r_u+1}, \\
\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma} &\leq C\|\tilde{\sigma}\|_{r_\sigma+1},
\end{aligned}$$

and choosing γ, ρ, and δ as in the hypothesis, we get the main assertion of the lemma:

$$\begin{aligned}
\mathcal{J}_h(\mathbf{w}_h^*) &\leq C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2(d_a+1)} + h^{2d_u}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\}
\end{aligned}$$

The proof of the lemma is completed.

Lemma 4.3.2 *Under the same hypothesis as in Lemma 4.3.1, we get the following bounds,*

$$\begin{aligned}\|\sigma_h^* - a_h^*\nabla u_h^*\|_0 &\leq \frac{C}{\sqrt{\alpha}}\{h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})\},\end{aligned}$$

$$\begin{aligned}\|\mathrm{div}\sigma_h^* + f\|_0 &\leq \frac{C}{\sqrt{\beta}}\{h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})\},\end{aligned}$$

$$\begin{aligned}\|g - a_h^*\nabla u_h^* \cdot \eta\|_{L^2(\Gamma_1)} &\leq \frac{C}{\sqrt{\xi}}\{h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})\},\end{aligned}$$

$$\begin{aligned}\|(\sigma_h^* - a_h^*\nabla u_h^*) \cdot \eta\|_{L^2(\Gamma_1)} &\leq \frac{C}{\sqrt{\mu}}\{h^{d_u+1} + \epsilon + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})\},\end{aligned}$$

Proof: This follows immediately from the definition of $\mathcal{J}_h(a, u, \sigma)$ and the previous lemma.

Corollary 4.3.1 *If all the conditions of Lemma 4.3.1 hold, and*

$$\begin{aligned}0 < \gamma &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma} + \\ &\quad \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\},\\ 0 < \rho &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma} + \\ &\quad \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\},\\ 0 < \delta &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_a+1)} + h^{2d_u} + h^{2(d_\sigma+1)}) + \beta h^{2d_\sigma} + \\ &\quad \xi(h^{2d_a} + h^{2(d_u-1)}) + \mu(h^{2d_\sigma} + h^{2d_a} + h^{2(d_u-1)})\}.\end{aligned}$$

then there exists a constant C such that,

$$\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C, \text{ for all } h.$$

Proof. The proof is identical to that in Corollary 4.2.1.

Corollary 4.3.2 *Assume the hypothesis of Corollary 4.3.1 and take H^2 to be the Sobolev space for the regularization of a_h. If $A_h = S_{h,l}^{d_a}$ with $l \geq 2$, then*

$$\|a_h^*\|_\infty < \infty.$$

Proof. The proof is identical to that in Corollary 4.2.2.

We state a preliminary result that we will need in the next Theorem.

Lemma 4.3.3 *If $\psi \in H^1(\Omega)$, then for any triangle $T \in \Delta_h$,*

$$\|\psi\|_{L^2(\partial T)} \leq C\{h^{-1/2}\|\psi\|_{L^2(T)} + h^{1/2}\|\nabla\psi\|_{L^2(T)}\},$$

where C is independent of h and T. Additionally,

$$\|\psi\|_{L^2(\partial\Omega)} \leq C\{h^{-1/2}\|\psi\|_0 + h^{1/2}\|\nabla\psi\|_0\},$$

Proof. Both inequalities are standard results from the theory of finite elements, they can be found in [13] for polygonal domains.

In our next Theorem, we obtain a weighted H^{-1} norm estimation for the error $\phi(x) = a_h^*(x) - \tilde{a}(x)$, when $g = 0$ in (4.3.33).

Theorem 4.3.1 *Suppose that all the conditions in Corollary 4.3.2 hold and $g = 0$ in (4.3.33). If $\mathbf{w}_h^* = (a_h^*, u_h^*, \sigma_h^*)$ is a minimizer of $\mathcal{J}_h(a, u, \sigma)$, then we get the following error estimation in the weighted H^{-1} norm,*

$$\begin{aligned}\|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1} \;\leq\; & C(\frac{1}{h} + \frac{1}{\sqrt{\alpha}} + \frac{1}{\sqrt{\beta}} + (\frac{1}{\sqrt{\xi}} + \frac{1}{\sqrt{\mu}}) \times (h^{-1/2} + h^{1/2}))\times \\ & (\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ & \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})).\end{aligned}$$

Proof. Due to the boundary conditions satisfied by $\tilde{u}$ in (4.3.33) and integration by parts, the following holds

$$\begin{aligned}\|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1} &= \sup_{v\in\tilde{H}^1(\Omega)} \frac{|(\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u}), v)|}{\|v\|_1}, \\ &= \sup_{v\in\tilde{H}^1(\Omega)} \frac{|((\tilde{a} - a_h^*)\nabla\tilde{u}, \nabla v)|}{\|v\|_1},\end{aligned}$$

The equality below

$$((\tilde{a}-a_h^*)\nabla\tilde{u},\nabla v)=(\tilde{a}\nabla\tilde{u}-a_h^*\nabla u_h^*,\nabla v)+(a_h^*\nabla(u_h^*-\tilde{u}),\nabla v),$$

holds for all $v\in\tilde{H}^1(\Omega)=\{v\in H^1(\Omega): v\mid_{\Gamma_0}=0\}$. Now we estimate the term

$$(\tilde{a}\nabla\tilde{u}-a_h^*\nabla u_h^*,\nabla v),$$

when $v\in\tilde{H}^1(\Omega)$. By using the weak form of equation (4.3.33), adding and subtracting various terms and integrating by parts, we get

$$\begin{aligned}(\tilde{a}\nabla\tilde{u}-a_h^*\nabla u_h^*,\nabla v) &= ((f+\mathrm{div}\sigma_h^*),v)+((\sigma_h^*-a_h^*\nabla u_h^*),\nabla v)\\ &\quad -((\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta,v)_{L^2(\Gamma_1)}-(a_h^*\nabla u_h^*\cdot\eta,v)_{L^2(\Gamma_1)},\end{aligned}$$

which holds for all test functions $v\in\tilde{H}^1(\Omega)$. Then

$$\begin{aligned}|(\tilde{a}\nabla\tilde{u}-a_h^*\nabla u_h^*,\nabla v)| &\leq \|f+\mathrm{div}\sigma_h^*\|_0\|v\|_0+\|\sigma_h^*-a_h^*\nabla u_h^*\|_0\|\nabla v\|_0+\\ &\quad C\{h^{-1/2}\|v\|_0\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}+\\ &\quad h^{1/2}\|\nabla v\|_0\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}\}+\\ &\quad h^{-1/2}\|v\|_0\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}+h^{1/2}\|\nabla v\|_0\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}\},\end{aligned}$$

whence,

$$\begin{aligned}|(\tilde{a}\nabla\tilde{u}-a_h^*\nabla u_h^*,\nabla v)| &\leq C\{\|f+\mathrm{div}\sigma_h^*\|_0+\|\sigma_h^*-a_h^*\nabla u_h^*\|_0+\\ &\quad h^{-1/2}\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}+h^{1/2}\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}+\\ &\quad h^{-1/2}\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}+h^{1/2}\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}\}\|v\|_1,\end{aligned}$$

Let $I_u^h\tilde{u}$ be the interpolant of $\tilde{u}$ in U_h. By the inverse inequality (2.2.16), we have

$$\begin{aligned}|(a_h^*\nabla(u_h^*-\tilde{u}),\nabla v)| &\leq \|a_h^*\nabla(u_h^*-\tilde{u})\|_0\|\nabla v\|_0\\ &\leq C(\|\nabla(u_h^*-I_u^h\tilde{u})\|_0+\|\nabla(I_u^h\tilde{u}-\tilde{u})\|_0)\|\nabla v\|_0\\ &\leq C(\|\nabla(u_h^*-I_u^h\tilde{u})\|_0+h^{d_u}\|\tilde{u}\|_{d_u+1})\|\nabla v\|_0\\ &\leq C(\frac{1}{h}\|u_h^*-I_u^h\tilde{u}\|_0+h^{d_u})\|\nabla v\|_0\\ &\leq C(\frac{1}{h}(\|u_h^*-z\|_0+\|z-\tilde{u}\|_0+\|\tilde{u}-I_u^h\tilde{u}\|_0)+h^{d_u})\|\nabla v\|_0\\ &\leq C(\frac{1}{h}\mathcal{J}_h^{1/2}(a_h^*,u_h^*,\sigma_h^*)+\frac{\epsilon}{h}+h^{d_u})\|\nabla v\|_0.\end{aligned}$$

Therefore,

$$\begin{aligned}|((\tilde{a}-a_h^*)\nabla\tilde{u},\nabla v)| &\leq C\{\|f+\mathrm{div}\sigma_h^*\|_0+\|\sigma_h^*-a_h^*\nabla u_h^*\|_0+\\ &\quad h^{-1/2}\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}+h^{1/2}\|(\sigma_h^*-a_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma_1)}+\\ &\quad h^{-1/2}\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}+h^{1/2}\|a_h^*\nabla u_h^*\cdot\eta\|_{L^2(\Gamma_1)}+\\ &\quad \frac{1}{h}\mathcal{J}_h^{1/2}(a_h^*,u_h^*,\sigma_h^*)+\frac{\epsilon}{h}+h^{d_u}\}\|v\|_1.\end{aligned}$$

Finally, applying the estimates in the previous Lemma to the above inequality leads to the estimate of this theorem.

4.3.4 An H^{-1} norm error estimation in 2d

According to the definition of the H^{-1} norm, we have that for the boundary value problem (4.3.33) with $g = 0$ and Γ_0 being the inflow boundary (i.e., $\dfrac{\partial \tilde{u}}{\partial \eta} < 0$ where η is the unit outer normal to $\partial\Omega$), the following inequality holds,

$$| \textstyle\int_\Omega (\tilde{a} - a_h^*)\nabla\tilde{u} \cdot \nabla v \, dx \; | \quad \leq \quad \|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1}\|v\|_1, \tag{4.3.43}$$

for all $v \in \tilde{H}^1(\Omega)$. In order to obtain an H^{-1} norm error estimation, we proceed in the following way. First, we have to make the following assumptions:

(A_1) Ω has to be a bounded Lipschitz domain with smooth boundary $\partial\Omega$.

(A_2) $\tilde{u} \in C^2(\bar{\Omega})$.

(A_3) $\inf_\Omega \|\nabla\tilde{u}\| > 0$.

(A_4) The inflow boundary Γ_0 and the outflow boundary Γ_1 (i.e., $\dfrac{\partial \tilde{u}}{\partial \eta} > 0$) have to be smooth and both can not be disconnected on $\partial\Omega$ in order to prevent $\tilde{u}$ from being discontinuous in Ω.

The next technical Lemma has been taken from [20]

Lemma 4.3.4 *For each $\psi \in H^1(\Omega)$, there exists a unique $v_\psi \in H^1(\Omega)$ satisfying the hyperbolic initial value problem*

$$\left\{ \begin{array}{rcl} \nabla\tilde{u} \cdot \nabla v & = & \psi, \quad \mathit{in}\ \Omega, \\ v\,|_{\Gamma_0} & = & 0. \end{array} \right. \tag{4.3.44}$$

where Γ_0 is the inflow boundary, along with the estimate

$$\|v_\psi\|_1 \leq \|\psi\|_1,$$

where C is independent of ψ and v_ψ.

Sketch of the proof: Since $\dfrac{\partial \tilde{u}}{\partial \eta} > 0$ on Γ_0, Γ_0 is not a characteristic. Hence for Γ_0 and $\tilde{u}$ sufficiently smooth, we get a unique smooth solution of this initial value problem. Since $\|\nabla\tilde{u}\| \geq \alpha > 0$ for all $x \in \Omega$, we can take as local coordinates $\tilde{u}$ and v, where v is a coordinate along the lines $\tilde{u}$ = constant. Write Γ_0 in the form $u = G(v)$. Then the solution is given by,

$$v_\psi(\tilde{u}, v) = \int_{G(v)}^{\tilde{u}} \frac{\psi_{(s,v)}}{\|\nabla\tilde{u}\|^2_{(s,v)}}\, ds$$

All the constants C after now are independent of v_ψ and ψ. By looking at Ω in the $(\tilde{u}, v)$-plane and using Hölder's inequality, the given hypothesis on $\tilde{u}$ and ψ yield that

$$\|v_\psi\|_0 \leq C\|\psi\|_0$$

Secondly, $\dfrac{\partial v_\psi}{\partial x}$ is given by

$$\frac{\partial v_\psi}{\partial x} = \int_{G(v)}^{\tilde{u}} \frac{\partial}{\partial x}\Big(\frac{\psi_{(s,v)}}{\|\nabla\tilde{u}\|^2_{(s,v)}}\Big)\, ds + \frac{\psi_{(\tilde{u},v)}}{\|\nabla\tilde{u}\|^2_{(\tilde{u},v)}}\frac{\partial\tilde{u}}{\partial x} - \frac{\psi_{(G(v),v)}}{\|\nabla\tilde{u}\|^2_{(G(v),v)}}\frac{\partial G(v)}{\partial x}$$

By using the hypothesis that $\tilde{u} \in C^2(\bar{\Omega})$ and Hölder's inequality, the $L^2(\Omega)$-norm of the first two terms in the expansion for $\dfrac{\partial v_\psi}{\partial x}$ are both bounded by $\|\psi\|_1$ up to a constant which is independent of v_ψ and ψ. The $L^2(\Omega)$-norm of the last term is bounded by $\|\psi\|_1$ by a Trace Theorem. Working in a similar way, it can be shown that

$$\|\frac{\partial v_\psi}{\partial y}\|_0 \leq C\|\psi\|_1,$$

holds. Finally putting all these estimations together, we get

$$\|\nabla v_\psi\|_0 \leq C\|\psi\|_1$$

which concludes with the proof.

Substituting $\nabla\tilde{u}\cdot\nabla v$ for ψ in (4.3.43) and using the H^1 norm estimation above, we get that for all $\psi \in H^1(\Omega)$, the following holds;

$$\begin{aligned} \Big| \int_\Omega(\tilde{a} - a_h^*)\psi\, dx \Big| &\leq \|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1}\|v_\psi\|_1, \\ &\leq C\|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1}\|\psi\|_1. \end{aligned} \tag{4.3.45}$$

Dividing both sides by $\|\psi\|_1$ and taking the supremum over all such ψ, we get the H^{-1} norm estimation,

$$\|(\tilde{a} - a_h^*)\|_{-1} \leq C\|\mathrm{div}((\tilde{a} - a_h^*)\nabla\tilde{u})\|_{-1}. \tag{4.3.46}$$

Theorem 4.3.2 *Under the same conditions as in Theorem (4.3.1), we have the following H^{-1} norm error estimation:*

$$\begin{aligned} \|\tilde{a} - a_h^*\|_{-1} \leq\ & C\Big(\frac{1}{h} + \frac{1}{\sqrt{\alpha}} + \frac{1}{\sqrt{\beta}} + \Big(\frac{1}{\sqrt{\xi}} + \frac{1}{\sqrt{\mu}}\Big) \times (h^{-1/2} + h^{1/2})\Big)\times \\ & (\epsilon + h^{d_u+1} + \sqrt{\alpha}(h^{d_\sigma+1} + h^{d_a+1} + h^{d_u}) + \sqrt{\beta}h^{d_\sigma} + \\ & \sqrt{\xi}(h^{d_a} + h^{d_u-1}) + \sqrt{\mu}(h^{d_\sigma} + h^{d_a} + h^{d_u-1})). \end{aligned}$$

Proof. The proof follows immediately from the inequality (4.3.46) and Theorem (4.3.1).

4.4 A nonlinear BVP with the Neumann boundary condition

In this section we are concerned with the identification of the unknown nonlinear diffusion coefficient $a(u)$ for the nonlinear elliptic boundary value problem with homogeneous Newmann boundary condition:

$$\begin{cases} -\mathrm{div}(a(u)\nabla u(x)) &= f(x), \quad \text{in } \Omega \subset I\!R^2 \\ a(u)\dfrac{\partial u}{\partial \eta}\,|_\Gamma &= 0. \end{cases} \tag{4.4.47}$$

given Ω a domain in $I\!R^2$, $f \in L^2(\Omega)$, is known exactly on Ω, and $\lambda \leq a(u) \leq \Lambda$ for some positive constants λ and Λ. The parameter identification (PI) problem is to recover a nonlinear parameter $a = a(u)$ within some suitable set of admissible elements, from an L^2 observation $z(x)$ of $u(x)$ such that $a(u)$ and $u(x)$ satisfy (4.4.47). In this section, we consider the nonlinear parameter $a(u)$ as a linear spatially varying parameter $b(x_1, x_2)$ defined below:

$$b(x_1, x_2) = a(u(x_1, x_2)), \tag{4.4.48}$$

along with the constraint equation,

$$\frac{\partial b}{\partial x_1}\frac{\partial u}{\partial x_2} = \frac{\partial b}{\partial x_2}\frac{\partial u}{\partial x_1}. \tag{4.4.49}$$

We will see in the next technical lemma, that this constraint is used to guarantee the identified parameter b is not a multi-valued function when it is regarded as a function of the state u. First we need to impose some smoothness on u. Assume that

$$u \in C^1(\bar{\Omega}),\ \|\nabla u\| \neq 0 \text{ on } \bar{\Omega} \tag{4.4.50}$$

The next technical Lemma can be found in [42], p.277.

Lemma 4.4.1 *Suppose that $d(x_1, x_2)$ and $w(x_1, x_2)$ are continuously differentiable functions of x_1 and x_2 such that*

$$\frac{\partial(w, d)}{\partial(x_1, x_2)} = \frac{\partial d}{\partial x_2}\frac{\partial w}{\partial x_1} - \frac{\partial d}{\partial x_1}\frac{\partial w}{\partial x_2} = 0, \tag{4.4.51}$$

holds in a neighborhood of the point $(x_{10}, x_{20}) \in \Omega$. Suppose that w satisfies condition (4.4.50). Let $u_0 = w(x_{10}, x_{20})$. Then there exists an interval of the u-axis centered at u_0, and a C^1 function $c(u)$ defined thereon, such that

$$d(x_1, x_2) = c(w(x_1, x_2))$$

throughout a neighborhood of (x_{10}, x_{20}).

Before taking up the proof, we comment on the Lemma. Condition (4.4.51) is a necessary condition for the existence of such a function c. It is a local result, for the Lemma does not deal with a fixed region R, but with neighborhoods of a single point. It can be shown that a Theorem similar to this technical Lemma holds for a fixed region.

Proof. Consider the equation $u = w(x_1, x_2)$. Without loss of generality we can assume that $\frac{\partial w}{\partial x_1} \neq 0$ at $w_0 = (x_{10}, x_{20})$; then the implicit-function theorem guarantees the existence of a solution $x_1 = f(u, x_2)$, $f \in C^1$ giving all triples (x_1, x_2, u) near (x_{10}, x_{20}, u_0) for which $u = w(x_1, x_2)$. Moreover,

$$\frac{\partial f}{\partial x_2} = -\frac{\frac{\partial w}{\partial x_2}}{\frac{\partial w}{\partial x_1}}. \tag{4.4.52}$$

All this applies when (x_1, x_2, u) is close enough to (x_{10}, x_{20}, u_0). Consider the function $d(x_1, x_2)$ as a function of u and x_2, with $x_1 = f(u, x_2)$. The partial derivative of d with respect to x_2 is

$$\frac{\partial d}{\partial x_1}\frac{\partial f}{\partial x_2} + \frac{\partial d}{\partial x_2} = \frac{\frac{\partial(w,d)}{\partial(x_1,x_2)}}{\frac{\partial w}{\partial x_1}} = 0.$$

because of (4.4.51) and (4.4.52). Thus $d(f(u, x_2), x_2)$ is actually independent of x_2. Let us write $c(u) = d(f(u, x_2), x_2)$. Since $x_1 = f(u, x_2)$ is equivalent to $u = w(x_1, x_2)$, for the values of the current variables, $c(u) = d(f(u, x_2), x_2)$ is equivalent to $d(x_1, x_2) = c(w(x_1, x_2))$, which holds throughout a neighborhood of (x_{10}, x_{20}).

Under this approach, our nonlinear equation (4.4.47) reduces to:

$$\left\{ \begin{array}{rcl} -\mathrm{div}(b\nabla u(x)) & = & f(x), \quad \text{in } \Omega \subset I\!R^2 \\ b\frac{\partial u}{\partial \eta} \mid_\Gamma & = & 0. \end{array} \right. \tag{4.4.53}$$

To describe our approach, we let $H^s(\Omega)$ be the standard Sobolev space on Ω along with the standard notation for its norm and seminorm. Following the same ideas as we did for the one dimensional case, we introduce the flux variable $\sigma(x) = b(x)\nabla u(x)$ as an additional unknown. In this way, the differential equation in (4.4.53) reduces to the following equivalent first order system.

$$\left\{ \begin{array}{rcl} -\mathrm{div}\sigma & = & f, \\ b\nabla u & = & \sigma. \end{array} \right.$$

Again, we try to find an approximation to the solution of the PI problem by minimizing the following cost functional,

$$\begin{array}{rcl} \mathcal{J}(b,u,\sigma) & = & \|u-z\|_0^2+\alpha\|\sigma-b\nabla u\|_0^2+\beta\|\mathrm{div}\sigma+f\|_0^2+ \\ & & \gamma\|b\|_{r_b}^2+\rho\|u\|_{r_u}^2+\delta\|\sigma\|_{r_\sigma}^2+ \\ & & \xi\|(\sigma-b\nabla u)\cdot\eta\|_{L^2(\Gamma)}^2+\mu\|b\nabla u\cdot\eta\|_{L^2(\Gamma)}^2+ \\ & & \theta\|u_{x_1}b_{x_2}-u_{x_2}b_{x_1}\|_0^2 \end{array} \tag{4.4.54}$$

over the space

$$\mathcal{H}=H^1(\Omega)\times(H^1(\Omega)\bigcap W^{1,\infty}(\Omega))\times(H^1(\Omega))^2,$$

where the integers r_b, r_u, and r_σ depend on the Sobolev spaces in which the regularization takes place. This method as in the one dimensional case is formulated in a least-squares mixed approach.

At this time, it is not clear to us how to identify $a(u)$ in the two dimensional case.

4.4.1 Finite element method

In this section, we formulate a finite element discretization of the cost functional (4.4.54) and discuss the related minimization problem. In our estimations, we regard C as a generic constant, which may vary in different contexts, but is always independent of h, unless specified otherwise. In order to define the finite element spaces, let $\{\Delta_h\}$, with $h\in(0,1)$, be a family of regular, quasi-uniform partitions of $\Omega\subset I\!R^n$. For fixed integers $r\geq 1$, $l\geq 0$, we let $S^r_{h,l}$ be the space defined in (2.2.14), $C^{-1}(\Omega)$ again is interpreted as $L^2(\Omega)$. By $S^{r,0}_{h,l}$, we denote the subspace of $S^r_{h,l}$ of functions which vanish on $\Gamma=\partial\Omega$. The bounds (2.2.15), (2.2.16), (2.2.17) hold, H^{-1} stands for the dual space $(H^1(\Omega))^*$ equipped with the norm,

$$\|v\|_{-1}=\sup_{0\neq\psi\in H^1(\Omega)}\frac{|(v,\psi)|}{\|\psi\|_1}. \tag{4.4.55}$$

Three different discretization spaces are needed for the computational procedure: U_h for the solution u, B_h for the parameter b and Σ_h for the flux variable σ, we assume that these spaces are:

$$\begin{array}{rcl} B_h & = & S^{d_b}_{h,l}, \\ U_h & = & S^{d_u}_{h,1}, \\ \Sigma_h & = & S^{d_\sigma}_{h,1}\times S^{d_\sigma}_{h,1}, \end{array} \tag{4.4.56}$$

where d_b, d_u, d_σ are integers.

Now we are ready to define the numerical method to solve the PI problem: (P_h) seek $b_h^* \in B_h$, $u_h^* \in U_h$, $\sigma_h^* \in \Sigma_h$, such that

$$\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) = \min_{b\in B_h, u\in U_h, \sigma\in\Sigma_h} \mathcal{J}_h(b, u, \sigma) \tag{4.4.57}$$

where $\mathcal{J}_h$ is defined by,

$$\begin{aligned}\mathcal{J}_h(\mathbf{w}_h) \quad = \quad & \|u_h - z\|_0^2 + \alpha\|\sigma_h - b_h\nabla u_h\|_0^2 + \beta\|\mathrm{div}\sigma_h + f\|_0^2 + \\ & \gamma\|b_h\|_{r_b}^2 + \rho\|u_h\|_{r_u}^2 + \delta\|\sigma_h\|_{r_\sigma}^2 + \\ & \xi\|(\sigma_h - b_h\nabla u_h)\cdot\eta\|_{L^2(\Gamma)}^2 + \mu\|b_h\nabla u_h\cdot\eta\|_{L^2(\Gamma)}^2 + \\ & \theta\|u_{h,x_1}b_{h,x_2} - u_{h,x_2}b_{h,x_1}\|_0^2\end{aligned} \tag{4.4.58}$$

where $\mathbf{w}_h = (b_h, u_h, \sigma_h) \in B_h \times U_h \times \Sigma_h$. The advantage of this approach is that we do not differentiate the noisy data z. The weights γ, ρ, and δ will be chosen in the error estimation so that all terms of $\mathcal{J}_h(\mathbf{w}_h)$ are balanced.

4.4.2 Some error estimates

In this section, we establish some error estimates for an approximate solution to the PI problem generated by the finite element method in the previous section. Let I_b^h, I_u^h, and I_σ^h be the interpolation operators in the finite element spaces B_h, U_h, and Σ_h, respectively. The regularization spaces are chosen so that

$$1 \le r_b \le d_b + 1, 1 \le r_u \le d_u + 1, 1 \le r_\sigma \le d_\sigma + 1.$$

Moreover, there exists a constant C such that the interpolation operators satisfy,

$$\|I_b^h b\|_{r_b} \le C\|b\|_{r_b+1}, \|I_u^h u\|_{r_u} \le C\|u\|_{r_u+1}, \|I_\sigma^h \sigma\|_{r_\sigma} \le C\|\sigma\|_{r_\sigma+1},$$

for any $b \in H^{r_b+1}(\Omega), u \in H^{r_u+1}(\Omega), \sigma \in (H^{r_\sigma+1}(\Omega) \times H^{r_\sigma+1}(\Omega))$. We first derive an estimate for a minimizer of $\mathcal{J}_h$.

Lemma 4.4.2 *Let us assume the following hypothesis,*

* *Let $z(x)$ be a distributed L^2 observation of the state $\tilde{u}$ with an observation error of the form*

$$\|\tilde{u} - z\|_0 \le \epsilon.$$

* *Assume that* $\tilde{b} = \tilde{a}(\tilde{u})$ *and* $\tilde{u}$ *satisfy the differential equation (4.4.53). Moreover,* $\tilde{b}$*,* $\tilde{u}$ *and* $\tilde{\sigma} = \tilde{b}\nabla\tilde{u}$ *have the regularities:*

$$\tilde{b} \in H^{d_b+1}(\Omega)\bigcap W^{1,\infty}(\bar{\Omega}), \tilde{u} \in H^{d_u+1}(\Omega)\bigcap W^{1,\infty}(\bar{\Omega}), \tilde{\sigma} \in H^{d_\sigma+1}(\Omega),$$

* *Assume that* $\tilde{b}$ *and* $\tilde{u}$ *satisfy the constraint (4.4.49).*

* *The finite element space* U_h *is such that*

$$\|\nabla(I_u^h\tilde{u})\|_\infty \leq C,$$

for h *small enough.*

* *The regularization parameters* γ*,* ρ *and* δ *are chosen such that*

$$\begin{aligned}\gamma, \rho, \text{ and } \delta \;\; &\sim \;\; C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\ &\xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}\end{aligned} \tag{4.4.59}$$

with a constant C *independent of* h*.*

Let $\mathbf{w}_h^* = (b_h^*, u_h^*, \sigma_h^*)$ *be a minimizer of* $\mathcal{J}_h(\mathbf{w}_h)$ *defined by (4.4.58) in* $B_h \times U_h \times \Sigma_h$*, then we have the following estimate for the cost functional*

$$\begin{aligned}\mathcal{J}_h(\mathbf{w}_h^*) \;\; &\leq \;\; \{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\ &\xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}\end{aligned}$$

here C *is independent of* ϵ *and* h*.*

Proof. Let $(b_h^*, u_h^*, \sigma_h^*)$ be a minimizer of $\mathcal{J}_h$, which will be denoted by $\mathbf{w}_h^*$, and let $I_b^h\tilde{b}$, $I_u^h\tilde{u}$ and $I_\sigma^h\tilde{\sigma}$ be the interpolates of $\tilde{b} = \tilde{a}(\tilde{u})$, $\tilde{u}$, and $\tilde{\sigma} = \tilde{a}(\tilde{u})\nabla\tilde{u}$ in B_h, U_h and Σ_h, respectively. Now

$$\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) \leq \mathcal{J}_h(I_b^h\tilde{b}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}),$$

and

$$\begin{aligned}\mathcal{J}_h(I_b^h\tilde{b}, I_u^h\tilde{u}, I_\sigma^h\tilde{\sigma}) \;\; &= \;\; \|I_u^h\tilde{u} - z\|_0^2 + \alpha\|I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}\nabla I_u^h\tilde{u}\|_0^2 + \beta\|\mathrm{div} I_\sigma^h\tilde{\sigma} + f\|_0^2 + \\ &\gamma\|I_b^h\tilde{b}\|_{r_b}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2 + \\ &\xi\|(I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}\nabla I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma)}^2 + \mu\|I_b^h\tilde{b}\nabla I_u^h\tilde{u}\cdot\eta\|_{L^2(\Gamma)}^2 + \\ &\theta\|I_u^h\tilde{u}_{x_1}I_b^h\tilde{b}_{x_2} - I_u^h\tilde{u}_{x_2}I_b^h\tilde{b}_{x_1}\|_0^2\end{aligned}$$

and so,

$$\begin{aligned}\|I_u^h\tilde{u} - z\|_0 \;\; &\leq \;\; \|I_u^h\tilde{u} - \tilde{u}\|_0 + \|\tilde{u} - z\|_0 \\ &\leq \;\; Ch^{d_u+1}\|\tilde{u}\|_{d_u+1} + \epsilon,\end{aligned}$$

$$
\begin{aligned}
\|I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}\nabla I_u^h\tilde{u}\|_0 &\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_0 + \|(\tilde{a}(\tilde{u}) - I_b^h\tilde{b})\nabla\tilde{u}\|_0 + \|I_b^h\tilde{b}\nabla(\tilde{u} - I_u^h\tilde{u})\|_0 \\
&\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_0 + \|\nabla\tilde{u}\|_{\infty,\Omega}\|\tilde{b} - I_b^h\tilde{b}\|_0 + \|I_b^h\tilde{b}\|_{\infty,\Omega}\|\nabla(\tilde{u} - I_u^h\tilde{u})\|_0 \\
&\leq Ch^{d_\sigma+1}\|\tilde{\sigma}\|_{d_\sigma+1} + Ch^{d_b+1}\|\tilde{b}\|_{d_b+1} + Ch^{d_u}\|\tilde{u}\|_{d_u+1} \\
&\leq C\{h^{d_\sigma+1} + h^{d_u} + h^{d_b+1}\},
\end{aligned}
$$

$$
\begin{aligned}
\|\mathrm{div} I_\sigma^h\tilde{\sigma} + f\|_0 &= \|\mathrm{div}(I_\sigma^h\tilde{\sigma} - \tilde{a}(\tilde{u})\nabla\tilde{u})\|_0 \\
&= \|\mathrm{div}(I_\sigma^h\tilde{\sigma} - \tilde{\sigma})\|_0 \\
&\leq Ch^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1} \\
&\leq Ch^{d_\sigma},
\end{aligned}
$$

$$
\begin{aligned}
\|(I_\sigma^h\tilde{\sigma} - I_b^h\tilde{b}\nabla I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma)} &\leq \|(I_\sigma^h\tilde{\sigma} - \tilde{\sigma})\cdot\eta\|_{L^2(\Gamma)} + \\
&\quad \|(\tilde{b} - I_b^h\tilde{b})\nabla\tilde{u}\cdot\eta\|_{L^2(\Gamma)} + \|I_b^h\tilde{b}\nabla(\tilde{u} - I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma)} \\
&\leq \|I_\sigma^h\tilde{\sigma} - \tilde{\sigma}\|_{L^2(\Gamma)} + \\
&\quad \|\nabla\tilde{u}\|_{\infty,\Gamma}\|I_b^h\tilde{b} - \tilde{b}\|_{L^2(\Gamma)} + \|I_b^h\tilde{b}\|_{\infty,\Gamma}\|\nabla(I_u^h\tilde{u} - \tilde{u})\|_{L^2(\Gamma)} \\
&\leq C\{h^{d_\sigma}\|\tilde{\sigma}\|_{d_\sigma+1} + h^{d_b}\|\tilde{b}\|_{d_b+1} + h^{d_u-1}\|\tilde{u}\|_{d_u+1}\},
\end{aligned}
$$

$$
\begin{aligned}
\|I_b^h\tilde{b}\nabla(I_u^h\tilde{u})\cdot\eta\|_{L^2(\Gamma)} &= \|I_b^h\tilde{b}\nabla(I_u^h\tilde{u} - \tilde{u})\cdot\eta\|_{L^2(\Gamma)} + \|(I_b^h\tilde{b} - \tilde{b})\nabla\tilde{u}\cdot\eta\|_{L^2(\Gamma)} \\
&\leq \|I_b^h\tilde{b}\|_{\infty,\Gamma}\|\nabla(I_u^h\tilde{u} - \tilde{u})\cdot\eta\|_{L^2(\Gamma)} + \|I_b^h\tilde{b} - \tilde{b}\|_{L^2(\Gamma)}\|\nabla\tilde{u}\cdot\eta\|_{\infty,\Gamma} \\
&\leq C\{h^{d_u-1}\|\tilde{u}\|_{d_u+1} + h^{d_b}\|\tilde{b}\|_{d_b+1}\}
\end{aligned}
$$

In this last estimation, we have used the results from section 4.3.1.

$$
\begin{aligned}
\|I_u^h\tilde{u}_{x_1}I_b^h\tilde{b}_{x_2} - I_u^h\tilde{u}_{x_2}I_b^h\tilde{b}_{x_1}\|_0^2 &\leq \|I_u^h\tilde{u}_{x_1}(I_b^h\tilde{b}_{x_2} - \tilde{b}_{x_2})\|_0 + \|\tilde{b}_{x_2}(I_u^h\tilde{u}_{x_1} - \tilde{u}_{x_1})\|_0 + \\
&\quad \|\tilde{b}_{x_1}(I_u^h\tilde{u}_{x_2} - \tilde{u}_{x_2})\|_0 + \|I_u^h\tilde{u}_{x_2}(I_b^h\tilde{b}_{x_1} - \tilde{b}_{x_1})\|_0 \\
&\leq \|\nabla I_u^h\tilde{u}\|_{\infty,\Omega}\|\nabla(I_b^h\tilde{b} - \tilde{b})\|_0 + \|\tilde{b}_{x_2}\|_{\infty,\Omega}\|\nabla(I_u^h\tilde{u} - \tilde{u})\|_0 + \\
&\quad \|\tilde{b}_{x_1}\|_{\infty,\Omega}\|\nabla(I_u^h\tilde{u} - \tilde{u})\|_0 + \|\nabla I_u^h\tilde{u}\|_{\infty,\Omega}\|\nabla(I_b^h\tilde{b} - \tilde{b})\|_0 \\
&\leq Ch^{d_b}\|\tilde{b}\|_{d_b+1} + Ch^{d_u}\|\tilde{u}\|_{d_u+1} \\
&\leq C\{h^{d_b} + h^{d_u}\}
\end{aligned}
$$

Then so far,

$$
\begin{aligned}
\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) &\leq C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \gamma\|I_b^h\tilde{b}\|_{r_b}^2 + \rho\|I_u^h\tilde{u}\|_{r_u}^2 + \delta\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma}^2 + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}.
\end{aligned}
$$

Since $\|I_b^h\tilde{b}\|_{r_b}$, $\|I_u^h\tilde{u}\|_{r_u}$ and $\|I_\sigma^h\tilde{\sigma}\|_{r_\sigma} \leq$ constant, choosing γ, ρ, δ such that,

$$
\begin{aligned}
\gamma, \lambda \text{ and } \delta &\simeq C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}
\end{aligned}
$$

leads to

$$
\begin{aligned}
\mathcal{J}_h(b_h^*, u_h^*, \sigma_h^*) &\leq C\{h^{2(d_u+1)} + \epsilon^2 + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}.
\end{aligned}
$$

The proof of the lemma is completed.

Lemma 4.4.3 *Under the same hypothesis as in Lemma 4.4.2, we have the following bounds,*

$$\begin{aligned}\|\sigma_h^* - b_h^* \nabla u_h^*\|_0 &\leq \frac{C}{\sqrt{\alpha}}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\},\end{aligned}$$

$$\begin{aligned}\|\mathrm{div}\sigma_h^* + f\|_0 &\leq \frac{C}{\sqrt{\beta}}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\},\end{aligned}$$

$$\begin{aligned}\|(\sigma_h^* - b_h^* \nabla u_h^*) \cdot \eta\|_{L^2(\Gamma)} &\leq \frac{C}{\sqrt{\xi}}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\},\end{aligned}$$

$$\begin{aligned}\|b_h^* \nabla u_h^* \cdot \eta\|_{L^2(\Gamma)} &\leq \frac{C}{\sqrt{\mu}}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\},\end{aligned}$$

$$\begin{aligned}\|\nabla(u_h^* - \tilde{u})\|_0 &\leq \frac{C}{h}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\}.\end{aligned}$$

Proof. The first four estimates follow immediately from the definition of $\mathcal{J}_h(b, u, \sigma)$ and the previous lemma. The the last one, by the inverse inequality (2.2.16), we have

$$\begin{aligned}\|\nabla(u_h^* - \tilde{u})\|_0 &\leq C\{\|\nabla(u_h^* - I_u^h \tilde{u})\|_0 + \|\nabla(I_u^h \tilde{u} - \tilde{u})\|_0\} \\ &\leq \{\frac{1}{h}\|u_h^* - I_u^h \tilde{u}\|_0 + h^{d_u}\|\tilde{u}\|_{d_u+1}\} \\ &\leq C\{\frac{1}{h}(\|u_h^* - z\|_0 + \|z - \tilde{u}\|_0 + \|\tilde{u} - I_u^h \tilde{u}\|_0) + h^{d_u}\} \\ &\leq C\{\frac{1}{h}(\|u_h^* - z\|_0 + \epsilon + h^{d_u+1}) + h^{d_u}\} \\ &\leq C\{\frac{1}{h}\mathcal{J}_h^{1/2}(b_h^*, u_h^*, \sigma_h^*) + \frac{\epsilon}{h} + h^{d_u}\} \\ &\leq \frac{C}{h}\{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta} h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\}\end{aligned}$$

This completes the proof of the Lemma.

It can be shown that the minimizer of $\mathcal{J}_h(\mathbf{w})$ is uniformly bounded with respect to h as for the one dimensional case.

Corollary 4.4.1 *If all the conditions of Lemma 4.4.2 hold, and*

$$\begin{aligned}
0 < \gamma &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\} \\
0 < \rho &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\} \\
0 < \delta &= C\{\epsilon^2 + h^{2(d_u+1)} + \alpha(h^{2(d_\sigma+1)} + h^{2d_u} + h^{2(d_b+1)}) + \beta h^{2d_\sigma} + \\
&\quad \xi(h^{2d_\sigma} + h^{2d_b} + h^{2(d_u-1)}) + \mu(h^{2(d_u-1)} + h^{2d_b}) + \theta(h^{2d_b} + h^{2d_u})\}
\end{aligned}$$

then there exists a constant C such that,

$$\|\mathbf{w}_h^*\|_{\mathcal{H}} \leq C, \text{ for all, } h.$$

Proof. The proof is identical to that in Corollary 4.2.1.

Corollary 4.4.2 *Assume the hypothesis of Corollary 4.4.1 and take H^2 to be the Sobolev space for the regularization of a_h. If $A_h = S_{h,l}^{d_a}$ with $l \geq 2$ then*

$$\|a_h^*\|_\infty < \infty.$$

Proof. The proof is identical to that in Corollary 4.2.2.

We will state a Lemma that will be needed in the next error estimate. It has been taken from [26].

Lemma 4.4.4 *Suppose that $\tilde{u}(x) \in C^2(\bar{\Omega})$ and $\|\nabla\tilde{u}\| \neq 0$ on $\bar{\Omega}$. Then for each $\psi \in H^1(\Omega)$, there exists a $v_\psi \in H^1(\Omega)$ satisfying*

$$\nabla\tilde{u} \cdot \nabla v_\psi = \psi \text{ in } \Omega,$$

and the estimate

$$\|v_\psi\|_1 \leq C\|\psi\|_1,$$

where C is independent of ψ.

Proof. For its proof, see Lemma 5 in [26].

We now present an estimate of the error in the space $H^{-1}(\Omega)$.

Theorem 4.4.1 *Suppose that all the conditions in Corollary 4.4.2 hold, along with the hypothesis in Lemma 4.4.4. If $\mathbf{w}_h^* = (b_h^*, u_h^*, \sigma_h^*)$ is a minimizer of $\mathcal{J}_h(b, u, \sigma)$, then we get the following estimation in the H^{-1} norm,*

$$\|b_h^* - \tilde{a}(\tilde{u})\|_{-1} \quad \leq \quad C\{c_0 + c_1\}$$

where c_0 and c_1 are defined by,

$$\begin{aligned} c_0 \quad &\leq \quad C(\frac{1}{\sqrt{\beta}} + h^{-1/2} \times (\frac{1}{\sqrt{\xi}} + \frac{1}{\sqrt{\mu}})) \times \{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\}, \\ c_1 \quad &\leq \quad C(\frac{1}{\sqrt{\alpha}} + \frac{1}{h} + h^{1/2} \times (\frac{1}{\sqrt{\xi}} + \frac{1}{\sqrt{\mu}})) \times \{h^{(d_u+1)} + \epsilon + \sqrt{\alpha}(h^{(d_\sigma+1)} + h^{d_u} + h^{(d_b+1)}) + \sqrt{\beta}h^{d_\sigma} + \\ &\quad \sqrt{\xi}(h^{d_\sigma} + h^{d_b} + h^{(d_u-1)}) + \sqrt{\mu}(h^{(d_u-1)} + h^{d_b}) + \sqrt{\theta}(h^{d_b} + h^{d_u})\}. \end{aligned}$$

Proof. Set $\phi(x) = b_h^*(x) - \tilde{a}(\tilde{u}(x))$. Now, we estimate the term $-\int_\Omega \phi \nabla\tilde{u} \cdot \nabla v\, dx$ when $v \in H^1(\Omega)$. By using the variational form of equation (4.4.47), we get

$$\begin{aligned} -\textstyle\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\, dx \quad &= \quad \textstyle\int_\Omega \tilde{a}(\tilde{u})\nabla\tilde{u}\cdot\nabla v\, dx - \int_\Omega b_h^*\nabla\tilde{u}\cdot\nabla v\, dx \\ &= \quad \textstyle\int_\Omega fv\, dx - \int_\Omega b_h^*\nabla\tilde{u}\cdot\nabla v\, dx \end{aligned}$$

Adding and subtracting various terms and integrating by parts, we get

$$\begin{aligned} -\textstyle\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\, dx \quad = \quad &\textstyle\int_\Omega (f + \mathrm{div}\sigma_h^*)v\, dx + \int_\Omega(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\nabla v\, dx + \\ &\textstyle\int_\Omega b_h^*\nabla(u_h^* - \tilde{u})\cdot\nabla v\, dx - \int_\Gamma(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta v\, ds - \int_\Gamma b_h^*\nabla u_h^*\cdot\eta v\, ds \end{aligned}$$

which holds for all test functions $v \in H^1(\Omega)$. To bound the left hand side, we use Lemma 4.3.3,

$$\begin{aligned} |\textstyle\int_\Omega \phi\nabla\tilde{u}\cdot\nabla v\, dx| \quad \leq \quad &\|f + \mathrm{div}\sigma_h^*\|_0\|v\|_0 + \|\sigma_h^* - b_h^*\nabla u_h^*\|_0\|\nabla v\|_0 + \\ &\|b_h^*\|_{\infty,\Omega}\|\nabla(u_h^* - \tilde{u})\|_0\|\nabla v\|_0 + \\ &C(h^{-1/2}\|v\|_0 + h^{1/2}\|\nabla v\|_0)(\|(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma)} + \\ &\|b_h^*\|_{\infty,\Omega}\|\nabla u_h^*\cdot\eta\|_{L^2(\Gamma)}) \\ \leq \quad &C\{\|f + \mathrm{div}\sigma_h^*\|_0 + h^{-1/2}(\|(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma)} + \\ &\|b_h^*\|_{\infty,\Omega}\|\nabla u_h^*\cdot\eta\|_{L^2(\Gamma)})\}\|v\|_0 + \\ &\{\|\sigma_h^* - b_h^*\nabla u_h^*\|_0 + \|b_h^*\|_{\infty,\Omega}\|\nabla(u_h^* - \tilde{u})\|_0 + \\ &h^{1/2}(\|(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma)} + \|b_h^*\|_{\infty,\Omega}\|\nabla u_h^*\cdot\eta\|_{L^2(\Gamma)})\}\|\nabla v\|_0 \\ \leq \quad &c_0\|v\|_0 + c_1\|\nabla v\|_0 \end{aligned}$$

for all $v \in H^1(\Omega)$, $\|b_h^*\|_\infty$ is bounded by the previous Corollary and c_0 and c_1 are defined below,

$$\begin{aligned} c_0 &= \|f + \mathrm{div}\sigma_h^*\|_0 + h^{-1/2}(\|(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma)} + \|b_h^*\|_{\infty,\Omega}\|\nabla u_h^*\cdot\eta\|_{L^2(\Gamma)}), \\ c_1 &= \|\sigma_h^* - b_h^*\nabla u_h^*\|_0 + \|b_h^*\|_{\infty,\Omega}\|\nabla(u_h^* - \tilde{u})\|_0 + \\ & \quad h^{1/2}(\|(\sigma_h^* - b_h^*\nabla u_h^*)\cdot\eta\|_{L^2(\Gamma)} + \|b_h^*\|_{\infty,\Omega}\|\nabla u_h^*\cdot\eta\|_{L^2(\Gamma)}). \end{aligned}$$

Equating $\nabla\tilde{u}\cdot\nabla v$ to ψ, invoking Lemma 4.4.4 and replacing v by v_ψ, we get that

$$|\int_\Omega \phi\psi\, dx| \leq C(c_0 + c_1)\|\psi\|_1 \text{ for all } \psi \in H^1(\Omega)$$

Dividing both sides by $\|\psi\|_1$, we get the following H^{-1} norm estimation,

$$\|b_h^* - \tilde{a}(\tilde{u})\|_{-1} \leq C(c_0 + c_1).$$

The proof is complete by applying Lemma 4.4.3 to the above inequality.

Theorem 4.4.2 *Under the hypothesis of Theorem 4.4.1, we have the following L^2 norm error estimation,*

$$\|b_h^* - \tilde{a}(\tilde{u})\|_0 \leq C(c_0 + c_1)^{1/2}$$

where c_0 and c_1 are defined in Theorem 4.4.1.

Proof. The proof follows immediately from the inequality below, which was taken from [26], pg. 146.

$$\|b_h^* - \tilde{a}(\tilde{u})\|_0 \leq C\|b_h^* - \tilde{a}(\tilde{u})\|_{-1}^{1/2}\|b_h^* - \tilde{a}(\tilde{u})\|_1^{1/2}, \tag{4.4.60}$$

which reduces to

$$\|b_h^* - \tilde{a}(\tilde{u})\|_0 \leq C\|b_h^* - \tilde{a}(\tilde{u})\|_{-1}^{1/2}. \tag{4.4.61}$$

Chapter 5

Future work

We would like to list some projects we think are worth mentioning and need further investigation.

The identification of two or more parameters in the most general elliptic partial differential equation

$$-\sum_{i,j=1}^{n}(a_{ij}u_{x_j})_{x_i} + \sum_{i=1}^{n} b_i u_{x_i} + cu = f \text{ in } \Omega, \tag{5.0.1}$$

with boundary conditions given by,

$$u \mid_{\Gamma} = 0, \text{ or } \frac{\partial u}{\partial \eta} \mid_{\Gamma} = 0, \text{ where } \Gamma = \partial\Omega, \tag{5.0.2}$$

is a problem we would like to work on to define an algorithm and get error estimates for the approximations to the unknown coefficients, in appropriate weighted norms, by using a finite element method. Here, Ω is a bounded domain in $I\!R^n$, with smooth boundary.

The identification of a discontinuous spatially varying parameter including the location of its discontinuities seems to be quite a challenging problem as to what numerical technique should be used to recover such parameter. We wonder whether we should use discontinuous finite elements with a step by step adjustable partition (or triangulation) to recover the coefficient or not.

Another topic that needs further study is an algorithm to define $a(u)$ for the nonlinear boundary value problem (4.4.47).

Another interesting problem is the identification of the parameter for the following parabolic partial differential equation:

$$\begin{aligned} \frac{\partial u}{\partial t} - \nabla \cdot (a(x)\nabla u) &= f(x,t) \quad \text{in } \Omega,\ t>0 \\ \frac{\partial u}{\partial \eta} &= g(x,t) \quad \text{on } \Gamma,\ t>0 \\ u(0,x) &= u_0(x) \quad \text{in } \Omega. \end{aligned} \tag{5.0.3}$$

where Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary Γ, f and u_0 are given data, $a : \mathbb{R} \to \mathbb{R}^+$ is the unknown coefficient function satisfying

$$\mu \le a(r) \le \Lambda \text{ for all } r \in \mathbb{R}^+,$$

for some positive constants μ and Λ.

The problem of identifying the coefficient $a = a(x)$ (or both $a = a(x)$ and $q = q(x)$) in the wave equation below,

$$\begin{aligned} q(x)\frac{\partial^2 u}{\partial t^2} - \nabla \cdot (a(x)\nabla u) &= f(x,t) \quad \text{in } \Omega \times [0,\tau], \\ u(x,0) &= g(x) \quad \text{in } \Omega, \\ \frac{\partial u}{\partial t}(x,0) &= h(x) \quad \text{in } \Omega, \\ u(x,t) &= k(x,t) \quad \text{on } \partial\Omega \times [0,\tau]. \end{aligned} \tag{5.0.4}$$

from additional observations $z_i(t) = u(x_i,t)$, $x_i \in \Omega$, $i = 1, \cdots, n$, $t \in [0,\tau]$, is interesting in its own right. Here $\Omega \subset \mathbb{R}^n$ denotes some bounded domain with a smooth boundary and q, f, g, h and k are given sufficiently smooth functions.

We would like to extend our mixed least squares numerical approach along with a discrete-time scheme to recover the coefficients in the boundary value problems (5.0.3) and (5.0.4) and get weighted error estimates for the approximation of the coefficients in the H^{-1} and L^2 norms. These last two problems are potential research topics, since they have applications in, for instance, seismology, glaciology, population dynamics, climatology and in the elastic body theory.

Finally, a rigorous proof for the weighted L^2 norm error estimation for the parameter identification problem with homogeneous Dirichlet boundary condition still remains an open issue that needs further investigation.

Bibliography

[1] R. C. Acar. Identification of coefficients in elliptic systems. *Thesis, Univ. of Wisconsin, USA*, 1987.

[2] G. Alessandrini. On the identification of the leading coefficient of an elliptic equation. *Bolletino U.M.I., Analisi Funzionale e Applicazioni, Serie* VI, IV-C:87–111, 1985.

[3] G. Alessandrini. An identification problem for an elliptic equation in two variables. *Ann. Mat. Pura. Appl.*, 145:265–296, 1986.

[4] E. L. Allgower, K. Böhmer, F.A. Potra, and W. C. Rheinboldt. A mesh-independence principle for operator equations and their discretizations. *SIAM J. Numer. Anal.*, 23(1):160–169, 1986.

[5] E. L. Allgower and S.F. McCormick. Newton's method with mesh refinements for numerical solution of nonlinear two-point boundary boundary value problems. *Numer. Math.*, 29:237–260, 1978.

[6] J. Baumeister. *Stable Solutions of Inverse Problems.* Vieweg, Wiesbaden, 1986.

[7] M.S. Bazaraa and C.M. Shetty. *Nonlinear Programing, Theory and Algorithms.* John Wiley & Sons, New York, 1979.

[8] J. Bear. *Dynamics of Fluids in Porous Media.* American Elsevier, New York, 1972.

[9] A. J. Bensoussan, L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures.* North-Holland, Amsterdam, 1978.

[10] P.B. Bocheve and M.D. Gunzburger. Accuracy of the least-squares methods for the navier-stokes equations. *Comput. Fluids*, 22:545–563, 1993.

[11] G.F. Carey, A. Pehlivanov, and Y. Shen. *Least-squares mixed finite elements.* Finite Element Methods. Marcel Dekker, Inc, 1994.

[12] G. Chavent. Identification of distributed parameter system: about the output least square method, its implementation, and identifiability. In Isermann R, editor, *Proc. 5th IFAC Symposium on Identification and System Parameter Estimation*, pages 85–97. Pergamon Press, 1980.

[13] P.G. Ciarlet. *Numerical Analysis of the Finite Element Method.* Montreal Press, 1975.

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

[15] P.G. Ciarlet. *Basic Error Estimates for Elliptic Problems*, volume II of *Handbook of Numerical Analysis, edited by P.G. Ciarlet and J.L. Lions.* North-Holland, Amsterdam, 1991.

[16] C.L.Chang and J.J. Nelson. Least-squares finite element method for the stokes problem with zero residual of mass conservation. *SIAM Num. Anal.*, 34:480–489, 1997.

[17] F. Colonius and K. Kunisch. Stability for parameter estimation in two point boundary value problems. *J. Reine Angew. Math.*, 370:1–29, 1986.

[18] L.C.W. Dixon, E. Spedicato, and G.P. Szego. *Nonlinear optimization, theory and algorithms.* (The proceedings of an international summer shool held at University of Bergamo, Italy, in September 1979 and repetead at the Hatfield Polytechnic, England, in July 1980). Birkhauser, 1980.

[19] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, 5:423–440, 1989.

[20] R.S. Falk. Error estimates for the numerical identification of a variable coefficient. *Math. Comp.*, 40(162):537–546, 1983.

[21] E.O. Frind and G.F. Pinder. Galerkin solution of the inverse problem for aquifer transmissivity. *Water Resources Research*, 9:1397–1410, 1973.

[22] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag, New York, 1977.

[23] K.H. Hoffman and J. Sprekels. On the identification of coefficients of elliptic problems by asymptotic regularization. *Num. Funct. Anal. and Optimiz.*, 7(2&3):157–177, 1984-85.

[24] T. Kärkkäinen. Distributed parameter identification in linear elliptic problems. Technical Report 5, Laboratory of Scientific Computing, Department of Mathematics, University of Jyväskylä, 1994.

[25] T. Kärkkäinen. Error estimates for distributed parameter identification in linear elliptic equations. *Journal of Mathematical Systems, Estimation, and Control*, 6(1):117–120, 1996.

[26] R.V. Kohn and B.D. Lowe. A variational method for parameter identification. *Rairo, Math. Model. and Num. Anal.*, 22(1):119–158, 1988.

[27] C. Kravaris and J.H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM J. Contr. Optimiz.*, 23:217–241, 1985.

[28] R. Lattes and J.L. Lions. *Méthode de quasi-réversibilité et applications*. Dunod, Paris, 1967.

[29] T. Lin. A c^0 finite element method for an inverse problem. *Applied Mathematics and Computation*, to appear, 1998.

[30] T. Lin and R. Ewing. Direct numerical method for an inverse problem of hyperbolic equations. *Numerical Methods for Partial Differential Equations*, 8:551–574, 1992.

[31] T. Lin and E. Ramirez. A numerical method for parameter identification of a two point boundary value problem. *preprint*, 1997.

[32] T. Lin and B. Zhang. A finite element method for an initial value inverse problem with overspecified boundary data. *Journal of Mathematical Systems, Estimation, and Control*, 7(3):305–342, 1997.

[33] G.P. McCormick. *Nonlinear Programming, Theory, Algorithms, and Applications*. John Wiley & Sons, New York, 1983.

[34] R. Rannacher. On the convergence of the newton-raphson method for strongly nonlinear. *Nonlinear Computational Mechanics, State of the Art*, pages 11–30, 1991.

[35] G.R. Richter. An inverse problem for the steady state diffusion equation. *SIAM J. Appl. Math.*, 41:210–221, 1981.

[36] G.R. Richter. Numerical identification of a spatially varying diffusion coefficient. *Math. Comp.*, 36:375–386, 1981.

[37] M.H. Schultz. *Spline Analysis*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1973.

[38] R. Scott. Interpolated boundary conditions in the finite element method. *SIAM J. Numer. Anal.*, 12:404–427, 1975.

[39] S. Taásan and P. Dutt. Parameter estimation problems for distributed systems using a multigrid method. *ICASE Report No. 86-65,NASA Langley Research Center, Hampton VA*, 1986.

[40] X.-C. Tai and P. Neittaanmäki. A linear approach for the nonlinear distributed parameter identification problem. *Numerical Methods for Free Boundary Problems*, 99:401–411, 1991.

[41] U. Tautenhahn. A new regularisation method for parameter identification in elliptic problems. *Inverse Problems*, 6:465–477, 1990.

[42] A.E. Taylor. *Advanced calculus*. Blaisdell Publishing Company, 1965.

[43] W.I. Zangwill. *Nonlinear Programming: A Unified Approach*. Prentice-Hall, Englewood Cliffs, N.J., 1969.

[44] Z.Cai, R.Lazarov, T.Manteuffel, and S.McCormick. First-order system least-squares for second-order partial differential equations. *Part I, SIAM Num. Anal.*, 31:1785–1802, 1994.

[45] Z.Cai, T.Manteuffel, and S.McCormick. First-order system least-squares for second-order partial differential equations. *Part II, SIAM Num. Anal.*, 34:425–454, 1997.

[46] E. Zeidler. *Nonlinear Functional Analysis and its Applications, II/B*. Springer-Verlag, New York, 1990.

Vita

I was born on June 11, 1966 in Mendoza, Argentina. During my childhood, I grew up in different places, the most exciting of which was by the foot of the Andes mountains. After wandering around for seven years, my family decided to settle down back in the city. By the time I completed my elementary school education at Domingo F. Sarmiento, I knew I was interested in seriously pursuing studies in a technical discipline and so attended the National Technical School Pablo Nogues for High School. I graduated from there in 1984. I graduated from San Luis State University in 1990 with the title ”Licenciado en Matemáticas”. At this time I moved east to Santa Fe State University to study Harmonic Analysis with Dr. Eleonora Harboure. From there, I came to the USA in August, 1992 to pursue my Ph.D. degree in Applied Mathematics at Virginia Tech. Since then I have been working for the Math Department as a Graduate Teaching Assistant. I received my M.S. in Applied Mathematics from Virginia Tech in August, 1993 and begun conducting research with Dr. Lin since August, 1995.