Azmy S. Ackleh; Ben G. Fitzpatrick

Estimation of discontinuous parameters in general nonautonomous parabolic systems

Kybernetika, Vol. 32 (1996), No. 6, 543--556

Persistent URL: http://dml.cz/dmlcz/125367

KYBERNETIKA — VOLUME 32 (1996), NUMBER 6, PAGES 543-556

ESTIMATION OF DISCONTINUOUS PARAMETERS IN GENERAL NONAUTONOMOUS PARABOLIC SYSTEMS

AZMY S. ACKLEH[1] AND BEN G. FITZPATRICK[2]

In this paper we present a unified convergence theory for estimating discontinuous parameters in a general class of nonautonomous parabolic systems. The application of this theory to estimate parameters in the Euler–Bernoulli beam equation, flow equations, and the Fokker–Planck population model is discussed.

1. INTRODUCTION

Identifying discontinuous parameters is crucial in many applications, including bioremediation of contaminated groundwater, in population biology problems, and in physical models for flexible structures. For example, the reproduction function of an individual in a population model is usually represented in terms of a discontinuous function of the form

$$\beta(t,y) = \begin{cases} 0 & y_B \le y < y_A \\ \beta(t,y) & y_A < y \end{cases}$$

where y_B is the birth size or age and y_A is the adult size or age (see for example, [11]). The stiffness of a beam in a flexible structure may decrease over a long period of time and could form a discontinuity at a certain time, due to a crack or other damage. Fluctuations in water tables due to precipitation may cause rapid changes in groundwater velocity field. The problems considered here involve applications of partial differential equation models, nonautonomous models with coefficients or other parameters which are (possibly) discontinuous in the time and/or space variable. From the point of view of computing parameter estimates, one must typically work with a numerical approximation scheme for integrating the differential equation, compare this simulation result to the observed data, and iterate over the parameter space until an acceptable least squares cost is obtained. A major issue in such computations is the impact of the numerical approximation on the estimation. We seek to analyze the convergence of minimizers of these approximate problems, as the discretization of the differential equation becomes finer.

[1]Research supported in part by AFOSR grant F49620–93–0198.
[2]Research supported in part by AFOSR grant F49620–95–1–0214.

General theory for parameter identification in an abstract setting can be found in [8]. In that work (and the many references contained therein) one finds that the main components in inverse problem analyses are compactness of parameter spaces, continuity of the system state with respect to the parameter, and convergence of numerical approximations that is uniform with respect to the parameters and consistent with the topology of the observation space. General theorems are available to guarantee the parameter convergence of interest, if one can verify compactness, continuity and uniform convergence as mentioned above (see in particular pp. 143–145 of [8]). For the autonomous case the authors in [7] developed a rather general abstract theoretical framework for parameter convergence and inverse problem. The sesquilinear form approach contained therein provides a unified way to handle a wide variety of problems, with conditions that can be verified in a straightforward manner. In the paper [1] results which extend the framework of [7] to nonautonomous parabolic problems were established, in order to allow general coverage of many problems, together with verifiable conditions on the sesquilinear form that determines the dynamics.

In the papers [16, 17] the authors developed convergence theory for estimating spatially discontinuous parameters in the autonomous scalar parabolic system, as well as, the autonomous second-order hyperbolic system arising from 1-D surface seismic problems. In both of these papers the parameters were restricted to the following 1-D spatially discontinuous form

$$q(x) = \varphi_0(x) + \sum_{i=1}^{\nu} H_{\xi_i}(x)\, \varphi_i(x)$$

where $\varphi_i,\ i = 0, \dots, \nu$, are continuous functions and $H_{\xi_i},\ i = 1, \dots, \nu$, is the usual Heaviside function ($H_\xi(x) = 1,\ x \in (\xi, 1]$, and $H_\xi(x) = 0$, otherwise) on $[0, 1]$. Our goal here is to further extend these continuity, compactness, and convergence results to general nonautonomous evolution systems with discontinuous (in time and/or space) parameters. We remark that in our subsequent analysis we use a more general parameter space then the one used above. In particular, our parameters can be time varying discontinuous functions defined on $[0, T] \times \Omega$, where Ω is a bounded subset of R^n. In addition, we do not restrict our parameters to have a specific discontinuous structure (see Section 2 for the definition of our parameter space).

Our theory is based on the weak version of the system in terms of sesquilinear forms used in [7] and [1]. The theory depends on the following properties of the time and parameter dependent sesquilinear form $\sigma(t, q)(\cdot\,, \cdot)$ describing the system: continuity with respect to the parameter, uniform boundedness (both in time and the parameter), and uniform coercivity in time and the parameter.

The paper is organized as follows. In Section 2, we present a theoretical framework for the approximation. The application of this theory to estimating discontinuous (both in time and space) parameters in the Euler–Bernoulli beam equation, flow equations, and the Fokker–Planck model is discussed in Section 3. Numerical findings are the topic of Section 4.

2. PARAMETER IDENTIFICATION

In this section, we consider an abstract, parameter dependent evolution system. We begin with some basic notation. Let H be a Hilbert space with inner product $\langle \cdot , \cdot \rangle$ and corresponding norm $| \cdot |$. Let V be a Hilbert space that is densely and continuously imbedded in H, with norm $\| \cdot \|$ and imbedding constant K : for each $\phi \in V$, we have $|\phi| \leq K\|\phi\|$. We use these spaces to form a Gelfand triple structure $V \hookrightarrow H = H^* \hookrightarrow V^*$.

We consider the following abstract differential equation on H

$$\begin{cases} \dot{u}(t,q) = A(t,q)\,u(t,q) + f(t,q), \\ u(0,q) = u_0(q). \end{cases} \tag{2.1}$$

We denote by q the parameter to be estimated. Since we are interested in discontinuous-in-time parameters, we take as our parameter set

$$BV_\infty = \{f \in L^1((0,T),\widetilde{Q}) : f \in BV([0,T],Q),\ \|f\|_{L^\infty([0,T],Q)} \leq M_1 \text{ and } TV(f) \leq M_2\}$$

where,

$$TV(f) = \sup \sum_{i=1}^{M} \|f(t_{i+1}) - f(t_i)\|_Q$$

the supremum taken over all finite partitions $0 = t_1 < \cdots < t_M = T$, with Q being compactly embedded in the normed linear space $\widetilde{Q}$. It is well known that $BV([0,T],Q)$ is compactly embedded in the space $L^1((0,T),\widetilde{Q})$ (see, [9]). Hence, since BV_∞ is a closed subset of $BV([0,T],Q)$ then it is compact in $L^1((0,T),\widetilde{Q})$. The generality of the "range space" Q allows us to consider various types of spatially dependent parameters.

The mapping $q \to f(\cdot,q)$ is assumed to be continuous from $BV_\infty \subset L^1(0,T;\widetilde{Q})$ into H. The operator A is assumed to be determined by a time and parameter dependent sesquilinear form on V; i.e., $\sigma(\cdot,\cdot)(\cdot,\cdot) : [0,\infty) \times BV_\infty \times V \times V \to C$, where $\sigma(t,q)(\cdot,\cdot)$ is sesquilinear for each $t \in [0,\infty)$ and $q \in BV_\infty$. Concerning σ, we make the following assumptions.

($\Sigma 0$) The function $\sigma(\cdot,q)(\phi,\psi)$ is measurable on $[0,\infty)$, for fixed $\phi,\psi \in V$ and $q \in BV_\infty$.

($\Sigma 1$) There exists $K_0 > 0$ such that $\sigma(t,q)(\phi,\psi)| \leq K_0\|\phi\| \cdot \|\psi\|\ \forall \phi,\ \psi \in V$, $q \in BV_\infty$ uniformly in t on each interval $[0,T]$.

($\Sigma 2$) There exists $c_0 > 0$, $\lambda_0 \in R$ such that $\sigma(t,q)(\phi,\phi) + \lambda_0|\phi|^2 \geq c_0\|\phi\|^2$, $\forall \phi \in V$, $q \in BV_\infty$ uniformly in t on each interval $[0,T]$.

($\Sigma 3$) For each $\phi \in V$, there is a sequence of functions $\mu^N(t)$ such that for any $q^N \to q$ in BV_∞, then for any interval $[0,T]$ we have

$$\lim_{N\to\infty} \int_0^T |\mu^N(t)|^2 \, dt = 0,$$

and for a.e. t and each $\psi \in V$,

$$\left|\sigma(t,q)(\phi,\psi) - \sigma(t,q^N)(\phi,\psi)\right| \le \mu^N(t)\cdot \|\psi\|.$$

We note here that the assumption (Σ3) above is less restrictive then the assumption (Σ3) announced earlier in [2]. Such modification allows the convergence theory for estimating discontinuous (both in time and space) parameters to hold. It is a well-known consequence of the Riesz Theorem that, under the assumptions (Σ0)–(Σ2) there exists a family of uniquely determined linear operators $A(t,q) : \mathrm{dom}(A(t,q)) \to H$, with dense domains, satisfying

$$\sigma(t,q)(\phi,\psi) = \langle -A(t,q)\,\phi,\psi\rangle,$$

for all $\phi \in \mathrm{dom}\ A(t,q),\ \psi \in V$.

The main goal of this paper involves the convergence of parameter estimates determined from approximations of the above dynamics. Toward that end, we consider an approximation method based on a sequence of Hilbert spaces $H^N,\ N = 1,2,\ldots$, with orthogonal projections $P^N : H \to H^N$. To obtain convergence results, the following assumption about these approximations will be needed.

(A1). The subspaces H^N are subsets of V, and $\forall\, v \in V$, we have that $\|P^N v - v\| \to 0$, as $N \to \infty$.

Assumption (A1) is satisfied by many finite element and spectral schemes (see [8, 10, 15]). The Galerkin approach to approximation involves restricting $\sigma(t,q)$ to $H^N \times H^N$, yielding bounded linear operators $A^N(t,q)$ satisfying

$$\sigma(t,q)(\phi^N,\psi^N) = -\langle A^N(t,q)\phi^N,\psi^N\rangle.$$

Using the above assumptions we have the following theorem:

Theorem 2.1. Suppose that (Σ0)–(Σ3), and (A1) hold, and that $\{q^N\}_{N=1}^{\infty} \subset BV_\infty$ satisfying $q^N \to q$ in the $L^1((0,T),\widetilde{Q})$ topology. Then we have that $u^N(t,q^N) \to u(t,q)$, in H, uniformly on $[0,T]$.

Proof. We first define $z^N = u^N - P^N u$. It is sufficient to show that $z^N(t) \to 0$, uniformly in t. Now, $\forall\, \phi \in H^N$, we have that

$$\begin{aligned}
\langle \dot z^N,\phi\rangle_{V^*,V} &= \langle \dot u^N - \dot u + (\mathrm{d}/\mathrm{d}t)\,(u - P^N u),\phi\rangle_{V^*,V}\\
&= \langle \dot u^N - \dot u,\phi\rangle + (\mathrm{d}/\mathrm{d}t)\,\langle u - P^N u,\phi\rangle_{V^*,V}\\
&= \langle \dot u^N - \dot u,\phi\rangle_{V^*,V} + (\mathrm{d}/\mathrm{d}t)\,\langle u - P^N u,\phi\rangle_H\\
&= \langle \dot u^N - \dot u,\phi\rangle_{V^*,V}.
\end{aligned}$$

Using this fact we have that

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|z^N|^2 &= \langle \dot z^N, z^N\rangle = \langle \dot u^N - \dot u, z^N\rangle\\
&= -\sigma(t,q^N)\,(u^N,z^N) + \sigma(t,q)\,(u,z^N)\\
&\quad - \langle f(t,q) - P^N f(t,q^N), z^N\rangle.
\end{aligned}$$

Hence adding and subtracting a few terms and using coercivity, we have that for a.e. $t \geq 0$,

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|z^N|^2 + c_0\|z^N\|^2 - \lambda_0|z^N|^2 \\
\leq\;& \sigma(t,q^N)\,(u - P^N u, z^N) + \sigma(t,q)\,(u,z^N) - \sigma(t,q^N)\,(u,z^N) \\
&\qquad + |f(t,\cdot,q) - P^N f(t,\cdot,q^N)|\,|z^N| \\
\leq\;& K_0\|u - P^N u\|\cdot\|z^N\| + \mu^N(t)\|z^N\| + |f(t,q) - P^N f(t,q^N)|\cdot|z^N|.
\end{aligned}$$

Choosing $\epsilon = c_0$ and integrating both sides and using standard inequalities of the form $ab \leq \frac{\epsilon}{2}a^2 + \frac{1}{2\epsilon}b^2$ we get.

$$\begin{aligned}
\frac{1}{2}|z^N(t)|^2 \;\leq\;& \frac{K_0^2}{2\epsilon}\int_0^t \|u - P^N u\|^2\,\mathrm{d}t + C\int_0^t |z^N|^2\,\mathrm{d}t + \frac{1}{2\epsilon}\int_0^t |\mu^N(t)|^2\,\mathrm{d}t \\
&+ \frac{1}{2\epsilon}\int_0^t |f(t,q) - P^N f(t,q^N)|^2\,\mathrm{d}t.
\end{aligned}$$

A standard application of Gronwall's inequality together with the dominated convergence theorem we get

$$|z^N(t)|^2 \to 0$$

uniformly on $[0,T]$. □

We have thus obtained, based on the assumptions given above, that $u^N(t;q^N) \to u(t;q)$ in H, when $q^N \to q$ in the $L^1((0,T),\tilde{Q})$ topology. To put this result into the context of least squares estimation, we consider a continuous map $\mathcal{C}\colon H \to Z$, where Z is a normed linear space. Given $z \in Z$, one determines an appropriate parameter value for the system by minimizing

$$J(q) = \|\mathcal{C}u(q) - z\|^2.$$

The continuous dependence results above indicate that a minimizer exists within the compact set BV_∞.

In order to compute minimizers, we must make some approximations. The approximation u^N of the state variable u, as discussed above, lead to a cost functional

$$J^N(q) = \|\mathcal{C}u^N(q) - z\|_Z^2$$

to be minimized. The above convergence results guarantee that if $\{q^N\}_{N=1}^\infty \subset BV_\infty$ satisfying $q^N \to q$ in $L^1((0,T),\tilde{Q})$ then $J^N(q^N) \to J(q)$, which will give us (see, e.g., pp. 143–145 of [8]) subsequential convergence of minimizers.

In the next section we present some examples, to illustrate the application of this general theory.

3. APPLICATIONS

In this section we present three examples to which the theoretical framework of the previous section apply. We hope, thus, to give the reader a sense of the wide applicability of this unified theory.

3.1. Euler–Bernoulli beam equation with Kelvin–Voigt damping

Consider the following second order hyperbolic equation which describes the transverse vibrations of a cantilevered Euler–Bernoulli beam with Kelvin–Voigt internal damping.

$$\begin{gathered} y_{tt} + (EIy_{xx} + c_D I y_{xxt})_{xx} = f(t,x) \\ y(t,0) = y_x(t,0) = 0 \\ EIy_{xx}(t,l) + c_D I y_{xxt}(t,l) = 0, \\ (EIy_{xx}(t,l) + c_D I y_{xxt}(t,l))_x = 0 \\ y(0,x) = y_0(x), \qquad y_t(0,x) = y_1(x). \end{gathered} \tag{3.1}$$

Here, the function $y(t,x)$ is the displacement along the beam at time t at position x. The parameter $EI(t,x)$ is the stiffness coefficient, and $c_D I(t,x)$ is the Kelvin–Voigt damping coefficient. The function f represents external distributed forces applied to the beam. For details on this model, see [6].

Equation (3.1) may be written as a first-order system in the following way: Define $w(t,x) = [y(t,x), \frac{\partial y}{\partial t}(t,x)]^T$ and $\hat{F} = [0, f(t,x)]^T$. Then denoting $\frac{\partial^2}{\partial x^2}$ by D^2, we see that (3.1) is equivalent to

$$w_t(t,x) = \mathcal{A}(t,q)w(t,x) + \hat{F}(t,x), \tag{3.2}$$

where (formally)

$$\mathcal{A} = \begin{bmatrix} 0 & I \\ -D^2(EID^2) & -D^2(c_D ID^2) \end{bmatrix}$$

with the initial condition $w(0) = w_0 = (y_0, y_1)$.

To write equation (3.2) in a weak formulation, we define the following spaces

$$\begin{aligned} H_L^2(0,l) &= \{u \in H^2(0,l) | u(0) = u_x(0) = 0\} \\ H &= H_L^2(0,l) \times L^2(0,l) \\ V &= H_L^2(0,l) \times H_L^2(0,l). \end{aligned}$$

The inner product on the $L^2(0,l)$ will be denoted by $(\cdot,\cdot)$ and for the space $H_L^2(0,l)$ we use the inner product

$$\langle\langle \phi, \psi \rangle\rangle = (\phi_{xx}, \psi_{xx})$$

and the associated norm $|||\cdot|||$. It is easily shown that this norm is equivalent to the usual $H_L^2(0,l)$ norm using Poincare's inequality. The inner products for the spaces H, V will be taken to be the usual product space inner products, and will be denoted as in the abstract formulation $\langle\cdot,\cdot\rangle$ and $|\cdot|$, and $\langle\cdot,\cdot\rangle_V$ and $\|\cdot\|$ respectively. Then as in [7] the weak form of the beam can be written in terms of the following sesquilinear form: with $w = (u,v)$ and $\mathcal{X} = (\phi,\psi)$ elements of V, define

$$\sigma(t,q)(w,\mathcal{X}) = -\langle\langle v, \phi \rangle\rangle + \sigma_1(t,q)(u,\psi) + \sigma_2(v,\psi),$$

where,

$$\begin{aligned}\sigma_1(t,q)(u,\psi) &= (EIu_{xx},\psi_{xx}),\\ \sigma_2(t,q)(v,\psi) &= (c_DIv_{xx},\psi_{xx}).\end{aligned}$$

Then with $w=(y,\dot{y})$, the weak form of equation (3.2) can be written

$$\langle \dot{w}(t),\mathcal{X}\rangle + \sigma(t,q)(w(t),\mathcal{X}) = \langle \hat{F}(t),\mathcal{X}\rangle,$$

for $\mathcal{X}\in V$.

In terms of the abstract formulation developed in the previous section we will set $q=(EI,c_DI)$, $\Omega=(0,l)$,

$$\widetilde{Q} = L^1(\Omega)\times L^1(\Omega)$$

and

$$\begin{aligned}Q \;=\; \{(EI,c_DI)\in (BV(\Omega))^2 : d_0\le EI(x)\le d_1,\;\; d_2\le c_DI(x)\le d_3,\\ TV(EI)\le d_4,\;\; TV(c_DI)\le d_5\}\end{aligned}$$

Compactness of the set Q in $\widetilde{Q}$ is a well known result (see for example [13]). Standard arguments and the fact that $EI(t,x)\ge d_0$ and $c_DI(t,x)\ge d_2$ easily verify the assumptions $(\Sigma 0)-(\Sigma 2)$. For verification of $(\Sigma 3)$ we let $w=(u,v),\mathcal{X}=(\phi,\psi)\in V$, and suppose that $q^N\to q$ in $L^1((0,T),\widetilde{Q})$. Then we have that

$$\begin{aligned}|\sigma(t,q)(w,\mathcal{X})-\sigma(t,q^N)(w,\mathcal{X})| &\le |\sigma_1(t,q)(u,\psi)-\sigma_1(t,q^N)(u,\psi)|\\ &\quad + |\sigma_2(t,q)(v,\psi)-\sigma_1(t,q^N)(v,\psi)|\\ \le\; & \left(\int_0^l |(EI(x))(t)-(EI^N(x))(t)|^2|u_{xx}|^2 dx\right)^{\frac{1}{2}}\left(\int_0^l |\psi_{xx}|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\ &+\left(\int_0^l |(c_DI(x))(t)-(c_DI^N(x))(t)|^2|v_{xx}|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_0^l |\psi_{xx}|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\end{aligned}$$

Hence we have that

$$|\sigma(t,q)(w,\mathcal{X})-\sigma(t,q^N)(w,\mathcal{X})| \le \mu(t).\|\mathcal{X}\|_V,$$

where

$$\begin{aligned}\mu(t) \;=\; & \left(\int_0^l |(EI(x))(t)-(EI^N(x))(t)|^2|u_{xx}|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\ &+\left(\int_0^l |(c_DI(x))(t)-(c_DI^N(x))(t)|^2|v_{xx}|^2\,\mathrm{d}x\right)^{\frac{1}{2}}.\end{aligned}$$

It remains to show now that

$$\lim_{N\to\infty}\int_0^T |\mu^N(t)|^2\,\mathrm{d}t = 0,$$

to verify that $(\Sigma 3)$ holds. To this end, we will restate the following theorem (see page 198 in [12]).

Theorem 3.1.1. Let (S, Σ_S, ν) and (E, Σ_E, λ) be measure spaces which are either both finite or both positive and σ-finite, and let (R, Σ_R, ρ) be their product. Let $1 \leq p \leq \infty$ and let F be a ν-integrable function from S to $L^p(E, \Sigma_E, \lambda, X)$ where X is a real or complex B-space. Then there is a ρ-measurable function f on R to X, which is uniquely determined except for a set of ρ-measure zero, and such that $f(s, \cdot) = F(s)$ for ν-almost all $s \in S$. Moreover for λ-almost all t, $f(\cdot, t)$ is ν-integrable on S, and $\int_S f(s, \cdot)\, \nu(\mathrm{d}s) = \int_S F(s)\, \nu(\mathrm{d}s)$ in $L^p(E, \Sigma_E, \lambda, X)$.

We will now prove the following lemma,

Lemma 3.1.1. Let Ω be a bounded set in R^n. Suppose that

(i) $f^N, f \in BV_\infty$

(ii) $\widetilde{Q} = L^1(\Omega)$

(iii) $Q = \{a \in \widetilde{Q} : ||a||_\infty \leq C_1,\ TV(a) \leq C_2\}$

(iv) $f^N \to f$ in the $L^1((0,T), \widetilde{Q})$ topology

(v) $g \in L^2((0,T), L^2(\Omega))$ fixed.

Then,

$$\lim_{N \to \infty} \int_0^T \left(\int_\Omega |f^N - f|^2 \, |g|^2 \, \mathrm{d}x \right) \mathrm{d}t = 0.$$

Proof. By Theorem 3.1.1 there exist a unique $\widetilde{f}^N, \widetilde{f} \in L^1((0,T) \times \Omega)$ such that the following hold for a. e. t

$$\int_\Omega |\widetilde{f}(t,x) - \widetilde{f}^N(t,x)| \, \mathrm{d}x = \int_\Omega |(f(x))\,(t) - (f^N(x))\,(t)| \, \mathrm{d}x.$$

Since $f^N \to f$, as $N \to \infty$ in the $L^1((0,T), \widetilde{Q})$ topology then we have that $\widetilde{f} - \widetilde{f}^N \to 0$ in $L^1((0,T) \times \Omega)$, which in turn implies that $\widetilde{f}^N \to \widetilde{f}$ in measure on the product space $(0,T) \times \Omega$.

Hence,

$$\begin{aligned} &\int_0^T \left(\int_\Omega |(f(x))\,(t) - (f^N(x))\,(t)|^2 \, |g|^2 \, \mathrm{d}x \right) \mathrm{d}t \\ = &\int_0^T \left(\int_\Omega |\widetilde{f}(t,x) - \widetilde{f}^N(t,x)|^2 \, |g|^2 \, \mathrm{d}x \right) \mathrm{d}t \end{aligned}$$

By the definition of Q, we have that $|(f(x))\,(t)|$, $|(f^N(x))\,(t)| \leq C_1$, for a. e. t and x. Thus, the functions $\widetilde{f}^N$ and $\widetilde{f}$ are uniformly bounded (a. e. in the product space); thus, the result follows by applying the dominated convergence theorem. □

Using Lemma 3.1.1, it is straightforward to see that

$$\lim_{N \to \infty} \int_0^T |\mu^N(t)|^2 \, \mathrm{d}t = 0$$

and hence ($\Sigma 3$) is verified.

3.2. Estimation of diffusion coefficient in flow equations

In this example we consider the following equation

$$\begin{cases} u_t - \nabla(a(t,x)\,\nabla u) = f(t,x) & \\ u(t,x) = 0, & (t,x) \in (0,T) \times \partial\Omega \\ u(0,x) = u_0(x) & x \in \Omega \end{cases}$$

which describes the flow of a fluid through the medium with permeability $a(t,x)$. Here Ω is assumed to be open bounded domain in R^n. We note that convergence theory for estimating diffusion coefficient in the above model for the autonomous case ($a = a(x)$) has been considered in [14].

In terms of the abstract theory if we define $H = L^2(\Omega)$, and $V = H_0^1(\Omega)$, the above equation can be written in the following weak form as follows

$$\langle u_t, \phi\rangle + \sigma(t,a)\,(u,\phi) = \langle f, \phi\rangle \qquad \text{for } \phi \in V$$

where,

$$\sigma(t,a)\,(u,\phi) = \langle a\nabla u, \nabla\phi\rangle.$$

In terms of our theoretical setting we define the set $\widetilde{Q} = L^1(\Omega)$ and

$$Q = \{a \in BV(\Omega) : 0 < d_0 \le a(x) \le d_1 \text{ a.e. in } \Omega,\ TV(a) \le c_2\}.$$

Clearly Q is a compact set $\in \widetilde{Q}$. Standard arguments and the fact that $a(t,x) \ge d_0$ easily verify the assumptions $(\Sigma 0)-(\Sigma 2)$. For verification of $(\Sigma 3)$ suppose that $a^N \to a$ in the $L^1((0,T),\widetilde{Q})$ topology. Then we have that

$$\begin{aligned} &|\sigma(t,a)\,(u,\psi) - \sigma(t,a^N)\,(u,\psi)| \\ \le\ & \left(\int_\Omega |(a(x))\,(t) - (a^N(x))\,(t)|^2\,|\nabla u|^2\,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_0^l |\nabla\psi|^2\,\mathrm{d}x\right)^{\frac{1}{2}}. \end{aligned}$$

We define

$$\mu^N(t) = \left(\int_\Omega |(a(x))\,(t) - (a^N(x))\,(t)|^2\,|\nabla u|^2\,\mathrm{d}x\right)^{1/2},$$

and again, by Lemma 3.1.1, we see that

$$\lim_{N\to\infty} \int_0^T |\mu^N(t)|^2\,\mathrm{d}t = 0.$$

3.3. The Fokker–Planck population model

Consider the following "Fokker–Planck" model which is used to model the dynamics of many structured populations (see, e. g., [3, 4, 7]):

$$u_t + (Gu)_x = -q_3 u \quad x_0 < x < x_1$$

where,

$$
\begin{gathered}
(Gu)(t,x) = q_1(t,x)\,u(t,x) - (q_2(t,x)\,u(t,x))_x,\\
(Gu)(t,x_0) = \int_{x_0}^{x_1} q_4(t,\xi)\,u(t,\xi)\,\mathrm{d}\xi,\\
(Gu)(t,x_1) = 0,\\
u(0,x) = u_0(x).
\end{gathered}
$$

Here, $u(t,x)$ is the density of individuals in the interval $(x, x+\mathrm{d}x)$ at time t. The term Gu is the population flux and the boundary condition at x_0 represents the recruitment of new individuals. The boundary at x_1 models the fact that no individuals grow beyond size x_1. For more details, please refer to [3, 4, 7].

By defining $H = L^2(x_0, x_1)$ and $V = H^1(x_0, x_1)$ the associated sesquilinear form is given by

$$
\sigma(t,q)(u,\psi) = -\langle Gu, \psi_x\rangle + \langle q_3 u, \psi\rangle - \psi(x_0)\int_{x_0}^{x_1} q_4(t,\xi)\,u(t,\xi)\,\mathrm{d}\xi.
$$

The above equation can be written in the following weak form

$$
\langle u_t, \psi\rangle + \sigma(t,q)(u,\psi) = 0 \quad \forall\, \psi \in V.
$$

In terms of the theory developed in Section 2 we define the sets

$$
\widetilde{Q} = L^1(x_0,x_1) \times W^{1,\infty}(x_0,x_1) \times L^1(x_0,x_1) \times L^1(x_0,x_1)
$$

and,

$$
\begin{aligned}
Q \;=\; & \{(q_1,q_2,q_3,q_4) \in \widetilde{Q} \mid TV(q_i) \le d_1,\ \|q_i\|_\infty \le d_2,\ i = 1,3,4,\\
& \|q_2\|_{W^{1,\infty}} \le d_3 \text{ and } q_2(x) \ge d_4 > 0\}.
\end{aligned}
$$

Standard arguments can be used to verify that $(\Sigma 0)-(\Sigma 2)$ hold for this problem (see for example, [3]). To verify $(\Sigma 3)$ similar techniques as in the above two examples can be used to show that

$$
|\sigma(t,q)(u,\phi) - \sigma(t,q^N)(u,\phi)| \le \mu(t)\,\|\phi\|_V
$$

where,

$$
\begin{aligned}
\mu^N(t) \;=\; & \left(\int_{x_0}^{x_1} |(q_1^N(x))(t) - (q_1(x))(t)|^2 |u|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\
& + \left(\int_{x_0}^{x_1} |(q_2^N(x))(t) - (q_2(x))(t)|^2 |u_x|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\
& + \left(\int_{x_0}^{x_1} |({q_2}_x{}^N(x))(t) - ({q_2}_x(x))(t)|^2 |u|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\
& + \left(\int_{x_0}^{x_1} |(q_3^N(x))(t) - (q_3(x))(t)|^2 |u|^2\,\mathrm{d}x\right)^{\frac{1}{2}}\\
& + \left(\int_{x_0}^{x_1} |(q_4^N(x))(t) - (q_4(x))(t)|^2 |u|^2\,\mathrm{d}x\right)^{\frac{1}{2}}.
\end{aligned}
$$

By Lemma 3.1.1 we have that

$$\lim_{N\to\infty}\int_0^T |\mu^N(t)|^2\,dt = 0.$$

4. NUMERICAL RESULTS

For the inverse problem numerical experiments, we focus on the model

$$\begin{aligned} & y_{tt} + (EI(t)\,y_{xx} + 0.01 y_{xxt})_{xx} \\ = \;& 100\sin(5\pi t) \times \frac{1}{.01}\chi_{[.495,.505]}(x) \\ & y(t,0) = y_x(t,0) = 0 \\ & EIy_{xx}(t,1) + c_D I y_{xxt}(t,1) = 0, \\ & (EIy_{xx}(t,1) + c_D I y_{xxt}(t,1))_x = 0 \\ & y(0,x) = y_0(x), \qquad y_t(0,x) = y_1(x). \end{aligned}$$

Note that this is a special case of the model (3.1). Here the forcing function approximates a δ function in the spatial variable.

In our computational methods, we estimated the stiffness parameter EI as a function of time from computationally generated data. For the simulations we present here, our FORTRAN program uses 15, uniformly spaced, cubic B-splines to approximate the solution of the Euler Bernoulli differential equation and the identification algorithm uses 10, uniformly spaced, piecewise constant functions to estimate the parameter. For the generated data, we used for EI the function

$$EI(t) = \begin{cases} 12 & t < 0.35 \\ 10 & t \geq 0.35 \end{cases}$$

which is constant with respect to the spatial variable.

For observation data, we sampled the displacement $u(t_i, x = 1)$ at 200 uniformly spaced time points t_i in the time interval $[0, 1]$, as generated with the above model. To test the behavior of the least squares identification procedure, we used as data the actual model generated signal, as well as the signal modified by Gaussian noise: $z_i = u(t_i, 1) \times (1 + \epsilon_i)$, was used for data, with ϵ_i a random sample from a zero mean Gaussian random number generator. We used $\sigma = .01$ and $\sigma = .05$ for standard deviations for the noise.

To implement the above mentioned compactness constraints, we used the following penalized least squares functional

$$J(EI) = \sum_{i=1}^{200} |z_i - u(t_i, 1; EI)|^2 + \beta\int_0^1 \sqrt{|\dot{EI}(t)|^2 + \alpha}\;dt$$

with α small positive constant and two different choices of β, depending on the noise level. Note that the integral term is (at least for smooth EI) a differentiable approximation to the total variation of EI.

In Figure 1, we presented the estimated (dashed-line) versus the true (solid-line) function EI. We used the constant function $EI = 12$ as our initial guess in the optimization, which was carried out using the package `lmdif1` from `netlib`. The regularization parameter β used was 10^{-4} and $\alpha = 10^{-5}$.

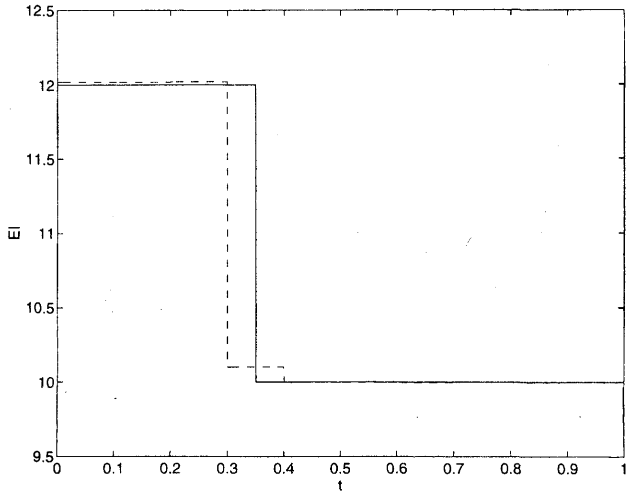

Fig. 1.

Figures 2 and 3 represent the same procedure as Figure 1 when the data was corrupted by the above described noise. The solid line is again the true EI; the dashed line is the estimated EI with $\sigma = .01$ and $\sigma = 0.05$, respectively. The regularization parameter values used for the two estimation runs were $\alpha = 10^{-5}$, $\beta = 10^{-4}$ and $\beta = 10^{-3}$, respectively.

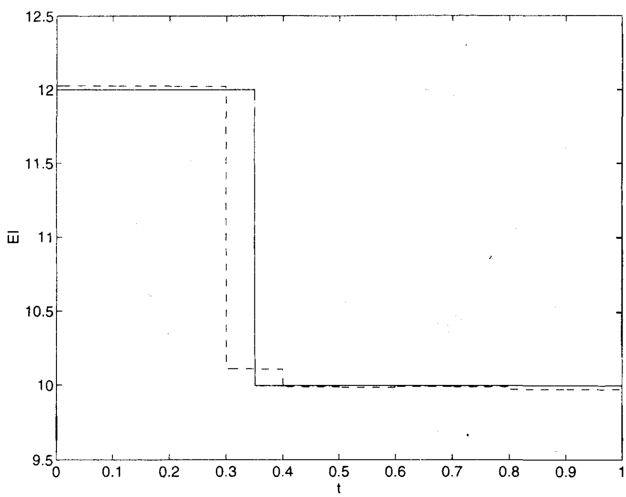

Fig. 2.

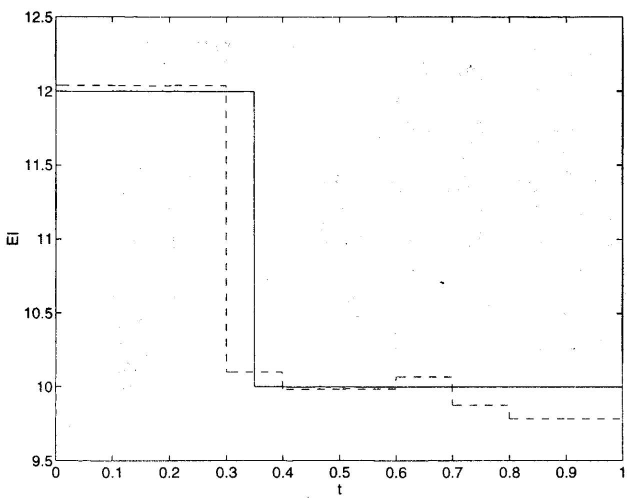

Fig. 3.

(Received February 14, 1996.)

REFERENCES

[1] A. S. Ackleh and B. G. Fitzpatrick: Estimation of time dependent parameters in general parabolic evolution systems. J. Math. Anal. Appl. *203* (1996), 464–480.

[2] A. S. Ackleh and B. G. Fitzpatrick: Estimation of temporally discontinuous parameters in general parabolic evolution systems. In: Proceedings of the 3rd IEEE Mediterranean Symposium on New Directions in Control and Automation, Vol. 1, pp. 280–286.

[3] H. T. Banks: Computational techniques for inverse problems in size structured stochastic population models. In: Control of Partial Differential Equations (A. Bermudez, ed.), (Lecture Notes in Control and Inform. Sci. 114), Springer–Verlag, Berlin 1989, pp. 3–10.

[4] H. T. Banks, L. W. Botsford, F. Kappel and C. Wang: Modeling and estimation in size structured population models. In: Mathematical Ecology, Proceeding of the Autumn Course Research Seminars (T. G. Hallam, L. J. Gross and S. A. Levin, eds.), World Scientific, 1988, pp. 521–541.

[5] H. T. Banks and J. M. Crowley: Parameter Estimation for Distributed Systems Arising in Elasticity. LCDS Report No. 81–24, Brown University. In: Proc. Symposium on Engineering Sciences and Mechanics, National Cheng Kung University, Tainan 1981, pp. 158–177.

[6] H. T. Banks, J. M. Crowley and I. G. Rosen: Methods for the identification of material parameters in distributed models for flexible structures. Mat. Apl. Comput. *5* (1986), 2, 139–168.

[7] H. T. Banks and K. Ito: A unified framework for approximation in inverse problems for distributed parameter systems. Control Theory Advanced Technology *4* (1988), 1, 73–90.

[8] H. T. Banks and K. Kunisch: Estimation Techniques for Distributed Parameter Systems. Birkhäuser, Boston – Basel 1989.

[9] V. Barbu: Convexity and optimization in Banach spaces. Editura Academiei, Holland 1986.
[10] C. Canuto and A. Quateroni: Approximation results for orthogonal polynomials in Sobolev spaces. Math. Comp. *38* (1982), 157, 67-86.
[11] O. Diekmann, M. Gyllenberg and H. R. Thieme: Perturbing evolutionary systems by step responses and cumulative outputs. Differential Integral Equations *8* (1995), 1205–1244.
[12] N. Dunford and T. Schwartz: Linear Operators Part I: General Theory. Interscience Publishers, New York 1958.
[13] E. Giusti: Minimal Surfaces and Functions of Bounded Variations. Birkhäuser, Boston - Basel 1976.
[14] S. Gutman: Identification of discontinuous parameters in flow equations. SIAM J. Control Optim. *28* (1990), 1049–1060.
[15] C. Johnson: Numerical Solution of Partial Differential Equations by the Finite Element Method. Cambridge Press, Cambridge 1987.
[16] P. K. Lamm: Estimation of discontinuous coefficients in parabolic systems: Applications to reservoir simulation. SIAM J. Control Optim. *25* (1987), 18–37.
[17] P. K. Lamm and K. A. Murphy: Estimation of discontinuous coefficients and boundary parameters for hyperbolic systems. Quart. Appl. Math. *46* (1988), 1–22.

Dr. Azmy S. Ackleh, Department of Mathematics, University of Southwestern Louisiana, Lafayette, LA 70504-1010. U. S. A.

Dr. Ben G. Fitzpatrick, Center for Research in Scientific Computation and Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. U. S. A.