Inverse Problems 4 (1988) 973–994. Printed in the UK

Identification of spatially discontinuous parameters in second-order parabolic systems by piecewise regularisation

Chang-Bock Chung† and Costas Kravaris

Department of Chemical Engineering, The University of Michigan, Ann Arbor, MI 48109-2136, USA

Received 6 May 1987, in final form 1 February 1988

Abstract. A novel piecewise regularisation approach has been developed for the identification of spatially discontinuous parameters in second-order parabolic systems. First, previous results on the regularisation identification approach are reviewed; regularisation is interpreted as enforced compactness of the parameter space. Then, a piecewise regularisation identification approach is rigorously formulated in a functional analytic framework. The theoretical results are applied to the history matching of one-dimensional fractured reservoirs. The performance of the proposed algorithm is evaluated by numerical experiments.

1. Introduction

The problem of estimating spatially varying parameters in partial differential equations (PDE) from noisy data arises in many areas of science and engineering. The present work has been primarily motivated by petroleum engineering applications, in which the spatially varying parameters to be estimated represent unknown reservoir properties such as permeability and porosity. These parameters are inaccessible to direct measurement and, therefore, have to be estimated on the basis of measured pressure and flow rate histories; this estimation process is commonly referred to as 'history matching'. In groundwater hydrology, similar problems arise; they are called 'inverse problems'.

Quite a few history-matching algorithms have been developed for petroleum reservoirs in the past two decades with the implicit or explicit assumption that the unknown parameters are continuous functions of position (see references in Seinfeld and Kravaris (1982) and Yeh (1986)). However, very little attention has been given to naturally fractured reservoirs, which are frequently encountered in practice, primarily because of the complexity involved in the reservoir description. In fractured reservoirs the unknown parameters are discontinuous functions of position. Furthermore, the location of the faults may be unknown.

The flow of oil in a fractured reservoir is described by a linear parabolic equation of the form:

$$\begin{aligned} &\partial u/\partial t = \nabla \cdot (\alpha(x)\nabla u) + f(x,t) && \text{in } \Omega \times]0, T[\\ &u(x,0) = u_0(x) && \text{in } \Omega \\ &\partial u/\partial \nu = 0 && \text{on } \partial\Omega \times]0, T[\end{aligned} \tag{1.1}$$

† Present address: College of Engineering, Chonnam National University, Kwangju, 500 South Korea.

0266-5611/88/040973+22 $02.50

where $\alpha(x)$ is a piecewise continuous function on a spatial domain Ω with boundary $\partial\Omega$. In the above model, the dependent variable u represents pressure, f accounts for the withdrawal or injection of fluid in the reservoir, α is the transmissivity of the porous medium and ν denotes an outward normal vector. The history-matching problem can be stated as follows.

Knowing the source term f and the initial condition and given a measurement of $u(x, t)$ at a discrete set of points $x_1, \ldots, x_\mu$, determine the spatially varying parameter $\alpha(x)$ (including the location of its discontinuities).

Before one develops a method for the identification of (continuous or discontinuous) spatially varying parameters, two key questions need to be examined.

(i) Do the measurements provide sufficient information to determine α uniquely? (identifiability).

(ii) Do small perturbations in the data result in correspondingly small errors in the solution? (stability).

If the answers to these questions are both affirmative, the identification problem is said to be *well-posed*; otherwise, it is said to be *ill-posed*.

When observation data are available for the entire spatial domain, identifiability can usually be easily established under relatively mild assumptions (see Kitamura and Nakagiri (1977) for linear parabolic systems). Establishing identifiability for pointwise observation data is an extremely involved problem. At present the results are available only for special cases (Kravaris and Seinfeld 1986b). On the other hand, it can be easily shown that the identification problem is, as a rule, unstable. In fact, the homogenisation theory (Bensoussan *et al* 1978) states that the solution of a PDE with highly oscillatory coefficients is virtually the *same* as the solution with a very *different* smooth coefficient. The following simple example would suffice to illustrate the instability argument implied by the homogenisation theory.

Consider a one-dimensional PDE of the form (1.1) in $\Omega =]0, 2\pi[$ with a highly oscillatory coefficient $\alpha = \alpha_\varepsilon(x) = (100 + 99 \sin x/\varepsilon)$, $0 < \varepsilon \ll 1$. The solution u_ε converges, as ε tends to 0, to the solution u_h corresponding to the *homogenised* uniform coefficient $\alpha = \alpha_h \simeq 14.1$.

The customary identification approach has been the least-squares method, which consists of minimising the functional

$$J_{LS}(\alpha) = \int_0^T \sum_{j=1}^{\mu} (u(x_j, t; \alpha) - u_j^{obs})^2 \, dt \tag{1.2}$$

subject to the model equation. In the numerical implementation of the least-squares approach, the unknown spatially varying parameter α is *a priori* approximated using a finite-dimensional representation (e.g. splines); then the problem reduces to the one of determining a finite number of constant parameters (e.g. spline coefficients) that minimises J_{LS} (Lamm 1987). When the number of constant parameters is kept small, a well behaved solution results. However, the modelling error introduced by the finite dimensional approximation can be significant, since the spatial variation of an arbitrary α cannot be fully described by a small number of constant parameters. As the number of parameters is increased, numerical instabilities appear, manifested by anomalous oscillations in the estimated α.

A popular approach to alleviate the numerical instabilities is to use *a priori* information about the parameters to be estimated. If the means and covariances of the unknown parameters are *a priori* known, the least-squares functional can be augmented with a term that penalises the weighted deviations of the parameters from their assumed mean values (Gavalas *et al* 1976, Tarantola 1984). While it is not difficult to show that *a priori* statistical information about the unknown parameters would lead to better conditioned estimates, such information is not always available in practice.

The ill-posed nature of the distributed parameter identification problem suggests a regularisation approach for its numerical solution. In the regularisation approach, an ill-posed problem is reformulated into a related well-posed problem, the solution of which is more *regular* (in a sense) than the solution of the original problem and approximates the solution of the original problem. The idea of regularisation, initially proposed by Tikhonov (1963a, b) as a method of solving ill-posed integral equations, has been rigorously extended to the identification of spatially continuous parameters in PDEs of parabolic type (Kravaris and Seinfeld 1985) and successfully implemented in the history matching of oil reservoirs with spatially continuous properties (Lee *et al* 1986).

In the present work we introduce a new piecewise regularisation approach for the identification of spatially discontinuous parameters in one-dimensional second-order parabolic equations. First, in § 2, the general regularisation identification theory is briefly reviewed with an emphasis on the illustration of the basic idea of regularisation. Then, in § 3, identification by piecewise regularisation is rigorously formulated in a functional analytic framework. Finally, in § 4, the proposed approach is numerically implemented for test problems, and its performance is compared with that of the conventional least-squares method.

2. Regularisation approach in distributed parameter identification

Consider a distributed parameter system modelled by a PDE and the problem of identifying the unknown parameter α in the PDE from an observation u^{obs} of the solution u. A conceptual framework for the identification problem can be set up as follows. First, we define three function spaces: the parameter space A, the solution space U and the observation space H, to which belong α, u and u^{obs} respectively. The set of physically admissible parameters is denoted by $A_{\text{ad}} \subset A$. For example, in history matching of reservoir properties, A_{ad} will contain all strictly positive elements $\alpha(x)$ of A. Then, solving the PDE for a given α (with the other pertinent information, such as given boundary and initial conditions) is represented by a solution operator Φ: $A \rightarrow U$ defined by

$$u = \Phi(\alpha). \tag{2.1}$$

The type of measurements available is characterised by an observation operator Θ: $U \rightarrow H$ defined by

$$u^{\text{obs}} = \Theta(u). \tag{2.2}$$

For a distributed observation, Θ is simply an identity operator; for a point observation, Θ is represented as an appropriate injection operator. Combining (2.1) and (2.2), u^{obs} is given by

$$u^{\text{obs}} = (\Theta \circ \Phi)(\alpha) \tag{2.3}$$

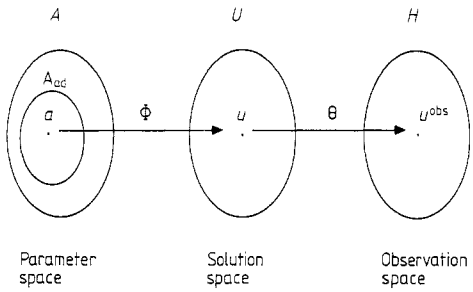

Figure 1. Function spaces for the identification problem.

where $(\Theta \circ \Phi)$ denotes the composite mapping of Φ and Θ. Now the identification problem can be posed as follows.

Knowing the mappings Φ: $A \to U$ and Θ: $U \to H$ and given an observation $u^{obs} \in H$, find $\alpha \in A$ satisfying (2.3).

The situation is depicted in figure 1.

Thus, the identification problem can be viewed as solving in A_{ad} the (nonlinear) operator equation

$$(\Theta \circ \Phi)(\alpha) = u^{obs}. \tag{2.4}$$

Then the ill-posedness of the identification problem, which is roughly defined in the introduction, can be characterised by the properties of the operator $(\Theta \circ \Phi)$. More precisely, the problem of solving (2.4) is said to be well-posed if

(i) (2.4) admits a unique solution, i.e. $(\Theta \circ \Phi)$ is injective (identifiability),

(ii) the solution depends continuously on the data u^{obs}, i.e. $(\Theta \circ \Phi)^{-1}$ is continuous (stability).

Otherwise, the problem is said to be ill-posed.

It was already pointed out that the identification problem is, as a rule, unstable. Given the continuity of the mapping $(\Theta \circ \Phi)$: $A \to H$ with respect to physically meaningful norms, the inherent instability of the identification problem implies that the set of admissible parameters A_{ad} is in general not compact. This is an immediate consequence of the well known topological lemma.

Lemma. Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ be metric spaces, and M a compact subset of X. If a mapping F: $M \to Y$ is continuous and injective, the inverse mapping F^{-1} is also continuous on the set $F(M)$.

In the customary least-squares approach, one attempts to construct a solution by minimising on A_{ad} the least-squares functional

$$J_{LS}(\alpha) = \|\Theta(\Phi(\alpha)) - u^{obs}\|_H^2. \tag{2.5}$$

Being a measure of the deviation of the model ouput from the actual observed output in the norm of the observation space, J_{LS} seems to be a natural choice for the performance index. Due to the non-compactness of A_{ad}, however, $J_{LS}(\alpha)$ will not in general have a minimum on A_{ad}. Furthermore, using the same non-compactness argument, it can be shown that even if a minimum of J_{LS} exists, it may not depend

continuously on u^{obs}. Thus, even when the observation is error-free, a minimising sequence $\{\alpha_n\}$ generated by a minimisation scheme does not necessarily converge to the true solution.

Roughly speaking, the non-compactness of A_{ad} as an underlying source of instability in the least-squares estimation implies that A_{ad} is too *big*, in the sense that one cannot expect to pick up a stable approximate solution instead of anomalous oscillatory ones, which also belong to A_{ad} and give smaller (hence *improved*) J_{LS}. (This is due to the infinite dimensionality of the distributed parameter system; a closed and bounded subset of an infinite dimensional space is not necessarily compact, as it is in a finite dimensional space.) Thus, in order to obtain a least-squares estimate which depends continuously on the observation, one has to constrain the space of unknown parameters in such a way that the mapping $(\Theta \circ \Phi)$ would have a continuous inverse on the image of this constrained set.

The regularisation approach for distributed parameter identification (Kravaris and Seinfeld 1985) is based on *enforced compactness* of the parameter space during the minimisation process. Specifically, this compactness can be attained as follows. First, we introduce a more regular space R which is compactly embedded in A, i.e. R is such that a closed sphere $R_d \subset R$

$$R_d = \{\alpha \in R \mid \|\alpha\|_R^2 \leq d,\ d > 0\}$$

is a compact subset of A in the norm of A. Then we define $R_{\text{ad}} = R \cap A_{\text{ad}}$, and introduce the *stabilising functional*

$$J_S(\alpha) = \|\alpha\|_R^2 \tag{2.6}$$

and the *smoothing functional*

$$\begin{aligned} J_\beta(\alpha) &= J_{\text{LS}}(\alpha) + \beta J_S(\alpha) \\ &= \|\Theta(\Phi(\alpha)) - u^{\text{obs}}\|_H^2 + \beta \|\alpha\|_R^2 \end{aligned} \tag{2.7}$$

where $\beta > 0$ is the *regularisation parameter*. Finally, we obtain a regularised estimate by minimising $J_\beta(\alpha)$ on R_{ad}. It is easy to show that a minimising sequence $\{\alpha_n\}$ of J_β indeed lies in a compact set as follows.

Without loss of generality we may assume that

$$\ldots \leq J_\beta(\alpha_{n+1}) \leq J_\beta(\alpha_n) \leq \ldots \leq J_\beta(\alpha_1).$$

Hence, for every α_n, $n \geq 1$,

$$\|\alpha_n\|_R^2 \leq \frac{1}{\beta} J_\beta(\alpha_n) \leq \frac{1}{\beta} J_\beta(\alpha_1)$$

i.e. $\{\alpha_n\}$ lies in R_d, with $d = 1/\beta J_\beta(\alpha_1)$.

As a result of the compact embedding of R into A, existence of a minimum of $J_\beta(\alpha)$ easily follows. This is a consequence of the well known fact that a continuous function has a minimum on a compact set. Further, as a key result of the regularisation approach, the continuous dependence of the minimum on the observation u^{obs} with an adequate choice of β can be shown using a similar compactness argument (Kravaris and Seinfeld 1985).

The regularisation parameter β can be selected as a function of an upper bound δ on the observation error (i.e. $\|u^{\text{obs}} - u_{\text{T}}^{\text{obs}}\|_H \leq \delta$). The following methods have been discussed in Kravaris and Seinfeld (1985).

Method 1. When an *a priori* upper bound on $\|\alpha\|_R$ is known, i.e. $\|\alpha_T\|_R \leq \Delta$, one can choose $\beta(\delta) = (\delta/\Delta)^2$.

Method 2. Choose $\beta(\delta)$ so that

$$\|\Theta(\Phi(\alpha_{\beta(\delta)})) - u^{\text{obs}}\|_H = \delta$$

where $\alpha_{\beta(\delta)}$ minimises

$$J_\beta(\alpha) = \|\Theta(\Phi(\alpha)) - u^{\text{obs}}\|_H^2 + \beta(\delta)\|\alpha\|_R^2.$$

In order to apply the above regularisation approach to concrete distributed parameter identification problems, all that is needed is an appropriate PDE framework that will permit one to select physically meaningful function spaces A, U, H and a set A_{ad} so that the mapping $(\Theta \circ \Phi)$: $A_{\text{ad}} \rightarrow H$ is well defined and continuous. Then one can choose R so that the embedding of R into A provides the necessary compactness of the constrained parameter set.

3. Identification of spatially discontinuous parameters in second-order parabolic systems by piecewise regularisation

Consider a one-dimensional parabolic equation

$$\begin{aligned} &\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) = f(x,t) \qquad \text{in }]0, l[\times]0, T[\\ &u(x,0) = u_0(x) \qquad \text{in }]0, l[\\ &\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(l,t) = 0 \qquad \text{in }]0, T[\end{aligned} \tag{3.1}$$

where $\alpha(x)$ is a piecewise continuous function having n discontinuities at $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ with $0 < \xi_1 < \ldots < \xi_n < l$ and $\xi_0 = 0$, $\xi_{n+1} = l$.

We consider the following identification problem.

Knowing f and u_0 and given observations $u_j^{\text{obs}}(t)$, $j = 1, \ldots, \mu$ of u at the points $x_1, \ldots, x_\mu$ and given the points $\xi_1, \ldots, \xi_n$ of possible parameter discontinuity, estimate $\alpha(x)$.

To rigorously formulate this identification problem, we first define the parameter space A as follows.

$A = C_\xi([0, l]) \equiv$ the space of functions which are uniformly continuous on each $]\xi_{i-1}, \xi_i[$, $i = 1, 2, \ldots, n+1$. In other words, $C_\xi([0, l])$ will contain those functions which are continuous on $[0, l]$ except possibly for jump discontinuities at ξ_i, $i = 1, 2, \ldots, n$. When endowed with the norm

$$\|u\|_{C_\xi([0,l])} = \max\{\|u\|_{C^0([\xi_{i-1}, \xi_i])}\}$$

$C_\xi([0, l])$ is a Banach space.

Since $\alpha(x)$ must be strictly positive to make physical sense, we define

$$A_{\text{ad}} = \{\alpha \in A \mid \alpha(x) \geq \alpha_0 > 0\}. \tag{3.2}$$

The second step is to define a solution space and establish continuous dependence of the solution u of the boundary value problem (3.1) on α. Consider the variational formulation of (3.1):

$$\begin{aligned} &\int_0^l \frac{\partial u}{\partial t} v + \int_0^l \alpha(x) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} = \int_0^l fv \qquad \forall v \in H^1(]0, l[) \\ &u(x, 0) = u_0(x). \end{aligned} \tag{3.3}$$

A standard result from PDE theory (see, for example, Lions 1971) states that if

$$\alpha(x) \in L^\infty(]0, l[) \qquad \text{and} \qquad \alpha(x) \geqslant \alpha_0 > 0 \qquad \text{in }]0, l[$$

$$u_0(x) \in L^2(]0, l[)$$

$$f(x, t) \in L^2(]0, l[\times]0, T[) \text{ (or more generally in } L^2(0, T; (H^1(]0, l[))')),$$

equation (3.3) admits a unique solution u in

$$W(0, T) = \left\{ u \middle| u \in L^2(0, T; H^1(]0, l[)), \frac{du}{dt} \in L^2(0, T; (H^1(]0, l[))') \right\}, \tag{3.4}$$

which depends continuously on f and u_0. Now applying the implicit function theorem to the variational formulation (3.3), it is not difficult to show that u depends continuously on α as well. More precisely, defining a mapping

$$\Psi: (\alpha, u) \in L^\infty(]0, l[) \times W(0, T)$$

$$\to \left(\int_0^l \frac{\partial u}{\partial t} v + \int_0^l \alpha(x) \frac{\partial u}{\partial x} \frac{\partial v}{\partial x}, u(0) \right) \in L^2(0, T; (H^1(]0, l[))') \times L^2(]0, l[)$$

(3.3) can be represented by the equation

$$\Psi(\alpha, u) = \left(\int_0^l fv, u_0 \right). \tag{3.5}$$

Since Ψ is of C^1-class and $(\partial \Psi / \partial u)(\alpha, u)$ is a linear homeomorphism of $W(0, T)$ onto $L^2(0, T; (H^1(]0, l[))') \times L^2(]0, l[)$, the implicit function $u = \Phi(\alpha)$ is also of C^1-class from $L^\infty(]0, l[)$ into W(0, T). Since $C_\xi([0, l]) \subset L^\infty(]0, l[)$, the result holds for $\alpha \in C_\xi([0, l])$.

Since the solution u is in $L^2(0, T; H^1(]0, l[))$, $u(x_j, t)$ will have meaning (and '$t \to u(x_j, t)$' $\in L^2(0, T)$) as a result of the embedding $H^1(]0, l[) \to C^0([0, l])$. Accordingly the point observation $u^{\text{obs}} = \{u(x_j, t), j = 1, \ldots, \mu\}$ is well defined for every $u \in W(0, T)$ and has values in $(L^2(0, T))^\mu$; so $H = (L^2(0, T))^\mu$ will be the observation space. Furthermore, it is straightforward to show that the point evaluation operation is continuous. Consequently, we have the continuous dependence of $u(x_j, t), j = 1, \ldots, \mu$ on α. In the following we will consider u to be a function of α and we will use the notation $u(x_j, t; \alpha)$ to indicate the dependence of u on α.

Now for a regular space R which is compactly embedded in $A = C_\xi([0, l])$, we introduce a candidate $H_\xi^m(]0, l[)$ defined below. The assertion on the compact embedding $H_\xi^m(]0, l[) \to C_\xi([0, l])$ is established by theorem 1. We define:

$R = H_\xi^m(]0, l[) \equiv$ the space of functions whose weak derivatives up to order m are square integrable over each $]\xi_{i-1}, \xi_i[$, $i = 1, 2, \ldots, n+1$. When endowed with the inner product

$$(u, v)_{H_\xi^m(]0,l[)} = \sum_{i=1}^{n+1} \int_{\xi_{i-1}}^{\xi_i} \sum_{k=0}^{m} \frac{\partial^k u}{\partial x^k} \frac{\partial^k v}{\partial x^k}$$

$H_\xi^m(]0, l[)$ is a Hilbert space.

Theorem 1. $H_\xi^m(]0, l[)$ is compactly embedded in $C_\xi([0, l])$ for $m = 1, 2, \ldots$.

Proof. From the Sobolev embedding theorem (Adams 1975), it readily follows that for $i = 1, \ldots, n+1$, $H^m(]\xi_{i-1}, \xi_i[)$ is compactly embedded in $C^0([\xi_{i-1}, \xi_i])$ iff $m > \frac{1}{2}$. This means that the set

$$X_i = \{u \in C^0([\xi_{i-1}, \xi_i]) \mid \|u\|^2_{H^m(]\xi_{i-1}, \xi_i[)} \leq M_i < \infty\}$$

is a compact subset of $C^0([\xi_{i-1}, \xi_i])$. Now consider the following product spaces:

$C \equiv \prod_{i=1}^{n+1} C^0([\xi_{i-1}, \xi_i])$, a Banach space endowed with the norm

$$\|u\|_C = \max_i \{\|u\|_{C^0(]\xi_{i-1}, \xi_i[)}\}$$

and

$H = \prod_{i=1}^{n+1} H^m(]\xi_{i-1}, \xi_i[)$, a Hilbert space endowed with the inner product

$$(u, v)_H = \sum_{i=1}^{n+1} (u, v)_{H^m(]\xi_{i-1}, \xi_i[)}.$$

From the embedding $H^m(]\xi_{i-1}, \xi_i[) \to C^0([\xi_{i-1}, \xi_i])$ it easily follows that the cartesian product H is embedded in C. Now using the well known Tikhonov theorem of topology (Simmons 1963), it is straightforward to verify that the cartesian product X

$$X = \prod_{i=1}^{n+1} X_i$$

$$= \{(u_1, \ldots, u_{n+1}) \in C \mid \|u_i\|_{H^m(]\xi_{i-1}, \xi_i[)} \leq M_i < \infty, i = 1, \ldots, n+1\}$$

is compact in the norm topology of C. Hence a closed subset $\hat{X}$ of X

$$\hat{X} = \{(u_1, \ldots, u_{n+1}) \in C \mid \|u\|_H^2 = \sum_{i=1}^{n+1} \|u_i\|^2_{H^m(]\xi_{i-1}, \xi_i[)} \leq M\}$$

with $M = \min\{M_1, \ldots, M_{n+1}\}$ is also a compact subset of C, which indicates that H is compactly embedded in C. But C and $C_\xi([0, l])$ are clearly isometrically isomorphic, and H and $H_\xi^m(]0, l[)$ are isomorphic Hilbert spaces. Therefore $H_\xi^m(]0, l[)$ is compactly embedded in $C_\xi([0, l])$.

So far we have laid down the groundwork for the identification of spatially discontinuous parameters in (3.1) by regularisation: a PDE framework with appropriate function spaces and the compact embedding of $H_\xi^m(]0, l[)$ into $C_\xi([0, l])$. A regularised solution of this problem can be obtained by minimising the smoothing functional

$$J_\beta(\alpha) = \sum_{j=1}^{\mu} \int_0^T (u(x_j, t; \alpha) - u_j^{\text{obs}}(t))^2 \, dt + \beta \|\alpha\|^2_{H_\xi^m(]0,l[)} \tag{3.6}$$

on the set

$$\begin{aligned} R_{\text{ad}} &= A_{\text{ad}} \cap R \\ &= \{\alpha \in H_\xi^m(]0, l[) | \alpha(x) \geq \alpha_0 > 0, x \in]0, l[\}. \end{aligned} \tag{3.7}$$

The performance index is clearly well defined on R_{ad}.

The next theorem establishes the existence of a global minimum of $J_\beta(\alpha)$ on R_{ad}. As pointed out in the previous section, the key elements of the proof are (i) the continuous dependence of $u(x_j, t)$ on α and (ii) the compact embedding of $H_\xi^m(]0, l[)$ into $C_\xi([0, l])$. The argument of the proof is very similar to the one of theorem 4.2 of Kravaris and Seinfeld (1985) and will be omitted for brevity. See Chung (1988) for details.

Theorem 2. The functional $J_\beta(\alpha)$ admits a global minimum α_β on R_{ad}.

The next theorem establishes that minima of J_β depend continuously on the observation; what it says is roughly the following.

Let α_T be the *true* value of the parameter and $u_T^{\text{obs}} = \{u(x_j, t; \alpha_T), j = 1, \ldots, m\}$, an *error-free* observation. Provided that

(i) α_T is the unique pre-image of u_T^{obs};
(ii) β is an appropriately chosen function of the observation error;

any minimum of J_β converges (in the norm of $C_\xi([0, l])$) to α_T as the observation error tends (in the norm of $(L^2(0, T))^\mu$) to zero.

Theorem 3. Assume

$u_T^{\text{obs}} \in (L^2(0, T))^\mu$

$\exists$ a unique $\alpha_T(x) \in H_\xi^m(]0, l[)$ such that $u(x_j, t; \alpha_T) = u_{j,T}^{\text{obs}}$

$B_1(\delta)$ and $B_2(\delta)$ are given non-negative, nondecreasing and continuous functions

satisfying $B_2(0) = 0$, $\delta^2 / B_1(\delta) \leq B_2(\delta)$.

Then $\forall \varepsilon > 0$ $\exists \delta_0(\varepsilon, B_1, B_2)$ such that $\forall u^{\text{obs}} \in (L^2(0, T))^\mu$ $\forall \delta \leq \delta_0$,

$$\|u^{\text{obs}} - u_T^{\text{obs}}\|_{(L^2(0,T))^\mu} \leq \delta \Rightarrow \|\alpha_\beta - \alpha_T\|_{C_\xi([0,l])} \leq \varepsilon$$

for all β satisfying $\delta^2 / B_1(\delta) \leq \beta \leq B_2(\delta)$.

The argument of the proof is very similar to that of theorem 4.3 of Kravaris and Seinfeld (1985) and will be omitted for brevity. See Chung (1988) for details.

Remark 1. The approach described above extends naturally to more general second-order parabolic systems of the form

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right) + b(x)\frac{\partial u}{\partial x} + c(x)u = f(x, t) \tag{3.8}$$

where any of $a(x)$, $b(x)$, $c(x)$ can be unknown and need to be identified. The appropriate function spaces will be given as follows:

$$A = C_\xi([0, l]) \times C_\xi([0, l]) \times C_\xi([0, l]) \tag{3.9}$$

with norm

$$\|\alpha\|_A = \max\{\|a\|_{C_\xi([0,l])}, \|b\|_{C_\xi([0,l])}, \|c\|_{C_\xi([0,l])}\}$$

$$U = W(0, T) \tag{3.10}$$

$$H = (L^2(0, T))^\mu \tag{3.11}$$

$$R = H_\xi^m(]0, l[) \times H_\xi^m(]0, l[) \times H_\xi^m(]0, l[) \qquad m \geqslant 1 \tag{3.12}$$

with inner product

$$(\alpha_1, \alpha_2)_R = (a_1, a_2)_{H_\xi^m(]0,l[)} + (b_1, b_2)_{H_\xi^m(]0,l[)} + (c_1, c_2)_{H_\xi^m}(]0, l[). \tag{3.13}$$

Now using the same Tikhonov argument it is straightforward to establish the compact embedding $A \rightarrow R$. Then one can obtain a regularised solution by minimising

$$J_\beta(a, b, c) = \sum_{j=1}^{\mu} \int_0^T (u(x_j, t; a, b, c) - u_j^{\text{obs}}(t))^2 \, dt$$

$$+ \beta(\|a\|^2_{H_\xi^m(]0,l[)} + \|b\|^2_{H_\xi^m(]0,l[)} + \|c\|^2_{H_\xi^m(]0,l[)}). \tag{3.14}$$

The arguments of theorems 2 and 3 carry over in this case.

Remark 2. When the location of the discontinuities $\xi = (\xi_1, \ldots, \xi_n)$ is unknown, ξ has to be estimated along with $\alpha(x)$. For this problem, a theoretical formulation is not available. However, a formal extension of the approach that consists of minimising

$$J_\beta(\alpha, \xi) = \sum_{j=1}^{\mu} \int_0^T (u(x_j, t; \alpha, \xi) - u_j^{\text{obs}}(t))^2 \, dt + \beta \sum_{i=1}^{n+1} \|\alpha\|^2_{H^m(]\xi_{i-1}, \xi_i[)} \tag{3.15}$$

is straightforward. It has been implemented in the numerical part of our work (see the next section).

4. Numerical implementation

4.1. Numerical experiments

The identification approach developed in the previous section was applied to the estimation of the spatially discontinuous parameter $\alpha(x)$ in the following PDE:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) \qquad x \in]0, 1[, t \in]0, 0.025[$$

$$u(x, 0) = 10 + 20x \tag{4.1}$$

$$\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0.$$

The performance of the proposed algorithm was evaluated through a series of numerical experiments. Specifically, on the basis of simulated data generated with a specified true parameter $\alpha_T(x)$, starting with an initial guess, we tried to recover the true parameter. In our simulation study, two different true parameters were considered:

$$\alpha_{T1}(x) = \begin{cases} 3[\exp(-25(x-0.25)^2) + \exp(-25(x-0.75)^2)] & x \in [0, 0.6] \\ 3[\exp(-25(x-0.25)^2) + \exp(-25(x-0.75)^2)] + 1.5 & x \in [0.6, 1] \end{cases} \tag{4.2}$$

$$\alpha_{T2}(x) = \begin{cases} 4 + 3\sin \pi(x-0.5) & x \in [0, 0.5] \\ 1 + \cos \pi(x-0.5) & x \in [0.5, 1]. \end{cases} \tag{4.3}$$

In order to numerically minimise the smoothing functional $J_\beta(\alpha)$, two kinds of discretisation are needed: parameter discretisation for a finite dimensional representation of $\alpha(x)$ and state discretisation for a numerical solution of the state equation (4.1). Figure 2 shows the two corresponding grid systems which are employed in our numerical experiments as explained below. Thus we will get an approximate minimum by minimising a discretised (both in space and time) smoothing functional over an approximate finite dimensional approximating subspace of $H_\xi^m(]0, l[)$ (Kravaris and Seinfeld 1986a). For $m \leqslant 3$, such an approximating subspace can be conveniently generated by piecewise cubic splines as follows (Lamm 1987).

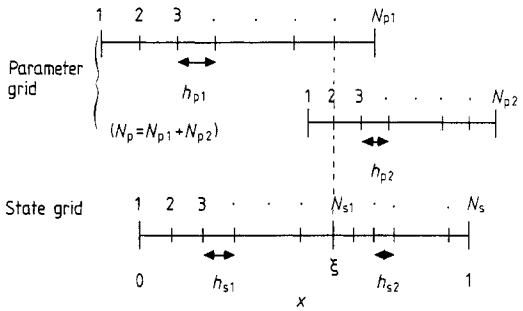

Figure 2. Two grid systems for the numerical minimisation of the smoothing functional.

In each subregion $[0, \xi]$ and $[\xi, 1]$, an equidistant grid is defined as shown in figure 2. When N_{p1} and N_{p2} are sufficiently large integers, an arbitrary element of $H_\xi^m(]0, l[)$ can then be approximated by

$$\alpha(x) = \begin{cases} \sum_{k=1}^{N_{p1}} w_{1k} b_{k1}(x) & x \in [0, \xi] \\ \sum_{k=1}^{N_{p2}} w_{2k} b_{k2}(x) & x \in [\xi, 1] \end{cases} \tag{4.4}$$

where $B_{ki}(x)$, $i = 1, 2$ is the cubic B-spline function on each subregion.

Remark 3. It is well known that the cubic spline is a convergent approximation in H^m (Aubin 1972). Therefore, since $H_\xi^m(]0, l[)$ is defined as a product of H^m on each subregion, the piecewise cubic spline of (4.4) will be a convergent approximation in $H_\xi^m(]0, l[)$. Also, it is straightforward to extend the above piecewise cubic-spline representation to $\alpha(x)$ with n discontinuities.

The state equation was solved numerically using the implicit finite difference scheme with the state grid shown in figure 2. An accurate solution was obtained using time step $\Delta t = 10^{-3}$ and $N_s = 33$ ($N_{s1} = 17$, $N_{s2} = 16$), and this level of state discretisation was used when the level of parameter discretisation was low ($N_p = 8$ and 14). For higher N_p, N_s was increased accordingly.

The data u_{jm}^{obs} used in our runs were generated as follows. First, the state equation (4.1) was solved numerically for $\alpha = \alpha_{T1}$ or α_{T2} with $N_s = 129$ and $\Delta t = 10^{-5}$. Then the set of values $\{u(x_j, t_m; \alpha_T);\ x_j = j/16,\ j = 1, \ldots, 15,\ t_m = 10^{-3}m,\ m = 1, \ldots, 25\}$ were perturbed by normally distributed random numbers with zero mean and standard deviation $0.1(\simeq 0.5\%$ error). Consequently, the approximate error level δ in the observation data was

$$\delta^2 = \sum_{j=1}^{15} \sum_{m=1}^{25} (u(x_j, t_m; \alpha_T, \xi) - u_{jm}^{obs})^2 \Delta t \simeq 3.75 \times 10^{-3}. \tag{4.5}$$

Thus in our simulation, we used the following discretised smoothing functional

$$J_\beta(\theta) = \sum_{j=1}^{15} \sum_{m=1}^{25} (u(x_j, t_m; \alpha, \xi) - u_{jm}^{obs})^2 \Delta t + \beta \left[\int_0^\xi \alpha^2 + \left(\frac{d\alpha}{dx}\right)^2 + \left(\frac{d^2\alpha}{dx^2}\right)^2 dx + \int_\xi^1 \alpha^2 + \left(\frac{d\alpha}{dx}\right)^2 + \left(\frac{d^2\alpha}{dx^2}\right)^2 dx \right] \tag{4.6}$$

where the unknown vector θ is given by

$$\theta = \begin{cases} w = (w_{11}, \ldots, w_{1N_{p1}}, w_{21}, \ldots, w_{2N_{p2}})^T & \text{when } \xi \text{ known} \\ (w|\xi)^T = (w_{11}, \ldots, w_{1Np1}, w_{21}, \ldots, w_{2N_{p2}}, \xi)^T & \text{when } \xi \text{ unknown} \end{cases} \tag{4.7}$$

and α is represented by (4.4). It is noteworthy that when ξ is assumed to be unknown and has to be estimated along with the spline coefficients ω, the mesh sizes h_p and h_s in each subdomain, $[0, \xi]$ and $[\xi, 1]$, may vary depending on the current estimate of ξ during the minimisation process. Such ξ-dependent grid systems were used previously in a similar fashion (Lamm 1987).

The function $J_\beta(\theta)$ was minimised iteratively as follows. As an initial guess of the spline coefficient vector ω, uniform values of 2 (i.e. $\alpha^{(0)} = 2.0$) were used. Then, at each major iteration, the descent direction was updated using the BFGS (Broyden–Fletcher–Goldfarb–Shanno) quasi-Newton minimisation method (Shanno and Phua 1978); the step length along this direction was found using the golden section line search method. The convergence criterion for stopping the major iterations was

$$0 \leqslant J_\beta^{(k+1)} - J_\beta^{(k)} \leqslant 10^{-7}. \tag{4.8}$$

4.2. *Results and discussion*

Our simulation runs are divided into two parts: (i) when ξ is known and (ii) when ξ is unknown and has to be estimated. In the first part, we have studied the effect of the level of discretisation N_p on the least-squares estimates and the effect of the regularisation parameter β on the regularised estimates. The performances of the two approaches are compared. We also studied the effect of the magnitude of the spatial gradient $\partial u/\partial x$ on the least-squares estimates. In the second part, however, we encountered severe convergence problems with both methods. These difficulties are attributed to the severely ill-posed nature of the problem of estimating discontinuities.

Estimation with ξ known. The ability of the least-squares method to recover each of α_{T1} and α_{T2} is investigated at various levels of discretisation. Table 1 shows the values of the performance indices evaluated at the final estimates. The corresponding estimates are shown in figures 3 and 4.

In estimating α_{T1}, J_{LS} can be reduced with higher N_p only at the expense of higher J_S. Figure 3 shows clearly that at low $N_p = 8$ the modelling error introduced by the finite dimensional approximation is significant while at high $N_p = 134$ ill conditioning in the form of anomalous oscillations is inevitable. In estimating α_{T2}, which is smoother than α_{T1}, a similar effect is observed except that the modelling error at low N_p is not so significant as in estimating α_{T2}. The difference can also be noticed by comparing the corresponding J_{LS} values.

As a result of a trade-off between modelling errors at low N_p and instabilities at high N_p, there seems to be an optimum level of discretisation. For our examples, the

Table 1. Least-squares estimation with ξ known: effect of the level of discretisation. (*a*) For α_{T1}, (*b*) for α_{T2}.

N_p	J_{LS}	J_S	Number of iterations
(*a*)			
8	8.801×10^{-3}	4.524×10^{2}	16
14	3.746×10^{-3}	1.018×10^{4}	20
134	3.588×10^{-3}	1.137×10^{8}	42
(*b*)			
8	4.272×10^{-3}	1.715×10^{2}	12
14	3.866×10^{-3}	3.202×10^{4}	24
134	3.582×10^{-3}	7.261×10^{6}	21

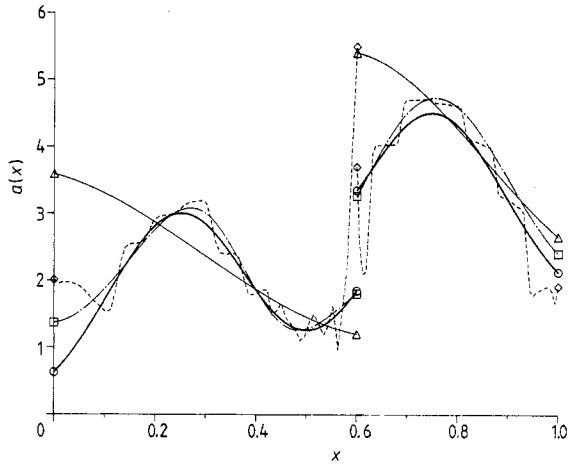

Figure 3. Least-squares estimates of α_{T1} with ξ known: N_p effect. ○, true; △, $N_p = 8$; □, $N_p = 14$; ◇, $N_p = 134$.

best performance was obtained with $N_p = 14$ for both α_{T1} and α_{T2}. In practice, however, the true parameter is completely unknown and. therefore, it will not be easy to determine the optimum level of discretisation on the basis of the least-squares estimates for various levels of discretisation. Hence there arises a need to use sufficiently high N_p in order to describe the spatial distribution of the unknown parameter reasonably well and, at the same time, a need to provide a countermeasure against numerical instabilities which will show up in the least-squares estimation.

In our regularisation identification approach, it is the presence of the penalty term βJ_S in the performance index that prevents the numerical instabilities of the least-squares estimates obtained at high level of discretisation. Accordingly, larger values of β would lead to estimates with smaller J_S, i.e. to smoother estimates. This effect of increasing regularisation parameter can be easily seen in table 2, and is well illustrated in figures 5 and 6. Also note in table 2 that decreasing values of J_S are

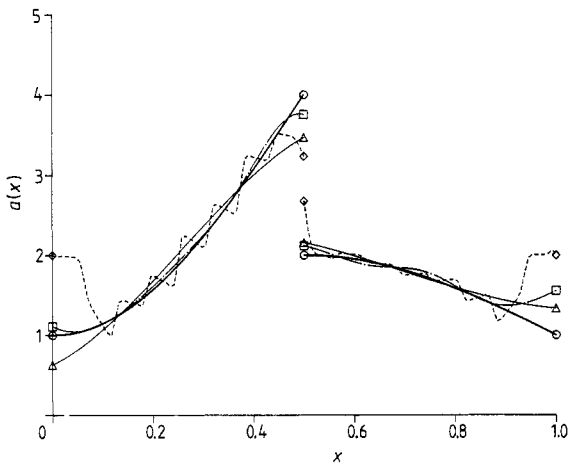

Figure 4. Least-squares estimates of α_{T2} with ξ known: N_p effect. ○, true; △, $N_p = 8$; □, $N_p = 14$; ◇, $N_p = 134$.

Table 2. Regularised estimates with ξ known at high level of discretisation ($N_p = 134$): effect of the regularisation parameter. (*a*) For α_{T1}, (*b*) for α_{T2}.

β	J_{β}	J_{LS}	J_S	Number of iterations
(*a*)				
0	3.588×10^{-3}	3.588×10^{-3}	1.137×10^{8}	42
10^{-10}	1.181×10^{-2}	6.992×10^{-3}	4.819×10^{7}	92
10^{-8}	3.723×10^{-3}	3.597×10^{-3}	1.256×10^{4}	202
10^{-6}	6.586×10^{-3}	5.104×10^{-3}	1.483×10^{3}	154
(*b*)				
0	3.582×10^{-3}	3.582×10^{-3}	7.261×10^{6}	21
10^{-8}	3.611×10^{-3}	3.553×10^{-3}	5.825×10^{3}	146
10^{-6}	3.699×10^{-3}	3.621×10^{-3}	7.826×10	151
10^{-3}	2.142×10^{-2}	9.520×10^{-3}	1.172×10	154

accompanied by increasing values of J_{LS} with a few exceptions ($\beta = 10^{-10}$ for α_{T1} and $\beta = 10^{-8}$ for α_{T2}), which are probably due to numerical conditioning problems in the minimisation process.

The regularisation parameters were chosen according to the discrepancy rule (see method 2 in § 2). Specifically, we made runs for different orders of magnitude of β and selected β so that the J_{LS} value corresponding to the estimates is close to the square of the observation error δ^2 (see equation (4.5)). Thus, for the cases considered here, we chose $\beta = 10^{-8}$ for α_{T1} and $\beta = 10^{-6}$ for α_{T2}. In the runs we made, this rule seems to pick the optimal β within a range in which the estimated surface does not change significantly.

Figures 7 and 8 compare the regularised estimates, for α_{T1} and α_{T2} respectively, corresponding to the optimal values of β with the least-squares estimates. These figures clearly demonstrate that regularisation is an effective means of obtaining a well behaved approximate estimate.

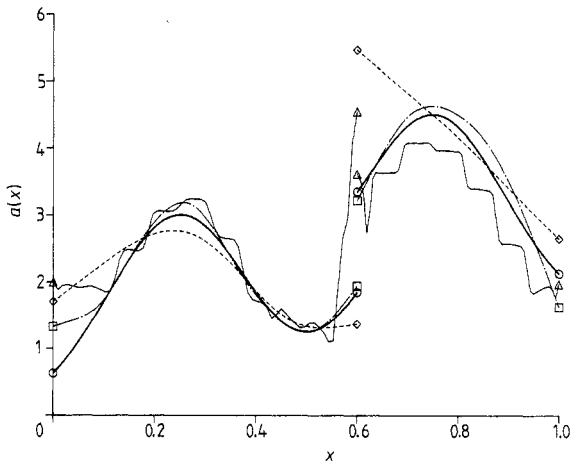

Figure 5. Regularised estimates of α_{T1} at $N_p = 134$ with ξ known: β effect. ○, true; △, $\beta = 10^{-10}$; □, $\beta = 10^{-8}$; ◇, $\beta = 10^{-6}$.

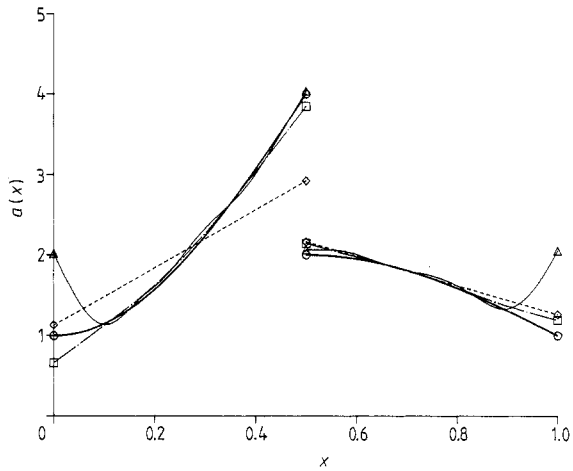

Figure 6. Regularised estimates of α_{T2} at $N_p = 134$ with ξ known: β effect. ○, true; △, $\beta = 10^{-8}$; □, $\beta = 10^{-6}$; ◇, $\beta = 10^{-3}$.

Remark 4. The problem of selecting the best regularisation parameter β is an open research problem. The discrepancy rule that was applied here is not necessarily the best method of selection of β. In fact, numerical problems arising in the actual implementation of the identification approach—such as numerical errors in solving PDEs, extremely slow convergence rate in the minimisation process, existence of local minima, etc—often make it impractical to minimise J_{LS} sufficiently close to the *a priori* known error level δ^2. However, alternative methods (e.g. order-of-magnitude rule, quasi-optimal value, etc) are available in the literature (Tikhonov 1963b). In practice, one may make preliminary test simulation runs to choose a β-selection scheme which is most appropriate for the identification problem at hand.

Next we investigated the effect of the magnitude of the spatial gradient of the solution, $\partial u/\partial x$, on the performance of the estimation algorithm. Since the diffusion

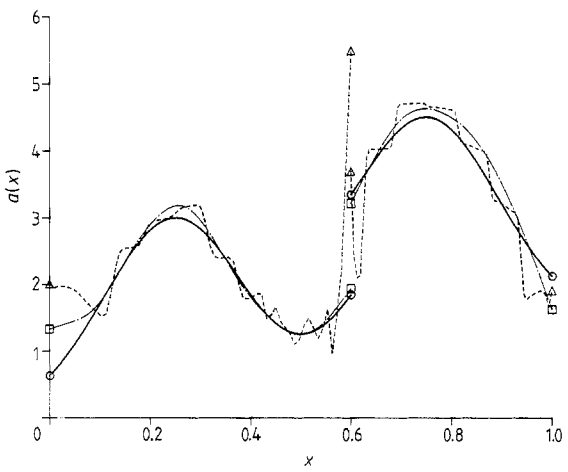

Figure 7. Least-squares and optimal regularised estimates of α_{T1} at $N_p = 134$ with ξ known. ○, true; △, $\beta = 0$; □, $\beta = 10^{-8}$.

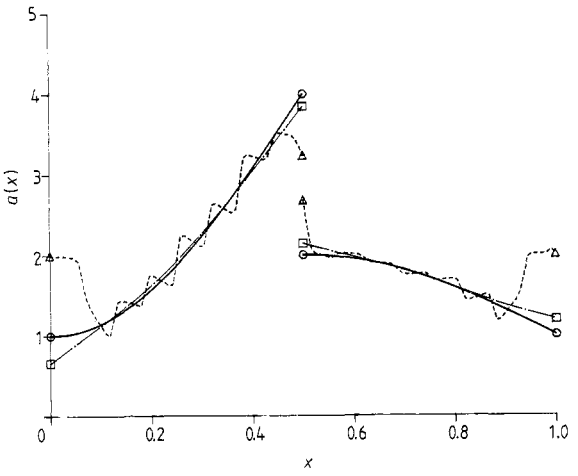

Figure 8. Least-squares and optimal regularised estimates of α_{T2} at $N_p = 134$ with ξ known. ○, true; △, $\beta = 0$; □, $\beta =$ [illegible].

coefficient $\alpha(x)$ in a parabolic PDE is a measure of the tendency of the system to dissipate its gradient, it follows that a system with a larger gradient $\partial u/\partial x$, as a rule, has a higher parameter sensitivity $\partial u/\partial \alpha$; this is a favourable situation for estimating α in the inverse problem. On the contrary, estimation of the systems with smaller or vanishing gradients is more sensitive to the noise in the data. To investigate the issue, we performed estimation runs for a system initially at rest ($u_0(x) = 20$). In order to collect meaningful observation data, we excited the system using a withdrawal term f of the following form

$$f(x, t) = -\sum_{j=1}^{N_w} \frac{q_j}{d_j} X_{\Omega_j} \tag{4.9}$$

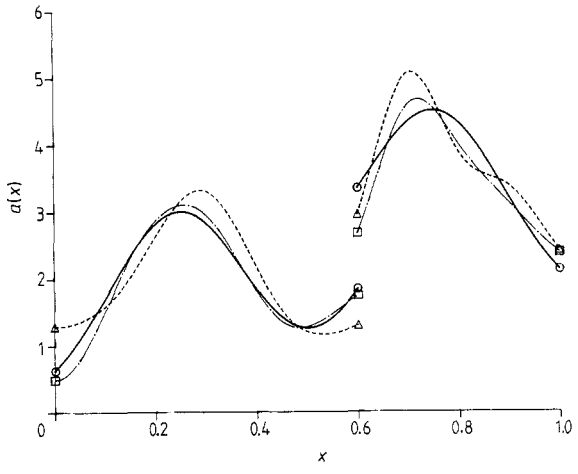

Figure 9. Least-squares estimates of α_{T1} at $N_p = 14$ with ξ known: q effect. ○, true; △, $q = 20$; □, $q = 80$.

Table 3. Least-squares estimation of α_{T1} with ξ known and $N_p = 14$: effect of withdrawal rates.

q	J_{LS}	J_S	Number of iterations
20	3.596×10^{-3}	3.167×10^{4}	10
80	3.708×10^{-3}	2.143×10^{4}	22

where q_j is the withdrawal rate at the jth source Ω_j, a small interval represented by

$$\Omega_j = \{x \mid |x - x_j| \leq r_j\} \qquad j = 1, \ldots, \mu_w \tag{4.10}$$

and X_{Ω_j} denotes the characteristic function. In our simulation, two constant rates, $q = 20$ and 80, were used at each of five point sources located at $x = 0.075$, 0.225, 0.525, 0.65 and 0.85; this amounts to the depletion of 12.5% and 50% of the material (or energy) of the system respectively at $t = 0.025$. The withdrawal builds up sharp spatial gradients around the point sources, thus leaving locally vanishing points in between. The results of least-squares estimation for α_{T1} obtained at $N_p = 14$ are shown in figure 9 and table 3. They clearly show that the better estimate is obtained with the higher q. This illustrates the point that a larger gradient resulting from the larger excitation provides more favourable estimation conditions. However there seems to be no indication that locally vanishing gradients affect the estimation performance significantly. This may be partly explained by considering that the least-squares estimation relies on the point observation, the locations of which do not necessarily coincide with the points where $\partial u/\partial x$ vanishes. But a full explanation should be furnished by detailed sensitivity analysis of the system under a given set of conditions.

Estimation with ξ unknown. The problem of identifying spatially discontinuous parameters becomes much more difficult when the locations of discontinuities ξ are unknown. Besides the difficulties of theoretical formulation mentioned in the previous section, we have potentially larger parameter spaces to search over than when ξ is known *a priori*. Our simulation study shows that the performance and convergence of the algorithm is severely dependent on the initial guess $\xi^{(0)}$ of ξ. This is well illustrated in table 4 and figure 10 which shows two different least-squares estimates for α_{T1} obtained with $N_p = 14$. The different initial guesses of ξ result in different estimates of ξ, and, consequently, the overall spatial distribution of the estimates looks distorted accordingly. It is apparent that these estimates are poorer than the one in figure 3 which was obtained with ξ known and at the same (optimum) level of discretisation. In fact, significantly larger J_{LS} values indicate premature local convergence.

Table 4. Least-squares estimation of α_{T1} with ξ unknown and $N_p = 14$: initial guess dependence.

$\xi^{(0)}$	$\hat{\xi}$	J_{LS}	J_S	Number of iterations
0.7	0.639	8.718×10^{-3}	1.078×10^{4}	11
0.4	0.405	5.920×10^{-3}	7.528×10^{3}	38

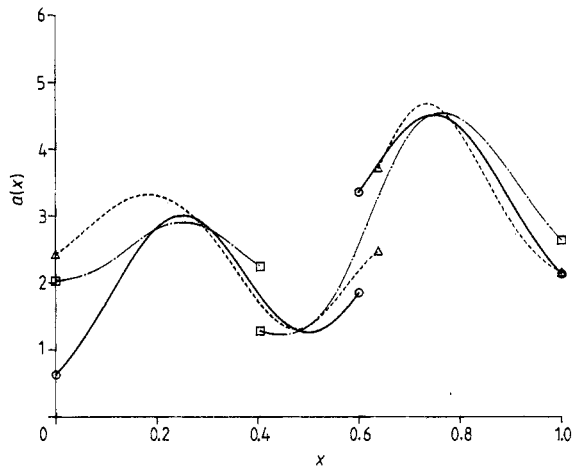

Figure 10. Least-squares estimates of α_{T1} at $N_p = 14$ with ξ unknown: $\xi^{(0)}$ dependence. ○, true; △, $\xi^{(0)} = 0.7$; □, $\xi^{(0)} = 0.4$.

Estimation runs at higher level of discretisation yielded poorer results, shown in figure 11 and table 5. Regularisation does not alleviate the convergence problems, though it reduces the anomalous oscillations in the least-squares estimates. The simulation runs for estimating α_{T2} also showed poor performance and are not presented here.

As pointed out earlier, the convergence problems encountered in estimation runs with ξ unknown reflect the difficulty of carrying out minimisation in the *larger* parameter space. Specifically, the parameter space can be conceived of as an indexed family of $C_\xi([0, 1])$, $\xi \in]0, 1[$. Thus, any optimisation process over this larger space, with its complicated structure, would suffer more from local convergence and/or a slow convergence rate than over $C_\xi([0, 1])$.

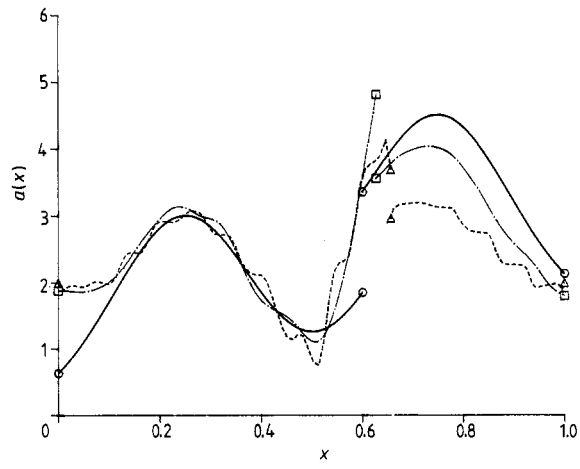

Figure 11. Least-squares and regularised estimates of α_{T1} at $N_p = 134$ with ξ unknown ($\xi^{(0)} = 0.7$). ○, true; △, $\beta = 0$; □, $\beta = 10^{-7}$.

Table 5. Least-squares and regularisation estimation of α_{T1} with ξ unknown ($\xi^{(0)} = 0.7$) and $N_p = 134$.

β	$\hat{\xi}$	J_β	J_{LS}	J_S	Number of iterations
0	0.656	1.764×10^{-2}	1.764×10^{-2}	6.751×10^{6}	19
10^{-7}	0.627	1.286×10^{-2}	8.038×10^{-3}	4.820×10^{4}	113

Another difficulty of the problem of estimating α with ξ unknown lies in its severe ill-posedness in the sense that point observations of u do not provide enough information to estimate ξ. In view of the stability of the boundary value problem (4.1) with respect to variations in the coefficient $\alpha \in L^\infty(]0, l[)$, the solution of (4.1) with discontinuous α_T is very close to the solution u_m (i.e. $\|u_m - u_T\| \leqslant \varepsilon$) with α_m, which is obtained by smoothing out α_T around ξ such that $\|\alpha_m - \alpha_T\| \leqslant \delta$. For the identification problem (which is an inverse problem of (4.1)), this stability may imply non-identifiability of ξ under a given level of uncertainty in the data.

This claim regarding the non-identifiability of ξ seems to be supported partly by the results of the following experiments. To reduce the adverse effect of variable ξ on the convergence of the optimisation, we fixed $\xi^{(0)}$ throughout the minimisation process. This is repeated for various values of $\xi^{(0)}$. Table 6 shows a partial list of the results of these experiments, and figure 12 the corresponding estimates. It can be seen that all the estimates in figure 12 exhibit the spatial distribution of α_{T1} well enough, except that the jump discontinuity is smoothed out in the estimates obtained with $\xi^{(0)} = 0.4$ and $\xi^{(0)} = 0.7$. Furthermore, it should be noted that the estimate obtained with $\xi^{(0)} = 0.6$ (i.e. with ξ known) does not yield the smallest J_{LS} in table 6.

5. Conclusion

A novel approach for the identification of spatially discontinuous parameters in second-order parabolic systems has been developed on the basis of a rigorous formulation of piecewise regularisation. Combined with piecewise cubic spline representations of unknown parameters, the approach provides a stable and efficient history-matching algorithm for one-dimensional fractured oil reservoirs. Our numerical experiments showed:

(i) regularisation is an effective method for obtaining a well behaved smooth approximate solution by alleviating the ill-conditioning present in the least-squares method;

Table 6. Least-squares estimation of α_{T1} with $\xi^{(0)}$ fixed and $N_p = 14$.

$\xi^{(0)}$	J_{LS}	J_S	Number of iterations
0.4	3.702×10^{-3}	2.406×10^{4}	19
0.6	3.746×10^{-3}	1.018×10^{4}	20
0.7	3.997×10^{-3}	2.411×10^{4}	23

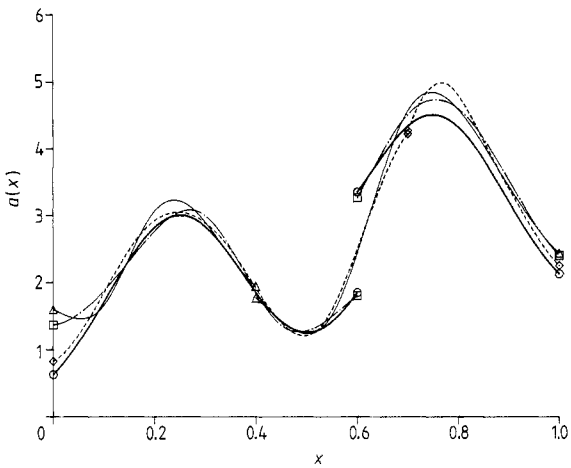

Figure 12. Least-squares estimates of α_{T1} at $N_p = 14$ with $\xi^{(0)}$ fixed: $\xi^{(0)}$ effect. ○, true; △, $\xi^{(0)} = 0.4$; □, $\xi^{(0)} = 0.6$; ◇, $\xi^{(0)} = 0.7$.

(ii) when the location of discontinuity is unknown, it is difficult to obtain an accurate estimate due to local convergence and unidentifiability of the discontinuity from point observations.

Acknowledgment

Acknowledgment is made to the Donors of The Petroleum Research Fund, administered by the American Chemical Society for the support of this research (PRF No 17911-AC2).

References

Adams R A 1975 *Sobolev Spaces* (New York: Academic)

Aubin J P 1972 *Approximation of Elliptic Boundary-Value Problems* (New York: Wiley-Interscience)

Bensoussan A J, Lions L and Papanicolaou G 1978 *Asymptotic Analysis for Periodic Structures* (Amsterdam: North-Holland)

Chung C B 1988 Identification of spatially varying parameters in distributed parameter systems *PhD Thesis* The University of Michigan, Ann Arbor

Gavalas G R, Shah P C and Seinfeld J H 1976 Reservoir history matching by Bayesian estimation *Soc. Pet. Eng. J.* **16** 337–50

Kravaris C and Seinfeld J H 1985 Identification of parameters in distributed systems by regularization *SIAM J. Control Optim.* **23** 217–41

—— 1986a Identification of spatially varying parameters in distributed systems by discrete regularization *J. Math. Anal Appl.* **119** 128–52

—— 1986b Identifiability of spatially-varying conductivity from point observation as an inverse Sturm–Liouville problem *SIAM J. Control Optim.* **24** 522–42

Kitamura S and Nakagiri S 1977 Identifiability of spatially-varying and constant parameters in distributed systems of parabolic type *SIAM J. Control Optim* **15** 785–802

Lamm P D 1987 An approximation technique for estimating discontinuous coefficients in distributed systems *SIAM J. Control Optim.* **25** 18–37

Lee T Y, Kravaris C and Seinfeld J H 1986 History matching by spline approximation and regularization on single-phase areal reservoirs *SPE Reservoir Eng.* **1** 521–34

Lions J L 1971 *Optimal Control of Systems Governed by Partial Differential Equations* (Berlin: Springer)

Seinfeld J H and Kravaris C 1982 Distributed parameter identification in geophysics—petroleum reservoirs and aquifers *Distributed Parameter Control Systems: Theory and Application* ed. S Tzafestas (Oxford: Pergamon) pp 367–90

Shanno D F and Phua K H 1978 Matrix conditioning and nonlinear optimization *Math. Programming* **14** 149–60

Simmons G F 1963 *Introduction to Topology and Modern Analysis* (New York: McGraw-Hill)

Tarantola A 1984 Inversion of seismic reflection data in the acoustic approximation *Geophys.* **49** 1259–66

Tikhonov A N 1963a Solution of ill-posed problems and the regularization method *Dokl. Akad. Nauk* **151** 501–4 (1963 *Sov. Math. Dokl.* **4** 1035–8)

—— 1963b Regularization of ill-posed problems *Dokl. Akad. Nauk* **153** 49–52 (1963 *Sov. Math. Dokl.* **4** 1624–7)

Yeh W W-G 1986 Review of parameter identification procedures in groundwater hydrology: the inverse problem *Water Resour. Res.* **22** 95–108