QUARTERLY OF APPLIED MATHEMATICS
VOLUME XLIV, NUMBER 3
OCTOBER 1986, PAGES 475-486

PARAMETER IDENTIFIABILITY UNDER APPROXIMATION*

By

K. KUNISCH

(*Technische Universität Graz and University of Oklahoma*)

AND

L. W. WHITE (*University of Oklahoma*)

Abstract. The problem of injectivity of the parameter-to-state map is discussed for Galerkin approximations of a linear parabolic equation. A necessary and sufficient condition is derived and illustrated by means of simple examples. Finally, output least squares identifiability of the Galerkin approximations is discussed.

1. Introduction. In recent years an increased amount of research has been directed toward parameter identification in distributed systems, see for instance [1–3, 7–11]. The investigations have been mainly concerned with two problems: parameter estimation and parameter identifiability. To briefly discuss these concepts, let us assume that a physical system follows the simple linear equation

$$\begin{gathered}(u_m)_t = (a_m(u_m)_x)_x, \quad \text{with } u_m = u_m(t,x), 0 < t, 0 \leqslant x \leqslant 1,\\ u_m(t,0) = u_m(t,1) = 0, \\ u_m(0,x) = \varphi(x),\end{gathered} \tag{1.1}$$

where $\varphi \in L^2(0,1)$. Assume that some part Cu_m of the state u_m is observable by the investigator. He is to reconstruct a_m from this observation by choosing it from a class of admissible coefficients

$$\mathscr{A}_{\text{ad}} = \{a \in L^\infty(0,1)\colon a(x) \geqslant \tilde{a} > 0, \text{for almost every } x \in [0,1]\}.$$

Here C is a mapping from the space of trajectories $C(0,\infty;\ L^2(0,1))$ to the observation space Z. The problem of *parameter identifiability* is concerned with the injectivity of the mapping $a \to Cu(t,x;\ a)$, where $u(t,x;\ a)$ is the solution of (1.1) with a_m replaced by a. To actually determine or approximate the value a_m from the observation, one considers the optimization problem

$$(P) \qquad \min_{a \in \mathscr{A}_{\text{ad}}} |Cu(t,x;\ a) - \hat{y}|^2, \quad \text{subject to } u(t,x;\ a) \text{ satisfying (1.1).}$$

*Received March 29, 1985.

Here $\hat{y}$ stands for the observation of the system that is modeled, and in an idealized situation one has $\hat{y} = Cu(t, x;\ a_m)$. (This would require that one has chosen an exact mathematical model for the physical system and a correct choice for a_m will precisely produce $\hat{y}$.) Since the state space of Eq. (1.1) and the coefficient space are infinite-dimensional, problem (P) is infinite-dimensional itself. Its approximation is called the *parameter estimation problem.* If H^N is a sequence of linear subspaces of $H_0^1(0,1)$ and $\mathscr{A}^M$ a sequence of linear subspaces of $L^\infty(0,1)$, then one considers the sequence of problems

$$(P^{N,M}) \qquad \min_{a^M \in \mathscr{A}^M} \left|Cu^N(t;\ a^M) - \hat{y}\right|^2, \quad \text{subject to } u^N \text{ satisfying}$$

$$\frac{d}{dt}\left(u^N(t;\ a^M), v^N\right) = -\left(a^M u_x^N(t;\ a^M), v_x^N\right), \tag{1.2}$$

$$\left(u^N(0;\ a^M), v^N\right) = \left(\varphi, v^N\right), \quad \text{for all } v^N \in H^N.$$

Here $a^M \in \mathscr{A}^M$ is some appropriately chosen approximation to $a \in \mathscr{A}_{\text{ad}}$. Several studies have been conducted to show that under appropriate additional assumptions convergence of $(P^{N,M})$ to (P) holds in the sense that solutions $\bar{a}^{N,M}$ of $(P^{N,M})$ converge to a solution $\bar{a}$ of (P) and that the associated trajectories converge in appropriate topologies. This alone cannot imply that $\bar{a} = a_m$, with a_m the "true model parameter." This is partially due to model and observation error [4] and partially due to the fact that solutions of (P) and $(P^{N,M})$ are in general not guaranteed to be unique or to depend continuously on the observation $\hat{y}$. For the original (infinite-dimensional) optimization problem this has been phrased precisely in [4], and it is shown that *parameter identifiability* is a necessary condition to argue continuous dependence of the solution of the infinite-dimensional optimization problem on the observations. Similar investigations for the approximating problems $(P^{N,M})$ are not available. In the present paper we discuss the injectivity of the mapping $a^M \to u^N(t;\ a^M)$. Even though we take the simplest case where C is the identity so that all of the state is assumed to be observable, the study of "identifiability under approximation" will be found to be quite involved when specific choices for H^N or A^M are made.

In Sec. 2 we formally define identifiability under approximation and prove a necessary and sufficient condition for the present model case. Subsequently we study some aspects of output least squares identifiability. Sec. 3 is devoted to two examples. Finally, a discussion of our results is contained in Sec. 4.

2. Parameter identifiability under approximation. We now fix a basis for the subspaces H^N and $\mathscr{A}^M$ that were introduced in Sec. 1, so that $H^N = \text{span}\{B_i^N\colon i = 1, \ldots, i_N\}$ with $i_N = \dim H^N$ and $a \in \mathscr{A}_{\text{ad}}$ in $\mathscr{A}^M$ must be of the form $a^M = \sum_{j=1}^{j_M} (a^M)_j \varphi_j^M$ and the solution $u^{N,M}(t;\ a^M) = \sum_{i=1}^{i_N} \gamma_i(t) B_i^N$ of (1.2) is characterized by

$$\frac{d}{dt}\left(u^N(t;\ a^M), B_i^N\right) = \left(a^M u_x^N(t;\ a^M), \left(B_i^N\right)_x\right) \quad \text{for } t > 0, \tag{2.1}$$

$$\left(u^N(0), B_i^N\right) = \left(\varphi, B_i^N\right) \quad \text{for all } i = 1, \ldots, i_N.$$

Here $(\cdot, \cdot)$ denotes the inner product in $L^2(0,1)$. Of course, γ_i depends on N and M, but it is convenient to drop these indices. The coordinate vector of a^M is denoted by (a^M). Let $u^{N,M}(t;\ a_m^M) = \sum_{i=1}^{i_N}(\gamma_m)_i(t)B_i^N$ denote the solution of (2.1) with a^M replaced by $a_m^M = \sum_{j=1}^{i_M}(a_m)_j^M\varphi_j^M$, where a_m^M is some appropriately defined approximation to a_m in $\mathscr{A}^M$. We note that for a and a_m in $\mathscr{A}_{\text{ad}}$, a^M and a_m^M might not be in $\mathscr{A}_{\text{ad}}$; for example, if $\mathscr{A}^M$ are subspaces of $L^\infty(0,1)$ of linear spline functions, and a^M is defined as the projection P_a^M [with respect to the $L^2(0,1)$-norm] of $a \in \mathscr{A}_{\text{ad}}$ onto $\mathscr{A}^M$, then a^M might not be in $\mathscr{A}_{\text{ad}}$ for all M, although $\lim_{M\to\infty} P^M a \in \mathscr{A}_{\text{ad}}$ in $L^\infty(0,1)$[1]. However, it is reasonable to assume that the approximations a^M to $a \in \mathscr{A}_{\text{ad}}$ satisfy for all M sufficiently large

$$a^M \in \mathscr{A}_{\text{ad}}^* = \left\{a \in L^\infty(0,1)\colon a(x) \geqslant \frac{\tilde{a}}{2} \text{ almost everywhere}\right\}. \tag{2.2}$$

Such an assumption will be made in our results. We next write (2.1) in matrix form and put $\gamma(t) = \operatorname{col}(\gamma_1(t), \ldots, \gamma_{i_N}(t))$. Then

$$G\dot{\gamma}(t) = \sum_{j=1}^{j_M}(a^M)_j A_j \gamma(t), \quad \text{for } t > 0,\ \gamma(0) = \gamma_0, \tag{2.3}$$

where $(G)_{ik} = (B_i^N, B_k^N), (A_j)_{ik} = -(\varphi_j^M(B_i^N)_x, (B_k^N)_x)$, and $(G\gamma_0)_i = (\varphi, B_i^N)$, for $1 \leqslant i, k \leqslant i_N$ and $1 \leqslant j \leqslant j_M$. Since identifiability will be discussed for fixed N and M, we again choose to drop these indices in the notation for the matrices A_j. If the dependence of γ on a is relevant we write $\gamma(t;\ a)$ to denote a solution of (2.3).

DEFINITION 2.1. The model parameter $a_m \in \mathscr{A}_{\text{ad}}$ is called *identifiable under approximation* for the indices N and M if $a_m^M \in \mathscr{A}^m \cap \mathscr{A}_{\text{ad}}^*$ and for all $a^M \in \mathscr{A}^M \cap \mathscr{A}_{\text{ad}}^*$ it follows that $\gamma(t;\ a_m^M) = \gamma(t;\ a^M)$ for all $t \geqslant 0$ implies $a_m^M = a^M$.

We need to introduce some notation before we state a necessary and sufficient condition that guarantees identifiability under approximation of a_m. Let λ_l, $1 \leqslant l \leqslant l_N \leqslant i_N$, denote the distinct eigenvalues of $G^{-1}\sum_{j=1}^{j_M}(a_m^M)_j A_j$ with associated eigenprojections P_l. We define the $i_N l_N \times j_M$ matrix $\mathscr{B}$ depending on M and N by

$$\mathscr{B} = \begin{pmatrix} A_1P_1\gamma_0 & A_2P_1\gamma_0 & \cdots & A_{j_M}P_1\gamma_0 \\ A_1P_2\gamma_0 & & & \\ \vdots & & & \\ A_1P_{l_N}\gamma_0 & & \cdots & A_{j_M}P_{l_N}\gamma_0 \end{pmatrix}.$$

THEOREM 2.1. *The parameter a_m is identifiable under approximation for the indices N and M, if $a_m^M \in \mathscr{A}^M \cap \mathscr{A}_{\text{ad}}^*$ and $\ker(\mathscr{B}) = \{0\}$. Conversely, if the parameter a_m is identifiable under approximation for the indices N and M, $l_N = i_N$, and $a_m^M(x) \geqslant \tilde{\alpha}/2 + \eta$ a.e. for some $\eta > 0$, then $\ker(\mathscr{B}) = 0$.*

---

[1]See J. Douglas, T. Dupont, and L. Wahlbin, *The stability in L^q of the L^2-projection into finite element function spaces*, Numer. Math. **23**, 193–197 (1975)

Proof. We start with some generalities. It is simple to verify that there exists a constant $\alpha > 0$ and that

$$\left\langle G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j x,\ x \right\rangle \leqslant -\alpha |x|^2_{\mathbf{R}^{l_N}}$$

for all $x \in \mathbf{R}^{l_N}$ and all $a^M \in \mathscr{A}^*_{\text{ad}} \cap \mathscr{A}^M$. Here $\langle \cdot , \cdot \rangle$ denotes the inner product in $\mathbf{R}^{l_N}$. The above dissipativity estimate implies that the spectrum of $G^{-1}\sum_{j=1}^{j_M}(a^M)_j A_j$ lies strictly in the left half-plane. We may take the Laplace transform in (2.3) with $(a^M)_j$ replaced by $(a^M_m)_j$ to obtain

$$\hat{\gamma}(s) = \left(sI - G^{-1} \sum_{j=1}^{j_M} (a^M_m)_j A_j \right)^{-1} \gamma_0, \quad \text{for } s \geqslant 0. \tag{2.4}$$

Next we define the error function

$$e(t) = \gamma(t;\ a^M) - \gamma(t;\ a^M_m), \quad \text{for } t \geqslant 0.$$

From (2.3) it follows that

$$\dot{e}(t) = G^{-1}\left(\sum_{j=1}^{j_M} \beta_j A_j \right) \gamma(t;\ a^M_m) + G^{-1} \sum_{j=1}^{j_M} (a^M)_j (A_j) e(t), \quad \text{for } t > 0, \tag{2.5}$$

$$e(0) = 0,$$

holds with $\beta_j = (a^M)_j - (a^M_m)_j$ and $e(t) \in \mathbf{R}^{l_N}$. We shall also make use of the meromorphic function [12; pp. 328, 331; 6, p. 38]

$$f(z) = \sum_{j=1}^{j_M} \beta_j A_j (zI - F)^{-1} \gamma_0, \quad \text{for } z \in \mathbf{C},$$

with possible poles at the eigenvalues λ_l of $F = G^{-1}\sum_{j=1}^{j_M}(a^M_m)_j A_j$. Let Γ_l, $1 \leqslant l \leqslant l_N$, be a positively oriented circle enclosing λ_l but excluding all other eigenvalues. Then

$$\frac{1}{2\pi i} \int_\Gamma f(z)\, dz = \sum_{j=1}^{j_M} \beta_j A_j P_l \gamma_0, \tag{2.6}$$

[12; p. 329], where P_l is the orthogonal projection onto the generalized eigenspace associated with the eigenvalue λ_l of F.

Now we prove the first part of Theorem 2.1 and assume that $\ker(\mathscr{B}) = 0$ and that $\gamma(t;\ a^M) = \gamma(t;\ a^M_m)$ for $t \geqslant 0$ and $a^M \in \mathscr{A}^*_{\text{ad}} \cap \mathscr{A}^M$. In this case $e(t) = 0$ for $t \geqslant 0$ and (2.5) implies

$$G^{-1}\left(\sum_{j=1}^{j_M} \beta_j A_j \right) \gamma(t;\ a^M_m) = 0, \quad \text{for } t \geqslant 0,$$

or equivalently

$$\left(\sum \beta_j A_j \right) \hat{\gamma}(s) = 0, \quad \text{for } s \geqslant 0.$$

Using this in (2.4), we obtain

$$\left(\sum_{j=1}^{j_M} \beta_j A_j \right) \left(sI - G^{-1} \sum_{j=1}^{j_M} (a^M_m)_j A_j \right)^{-1} \gamma_0 = 0 \quad \text{for } s \geqslant 0.$$

Since f is analytic on the resolvent set $\rho(F)$ of F we obtain by analytic continuation that $f = 0$ on $\rho(F)$. Using this in (2.6), we thus have

$$\left(\sum_{j=1}^{j_M} \beta_j A_j\right) P_l \gamma_0 = 0 \quad \text{for } 1 \leqslant l \leqslant l_N. \tag{2.7}$$

Since we assumed $\ker(\mathscr{B}) = \{0\}$ we arrive at $\beta = 0$. To prove the second part of the theorem let $\beta \in \ker(\mathscr{B})$, so that $\beta = \operatorname{col}(\beta_1, \ldots, \beta_{j_M})$ satisfies (2.7). Let us define $a^M = \sum_{j=1}^{j_M} \beta_j \varphi^M + a_m^M$. By changing β to $\varepsilon\beta$ with ε sufficiently small we can guarantee that $a^M(x) \geqslant \tilde{\alpha}/2$ almost everywhere, since by assumption $a_m^M(x) \geqslant \tilde{\alpha}/2 + \eta$ almost everywhere. Since we intend to show that $\beta = 0$ we may assume that β itself has been chosen sufficiently small. Let λ_l be an arbitrary eigenvalue of F. Then

$$f(z) = \sum_{j=1}^{j_M} \beta_j A_j (zI - F)^{-1} \gamma_0 = \sum_{j=1}^{j_M} \beta_j A_j \sum_{n=-1}^{\infty} (z - \lambda_l)^n B_n \gamma_0,$$

where $B_n = (1/2\pi i)\int_{\Gamma_l} (z - \lambda_l)^{-n-1}(zI - F)^{-1}\, dz$. Observe that (2.7) implies $\sum_{i=1}^{j_M} \beta_j A_j B_{-1} \gamma_0 = 0$. Therefore f is analytic inside Γ_l. But $f(z)$ is also bounded for $|z| \to \infty$ ([12, pp. 278, 321]) and therefore $f \equiv c_1$, $c_1 \in \mathbf{R}^{i_N}$. Integration of f around Γ_l for some l and (2.7) implies $c_1 = 0$ and therefore $f \equiv 0$. By (2.4) we have

$$G^{-1} \sum_{j=1}^{j_M} \beta_j A_j \hat{\gamma}(s) = 0 \quad \text{for } s \geqslant 0$$

and moreover, $G^{-1} \sum_{j=1}^{j_M} \beta_j A_j \gamma(t;\ a_m^M) = 0$ for $t \geqslant 0$. Now (2.5) implies $e(t) = \gamma(t;\ a^M) - \gamma(t;\ a_m^M) \equiv 0$. Since a_m is identifiable under approximation for the indices N and M this implies $(a^M)_j - (a_m^M)_j = \beta_j = 0$, $j = 1, \ldots, j_m$, and the proof is complete.

REMARK 2.1. A necessary condition for parameter identifiability under approximation for indices N and M of a_m is given by $j_M \leqslant i_N l_N$.

The proof of Theorem 2.1 reveals the following corollaries concerning sufficient conditions for parameter identifiability under approximation.

COROLLARY 2.1. Let $\{A_j P_l \gamma_0\colon\ j = 1, \ldots, j_M\}$ be linearly independent for some l. Then a_m is parameter identifiable under approximation for indices N and M, if $a_m^M \in \mathscr{A}^M \cap \mathscr{A}_{\text{ad}}^*$.

The proof of this corollary follows directly from (2.7). Note also, that P_l *depends on* a_m^M.

COROLLARY 2.2. Let $\{A_j \gamma_0\colon\ j = 1, \ldots, j_M\}$ be linearly independent. Then *any* $a_m \in \mathscr{A}_{\text{ad}}$ is parameter identifiable under approximation for indices N and M, if $a_m^M \in \mathscr{A}^M \cap \mathscr{A}_{\text{ad}}^*$.

Proof. If we sum (2.7) over l and use the fact that P_l forms a resolution of the identity, $I = \sum_{l=1}^{l_N} P_l$, then $\sum_{j=1}^{j_M} \beta_j A_j \gamma_0 = 0$. Since $\{A_j \gamma_0\}_{j=1}^{j_M}$ is a linearly independent set of vectors, $\beta_j = 0$ for all j and the claim follows.

REMARK 2.2. The conditions in Corollaries 2.1 and 2.2 require $j_M \leqslant i_N$.

We end this section with some remarks concerning output least squares identifiability (OLSI) of the approximating equations (2.1). First we recall [4] the definition of OLSI for the problem under consideration. Let $Z = C(0, \infty; \mathbf{R}^{i_N})$, the space of continuous bounded functions endowed with the supremum norm, and consider a convex and compact subset

$\tilde{\mathscr{A}}$ of the set of admissible parameters $\mathscr{A}_{ad}^* \cap \mathscr{A}^M$. By $u^N(\cdot;\ \tilde{\mathscr{A}})$ we denote the set of admissible trajectories as a^M varies in $\tilde{\mathscr{A}} \subset \mathscr{A}_{ad}^* \cap \mathscr{A}^M$.

DEFINITION 2.2. If there exists a neighborhood V of $u^N(\cdot;\ \tilde{\mathscr{A}})$ in $C(0,\infty;\ L^2(0,1))$, so that for every $\hat{y} \in V$ the problem $(P^{N,M})$ has a unique solution depending continuously on $\hat{y}$, then $(P^{N,M})$ is called OLSI.

We note that if $u^N(\cdot;\ \tilde{\mathscr{A}})$ were closed and convex, then injectivity of $a^M \to u^N(\cdot;\ a^M) \in C(0,\infty;\ L^2(0,1))$ as discussed in Theorem 2.1 above would guarantee OLSI. However, $u^N(\cdot;\ \tilde{\mathscr{A}})$ is not convex. For take a^M and b^M in $\tilde{\mathscr{A}}$. Then there ought to exist c^M in $\tilde{\mathscr{A}}$ so that for any $\vartheta \in (0,1)$

$$\vartheta u^N(\cdot;\ a^M) + (1-\vartheta) u^N(\cdot;\ b^M) = u^N(\cdot;\ c^M).$$

Such a c^M will not exist in general, as can be seen from the following special case. Let $M = 1$ and $\mathscr{A}^1$ be the subspace of $L^\infty(0,1)$ spanned by the constant function with value 1. Then (2.3) becomes

$$\dot{\gamma}(t) = aG^{-1}A\gamma(t),$$

where $(A)_{i,j} = -((B_i^N)_x,\ (B_j^N)_x)$ and $a \in [\tilde{\alpha}/2, \infty)$. We put $H = G^{-1}A$. Note that $a \to \gamma(\cdot;\ a)$ from $[\tilde{\alpha}/2,\infty)$ to $C(0,\infty;\ \mathbf{R}^{i_N})$ is not convex. If it were, then for each pair $a, b \in [\tilde{\alpha}/2,\infty)$ with $a \neq b$ there exists $c \in [\tilde{\alpha}/2,\infty)$ such that

$$\tfrac{1}{2}\exp(aHt)\gamma_0 + \tfrac{1}{2}\exp(bHt)\gamma_0 = \exp(cHt)\gamma_0.$$

Differentiating this equation twice with respect to t and evaluating at $t = 0$ implies

$$(a+b)H\gamma_0 = 2cH\gamma_0 \quad \text{and} \quad (a^2+b^2)H^2\gamma_0 = 2c^2H^2\gamma_0.$$

If $\gamma_0 \notin \ker H^2$, then no $c \in (-\infty,\infty)$ satisfies these two equations simultaneously.

Recall that $u^N(t;\ a^M) = \sum_{i=1}^{i_N} \gamma_i(t;\ a^M) B_i^N$. We denote the Fréchet derivative of $\gamma(\cdot;\ a^M)$ with respect to a^M by $\partial\gamma(\cdot;\ a^M)$; similarly $\partial^2\gamma(\cdot;\ a^M)$ stands for the second Fréchet derivative. In [4] it is proved that if for some $\beta_1 > 0$ and $\beta_2 > 0$

$$\beta_1 |a|_{\mathbf{R}^{M_j}} \leqslant |\partial\gamma(\cdot;\ a^M)a|_{C(0,\infty;\mathbf{R}^{N_i})}, \tag{2.8}$$

$$|\partial^2\gamma(\cdot;\ a^M)(a,a)|_{C(0,\infty;\mathbf{R}^{N_i})} \leqslant \beta |a|^2_{\mathbf{R}^{M_j}}, \tag{2.9}$$

for all $a \in \mathbf{R}^{M_j}$ and all $a^M \in \tilde{\mathscr{A}}$, and

$$\operatorname{diam}(\tilde{\mathscr{A}}) \leqslant 2\sqrt{2}\,\beta_1\beta_2^{-1}, \tag{2.10}$$

then $(P^{N,M})$ is OLSI.

In general it seems quite intractable to calculate β_1 in (2.8). However, under certain conditions $\partial\gamma(\cdot;\ a^M)$ is injective for every $a^M \in \mathscr{A}_{ad}^*$. If (2.9) holds as well, then the finite-dimensionality will guarantee OLSI for $\operatorname{diam}(\tilde{\mathscr{A}})$ sufficiently small. The Fréchet differential $\partial\gamma(\cdot;\ a^M)(h)$ of $\gamma(\cdot;\ a^M)$ at a^M in direction of $(h) = \operatorname{col}(h_1, \ldots, h_{j_M})$ is given by

$$\frac{d}{dt}\partial\gamma(t;\ a^M)(h) = G^{-1}\sum_{j=1}^{j_M}(h)_j A_j \gamma(t;\ a^M) + G^{-1}\sum_{j=1}^{j_M}(a^M)_j A_j \partial\gamma(t;\ a^M)(h), \tag{2.11}$$

$$\partial\gamma(0;\ a^M)(h) = 0.$$

For a proof we refer to the appendix (Sec. 5). Similarly, one can show that $z(t) = \partial^2\gamma(\cdot; a^M)((h),(k))$ is the unique solution of

$$\begin{cases} \dot{z}(t) = G^{-1}\left(\sum_{j=1}^{j_M}(a^M)_j A_j\right)z(t) + G^{-1}\left(\sum_{j=1}^{j_M}(k)_j A_j\right)\delta\gamma(t;\, a^M)(h) \\ \qquad\qquad + G^{-1}\left(\sum_{j=1}^{j_M}(h)_j A_j\right)\delta\gamma(t;\, a^M)(k), \\ z(0) = 0. \end{cases} \tag{2.12}$$

From (2.11) it follows that $\partial\gamma(\cdot;\, a^M) = 0$ for $t > 0$ if and only if

$$\sum_{j=1}^{j_M}(h)_j A_j \gamma(t;\, a^M) = 0 \quad \text{for all } t \geqslant 0. \tag{2.13}$$

To discuss injectivity of $\partial\gamma(\cdot;\, a^M)$ we let $\lambda_l(a^M)$, $1 \leqslant l \leqslant l_{N(a^M)} \leqslant i_N$ denote the distinct eigenvalues of $G^{-1}\sum_{j=1}^{j_M}(a_j^M)A_j$ with associated eigenprojections $P_l(a^M)$, where $a^M \in \tilde{\mathscr{A}}$. Then $i_N l_{N(a^M)} \times j_M$ matrices $\mathscr{B}(a^M)$ are defined analogously to $\mathscr{B}$ with P_l replaced by $P_l(a^M)$.

PROPOSITION 2.1. If $\tilde{\mathscr{A}} \subset \mathscr{A}_{\text{ad}}^* \cap \mathscr{A}^M$ is compact and convex and $\ker\mathscr{B}(a^M) = \{0\}$ for every $a^M \in \tilde{\mathscr{A}}$, then $(P^{N,M})$ with $\mathscr{A}_{\text{ad}}^*$ replaced by $\tilde{\mathscr{A}}$ is OLSI, provided $\text{diam}(\tilde{\mathscr{A}})$ is sufficiently small. The condition on $\ker\mathscr{B}(a^M)$ can be replaced by requiring that for every $a^M \in \tilde{\mathscr{A}}$ there exists an $l(a^M)$ such that $\{A_j P_{l(a^M)}\gamma_0\colon 1,\dots, j_M\}$ is linearly independent or by requiring that $\{A_j\gamma_0\colon j = 1,\dots, j_M\}$ is linearly independent.

Proof. From (2.12) and (A.5) (see the appendix, Sec. 5), it follows that (2.9) is satisfied and it therefore suffices to verify (2.8) or, equivalently, the injectivity of $(h) \to \partial\gamma(\cdot;\, a^M)(h)$. As observed above, $\partial\gamma(\cdot;\, a^M)(h) = 0$ if and only if

$$\sum_{j=1}^{j_M}(h)_j A_j \gamma(t;\, a^M) = 0 \quad \text{for all } t \geqslant 0. \tag{2.14}$$

Let us assume that (2.14) holds. Then as in the proof of Theorem 2.1 we argue that

$$\sum_{j=1}^{j_M}(h)_j A_j P_l(a^M)\gamma_0 = 0 \quad \text{for } l = 1,\dots, l_{N(a^M)}. \tag{2.15}$$

Since $\ker\mathscr{B}(a^M) = 0$, (2.15) implies $(h) = 0$. The remaining two assertions are the analogues of Corollaries 2.1 and 2.2.

3. Examples. In this section we present two simple examples to illustrate the theory that was developed in Sec. 2. Throughout we choose $M = 2$ and we drop this index. Let

$$\varphi_1(x) = \chi_{(0,\frac{1}{2})}(x) \quad \text{and} \quad \varphi_2(x) = \chi_{(\frac{1}{2},1)}(x),$$

be the characteristic functions of the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$, respectively. Further, we define the linear spline basis elements in $H_0^1(0,1)$ by

$$B_i^N(x) = \begin{cases} Nx - (i-1), & \text{for } (i-1)/N \leqslant x \leqslant i/N, \\ -Nx + (i+1), & \text{for } i/N \leqslant x \leqslant (i+1)/N, \\ 0, & \text{otherwise,} \end{cases}$$

for $i = 1, 2, \ldots, N-1$.

First we consider the case $N = 4$ and derive a sufficient condition for parameter identifiability under approximation for indices $N = 4$, $M = 2$. We compute A_1 and A_2 and obtain

$$A_1 = \begin{pmatrix} -8 & 4 & 0 \\ 4 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -4 & 4 \\ 0 & 4 & -8 \end{pmatrix}.$$

Further, if we set $(\tilde{\mu})_i = \mu_i$, where

$$\mu_i = \int_0^1 \varphi(x) B_i^4(x)\, dx, \quad i = 1, 2, 3,$$

then $\gamma_0 = G^{-1}\tilde{\mu}$. In the present case we have

$$G^{-1} = \frac{3}{7}\begin{pmatrix} 15 & -4 & 1 \\ -4 & 16 & -4 \\ 1 & -4 & 15 \end{pmatrix}.$$

Moreover, we find

$$A_1\gamma_0 = \frac{3}{7}\begin{pmatrix} -136\mu_1 + 96\mu_2 - 24\mu_3 \\ 76\mu_1 - 80\mu_2 + 20\mu_3 \\ 0 \end{pmatrix}$$

and

$$A_2\gamma_0 = \frac{3}{7}\begin{pmatrix} 0 \\ 20\mu_1 - 80\mu_2 + 76\mu_3 \\ -24\mu_1 + 96\mu_2 - 136\mu_3 \end{pmatrix}.$$

If $A_1\gamma_0$ and $A_2\gamma_0$ are linearly independent, then any $a_m \in \mathscr{A}_{\text{ad}}^*$ is parameter identifiable under approximation for indices $N = 4$, $M = 2$. We observe that if $-136\mu_1 + 96\mu_2 - 24\mu_3 \neq 0$ and $-24\mu_1 + 96\mu_2 - 136\mu_3 \neq 0$, then $A_1\gamma_0$ and $A_2\gamma_0$ are linearly independent.

Next we consider the case $N = 3$ and study parameter identifiability under approximation of a_m. We assume $a_m^2 = (1, 1)$. In this case the matrix G is given by

$$G = \frac{1}{18}\begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad G^{-1} = \frac{6}{5}\begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix},$$

while the matrices A_1 and A_2 are found to be

$$A_1 = \frac{3}{2}\begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad A_2 = \frac{3}{2}\begin{pmatrix} -1 & 1 \\ 1 & -3 \end{pmatrix}.$$

From this we see that

$$G^{-1}A_1 = \frac{9}{5}\begin{pmatrix} -13 & 5 \\ 7 & -5 \end{pmatrix} \quad \text{and} \quad G^{-1}A_2 = \frac{9}{5}\begin{pmatrix} -5 & 7 \\ 5 & -13 \end{pmatrix}.$$

Then

$$B = G^{-1}\sum_{j=1}^{2}\left(a_m^2\right)_j A_j = \frac{9}{5}\begin{pmatrix} -18 & 12 \\ 12 & -18 \end{pmatrix}.$$

The eigenvectors of B are

$$u_1 = \mathrm{col}(-1,1) \quad \text{and} \quad u_2 = \mathrm{col}(1,1),$$

and the projection operators onto the eigenspaces associated with these eigenvectors are

$$P_1 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad P_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Let $\tilde{\mu}_i = \int_0^1 \varphi(x)B_i^3(x)\,dx$, $i = 1,2$, and $\tilde{\mu} = \mathrm{col}(\tilde{\mu}_1, \tilde{\mu}_2)$. Then we find

$$A_1P_1G^{-1}\tilde{\mu} = \frac{9}{5}\begin{pmatrix} -10(\tilde{\mu}_1 - \tilde{\mu}_2) \\ 5(\tilde{\mu}_1 - \tilde{\mu}_2) \end{pmatrix}, \qquad A_2P_1G^{-1}\tilde{\mu} = \frac{9}{5}\begin{pmatrix} -5(\tilde{\mu}_1 - \tilde{\mu}_2) \\ 10(\tilde{\mu}_1 - \tilde{\mu}_2) \end{pmatrix},$$

$$A_1P_2G^{-1}\tilde{\mu} = \frac{9}{5}\begin{pmatrix} -3(\tilde{\mu}_1 + \tilde{\mu}_2) \\ 0 \end{pmatrix}, \qquad A_2P_2G^{-1}\tilde{\mu} = \frac{9}{5}\begin{pmatrix} 0 \\ -3(\tilde{\mu}_1 + \tilde{\mu}_2) \end{pmatrix}.$$

Observe that

$$\det\left(A_1P_1G^{-1}\tilde{\mu},\ A_2P_1G^{-1}\tilde{\mu}\right) = -75(\tilde{\mu}_1 - \tilde{\mu}_2)^2$$

and

$$\det\left(A_1P_2G^{-1}\tilde{\mu},\ A_2P_2G^{-1}\tilde{\mu}\right) = -9(\tilde{\mu}_1 + \tilde{\mu}_2)^2.$$

Hence if $\tilde{\mu}_1 \neq \tilde{\mu}_2$ then $A_1P_1G^{-1}\tilde{\mu}$ and $A_2P_1G^{-1}\tilde{\mu}$ are linearly independent, while if $\tilde{\mu}_1 \neq -\tilde{\mu}_2$ then $A_1P_2G^{-1}\tilde{\mu}$ and $A_2P_2G^{-1}\tilde{\mu}$ are linearly independent. In either case $a_m^M = (1,1)$ is identifiable under approximation. Finally, observe that

$$\begin{pmatrix} -10(\tilde{\mu}_1 - \tilde{\mu}_2) & -5(\tilde{\mu}_1 - \tilde{\mu}_2) \\ 5(\tilde{\mu}_1 - \tilde{\mu}_2) & 10(\tilde{\mu}_1 - \tilde{\mu}_2) \\ -3(\tilde{\mu}_1 + \tilde{\mu}_2) & 0 \\ 0 & 3(\tilde{\mu}_1 + \tilde{\mu}_2) \end{pmatrix}$$

has rank 2 if $(\tilde{\mu}_1, \tilde{\mu}_2) \neq (0,0)$, and hence by Theorem 2.1 the parameter $a_m^2 = (1,1)$ is identifiable under approximation in this case.

4. Discussion. In this paper we studied the problem of parameter identifiability in Galerkin approximations for a class of parabolic distributed systems. To the author's knowledge, this is the first attempt to achieve an understanding of the difficulties that arise in the investigation of the injectivity of the parameter-to-output map determined by the specific structure of the ordinary differential equations that arise in Galerkin approximations. Although we have only treated here parameter identifiability under approximation of the diffusion coefficient as a specific case, the method that we propose seems to be quite general and will be applicable to deriving similar results for other parameters in distributed systems. Our interest in this problem was stimulated by the increased amount of research that is currently being pursued to develop approximation methods for the parameter identification problem in infinite-dimensional systems (see [1–3, 8], for example).

The question of identifiability of the parameter in the original system has been studied for some time. Much of this work has its roots in a paper by Gelfand and Levitan [5] on the reconstruction of the function q in the scalar equation

$$\begin{cases} L(y) = y'' + (\lambda - q)y = 0, \\ y(0) = 1, \qquad y'(0) = h, \end{cases} \tag{4.1}$$

with $\lambda \in \mathbf{R}$, $h \in \mathbf{R}$, from the spectral function associated with (4.1). Among the recent investigators of parameter identifiability in parabolic systems are Kitamura and Nakagiri [7], who study identifiability of constant and variable coefficients from point and distributed observations, respectively. Pierce [9] gives an example in which he shows how knowledge of the spectral function together with the results of Gelfand and Levitan can be used to identify q in $y_t = Ly$ together with mixed boundary conditions in $x = 0$ and $x = 1$, from boundary observations at $x = 0$. Suzuki's technique [10] (using integral representation of the stationary solution) gives necessary and sufficient conditions for identifiability of parameters in the equation, in the boundary condition as well as in the initial data. Chavent and co-workers [1, 4] have introduced the concept of output least squares identifiability, which is defined as the problem of continuous dependence of the unique solution (parameter) minimizing the quadratic fit-to-data criterion which measures the distance between "observation" and trajectory evaluated for a specific parameter value [compare (P)]. Up to now there have been only a few equations of finite-dimensional or hyperbolic nature, where output least squares identifiability could be verified [1, 4, 11]. Once identifiability under approximation is shown, it is quite simple to argue output least squares identifiability of the Galerkin approximation on a sufficiently small parameter set. Generally it is not possible to give precise estimates of the first and second Fréchet derivatives of the parameter-to-observation map, and consequently it is difficult to determine explicitly the size of the sets of parameters such that output least squares identifiability is guaranteed [compare (2.10)].

5. Appendix. Here we prove (2.11), the characterization of the Fréchet derivative of $\gamma(t;\, a)$ with respect to a. As in the proof of Theorem 2.1, we can agree that there exist constants K_1 and K_2 such that

$$\|\gamma(t;\, a^M)\|_{\mathbf{R}^{i_N}} \leqslant K_1 e^{-K_2 t} \|\gamma_0\|_{\mathbf{R}^{i_N}} \tag{A.1}$$

for all $\gamma_0 \in \mathbf{R}^{i_N}$ and $a^M \in \mathscr{A}^M \cap \{a \in L^\infty(0,1) : a(x) \geqslant \tilde{\alpha}/4 \text{ a.e.}\} = \mathscr{A}^+$. Let $\varepsilon \in \mathbf{R}$, $h \in L^\infty(0,1)$ with $h = \sum_{j=1}^{j_M} (h)_j \varphi_j^M$ and $a^M \in \mathscr{A}^M \cap \mathscr{A}_{\text{ad}}^*$. Then $a^M + \varepsilon h \in \mathscr{A}^+ \cap \mathscr{A}^M$ for ε sufficiently small. It is quite simple to see that

$$\lim_{\varepsilon \downarrow 0} \|\gamma(\cdot\,;\, a^M + \varepsilon h) - \gamma(\cdot\,;\, a^M)\|_{C(0,\infty;\mathbf{R}^{i_N})} = 0. \tag{A.2}$$

We put $w(t;\, \varepsilon) = \gamma(t;\, a^M + \varepsilon h) - \gamma(t;\, a^M)$ for $t \geqslant 0$. Then

$$\dot{w}(t;\, \varepsilon) = \varepsilon G^{-1} \sum_{j=1}^{j_M} (h)_j A_j \gamma(t;\, a^M + \varepsilon h) + G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j w(t;\, \varepsilon) \quad \text{for } t > 0,$$

$$w(0;\, \varepsilon) = 0,$$

and therefore

$$w(t;\, \varepsilon) = \varepsilon \int_0^t \exp\left[G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j (t-s) \right] G^{-1} \sum_{j=1}^{j_M} (h)_j A_j \gamma(s;\, a^M + \varepsilon h)\, ds. \tag{A.3}$$

By (A.1) and (A.2) we observe that the integral in (A.3) converges in $C(0, \infty;\, \mathbf{R}^{i_N})$ to

$$\int_0^t \exp\left[G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j (t-s) \right] G^{-1} \sum_{j=1}^{j_M} (h)_j A_j \gamma(s;\, a^M)\, ds$$

as $\varepsilon \to 0$. Consequently $\lim_{\varepsilon \to 0} \varepsilon^{-1} w(\cdot;\, \varepsilon) = \delta\gamma(\cdot;\, a^M)(h)$ exists in $C(0, \infty;\, \mathbf{R}^{i_N})$ and

$$\delta\gamma(t;\, a^M)(h) = \int_0^t \exp\left[G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j (t-s) \right] G^{-1} \sum_{j=1}^{j_M} (h)_j A_j \gamma(s;\, a^M)\, ds.$$

Finally, we obtain

$$\frac{d}{dt} \delta\gamma(t;\, a^M)(h) = G^{-1} \sum_{j=1}^{j_M} (a^M)_j A_j \delta\gamma(t;\, a^M)(h) + G^{-1} \sum_{j=1}^{j_M} (h)_j A_j \gamma(t;\, a^M) \quad \text{for } t > 0$$

$$\delta\gamma(0;\, a^M)(h) = 0. \tag{A.4}$$

We observe that (A.4) implies

$$\| \delta\gamma(\cdot;\, a^M)(h) \|_{C(0,\infty;\mathbf{R}^{i_N})} \leqslant C_1 |(h)|_{\mathbf{R}^{j_M}} |\gamma_0|_{\mathbf{R}^{i_N}}, \tag{A.5}$$

where C_1 depends on $\tilde{\alpha}$ and A_i but not on γ_0 and $a^M \in \mathscr{A}_{\mathrm{ad}}^* \cap \mathscr{A}^M$. Let $B = G^{-1} \Sigma_{j=1}^{j_M} a_j M_j$. Then for $a, \bar{a} \in \mathscr{A}_{\mathrm{ad}}^* \cap \mathscr{A}^M$ and $(h), (\bar{h}) \in \mathbf{R}^{j_M}$ we have

$$\begin{aligned} \delta\gamma(t;\, \bar{a})(\bar{h}) - \delta\gamma(t;\, a)(h) &= \int_0^t e^{B(t-s)} G^{-1} \left(\sum_{j=1}^{j_M} \big(((\bar{a})_j - (a)_j) A_j \big) \delta\gamma(s;\, \bar{a})(h) \right) ds \\ &\quad + \int_0^t e^{B(t-s)} G^{-1} \sum_{j=1}^{j_M} (\bar{h})_j A_j \big(\gamma(s;\, \bar{a}) - \gamma(s;\, a) \big)\, ds \\ &\quad + \int_0^t e^{B(t-s)} G^{-1} \left(\sum_{j=1}^{j_M} \big((\bar{h})_j - (h)_j \big) A_j \right) \gamma(s;\, a)\, ds. \end{aligned}$$

Similarly to (A.2), one can show that

$$\| \gamma(\cdot;\, \bar{a}) - \gamma(\cdot;\, a) \|_{C(0,\infty;\mathbf{R}^{i_N})} \leqslant C_2 |(\bar{a}) - (a)|_{\mathbf{R}^{j_M}} |\gamma_0|_{\mathbf{R}^{i_N}}, \tag{A.6}$$

with C_2 depending on $\tilde{\alpha}$ and A_j, but not on $\bar{a}$, a, and γ_0. Now (A.1), (A.5), and (A.6) imply

$$\|\delta\gamma(\cdot;\ \bar{a})(\bar{h}) - \delta\gamma(\cdot;\ a)(h)\|_{C(0,\infty;\mathbf{R}^{N})} \leqslant C_3|\gamma_0|_{\mathbf{R}^{N}}\big(|(\bar{a}) - (a)|_{\mathbf{R}^{JM}}|(h)|_{\mathbf{R}^{JM}} + |(\bar{a}) - (a)|_{\mathbf{R}^{JM}}|(h)|_{\mathbf{R}^{JM}} + |(\bar{h}) - (h)|_{\mathbf{R}^{JM}}\big), \quad (A.7)$$

with C_3 independent of $\bar{a}$, a, h, and $\bar{h}$ (and γ_0). From (A.7) it finally follows that $a \to \gamma(\cdot;\ a)$ is continuously Fréchet-differentiable and the differential in direction h is given by the solution of (A.4).

Acknowledgment. We thank the referee for some helpful comments. K. Kunisch gratefully acknowledges support from the Max Kade Foundation. Research supported in part by the Fonds zur Förderung der Wissenschaftlichen Forschung, Austria, No. P4534, and by the Steiermärkische Wissenschafts und Forschungsförderungsfonds. L. W. White was supported in part by a grant from the Venezuelan Petroleum Institute and by the Cooperative Institute for Mesoscale Meteorological Studies. Parts of the research were carried out while the authors were visitors at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, Virginia, which is operated under NASA contracts NAS1-15810 and NAS1-16394.

References

[1] A. Bamberger, G. Chavent, and P. Lailly, *About the stability of the inverse problem in a 1-D wave equation—application to the interpretation of seismic profiles*, J. Appl. Math. Optim. **5**, 1–47 (1979)

[2] H. T. Banks and P. L. Daniel, *Estimation of variable coefficients in parabolic systems*, LCDS Report 82-22, Brown University, Sept. 1982; IEEE Trans. Autom. Control. **30**, 386–398 (1985)

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

[4] G. Chavent, *Local stability of the output least square parameter estimation technique*, INRIA Report No. 136, Paris (1982); Mat. Apl. Comput. **2**, 3–22 (1983)

[5] I. M. Gelfand and B. M. Levitan, *On the determination of a differential equation from its spectral function*, Am. Math. Soc. Transl. **1**, 253–304 (1955)

[6] T. Kato, *Perturbation theory for linear operators*, Springer, New York, 1966

[7] S. Kitamura and S. Nakagiri, *Identifiability of spatially varying and constant parameters in distributed systems of parabolic type*, SIAM J. Control Optimization **15**, 785–802 (1979)

[8] K. Kunisch and L. W. White, *The parameter estimation problem for parabolic equations and discontinuous observation operators*, SIAM J. Control Optimization **23**, 900–927 (1985)

[9] A. Pierce, *Unique identification of eigenvalues and coefficients in a parabolic problem*, SIAM J. Control Optimization **17**, 494–499 (1979)

[10] T. Suzuki, *Uniqueness and nonuniqueness in an inverse problem for the parabolic equation*, J. Differ. Eq. **47**, 296–316 (1983)

[11] L. W. White, *Identification of a friction parameter in a first order linear hyperbolic equation*, Proc. 22nd IEEE Conf. Decision and Control, San Antonio, 56–59 (1983)

[12] A. E. Taylor and D. C. Lay, *Introduction to functional analysis*, Wiley, New York, 1980