QUARTERLY OF APPLIED MATHEMATICS

VOLUME XLVI MARCH · 1988 NUMBER 1

MARCH 1988, PAGES 1-22

ESTIMATION OF DISCONTINUOUS COEFFICIENTS AND BOUNDARY PARAMETERS FOR HYPERBOLIC SYSTEMS*

By

PATRICIA K. LAMM (*Southern Methodist University*)

AND

KATHERINE A. MURPHY (*Brown University*)

Abstract. We consider the problem of estimating discontinuous coefficients, including locations of discontinuities, that occur in second-order hyperbolic systems typical of those arising in 1-D surface seismic problems. In addition, we treat the problem of identifying unknown parameters that appear in boundary conditions for the system. A spline-based approximation theory is presented, together with related convergence findings and representative numerical examples.

1. Introduction. In this paper we consider a one-dimensional seismic inverse problem. Our goal is to contruct a parameter estimation scheme for the hyperbolic model equations treated in [7] extending these ideas to allow the estimation of discontinuous coefficients (including the location of discontinuities). The approach taken here contrasts with that taken in [7] in that we consider a different decomposition of the wave equation, yielding a different operator and state space for theoretical arguments. We combine these ideas with a variation of the approximation scheme in [15] which was developed to treat the problem of estimating discontinuous coefficients in parabolic systems.[1]

The underlying theoretical approach to the identification problem follows the outline of other related papers (eg. [3], [4], [5], [6], [7], [15]) in that we define the (infinite-dimensional) identification problem and construct associated approximate identification problems; under compactness assumptions on our parameter space, we investigate questions of convergence for parameter estimates and approximating state variables. Our arguments here are based on an application of the Trotter–Kato theorem. In Sec. 5 we describe the numerical algorithm and conclude with some preliminary computational examples.

*Received February 28, 1986.

[1] Research supported in part by NSF grants DMS-8200883, DMS-8601968, and AFOSR contract ISSA-860079 (P.K.L.), and NSF grants DMS-8205355, DMS-8504316 and AFOSR contracts 84-0398 and 86-0256 (K.A.M.). Parts of the research were carried out while both authors were visitors at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA which is operated under NASA Contract Nos. NAS1-16394 and NAS1-17130.

We shall employ standard notation throughout, using, for example, $H^p(\Omega)$ and $W_p^m(\Omega)$ to denote the usual Sobolev spaces on Ω. If Ω is not specified, it is assumed that $\Omega = [0, 1]$. Further, given any $w \in L^\infty(0, 1)$ satisfying $0 < \underline{w} \le w(x) \le \overline{w}$, for almost all x, we shall define the $\frac{1}{w}$-weighted $H^0(0, 1)$ space, denoted $H^0(w)$, with inner product $\langle u, v\rangle_w \equiv \int_0^1 \frac{1}{w} uv$ (and associated norm $|\cdot|_w$). Similarly, given any $c > 0$ we define the c-weighted real line by $\mathbf{R}(c)$ with inner product of two elements $u,\ v \in \mathbf{R}(c)$ defined by cuv. Finally, throughout we shall write $I: Y \to Y$ to denote the identity operator, where the space Y will be clear from the context.

2. The identification problem. As in [7], we consider the problem of estimating unknown parameters that appear in the system

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = \frac{1}{\rho(x)} \frac{\partial}{\partial x}\left(E(x)\frac{\partial u}{\partial x}\right), & t \in (0, \bar{t}],\ x \in [0, 1], \\ \frac{\partial u}{\partial x}(t, 0) - k_1 u(t, 0) = s(t; \tilde{k}), & \\ \frac{\partial u}{\partial t}(t, 1) + k_2 \frac{\partial u}{\partial x}(t, 1) = 0, & \\ u(0, x) = \phi(x), \qquad \frac{\partial u}{\partial t}(0, x) = \psi(x), & \end{cases} \tag{2.1}$$

where, for the seismic problem, u represents subsurface particle displacement resulting from a disturbance at the earth's surface ($x = 0$). We model the disturbance here using the source term $s(t; \tilde{k})$ ($s \in H^1(0, \bar{t})$) which may involve an unknown vector parameter $\tilde{k}$. Other parameters of interest include ρ and E ($\rho > 0,\ E > 0$), which represent the density and elastic modulus, respectively, of the earth, and k_1 and k_2 ($k_1 > 0,\ k_2 > 0$), which occur in the elastic boundary condition at $x = 0$ and the absorbing boundary condition at $x = 1$, respectively. A more thorough discussion of the use of model equations (2.1) in the context of the seismic problem may be found in [7]; similar models are treated in [9] or [11], where boundary conditions are imposed instead on a semi-infinite strip.

The estimation problem of interest then is to identify the unknown parameters using observations of the system, which may be given in the form of displacement, velocity, or pressure (stress) measurements (corresponding to $u,\ u_t$, or Eu_x in (2.1)). Throughout we shall assume that unknown parameters are expressed in the form

$$q = (q_1(\cdot), q_2(\cdot), q_3, q_4, q_5) \equiv \left(\frac{1}{\rho(\cdot)}, E(\cdot), k_1, k_2, \tilde{k}\right)$$

where q belongs to the admissible set Q,

$$\begin{aligned} Q = \{(q_1, q_2, q_3, q_4, q_5) \in H^0 \times H^0 \times \mathbf{R}^{2+k} \mid \\ \text{for } i = 1, 2:\ 0 < \underline{q}_i \le q_i(x) \le \overline{q}_i \text{ for almost all } x \in [0, 1] \\ \text{and } q_i \text{ is continuous at } x = 0, 1;\ 0 < \underline{q}_3 \le q_3 \le \overline{q}_3; q_4 > 0\}. \end{aligned}$$

Due to the fact that the spatial domain [0, 1] represents an inhomogeneous, layered medium, it is likely that the two spatially varying parameters, q_1 and q_2, will be discontinuous with locations of discontinuities corresponding to abrupt changes in subsurface structure. In keeping with the objectives of the seismic inverse problem, i.e., to assemble as much information as possible about the nature of the subsurface medium, our goal will be to determine the location of discontinuities, as well as to identify the spatial variation in q_1 and q_2.

Before considering a precise statement of the parameter estimation problem, we first turn to an abstract formulation of (2.1), which is necessary for the convergence theory we develop in Sec. 4. To this end, we (formally) rewrite (2.1) as a first-order system in $v(t) = (v_1(t), v_2(t), v_3(t)) \sim (u(t,0), u_t(t,\cdot), q_2(\cdot)u_x(t,\cdot)) \in \mathbf{R} \times H^1 \times H^1$, namely,

$$\begin{cases} \dot{v}(t) = \begin{pmatrix} 0 & I_0 & 0 \\ 0 & 0 & q_1 D \\ 0 & q_2 D & 0 \end{pmatrix} v(t) \\ v_3(t)|_{x=0} - q_3 q_2(0) v_1(t) = q_2(0) s(t; q_5), \\ (q_2(1) v_2(t) + q_4 v_3(t))|_{x=1} = 0, \\ v(0) = (\phi(0), \psi(\cdot), q_2(\cdot)\phi_x(\cdot))^{\mathrm{T}}, \end{cases} \tag{2.2}$$

where $I_0 : C[0,1] \to \mathbf{R}$ is defined by $I_0 w = w(0)$ and D represents the spatial differentiation operator. Such a decomposition of state variables can be found elsewhere in the literature (see, for example, [10] in a different context). It is a natural one for the application we have in mind because the components $u(t,0)$, u_t, and $q_2 u_x$ are quantities needed to express the total energy of this system. In particular, the quantity $u(t,0)$ is used in computing the energy associated with the elastic deformation of the surface, due to the source disturbance s. In addition, these quantities are often readily observable in practice and thus their inclusion as components of the state variable facilitates our study of the estimation problem.

We use the form of (2.2) to define a parameter-dependent operator $A = A(q)$ and associated Hilbert space $X = X(q)$, where $\operatorname{dom} A \subseteq X$ and X is chosen in such a way to ensure that A is dissipative in X. In particular, for any value of $q \in Q$ we define $X(q) \equiv \mathbf{R}(q_3 q_2(0)) \times H^0(q_1) \times H^0(q_2)$ (see Sec. 1 concerning notation) with associated norm denoted by $\|\cdot\|_q$. It is easily seen that a positive constant $\mu = \mu(\underline{q}, \overline{q}_i;\ i = 1,2,3)$ may be found so that

$$\frac{1}{\mu}\|z\| \le \|z\|_q \le \mu\|z\| \tag{2.3}$$

for all $q \in Q$, where $\|\cdot\|$ denotes the $\mathbf{R} \times H^0 \times H^0$ norm; it thus follows that the $X(q)$ norms are uniformly equivalent as q ranges over Q. Due to the set-wise equivalence of $X(q)$ for all $q \in Q$, we can consider an element z in $X(q)$ to be an element in $X(\tilde{q})$, for $\tilde{q} \in Q$, or even consider z in $\mathbf{R} \times H^0 \times H^0$, and do so without a change in notation for z.

The abstract differential equation associated with (2.2) may be stated precisely as

$$\begin{aligned} \dot{z}(t) &= A(q) z(t) + F(t; q), \quad t \in (0, \bar{t}], \\ z(0) &= z_0(q), \end{aligned} \tag{2.4}$$

where the transformation $(z_1, z_2, z_3) = (v_1, v_2, (v_3 + (x-1)s))$ has been made in order to obtain homogeneous boundary conditions. In (2.4), $z(t) \in X(q)$, and

$$A(q) = \begin{pmatrix} 0 & I_0 & 0 \\ 0 & 0 & q_1 D \\ 0 & q_2 D & 0 \end{pmatrix},$$

with $\operatorname{dom} A(q) = \{z \in X(q) \cap (\mathbf{R} \times H^1 \times H^1) | z_3(0) - q_2(0) q_3 z_1 = 0,\ q_2(1) z_2(1) + q_4 z_3(1) = 0\}$; the nonhomogeneous term (due to the transformation to zero boundary conditions) and initial condition are given by

$$F(t;q) = \begin{pmatrix} 0 \\ -q_1(\cdot)s(t;q_5) \\ (\cdot - 1)\dot{s}(t;q_5) \end{pmatrix}, \qquad z_0(q) = \begin{pmatrix} \phi(0) \\ \psi(\cdot) \\ q_2(\cdot)\phi_x(\cdot) + (\cdot - 1)s(0;q_5) \end{pmatrix},$$

respectively.

THEOREM 2.1. $A(q)$ is the infinitesimal generator of a C_0-semigroup of contractions $T(t;q)$ on $X(q)$; thus for each $q \in Q$, there exists a unique mild solution of (2.4) with the representation

$$z(t;q) = T(t;q)z_0(q) + \int_0^t T(t-s;q)F(s;q)\,ds. \tag{2.5}$$

The proof of the theorem follows from standard results from the theory of semigroups (see, e.g., [17; pp. 14, 106]) once it has been verified that $A(q)$ is densely defined and dissipative in $X(q)$ and that for some $\lambda > 0$, $\mathscr{R}(\lambda I - A(q)) = X(q)$. It is easy to see that $\overline{\operatorname{dom} A(q)} = X(q)$ since $X(q)$ is equivalent to $\mathbf{R} \times H^0 \times H^0$, while dissipativity of $A(q)$ is a consequence of the topology chosen for $X(q)$. Showing that $\mathscr{R}(\lambda I - A(q)) = X(q)$ is equivalent to demonstrating the existence of a solution to a particular two-point boundary value problem. This can be accomplished using standard results (e.g., [20], pp. 194–197).

REMARK 2.1. Strong solutions of (2.4) (i.e., solutions where $z_0(q) \in \operatorname{dom} A(q)$) will satisfy (2.4) as well as (2.5). In this case, $(v_1, v_2, v_3) = (z_1, z_2, z_3 + (1-\cdot)s)$ satisfies the formal equations (2.2) so that the identification may be made between z_1 and $u(t,0)$, z_2 and $u_t(t,\cdot)$, and $z_3 + (1-\cdot)s$ and $q_2(\cdot)u_x(t,\cdot)$, where u is a solution of the original system (2.1). Hence, we shall develop a convergence theory based on formulations (2.4), (2.5), while keeping these relationships in mind throughout.

REMARK 2.2. That our approach is advantageous from both a theoretical and computational viewpoint is now apparent. Physical principles demand that $q_2 u_x$ be continuous even though q_2, and thus u_x, are discontinuous. By defining our operator as above, we satisfy this continuity condition simply by requiring that the third component of the state variable belong to $H^1(0,1)$.

We turn now to a precise statement of the parameter estimation problem. The unknown parameter, q, of interest contains for its first two components the piecewise continuous coefficients q_1 and q_2; we further parametrize q_1 and q_2 by

$$\begin{aligned} q_1(x) &= \alpha_0(x) + \sum_{i=1}^{\nu} \alpha_i(x) H_{\xi_i}(x), \\ q_2(x) &= \beta_0(x) + \sum_{i=1}^{\nu} \beta_i(x) H_{\xi_i}(x), \end{aligned} \tag{2.6}$$

where H_ξ denotes the usual Heaviside function ($H_\xi(x) = 1$, $x \in (\xi, 1]$, and $H_\xi(x) = 0$ otherwise) on $[0,1]$. Using this representation, the problem of estimating q_1 and q_2 is equivalent to the problem of identifying $2(\nu+1)$ continuous "pieces", $\alpha_0(x), \ldots,$ $\alpha_\nu(x)$, $\beta_0(x), \ldots, \beta_\nu(x)$, and the location of ν discontinuities, $\xi_1, \ldots, \xi_\nu$. Given ν, our goal then is to estimate $p \equiv (\alpha_0, \alpha_1, \ldots, \alpha_\nu, \beta_0, \beta_1, \ldots, \beta_\nu, \xi_1, \ldots, \xi_\nu, q_3, q_4, q_5)$ where p belongs to a prescribed constraint set $\mathscr{P} \subseteq \tilde{\mathscr{P}}$,

$$\tilde{\mathscr{P}} = \left\{ (\alpha_0, \ldots, \alpha_\nu, \beta_0, \ldots, \beta_\nu, \xi_1, \ldots, \xi_\nu, q_3, q_4, q_5) \in \left(\prod_{i=1}^{2(\nu+1)} C[0,1] \right) \times \mathbf{R}^{\nu+2+k} \,\middle|\, (\alpha_0 + \sum_{i=1}^{\nu} \alpha_i H_{\xi_i}, \beta_0 + \sum_{i=1}^{\nu} \beta_i H_{\xi_i}, q_3, q_4, q_5) \in Q \right\}. \quad (2.7)$$

Although our estimation theory will be developed for the fully parametrized set $\mathscr{P}$, it is clear that any choice of $p \in \mathscr{P}$ generates a physically meaningful entity $q = q(p) \in Q$ through the construction (2.6) for q_1 and q_2, and that sequential convergence $p^j \to p$ in the topology for $\tilde{\mathscr{P}}$ implies the associated parameters $q^j = q^j(p^j)$, $q = q(p) \in Q$ satisfy $q^j \to q$ in the topology on Q. We shall often make use of this association in the sections to follow.

The parameter identification problem may now be stated as follows:

(ID) Minimize $J(p) = \sum_j \|\mathscr{C} z(t_j; q(p)) - y_j\|^2$ over $p \in \mathscr{P}$, where $q = q(p) \in Q$, $z(t;q)$ is the solution to (2.4), $\mathscr{C} : X(q) \to \mathbf{R} \times H^0 \times H^0$ is a (continuous) observation operator and $y_j \in \mathbf{R} \times H^0 \times H^0$ denotes observed data corresponding to time t_j.

Remark 2.3. The estimation problem defined above allows for observations of particle displacement (at the surface), velocity, or pressure. Although it is not explicitly stated, our theory also includes the case of spatially distributed displacement observations below the earth's surface; in this case, the observed quantity is given by $\tilde{z}(t_j; q) = \phi(\cdot) + \int_0^{t_j} z_2(s; q)\, ds$, and observations are given by $\tilde{y}_j$, where $\tilde{y}_j$, $\tilde{z}(t_j; q) \in H^0$.

Remark 2.4. The parameter estimation problem described above is an example of a large class of problems ("inverse" problems) that are widely known to be ill-posed (see, for example [13], [16]) from both a theoretical and computational standpoint. We shall not address here the nature of the difficulties that arise, among them a lack of continuous dependence of parameters p on observations y_j, and the general nonuniqueness of an optimal parameter $\tilde{p}$. However, it is appropriate to note that one may circumvent some of these inherent problems by guaranteeing a type of "problem stability" (see [2], [4], [11], and [12, Remark 5.1]). There are a number of ways in which such stability may be obtained for the parameter estimation problem. These include such techniques as regularization [13], embedding methods [1], or parameter set compactness assumptions [2], [5]. As will be evident in later sections, we shall take the latter approach through the assumption that $\mathscr{P}$ is compact. An example of a compact constraint set $\mathscr{P}$ may be constructed as follows. Let

$$\mathscr{P}' = \Big\{ p = (\alpha_0, \ldots, \alpha_\nu, \beta_0, \ldots, \beta_\nu, \xi_1, \ldots, \xi_\nu, q_3, q_4, q_5) \in \prod_{i=1}^{2(\nu+1)} C^1[0,1] \times \mathbf{R}^{\nu+2+k} \Big| \; |\alpha_i|_{C^1[0,1]} \le K_{\alpha,i}, \; |\beta_i|_{C^1[0,1]} \le K_{\beta,i}, \; |\xi_i| \le K_{\xi,i}, \; i = 0, 1, \ldots, \nu; \; 0 < \underline{q}_i \le q_i \le \overline{q}_i, \; i = 3, 4; \; |q_5|_{\mathbf{R}^k} \le \overline{q}_5 \Big\}$$

where $K_{\alpha,i}, K_{\beta,i}, K_{\xi,i}, \underline{q}_4, \overline{q}_4$, and $\overline{q}_5$ are positive constants, and $\underline{q}_3, \overline{q}_3$ are given in the definition of Q. This set $\mathscr{P}'$ is a compact subset of $\tilde{P}$, but we must also require that each $p \in \mathscr{P}$ can be identified with some $q \in Q$ (recall that Q is defined in terms of the physical parameters, while $\tilde{\mathscr{P}}$, for mathematical convenience, has been defined using further parametrizations). To this end, define

$$\mathscr{P} = \Big\{ p \in \mathscr{P}' | 0 < \underline{q}_1 \le \alpha_0 + \sum_{i=1}^{\nu} \alpha_i H_{\xi_i} \le \overline{q}_1 \text{ and } 0 < \underline{q}_2 \le \beta_0 + \sum_{i=1}^{\nu} \beta_i H_{\xi_i} \le \overline{q}_2 \Big\}$$

where $\underline{q}_i, \overline{q}_i, \; i = 1, 2$, are as in the definition of Q. It can be shown that $\mathscr{P}$ is a closed subset of $\mathscr{P}'$, and therefore, $\mathscr{P}$ is a compact subset of $\tilde{\mathscr{P}}$ with the desired property that any $p \in \mathscr{P}$ gives rise to a $q \in Q$.

3. Approximation. In this section we construct a spline-based framework for the approximation of the state variable $z(t)$, where z satisfies (2.5). As these approximation spaces will be parameter-dependent, we shall assume that $p \in \mathscr{P}$ is given, where $p = (\alpha_0, \alpha_1, \beta_0, \beta_1, \xi, q_3, q_4, q_5)$ and $\xi \equiv \xi_1 \in (0,1)$ (for simplicity and without loss of generality, we shall take $\nu = 1$ throughout). Using (2.6) to construct $q_1, \; q_2$, we shall as usual associate with p the corresponding $q = q(p) \in Q$.

The construction of approximation spaces $X^N(q)$ is as follows. For $k = 0, \ldots, 2N$ we define uniform spatial mesh points by $s_k^N = k/(2N)$ and denote by S_k^N the kth (standard) basis element for the space of continuous piecewise-linear B-splines with knots at $\{s_k^N\}$; i.e., S_k^N is characterized by $S_k^N(s_j^N) = \delta_{jk}, \; j, k = 0, \ldots, 2N$ [18]. We then transform to a parameter-dependent basis through the invertible mapping $g: \; [0,1] \to [0,1]$,

$$g(x) = \begin{cases} \frac{x}{2\xi}, & 0 \le x < \xi, \\ \frac{x+1-2\xi}{2(1-\xi)}, & \xi \le x \le 1, \end{cases} \tag{3.1}$$

and define $\tilde{B}_k^N(x) = S_k^N(g(x))$. Thus $\operatorname{sp}\{\tilde{B}_k^N\}$ is the space of linear B-splines with knots at $x_k^N \equiv g^{-1}(s_k^N)$ ($x_k^N = k\xi/N, \; k = 0, \ldots, N; \; x_k^N = \xi + (k-N)(1-\xi)/N, \; k = N+1, \ldots, 2N$). The computational advantages of using ξ-dependent elements will be discussed in Sec. 5. Finally, we define approximation spaces $X^N(q) \equiv \operatorname{sp}\{B_j^N\}$ where the basis elements $B_j^N, \; j = 0, \ldots, 4N$, are defined as follows (we use $\mathscr{O}$ to represent the identically zero function on $[0,1]$):

$$\begin{aligned} B_j^N &= (0, \tilde{B}_j^N, \mathscr{O}), \quad j = 0, \ldots, 2N-1; \\ B_{2N}^N &= (0, q_4 \tilde{B}_{2N}^N, -q_2(1)\tilde{B}_{2N}^N); \\ B_j^N &= (0, \mathscr{O}, \tilde{B}_{4N-j}), \quad j = 2N+1, \cdots, 4N-1; \\ B_{4N}^N &= (1, \mathscr{O}, q_3 q_2(0) \tilde{B}_0^N). \end{aligned}$$

It is clear from the above construction that $X^N(q) \subseteq \operatorname{dom} A(q)$. It is also clear that if $q \neq \hat{q}$ then $X^N(q) \neq X^N(\hat{q})$ and, in fact, $X^N(\hat{q}) \nsubseteq \operatorname{dom} A(q)$. Thus, as we iterate on q, we must take some care with the changing domains of the operators; we shall address this difficulty in the next section.

We define approximating equations associated with (2.4) by defining operators $A^N(q)$ to "approximate" $A(q)$. Here we take $A^N(q) \equiv P^N(q)A(q)P^N(q)$ where $P^N(q): X(q) \to X^N(q)$ denotes the orthogonal projection (with respect to the topology on $X(q)$ along $(X^N(q))^\perp$). The differential equation on $X^N(q)$, equivalent to a system of ordinary differential equations, is given by

$$\begin{cases} \dot{z}^N(t) = A^N(q)z^N(q) + P^N(q)F(t;q), & t \in (0, \bar{t}], \\ z^N(0) = P^N(q)z_0(q), \end{cases} \tag{3.2}$$

where $z^N(t) \in X^N(q)$. Using standard arguments (see, e.g., [7]) it is easy to show that $A^N(q)$ is bounded and generates a C_0-semigroup of contractions $T^N(t, q)$ on $X(q)$; thus, there exists a unique mild solution $z^N(\cdot; q) \in C(0, \bar{t}; X^N(q))$ of (3.2), expressed as

$$z^N(t;q) = T^N(t;q)P^N(q)z_0(q) + \int_0^t T^N(t-s;q)P^N(q)F(s;q)\,ds \tag{3.3}$$

for $t \in [0, \bar{t}]$.

The approximate identification problem associated with (3.3) is the following:

$$(\mathrm{ID}^N) \text{ Minimize } J^N(p) = \sum_j \|\mathscr{C} z^N(t_j; q(p)) - y_j\|^2$$

over $p \in \mathscr{P}$, subject to $z^N(\cdot; q(p))$ satisfying (3.3).

It is not difficult to argue, using the matrix representation for (3.2) (see Sec. 5), that, for $z \in \mathbf{R} \times H^0 \times H^0$, the mappings $p \to P^N(q(p))z$ and $p \to T^N(t; q(p))z$ are continuous, the latter uniformly in $t \in [0, \bar{t}]$. Under reasonable assumptions of continuity on the mappings $q_5 \to s(t; q_5)$ and $q_5 \to \dot{s}(t; q_5)$ (again uniformly in t) we also obtain the continuity of $p \to z_0(q(p))$, $p \to F(t; q(p))$. We thus obtain, from the continuity of $p \to J^N(p)$, the following.

THEOREM 3.1. Assume $\mathscr{P}$ is compact. Then for each N there exists a solution $\tilde{p}^N \in \mathscr{P}$ to (ID^N).

4. Convergence. In this section, we will establish that a subsequence of the parameter estimates generated by solving the approximating identification problems (ID^N) does indeed converge to a solution for the original identification problem. (We note that in practice, we and others working with similar schemes, have observed direct convergence of the estimates; in fact, under the assumption of a unique solution of (ID), full sequential convergence is guaranteed.)

We will use the theoretical framework developed in [5], and discussed in the context of the seismic problem in [7]; we refer the interested reader to these references for proofs and details, and here simply state relevant results without proof, focusing instead on the details which are new to our formulation of the problem.

Throughout this section we will assume $\mathscr{P}$ is compact and that an arbitrary sequence $\{p^N\}$ in $\mathscr{P}$ has been given with $p^N \to \tilde{p} \in \mathscr{P}$, where $p^N \equiv (\alpha_0^N, \alpha_1^N, \beta_0^N, \beta_1^N, \xi^N,$

$q_3^N, q_4^N, q_5^N)$ and $\tilde{p} = (\tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\beta}_0, \tilde{\beta}_1, \tilde{\xi}, \tilde{q}_3, \tilde{q}_4, \tilde{q}_5)$, and $\tilde{\xi}$ satisfies $\delta < \tilde{\xi} < 1 - \delta$ for some $\delta > 0$. We recall that the associated parameters $q^N = q^N(p^N)$ and $\tilde{q} = \tilde{q}(\tilde{p})$ in Q then satisfy $q^N \to \tilde{q}$ in Q (i.e., in $H^0 \times H^0 \times \mathbf{R}^{2+k}$).

An important intermediate result in our convergence theory is to show that $z^N(q^N) \to z(\tilde{q})$ in $X(\tilde{q})$, where $z^N(q^N)$ is a solution of (3.3) and $z(\tilde{q})$ is a solution of (2.5). To this end, we first demonstrate convergence, in an appropriate sense, of the semigroups. We will use the following version of the Trotter–Kato theorem (see [5]).

THEOREM 4.1. Let $(B, |\cdot|)$ and $(B^N, |\cdot|_N)$, $N = 1, 2, \ldots$, be Banach spaces and let $\pi^N: B \to B^N$ be bounded linear operators. Further assume that $T(t)$ and $T^N(t)$ are C_0-semigroups on B and B^N with infinitesimal generators $\tilde{A}$ and $\tilde{A}^N$, respectively. If

(i) $\lim_{N\to\infty} |\pi^N z|_N = |z|$ for all $z \in B$,

(ii) there exist constants M, ω independent of N such that

$$|T^N(t)|_N \leq Me^{\omega t}, \quad \text{for } t \geq 0, \text{ and}$$

(iii) there exists a set $\mathscr{D} \subset B$, $\mathscr{D} \subset \operatorname{dom} \tilde{A}$, with $\overline{(\lambda_0 - \tilde{A})\mathscr{D}} = B$ for some $\lambda_0 > 0$, such that for all $z \in \mathscr{D}$ we have

$$|\tilde{A}^N \pi^N z - \pi^N \tilde{A} z|_N \to 0 \text{ as } N \to \infty,$$

then $|T^N(t)\pi^N z - \pi^N T(t) z|_N \to 0$ as $N \to \infty$, for all $z \in B$, uniformly in t on compact intervals in $[0, \infty)$.

We shall use the above theorem with $B = X(\tilde{q})$ and $B^N = X(q^N)$. For $N = 1, 2, \ldots$, the linear operators $\pi^N: X(\tilde{q}) \to X(q^N)$ are defined by $\pi^N(z_1, z_2, z_3) = (\pi_1^N z_1, \pi_2^N z_2, \pi_3^N z_3)$, where the positive constants π_i^N are defined by

$$\pi_1^N = \tilde{q}_2(0)\tilde{q}_3/(q_2^N(0)q_3^N), \qquad \pi_2^N = q_4^N \tilde{q}_2(1)/(\tilde{q}_4 q_2^N(1)), \qquad \text{and } \pi_3^N = 1.$$

An essential property of π^N (which can be easily verified) is that π^N maps elements of $\operatorname{dom} A(\tilde{q})$ to $\operatorname{dom} A(q^N)$. It is clear that $\pi_i^N \to 1$ as $N \to \infty$ for $i = 1, 2, 3$, so that for $z \in X(\tilde{q})$,

$$\|(\pi^N - I)z\|_{q^N} \to 0 \text{ as } N \to \infty, \tag{4.1}$$

and that Theorem 4.1(i) is satisifed; further, the operators are bounded uniformly in N. We take $T(t)$ and $T^N(t)$ to be the C_0-semigroups generated by $A(\tilde{q})$, $A^N(q^N)$, respectively; condition (ii) of the theorem follows from the fact that for all N, T^N is a contraction semigroup. As a first step in the verification of (iii), we select $\mathscr{D} \equiv \operatorname{dom} A(\tilde{q}) \cap (\mathbf{R} \times W_\infty^1 \times W_\infty^1)$ and again appeal to the theory of two-point boundary value problems to claim that, for any $\lambda > 0$, $(\lambda I - A(\tilde{q}))$ maps $\mathscr{D}$ onto $\mathbf{R} \times L^\infty \times L^\infty$, which is dense in $X(\tilde{q})$. To satisfy the hypotheses of Theorem 4.1 it only remains to show that for $z \in \mathscr{D}$, we have

$$\|A^N(q^N)\pi^N z - \pi^N A(\tilde{q})z\|_{q^N} \to 0 \quad \text{as } N \to \infty.$$

It is helpful to establish several lemmas, the first of which is a generalization of standard estimates regarding the usual linear spline interpolating operator, i^N, defined on the mesh $\{x_k^N, k = 0, \ldots, 2N\}$ (i.e., $i^N f = \sum_{k=0}^{2N} f(x_k^N)\tilde{B}_k^N$, in the notation of Sec. 3). The corresponding state space interpolation operator of interest is the mapping $\mathcal{I}^N: (\mathbf{R} \times H^1 \times H^1) \to (\mathbf{R} \times H^1 \times H^1)$, where we define $\mathcal{I}^N \equiv I \times i^N \times i^N$.

LEMMA 4.1. There is a constant c_0 independent of N and q^N such that for any $z \in \operatorname{dom} A(\tilde{q})$,

$$\|\mathcal{I}^N z - z\|_{q^N} \le c_0 N^{-1} \|D(\mathcal{I}^N z - z)\|_{q^N}, \tag{4.2}$$

where we define $D(z_1, z_2, z_3) = (0, Dz_2, Dz_3)$. Further,

$$\|D(\mathcal{I}^N z - z)\|_{q^N} \to 0 \quad \text{as } N \to \infty. \tag{4.3}$$

The proof of this lemma follows easily from the dense inclusion of H^2 in H^1, and from a number of estimates ((2.10), (2.16), (2.17)) in [19]; in particular, we shall need one of these estimates in a later proof, so we state the inequality here for convenience:

$$|i^N f - f|_0 \le (\pi N)^{-1} |D(i^N f - f)|_0. \tag{4.4}$$

LEMMA 4.2. There exists a constant c_1 independent of q^N and N such that, for any $z \in X^N(q^N)$,

$$\|Dz\|_{q^N} \le c_1 N \|z\|_{q^N}, \tag{4.5}$$

for N sufficiently large.

Proof. From (2.3), $\|Dz\|_{q^N}^2 \le \mu^2 |Dz_2|_0^2 + \mu^2 |Dz_3|_0^2$ where, for $j = 2, 3$,

$$|Dz_j|_0^2 = \sum_{k=1}^{2N} \int_{x_{k-1}^N}^{x_k^N} |Dz_j|^2 \, dx \le 12 \sum_{k=1}^{2N} (x_k^N - x_{k-1}^N)^{-2} \int_{x_{k-1}^N}^{x_k^N} |z_j|^2 \, dx,$$

from the Schmidt inequality [19; p. 7]. Since $\tilde{\xi}$ satisfies $\delta < \tilde{\xi} < 1 - \delta$, it is clear that, for N sufficiently large, ξ^N satisfies $\delta \le \xi^N \le 1 - \delta$ and that $x_k^N - x_{k-1}^N \ge \delta/N$. Thus, $\|Dz\|_{q^N}^2 \le 12(\mu N/\delta)^2 (|z_2|_0^2 + |z_3|_0^2)$ and, using (2.3) again, the result in (4.5) is obtained. □

Lemma 4.2 is a generalization of the Schmidt inequality for our product space and topology. The next lemma is such a generalization of the First Integral Relation for linear splines (i.e., $|D(i^N f)|_0 \le |Df|_0$ for $f \in H^1$ [19; p. 16]); the proof in our case follows directly from that relation and the statement of equivalence of norms (2.3).

LEMMA 4.3. There is a constant c_2 independent of N and q^N such that

$$\|D(\mathcal{I}^N z)\|_{q^N} \le c_2 \|Dz\|_{q^N} \tag{4.6}$$

for any $z \in X(q)$, and any $q \in Q$.

The spline estimates in the preceding lemmas may now be used to obtain convergence of $P^N(q^N)$ to I, in an appropriate sense. To simplify notation we shall write $P^N \equiv P^N(q^N)$, where no confusion exists.

LEMMA 4.4. If $z \in \operatorname{dom} A(\tilde{q})$, then

$$\begin{aligned} \|(P^N - I)\pi^N z\|_{q^N} &\le \|\mathcal{I}^N \pi^N z - \pi^N z\|_{q^N} \\ &\le \frac{c_3}{N} (\|(\pi^N - I)Dz\|_{q^N}^2 + \|D(\mathcal{I}^N z - z)\|_{q^N}^2)^{1/2} \end{aligned} \tag{4.7}$$

for a constant c_3 independent of N, q^N, and z; moreover, for $z \in X(\tilde{q})$,

$$\|(P^N - I)\pi^N z\|_{q^N} \to 0 \quad \text{as } N \to \infty. \tag{4.8}$$

Proof. For any $z \in \operatorname{dom} A(\tilde{q})$, the definition of π^N guarantees $\pi^N z \in \operatorname{dom} A(q^N)$ and hence $\mathcal{I}^N \pi^N z \in X^N(q^N)$. Thus

$$\begin{aligned}\|(P^N - I)\pi^N z\|_{q^N}^2 &\le \|(\mathcal{I}^N - I)\pi^N z\|_{q^N}^2 \\ &\le \mu^2(|i^N(\pi^N z)_2 - (\pi^N z)_2|_0^2 + |i^N(\pi^N z)_3 - (\pi^N z)_3|_0^2).\end{aligned}$$

From (4.4), it follows that

$$|i^N(\pi^N z)_j - (\pi^N z)_j|_0^2 \le (\pi N)^{-2} |D(i^N(\pi^N z)_j - (\pi^N z)_j)|_0^2,$$

for $j = 1, 2$, where

$$\begin{aligned}&|D(i^N(\pi^N z)_j - (\pi^N z)_j)|_0 \\ &\qquad \le |D(i^N(\pi^N z)_j - i^N z_j)|_0 + |D(i^N z_j - z_j)|_0 + |D(z_j - (\pi^N z)_j)|_0 \\ &\qquad \le 2|D((\pi^N z)_j - z_j)|_0 + |D(i^N z_j - z_j)|_0.\end{aligned}$$

(We have used the First Integral Relation for linear splines in the last inequality.) Combining these estimates with (2.3), and noting that $D\pi^N z = \pi^N Dz$, we obtain (4.7). Finally, the convergence in (4.8) for arbitrary $z \in X(\tilde{q})$ is guaranteed by (4.7), the dense inclusion of $\operatorname{dom} A(\tilde{q})$ and $X(\tilde{q})$, and the uniform (in N) boundedness of P^N and π^N. □

LEMMA 4.5. If $z \in \operatorname{dom} A(\tilde{q})$, then

$$\|D(P^N - I)\pi^N z\|_{q^N} \to 0, \quad \text{as } N \to \infty. \tag{4.9}$$

Proof. We use the triangle inequality to show

$$\begin{aligned}\|D(P^N - I)\pi^N z\|_{q^N} &\le \|D(P^N - \mathcal{I}^N)\pi^N z\|_{q^N} + \|D[\mathcal{I}^N(\pi^N z - z)]\|_{q^N} \\ &\quad + \|D(\mathcal{I}^N z - z)\|_{q^N} + \|D(z - \pi^N z)\|_{q^N} \\ &\le c_1 N \|(P^N - \mathcal{I}^N)\pi^N z\|_{q^N} + (c_2 + 1)\|D(\pi^N z - z)\|_{q^N} \\ &\quad + \|D(\mathcal{I}^N z - z)\|_{q^N},\end{aligned}$$

where the last inequality is due to estimates (4.5) (applicable since $\mathcal{I}^N \pi^N z \in X^N(q^N)$) and (4.6). Thus we have

$$\begin{aligned}\|D(P^N - I)\pi^N z\|_{q^N} &\le c_1 N(\|(P^N - I)\pi^N z\|_{q^N} + \|\pi^N z - \mathcal{I}^N \pi^N z\|_{q^N}) \\ &\quad + (c_2 + 1)\|(\pi^N - I)Dz\|_{q^N} + \|D(\mathcal{I}^N z - z)\|_{q^N} \\ &\le 2c_1 c_3(\|(\pi^N - I)Dz\|_{q^N}^2 + \|D(\mathcal{I}^N z - z)\|_{q^N}^2)^{1/2} \\ &\quad + (c_2 + 1)\|(\pi^N - I)Dz\|_{q^N} + \|D(\mathcal{I}^N z - z)\|_{q^N},\end{aligned}$$

where we have used (4.7) to argue the last inequality. Therefore, the estimates in (4.1) and (4.3) may be used to argue $\|D(P^N - I)\pi^N z\|_{q^N} \to 0$ as $N \to \infty$. □

We now have the machinery at hand to complete our verification of the hypotheses of Theorem 4.1 to obtain a statement of semigroup convergence.

THEOREM 4.2. For each $z \in X(\tilde{q})$ we have

$$\|T^N(t;q^N)\pi^N z - \pi^N T(t;\tilde{q})z\|_{q^N} \to 0 \quad \text{as } N \to \infty, \text{ uniformly in } t \in [0,\bar{t}].$$

Proof. It remains to argue that

$$\|A^N(q^N)\pi^N z - \pi^N A(\tilde{q})z\|_{q^N} \to 0 \quad \text{as } N \to \infty$$

for $z \in \mathscr{D}$. We see that

$$\begin{aligned}
\|A^N(q^N)\pi^N z - \pi^N A(\tilde{q})z\|_{q^N} &= \|P^N A(q^N)P^N\pi^N z - \pi^N A(\tilde{q})z\|_{q^N} \\
&\le \|P^N A(q^N)(P^N\pi^N z - \pi^N z)\|_{q^N} \\
&\quad + \|P^N[A(q^N)\pi^N z - \pi^N A(\tilde{q})z]\|_{q^N} + \|(P^N - I)\pi^N A(\tilde{q})z\|_{q^N} \\
&\le \|A(q^N)(P^N\pi^N z - \pi^N z)\|_{q^N} + \|A(q^N)\pi^N z - \pi^N A(\tilde{q})z\|_{q^N} \\
&\quad + \|(P^N - I)\pi^N A(\tilde{q})z\|_{q^N} \\
&\equiv \varepsilon_1(N) + \varepsilon_2(N) + \varepsilon_3(N).
\end{aligned}$$

That $\varepsilon_3(N) \to 0$ as $N \to \infty$ follows directly from (4.8). We consider $\varepsilon_1(N)$ and note that

$$\begin{aligned}
[\varepsilon_1(N)]^2 &= \|((P^N\pi^N z - \pi^N z)_2(0), q_1^N D(P^N\pi^N z - \pi^N z)_3, q_2^N D(P^N\pi^N z - \pi^N z)_2)^{\mathrm{T}}\|_{q^N}^2 \\
&\le q_3^N q_2^N(0)|(P^N\pi^N z - \pi^N z)_2(0)|^2 + (\overline{q}_1\overline{q}_2)\|D(P^N\pi^N z - \pi^N z)\|_{q^N}^2
\end{aligned}$$

where convergence of the second term follows from Lemma 4.5; in addition, $|(P^N\pi^N z - \pi^N z)_2(0)|^2 \to 0$ as $N \to \infty$ because Lemmas 4.4 and 4.5 guarantee the $H^1(0,1)$ convergence of $(P^N\pi^N z - \pi^N z)_2$. Thus $\varepsilon_1(N) \to 0$ as $N \to \infty$. Regarding $\varepsilon_2(N)$, we use the equivalence of norms to write

$$\begin{aligned}
[\varepsilon_2(N)]^2 &\le \mu^2\|((\pi_2^N - \pi_1^N)z_2(0),\ (q_1^N\pi_3^N - \pi_2^N\tilde{q}_1)Dz_3,\ (q_2^N\pi_2^N - \pi_3^N\tilde{q}_2)Dz_2)^{\mathrm{T}}\|^2 \\
&\le \mu^2\{|\pi_2^N - \pi_1^N|^2|z_2(0)|^2 + |q_1^N\pi_3^N - \pi_2^N\tilde{q}_1|_0^2|Dz_3|_\infty^2 \\
&\quad + |q_2^N\pi_2^N - \pi_3^N\tilde{q}_2|_0^2|Dz_2|_\infty^2\},
\end{aligned}$$

so that, from the convergence of $\pi_i^N \to 1$ $(i = 1,2,3)$, and $q^N \to \tilde{q}$ in Q, we have $\varepsilon_2(N) \to 0$ as $N \to \infty$. □

Just as in [7], we may use the properties of π^N, T^N, and T to obtain a statement of semigroup convergence which does not involve the π^N; similarly, we can obtain an analog of (4.8), i.e., strong convergence of P^N to I for any $z \in X(\tilde{q})$. Then, under the assumption that the mapping $q_5 \to s(\cdot;q_5) : \mathbf{R}^k \to H^1(0,\bar{t})$ is continuous (so that both $q \to z_0(q)$ and $q \to F(t;q)$ are continuous), we may obtain state variable convergence via an argument based on the "variation of constants" formulation of solutions.

COROLLARY 4.1. Assume $\{q^N\} \subseteq Q$, $\tilde{q} \in Q$ satisfy $q^N \to \tilde{q}$ in Q, and let $z^N(\cdot;q^N)$ and $z(\cdot;\tilde{q})$ denote the solutions to (3.3) and (2.5), respectively. Then

$$\|z^N(t;q^N) - z(t;\tilde{q})\|_{q^N} \to 0 \quad \text{as } N \to \infty$$

uniformly in $t \in [0,\bar{t}]$.

In addition, we may use the established state variable convergence to conclude with the desired result of this section (a slight modification of [5; p. 820]).

THEOREM 4.3. Assume $\mathscr{P}$ is compact in the $\mathscr{P}$-topology (see (2.7)) and that $\tilde{p}^N$ is a solution of (ID^N) for each N. Then $\{\tilde{p}^N\}$ contains a subsequence $\{\tilde{p}^{N_k}\}$ satisfying

(i) $\tilde{p}^{N_k} \to \tilde{p} \in \mathscr{P}$,

(ii) $z^{N_k}(\cdot;\tilde{p}^{N_k}) \to z(\cdot;\tilde{p})$ (in an appropriate sense), and,

(iii) $\tilde{p}$ is a solution of the original parameter estimation problem (ID).

Our approximation theory is not complete in that we have only considered the problem of state space approximation and have not addressed the problem of estimating truly variable parameters (where the infinite-dimensional parameter space $\mathscr{P}$ must also be discretized). In keeping with the ideas of [6], [7], and [15], we shall avoid *a priori* parametrizations of $\mathscr{P}$ and take a more general approach that involves (linear) spline-based approximations for $\mathscr{P}$ that are independent of the level of state variable approximation. To this end we define $\mathscr{P}^M \equiv I^M(\mathscr{P})$ where I^M takes elements $p = (\alpha_0, \alpha_1, \beta_0, \beta_1, \xi, q_3, q_4, q_5)$ from $\mathscr{P}$ and interpolates the spatially varying components $(\alpha_0, \alpha_1, \beta_0, \beta_1)$ to a ξ-dependent mesh; that is, $I^M p = (\alpha_0^M, \alpha_1^M, \beta_0^M, \beta_1^M, \xi, q_3, q_4, q_5)$ where, e.g., for the case of α_0^M, $\alpha_0^M = \alpha_0^{M,\xi} = \sum_{k=0}^{2M} \alpha_0(x_k^M)\tilde{B}_k^M$ (the ξ-dependent knots x_k^M and linear elements $\tilde{B}_k^M$ are defined in Sec. 3). It is easy to see that I^M is continuous in the topology on $\mathscr{P}$, so that compactness on $\mathscr{P}$ guarantees compactness on $\mathscr{P}^M$.

We now review our convergence theory in light of these finite-dimensional parameter spaces. Let $J^N(\tilde{p}^{N,M}) = \min_{p\in\mathscr{P}^M} J^N(p)$. From the construction of $\mathscr{P}^M$ there exists a sequence $\{p^{N,M}\}$ in $\mathscr{P}$ such that $\tilde{p}^{N,M} = I^M p^{N,M}$; further, using compactness of $\mathscr{P}$, a subsequence $\{p^{N_j,M_k}\}$ of $\{p^{N,M}\}$ may be found so that $p^{N_j,M_k} \to \tilde{p} \in \mathscr{P}$; using properties of I^M and by making additional smoothness assumptions for $\mathscr{P}$, we may argue that $\tilde{p}^{N_j,M_k} \to \tilde{p} \in \mathscr{P}$. Simple modifications in arguments made earlier in this section may be made to obtain corresponding state variable convergence as $N_j, M_k \to \infty$. We thus obtain an analog to Theorem 4.3 where a subsequence of solutions $\tilde{p}^{N,M}$ to the problem of minimizing J^N over $\mathscr{P}^M$ converges to a solution $\tilde{p}$ to the original problem (ID).

5. Numerical implementation.

In this section we discuss some aspects of the computational algorithm for solving the approximate parameter estimation problem (ID^N) over the finite-dimensional parameter space $\mathscr{P}^M$. To simplify the presentation we assume that $\nu = 1$ and that $q_1 = \frac{1}{\rho} = 1$ is fixed so that an arbitrary parameter $p \in \mathscr{P}$ has the form $p = (\beta_0, \beta_1, \xi, q_3, q_4, q_5)$ with corresponding parameter in Q, $q = (q_2, q_3, q_4, q_5)$, $q_2(x) = \beta_0(x) + \beta_1(x)H_\xi(x)$. In order to more easily describe the numerical algorithm in the case of truly variable q_2, it will be more convenient to temporarily work with the parameter $\frac{1}{q_2}$ instead of q_2. To this end, we note that $\frac{1}{q_2}$ may be written

$$\frac{1}{q_2}(x) = \gamma_0(x) + \gamma_1(x)H_\xi(x),$$

where $\gamma_0 = 1/\beta_0$, $\gamma_1 = 1/(\beta_0+\beta_1) - \gamma_0$; in practice (when we search for a nonconstant q_2), the $C[0,1]$ functions γ_0 and γ_1 will be estimated in place of β_0 and β_1 so we shall use the notation $\hat{p} = (\gamma_0, \gamma_1, \xi, q_3, q_4, q_5)$ to designate the unknown parameter of interest and constrain $\hat{p}$ to belong to the usual parameter set $\mathscr{P}$. Approximations

$\hat{p}^M$ to $\hat{p}$, $\hat{p}^M = (\gamma_0^M, \gamma_1^M, \xi, q_3, q_4, q_5) \in \mathscr{P}^M$ will be constructed as in Sec. 4, i.e., we express γ_0^M, γ_1^M in terms of ξ-dependent linear spline elements $\tilde{B}_k^M$ by writing $\gamma_i^M = \sum_{k=0}^{2M} b_{i,k}^M \tilde{B}_k^M$ for $i = 0, 1$. The approximation to $\frac{1}{q_2}$ then becomes $\left(\frac{1}{q_2}\right)^M \equiv \gamma_0^M + \gamma_1^M H_\xi$, or,

$$\left(\frac{1}{q_2}\right)^M (x) = \begin{cases} \sum_{k=0}^{2M} b_{0,k}^M \tilde{B}_k^M(x), & x \in [0, \xi] \\ \sum_{k=0}^{2M} (b_{0,k}^M + b_{1,k}^M) \tilde{B}_k^M(x), & x \in (\xi, 1], \end{cases}$$

where ξ, $b_{0,k}^M$, and $b_{1,k}^M$ $(k = 0, \dots, 2M)$ are unknown and to be determined. (We remark that the use of γ_0, γ_1 in place of β_0, β_1 does not change our convergence findings due to the fact that the convergence of $\hat{p}^M$ to $\hat{p}$ in $\hat{\mathscr{P}}$ yields the needed convergence of p^M to p in $\mathscr{P}$, where now $p^M \equiv (1/\gamma_0^M, -\gamma_1^M/((\gamma_0^M + \gamma_1^M)\gamma_0^M), \xi, q_3, q_4, q_5) \in \mathscr{P}$.)

For given values of N and parameters, a solution z^N of approximating system (3.2) may be written $z^N(t;q) = \sum_{j=0}^{4N} w_j^N(t;q) B_j^N$ where $w^N = \operatorname{col}(w_j^N)$ satisfies

$$\begin{cases} Q^N \dot{w}^N(t) = K^N(q) w^N(t) + F^N(t;q), & t \in (0, \bar{t}], \\ Q^N w^N(0) = w_0^N. \end{cases} \tag{5.1}$$

Here $(F^N(t;q))_i = \langle F(t;q), B_i^N \rangle_q$ and $(w_0^N)_i = \langle z_0, B_i^N \rangle_q$, for $i = 0, \dots, 4N$, while the (i, j)-elements of the $(4N + 1)$-square matrices Q^N and K^N are given by

$$\begin{aligned} Q_{ij}^N &= \langle B_j^N, B_i^N \rangle_q \\ &= q_3 q_2(0) (B_j^N)_1 (B_i^N)_1 + \int_0^1 (B_j^N)_2 (B_i^N)_2 + \int_0^1 \frac{1}{q_2} (B_j^N)_3 (B_i^N)_3, \\ K_{ij}^N &= \langle A(q) B_j^N, B_i^N \rangle_q \\ &= q_3 q_2(0) (B_j^N)_2(0) (B_i^N)_1 + \int_0^1 D(B_j^N)_3 (B_i^N)_2 + \int_0^1 \frac{1}{q_2} q_2 D(B_j^N)_2 (B_i^N)_3 \end{aligned}$$

(throughout $\langle \cdot, \cdot \rangle_q$ denotes the $X(q)$ inner product). For a typical i and j, the last term in Q_{ij}^N satisfies (for some $0 \le l, m \le 2N$),

$$\begin{aligned} \int_0^1 \frac{1}{q_2} (B_j^N)_3 (B_i^N)_3 &= \int_0^1 \frac{1}{q_2} \tilde{B}_l^N \tilde{B}_m^N \\ &= \sum_{k=0}^{2M} b_{0,k}^M \int_0^\xi \tilde{B}_k^M \tilde{B}_l^N \tilde{B}_m^N + \sum_{k=0}^{2M} (b_{0,k}^M + b_{1,k}^M) \int_\xi^1 \tilde{B}_k^M \tilde{B}_l^N \tilde{B}_m^N, \end{aligned}$$

where the parametrization for approximations to $1/q_2$ has been used. If, in contrast to the approach we take here, the spline elements $\tilde{B}_k^M$, $\tilde{B}_l^N$, $\tilde{B}_m^N$ are defined on a *uniform* mesh, the appearance of ξ in the range of integration for the above integrals leads to a large amount of computational work. Such quadratures must be recomputed *each time* that ξ changes (i.e., every time the parameter p^M is updated in an iterative scheme to minimize J^N over $\mathscr{P}^M$). An advantage (for scalar, as opposed to vector, computers) of our formulation is that we need only compute such quadratures once. That is, we may use the coordinate transformation g in (3.1) to write

$$\begin{aligned} \int_0^\xi \tilde{B}_k^M \tilde{B}_l^N \tilde{B}_m^N &= 2\xi \int_0^{1/2} S_k^M S_l^N S_m^N, \\ \int_\xi^1 \tilde{B}_k^M \tilde{B}_l^N \tilde{B}_m^N &= 2(1 - \xi) \int_{1/2}^1 S_k^M S_l^N S_m^N, \end{aligned}$$

where $\int_0^{1/2} S_k^M S_l^N S_m^N$, $\int_{1/2}^1 S_k^M S_l^N S_m^N$ need only be evaluated once, at the outset, then stored for recall during iteration when the coefficients 2ξ and $2(1-\xi)$ are updated. This savings in computational effort, realized on traditional machines where data transfer costs are small relative to costs of computationally intensive algorithms, is even more substantial in the case of multiple discontinuities $\xi_1, \ldots, \xi_\nu$, and in the case of unknown q_1.

We consider now several numerical examples which illustrate the ideas presented thus far. In the examples that follow we return to the use of standard notation (β_i will be used instead of γ_i) and we assume that all parameters are known except for β_0, β_1, and ξ, so that only $q_2 = \beta_0 + H_\xi \beta_1$ is to be determined. The reader is referred to [7] for examples with constant parameters in the source term $s(t; q_5)$ and in boundary conditions, and to [15] for multiple discontinuity examples in the context of parabolic systems. Indeed, both of these references provide a more complete numerical study than we present here. In [7], more realistic seismic examples are considered, as is the problem of surface observations (at $x = 0$) only, while in [15], examples are given to illustrate that the assumption that the *number* of discontinuities is known *a priori* is not unnecessarily restrictive (one may both overestimate and underestimate the number and still get useful information).

In the examples that follow, the "true" parameter, $\tilde{p} = (\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\xi})$, is known and is used for comparison with our approximations, $p^{N,M} = (\beta_0^{N,M}, \beta_1^{N,M}, \xi^{N,M})$. Initial data for (2.1) is given, as are both velocity and pressure observations (u_t and $\tilde{q}_2 u_x$ are determined from either analytic or finite difference solutions for (2.1), with random noise added in some cases) at discrete time and spatial locations. For our numerical examples we use a "pointwise" fit-to-data criterion,

$$\tilde{J}^N(p) = \sum_{i,j} |Cz^N(t_j; q(p))|_{x=x_i} - y_{ij}|^2, \tag{5.2}$$

where $C(z_1, z_2, z_3)^{\mathrm{T}} = (0, z_2, z_3)^{\mathrm{T}}$ and $y_{ij} = (0, u_t(t_j, x_i), (\tilde{q}_2 u_x)(t_j, x_i))^{\mathrm{T}}$; we note that in using $\tilde{J}^N$ (instead of the distributed criterion J^N defined in Sec. 3) for our examples, we are actually illustrating stronger convergence of $z^N(t)$ to $z(t)$ than is guaranteed in Sec. 4 (where only $\mathbf{R} \times H^0 \times H^0$ convergence is found). We thus exploit the fact that in practice $\mathbf{R} \times C \times C$ convergence is observed, and only present examples that use $\tilde{J}^N$ (numerous examples with distributed criterion such as J^N exist in the literature; e.g. [14] for an elliptic problem).

We initiate the parameter estimation process by supplying an initial guess $p^0 = (\beta_0^0, \beta_1^0, \xi^0)$ to IMSL's minimization routine ZXSSQ (a Levenberg–Marquardt algorithm) which is used in the numerical minimization of $\tilde{J}^N$. For each updated value of p, the Nth approximating system (5.1) is solved using IMSL's DGEAR, an ordinary differential equation solver. All calculations were performed on the IBM 3081D at Southern Methodist University.

Our first example is one in which an analytic solution u is available. The construction of initial data and forcing function is somewhat artificial and serves only to guarantee an analytic solution; nevertheless, this example is instructive in that it is the only one in which data $Cz^N(t)$ for the approximate solution is compared to *exact* observations $(Cz(t))$. In later examples, random noise is added, or finite differences are used to solve for $u(t)$ and to construct observations $Cz(t)$.

Example 5.1. We consider the problem of estimating the piecewise constant parameter q_2 with true value

$$\tilde{q}_2 = \begin{cases} 5, & x \in [0, .4] \\ 10, & x \in (.4, 1] \end{cases}$$

(thus $\tilde{\beta}_0 = 5$, $(\tilde{\beta}_0 + \tilde{\beta}_1) = 10$, $\tilde{\xi} = .4$). Observed data is calculated using the true solution

$$u(t, x) = \begin{cases} \frac{1}{7}(-102.5x^2 + 96x + 48) + 12.5tx^2, & x \in [0, .4], \\ .05(-31.25x^2 + 45x + 187) - \frac{5}{6}t(5x^2 - 10x + .8), & x \in (.4, 1], \end{cases}$$

and is available at $t_j = .5,\ 1.0,\ 1.5,\ 2.0$, and $x_i = .1i,\ i = 0, 1, \ldots, 10$. The actual model system we use here is a nonhomogeneous version of (2.1) (with $\rho = 1$), i.e.,

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(E(x)\frac{\partial u}{\partial x}\right) + f(t, x), \quad t \in (0, \bar{t}],\ x \in [0, 1],$$

where boundary and initial conditions are given in (2.1). For this nonhomogeneous system we take $f = u_{tt} - \tilde{q}_2 u_{xx}$, u and $\tilde{q}_2$ given above, and use $q_3 = 2$, $q_4 = 4$, and $s \equiv 0$ to construct the boundary conditions. Initial data and forcing function for equation (3.2) are computed using u, $\tilde{q}_2$, and f, i.e., $z_0(x) = (u(0, 0), u_t(0, x), \tilde{q}_2(x)u_x(0, x))$ and $F(t, x) = (0, f(t, x), 0)$. We search for constant values of β_0 and β_1, using the initial guess

$$q_2^0 = \begin{cases} .001, & x \in [0, .8] \\ .001, & x \in (.8, 1] \end{cases}$$

for $N = 2$; we then use the converged values of q_2^N to begin the iteration on q_2 for the next value of N. For this reason, CP time actually decreases as N increases, as may be seen in Table 5.1 where our results are summarized.

Example 5.1.a. We repeat Example 5.1 but add Gaussian noise to our observations at a level of 5% relative error. For example, if $\overline{u}_t \equiv$ average value of $u_t(t_j, x_i)$ over all i, j, then the new data for u_t is $(\hat{u}_t)_{ij} = u_t(t_j, x_i) + r_{ij}$, where r_{ij} falls in the range $[-.05\overline{u}_t, .05\overline{u}_t]$ with 99% certainty. These findings are reported in Table 5.1.a.

Example 5.1.b. We repeat Example 5.1 but add 10% relative noise to the observed data; the results may be found in Table 5.1.b.

Example 5.2. We present here an example that is a modification/rescaling (in order to include our boundary conditions) of an example found in [8; p. 381]. For this example, the true value of q_2 is given by

$$\tilde{q}_2 = \begin{cases} 6.25, & x \in [0, \frac{1}{2}], \\ 36., & x \in (\frac{1}{2}, 1], \end{cases}$$

(so $(\tilde{\beta}_0, \tilde{\beta}_0 + \tilde{\beta}_1, \tilde{\xi}) = (6.25, 36., .5))$ while parameters $q_3 = 1$, $q_4 = 6$ are held fixed. In keeping with [8], we use $z_0 = (\phi(0), \psi, \tilde{q}_2\phi_x)$, where $\phi(x) = \exp[-160(2x - .5)^2]$ and $\psi(x) = 1600(2x - .5)\exp[-160(2x - .5)^2]$, for initial data, and set $F \equiv 0$. Observations at times $t_j = .05,\ .1,\ .15,\ .2$, and spatial locations $x_i = .05i$, $i = 0, \ldots, 20$, are determined by solving (2.1) for u using a finite difference scheme (with $\tilde{q}_2$, q_3, q_4, ϕ, and ψ as above; $s = 0$). We note that this finite difference solution for $t \in [0, .2]$ matches the graph in [8; Fig. 2] for $t \in [0, 2.0]$ (our problem has been rescaled), and that we obtain similar wave reflection/transmission behavior at the interface (see Fig. 5.1).

Table 5.1. Example 5.1

N	β_0^N	$(\beta_0+\beta_1)^N$	ξ^N	$\tilde{J}^N$	CP time (secs)
(init)	(0.001)	(0.001)	(.8000)		
2	4.942	8.051	.4069	109.05	20.
4	4.959	8.939	.4034	24.21	67.
8	4.974	9.382	.4019	5.21	410.
16	4.987	9.719	.4009	1.27	1309.
32	4.991	9.846	.4005	0.31	1141.
(true)	(5.000)	(10.000)	(.4000)		

Table 5.1.a. Example 5.1.a (5% Noise Level)

N	β_0^N	$(\beta_0+\beta_1)^N$	ξ^N	$\tilde{J}^N$	CP time (secs)
2	4.952	8.064	.4075	109.64	20.
4	4.981	8.999	.4042	30.06	65.
8	4.953	8.934	.4042	24.87	203.
16	4.959	9.124	.4039	14.88	638.
32	5.001	9.877	.4010	7.71	1901.

Table 5.1.b. Example 5.1.b (10% Noise Level)

N	β_0^N	$(\beta_0+\beta_1)^N$	ξ^N	$\tilde{J}^N$	CP time (secs)
2	4.986	8.150	.4092	202.39	19.
4	4.957	8.001	.4098	194.80	61.
8	4.974	8.703	.4071	153.74	185.
16	4.974	8.704	.4071	133.75	300.
32	5.072	8.640	.4280	123.16	659.

To estimate constant values of β_0 and β_1, we use an initial guess of

$$q_2^0 = \begin{cases} 15., & x \in [0, .4], \\ 15., & x \in (.4, 1.], \end{cases}$$

for $N = 4$ and use previous converged values to begin the $N = 8,\ 16,\ 32,\ 40$ iterations. Our findings are given in Table 5.2 below.

Table 5.2. Example 5.2

N	β_0^N	$(\beta_0+\beta_1)^N$	ξ^N	$\tilde{J}^N$	CP time (secs)
(init)	(15.00)	(15.00)	(.4000)		
4	13.78	16.92	.4667	158.09	14.
8	9.78	23.09	.5105	147.85	174.
16	5.95	32.10	.4641	60.07	130.
32	6.30	35.93	.5000	1.21	260.
40	6.26	35.95	.5004	0.44	245.
(true)	(6.25)	(36.00)	(.5000)		

Example 5.3. Finally, we consider an example with truly spatially varying q_2,

$$\tilde{q}_2 = \begin{cases} 2.5x + 1.5, & x \in [0, .6], \\ 15.625x^2 - 24.375x + 9.75, & x \in (.6, 1], \end{cases}$$

(so that true $\tilde{\xi} = .6$). We fixed $q_3 = 1$, $q_4 = 1$, initial data $z_0 = (\phi(0), \psi, \tilde{q}\phi_x)$, where $\phi(x) = e^x$ and $\psi(x) = -e^x$, and set $F \equiv 0$ (i.e., $s \equiv 0$). Velocity and pressure observations were obtained by solving (2.1) (with the above parameter values and initial data) using finite difference techniques; observations were available at $t_j = .5,\ 1.0,\ 1.5,\ 2.0$, and $x_i = .05i,\ i = 0, 1, \ldots, 20$. The results of the parameter estimation process are illustrated in Figures 5.2–5.4 where we compare the graphs of $\frac{1}{\tilde{q}_2}$ with those of the converged parameter $(\frac{1}{q_2})^{N,M}$ (we recall that $\frac{1}{q_2}$ is used as the parameter in case of truly variable q_2). Figure 5.2 depicts the outcome of the estimation process when an initial guess of

$$q_2^0 = \begin{cases} 2., & x \in [0, .7] \\ 2., & x \in (.7, 1] \end{cases}$$

is used, and search is made in the space $\mathscr{P}^M$, $M = 3$. Figures 5.3 and 5.4 show results for $M = 4$ and $M = 5$, respectively, where the initial guess for $M = 4$ ($M = 5$) is the $M = 4$ ($M = 5$) linear interpolation of the converged value of $(1/q_2)^{N,M}$ for $M = 3$ ($M = 4$) and $N = 40$.

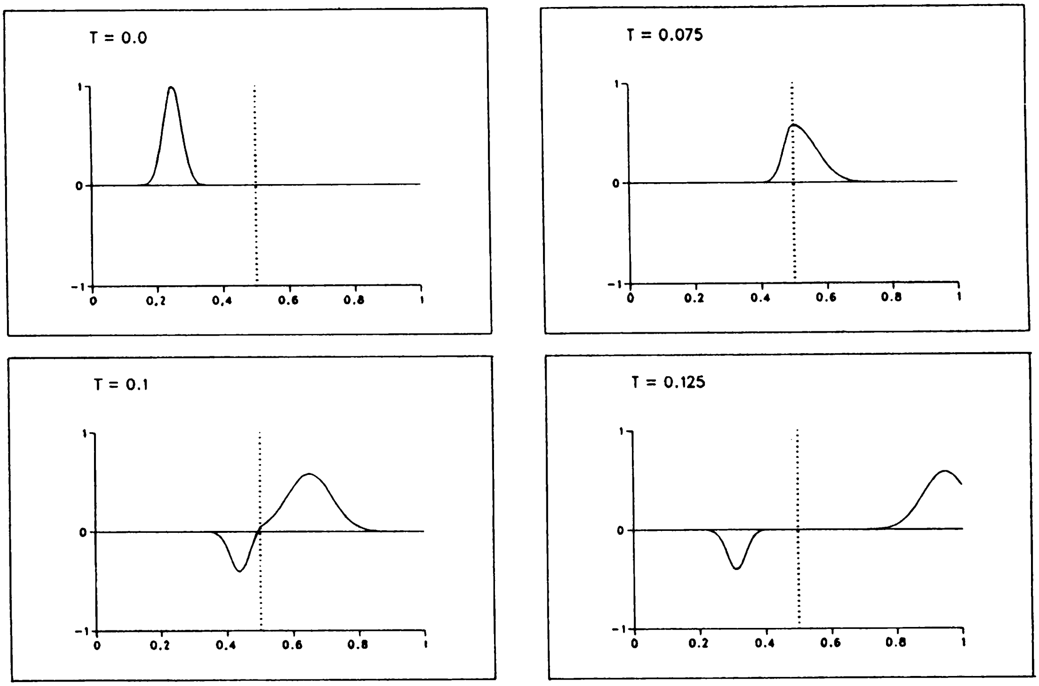

Fig. 5.1. $u(t, \cdot; \hat{q})$, $t = 0.0,\ 0.075,\ 0.1,\ 0.125$

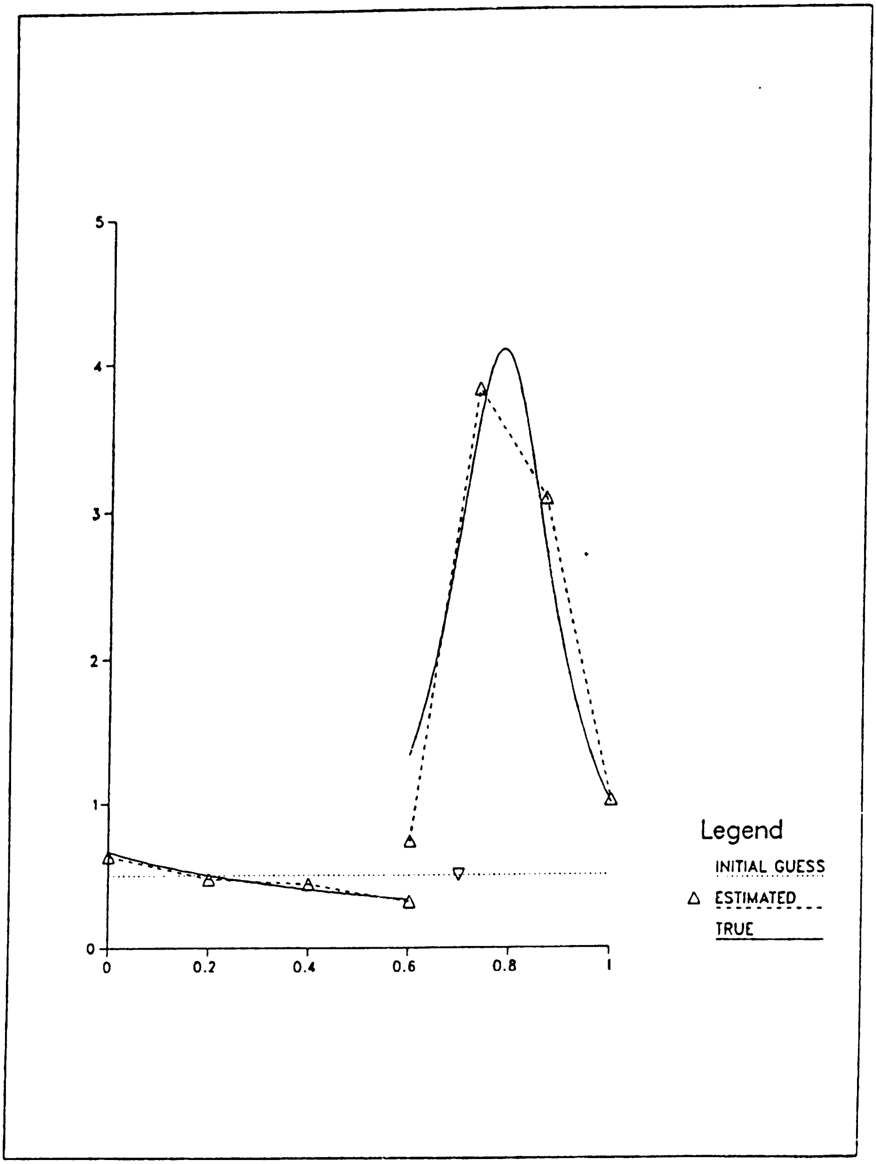

Fig. 5.2. Graphs of $\frac{1}{\bar{q}_2}$ and $\left(\frac{1}{q_2}\right)^{N,M}$, $N = 40$, $M = 3$.

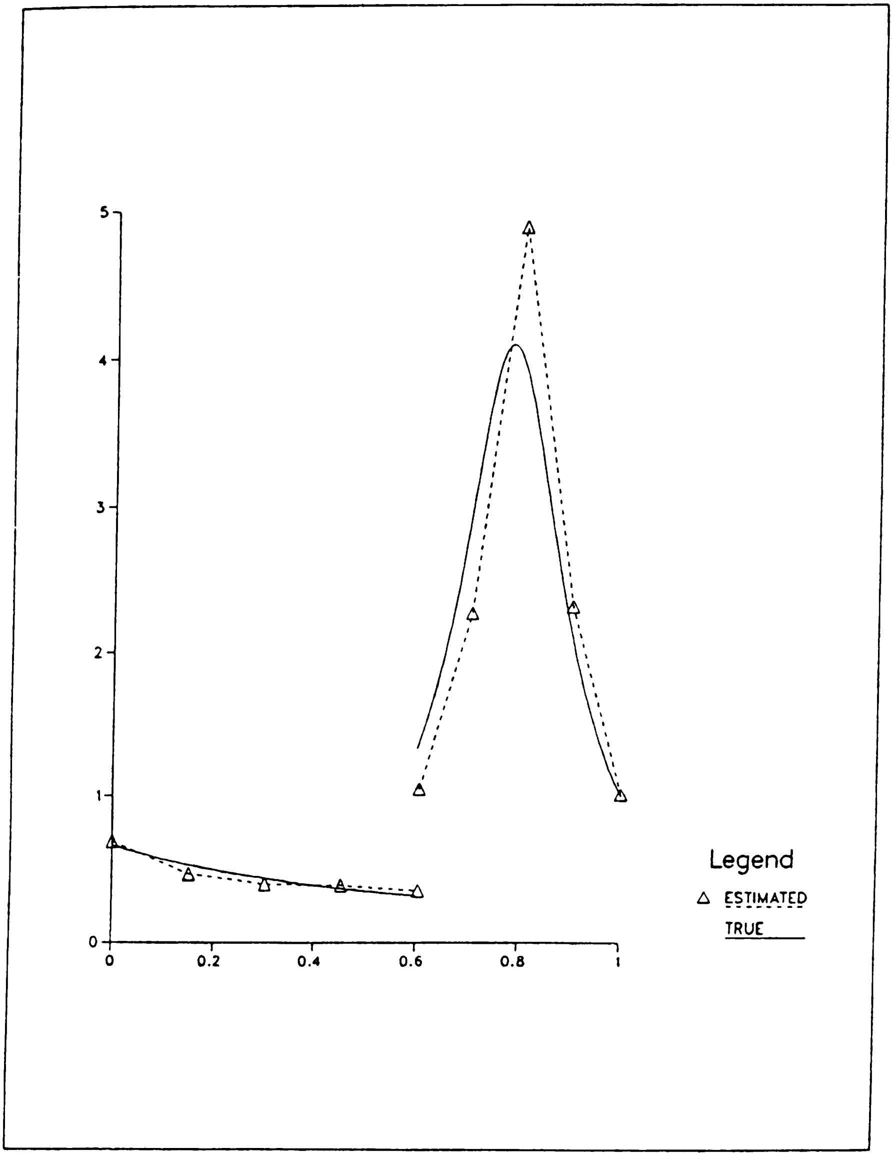

Fig. 5.3. Graphs of $\frac{1}{\bar{q}_2}$ and $\left(\frac{1}{q_2}\right)^{N,M}$, $N = 40$, $M = 4$.

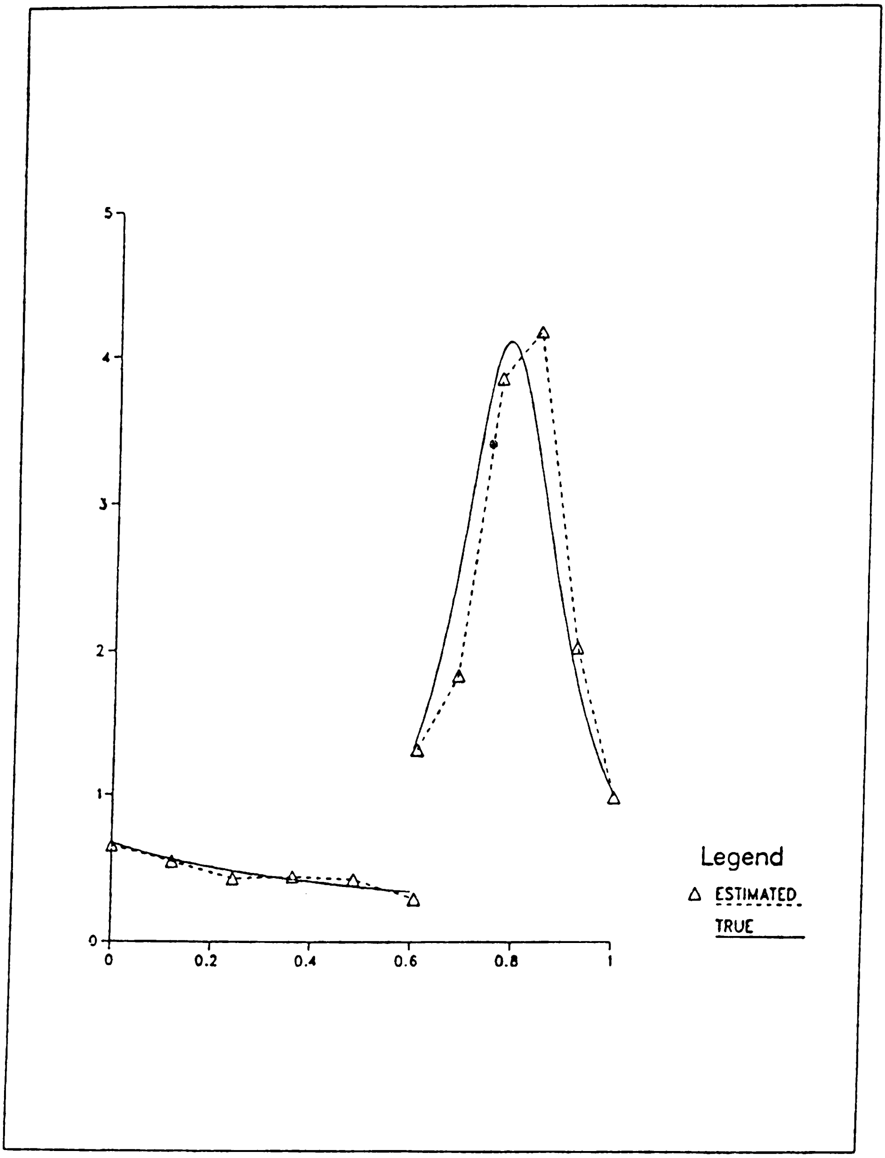

Fig. 5.4. Graphs of $\frac{1}{q_2}$ and $\left(\frac{1}{q_2}\right)^{N,M}$, $N = 40$, $M = 5$.

References

[1] H. W. Alt, K.-H. Hoffman, and J. Sprekels, *A numerical procedure to solve certain identification problems*, Int. Ser. Numer. Math. **68**, 11–43 (1984)

[2] H. T. Banks, *On a variational approach to some parameter estimation problems*, ICASE Rep. No. 85–32, ICASE, M.S. 132C, NASA Langley Research Center, Hampton, VA 23665, June 1985

[3] H. T. Banks, J. M. Crowley, and K. Kunisch, *Cubic spline approximation techniques for parameter estimation in distributed systems*, IEEE Trans. Auto. Control **28**, 773–786 (1983)

[4] H. T. Banks, P. M. Kareiva, P. K. Lamm, *Modeling insect dispersal and estimating parameters when mark-release techniques may cause initial disturbances*, J. Math. Biology **22**, 259–277 (1985)

[5] H. T. Banks and K. Kunisch, *An approximation theory for nonlinear partial differential equations with applications to identification and control*, SIAM J. Control and Optimization **20**, 815–849 (1982)

[6] H. T. Banks and Patricia Daniel Lamm, *Estimation of variable coefficients in parabolic distributed systems*, IEEE Trans. Auto. Control. **30**, 386–398 (1985)

[7] H. T. Banks and K. A. Murphy, *Estimation of coefficients and boundary parameters in hyperbolic systems*, LCDS Rep. 84-5, February 1984, Brown University; SIAM J. Control and Optimization **24**, 926–950 (1986)

[8] D. L. Brown, *A note on the numerical solution of the wave equation with piecewise smooth coefficients*, Math. Comp. **42**, 369–391 (1984)

[9] K. P. Bube and R. Burridge, *The one-dimensional inverse problem of reflection seismology*, SIAM Review **25**, 497–559 (1983)

[10] J. A. Burns and E. M. Cliff, *An approximation technique for the control and identification of hybrid systems*, Dyn. & Control of Large Flexible Spacecraft, 3rd VPI & SU/AIAA Symposium, pp. 269–284, 1981

[11] G. Chavent, *About the stability of the optimal control solution of inverse problems*, Inverse and Improperly Posed Problems in Differential Equations, G. Anger, editor, Akademie-Verlag, Berlin, pp. 45–58, 1979

[12] F. Colonius and K. Kunisch, *Stability for parameter estimation in two point boundary value problems*, Inst. für Math. Bericht No. 50–1984, Tech. Univ. Graz, October, 1984

[13] C. Kravaris and J. Seinfeld, *Identification of parameters in distributed parameter systems*, SIAM J. Control and Optimization, to appear

[14] K. Kunisch and L. W. White, *Identification under approximation for an elliptic boundary value problem*, April, 1985; SIAM J. Control and Optimization, submitted

[15] P. K. Lamm, *Estimation of discontinuous coefficients in parabolic systems: applications to reservoir simulation*, Jan. 1984; SIAM J. Control and Optimization **25**, 18–37 (1987)

[16] J. L. Lions, *Some aspects of modeling problems in distributed parameter systems*, Proc. IFIP Working Conf. (Rome 1976) Ruberti, ed., Lecture notes in Control and Info. **1**, pp. 11–41, Springer-Verlag, Berlin, 1978

[17] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences **44**, Springer-Verlag, N. Y., 1983

[18] P. M. Prenter, *Splines and variational methods*, John Wiley & Sons, N.Y., 1975

[19] M. H. Schultz, *Spline analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1973

[20] I. Stakgold, *Green's Functions and Boundary Value Problems*, John Wiley & Sons, 1979