A Discrete-Time Adaptive Filter for Stochastic Distributed Parameter Systems

K. Watanabe
Research Associate,
Department of Aeronautical Engineering,
Kyushu University,
Fukuoka, Japan

T. Yoshimura
Associate Professor.

T. Soeda
Professor.

Department of Mechanical Engineering,
Tokushima University,
Tokushima, Japan

A discrete-time adaptive filter is derived for a distributed system described by a linear partial differential equation with some unknown random constants whose a priori probabilities are known. The system concerned contains the Gaussian white noise in time, and its measurement system is treated as a so-called pointwise observation in which the measurement is taken at the finite discrete subdomains in the coordinate spaces. The use is conceptually made of an adaptive technique based on the Bayesian method and it is shown that the optimal distributed filter proposed here can be partitioned into two parts, a linear nonadaptive part that consists of a bank of distributed Kalman-Bucy type filters and a nonlinear part that incorporates the learning nature of the estimator. For the derivation of each "elemental filter," the discrete-time innovation theory is utilized. The eigenfunction expanding method in a complete orthonormal system is applied for the numerical procedure of the proposed filter. From the simulation of the estimation problem for the neutron flux distribution in a slab type nuclear reactor, the proposed adaptive filter is shown to have attractive characteristics and therefore can be recommended for practical on-line adaptive estimation of distributed parameter systems.

1 Introduction

The estimation problem of stochastic distributed dynamical systems is one of important and necessary problems for modeling and adaptive control in a practical field. In recent application studies in this field, for instance, the distributed square-root filter for estimation of air pollution concentrations has been proposed by Koda and Seinfeld [1], the least squares filter for estimation of transient temperature profile in an alminum slab subject to heating and cooling has been derived by Ajinkya, et al. [2], the estimation of temperature distribution for radial and axial direction in a cylindrical ingot has been implemented by Lausterer, et al. [3, 4], and the problem of imaging for gamma rays in the medical diagnosis used radioactive rays has been solved by Tzafestas [5,6].

It should be noted that most of the practical physical systems exhibit a randomness over rather broad scales of time and space. Especially, the case of parameter uncertainties frequently appears in practice. Since the system parameters (for example, in thermal conductive systems, the thermal conductivity, thermal radiation coefficient, etc.) and noise statistics must be completely known at the stage of its synthesis, any adaptive techniques may be required to compensate such uncertainties if the implementing of optimal filtering is hoped.

It is well known that the nonlinear filter techniques [7], which perform simultaneous state estimation and parameter identification, are effective on such a problem. For lumped parameter systems, though the properties of the extended Kalman filter as the parameter estimator are studied rather strictly [8], there is a lack of theoretical aspects or a so-called "curse of dimensionality" for such augmentation procedures. Therefore, comparatively little has been reported on a distributed adaptive filter up to now. Sunahara, et al. [9], have already tackled these problems in continuous-time stochastic distributed systems from the viewpoint of the Bayesian method, but they treated only the estimation of force term (i.e., thermal sources or pollutant ones) which is ineffective on the system stability.

In this paper, an adaptive filter for a distributed parameter system with unknown time and space-invariant parameters will be developed. The measurement system is considered as a so-called "pointwise observation" [10-12] in which the noisy measurement is carried out at several fixed points of the spatial domain and the boundary. The unknown parameters are assumed to be random constants over the operating range and therefore the Bayesian method is selected as the candidate for an adaptive procedure. The adaptive filter is derived in discrete-time, which is suitable for on-line estimation by electronic-computer, and it is shown that consequently this filter reduces to the parallel distributed Kalman filter consising of some "elemental distributed filters." That is, the present method is a direct extension to distributed parameter systems of the "partition theorem" [13,14] given in Lainiotis' works. In the derivation of each elemental distributed filter, the discrete-time innovation theory [15] for distributed systems with pointwise observation is utilized. For numerical aspects of this filter, the eigenfunction expansion method [2,10] is introduced. Thus, the main feature of this paper is to

Contributed by the Dynamic Systems and Control Division for publication in the JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL. Manuscript received by the Dynamic Systems and Control Division, June 23, 1980.

discretize the system (in time and also in space) and to show that the adaptive filter algorithms similar to those of lumped parameter systems [13,14,16-19] can be obtained. Finally, some qualitative and quantitative features of the proposed adaptive filter are examined by the estimation problem for the neutron flux distribution in a slab type nuclear reactor.

2 Problem Formulation

Let D be a bounded open domain in an r-dimensional Euclidean space $\mathbf{E}^r$, and ∂D, the boundary of D, be a finite number of $(r-1)$-dimensional hypersurface of class C^3. Let $\bar{D}$ be the closure of D, $\bar{D}=D\cup\partial D$ and $T=\{t;t\geq t_0\}$ be a real time interval, and define the product space with respect to coordinate space D by $\Sigma_s=D\times T$. Furthermore, denote the spatial coordinate vector by $x=(x_1,x_2,\cdots,x_r)\epsilon D$.

Consider a distributed parameter system described by a linear spatial differential equation,

$$\frac{\partial u(x,t)}{\partial t}=\mathcal{L}_x(\theta)u(x,t)+G(x,t;\theta)\eta(x,t),\ (x,t)\epsilon\Sigma_s \quad (2\text{-}1)$$

where $u(x,t)$ is an n-dimensional state vector, $\mathcal{L}_x(\cdot)$ is an $n\times n$ matrix valued linear spatial differential operator whose parameters are bounded and may depend upon x and/or t for all $(x,t)\epsilon\Sigma_s$, $G(x,t,\cdot)$ is an $n\times q$ system noise intensity matrix and $\eta(x,t)$ is a q-dimensional white Gaussian noise process in time. Here, a "subscript θ" denotes that there exists unknown parameters in the corresponding operators or matrices and the unknown parameter $\theta\epsilon\Omega_\theta$ is assumed to be a p'-dimensional random constant vector whose a priori probability is known and takes the values in fixed discrete parameter space Ω_θ, and $G(x,t,\cdot)\epsilon L^\infty(0,\infty;L^2(D))$ where $L^2(D)$ denotes the Hilbert space all square integrable real-valued functions with the inner product given by

$$<u_1,u_2>=\int_D u_1^T(x)u_2(x)dx,\quad \text{for all } u_i,\ i=1,2\epsilon L^2(D).$$

The observation process is given as so-called pointwise type [10-12] in which the measurement is taken at the finite discrete m points x^1, x^2, $\cdots$, x^m of the coordinate spaces $\bar{D}$ such as

$$z(t)=H(t;\theta)u_m(t)+\zeta(t),\ t\epsilon T \quad (2\text{-}2)$$

where $z(t)$ is a p-dimensional measurement vector at the points $x^1,\cdots,x^m$, $H(t;\cdot)$ is a $p\times nm$ matrix, $u_m(t)$ is an nm-dimensional state vector defined on finite discrete locations $x^i\epsilon\bar{D}$, $i=1,2,\cdots,m$ and defined as

$$u_m(t)=\mathrm{Col}[u(x^1,t),\cdots,u(x^m,t)] \quad (2\text{-}3)$$

where Col denotes the column vector, and $\zeta(t)$ is a p-dimensional white Gaussian measurement noise process.

The boundary condition for the system (2-1) to be governed by homogenous Dirichlet condition is given by

$$\beta_\xi(\theta)u(\xi;t)=0,(\xi,t)\epsilon\Sigma_b \quad (2\text{-}4)$$

where $\Sigma_b=\partial D\times T$ and $\beta_\xi(\cdot)$ is an $n\times n$ matrix valued linear spatial operator which is defined on the boundary surface ∂D at a point $\xi\epsilon\partial D$. This is a matter of convenience cases where Neumann and mixed conditions on the boundary can also be handled.

The initial condition is given by

$$\lim_{t\to+0} u(x,t)=u_0(x)\ \text{in}\ L^2(D),u_0(x)\epsilon L^2(D) \quad (2\text{-}5)$$

and further the system noise $\eta(x,t)$ and measurement noise $\zeta(t)$ are assumed to be stochastically independent of each other and also independent of the random initial condition $u_0(x)$, and their mean and covariance values are given by

$$\left.\begin{aligned}&E[\eta(x,t)]=0,\ E[\eta(x,t)\eta^T(y,\tau)]=Q(x,y,t)\delta(t-\tau)\\&E[\zeta(t)]=0,\ E[\zeta(t)\zeta^T(\tau)]=R(t)\delta(t-\tau),\\&\qquad (x,y,t)\epsilon D\times\Sigma_s\\&E[u_0(x)]=\bar{u}_0(x),\ E[\{u_0(x)-\bar{u}_0(x)\}\{u_0(y)-\bar{u}_0(y)\}^T]=P_0(x,y)\end{aligned}\right\}\quad (2\text{-}6)$$

where $E[\cdot]$ denotes an expectation operator and $\delta(\cdot)$ is the Dirac delta function. It is assumed that $Q(x,y,t)$ is symmetric, positive semidefinite, and bounded for all $x,y\epsilon D$, and that $R(t)$ is symmetric and positive definite.

Now, it is assumed that on a practical standpoint, the continuous-time observation process (2-2) is rewritten by the discrete-time one:

$$z(t_i)=H(t_i;\theta)u_m(t_i)+\zeta(t_i),\ i=0,1,\cdots,k,\cdots \quad (2\text{-}7)$$

where t_i denotes the sampling time at iteration number i. Then, given a record of measured data $Y_k=\{z(t_l);l\epsilon[0,1,\cdots,k]\}$ under the parameter uncertainty $\theta\epsilon\Omega_\theta$, an optimal conditional estimate $\hat{u}(x,t_k/t_k)\triangleq E[u(x,t_k)/Y_k]$ is desired in the mean-squared error sense:

$$\mathrm{MIN}\{E\|u(x,t_k)-\hat{u}(x,t_k/t_k)\|^2\}\ \text{for all}\ x\epsilon\bar{D} \quad (2\text{-}8)$$

where $\|\cdot\|$ represents a Euclidean norm.

3 Discrete-Time Model

In general, the aforementioned problem can be analyzed as a continuous-time model with discrete-time observation [12], but in this paper, in order to realize a recursive distributed filter by using the digital computer, a discrete-time model for the system (2-1) will be considered.

To do so, the following conditions are further assumed:

(*a*) The problem given by (2-1), (2-4), and (2-5) is well-posed in the sense of Hadamard, i.e., the solution exists uniquely and depends continuously on the initial and boundary data.

(*b*) There exists a fundamental solution[1] (or, Green's kernel function) $n\times n$ matrix $\Phi(x,t;y,\tau,\theta)$ of equations (2-1) and (2-4) defined for $t\geq\tau\geq0$, $x,y\epsilon\bar{D}$ and $\theta\epsilon\Omega_\theta$ such that

$$\left.\begin{aligned}&\Phi(x,t;y,\tau,\theta)=\Phi(y,t;x,\tau,\theta)\\&\frac{\partial\Phi(x,t;y,\tau,\theta)}{\partial t}=\mathcal{L}_x(\theta)\Phi(x,t;y,\tau,\theta)\\&\Phi(x,\tau;y,\tau,\theta)=I\delta(x-y)\\&\beta_\xi(\theta)\Phi(\xi,t;y,\tau,\theta)=0,\ \xi\epsilon\partial D\end{aligned}\right\}\quad (3\text{-}1a)$$

where I indicates the $n\times n$ identity matrix and the solution of equations (2-1), (2-4), and (2-5) is expressed as [33]

$$u(x,t)=\int_D\Phi(x,t;y,t_0,\theta)u(y,t_0)dy+\int_D\int_{t_0}^{t}\Phi(x,t;y,\tau,\theta)G(y,\tau;\theta)\eta(y,\tau)d\tau dy;\ (x,t)\epsilon\Sigma_s. \quad (3\text{-}1b)$$

(*c*) Between sampling instants, the optimal estimate in the sense of equation (2-8) satisfies the unforced system (2-1).

If the sampling interval $\Delta t=t_{k+1}-t_k$ is sufficiently small, then equation (3-1*b*) for $t=t_{k+1}$ may be rewritten as

$$u(x,t_{k+1})=\int_D\Phi(x,t_{k+1};y,t_k,\theta)u(y,t_k)dy+\int_D\int_{t_k}^{t_{k+1}}\Phi(x,t_{k+1};y,\tau,\theta)G(y,\tau,\theta)\eta(y,\tau)d\tau dy$$

[1] For mathematically more rigorous discussions, see [31-33].

$$\simeq \int_D \Phi(x,t_{k+1};y,t_k,\theta)u(y,t_k)dy$$

$$+\int_D \Phi(x,t_{k+1};y,t_k,\theta)G(y,t_k,\theta)\int_{t_k}^{t_{k+1}}\eta(y,\tau)d\tau dy, x,y\epsilon\bar{D}. \tag{3-2}$$

Applying the same procedure to the measurement system, equation (2-2) for $t=t_k$ becomes

$$z(t_k)=\frac{1}{t_k-t_{k-1}}\int_{t_{k-1}}^{t_k}[H(\tau;\theta)u_m(\tau)+\zeta(\tau)]d\tau$$

$$\simeq H(t_k;\theta)u_m(t_k)+\frac{1}{t_k-t_{k-1}}\int_{t_{k-1}}^{t_k}\zeta(\tau)d\tau. \tag{3-3}$$

Hence, the continuous-time system can be approximately expressed by the following discrete-time integral equations:

$$u_{k+1}(x)=A_x(\theta)u_k(x)+B_x(\theta)w_k(x), x\epsilon\bar{D} \tag{3-4}$$

$$z_k=H_k(\theta)u_{km}+v_k \tag{3-5}$$

$$\beta_\xi(\theta)u_{k+1}(\xi)=0, \xi\epsilon\partial D \tag{3-6}$$

where

$$u_k(x)\triangleq u(x,t_k), w_k(x)\triangleq\int_{t_k}^{t_{k+1}}\eta(x,\tau)d\tau$$

$$v_k\triangleq\frac{1}{t_k-t_{k-1}}\int_{t_{k-1}}^{t_k}\zeta(\tau)d\tau, H_k(\theta)\triangleq H(t_k,\theta) \tag{3-7}$$

$$u_{km}\triangleq u_m(t_k)$$

and the integral operators $A_x(\theta)$ and $B_x(\theta)$ are, respectively, given by

$$A_x(\theta)[\bullet]=\int_D \Phi(x,t_{k+1};y,t_k,\theta)[\bullet]dy, x,y\epsilon\bar{D} \tag{3-8}$$

$$B_x(\theta)[\bullet]=\int_D \Phi(x,t_{k+1};y,t_k,\theta)G(y,t_k;\theta)[\bullet]dy, x,y\epsilon\bar{D}. \tag{3-9}$$

The statistics of white Gaussian noise processes $w_k(x)$ and v_k, and the initial state $u_{k_0}(x)$ with respect to discrete-time are, respectively, given by

$$\left.\begin{aligned}&E[w_k(x)]=E[v_k]=E[w_k(x)v_k^T]\\&\qquad=E[u_{k_0}(x)w_k^T(x)]=E[u_{k_0}(x)v_k^T]=0\\&E[w_k(x)w_j^T(y)]=Q_k(x,y)\delta_{kj}, E[v_kv_j^T]=R_k\delta_{kj}\end{aligned}\right\} \tag{3-10}$$

$$\left.\begin{aligned}&E[u_{k_0}(x)]=\bar{u}_{k_0}(x)\\&E[(u_{k_0}(x)-\bar{u}_{k_0}(x))(u_{k_0}(y)-\bar{u}_{k_0}(y))^T]=P_{k_0}(x,y)\end{aligned}\right\} \tag{3-11}$$

where subscript k_0 represents the initial point in discrete-time case, $Q_k(x,y)$ is symmetric, positive semidefinite and bounded for all $x,y\epsilon D,R_k$ is symmetric positive definite and δ_{kj} is the Kronecker delta function. The discrete-time nuclear functions $Q_k(x,y)$ and R_k are related to those for continuous-time case as follows:

$$Q_k(x,y)=\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}E[\eta(x,t)\eta^T(y,\tau)]dtd\tau$$

$$=\int_{t_k}^{t_{k+1}}Q(x,y,\tau)d\tau$$

$$\simeq\Delta tQ(x,y,t_k) \tag{3-12}$$

$$R_k=\frac{1}{(\Delta t)^2}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k}E[\zeta(t)\zeta^T(\tau)]dtd\tau$$

$$=\frac{1}{(\Delta t)^2}\int_{t_{k-1}}^{t_k}R(\tau)d\tau$$

$$\simeq\frac{1}{\Delta t}R(t_k). \tag{3-13}$$

4 Wiener-Hopf Theory and Innovation Theorem

In the following derivation, it is assumed for simplicity of the problem that the time-invariant unknown parameter θ is completely known, namely, "θ" is omitted. Then, the least squares estimates for a discrete-time distributed parameter system with pointwise observation will be obtained by using the innovation theorem [15].

In the pointwise observation system (3-5), the sample function of zero-mean signal process is defined as

$$X_{km}\triangleq H_ku_{km} \tag{4-1}$$

and it is assumed to have a finite variance:

$$E[\|X_{km}\|^2]<\infty, k_0\le k<k_2 \tag{4-2}$$

and to be uncorrelated with v_k:

$$E[v_kX_{lm}^T]=0, k_0\le l<k<k_2. \tag{4-3}$$

If a set of measured data $Y_{k_2}=\{z_k; k\epsilon[k_0,\cdots,k_2]\}$ is given, the conditional mean value:

$$\hat{u}_{k_1/k_2}(x)=E[u_{k_1}(x)/Y_{k_2}], k_1\le k_2 \text{ or } k_1>k_2 \tag{4-4}$$

can be obtained by minimizing the following performance criterion:

$$J=E\|\tilde{u}_{k_1/k_2}(x)\|^2 \tag{4-5}$$

where

$$\tilde{u}_{k_1/k_2}(x)\triangleq u_{k_1}(x)-\hat{u}_{k_1/k_2}(x) \tag{4-6}$$

$$\hat{u}_{k_1/k_2}(x)\triangleq\sum_{j=k_0}^{k_2}\bar{N}(x,k_1j)z_j, x\epsilon\bar{D} \tag{4-7}$$

where $\bar{N}(\bullet)$ is an $n\times p$ matrix kernel function whose elements are continuously differentiable in x. Note that the least squares estimate for the system (3-4) – (3-6) is represented as a linear combination of a kernel function $\bar{N}(\bullet)$ and all the data record Y_{k_2} from the facts of [10-12]. Thus, the following two theorems can be deduced from the result in the continuous-time distributed parameter system with pointwise observation [11].

Theorem 1. The Projection Theorem and Wiener-Hopf Equation. The necessary and sufficient condition for the estimate (4-7) to be optimal is that the Wiener-Hopf equation:

$$\sum_{j=k_0}^{k_2}\bar{N}(x,k_1j)E[z_jz_l^T]=E[u_{k_1}(x)z_l^T] \tag{4-8}$$

holds for $k_0\le l<k_2$, $x\epsilon\bar{D}$. The foregoing equation is also equivalent to the following projection theorem:

$$E[\tilde{u}_{k_1/k_2}(x)z_l^T]=0, k_0\le l<k_2, x\epsilon\bar{D}. \tag{4-9}$$

Proof: This theorem can be proved by using the calculus of variations [10,25]. That is, the weak variation with respect to kernel function $\bar{N}(\bullet)$, $\delta\bar{N}=\epsilon M$, is taken in equation (4-7), where $M(x,k_1,j)$ is any $n\times p$ matrix kernel function and ϵ is a scalar parameter. If

$$\hat{u}_{\epsilon,k_1/k_2}(x)\triangleq\sum_{j=k_0}^{k_2}[\bar{N}(x,k_1j)+\epsilon M(x,k_1j)]z_j \tag{4-10}$$

then the necessary and sufficient condition for the estimate given by (4-8) to be optimal is that

$$\frac{dJ_\epsilon}{d\epsilon}\Big|_{\epsilon=0} \triangleq \frac{d\{E\|u_{k_1}(x)-\hat{u}_{\epsilon,k_1/k_2}(x)\|^2}{d\epsilon}\Big|_{\epsilon=0}=0. \tag{4-11}$$

After carrying out the indicated operation, it is found that

$$\frac{dJ_\epsilon}{d\epsilon}\Big|_{\epsilon=0}=2E\Big[\tilde{u}^T_{k_1/k_2}(x)\sum_{j=k_0}^{k_2}M(x,k_1,j)z_j\Big]=0 \tag{4-12}$$

or, equivalently,

$$tr\Big[\sum_{j=k_0}^{k_2}E\{\tilde{u}_{k_1/k_2}(x)z_j^T\}M^T(x,k_1,j)\Big]=0 \tag{4-13}$$

where "tr" denotes the trace operation of a matrix. If $M(x,k_1,j)=E[\tilde{u}_{k_1/k_2}(x)z_j^T]$ is assumed, equation (4-9) may be obtained as the necessary condition. On the other hand, the sufficiency of equation (4-9) is apparent from equation (4-13).

Theorem 2. The Innovation Theorem. The innovation process for the discrete-time distributed parameter system with pointwise observation is defined as:

$$\nu_k \triangleq z_k-\hat{X}_{k/k-1,m},\ -\infty<k_1\leq k<k_2<\infty \tag{4-14}$$

where

$$\hat{X}_{k/k-1,m}\triangleq H_k\hat{u}_{k/k-1,m}. \tag{4-15}$$

Then, ν_k is zero-mean and is a white but with a variance different from that for the measurement process.
Proof: This proof is given by Kailath [15] and hence it is omitted here.

It is noted from the use of previous Theorem 1 that the estimation error $\tilde{u}_{k_1/k_2}(x)$, $x\epsilon\bar{D}$ and the pointwise observation sequence z_l for $k_0\leq l<k_2$ are orthogonal. Moreover, since z_k and v_e are stochastically equivalent from Theorem 2, it is obvious that $\tilde{u}_{k_1/k_2}(x)$ and v_l for $k_0\leq l<k_2$ must be orthogonal. Hence, the optimal estimator equation (4-7) can be rewritten as a linear function of ν_k as

$$\hat{u}_{k_1/k_2}(x)=\sum_{j=k_0}^{k_2}\bar{N}(x,k_1,j)\nu_j, x\epsilon\bar{D}. \tag{4-16}$$

From the previous discussions and equations (4-16), the Wiener-Hopf equation (4-8) can be rewritten as

$$E[\tilde{u}_{k_1/k_2}(x)\nu_l^T]=\text{Cov}[u_{k_1}(x)-\hat{u}_{k_1/k_2}(x),\nu_l]$$

$$=E[u_{k_1}(x)\nu_l^T]-\sum_{j=k_0}^{k_2}\bar{N}(x,k_1,j)E[\nu_j\nu_l^T]=0 \tag{4-17}$$

or, equivalently,

$$E[u_{k_1}(x)\nu_l^T]=\sum_{j=k_0}^{k_2}\bar{N}(x,k_1,j)E[\nu_j\nu_l^T],$$

$$k_0\leq l<k_2,\ x\epsilon\bar{D} \tag{4-18}$$

where Cov[•,•] indicates the covariance operator. Since the covariance of the innovation process in the finite dimensional space is given by [15] as

$$E[\nu_j\nu_l^T]=[R_l+H_lP_{l/l-1}H_l^T]\delta_{jl} \tag{4-19}$$

Equation (4-18) yields the following relation:

$$E[u_{k_1}(x)\nu_l^T]=\bar{N}(x,k_1,l)[R_l+H_lP_{l/l-1}H_l^T] \tag{4-20}$$

where

$$P_{l/l-1}\triangleq E[\tilde{u}_{l/l-1,m}\tilde{u}^T_{l/l-1,m}]:nm\times nm \text{ matrix} \tag{4-21}$$

$$=\begin{bmatrix}P_{l/l-1}(x^1,x^1), & \ldots, & P_{l/l-1}(x^1,x^m)\\ \vdots & & \vdots\\ P_{l/l-1}(x^m,x^l), & \ldots, & P_{l/l-1}(x^m,x^m)\end{bmatrix}, x^i,\ i=1,2,\ldots,m\epsilon\bar{D}$$

$$P_{l/l-1}(x,y)\triangleq E[\tilde{u}_{l/l-1}(x)\tilde{u}^T_{l/l=1}(y)], x,y\epsilon\bar{D}.$$

Hence, using equations (4-16) and (4-20) gives the general basic formula for the optimal estimation with pointwise observation as

$$\hat{u}_{k_1/k_2}(x)=\sum_{j=k_0}^{k_2}E[u_{k_1}(x)\nu_j^T][R_j+H_jP_{j/j-1}H_j^T]^{-1}\nu_j, x\epsilon\bar{D}. \tag{4-22}$$

Premultiplying equation (4-18) by β_ξ and using the fact of equation (3-6) gives

$$E[\beta_\xi u_{k_1}(\xi)\nu_l^T]=\sum_{j=k_0}^{k_2}\beta_\xi\bar{N}(\xi,k_1,j)E[\nu_j\nu_l^T]$$

$$=0. \tag{4-23}$$

Since $E[\nu_j\nu_l^T]$ is not identically zero, it follows that

$$\beta_\xi\bar{N}(\xi,k_1,j)=0,\ \xi\epsilon\partial D. \tag{4-24}$$

Henceforth, using equations (4-16) and (4-24) gives the following boundary condition:

$$\beta_\xi\hat{u}_{k_1/k_2}(\xi)=0,\ \xi\epsilon\partial D. \tag{4-25}$$

5 Optimal Filter With Pointwise Observation

In this section, the filter mechanism with pointwise observation in discrete-time will be approached by the innovation theorem. Namely, putting $k_1=k_2=k$ in equation (4-22) gives

$$\hat{u}_{k/k}(x)=\sum_{j=k_0}^{k}E[u_k(x)\nu_j^T][R_j+H_jP_{j/j-1}H_j^T]^{-1}\nu_j. \tag{5-1}$$

Using equation (4-4), equation (5-1) can be rewritten as follows,

$$\hat{u}_{k/k}(x)=E[u_k(x)/Y_k]$$

$$=E[u_k(x)/Y_{k-1},z_k]$$

$$=E[u_k(x)/Y_{k-1}]+[u_k(x)/\nu_k] \tag{5-2}$$

where Theorem 2 and the orthogonal property between the set of measured data Y_k and pointwise observation z_k have been employed in the last equality in equation (5-2). Since the first term on the right-hand side of the last equality in equation (5-2) is apparently a one-stage prediction, the independence of $w_{k-1}(x)$ and Y_{K-1} gives

$$\hat{u}_{k/k-1}(x)\triangleq E[u_k(x)/Y_{k-1}]$$

$$=E[A_xu_{k-1}(x)+B_xw_{k-1}(x)/Y_{k-1}]$$

$$=A_x\hat{u}_{k-1/k-1}(x). \tag{5-3}$$

Using the relation of equations (5-2) and (5-3), equation (5-1) may reduce to

$$\hat{u}_{k/k}(x)=A_x\hat{u}_{k-1/k-1}(x)$$

$$+E[u_k(x)\nu_k^T][R_k+H_kP_{k/k-1}H_k^T]^{-1}\nu_k. \tag{5-4}$$

Furthermore, when the independence of $u_k(x)$ and v_k, and the orthogonal property between $\hat{u}_{k/k-1}(x)$ and $\tilde{u}_{k/k-1,m}$ have been employed, then

$$E[u_k(x)\nu_k^T]=\text{Cov}[u_k(x),H_k\tilde{u}_{k/k-1,m}+v_k]$$

$$=E[u_k(x)\tilde{u}^T_{k/k=1,m}]H_k^T$$

$$=E[\tilde{u}_{k/k-1}(x)\tilde{u}^T_{k/k-1,m}]H_k^T$$

$$=P_{k/k-1,m}(x)H_k^T \tag{5-5}$$

where

$$\tilde{u}_{k/k-1,m}\triangleq u_{k,m}-\hat{u}_{k/k-1,m} \tag{5-6}$$

$$P_{k/k-1,m}(x) \triangleq [P_{k/k-1}(x,x^1), P_{k/k-1}(x,x^2),$$

$$\ldots, P_{k/k-1}(x,x^m)] \tag{5-7}$$

$$P_{k/k-1}(x,y) \triangleq E[\tilde{u}_{k/k-1}(x)\tilde{u}^T_{k/k-1}(y)], x,y\epsilon\bar{D}. \tag{5-8}$$

Hence, the following mechanism, the optimal filter with pointwise observation, is given by

$$\hat{u}_{k/k}(x) = A_x\hat{u}_{k-1/k-1}(x) + P_{k/k-1,m}(x)H_k^T[R_k + H_kP_{k/k-1}H_k^T]^{-1}\nu_k, x\epsilon\bar{D}. \tag{5-9}$$

with the boundary condition

$$\beta_\xi\hat{u}_{k/k}(\xi) = 0,\ \xi\epsilon\partial D. \tag{5-10}$$

Next, the error covariance equations will be discussed. First, from equations (3-4) and (5-3), the one-stage prediction error is given as

$$\tilde{u}_{k/k-1}(x) \triangleq u_k(x) - \hat{u}_{k/k-1}(x) = A_x\tilde{u}_{k-1/k-1}(x) + B_xw_{k-1}(x) \tag{5-11}$$

with the boundary condition

$$\beta_\xi\tilde{u}_{k/k-1}(\xi) = 0,\ \xi\epsilon\partial D. \tag{5-12}$$

Then, the one-stage prediction error covariance is obtained as

$$\begin{aligned} P_{k/k-1}(x,y) &= E[\tilde{u}_{k/k-1}(x)\tilde{u}^T_{k/k-1}(y)] \\ &= \mathrm{Cov}[A_x\tilde{u}_{k-1/k-1}(x) + B_xw_{k-1}(x), \\ &\quad A_y\tilde{u}_{k-1/k-1}(y) + B_yw_{k-1}(y)] \\ &= A_xP_{k-1/k-1}(x,y)A_y^T \\ &\quad + B_xQ_{k-1}(x,y)B_y^T \end{aligned} \tag{5-13}$$

where the independence of $\tilde{u}_{k-1/k-1}(x)$ and $w_{k-1}(x)$ is utilized. Premultiplying the first equality of equation (5-13) by β_ξ and using equation (5-12), the boundary condition is derived as

$$\beta_\xi P_{k/k-1}(\xi,y) = 0,\ \xi\epsilon\partial D, y\epsilon\bar{D}. \tag{5-14}$$

On the other hand, the filtering error equation from equation (5-9) is derived as

$$\begin{aligned} \tilde{u}_{k/k}(x) &\triangleq u_k(x) - \hat{u}_{k/k}(x) \\ &= u_k(x) - \hat{u}_{k/k-1}(x) - K_{k,m}(x)\nu_k \\ &= \tilde{u}_{k/k-1}(x) - K_{k,m}(x)\nu_k \end{aligned} \tag{5-15}$$

where

$$K_{k,m}(x) \triangleq P_{k/k-1,m}(x)H_k^T[R_k + H_kP_{k/k-1}H_k^T]^{-1}. \tag{5-16}$$

Using equations (3-6) and (5-10), the boundary condition for equation (5-15) is obtained as

$$\beta_\xi\tilde{u}_{k/k}(\xi) = 0, \xi\epsilon\partial D. \tag{5-17}$$

Therefore, using equations (4-17), (4-19), (5-15), and (5-16), the filtering error covariance is given as

$$\begin{aligned} P_{k/k}(x,y) &\triangleq E[\tilde{u}_{k/k}(x)\tilde{u}^T_{k/k}(y)] \\ &= \mathrm{Cov}[\tilde{u}_{k/k-1}(x) - K_{k,m}(x)\nu_k, \tilde{u}_{k/k-1}(y) \\ &\quad - K_{k,m}(y)\nu_k] \\ &= P_{k/k-1}(x,y) + K_{k,m}(x)E[\nu_k\nu_k^T]K^T_{k,m}(y) \\ &\quad - K_{k,m}(x)E[\nu_k\tilde{u}^T_{k/k-1}(y)] \\ &\quad - E[\tilde{u}_{k/k-1}(x)\nu_k^T]K^T_{k,m}(y) \\ &= P_{k/k-1}(x,y) - K_{k,m}(x)H_kP^T_{k/k-1,m}(y) \end{aligned} \tag{5-18}$$

and premultiplying the first equality of equation (5-18) by β_ξ and using equation (5-17), the boundary condition reduces to

$$\beta_\xi P_{k/k}(\xi,y) = 0,\ \xi\epsilon\partial D, y\epsilon\bar{D}. \tag{5-19}$$

Thus, the optimal filter is given by equations (5-3), (5-9), and (5-10), and the corresponding error covariance satisfies equations (5-13), (5-14), (5-16), (5-18), and (5-19). Since one cannot make continuous measurements in space on a practical situation [34], it is noted that the algorithms for the case of pointwise observation derived in this paper are more practical than those for the case of distributed observation studied in Tzafestas [5,6,20].

6 Derivation of an Adaptive Distributed Filter

Viewing the unknown time-invariant parameter as a point of a finite-dimensional vector space Ω_θ having dimension equal to the number of unknown parameters, Magill [21] developed an optimal estimator for a lumped parameter system, but the derived algorithm has useless memory allocations, and therefore several adaptive techniques [13,14, 16–19] which have less memory and computational requirements compared with that of Magill will be adopted in this paper.

Suppose that $\{u_k(x), z_k\}$ comes from a finite number collection of possible processes, in which the state and measurement spaces are defined on $L^2(D)$ and $\mathbf{E}^p$, respectively, with an unknown time and space-invariant parameter vector θ. Then the optimal conditional mean value (4-4) can be derived by using the smoothing property of conditional expectation as

$$\hat{u}_{k/k}(x) = E[E[u_k(x)/Y_k,\theta]/Y_k], x\epsilon\bar{D}. \tag{6-1}$$

From the assumptions for the unknown parameter θ which have been stated in Section 2, the a priori probability for θ can be expressed as

$$P_r(\theta) = \sum_{i=1}^{M} P_r(\theta_i)\delta(\theta-\theta_i), \theta_i\epsilon\Omega_\theta \tag{6-2}$$

where M denotes the event number for the unknown random constant parameter θ, $P_r(\theta)$ is a priori probability for θ, and $P_r(\theta_i)$ is the a priori probability for the event where θ takes a parameter $\theta_i\epsilon\Omega_\theta$ from the finite discrete parameter set $\Omega_\theta = \{\theta_1, \theta_2, \cdots, \theta_M\}$.

If the a posteriori probabilities for any parameter $\theta_i\epsilon\Omega_\theta$ are able to be calculated, then the optimal conditional mean value (4-4) may be expressed as

$$\hat{u}_{k/k}(x) = \sum_{i=1}^{M} \hat{u}_{k/k}(x;\theta_i)P_r(\theta_i/Y_k), x\epsilon\bar{D}, \theta_i\epsilon\Omega_\theta \tag{6-3}$$

$$\beta_\xi\hat{u}_{k/k}(\xi) = \sum_{i=1}^{M}\beta_\xi\hat{u}_{k/k}(\xi;\theta_i)P_r(\theta_i/Y_k) = 0, \xi\epsilon\partial D \tag{6-4}$$

where θ-conditioned mean value,

$$\hat{u}_{k/k}(x;\theta_i) \triangleq E[u_k(x)/Y_k,\theta_i],\ x\epsilon\bar{D} \tag{6-5}$$

$$\beta_\xi\hat{u}_{k/k}(\xi;\theta_i) = 0,\ \xi\epsilon\partial D \tag{6-6}$$

will be obtained by using the previous discrete-time distributed Kalman filter matched to the system with specified parameter θ_i. The conditional a posteriori probability is calculated by applying a recursive Bayesian algorithm [16] as

$$\begin{aligned} P_r(\theta_i/Y_k) &= \frac{p(Y_k,\theta_i)}{p(Y_k)} \\ &= \frac{p(z_k,Y_{k-1},\theta_i)}{p(z_k,Y_{k-1})} = \frac{p(z_k,\theta_i/Y_{k-1})p(Y_{k-1})}{p(z_k/Y_{k-1})p(Y_{k-1})} \\ &= \frac{p(z_k,\theta_i/Y_{k-1})}{p(z_k/Y_{k-1})} \end{aligned}$$

$$= \frac{p(z_k/Y_{k-1},\theta_i)P_r(\theta_i/Y_{k-1})}{\sum_{j=1}^{M} p(z_k/Y_{k-1},\theta_j)P_r(\theta_j/Y_{k-1})}. \quad (6\text{-}7)$$

Here, it is noted that the conditional probability density $p(z_k/Y_{k-1},\theta_i)$ must be calculated in the functional space $L^2(D)$ of infinite dimensional state if the distributed observation over $\bar{D}$ is concerned ideally. However, in the case of the pointwise observation as discussed here, the observation space reduces to a finite one of Euclidean space $\mathbf{E}^p$ so that the conditional probability density $p(z_k/Y_{k-1},\theta_i)$ can be evaluated by using the similar derivation to lumped parameter systems [16].

Since the probabilities $p(z_k/Y_{k-1},\theta_i)$ for $i=1,2,\cdots, M$ are Gaussian ones, the associated one-stage prediction is given by

$$E[z_k/Y_{k-1},\theta_i]=H_k(\theta_i)\hat{u}_{k/k-1,m}(\theta_i) \quad (6\text{-}8)$$

and the conditional covariance matrix for z_k is obtained by

$$\begin{aligned}\mathrm{Cov}[z_k/Y_{k-1},\theta_i]&=\mathrm{Cov}[\nu_k(\theta_i),\nu_k(\theta_i)]\\ &=R_k(\theta_i)+H_k(\theta_i)P_{k/k-1}(\theta_i)H_k^T(\theta_i) \quad (6\text{-}9)\end{aligned}$$

hence $p(z_k/Y_{k-1},\theta_i)$ reduces to

$$p(z_k/Y_{k-1},\theta_i)$$

$$=C\cdot\exp\left\{-\frac{1}{2}<\nu_k(\theta_i),\mathrm{Cov}[\nu_k(\theta_i),\nu_k(\theta_i)]^{-1}\nu_k(\theta_i)>\right\} \quad (6\text{-}10)$$

where the scale factor C is

$$C=(2\pi)^{-p/2}|\mathrm{Cov}[\nu_k(\theta_i),\nu_k(\theta_i)]|^{-1/2} \quad (6\text{-}11)$$

here $|(\cdot)|$ denotes the determinant of a matrix $(\cdot)$. Then, the optimal filtering error covariance is obtained in terms of θ-conditioned error covariance by the same way of deriving as used in the study of continuous-time cases [22,23]. That is, obtained results are as follows:

$$P_{k/k}(x,y)=E\{[u_k(x)-\hat{u}_{k/k}(x)][u_k(y)-\hat{u}_{k/k}(y)]^T/Y_k\}$$

$$=\sum_{i=1}^{M}P_{k/k,a}(x,y;\theta_i)P_r(\theta_i/Y_k) \quad (6\text{-}12)$$

$$P_{k/k,a}(x,y,;\theta_i)\triangleq P_{k/k}(x,y;\theta_i)+[\hat{u}_{k/k}(x;\theta_i)-\hat{u}_{k/k}(x)]\times[\hat{u}_{k/k}(y;\theta_i)-\hat{u}_{k/k}(y)]^T \quad (6\text{-}13)$$

$$\beta_\xi P_{k/k,a}(\xi,y;\theta_i)=0,\xi\epsilon\partial D,y\epsilon\bar{D} \quad (6\text{-}14)$$

$$\begin{aligned}\beta_\xi P_{k/k}(\xi,y)&=\sum_{i=1}^{M}\beta_\xi P_{k/k,a}(\xi,y;\theta_i)P_r(\theta_i/Y_k)\\ &=0,\xi\epsilon\partial D,y\epsilon\bar{D}. \quad (6\text{-}15)\end{aligned}$$

Thus, the optimal distributed filter in discrete-time can constitute the parallel distributed Kalman filter when the distributed parameter systems with pointwise observation contain some time and space-invariant unknown parameters. This filter consists of some elemental filter matched to the system with specified parameter θ_i, and the M-distributed Kalman filter is implemented to produce the estimates $\hat{u}_{k/k}(x;\theta_i)$ for the states $u_k(x;\theta_i)$, $x\epsilon D$, $\theta_i\epsilon\Omega_\theta$. The implementation of $\hat{u}_{k/k}(x;\theta_i)$ yields automatically as a byproduct θ_i-conditioned finite-dimensional innovation process $\nu_k(\theta_i)$ and the inverse of its covariance $\mathrm{Cov}[\nu_k(\theta_i), \nu_k(\theta_i)]^{-1}$. Moreover, it is interesting to note that the adaptive realization of the optimal nonlinear distributed parameter estimate is given in terms of the parameter θ-conditional estimates $\hat{u}_{k/k}(x;\theta_i)$ and the a posteriori parameter probabilities $P_r(\theta_i/Y_k)$, and that the a posteriori probability $P_r(\theta_i/Y_k)$ is given by the ratio of two likelihood ratios of finite dimensional state space in discrete-time.

7 Approximation to Finite-Dimensional Subspace Via the Eigenfunction Expansion

The transformation of the infinite-dimensional space to a finite-dimensional one is necessitated when the derived adaptive filter is implemented in the practical situation. Under the assumption that the fundamental solution matrix and other vectors and matrices are square integrable in each space, in this study the eigenfunction expansion technique, which is effective on the initial and boundary value problems, is introduced. For simplicity of the problem, the coefficients of differential operators $\mathcal{L}_x(\theta)$ and $\beta_\xi(\theta)$ are assumed to be independent of time t.

Now, it is assumed that there exists a sequence $\{\phi_i(x)\}, x\epsilon\bar{D}$ of eigenfunctions and a sequence $\{\lambda_i\}$ of eigenvalues such that

$$\mathcal{L}_x(\theta)\phi_i(x)=-\lambda_i\phi_i(x),x\epsilon\bar{D} \quad (7\text{-}1)$$

$$\lambda_1\le\lambda_2\le\cdots\to\infty \quad (7\text{-}2)$$

$$\beta_\xi(\theta)\phi_i(\xi)=0,\ \xi\epsilon\partial D \quad (7\text{-}3)$$

where $\|u(x,t)\|\to0$ as $t\to\infty$ if $\lambda_1>0$ and $\|u(x,t)\|\to\infty$ as $t\to\infty$ if $\lambda_1<0$ in the homogeneous equation of (2-1). Furthermore, $\{\phi_i(x)\}$ is assumed to be complete orthonormal system in $L^2(D)$ as

$$\int_D\phi_i^T(x)\phi_j(x)dx=\delta_{ij},\ i,j=1,2,\ \ldots \quad (7\text{-}4)$$

and the assumption of completeness yields

$$\hat{u}_{k/k}(x;\theta)=\sum_{i=1}^{\infty}\hat{u}_i(k/k;\theta)\phi_i(x) \quad (7\text{-}5)$$

$$\hat{u}_{k/k-1}(x;\theta)=\sum_{i=1}^{\infty}\hat{u}_i(k/k-1;\theta)\phi_i(x) \quad (7\text{-}6)$$

$$\Phi(x,k+1;y,k;\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}(\theta)\phi_i(x)\phi_j^T(y) \quad (7\text{-}7)$$

$$P_{k/k}(x,y;\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{ij}(k/k;\theta)\phi_i(x)\phi_j^T(y) \quad (7\text{-}8)$$

$$P_{k/k-1}(x,y;\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{ij}(k/k-1;\theta)\phi_i(x)\phi_j^T(y) \quad (7\text{-}9)$$

$$P_{k/k-1,m}(x;\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{ij}(k/k-1;\theta)\phi_i(x)\Phi_j^T \quad (7\text{-}10)$$

where the coefficient functions, for example, $\hat{u}_i(k/k;\theta)$ and $a_{ij}(\theta)$ are, respectively, given by

$$\hat{u}_i(k/k;\theta)=\int_D\phi_i^T(x)\hat{u}_{k/k}(x;\theta)dx \quad (7\text{-}11a)$$

$$a_{ij}(\theta)=\int_D\int_D\phi_i^T(x)\Phi(x,k+1;y,k,\theta)\phi_j(y)dxdy \quad (7\text{-}11b)$$

and

$$\Phi_j^T=[\phi_j^T(x^1),\ \ldots\ ,\phi_j^T(x^m)] \quad (7\text{-}12)$$

substituting equations (3-8) and (7-5) into equation (5-3) and using the orthonormality condition (7-4) yields

$$\hat{u}_i(k/k-1;\theta)=a_{ii}(\theta)\hat{u}_i(k-1/k-1;\theta). \quad (7\text{-}13)$$

It is noted from equations (4-21) and (7-12) that

$$H_k(\theta)P_{k/k-1}(\theta)H_k^T(\theta)$$

$$=\sum_{i=1}^{\infty}\sum_{l=1}^{\infty}P_{il}(k/k-1;\theta)H_k(\theta)\Phi_i\Phi_l^TH_k^T(\theta) \tag{7-14}$$

and the innovation sequence (4-14) can be rewritten as

$$\nu_k(\theta)=z_k-H_k(\theta)\sum_{l=1}^{\infty}\hat{u}_l(k/k-1;\theta)\Phi_l. \tag{7-15}$$

Hence, $\hat{u}_i(k/k;\theta)$, $i=1,2,\cdots$ are given by

$$\hat{u}_i(k/k;\theta)=\hat{u}_i(k/k-1;\theta)$$

$$+\sum_{j=1}^{\infty}P_{ij}(k/k-1;\theta)\Phi_j^TH_k^T(\theta)$$

$$\times[R_k+\sum_{h=1}^{\infty}\sum_{l=1}^{\infty}P_{hl}(k/k-1;\theta)H_k(\theta)\Phi_h\Phi_l^TH_k^T(\theta)]^{-1}$$

$$\times\left[z_k-H_k(\theta)\sum_{l=1}^{\infty}\hat{u}_l(k/k-1;\theta)\Phi_l\right]. \tag{7-16}$$

Using equations (3-8), (3-9), (5-13) and (7-7)–(7-9), noting that

$$G_k(y;\theta)Q_k(y,x)G_k^T(x;\theta)$$

$$=\sum_{h=1}^{\infty}\sum_{l=1}^{\infty}q_{hl}(k;\theta)\phi_h(y)\phi_l^T(x) \tag{7-17}$$

$$G_k(y;\theta)\triangleq G(y,t_k;\theta)$$

and utilizing the orthonormality condition (7-4), then $P_{ij}(k+1/k;\theta)$, $i,j=1,2,\cdots$ become

$$P_{ij}(k+1/k;\theta)$$

$$=a_{ii}(\theta)[P_{ij}(k/k;\theta)+q_{ij}(k;\theta)]a_{jj}(\theta). \tag{7-18}$$

Defining the filter gain equation (5-16) as

$$K_{k,m}(x)\triangleq\sum_{i=1}^{\infty}\phi_i(x)K_i(k;\theta) \tag{7-19}$$

then its coefficient functions reduce to

$$K_i(k;\theta)=$$

$$\sum_{l=1}^{\infty}P_{il}(k/k-1;\theta)\Phi_l^TH_k^T(\theta)[R_k+\sum_{h=1}^{\infty}\sum_{h'=1}^{\infty}P_{hh'}(k/k-1;\theta)$$

$$H_k(\theta)\Phi_h\Phi_{h'}^TH_k^T(\theta)]^{-1} \tag{7-20}$$

where $K_i(k;\theta)$ is a p-dimensional row vector. The filter error covariance coefficient functions are derived by using equations (5-18), (7-9), (7-10) and (7-19) as

$$P_{ij}(k/k;\theta)=P_{ij}(k/k-1;\theta)$$

$$-\sum_{l=1}^{\infty}K_i(k;\theta)H_k(\theta)P_{jl}(k/k-1;\theta)\Phi_l. \tag{7-21}$$

By analogy with the foregoing results, the following expansions:

$$\hat{u}_{k/k}(x)=\sum_{i=1}^{\infty}\hat{u}_i(k/k)\phi_i(x) \tag{7-22}$$

$$P_{k/k,a}(x,y;\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{aij}(k/k;\theta)\phi_i(x)\phi_j^T(y) \tag{7-23}$$

$$P_{k/k}(x,y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{ij}(k/k)\phi_i(x)\phi_j^T(y) \tag{7-24}$$

and the following relations can be easily obtained:

$$P_{aij}(k/k;\theta)=P_{ij}(k/k;\theta)+[\hat{u}_i(k/k;\theta)$$

$$-\hat{u}_i(k/k)][\hat{u}_j(k/k;\theta)-\hat{u}_j(k/k)] \tag{7-25}$$

$$P_{ij}(k/k)=\sum_{l=1}^{M}P_{aij}(k/k;\theta_l)P_r(\theta_l/Y_k) \tag{7-26}$$

$$\hat{u}_i(k/k)=\sum_{l=1}^{M}\hat{u}_i(k/k;\theta_l)P_r(\theta_l/Y_k) \tag{7-27}$$

$$\hat{\theta}(k)=\sum_{l=1}^{M}\theta_lP_r(\theta_l/Y_k). \tag{7-28}$$

Approximating these equations by the first N terms based on the practical viewpoint, the approximated ordinary Kalman filter equations for those coefficient functions can be implemented. That is, the one-stage predictor for the series expanded elemental Kalman filter with θ_i is given as follows:

$$\hat{\mathbf{U}}(k/k-1;\theta_i)\triangleq\mathbf{A}(\theta_i)\hat{\mathbf{U}}(k-1/k-1;\theta_i) \tag{7-29}$$

$$\mathbf{P}(k/k-1;\theta_i)\triangleq\mathbf{A}(\theta_i)[\mathbf{P}(k-1/k-1;\theta_i)$$

$$+\mathbf{Q}(k-1;\theta_i)]\mathbf{A}(\theta_i) \tag{7-30}$$

where

$$\hat{\mathbf{U}}(k/k-1;\theta_i)\triangleq\mathrm{Col}[\hat{u}_1(k/k-1;\theta_i),\ \ldots\ ,\hat{u}_N(k/k-1;\theta_i)] \tag{7-31}$$

$$\mathbf{A}(\theta_i)\triangleq\mathrm{diag}[a_{11}(\theta_i),\ \ldots\ ,a_{NN}(\theta_i)] \tag{7-32}$$

$$\mathbf{P}(k/k-1;\theta_l)\triangleq\{P_{ij}(k/k-1;\theta_l\}i,j=1,2,\cdots,N \tag{7-33}$$

$$\mathbf{Q}(k-1;\theta_l)\triangleq\{q_{ij}(k-1;\theta_l)\}i,j=1,2,\cdots,N. \tag{7-34}$$

Defining the matrix

$$\Psi=[\Phi_1,\Phi_2,\ \ldots\ ,\Phi_N]:nm\times N \tag{7-35}$$

yields the following series expanded elemental Kalman filter with θ_i

$$\mathrm{Cov}[\nu_k(\theta_i),\nu_k(\theta_i)]\triangleq\bar{\mathbf{P}}(k/k-1;\theta_i):p\times p$$

$$=R_k+H_k(\theta_i)\Psi\mathbf{P}(k/k-1;\theta_i)\Psi^TH_k^T(\theta_i) \tag{7-36}$$

$$\mathbf{K}(k;\theta_i)=\mathbf{P}(k/k-1;\theta_i)\Psi^TH_k^T(\theta_i)$$

$$\bar{\mathbf{P}}(k/k-1;\theta_i)^{-1}:N\times p \tag{7-37}$$

$$\nu_k(\theta_i)=z_k-H_k(\theta_i)\Psi\hat{\mathbf{U}}(k/k-1;\theta_i):p\times 1 \tag{7-38}$$

$$\hat{\mathbf{U}}(k/k;\theta_i)=\hat{\mathbf{U}}(k/k-1;\theta_i)+\mathbf{K}(k;\theta_i)\nu_k(\theta_i):N\times 1 \tag{7-39}$$

$$\mathbf{P}(k/k;\theta_i)=[\tilde{I}-\mathbf{K}(k;\theta_i)H_k(\theta_i)\Psi]\mathbf{P}(k/k-1;\theta_i) \tag{7-40}$$

where $\tilde{I}$ denotes the $N\times N$ identity matrix. Then the optimal filter and the corresponding error covariance for the "weighted coefficient functions" reduce to

$$\hat{\mathbf{U}}(k/k)=\sum_{i=1}^{M}\hat{\mathbf{U}}(k/k;\theta_i)P_r(\theta_i/Y_k) \tag{7-41}$$

$$\mathbf{P}(k/k)=\sum_{i=1}^{M}\mathbf{P}_a(k/k;\theta_i)P_r(\theta_i/Y_k) \tag{7-42}$$

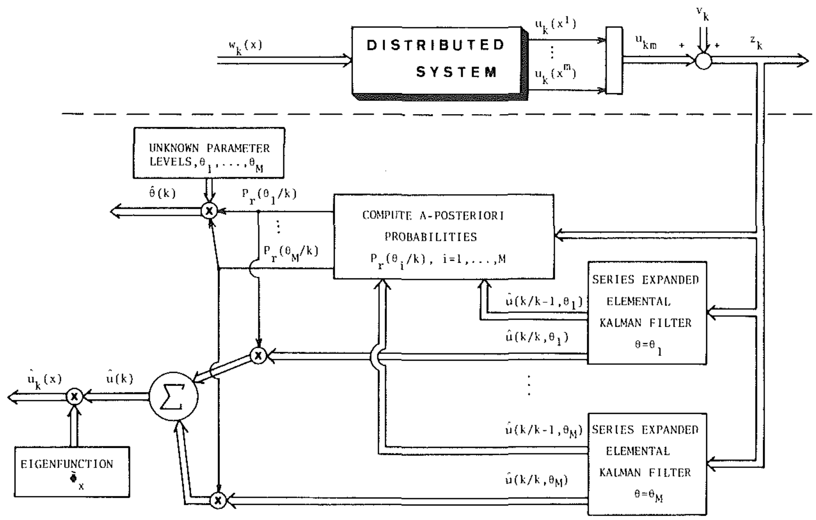

Fig. 1 Adaptive distributed-parameter filter structure

where

$$\mathbf{P}_a(k/k;\theta_i) = \mathbf{P}(k/k;\theta_i)$$

$$+[\hat{\mathbf{U}}(k/k;\theta_i) - \hat{\mathbf{U}}(k/k)][\hat{\mathbf{U}}(k/k;\theta_i) - \hat{\mathbf{U}}(k/k)]^T. \qquad (7\text{-}43)$$

Furthermore, if the matrix which represents the eigenfunction for the coordinate of any estimated location is defined as:

$$\bar{\Phi}_x \triangleq [\phi_1(x), \phi_2(x), \ldots, \phi_N(x)]: n \times N, x\epsilon\bar{D} \qquad (7\text{-}44)$$

after all, the state estimate and its covariance at any coordinate $x\epsilon\bar{D}$ are obtained by

$$\hat{u}_{k/k}(x) = \bar{\Phi}_x \hat{\mathbf{U}}(k/k), x\epsilon\bar{D} \qquad (7\text{-}45)$$

$$P_{k/k}(x,y) = \bar{\Phi}_x \mathbf{P}(k/k)\bar{\Phi}_y^T, \; x,y\epsilon\bar{D}. \qquad (7\text{-}46)$$

It is interesting to note that the filter algorithms obtained by equations (7-36)-(7-40) are similar to those of lumped parameter systems studied by Kalman [24], Kailath [15], and Meditch [25], etc. In addition, note that the present adaptive filtering method is a direct extension to distributed parameter systems of the partition theorem [13,14] given in Laintiotis' works. The proposed adaptive distributed parameter filter structure is also depicted in Fig. 1.

8 Illustrative Example

To illustrate the application of the proposed adaptive distributed filter, the estimation for the flux pattern in a slab type nuclear reactor [26] will be considered. Assume that the reactor is operating at a given neutron flux level with a steady state spatial pattern.

The diffusion equation to be treated is described by

$$\frac{\partial u(x,t)}{\partial t} = a\frac{\partial^2 u(x,t)}{\partial x^2} + cu(x,t) + w(x,t), \; 0<x<1 \qquad (8\text{-}1)$$

where the state $u(x,t)$ denotes the deviation from that in a steady state. The boundary conditions for this equation are the following Dirichlet type:

$$u(x,t)\,|_{x=0} = u(x,t)\,|_{x=1} = 0 \text{ for all } t. \qquad (8\text{-}2)$$

The conditions for the initial distribution are assumed that

$$u_{k_0}(x) = \bar{u}_{k_0}(x) = 0 \qquad (8\text{-}3)$$

$$P_{k_0}(x,y) = 0.5\delta(x-y). \qquad (8\text{-}4)$$

The diffusivity coefficient a and the absorption factor c are, respectively, given by

$$a = 0.0256, \; c = 0.252. \qquad (8\text{-}5)$$

The eigenfunctions for equations (8-1) and (8-2) are selected as

$$\phi_i(x) = \sqrt{2}\,\sin(i\pi x), \; i = 1,2, \ldots, N \qquad (8\text{-}6)$$

and then the fundamental solution reduces to[2]

$$\Phi(x,t_{k+1};y,t_k;\theta)$$

$$= 2\sum_{i=1}^{N} \exp\{-\lambda_i(\theta)\Delta t\}\sin(i\pi x)\sin(i\pi y) \qquad (8\text{-}7)$$

$$\lambda_i(\theta) = (i^2\pi^2 a - c), \theta^T = [a,c]. \qquad (8\text{-}8)$$

For this case, $a_{ii}(\theta)$ in equation (7-32) becomes

$$a_{ii}(\theta) = \exp\{-\lambda_i(\theta)\Delta t\}. \qquad (8\text{-}9)$$

The system noise covariance function for $w_k(x)$ is assumed to be

$$Q_k(x,y) = 0.04\delta(x-y) \qquad (8\text{-}10)$$

and therefore q_{ij} in equation (7-34) reduces to

$$q_{hl} = \int_0^1\int_0^1 0.04\delta(y'-x')\phi_h(y')\phi_l(x')dy'dx'$$

$$= \int_0^1 0.04\phi_h(y')\phi_l(y')dy'$$

$$= 0.04\delta_{hl}. \qquad (8\text{-}11)$$

The measurement of the flux level $u(x,t)$ is observed at several locations at $0<x<1$, but it is corrupted by the additive measurement noise.

9 Simulation Results and Discussion

To simulate the system (8-1) (8-3), nine node approximation is used and the coordinate division and the sampling interval are $\Delta x = 0.1$ and $\Delta t = 0.1$, respectively.

[2]If the boundary condition (8-2) is assumed to be the Neuman type, then equations (8-6)-(8-8) reduce to
$\phi_0(x) = 1$, $\phi_i(x) = \sqrt{2}\cos(i\pi x)$, $i = 1,2,\ldots,N$,

$$\Phi(x,t_{k+1};y,t_k;\theta)$$

$$= \exp\{c\Delta t\} + 2\sum_{i=1}^{N}\exp\{-\lambda_i(\theta)\Delta t\}\cos(i\pi x)\cos(i\pi y)$$

and $\lambda_i(\theta) = (i^2\pi^2 a - c)$.

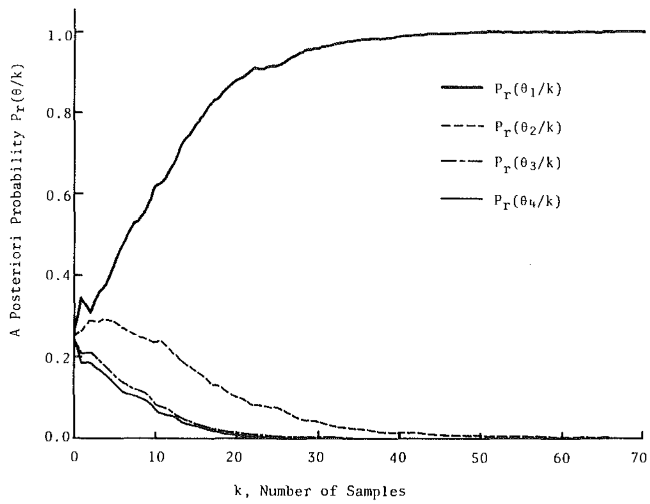

Fig. 2 A posteriori probabilities

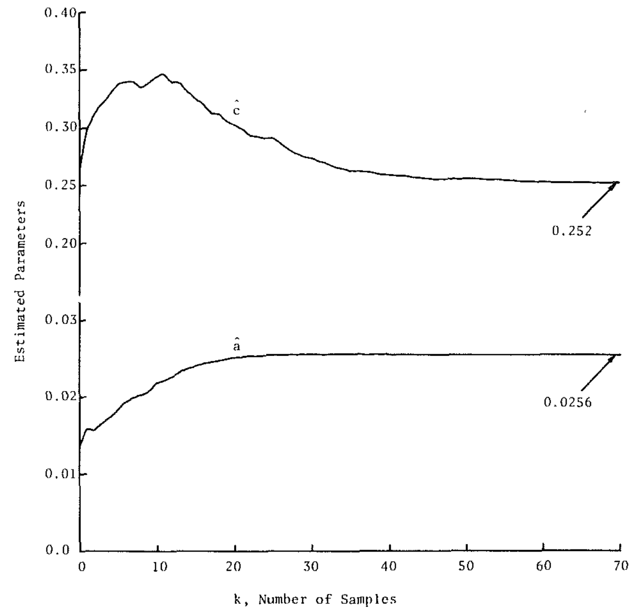

Fig. 3 Estimated parameter evolution

Table 1 Unknown parameter events

Filter	$\theta^T = [\quad a$	$c \quad]$	$P_r(\theta/0)$
1	$\theta_1^T = [0.0256$	$0.2520\ ]$	0.25
2	$\theta_2^T = [0.0256$	$0.8000\ ]$	0.25
3	$\theta_3^T = [0.0010$	$0.0098\ ]$	0.25
4	$\theta_4^T = [0.0010$	$0.0008\ ]$	0.25

(*i*) *Case of the Noise Statistics Are Known.* The following typical example will be considered to grasp the properties of the proposed adaptive distributed filter. The number of points measured is assumed as $m=3$ which is located at

$$x^1 = 0.3, x^2 = 0.5, x^3 = 0.7$$

and the number of points to be estimated and point locations are assumed to be identical. The measurement noise is assumed to be

$$R_k = \mathrm{diag}(0.01, 0.01, 0.01).$$

Four distributed Kalman filters given in Table 1 were used for estimating the unknown parameters, a and c, one of which matched the pure diffusion model. That is, Filter 1 in the bank of distributed Kalman filters has the true parameter values. Here, the truncation number of the expansion coefficients was applied as $N=4$.

The results after 100 samples are shown in Figs. 2–7. It is seen from Figs. 2 and 3 that the convergence was attained at about 60 sample points. Figures 4–6 illustrate the estimated flux pattern in this case, and the associated optimal error covariance is depicted in Fig. 7. Note that the quantization of unknown parameters, which is lying in the continuous parameter space, must be done so as to keep the first eigenvalue of equation (7-2) discussed in Section 7 within a positive real domain. However, in order to show a typical action of Filter 1 with the true parameters, Filter 2 was intentionally unstabilized. Indeed, quantizing the unknown parameters for the thermal problem as discussed in [2] may be an easier task than for a neutron one because in the former case the eigenvalue (8-8) is always a positive value due to replacing c in equation (8-1) by $-c$.

(*ii*) *Case of the Noise Statistics Are Unknown.* Hereafter, as a measure of the filter performance, the following two criteria will be introduced [27]. Namely,

$$J_1 = E\left\{ \frac{1}{m'} \left[\sum_{i=1}^{m'} u_k(x^i) - \sum_{i=1}^{m'} \hat{u}_k(x^i) \right]^2 \right\}$$

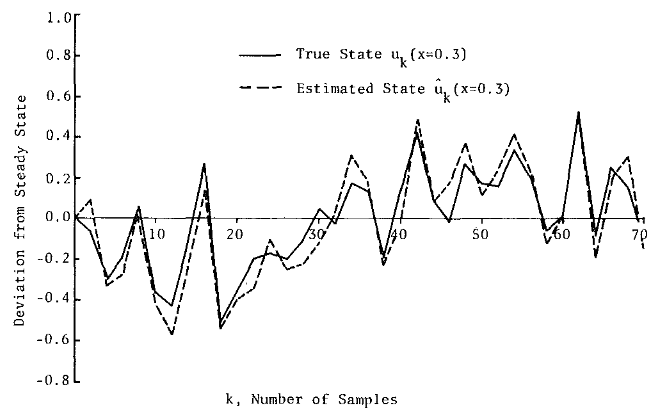

Fig. 4 Estimated flux pattern at $x=0.3$

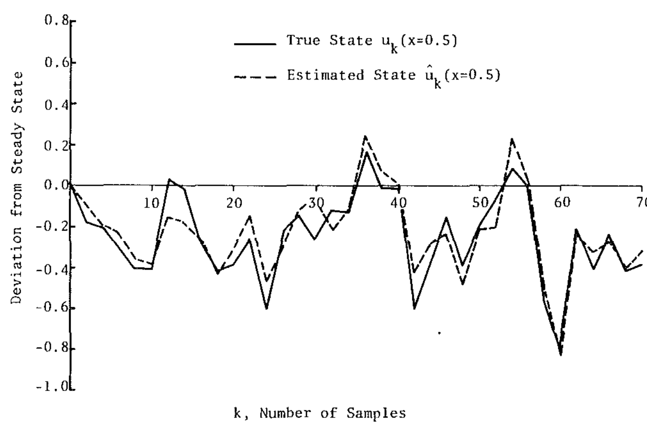

Fig. 5 Estimated flux pattern at $x=0.5$

$$J_2 = \frac{1}{m'} \sum_{i=1}^{m'} E[u_k(x^i) - \hat{u}_k(x^i)]^2$$

where J_1 expresses an estimate of the real variance of the average error and J_2 denotes an average of the mean-square errors between the actual states and the filtered ones. The

Table 2 Unknown parameter events

Filter	$\theta^T = [$ a	c	q	r]	$P_r(\theta/0)$
1	$\theta_1^T = [0.03000$	0.2750	0.05	0.015]	0.125
2	$\theta_2^T = [0.02875$	0.2750	0.04	0.010]	0.125
3	$\theta_3^T = [0.02750$	0.2500	0.05	0.015]	0.125
4	$\theta_4^T = [0.02625$	0.2500	0.04	0.010]	0.125
5	$\theta_5^T = [0.02500$	0.2250	0.05	0.015]	0.125
6	$\theta_6^T = [0.02375$	0.2250	0.04	0.010]	0.125
7	$\theta_7^T = [0.02250$	0.2000	0.05	0.015]	0.125
8	$\theta_8^T = [0.02125$	0.2000	0.04	0.010]	0.125

Table 3 Parameter estimates and performance criteria for Case (ii)

	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$\hat{q}(\times 10^{-2})$	$\hat{r}(\times 10^{-2})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-2})$
Completely Known	*	*	*	*	1.2465	9.1106
Unknown	2.8293	2.7043	1.00	4.00	1.2468	9.1031

Table 4 Unknown parameter events

Filter	$\theta^T = [$ a	c]	$Pr(\theta/0)$
1	$\theta_1^T = [0.0300$	0.275]	0.25
2	$\theta_2^T = [0.0275$	0.250]	0.25
3	$\theta_3^T = [0.0250$	0.225]	0.25
4	$\theta_4^T = [0.0225$	0.200]	0.25

expected value is obtained by taking the average over 100 samples. The measurement locations are the same as those in Case (*i*) and the points to be estimated are taken at nine node points $m' = 1,2,\cdots,9$ except for two boundaries $x=0$ and $x=1$. The events of unknown parameters $\theta^T = [a,c,q,r]$ for each elemental distributed Kalman filter are shown in Table 2. The variances for measurement noises at three observation points are assumed to be $R_k = \text{diag}(r,r,r)$ where $r=0.01$. Table 3 shows that the filter performances for the cases of completely known parameters and unknown parameters resemble although the estimates for the diffusivity coefficient a and absorption factor c have still some biases.

(*iii*) *Effect of Measurement Point Numbers.* After replacing the events of unknown parameters for Case (ii) by ones shown in Table 4, the effect of measurement point numbers is examined.

It is seen from Table 5 that the filter performance is more improved if the measurement point numbers increase. For the performance criterion J_2, although increasing the measurement point numbers seems to be insignificant for its filter performance, as may be seen from later simulation, this effect will be changed due to the difference of selection for measurement locations. If the note is restricted to the parameter identification, it may be desirable for its purpose to have one or two measurement points. This means that if the unknown parameters are regarded as some state values by using a usual nonlinear filter technique, increasing the information on filtering may improve the correctness of parameter estimation, but in the proposed method, the number of measurement points is ineffective on its parameter estimation directly; rather, the proposed method depends upon the successful quantization levels for the unknown parameter values.

(*iv*) *Effect of the Location of the Measurement Points.* Under the unknown parameter events of Case (iii), three cases were considered, $m=1$, $m=2$, and $m=3$ as Tables 6, 7, and 8. Table 6 shows the results of the case where $m=1$. It is noted that the optimal measurement location is $x=0.5$ for the systems which have the parabolic operator as equation (8-1) with the constant system and measurement noise variances

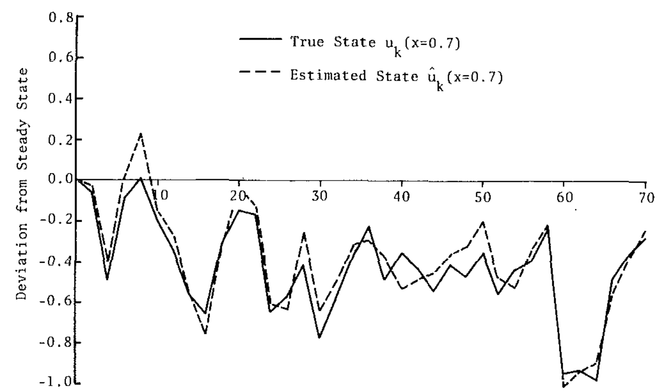

Fig. 6 Estimated flux pattern at $x=0.7$

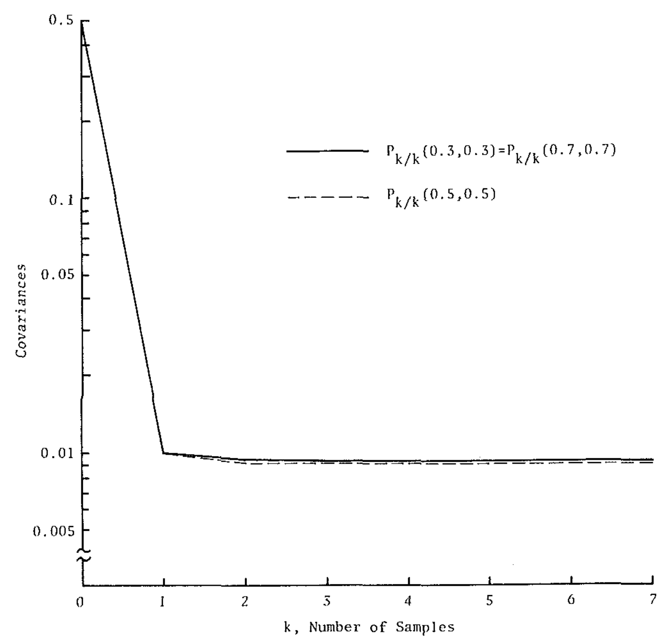

Fig. 7 Optimal state-error covariances trajectory

(over the time and coordinate space). This result may be conjectured easily from the problem for the optimal sensor allocations [28].

It is seen from Table 7 that $x^1 = 0.4$ and $x^2 = 0.6$ are suitable for filtering performance and $x^1 = 0.4$ and $x^2 = 0.5$ are also desirable for parameter estimation in two measurement

Table 5 Parameter estimates and performance criteria for Case (iii)

m	Location $x^i, i=1, \ldots, m$	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-1})$
1	0.1	2.8820	2.6320	1.7793	0.7933
2	0.1,0.9	2.9396	2.6896	2.8080	1.0934
3	0.1,0.5,0.9	2.9618	2.7118	1.3766	1.0884
4	0.1,0.4,0.6,0.9	2.9932	2.7432	0.8286	1.2723
5	0.1,0.3,0.5,0.7,0.9	2.9935	2.7435	0.5311	0.7731
6	0.1,0.2,0.4,0.6,0.8,0.9	2.9964	2.7464	0.5196	0.9899
7	0.1,0.3,0.4,0.5,0.6,0.7,0.9	2.9961	2.7461	0.4401	1.2701
8	0.1,0.2,0.3,0.4,0.6,0.7,0.8,0.9	2.9976	2.7476	0.2930	0.8168
9	0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9	2.9976	2.7476	1.1761	0.7207

Table 6 Parameter estimates and performance criteria for one measurement point

Location x^i, $i=1$	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-2})$
0.1	2.8820	2.6320	1.7793	7.9333
0.4	2.7794	2.5294	1.6900	7.8223
0.5	2.7158	2.4658	1.5851	7.6005
0.6	2.7730	2.5230	2.1558	7.9769
0.9	2.8401	2.5901	4.5855	13.128

Table 7 Parameter estimates and performance criteria for two measurement points

Location $x^i, i=1,2$	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-2})$
0.1,0.9	2.9396	2.6896	2.8080	10.934
0.1,0.4	2.9316	2.6816	1.2834	6.4290
0.4,0.5	2.8452	2.5952	1.2311	9.1689
0.4,0.6	2.8990	2.6490	1.0272	7.5913
0.4,0.9	2.9214	2.6714	1.4038	9.3386

Table 8 Parameter estimates and performance criteria for three measurement points

Location $x^i, i=1,2,3$	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-2})$	$J_2(\times 10^{-2})$
0.1,0.5,0.9	2.9618	2.7118	13.766	10.884
0.2,0.5,0.8	2.9554	2.7054	6.2455	8.1225
0.1,0.4,0.7	2.9499	2.6999	10.323	7.9074
0.3,0.6,0.9	2.9679	2.7279	9.5152	11.153
0.3,0.5,0.7	2.9432	2.6932	12.476	9.0932

Table 9 Parameter estimates and performance criteria for eigenfunction up to 10-term approximation

Terms N	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-2})$
1	2.6241	2.3741	1.1210	11.567
2	2.7657	2.5157	1.1764	8.6228
3	2.9253	2.6753	1.2670	9.1228
4	2.9432	2.6932	1.2476	9.0932
5	2.9865	2.7365	1.0826	8.6607
6	2.9888	2.7388	1.0813	8.6391
7	2.9912	2.7412	1.0814	8.6173
8	2.9930	2.7430	1.0810	8.6049
9	2.9939	2.7439	1.0730	8.6035
10	2.9939	2.7439	1.0730	8.6035

Table 10 Parameter estimates and performance criteria for variation of measurement noise variance

Variance r	$\hat{a}(\times 10^{-2})$	$\hat{c}(\times 10^{-1})$	$J_1(\times 10^{-1})$	$J_2(\times 10^{-2})$
0.005	2.9476	2.6976	1.1737	8.8598
0.010	2.9432	2.6932	1.2476	9.0932
0.020	2.9364	2.6864	1.3787	9.5046
0.030	2.9310	2.6810	1.4950	9.8589
0.040	2.9263	2.6763	1.6013	10.190
0.050	2.9220	2.6720	1.6999	10.508
0.100	2.9033	2.6533	2.1124	11.803

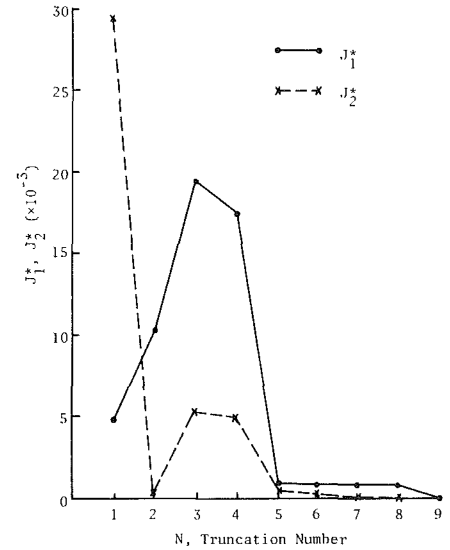

Fig. 8 Effect of the truncation number of eigenfunction

locations. The former tendency agrees nearly with the result of [29]. It is seen that the filtering performance is good at $x^1 = 0.2$, $x^2 = 0.5$ and $x^3 = 0.8$ or $x^1 = 0.1$, $x^2 = 0.4$ and $x^3 = 0.7$ for the case of $m = 3$ in Table 8, and the latter is superior to the former slightly in the viewpoint of parameter estimation.

(*v*) *Effect of the Truncation Number of Eigenfunction.* In order to examine the effect of the truncation number of eigenfunction, the simulations in the case of the adaptive filter with approximation of up to 10-term were achieved. It is seen from Table 9 that 2-term approximation is enough for criterion J_2, but if another criterion J_1 is considered, too, a 5-term approximation may be necessary for the guarantee of filter performance.

If the following two simplified criteria are utilized, the conclusions previously discussed may become more apparent ones (see Fig. 8):

$$J_1^* = J_1 - J_{1,10} \tag{9-3}$$

$$J_2^* = J_2 - J_{2,10} \tag{9-4}$$

where subscript "10" indicates the performance J_i, $i = 1,2$ at $N = 10$.

(*vi*) *Effect of the Measurement Noise Level.* The correctness of parameter estimation and the filtering performance due to the variation of measurement noise level are shown in Table 10. It is seen that increasing the noise level degrades the filter performance in Table 10. This tendency may be understood from the information matrix for distributed systems (e.g., see equation (37) in reference [29]). It is noted that there exists an opposite tendency to the case of parameter identification by using a usual gradient search method [30]. If the unknown parameter is estimated by using a usual nonlinear filter, the correctness of parameter estimation will deteriorate as the noise level increases. The proposed adaptive filter, however, is independent of the mechanism for parameter estimation and the identified parameters are not fed back to the filter mechanism (see Fig. 1). Thus it is concluded that the best weight for the unknown parameters always becomes the best one for the elemental distributed Kalman filter, but not vice versa.

10 Conclusions

An adaptive filter in discrete-time has been developed for a stochastic distributed parameter system with unknown time and space-invariant parameters. The basic idea presented here is the separation principle of the parameter identification scheme from the state estimation discussed in a continuous-time case [9]. That is, the essentially nonlinear adaptive distributed estimators are shown to be partitioned into two parts, a linear nonadaptive part consisting of a bank of distributed Kalman filters and a nonlinear part incorporating the learning nature of the estimator. This also implies that the Lainiotis' partition theorem for lumped parameter systems [20] holds at distributed parameter systems.

It is found from numerical results that this adaptive filter approximated by the eigenfunction expansion technique is highly effective in reducing the parameter uncertainties. Moreover, the problems for the effectiveness of measurement point numbers, the location of the measurement points, and the truncation number of eigenfunction are pointed out through those simulations. The problems of finding a successful technique for quantizing the unknown parameters [17], of extending the results obtained here to the case when the parameters are spatially and/or temporally [18] dependent, and of applying them to the detection-estimation are current research areas.

Acknowledgments

The authors wish to thank Professor M. Iwasaki and Research Associate M. Hashiguchi of Kyushu University for a number of enlightening discussions on distributed parameter systems.

References

1 Koda, M., and Seinfeld, J. H., "Estimation of Urban Air Pollution," *Automatica*, Vol. 14, 1978, pp. 583–595.

2 Ajinkya, M. B., Köhne, M., Mäder, H. F., and Ray, W. H., "The Experimental Implementation of a Distributed Parameter Filter," *Automatica*, Vol. 11, 1975, pp. 571–577.

3 Lausterer, G. K., Ray, W. H., and Martens, H. R., "Real Time Distributed Parameter State Estimation Applied to a Two Dimensional Heated Ingot," *Automatica*, Vol. 14, 1978, pp. 335–344.

4 Lausterer, G. K., and Ray, W. H., "Distributed Parameter State Estimation and Optimal Feedback Control—An Experimental Study in Two Space Dimensions," *IEEE Trans. Aut. Contr.*, Vol. AC-24, No. 2, Apr. 1979, pp. 179–190.

5 Tzafestas, S. G., "Innovation Approach to Distributed-Parameter Detection and Estimation," *Int. J. Syst. Sci.*, Vol. 9, No. 2, 1978, pp. 147–184.

6 Tzafestas, S. G., and Nightingale, J. M., "Stochastic Distributed Filtering Approach to Gamma-Ray Imaging," *Int. J. Syst. Sci.*, Vol. 7, No. 11, 1976, pp. 1249-1274.

7 Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

8 Ljung, L., "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," *IEEE Trans. Aut. Contr.*, Vol. AC-24, No. 1, Feb. 1979, pp. 36–50.

9 Sunahara, Y., Ohsumi, A., and Imamura, M., "A Method of Parameter Identification for Linear Distributed Parameter Systems," *Automatica*, Vol. 12, 1976, pp. 245–256.

10 Sakawa, Y., "Optimal Filtering in Linear Distributed-Parameter Systems," *Int. J. Control*, Vol. 16, No. 1, 1972, pp. 115–127.

11 Watanabe, K., "Optimal Filtering and Smoothing Algorithms for Linear Distributed-Parameter Systems With Pointwise Observation," *Int. J. Syst. Sci.*, to be presented, 1981.

12 Thau, F. E., "On Optimum Filtering for a Class of Linear Distributed-Parameter Systems," ASME *Journal of Basic Engineering*, Vol. 91, June 1969, pp. 173–178.

13 Lainiotis, D. G., "Optimal Adaptive Estimation: Structure and Parameter Adaptation," *IEEE Trans. Aut. Contr.*, Vol. AC-16, Apr. 1971, pp. 160–170.

14 Deshpande, J. G., Upadhyay, T. N., and Lainiotis, D. G., "Adaptive Control of Linear Systems," *Automatica*, Vol. 9, 1973, pp. 107–115.

15 Kailath, K., "An Innovation Approach to Least-Squares Estimation Part I: Linear Filtering in Additive White Noise," *IEEE Trans. Aut. Contr.*, Vol. AC-13, No. 6, 1968, pp. 646–655.

16 Sims, F. L., Lainiotis, D. G., and Magill, D. T., "Recursive Algorithm for the Calculation of the Adaptive Kalman Filter Weighting Coefficients," *IEEE Trans. Aut. Contr.*, Vol. AC-14, Apr. 1969, pp. 215–218.

17 Sengubush, R. L., and Lainiotis, D. G., "Simplified Parameter Quan-

tization Procedure for Adaptive Estimation," *IEEE Trans. Aut. Contr.*, Vol. AC-14, Aug. 1969, pp. 424-425.

18 Hilborn, C. G., and Lainiotis, D. G., "Optimal Adaptive Filter Realization for Sample Stochastic Processes with an Unknown Parameter," *IEEE Trans. Aut. Contr.*, Vol. AC-14, Dec. 1969, pp. 767-770.

19 Lainiotis, D. G., and Sims, F. L., "Performance Measure for Adaptive Kalman Estimators," *IEEE Trans. Aut. Contr.*, Vol. AC-15, Apr. 1970, pp. 249-250.

20 Tzafestas, S. G., "Bayesian Approach to Distributed-Parameter Filtering and Smoothing," *Int. J. Control*, Vol. 15, No. 2, 1972, pp. 273-295.

21 Magill, D. T., "Optimal Adaptive Estimation of Sampled Stochastic Processes," *IEEE Trans. Aut. Contr.*, Vol. AC-10, Oct. 1965, pp. 434-439.

22 Watanabe, K., Yoshimura, T., and Soeda, T., "Optimal Adaptive Estimation and Stochastic Control for a Distributed-Parameter Systems" (submitted for publication).

23 Watanabe, K., Yoshimura, T., and Soeda, T., "Optimal Nonlinear Estimation for Distributed-Parameter Systems via the Partition Theorem," *Int. J. Syst. Sci.*, Vol. 11, No. 9, 1980, pp. 1113-1130.

24 Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," ASME *Journal of Basic Engrg.*, Vol. 82, 1960, pp. 35-45.

25 Meditch, J. S., *Stochastic Optimal Linear Estimation and Control*, McGraw-Hill, 1969.

26 Hassan, M. A., and Solberg, K. O., "Discrete Time Control of Linear Distributed Parameter Systems," *Automatica*, Vol. 6, 1970, pp. 409-417.

27 Hamza, M. H., and Sheirah, M. A., "A Self-Tuning Regulator for Distributed Parameter Systems," *Automatica*, Vol. 14, 1978, pp. 453-463.

28 Aidarous, S. E., Gevers, M. R., and Installe, M. J., "Optimal Sensors' Allocation Strategies for a Class of Stochastic Distributed Systems," *Int. J. Control*, Vol. 22, No. 2, 1975, pp. 197-213.

29 Kumar, S., and Seifeld, J. H., "Optimal Location of Measurements for Distributed Parameter Estimation," *IEEE Trans. Aut. Contr.*, Vol. AC-23, Aug. 1978, pp. 690-698.

30 Crescitelli, S., and Nicoletti, B., "An Algorithm for the Identification of Parameters in Chemical Reactor Processes," *Int. J. Syst. Sci.*, Vol. 7, No. 9, 1976, pp. 1017-1028.

31 Lions, J. L., *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1971.

32 Curtain, R. F., and Pritchard, A. J., *Infinite Dimensional Linear Systems Theory*, Lecture Notes in Contr. and Inf. Sci. 8, Springer-Verlag, 1978.

33 Sawaragi, Y., Soeda, T., and Omatu, S., *Modeling, Estimation, and Their Applications for Distributed Parameter Systems*, Lecture Notes in Contr. and Inf. Sci. 11, Springer-Verlag, 1978.

34 Meditch, J. S., "Least-Squares Filtering and Smoothing for Linear Distributed Parameter Systems," *Automatica*, Vol. 7, 1971, pp. 315-322.