Department
of
APPLIED MATHEMATICS

Identification of a nonlinear parameter in a parabolic equation by a linear approach

by

Xue-Cheng Tai and Tommi Kärkkäinen

Report No. 93 March 1993

UNIVERSITY OF BERGEN
Bergen, Norway

Department of Mathematics
University of Bergen
5007 Bergen
Norway

ISSN 0084-778x

Identification of a nonlinear parameter in a parabolic equation by a linear approach

by

Xue-Cheng Tai and Tommi Kärkkäinen

Report No. 93

March 1993

ACKNOWLEDGEMENTS

This research was supported in part by the university of Bergen, Norway, and in part by the university of Jyväskylä, Finland.

IDENTIFICATION OF A NONLINEAR PARAMETER IN A PARABOLIC EQUATION BY A LINEAR APPROACH

X.-C. TAI*, T. KÄRKKÄINEN**

* Department of applied mathematics,
University of Bergen, 5007, Bergen,
Norway.
** Department of Mathematics,
University of Jyväskylä,
P. O. Box. 35, Fin-40351,
Jyväskylä, Finland.

CONTENTS

1. Introduction

In this article we consider the nonlinear parabolic system of the form

$$\begin{aligned} \frac{\partial u}{\partial t} - \nabla \cdot (a(u) \nabla u) &= f(t,x) \quad \text{in } (0,T] \times \Omega\,, \\ a(u)\frac{\partial u}{\partial n} &= g(t,x) \quad \text{on } \partial\Omega\,, \\ u(0,x) &= u_0(x) \quad \text{in } \Omega\,, \end{aligned} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$, $n = 1,2,3$. Our problem is that we want to find the nonlinear parameter $a(u)$ in (1.1). We assume, that we have some experimental data, which gives some information about the state solution u of (1.1), and we want to use this information of u to recover the unknown parameter $a(u)$. In practical applications the equation (1.1) can describe for example a heat conduction process. In this case the nonlinear parameter is the thermal conductivity, which depends only on the temperature u. Of course equation (1.1) can also describe other physical phenomena too, for example, a diffusion model of population [**2**], [**3**]. In the population model, the nonlinear parameter $a(u)$ means the diffusion parameter, which depends only on the number of population. In these practical applications it is always not easy to measure the parameter. However, to understand the dependence of the parameter on the solution u is important for many applications. Therefore, we need some ways to determine this dependence.

To be more concrete, we give one experimental setup, for which our proposed method can be applied. This experiment represents a process of casting of metal. A number of thermal couples is placed at the centerline of the isolated cylinder at varying distances from the interfaces of the melted metal and mould, the thermal conductivity of which is to be determined, which we need to use for the continuous casting of the metal [**12**], [**13**] etc. After conforming that all thermal couples indicate room temperature, melted metal is poured into the mould. Data of the mould temperature is then stored in an array $\tilde{u}(x_i, t_j)$, where $x_i, i = 1, ..., k$ is the distance of the i^{th} thermal couple from the casting and $t_j, j = 0, ..., \eta$ is the time after pouring the metal. The temperatures are measured at every time level until the final measuring time is reached. The actual distance of each thermocouple from the interphase is measured with a depthmicrometer after solidification by carefully scraping away the metal until the thermocouples are exposed. The above heat process can be described by the equation (1.1).

In the above experiment, we have a point-observation at different time levels. After interpolating this point data we get a distributed observation for the temperature and the velocity of its change, the initial value $u_0(x)$ and also boundary value $g(t,x)$. Now, we are going to identify the nonlinear parameter from this distributed observation. It is well known, that the identification problems are usually illposed. This gives us a reason to transform the identification problem into a minimization problem. To do this we will use the output-least-squares method. However, the method we are going to propose, is different from the ones described in literature, see [**1**], [**3**], [**4**], [**15**], etc. The output-least-squares method in the identification will lead to a minimization of a cost functional of output- error. The partial differential equation (1.1) is regarded as a constraint in the minimization. In order to compute the gradient or the Hessian in the minimization procedure, we need to solve the equation many times. For the nonlinear equation (1.1), the cost of the CPU-time will become unfavorable even for simple application problems.

The special approach we are going to propose here has been reported in a preliminary paper [**17**]. In [**17**], the idea was used for an elliptic problem, but the numerical tests there were done for both parabolic and elliptic systems. However, in [**17**], we did not treat observation errors. The error analysis and numerical tests were done only for the case that the the observation z is without observation error, i.e. $z = u$. The essential idea we are going to use is as follows: For inverse problem, we assume that we approximately know the state u and we will identify $a(u)$. To do this we write

$$b(t,x) = a(u(t,x)) \,. \tag{1.2}$$

If we can recover the parameter $b(t,x)$, we can easily recover parameter $a(u)$ from relation (1.2). Using this definition and substituting $b(t,x)$ into (1.1), we get an equation

$$\begin{aligned} \frac{\partial u}{\partial t} - \nabla \cdot (b \nabla u) &= f(t,x) \quad \text{in } (0,T] \times \Omega \,, \\ b\frac{\partial u}{\partial n} &= g(t,x) \quad \text{on } \partial\Omega \,, \\ u(0,x) &= u_0(x) \quad \text{in } \Omega \,. \end{aligned} \tag{1.3}$$

If we have some observation for u, we can recover parameter $b(t,x)$ from (1.3). However, the parameter $b(t,x)$ is not an ordinary linear parameter but it satisfies also relations (1.1) and (1.2). How can we now guarantee that for the identified parameter $b(t,x)$ from (1.3) there exists a function $a(u)$ such that (1.2) is valid? This question is answered in Lemma 2.1. We assume also, that in the computation of the identification we have some observation error as well as computational error. In Lemmas 2.2 and 2.3 we demonstrate, how condition (1.2) can be guaranteed "approximately".

By using the above idea, we can transform the identification of a nonlinear parameter to the identification of a linear parameter. The reflection of this change in the computation is evident. We don't have to solve any nonlinear equations but only linear equation need to be solved in the computation process. For different approaches about nonlinear distributed parameter identification, we refer to papers, **[4]**, **[5]**, **[6]**, **[7]**, **[14]**, **[15]**, etc.

2. Notations and preliminaries

We assume $\Omega \subset \mathbb{R}^n$ is convex and bounded, $\partial\Omega \in C^1$, and denote by $Q = (0,T] \times \Omega \subset \mathbb{R}^{n+1}$. We will use $D_i = \frac{\partial}{\partial x_i}$, $i = 1,2 \cdots n$ to denote the partial derivatives, and use $D_t = \frac{\partial}{\partial t}$ to denote the velocity. For convenience, we will use $\nabla u = (D_1 u, D_2 u, \cdots, D_n u)$ to denote the gradient. In proofs, we need sometimes to treat the time variable t and the space variable x equally. In such situations, we will denote $x_0 = t$, $D_0 = D_t$, $(t,x) = (x_0, x_1, x_2, \cdots x_n)$ and do not distinguish between t and x_0. Standard notations for Sobolev spaces will be used. In space $L^2(\Omega)$, we will use $(\cdot,\cdot)$ to denote the inner product, and use $\|\cdot\|$ to denote its norm. For a given domain S, we will use $\|\cdot\|_{W^{k,p}(S)}$ to denote the norm for $W^{k,p}(S)$. If $p = 2$, we use $\|\cdot\|_{H^k(S)}$ to denote the norm, and if $k = 0$, we use $\|\cdot\|_{L^p(S)}$ to denote the norm. Due to the appearance of the Neumann boundary condition, we will use $\langle \cdot,\cdot \rangle$ to denote the L^2–inner product on $\partial\Omega$. As for constant, we will use C to denote a generic positive constant, which may differ from context to context.

For parabolic equation (1.1) we will use finite element method for the space discretization. As we do not want to confine ourself to a specific time discretization method for the time variable, we will concentrate only on the semidiscrete analysis of (1.1). In order to define the finite element spaces, we let $\mathcal{T}_h, 0 < h < 1$, be a family of triangulation of $\bar{\Omega}$. If the boundary of Ω is curved, we shall use triangles at the boundary with one edge replaced by a curved segment of the boundary. We assume, that the family $\mathcal{T}_h$ is regular and quasi-uniform. For a fixed integer $r \geq 1$, we define

$$S_h^r = \left\{ v \,|\, v \in C^0(\Omega),\ \ v|_e \in P_r,\ \forall\, e \in \mathcal{T}_h \right\},$$

where P_r is the space of polynomials of degree less or equal to r. Associated with the finite element space, we will define one norm for piecewise Sobolev functions. We define

$$|||\cdot|||_{W^{k,p}(\Omega)} = \sum_{e \in \mathcal{T}_h} \|\cdot\|_{W^{k,p}(e)} \ ,$$

and if $p = 2$, we write it as $|||\cdot|||_{H^k(\Omega)}$.

In order to get an error estimate for the identified parameter, we need the following assumptions about the u, first we assume $u \in C^1(\bar{Q})$ and also, that there exists a constant vector $\vec{\nu}$ and a constant $\delta > 0$ such that

$$\text{(A1)} \qquad \nabla u(\cdot,x)\cdot\vec{\nu} \geq \delta > 0 \quad \forall\, x \in \Omega\,.$$

This assumption is essential in our analysis, but does not appear to be necessary in practical applications.

The three Lemmas, we are going to prove, reflect the existence of the dependency (1.2) for linear parameter $b(t,x)$ and nonlinear parameter $a(u)$. In the proof of the following lemmas, we regard t and x_0 as equivalent.

Lemma 2.1. *Let $d(t,x), w(t,x) \in C^1(Q)$ be given functions and let $w(t,x)$ satisfy the condition (A1). The necessary and sufficient condition for the existence of a function $c \in C^1(\mathbb{R})$ such that*

$$d(t,x) = c(w(t,x)) \tag{2.1}$$

is given by

$$D_i d D_j w = D_j d D_i w \quad i,j = 0,1,...,n, i \neq j \,. \tag{2.2}$$

Proof. From the condition (2.1) we get, that $\forall\, 0 \leq i,j \leq n$

$$D_i d = c'(w) D_i w, \;\; D_j d = c'(w) D_j w \,. \tag{2.3}$$

A direct calculation using (2.3) leads to

$$D_i d D_j w = c'(w) D_i w D_j w = c'(w) D_j w D_i w = D_j d D_i w \,.$$

So if (2.1) holds, then the functions will satisfy condition (2.2). We now turn our attention to prove that this is also a sufficient condition. Since $w(t,x)$ satisfy the condition (A1), we can without loss of generality assume, that $D_1 w \geq \delta > 0$ (by changing the x_1-axis to $\vec{\nu}-$direction). From the implicit function theorem we then know, that for the function

$$w = w(t, x_1, x_2, ..., x_n) \tag{2.4}$$

there exists a unique inverse function e such that

$$x_1 = e(t, w, x_2, ..., x_n) \,.$$

By substituting this into $d(t, x_1, ..., x_n)$, i.e. by changing the independent variable of d into $(t, w, x_2, ..., x_n)$ and considering the mapping of $(t, w, x_2, ..., x_n)$ to d as a function $c(t, w, x_2, ..., x_n)$ we get

$$d = d(t, x_1, ..., x_n) = d(t, e(t, w, x_2, ..., x_n), x_2, ..., x_n) = c(t, w, x_2, ..., x_n) \,. \tag{2.5}$$

Next we prove, that c only depends on w. By taking w as a variable and regarding x_1 as a function of $(t, w, x_2, ..., x_n)$ we get by differentiating (2.4) with respect to $x_i, i \neq 1$,

$$0 = D_1 w D_i e + D_i w \,, \tag{2.6}$$

and this gives

$$D_i e = -D_i w (D_1 w)^{-1} \,. \tag{2.7}$$

By differentiating (2.5) with respect to x_i, $i \neq 1$, we get

$$D_i c = D_1 d D_i e + D_i d \,. \tag{2.8}$$

By substituting (2.7) into (2.8) and using the condition (2.2) we get

$$D_i c = -D_1 d D_i w (D_1 w)^{-1} + D_i d = 0 \,. \tag{2.9}$$

As Ω is convex, this means, that $c(t, w, x_2, ..., x_n)$ is independent on $x_i, i \neq 1$ and hence, we have proved the existence of c that satisfies (2.1). In fact this proof gives us, that $c \in C^1([\min_{(t,x)\in Q} w(t,x), \max_{(t,x)\in Q} w(t,x)])$. Using [**11**, Lemma 6.37] we get, that there exists an extension of this into $C^1(\mathbb{R})$. This proves the result.

Lemma 2.2. *Let $d(t,x), w(t,x) \in C^1(Q)$ be two given functions and let $w(t,x)$ satisfy the condition (A1). If*

$$D_i d D_j w = D_j d D_i w + \beta_{ij}(t,x) \quad i,j = 0,1,2,...,n,\ i \neq j \ , \tag{2.10}$$

where $\beta_{ij}(t,x) \in C^0(Q)$ are given, then there exists functions $c \in C^1(\mathbb{R})$ and $\beta(t,x) \in C^1(Q)$ such that

$$d(t,x) = c(w(t,x)) + \beta(t,x) \ . \tag{2.11}$$

Moreover, function $\beta(t,x)$ satisfies

$$\|\beta(t,x)\|_{L^2(Q)} \leq C \sum_{\substack{i,j=0 \\ i\neq j}}^{n} \|\beta_{ij}\|_{L^2(Q)} \ , \tag{2.12}$$

and if $d(t,x), w(t,x) \in C^k(S)$, $\forall S \subset Q$, $\forall k \geq 1$, then

$$\|\beta(t,x)\|_{H^k(S)} \leq C\|w\|_{W^{k,\infty}(S)} \sum_{\substack{i,j=0 \\ i\neq j}}^{n} \|\beta_{ij}\|_{H^{k-1}(S)} \ . \tag{2.13}$$

Proof. Because $w(t,x)$ satisfy the condition (A1), we have like in Lemma 2.1, that there exists a function $\bar{c}$ such that

$$d(t,x) = \bar{c}(t, w, x_2, ..., x_n) \ . \tag{2.14}$$

Let $(t^0, x_1^0, ..., x_n^0)$ be an arbitrary point in Q, regard w as a variable in (2.14) and denote

$$c(w) = \bar{c}(t^0, w, x_2^0, ..., x_n^0) \ , \tag{2.15}$$

we see that

$$\begin{aligned} \beta(t,x) &= \beta(t, w(t,x), x_2, ..., x_n) \\ &= \bar{c}(t, w(t,x), x_2, ..., x_n) - \bar{c}(t^0, w(t,x), x_2^0, ..., x_n^0) \ . \end{aligned} \tag{2.16}$$

From this we see that $d(t,x) = c(w) + \beta(t,x)$ and these functions have the desired regularity. (The extension of c into $\mathbb{R}$ can be done as in Lemma 2.1). Next, we use condition (2.10) and relation (2.9) to get

$$\begin{aligned} D_i\beta &= D_i d - D_1 d (D_1 w)^{-1} D_i w \\ &= (D_1 w)^{-1}(D_i d D_1 w - D_1 d D_i w) \\ &= \frac{\beta_{i1}}{\frac{\partial w}{\partial x_1}} \ . \end{aligned} \tag{2.17}$$

Notice that

$$
\begin{aligned}
\beta(t,x) &= \bar{c}(t,w(t,x),x_2,...,x_n) - \bar{c}(t^0,w(t,x),x_2^0,...,x_n^0) \\
&= \bar{c}(t,w(t,x),x_2,...,x_n) - \bar{c}(t^0,w(t^0,x^0),x_2^0,...,x_n^0) + \\
&\quad + \bar{c}(t^0,w(t^0,x^0),x_2^0,...,x_n^0) - \bar{c}(t^0,w(t,x),x_2^0,...,x_n^0) \\
&= \int_0^1 \Big[\frac{d}{d\lambda}\Big(\bar{c}(t^0+\lambda(t-t^0),w(t^0+\lambda(t-t^0),x^0+\lambda(x-x^0)), \\
&\qquad\qquad x_2^0+\lambda(x_2-x_2^0),\cdots,x_n^0+\lambda(x_n-x_n^0))\Big) \\
&\quad - \frac{d}{d\lambda}\left(\bar{c}(t^0,w(t^0+\lambda(t-t^0),x^0+\lambda(x-x^0)),x_2^0,\cdots,x_n^0)\right)\Big]d\lambda \\
&= \int_0^1 \sum_{i\neq 1}\left(\frac{\beta_{i1}}{\frac{\partial w}{\partial x_1}}\right)\Big|_{(t^0+\lambda(t-t^0),x^0+\lambda(x-x^0))} \cdot (x_i - x_i^0)d\lambda\ .
\end{aligned}
\tag{2.18}
$$

Because Q is bounded and $\partial\Omega \in C^1$, we deduce that

$$
\begin{aligned}
&\|\beta(t,x)\|^2_{L^2(Q)} \\
&= \int_Q \left|\int_0^1 \sum_{i\neq 1}\left(\frac{\beta_{i1}}{\frac{\partial w}{\partial x_1}}\right)\Big|_{(t^0+\lambda(t-t^0),x^0+\lambda(x-x^0))} \cdot (x_i - x_i^0)d\lambda\right|^2 dxdt \\
&\leq C\delta^{-2}\sum_{i\neq 1}\int_Q |\beta_{i1}|^2 dxdt \leq C\sum_{i\neq 1}\int_Q |\beta_{i1}|^2 dxdt
\end{aligned}
\tag{2.19}
$$

Estimate (2.19) is true if $\vec{\nu}$ is in the x_1 direction. If $\vec{\nu}$ is not in the x_1 direction, we get (2.12) from (2.19).

We see that if $d, w \in C^k(S)$, then $\beta \in C^k(S)$, and by differentiating (2.17) k times, we get that

$$
\|\beta\|_{H^k(S)} \leq C\|w\|_{W^{k,\infty}(S)} \sum_{\substack{i,j=0 \\ i\neq j}}^{n} \|\beta_{ij}\|_{H^{k-1}(S)}\ .
$$

Lemma 2.3. *Assume that Q_l, $l = 1,2,...,m$ are nonoverlapping subdomains of Q, and $\bar{Q} = \cup_{l=1}^m \bar{Q}_l$. Let $d(t,x) \in C^0(Q)\cap C^1(Q_l)\cap H^k(Q_l)$, $w(t,x) \in W^{k,\infty}(Q_l)\cap C^1(\overline{Q})$, $l = 1,2,\cdots,m$, $k > 1$. Moreover, assume that $w(t,x)$ satisfies condition (A1), and*

$$
D_i d D_j w = D_j d D_i w + \beta_{ij}(t,x), \quad i,j = 0,1,2,...,n,\ i\neq j\ . \tag{2.20}
$$

Then there exist functions $c \in W^{1,\infty}(\mathbb{R})$ and $\beta(t,x) \in C^0(Q)$ such that

$$
d(t,x) = c(w(t,x)) + \beta(t,x)\ , \tag{2.21}
$$

and

$$\|\beta(t,x)\|_{L^2(Q)} \leq C \sum_{\substack{i,j=0\\ i\neq j}}^{n} \|\beta_{ij}\|_{L^2(Q)} , \tag{2.22}$$

$$\sum_{l=1}^{m} \|\beta(t,x)\|_{H^k(Q_l)} \leq C \sum_{l=1}^{m} \|w\|_{W^{k,\infty}(Q_l)} \sum_{\substack{i,j=0\\ i\neq j}}^{n} \|\beta_{ij}\|_{H^{k-1}(Q_l)} . \tag{2.23}$$

Proof. As $w(t,x) \in C^1(Q)$, $d(t,x) \subset C^0(Q)$, we can as in Lemma 3.1 to prove that there exists $\bar{c}$ such that

$$d(t,x) = \bar{c}(t,w,x_2,...,x_n) .$$

By choosing $c(w)$ in a same way, we shall get

$$d(t,x) = c(w) + \beta(t,x) .$$

As $d(t,x) \in C^0(Q)$, and $d(t,x) \in C^1(Q_l)$, $\forall l$, we can see that $c \in W^{1,\infty}(I)$

$$I = (\min_{x\in Q} w, \max_{x\in Q} w) \quad \text{and} \quad \beta \in C^0(Q).$$

By one extension, we can have $c \in W^{1,\infty}(\mathbb{R})$. As d is piecewise differentiable, so we use the technique as in (2.17)–(2.19) to have that

$$\|\beta(t,x)\|_{L^2(Q_l)} \leq C \sum \|\beta_{ij}(t,x)\|_{L^2(Q_l)} , \quad \forall l .$$

Summing up for $l = 1,2,...,m$, we get (2.22). Relation (2.16) shows that

$$\beta(t,x) = d(t,x) - d(t^0, e(t^0, w(t,x), x_2^0, ..., x_n^0), x_2^0, ..., x_n^0) .$$

Under the given conditions, the chain-rule and the rule for differentiation for product of functions are valid for the weak derivatives, so we can have

$$\begin{aligned} D_i\beta &= D_i d - D_1 d \cdot \frac{\partial e}{\partial w} \cdot D_i w \\ &= \frac{D_i d D_1 w - D_1 d D_i w}{\frac{\partial w}{\partial x_1}} = \frac{\beta_{i1}}{\frac{\partial w}{\partial x_1}} . \end{aligned} \tag{2.24}$$

The above relation is valid because $w \in C^1(Q)$, and so $\frac{\partial e}{\partial w} = (\frac{\partial w}{\partial e})^{-1} = (\frac{\partial w}{\partial x_1})^{-1}$. In each subdomain Q_i, as $d \in H^k(Q_i)$, we can differentiate (2.24) in each Q_i and as the rule for differentiation of product of functions is also correct for weak derivatives, we get that estimate (2.23) is true.

3. The identification problem and its error estimate

In this section, we first formulate the output-least-squares method for our identification problem and then prove a corresponding error estimate for this method. As we do not want to confine ourselves to a specific time discretization method, we will treat semi-discrete problems in the error analysis. However, in order to have an error estimate for the parameter, this will also force us to have two observations in the identification. We need one observation for the state, and another observation for the velocity. Fortunately, in application problems, often it is difficult to increase the observation points for the space variables, but it is relatively easy to shorten the observation time interval, and so we can get a believable velocity observation.

Let $z(t,x)$ be a distributed $L^2(Q)$ observation of the state u and let ϕ be an observation for the velocity $D_t u$. We assume the observation errors are as the following:

$$\|z-u\|_{L^2(Q)} \le \varepsilon_1, \quad \|\phi - D_t u\|_{L^2(Q)} \le \varepsilon_2 \ . \tag{3.1}$$

We remind, that the state equation in weak form states as follows:

$$\begin{cases} (D_t u, v) + (b\nabla u, \nabla v) = (f, v) + \langle g, v\rangle & \forall v \in H^1(\Omega)\ , \\ u(0,x) = u_0(x) & \text{in } \Omega\ . \end{cases} \tag{3.2}$$

The semidiscrete finite element approximation of (3.2) is of the form

$$\begin{cases} (D_t u_h, v_h) + (b\nabla u_h, \nabla v_h) = (f, v_h) + \langle g, v_h\rangle & \forall v_h \in S_h^{r+1}\ , \\ u(0,x) = L_h u_0(x) & \text{in } \Omega\ . \end{cases} \tag{3.3}$$

where $L_h u_0(x)$ is the $L^2(\Omega)$ projection of u_0 into S_h^{r+1}. In the following, for a given b, we always use $u_h(b)$ to denote the corresponding solution of (3.3).

As proposed in the introduction, we simplify the identification of the nonlinear parameter by replacing it with a linear one. We then recover the nonlinear parameter from the calculated linear parameter. We also use a semidiscrete scheme for the parameter and the analysis carries over to the case that we discretize the parameter in the time direction by a r-order spline function. In order to identify the parameter, we try to find a minimizer for:

$$\begin{aligned} \mathcal{J}_\mu(b) = & \int_Q |u_h(b) - z|^2\, dx\, dt + h^2 \int_Q |D_t u_h(b) - \phi|^2\, dx\, dt \\ & + \mu h^2 \sum_{i,j=0}^{n} \int_Q |D_i u_h(b) D_j b - D_j u_h(b) D_i b|^2\, dx\, dt\ . \end{aligned} \tag{3.4}$$

In (3.4), $\mu > C_0 > 0$ is a penalization parameter, $u_h(b)$ is a solution of equation (3.3) with corresponding parameter b. Due to different amount of differentiation involved in different terms of the cost function, we will loss convergence order in the error analysis. We put h^2 in for some terms in the cost function. This will help us to get one order of convergence back. For the sake of error analysis, we will try to identify b from S_h^r and the actual identification problem is defined as:

(P) Find $b_h(t,\cdot) \in S_h^r \cap M$ such that $\mathcal{J}_\mu(b_h) \le \mathcal{J}_\mu(\tilde{b}_h)$, $\forall \tilde{b}_h(t,\cdot) \in S_h^r \cap M$, where

$$M = \{b \mid b \in C^1([0,T]; L^\infty(\Omega)),\ 0 < \lambda_1 \le b \le \lambda_2 < +\infty \text{ a.e. in } \Omega,\ \forall t \in [0,T]\}\ .$$

Concerning the smoothness of the functions in (1.1), we assume $u(t,x) \in C^1([0,T]; W^{r+2,\infty}(\Omega))$ and $b = a(u) \in C^1([0,T]; W^{r+1,\infty}(\Omega))$, where $r \ge 1$. For a given function v, we will use v_I to denote the interpolate of v in S_h^{r+1}. Our main results are in Theorem 3.6. Before the proof of the main theorem, we prove a few more lemmas.

Lemma 3.1. *Let θ_h be the $L^2(\Omega)$-projection of b into S_h^r for each $t \in [0,T]$. Then there exists a $h_0 > 0$ such that if $h < h_0$, we have*

$$\max_{t>0} \|u_h(\theta_h) - u\|_{L^2(\Omega)} + \|\nabla(u_h(\theta_h) - u)\|_{L^2(Q)} \le Ch^{r+1}\ , \tag{3.5}$$

$$\|D_t(u_h(\theta_h) - u)\|_{L^2(Q)} \le Ch^r\ . \tag{3.6}$$

Proof. As θ_h is the projection of b into S_h^r for each time level $t \in [0,T]$, it follows from [**9**] that

$$\|b - \theta_h\|_{L^p(\Omega)} \le C\|b\|_{W^{r+1,p}(\Omega)} h^{r+1}\ , \quad \forall 1 \le p \le \infty,\ \forall t \in [0,T]\ . \tag{3.7}$$

Therefore, taking $p = \infty$, we see that there exists a $h_0 > 0$ small enough, such that if $h < h_0$, then $\theta_h \in M$ and this h_0 depends only on b and λ_1, λ_2.

Let b_I be the interpolant of b in S_h^r for each time level $t \in [0,T]$, then (see Scott [**16**])

$$|||b - b_I|||_{W^{k,p}(\Omega)} \le C\|b\|_{W^{r+1,p}(\Omega)} h^{r+1-k}\ , \quad 0 \le k \le r,\ 1 \le p \le \infty,\ \forall t \in [0,T]\ .$$

By using inverse inequality, we see that

$$\begin{aligned}
|||b - \theta_h|||_{W^{k,p}(\Omega)} &\le |||b - b_I|||_{W^{k,p}(\Omega)} + |||b_I - \theta_h|||_{W^{k,p}(\Omega)} \\
&\le C\|b\|_{W^{r+1,p}(\Omega)} h^{r+1-k} + Ch^{-k}\|b_I - \theta_h\|_{L^p(\Omega)} \\
&\le C\|b\|_{W^{r+1,p}(\Omega)} h^{r+1-k} + Ch^{-k}\|b - b_I\|_{L^p(\Omega)} \\
&\quad + Ch^{-k}\|b - \theta_h\|_{L^p(\Omega)} \\
&\le C\|b\|_{W^{r+1,p}(\Omega)} h^{r+1-k}\ .
\end{aligned}$$

Thus, as $b \in C^1([0,T]; W^{r+1,\infty}(\Omega))$, we can assume that

$$\begin{aligned}
&\|D_t\theta_h\|_{L^\infty(\Omega)} \le C, \quad \forall h < h_0,\ \forall t \in [0,T]\ , \\
&|||\theta_h|||_{W^{k,p}(\Omega)} \le C, \quad 0 \le k \le r,\ 1 \le p \le \infty,\ \forall h < h_0,\ \forall t \in [0,T]\ .
\end{aligned} \tag{3.8}$$

From (3.3), we see that $u_h(\theta_h)$ is the solution of

$$\begin{cases} (D_t u_h(\theta_h), v_h) + (\theta_h \nabla u_h(\theta_h), \nabla v_h) = (f, v_h) + \langle g, v_h\rangle, & \forall v_h \in S_h^{r+1}\ , \\ u_h(\theta_h)(0,x) = L_h u_0(x) & \text{in } \Omega\ . \end{cases} \tag{3.9}$$

By combining (3.2) and (3.9), we get that

$$(D_t(u - u_h(\theta_h)), v_h) + (\theta_h \nabla(u - u_h(\theta_h)), \nabla v_h)$$
$$= ((\theta_h - b)\nabla u_h(\theta_h), \nabla v_h), \quad \forall v_h \in S_h^{r+1} ,$$

and so

$$\begin{aligned} &(D_t(u_I - u_h(\theta_h)), v_h) + (\theta_h \nabla(u_I - u_h(\theta_h)), \nabla v_h) \\ =&(D_t(u_I - u), v_h) + (\theta_h \nabla(u_I - u), \nabla v_h) + ((\theta_h - b)\nabla u, \nabla v_h), \\ &\forall v_h \in S_h^{r+1} . \end{aligned} \tag{3.10}$$

Taking $v_h = u_I - u_h(\theta_h) \in S_h^{r+1}$, we see that

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|u_I - u_h(\theta_h)\|^2 + (\theta_h \nabla(u_I - u_h(\theta_h)), \nabla(u_I - u_h(\theta_h))) \\ &= (D_t(u_I - u), u_I - u_h(\theta_h)) + (\theta_h \nabla(u_I - u), \nabla(u_I - u_h(\theta_h))) \\ &\quad + ((\theta_h - b)\nabla u, \nabla(u_I - u_h(\theta_h))) . \end{aligned}$$

By using the simple kick-back techniques, we will have:

$$\begin{aligned} &\frac{1}{2}\|u_I - u_h(\theta_h)\|^2 + \lambda_1 \int_0^t \|\nabla(u_I - u_h(\theta_h)\|^2 \, ds \\ &\leq \frac{1}{2}\int_0^t \|u_I - u_h(\theta_h)\| ds + \frac{1}{2}\int_0^t \|D_t(u_I - u)\|^2 \, ds \\ &+ \alpha \int_0^t \|\nabla(u_I - u_h(\theta_h)\|^2 \, ds + \frac{C}{4\alpha}\int_0^t \|\nabla(u_I - u)\|^2 \, ds \\ &+ \alpha \int_0^t \|\nabla(u_I - u_h(\theta_h))\|^2 \, ds + \frac{1}{4\alpha}\|\nabla u\|^2_{L^\infty(Q)} \int_0^t \|\theta_h - b\|^2 \, ds \\ &+ \frac{1}{2}\|u_I(0,x) - L_h u_0(x)\|^2 , \quad \forall \alpha > 0 . \end{aligned} \tag{3.11}$$

As u_I is the interpolant of u in S_h^{r+1}, we have

$$\begin{aligned} \|u - u_I\|_{H^k(\Omega)} &\leq Ch^{r+2-k}\|u\|_{H^{r+2}(\Omega)}, \quad && k = 0, 1, \ \forall t \in [0, T] , \\ \|D_t(u - u_I)\|_{H^k(\Omega)} &\leq Ch^{r+2-k}\|D_t u\|_{H^{r+2}(\Omega)}, \quad && k = 0, 1, \ \forall t \in [0, T] . \end{aligned}$$

Thus, by choosing α suitably, from (3.11) and the Gronwall's inequality, we get the following estimate:

$$\max_{t>0} \|u - u_h(\theta_h)\|^2 + \int_0^T \|\nabla(u - u_h(\theta_h))\|^2 \, ds \leq Ch^{2(r+1)} .$$

To prove (3.6), we take $v_h = D_t(u_I - u_h(\theta_h))$ in (3.10), and obtain:

$$\begin{aligned} &\|D_t(u_I - u_h(\theta_h))\|^2 + \frac{1}{2}\frac{d}{dt}\int_\Omega \theta_h |\nabla(u_I - u_h(\theta_h))|^2 \, ds \\ &+ \frac{1}{2}\int_\Omega \frac{d\theta_h}{dt} |\nabla(u_I - u_h(\theta_h))|^2 \, ds \\ &= (D_t(u - u_I), D_t(u_I - u_h(\theta_h))) + (\theta_h \nabla(u - u_I), \nabla D_t(u_I - u_h(\theta_h))) \\ &\quad + ((\theta_h - b)\nabla u, \nabla D_t(u_I - u_h(\theta_h))) \end{aligned} \tag{3.12}$$

Using the inverse inequality, it is true

$$\|\nabla D_t(u_I - u_h(\theta_h))\| \le Ch^{-1}\|D_t(u_I - u_h(\theta_h))\|, \quad \forall t > 0$$

and from $ab \le \alpha a^2 + \frac{1}{4\alpha}b^2$ and (3.12), it follows

$$\begin{aligned}
&\int_0^T \|D_t(u_I - u_h(\theta_h))\|^2\,ds + \max_{t>0}\|\nabla(u_I - u_h(\theta_h))\|^2 \\
&\le C\Big[\int_0^T \|D_t(u - u_I)\|^2\,ds + h^{-2}\int_0^T \|\nabla(u - u_I)\|^2\,ds \\
&\quad + \int_0^T \|\nabla(u_I - u_h(\theta_h))\|_0^2 ds + h^{-2}\int_0^T \|\theta_h - b\|^2\,ds + \|\nabla(u_I(0,x) - L_h u_0(x))\|^2\Big] \\
&\le Ch^{2r}\,,
\end{aligned}$$

which proves (3.6).

Lemma 3.2. *Let b_h be the minimizer of (P), and w_h be the corresponding state to b_h, i.e. $w_h = u_h(b_h)$. Then for $h < h_0$, we have*

$$\sum_{i,j=0}^{m} \int_Q |D_i w_h D_j b_h - D_j w_h D_i h_h|^2\,dx\,ds \le Ch^{2r} + \frac{2\varepsilon_1^2}{\mu h^2} + \frac{2\varepsilon_2^2}{\mu}\,, \tag{3.13}$$

$$\|w_h - u\|_{L^2(Q)}^2 \le C\mu h^{2(r+1)} + 6\varepsilon_1^2 + 4h^2\varepsilon_2^2\,, \tag{3.14}$$

$$\|D_t(w_h - u)\|_{L^2(Q)}^2 \le C\mu h^{2r} + \frac{4\varepsilon_1^2}{h^2} + 6\varepsilon_2^2\,. \tag{3.15}$$

Proof. As b_h is the minimizer of (P) and for $h < h_0$, it is true that $\theta_h \in M$, we have

$$\begin{aligned}
&\|w_h - z\|_{L^2(Q)}^2 + h^2\|D_t w_h - \phi\|_{L^2(Q)}^2 \\
&+ \mu h^2 \sum_{i,j=0}^{n} \int_Q |D_i w_h D_j b_h - D_j w_h D_i b_h|^2\,dx\,ds \\
&\le \|u_h(\theta_h) - z\|_{L^2(Q)}^2 + h^2\|D_t u_h(\theta_h) - \phi\|_{L^2(Q)}^2 \\
&\quad + \mu h^2 \sum_{i,j=0}^{n} \int_Q |D_i u_h(\theta_h) D_j\theta_h - D_j u(\theta_h) D_i\theta_h|^2 dx\,ds\,.
\end{aligned} \tag{3.16}$$

Let us denote

$$\begin{aligned}
I_1 &= \|u_h(\theta_h) - z\|_{L^2(Q)}^2\,, \\
I_2 &= \|D_t u_h(\theta_h) - \phi\|_{L^2(Q)}^2\,, \\
I_3 &= \sum_{i,j=0}^{n} \int_Q |D_i u_h(\theta_h) D_j\theta_h - D_j u_h(\theta_h) D_i\theta_h|^2\,dx\,ds\,.
\end{aligned}$$

From (3.5) and (3.6) of Lemma 3.1 we have

$$
\begin{aligned}
I_1 &\le 2\|u_h(\theta_h) - u\|^2_{L^2(Q)} + 2\|u - z\|^2_{L^2(Q)} \\
&\le Ch^{2(r+1)} + 2\varepsilon_1^2 , \\
I_2 &\le 2\|D_t(u_h(\theta_h) - u)\|^2_{L^2(Q)} + 2\|D_t u - \phi\|^2_{L^2(Q)} \\
&\le Ch^{2r} + 2\varepsilon_2^2 .
\end{aligned}
$$

As $b = a(u)$, so $D_i u D_j b - D_j u D_i b = 0$, $\forall i, j$, and (3.8) is true, we deduce

$$
\begin{aligned}
I_3 &\le C \sum_{i,j=0}^{n} \int_Q \big(|D_i(u_h(\theta_h) - u)D_j\theta_h|^2 + |D_j(u_h(\theta_h) - u)D_i\theta_h|^2 \\
&\quad + |D_i u D_j(b - \theta_h)|^2 + |D_j u D_i(b - \theta_h)|^2\big)\, dx\, dt \\
&\le C\Big(\int_0^T \|\nabla(u_h(\theta_h) - u)\|^2\, ds + \int_0^T \|D_t(u_h(\theta_h) - u)\|^2\, ds \\
&\quad + \int_0^T \|\nabla(b - \theta_h)\|^2\, ds + \int_0^T \|D_t(b - \theta_h)\|^2\, ds\Big) \\
&\le Ch^{2r} .
\end{aligned}
$$

Let us denote $I = I_1 + h^2 I_2 + \mu h^2 I_3$. As $\mu > C_0 > 0$, it is clear that

$$
\begin{aligned}
I &\le Ch^{2(r+1)} + 2\varepsilon_1^2 + Ch^{2(r+1)} + 2\varepsilon_2^2 h^2 + \mu h^{2(r+1)} \\
&\le C\mu h^{2(r+1)} + 2\varepsilon_1^2 + 2\varepsilon_2^2 h^2 .
\end{aligned}
$$

From (3.16), it is simple to observe that

$$
\begin{aligned}
&\sum_{i,j=0}^{n} \int_Q |D_i w_h D_j b_h - D_j w_h D_i b_h|^2\, dx\, ds \le \mu^{-1} h^{-2} I, \\
&\|w_h - u\|^2_{L^2(Q)} \le 2\|w_h - z\|^2_{L^2(Q)} + 2\|u - z\|^2_{L^2(Q)} \le 2I + 2\varepsilon_1^2,
\end{aligned}
$$

and

$$
\begin{aligned}
\|D_t(w_h - u)\|^2_{L^2(Q)} &\le 2\|D_t w_h - \phi\|^2_{L^2(Q)} + 2\|D_t u - \phi\|^2_{L^2(Q)} \\
&\le 2h^{-2} I + 2\varepsilon_2^2 .
\end{aligned}
$$

This proves the lemma.

For simplicity, in the following we will denote

$$
e(h, \mu, \varepsilon) = \mu h^{2r} + \frac{\varepsilon_1^2}{h^2} + \varepsilon_2^2 . \tag{3.17}
$$

Corollary 3.3. *For the minimizer b_h of (P), we have*

$$\sum_{i,j=0}^{n} \int_Q |D_i u D_j b_h - D_j u D_i b_h|^2 \, dx\, ds \le C h^{-2} e(h, \mu, \varepsilon) \ . \tag{3.18}$$

Proof. We see that

$$\begin{aligned}
&\int_Q |D_i u D_j b_h - D_j u D_i b_h|^2 \, dx\, ds \\
&\le 2 \int_Q |D_i(u - w_h) D_j b_h - D_j(u - w_h) D_i b_h|^2 \, dx\, ds \\
&\quad + 2 \int_Q |D_i w_h D_j b_h - D_j w_h D_i b_h|^2 \, dx\, ds, \quad \forall i, j \ .
\end{aligned}$$

We use inverse equality to estimate

$$\begin{aligned}
&\|D_j b_h\|_{L^\infty(Q)} \le C h^{-1} \|b_h\|_{L^\infty(Q)} \le C h^{-1}, \quad \forall j, \\
&\int_Q |D_i(u - w_h)|^2 \, dx\, ds \\
&= \int_0^T \|D_i(u - w_h)\|^2 \, ds \\
&\le 2 \int_0^T \|D_i(u - u_I)\|^2 \, ds + 2 \int_0^T \|D_i(u_I - w_h)\|^2 \, ds \\
&\le 2 \int_0^T \|D_i(u - u_I)\|^2 \, ds + C h^{-2} \int_0^T \|u_I - w_h\|^2 \, ds \\
&\le C h^{2(r+1)} + C h^{-2} \int_0^T \|u - w_h\|^2 \, ds \\
&\le C e(h, \mu, \varepsilon) \ , \\
&\int_Q \|D_t(u - w_h)\|^2 \, ds \le C e(h, \mu, \varepsilon) \ ,
\end{aligned}$$

and so we get (3.18).

Remark 3.4. As we do not know the boundness of $\|D_i b_h\|_{L^\infty(Q)}$, we use inverse property to estimate it. This makes the error estimate to lose one order of convergence. Due to the different amount of differentiation used in the cost function in (3.4), we lose another order of convergence. We will see in the final error estimate that we loss totally together two orders of convergence.

Corollary 3.5. *For the minimizer b_h of (P), there exists $a_h(\cdot) \in W^{1,\infty}(\mathbb{R})$ and $\beta_h(t,x) \in C^0(Q)$ such that*

$$b_h(t,x) = a_h(u(t,x)) + \beta_h(t,x) \ , \tag{3.19}$$

and

$$\|\beta_h(t,x)\|^2_{L^2(Q)} \le Ch^{-2}e(h,\mu,\varepsilon) \tag{3.20}$$

$$|||\beta_h|||^2_{H^{r+1}(\Omega)} \le C + Ch^{-2(r-1)}\|b_I - b_h\|^2 + Ch^{-2r}\sum_{i,j=0}^{n}\|D_iuD_jb_h - D_juD_ib_h\|^2 \ , \quad \forall t \in [0,T]. \tag{3.21}$$

Proof. As $u \in C^1([0,T]; W^{r+2,\infty}(\Omega)) \subset C^1(\overline{Q})$, $b_h \in C^1([0,T]; C^\infty(e))$, $\forall e \in \mathcal{T}_h$, we see from Lemma 2.3 that there exists $a_h(\cdot) \in W^{1,\infty}(\mathbb{R})$ and $\beta_h(t,x) \in C^0(Q)$ such that (3.19) holds. Estimate (3.20) follows from (2.22) and (3.18). To prove (3.21), let us notice that from (2.24) we can have that

$$|||\beta_h|||_{H^{r+1}(\Omega)} \le C\|u\|_{W^{r+1,\infty}(\Omega)}\sum_{i,j=0}^{n}|||\beta_{ij}|||_{H^r(\Omega)}$$

with $\beta_{ij} = D_iuD_jb_h - D_juD_ib_h$. Next, we estimate $|||\beta_{ij}|||_{H^r(\Omega)}$. We will use inverse property.

$$\begin{aligned}
|||\beta_{ij}|||^2_{H^r(\Omega)} &= \sum_{e\in\mathcal{T}_h}\|\beta_{ij}\|^2_{H^r(e)} \\
&= \sum_{e\in\mathcal{T}_h}\|D_iuD_jb_h - D_juD_ib_h\|^2_{H^r(e)} \\
&\le C\sum_{e\in\mathcal{T}_h}\left[\|u-u_I\|^2_{W^{r+1,\infty}(e)}\|b_h\|^2_{H^r(e)} + \|D_iu_ID_jb_h - D_ju_ID_ib_h\|^2_{H^r(e)}\right] \\
&\le C\sum_{e\in\mathcal{T}_h}\left[\|u-u_I\|^2_{W^{r+1,\infty}(e)}\|b_I\|^2_{H^r(e)} + \|u-u_I\|^2_{W^{r+1,\infty}(e)}\|b_I - b_h\|^2_{H^r(e)}\right. \\
&\quad \left. + \|D_iu_ID_jb_h - D_ju_ID_ib_H\|^2_{H^r(e)}\right] \\
&\le C\big(h^2\|u\|_{W^{r+2,\infty}(\Omega)}\|b\|^2_{H^{r+1}(\Omega)} + h^2\|u\|_{W^{r+2,\infty}(\Omega)}h^{-2r}\|b_I - b_h\| \\
&\quad + h^{-2r}\|D_iu_ID_jb_h - D_ju_ID_ib_h\|^2\big) \\
&\le C\big(h^2 + h^{-2(r-1)}\|b_I - b_h\|^2 + h^{-2r}\|D_iuD_jb_h - D_juD_ib_h\|^2 \\
&\quad + h^{-2r}\|D_i(u_I-u)D_jb_h\|^2 + h^{-2r}\|D_j(u_I-u)D_ib_h\|^2\big) \\
&\le C + Ch^{-2(r-1)}\|b_I - b_h\|^2 + Ch^{-2r}\|D_iuD_jb_h - D_juD_ib_h\|^2 \ .
\end{aligned}$$

Theorem 3.6. *For $h < h_0$ the identified parameter $a_h(\cdot)$ and the real parameter $a(\cdot)$ satisfy the following error estimate:*

$$\|a(u) - a_h(u)\|_{L^2(Q)} \le C(h^{r-1} + h^{-2}\varepsilon_1 + h^{-1}\varepsilon_2) \ . \tag{3.22}$$

Proof. As $S_h^{r+1} \subset H^1(\Omega)$, we see that

$$(D_tw_h, v_h) + (b_h\nabla w_h, \nabla v_h) = (f, v_h) + \langle g, v_h\rangle, \quad \forall v_h \in S_h^{r+1} \ , \tag{3.23}$$

$$(D_tu, v_h) + (b\nabla u, \nabla v_h) = (f, v_h) + \langle g, v_h\rangle, \quad \forall v_h \in S_h^{r+1}. \tag{3.24}$$

Subtracting (3.23) from (3.24), we get

$$((b-b_h)\nabla u, \nabla v_h) = (D_t(w_h-u), v_h) + (b_h \nabla(w_h-u), \nabla v_h) \ . \tag{3.25}$$

As $b = a(u)$, $b_h = a_h(u) + \beta_h$, relation (3.25) gives

$$((a(u)-a_h(u))\nabla u, \nabla v_h) = (D_t(w_h-u), v_h) + (b_h \nabla(w_h-u), \nabla v_h) + (\beta_h \nabla u, \nabla v_h) \ ,$$

and so

$$\begin{aligned} ((a(u)-a_h(u))\nabla u, \nabla v) &= (D_t(w_h-u), v_h) + (b_h \nabla(w_h-u), \nabla v_h) \\ &+ (\beta_h \nabla u, \nabla v_h) + ((a(u)-a_h(u))\nabla u, \nabla(v-v_h)) \ . \end{aligned} \tag{3.26}$$

We take

$$v(t,x) = \int_{u_Q}^{u(t,x)} (a(s)-a_h(s))\, ds$$

where, u_Q is a constant to be chosen. We note that

$$\begin{aligned} D_t v &= (a(u)-a_h(u))D_t u \ , \\ \nabla v &= ((a(u)-a_h(u))\nabla u \ . \end{aligned} \tag{3.27}$$

Let v_I be the interpolate of v in S_h^{r+1}. Then as in Falk [**10**], we can show that

$$\begin{aligned} &\|\nabla(v-v_I)\|^2 + \|v-v_I\|^2 \\ &= \sum_{e\in\mathcal{T}_h} (\|\nabla(v-v_I)\|_{L^2(e)}^2 + \|v-v_I\|_{L^2(e)}^2) \\ &\le \sum_{e\in\mathcal{T}_h} Ch^{2(r+1)} |v|_{H^{r+2}(e)}^2 \ , \quad \forall t\in[0,T]. \end{aligned} \tag{3.28}$$

Here, $|\cdot|_{H^{r+2}(e)}$ means the seminorm of $H^{r+2}(e)$. From (3.27) we know that

$$\begin{aligned} &\sum_{e\in\mathcal{T}_h} |v|_{H^{r+2}(e)}^2 \\ &\le \sum_{e\in\mathcal{T}_h} \|u\|_{W^{r+2,\infty}(e)}^2 \|a(u)-a_h(u)\|_{H^{r+1}(e)}^2 \\ &\le 2\|u\|_{W^{r+2,\infty}(\Omega)}^2 \sum_{e\in\mathcal{T}_h} (\|b-b_h\|_{H^{r+1}(e)}^2 + \|\beta_h\|_{H^{r+1}(e)}^2) \\ &\le C(|||b-b_I|||_{H^{r+1}(\Omega)}^2 + |||b_I-b_h|||_{H^{r+1}(\Omega)}^2 + |||\beta_h|||_{H^{r+1}(\Omega)}^2) \\ &\le C + C|||b_I-b_h|||_{H^r(\Omega)}^2 + C|||\beta_h|||_{H^{r+1}(\Omega)}^2 \\ &\le C + Ch^{-2r}\|b_I-b_h\|^2 + Ch^{-2r}\sum_{i,j=0}^{n} \|D_i u D_j b_h - D_j u D_i b_h\|^2 \ . \end{aligned} \tag{3.29}$$

Therefore

$$
\begin{aligned}
&\int_0^T \left(\|\nabla(v-v_I)\|^2 + \|v-v_I\|^2\right) ds \\
&\le Ch^{2(r+1)} + Ch^2\|b_I - b_h\|^2_{L^2(Q)} + Ch^2 \sum_{i,j=0}^{n} \|D_iuD_jb_h - D_juD_ib_h\|^2_{L^2(Q)} \\
&\le Ch^{2(r+1)} + Ch^2\|b_I - b_h\|^2_{L^2(Q)} + Ce(h,\mu,\varepsilon) \\
&\le Ch^{2(r+1)} + Ch^2\|a(u) - a_h(u)\|^2_{L^2(Q)} + Ce(h,\mu,\varepsilon) \, .
\end{aligned} \tag{3.30}
$$

Next, we take the constant u_Q such that

$$
\|v\|^2_{L^2(\Omega)} \le C\Big(\int_0^T \|\nabla v\|^2\, ds + \int_0^T \|D_t v\|^2\, ds\Big) \, .
$$

More precisely, taking $u_Q = \frac{1}{|Q|}\int_Q u\, dx\, ds$, we have

$$
\min_{x\in Q} u \le u_Q \le \max_{x\in Q} u \, .
$$

As $u \in C^1([0,T], W^{r+2}(\Omega)) \subset C^1(\overline{Q})$, and $\nabla u \cdot v \ge \delta > 0$, there exists a smooth curve $\Gamma_Q \subset Q$, whose n-dimensional measure is non-zero, such that

$$
u\big|_{\Gamma_Q} = u_Q \, .
$$

From the definition of v, we know

$$
v\big|_{\Gamma_Q} = 0 \, .
$$

By using the Poincare's inequality we get

$$
\begin{aligned}
&\|v\|^2_{L^2(Q)} \\
&\le C\Big(\int_0^T \|\nabla v\|^2\, ds + \int_0^T \|D_t v\|^2\, ds\Big) \\
&\le C\int_0^T \|a(u) - a_h(u)\|^2\, ds \, .
\end{aligned} \tag{3.31}
$$

Estimate (3.31) shows that

$$
\begin{aligned}
&\int_0^T \|v_I\|^2\, ds + \int_0^T \|\nabla v_I\|^2\, ds \\
&\le C\int_0^T \|a(u) - a_h(u)\|^2\, ds + C\int_0^T \|v_I - v\|^2_{H^1(\Omega)}\, ds \, .
\end{aligned} \tag{3.32}
$$

Using estimates (3.30) and (3.32), and taking $v_h = v_I$ in (3.26), we conclude by a simple kick-back technique that for any $\alpha > 0$

$$
\int_0^T \|a(u) - a_h(u)\|^2\, ds
$$

$$
\begin{aligned}
&\le \alpha \int_0^T \|v_I\|^2\, ds + \frac{1}{4\alpha} \int_0^T \|D_t(w_h - u)\|^2\, ds \\
&\quad + \alpha \int_0^T \|\nabla v_I\|^2\, ds + \frac{C}{4\alpha} \int_0^T \|\nabla (w_h - u)\|^2\, ds \\
&\quad + \alpha \int_0^T \|\nabla v_I\|^2\, ds + \frac{C}{4\alpha} \int_0^T \|\beta_h\|^2\, ds \\
&\quad + C \int_0^T \|a(u) - a_h(u)\|\, \|\nabla (v - v_I)\|\, ds \\
&\le C\alpha \int_0^T \|a(u) - a_h(u)\|^2\, ds + C\alpha \int_0^T \|v_I - v\|^2\, ds \\
&\quad + \frac{C}{4\alpha} \int_0^T \left(\|D_t(w_h - u)\|^2 + \|\nabla (w_h - u)\|^2\right) ds \\
&\quad + \frac{C}{4\alpha} h^{-2} e(h,\mu,\varepsilon) + C \int_0^T \|a(u) - a_h(u)\| \\
&\quad \cdot \left(h^{(r+1)} + h^{-1}\sqrt{e(h,\mu,\varepsilon)} + h\|a(u) - a_h(u)\|\right) ds \\
&\le C\alpha \int_0^T \|a(u) - a_h(u)\|^2\, ds + C\alpha \int_0^T \|v_I - v\|^2\, ds \\
&\quad + \frac{C}{4\alpha} \int_0^T \left(\|D_t(w_h - u)\|^2 + \|\nabla (w_h - u)\|^2\right) ds \\
&\quad + \frac{C}{4\alpha} h^{-2} e(h,\mu,\varepsilon) + C\alpha \int_0^T \|a(u) - a_h(u)\|^2\, ds \\
&\quad + \frac{C}{4\alpha} h^{2(r+1)} + \frac{C}{4\alpha} h^{-2} e(h,\mu,\varepsilon) + Ch \int_0^T \|a(u) - a_h(u)\|^2\, ds\ .
\end{aligned}
$$

By choosing α suitably and assuming h_0 small enough, we get

$$
\begin{aligned}
&\int_0^T \|a(u) - a_h(u)\|^2\, ds \\
&\le C \int_0^T \left(\|D_t(w_h - u)\|^2 + \|\nabla (w_h - u)\|^2\right) ds \\
&\quad + Ch^{-2} e(h,\mu,\varepsilon) + Ch^{2(r+1)}\ .
\end{aligned}
\tag{3.33}
$$

This proves Theorem 3.6. Due to the reason stated in Remark 3.4, we lose two orders of convergence in this estimate. For one dimensional problems, optimal convergence order was obtained in [**18**] and [**19**].

4. Numerical approximation and test results

In this chapter we present some numerical results, which are based on the theory of previous chapters. We will concentrate only on 1-d case, since it already contains all the important aspects of our method. Moreover, because we are dealing with the parabolic equations, the number of unknows for calculating 2-d and 3-d cases would become very high. Then, in those cases, we would be testing the performance of different optimization algorithms in solving our parameter identification problems, and try to find one efficient minimization routine.

Let us first recall the usual semidiscrete approximation of the state equation using finite element method. We seek for a weak solution $u_h(\cdot, x) \in S_h^{r+1}$ of the form $u_h = \sum_i \tilde{u}_i(t)\,\phi_i(x)$, where $\tilde{u}_i(t)$ are the unknown time-dependent coefficients and $\phi_i(x)$ are the basis functions of the expansion with a fixed discretization parameter h. For the time discretization, we will use the Crank-Nicolson scheme. For $1 \leq l \in \mathbb{N}$, we define $\Delta t = \frac{T}{l}$ and $t_k = k\,\Delta t, k = 0, ..., l$, and $t_{k+\frac{1}{2}} = (k + \frac{1}{2})\,\Delta t$ for $0 \leq k \leq l-1$. Moreover, we define $f_k = f(t_k), u_h^k = \sum_i \tilde{u}_i(t_k)\phi(x), k = 1, ..., l$ recursively by

$$\begin{aligned}&\frac{1}{2}(\frac{u_h^{k+1} - u_h^k}{\Delta t}, v_h) + \frac{1}{2}(b\,\nabla(u_h^{k+1} + u_h^k), \nabla v_h) = \\ &\quad \frac{1}{2}(f(t_{k+1}) + f(t_k), v_h) + \frac{1}{2}\langle g(t_{k+1}) + g(t_k), v_h\rangle \quad \forall v_h \in S_h^{r+1}\,,\end{aligned} \tag{4.1}$$

where $u_h^0 = L_h u_0(x)$.

For the parameter function b we also introduce a discrete form from $S_{\bar{h}}^r$-space with a discretization parameter $\bar{h}$, which can be different from h. For inverse problems, it is always better to take a smaller mesh size for the solution u than for the parameter b. In our computations, we will take h equal to $\bar{h}$ divided by an integer number. In this case, all our error estimates hold.

Next we introduce the numerical realization of the cost-functional. Recall, that the continuos cost functional is of the form

$$\begin{aligned}\mathcal{J}_\mu(b_{\bar{h}}) = &\int_0^T \int_\Omega (|u_h(b_{\bar{h}}) - z|^2 + h^2\,|D_t u_h(b_{\bar{h}}) - \phi|^2)\,dx\,dt \\ &+ \mu\, h^2 \sum_{\substack{i,j=0 \\ i\neq j}}^{n} \int_0^T \int_\Omega |D_i u_h D_j b_{\bar{h}} - D_j u_h D_i b_{\bar{h}}|^2\,dx\,dt\,.\end{aligned} \tag{4.2}$$

Now we must apply some suitable quadrature formulas for the calculation of the integrals. For the integration with respect to time variable t we use the trapezoid rule in the first term of (4.2) and for the other two term of the cost functional we use the mid-point rule.

For the space discretization with $n = 1$ the domain Ω reduces to an interval $[a, b]$. In the sequel we assume this to be the standard unit interval $[0, 1]$. For this interval we formulate the equidistant discretization points as $x_j = jh, j = 0, ..., n,\ h = \frac{1}{n}$ for the approximation u_h. We apply the trapezoid rule also for the space integrals. We will drop the common factor $h\Delta t$ from each term of the cost functional. After doing this, the discrete cost functional to be minimized reads as

$$\mathcal{J}_{\mu,h}(b_{\bar{h}}) = \mathcal{J}_1 + h^2\mathcal{J}_2 + \mu h^2\mathcal{J}_3\,, \tag{4.3}$$

where

$$
\begin{aligned}
\mathcal{J}_1 = & \sum_{k=1}^{l-1} \Big(\sum_{i=1}^{n-1} |u_h(t_k, x_i) - z(t_k, x_i)|^2 \\
& + \frac{1}{2}(|u_h(t_k, 0) - z(t_k, 0)|^2 + |u_h(t_k, 0) - z(t_k, 0)|^2) \Big) \\
& + \frac{1}{2} \sum_{i=1}^{n-1} |u_h(T, x_i) - z(T, x_i)|^2 \\
& + \frac{1}{4}(|u_h(T, 0) - z(T, 0)|^2 + |u_h(T, 1) - z(T, 1)|^2) ,
\end{aligned}
\tag{4.4}
$$

$$
\begin{aligned}
\mathcal{J}_2 = & \sum_{k=0}^{l-1} \Big(\sum_{i=1}^{n-1} |\partial_t u_h(t_{k+\frac{1}{2}}, x_i) - \phi(t_{k+\frac{1}{2}}, x_i)|^2 \\
& + \frac{1}{2}(|\partial_t u_h(t_{k+\frac{1}{2}}, 0) - \phi(t_{k+\frac{1}{2}}, 0)|^2 \\
& + |\partial_t u_h(t_{k+\frac{1}{2}}, 1) - \phi(t_{k+\frac{1}{2}}, 1)|^2) \Big) ,
\end{aligned}
\tag{4.5}
$$

$$
\begin{aligned}
\mathcal{J}_3 = & \sum_{k=0}^{l-1} \Big(\sum_{i=1}^{n-1} |\partial_t u_h(t_{k+\frac{1}{2}}, x_i)\, D_x b_{\bar{h}}(t_{k+\frac{1}{2}}, x_i) \\
& - D_x u_h(t_{k+\frac{1}{2}}, x_i)\, \partial_t b_{\bar{h}}(t_{k+\frac{1}{2}}, x_i)|^2 \\
& + \frac{1}{2}(|\partial_t u_h(t_{k+\frac{1}{2}}, 0)\, D_x b_{\bar{h}}(t_{k+\frac{1}{2}}, 0) - D_x u_h(t_{k+\frac{1}{2}}, 0)\, \partial_t b_{\bar{h}}(t_{k+\frac{1}{2}}, 0)|^2 \\
& + |\partial_t u_h(t_{k+\frac{1}{2}}, 1)\, D_x b_{\bar{h}}(t_{k+\frac{1}{2}}, 1) - D_x u_h(t_{k+\frac{1}{2}}, 1)\, \partial_t b_{\bar{h}}(t_{k+\frac{1}{2}}, 1)|^2) \Big) .
\end{aligned}
\tag{4.6}
$$

Above we have denoted by $\partial_t u_h = \frac{u_h(t_{k+1}) - u_h(t_k)}{\Delta t}$ and $\partial_t b_{\bar{h}} = \frac{b_{\bar{h}}(t_{k+1}) - b_{\bar{h}}(t_k)}{\Delta t}$. In the cost functional, we only need the value of $D_x b_{\bar{h}}$ at the nodal points, we take its value as the average of its values in the two neighboring elements.

The following pictures illustrate the error, which is calculated by minimizing the discrete cost functional (4.3) with E04JAF-optimization routine from NAG-library using different values for the discretization parameters.

Example 1. *We take $u(t, x) = \exp(-t)\exp(\frac{x}{2})$ and $a(u) = \exp(u)$. We use third-order Lagrange basis for the parameter in $x-$direction and second order Lagrange approximation with respect to t. A third-order Hermite basis is used for the solution u. In all cases of this example the discretization parameters are $h = \frac{1}{12}, \bar{h} = \frac{1}{4}$ and $\Delta t = \frac{1}{9}$. The value of the observation z and ϕ is taken as the value of u and u_t without observation errors. Figures 1–4 show the computed results with different values of μ and initial guesses.*

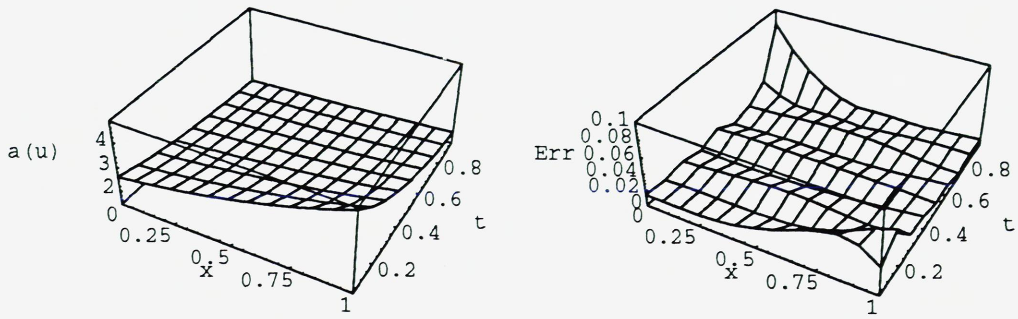

Figure 1: $a(u(t,x))$ and error $a(u(t,x)) - a_{\bar{h}}(u(t,x))$ at nodal points with $\mu = 1$ and init. guess 5.0

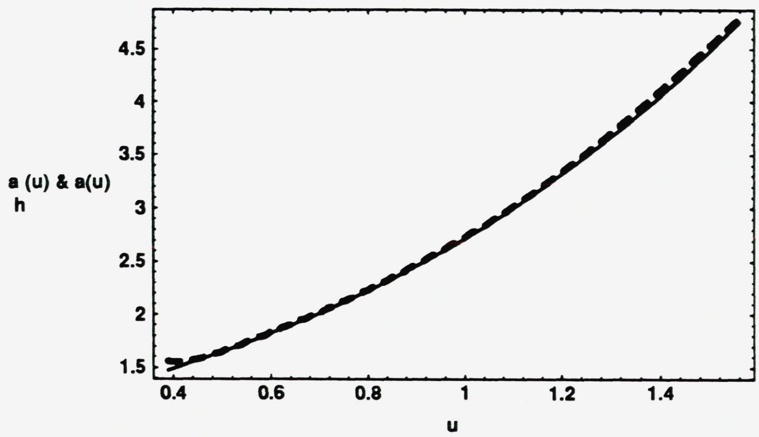

Figure 2: $a_h(u)$ and $a(u)$ with $\mu = 1$ and init. guess 5.0; max.err: 0.08.

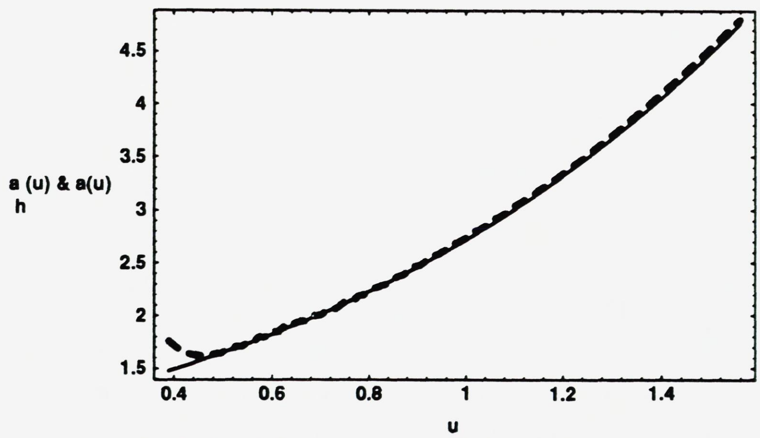

Figure 3: $a_h(u)$ and $a(u)$ with $\mu = 100$ and init. guess 3.0; max.err: 0.28

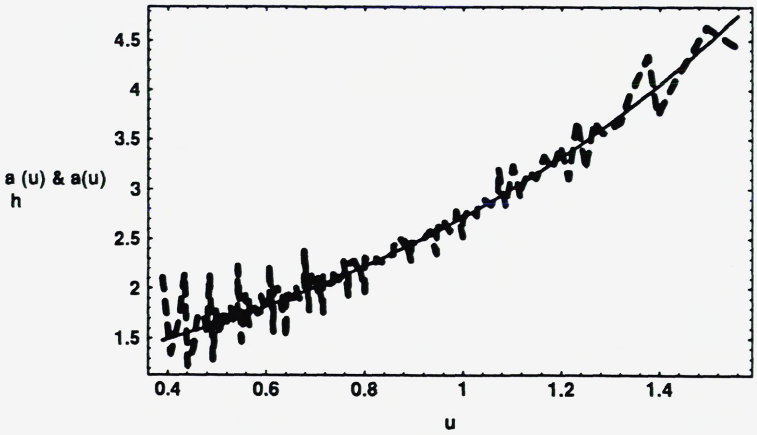

Figure 4: $a_h(u)$ and $a(u)$ with $\mu = 0$ and init. guess 3.0; max.err: 0.63

For the following examples we change the cost functional. It is more realistic to assume, that we have observations only for the value of the unknown solution u in fixed time levels. This means, that we don't have to apply any integration formula for time-axis $[0, T]$. We only need to sum up the output-error for all the observation levels. Also, we approximate ϕ by the difference quotient between the observations of two consecutive time levels, i.e. $\phi(t_{k+\frac{1}{2}}) = \frac{z(t_{k+1})-z(t_k)}{\Delta t}$, here t_k means the observation levels. The integration with respect to $x-$variable is done as in (4.3). We remark that in the previous example, we assume that we have an observation ϕ for the velocity.

Example 2. *Same functions and discrete basis as in example 1, but this time we assume that we only have observation at 9 time levels: $t = 0, 1/9, \cdots, 1$, and the observation is taken as the value of u without observation errors. The cost functional is calculated just as explained.*

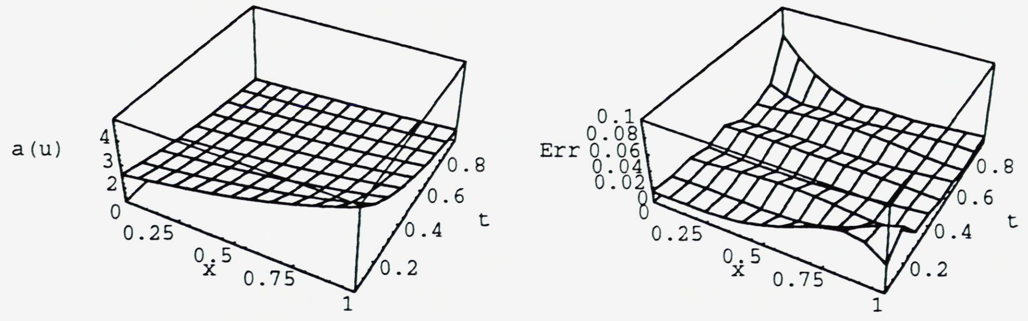

Figure 5: $a(u(t,x))$ and error $a(u(t,x)) - a_{\bar{h}}(u(t,x))$
on nodal points with $\mu = 1$ and init. guess 5.0

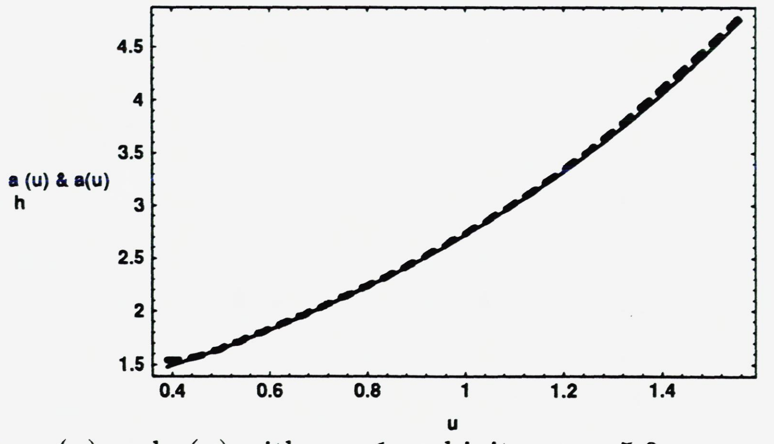

Figure 6: $a_h(u)$ and $a(u)$ with $\mu = 1$ and init. guess 5.0; max.err: 0.06.

Example 3. *The third example is calculated in the same framework as Example 2, but now we have an observation error in the computations. We still assume that we have observation at 10 time levels. We take the observation values from the values of function* z *of the form* $z(t,x) = u(t,x) + c\,\sin(\pi t)\sin(\pi x)$. *For a given c, the observation error is defined by* $\varepsilon = \|z - u\|_{L^2(Q)}$.

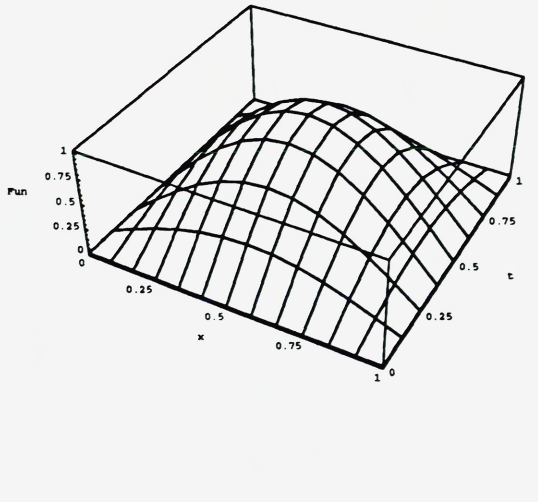

Figure 7: Error function $\sin(\pi t)\sin(\pi x)$.

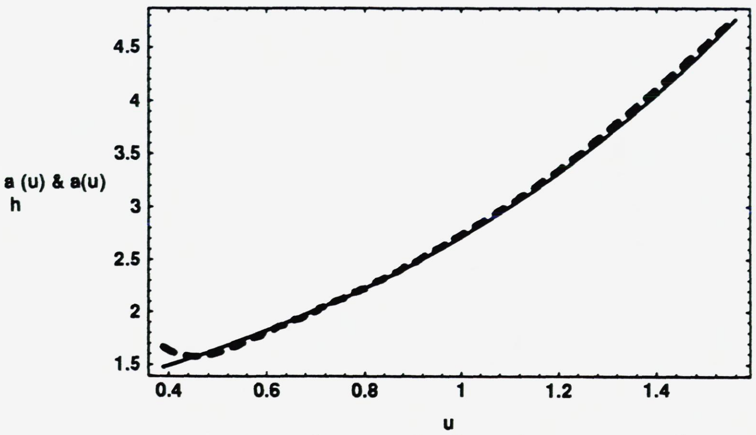

Figure 8: $a_h(u)$ & $a(u)$ with $\Delta t = \frac{1}{9}, h = \frac{1}{12}, \bar{h} = \frac{1}{4}$ and $\varepsilon = \frac{1}{100}$; max. err. 0.18

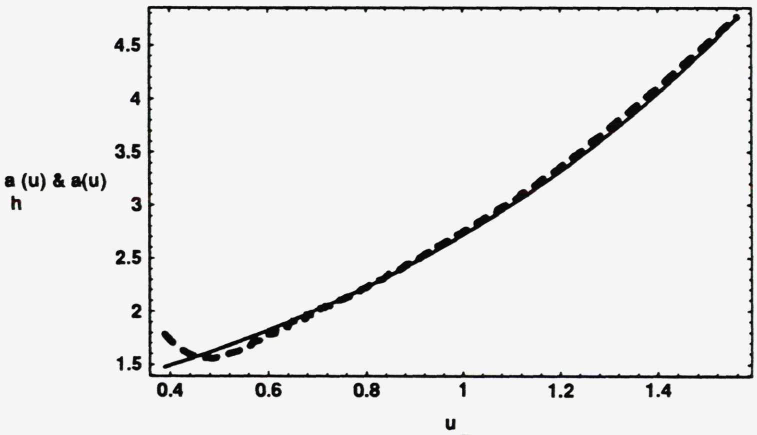

Figure 9: $a_h(u)$ & $a(u)$ with $\Delta t = \frac{1}{9}, h = \frac{1}{12}, \bar{h} = \frac{1}{4}$ and $\varepsilon = \frac{1}{50}$; max. err. 0.31

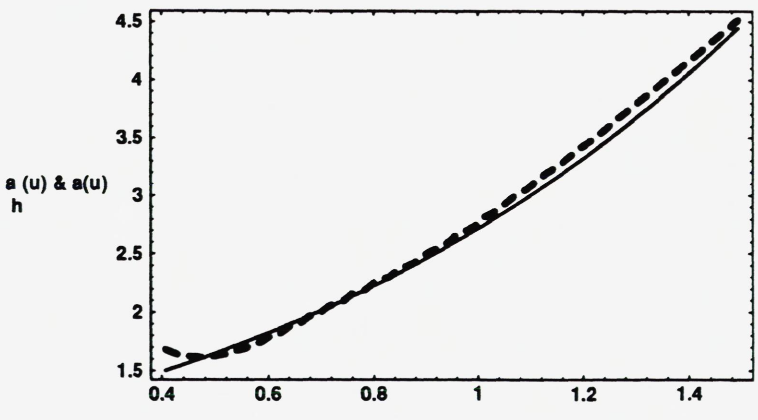

Figure 10: $a_h(u)$ & $a(u)$ with $\Delta t = \frac{1}{5}, h = \frac{1}{9}, \bar{h} = \frac{1}{3}$ and $\varepsilon = \frac{1}{50}$; max. err. 0.18

Example 4. *The last example is calculated in the same way as in Example 3, but here we assume that for the space variable and time variable, we only have observations at points* (x_i, t_k) *with* $t_k = 0, 1/5, 2/5, 3/5, 4/5, 1$ *and* $x_i = 0, 1/6, 1/3, 1/2, 2/3, 5/6, 1$. *We take* $h = \frac{1}{12}$, *and the cost functional is the sum of the output-error at the discretization points* x_j *with odd index* j *for all the observation levels. We also fix* $\bar{h} = \frac{1}{2}$ *and* $\Delta t = \frac{1}{5}$.

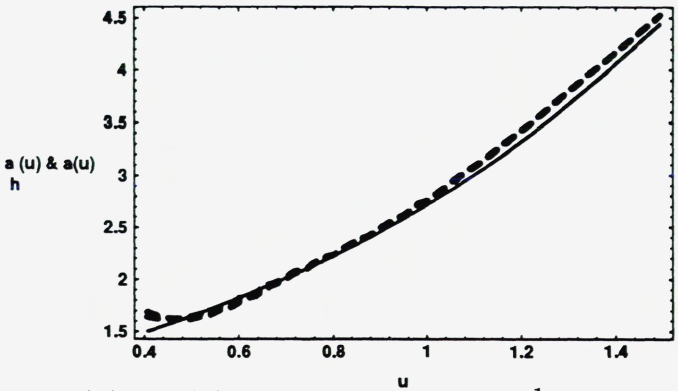

Figure 11: $a_h(u)$ & $a(u)$ with $\mu = 1$ and $\varepsilon = \frac{1}{100}$; max. err. 0.14

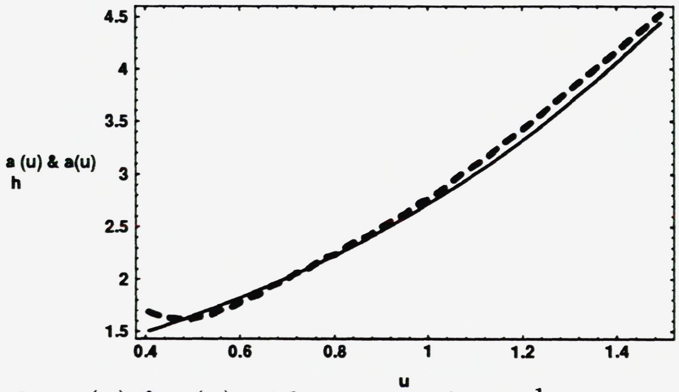

Figure 12: $a_h(u)$ & $a(u)$ with $\mu = 1$ and $\varepsilon = \frac{1}{50}$; max. err. 0.18

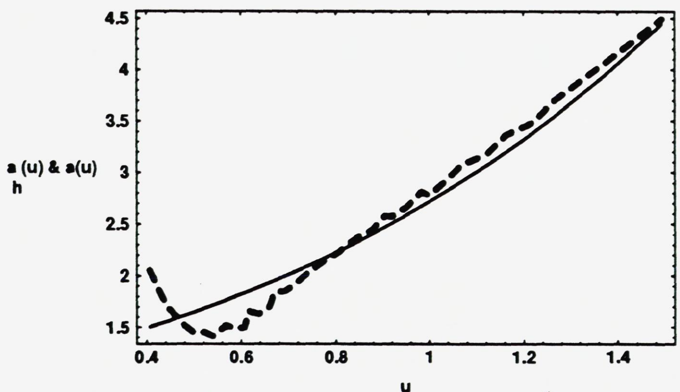

Figure 13: $a_h(u)$ & $a(u)$ with $\mu = 1$ and $\varepsilon = \frac{1}{10}$; max. err. 0.6

References

1. H. T. Banks and K. Kunisch, *Estimation techniques for distributed parameter systems*, Birkhäuser, Basel, 1989.
2. H. T. Banks and A. Murphy., *Estimation of nonlinearities in parabolic models for growth, predation, and dispersal of populations*, J. Math. Anal. Appl. **141** (1989), 580-602.
3. H. T. Banks, P. M. Kareiva and K. A. Murphy, *Parameter estimation techniques for interaction and redistribution models: a predator-prey example*, Oecologia **74** (1987), 356-362.
4. V. Barbu and K. Kunisch, *Identification of nonlinear elliptic equations*, Preprint No. 250, Technical University of Graz (1993).
5. J. R. Cannon and Yanping Lin., *An inverse problem of finding a parameter in a semi-linear heat equation*, J. Math. Anal. Appl. **145** (1990), 470-484.
6. J. R. Cannon, P. Duchateau and K. Steube, *Unknown ingredient inverse problems and trace–type functional differential equations*, Inverse problems in partial differential equations (D. Colton, R. Ewing and W. Rundell, eds.), SIAM, Philadelphia, 1990, pp. 185–200.

7. G. Chavent and P. Lemonnier, *Identification de la non-linearite D'Une equation parabolique quasilineaire*, Appl. Math. Optim. **1** (1974), 121-162.
8. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.
9. J. Douglas Jr., J. T. Dupont and L. B. Wahlbin., *The stability in L^p of the L^2-projection into finite element function spaces*, Numer. Math. **23** (1975), 193-197.
10. R. S. Falk, *Error estimates for the numerical identification of a variable coefficient*, Math. Comp. **40** (1983), 537-546.
11. D. Gilbarg and Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, 1977.
12. E. Laitinen and P. Neittaanmäki, *On numerical simulation of continuous casting process,*, J. Eng. Math. **22** (1988,), 335–345.
13. E. Laitinen and P. Neittaanmäki, *On numerical solution of the problems connected with the control of the secondary cooling in the continuous casting process,*, Control: Theory and Advances Technology, vol. 4, 1988, pp. 285–305.
14. M. Pilant and W. Rundell, *A convolution method for inverse heat conduction problems*, Math. Syst. Theor. **21** (1988), 49-60.
15. M. Pilant and W. Rundell, *Multiple undetermined coefficient problems for quasi-linear parabolic equations*, Numer. Meth. for Part. Diff. Equat. **5** (1989), 297-312.
16. R. Scott, *Interpolated boundary conditions in the finite element method*, SIAM J. Numer. Anal. **12** (1975), 404–427.
17. X.-C. Tai and P. Neittaanmäki, *A linear approach for the nonlinear distributed parameter identification.*, Numerical methods for free boundary problems, International series of numerical mathematics, vol. 99 (Neittaanmaki, ed.), Birkhäuser Verlag Basel, 1991, pp. 401-411.
18. X.-C. Tai and P. Neittaanmäki, *On the numerical solution of the distributed parameter parameter identification problem*, Control and estimation of distributed parameter systems, International series of numerical mathematics, Vol. 100, Birkhäuser Verlag Basel, 1991, pp. 317-329.
19. X.-C. Tai and P. Neittaanmäki, *Error estimates for numerical identification of distributed parameters*, J. Comp. Math., supplenmentary issue (1992), 66–78.