ESTIMATION IN DISTRIBUTED PARAMETER SYSTEMS

A THESIS
Presented to
The Faculty of the Division of Graduate
Studies and Research
By
Raman K. Sahgal

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in the School of Electrical Engineering

Georgia Institute of Technology
March 1976

ESTIMATION IN DISTRIBUTED PARAMETER SYSTEMS

Approved:

R. P. Webb, Chairman

A. S. Debs

J. L. Hammond

Date Approved by Chairman 1/5/76

ACKNOWLEDGEMENTS

I wish to express my appreciation to Dr. Roger P. Webb, my thesis advisor, for his encouragement, assistance and support throughout the development of this dissertation. I would also like to extend my appreciation to my reading committee members, Dr. Atif S. Debs and Dr. Joseph L. Hammond, for their comments on this dissertation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

The objective of this research is to develop estimation algorithms for noisy dynamical systems represented by partial differential equations, and to evaluate their performance on a digital computer. Filtering and smoothing algorithms are developed for noisy vector distributed systems, each of whose components is at most second order. Two classes of systems are assumed--those for which noise statistics are known and those whose noise statistics are unknown. For the first class distributed white disturbances in the dynamics and measurements are assumed and the innovations approach is used. For systems where the noise statistics are unknown, a least squares fit estimation approach is used. The estimate is required to minimize a positive definite functional which results in a two-point boundary value problem. The solution to this boundary value problem is obtained using invariant embedding, converting it into an initial (or final) value problem. The results of these algorithms are applied to representative problems in engineering, and compared with previously existing algorithms.

CHAPTER I

INTRODUCTION

The problem of estimating states in distributed systems arises in almost every kind of dynamic system. Every dynamic system is by nature distributed, but if in the course of its motion its energy is sufficiently concentrated at one point, it can be described by a lumped system. A typical example is the characterization of the behavior of a homogeneous vibrating beam by an ordinary differential equation describing the motion of the lowest eigen-mode. However, there are other systems which at times cannot be thus described, and must be characterized by partial differential equations. Noisy terms in distributed systems arise due to a number of factors, like incomplete information on the system behavior, improper characterization, etc. An example shall illustrate this. Consider the problem of modeling neutron flow in a nuclear reactor. A neutron is a particle, but its flow is, for purposes of keeping analysis simple, characterized by that of a fluid. The discrepancy between the characterization of neutron flow as fluid motion vis-á-vis particle motion gives arise to a stochastic noise term which can be measured. Measurement in itself introduces uncertainty, so we see that our system is characterized by both dynamic and measurement uncertainties. So it is desirable to know the actual state in the absence of noise, and it is to these points that this thesis addresses itself.

1.1. History of the Problem

A considerable body of literature exists pertaining to the estimation of states for distributed systems. These works have widely different symbolisms, measurement processes and boundary conditions associated with them. Since no rigorous extension of the Itô calculus to infinite dimensional systems is at present available, although it is being researched [1], such problems are attacked by indirect means. Falb [2] considers the linear filtering problem in Banach space formalism. The behavior of the system is assumed represented by an abstract differential equation with additive "white noise." Homogeneous boundary conditions are assumed, and an abstract orthogonal projection lemma obtained. Using an extension of Kalman and Bucy's work [12], the minimum-variance estimate in a Banach space is obtained. Tzafestas [3] and Thau [4] obtain identical results using the more conventional formalism for partial differential equations, for continuous and discrete measurements, respectively.

A different approach to the linear filtering problem is taken by Balakrishnan and Lions [5]. Based on the observation that for a lumped system with Gaussian disturbances, the estimation problem is the dual of the optimal regulator problem, they derive a minimum mean square filter in case there are noisy measurements, but no dynamic noise is present. Meditch [6] also considers the filtering problem in an identical manner for linear systems with dynamic noise present.

Tzafestas considers the filtering problem for linear Gaussian systems using the maximum likelihood approach [9], and the maximum a posteriori (MAP) approach [10].

The linear smoothing problem has not been so extensively studied as the filtering problem. Tzafestas derives a 'two-filter' version of the fixed interval smoothing algorithm for distributed Gaussian systems (3], [9]. A different algorithm is derived using the MAP approach [10]. The 'two-filter' algorithm is far more complicated than the MAP algorithm. Using the rather complex limit procedure of Meditch [12], Tzafestas derives fixed point and fixed lag smoothing algorithms for linear Gaussian systems.

The nonlinear estimation problem has been studied by Tzafestas [9], [10] for additive Gaussian disturbances. Using a first order linearization of the model, extended Kalman algorithms are obtained. Lamont [8] using a least squares estimation criterion and the invariant embedding approach has derived a nonlinear algorithm for filtering. Seinfeld et al. [14] use an identical approach, but discretize the system to a lumped system. Gavalas et al. [18] extend Lamont's results to a very general distributed-lumped system using the same technique. The nonlinear smoothing problem has not been studied in any detail.

1.2. Statement of the Problem

The above review of the literature can be summarized thus: The linear Gaussian filtering algorithm have been derived using orthogonal projections [2] - [4], least squares [5], maximum-likelihood [9], MAP [10]. For arbitrary disturbances of unknown statistics, least squares have been used [6]. The linear Gaussian smoothing problem has been approached using MAP [10]. Extended Kalman estimation algorithms exist for nonlinear systems with additive Gaussian noises [10]. For noises of arbitrary statistics, an invariant embedding filtering algorithm exists [7], [8], [14]. No smoothing algorithms exist.

This points the need for having smoothing algorithms, for arbitrary noises, for both linear and nonlinear systems. Smoothing is of interest in general dynamical systems because of the improvement in the estimate as evidenced by the reduction in error covariance.

The computational problem associated with any estimation algorithm is of supreme importance because eventually the algorithm is going to be processed on a computer. Very little attention appears to have been devoted to the computational aspects of the estimation problem, other than the results of Seinfeld et al. [14], who approached the problem by discretizing the system in space alone to get a set of ordinary differential equations for the estimator. These are cumbersome to solve on a digital computer.

With this short survey into the previous literature existing on the subject, we can now detail the problems to be considered in this research:

(a) Definition of a complete set of problem constraints which do not exclude nonlinearities or nongaussian statistics, and derivation of least squares estimators for the various smoothing problems. Filtering algorithms similar to the ones obtained by Lamont [8], Seinfeld [14], are necessary adjuncts to the smoothing algorithms.

(b) Derivation of smoothing algorithms for systems with white noise disturbances via the innovations approach.

(c) An examination of the computational feasibility of the above developed algorithms for computer execution time, storage requirements and accuracy vis-á-vis the previously existing algorithms of Seinfeld et al. [14] and analyses of practicality by application to representative engineering problems.

CHAPTER II

THE ESTIMATION PROBLEM WITH UNKNOWN NOISE STATISTICS

As stated in Chapter I, classical least squares coupled with the invariant embedding approach shall be used in this section to derive smoothing algorithms for systems with unknown noise statistics. In this chapter, the class of systems to be considered are defined and the Euler-Lagrange equations governing the minimization of the weighted-least squares functional obtained; prior to detailing the solution to the various estimation problems outlined below.

2.1. System Model and Problem Formulation

Let Ω be a bounded, connected, open subset of n-dimensional Euclidean space E^n with piecewise smooth boundary $\partial\Omega$. Let t denote time on the fixed interval $[t_o,t_f]$.

We consider systems governed by the noisy vector partial differential equation of the form

$$\frac{\partial x(y,t)}{\partial t} = N(x,y,t,q(t)) + w(y,t),\ y \in \Omega,\ E\left\{x(y,t_o)\right\} = x_o \tag{2.1}$$

where $x(y,t)$ is the n-dimensional state vector;

$w(y,t)$ is the n-dimensional disturbance vector of unknown statistics;

N is an n-dimensional nonlinear spatial differential or integro-differential operator of order k defined on Ω.

For each $t \in [t_o,t_f]$ we assume $x(y,t) \in L_2^n(\Omega)$ and $w(y,t) \in L_2^n(\Omega)$. The n_1-vector $q(t)$ (independent of y) is defined by

$$\frac{dq(t)}{dt} = N_q(t,q) + \xi_1(t) \tag{2.2}$$

and the boundary condition is of the form

$$N_b(x,y,t,r(t)) = \xi_2(y,t) \quad y \in \partial\Omega \tag{2.3}$$

The m_3- vector $r(t)$ (independent of y) is governed by

$$\frac{dr(t)}{dt} = N_r(t,r) + \xi_3(t) \tag{2.4}$$

where N_b is an n_2-dimensional spatial operator on $\partial\Omega$ with $n_2 \leq n$ and operator order $< k$;

$\xi_1(t)$, $\xi_2(y,t)$, $\xi_3(t)$ are independent random processes with unknown statistical characteristics;

N_q and N_r are nonlinear n_1 and n_3-dimensional vector functions.

The state of the system is $S_t = \{x(y,t), q(t), r(t) : y \in \Omega, t \in [t_o,t_f]\}$.

The observation is assumed to be of the form*

$$z(y,t) = h(x,y,t) + v(y,t) \qquad y \in \Omega \tag{2.5}$$

where h is a bounded, continuous, ℓ-vector spatial operator;

$v(y,t)$ is the unknown additive measurement noise.

For all $t \in [t_o,t_f]$, $v(y,t) \in L_2^\ell(\Omega)$ is assumed.

---

*Observability of the system is discussed in Appendix C.

The model as postulated here is rather general and encompasses a very wide class of systems. In reality, by assuming broad generality for the system model, we would be sacrificing some very powerful mathematical tool which would not otherwise easily extend to systems of arbitrary order. So, in particular, to facilitate application of Green's theorem and the concept of adjoint systems, we restrict our systems such that

$$N(x,y,t,q(t)) = f\left(x,y,\frac{\partial x}{\partial y}, \frac{\partial^2 x}{\partial y^2}, t, q(t)\right)$$

The estimation algorithms require a linearization about a nominal trajectory and by making the assumption that a Frechet derivative of N exists, the application of Green's theorem greatly simplifies the analysis.

This also requires that we restrict our boundary operator to

$$N_b(x,y,t,r) = g\left(x,\frac{\partial x}{\partial \sigma}, y,t,r\right)$$

where $\partial[\ ]/\partial\sigma$ denotes the derivative w.r.t. the outward normal n. Such an assumption is not particularly restrictive, as it encompasses within it all quasilinear parabolic and hyperbolic systems. The only system of any importance left out is the transient biharmonic equation.

The model as formulated then is a very general model describing both distributed and lumped systems. Such systems arise, for instance, in describing the heat conduction process where the temperature of the medium is controlled by the fuel flow [27]. Purely distributed systems arise so often in nature that no mention need be made of the field of application of this model.

Note that the measurements on the system are considered only in the volume Ω. No measurements on the boundary are assumed. Boundary measurements can be incorporated into the system model by a slight extension of the present analysis.

The state estimation problems to be considered here are: (a) The Filtering Problem. Given measured data $z(y,\tau) : y \in \Omega$ from $\tau = t_o$ to the present time $\tau = t$ and an initial estimate $\hat{x}(y,t_o|t_o)$ together with its covariance matrix

$$P_o(y,y_1) = P(y,y_1,t_o|t_o)$$

find an estimate $S_{t|t}$ of the current state S_t. (b) The Fixed-Interval Smoothing Problem. Given measured data $z(y,\tau) : y \in \Omega \,\forall\, \tau \in [t_o,t_f]$ and an initial estimate $\hat{x}(y,t_o|t_o) : y \in \Omega$ together with its covariance matrix, find a smoothed estimate $S_{t|t_f}$ over the entire fixed-interval $t \in [t_o,t_f]$. (c) The Fixed-Point Smoothing Problem. Given data as above, find a smoothed estimate $S_{t|T}$ at only one particular time $t = T$ within a time interval $[t_o,t_f]$. (d) The Fixed-Lag Smoothing Problem. Given data as above, find a running smoothed estimate $S_{t|t+\theta}$ which lags behind the time of the most recent observation by a constant amount of time θ.

The estimation criterion for any admissible estimate $S_t = \{x(y,t),\ q(t),\ r(t) : y \in \bar{\Omega},\ t \in [t_o,t_f]\}$ is the minimization of the least-squares error functional

$$J = \frac{1}{2}\int_{t_o}^{t_f}\Big\{\int_\Omega\int_\Omega (z(y,t) - h(x,y,t))^T Q(y,y_1,t)(z(y_1,t)$$

$$- h(x,y_1,t))d\Omega d\Omega_1 + \int_\Omega\int_\Omega\left[\frac{\partial x(y,t)}{\partial t} - N(x,y,t)\right]^T R(y,y_1,t)$$

$$\left[\frac{\partial x(y_1,t)}{\partial t} - N(x,y_1,t)\right]d\Omega d\Omega_1 + \int_{\partial\Omega}\int_{\partial\Omega}(N_b(x,y,t))^T S(y,y_1,t)$$

$$N_b(x,y_1,t)d(\partial\Omega)d(\partial\Omega_1) + ||q(t) - N_q(t)||^2_{T_1(t)} + ||r(t)$$

$$- N_r(t)||^2_{T_2(t)}\Big\}dt + \int_\Omega\int_\Omega(x(y,t_o) - x_o(y))^T P_o^{-1}(y,y_1)(x(y_1,t_o)$$

$$- x_o(y_1)\,d\Omega d\Omega_1 \qquad (2.6)$$

In equation (2.6) the first three terms represent the weighted observation residual over Ω and the weighted dynamic residuals over Ω and $\partial\Omega$ respectively. The next two terms represent the weighted error residuals in estimating q(t) and r(t) respectively. The last term is the weighted residual for the initial state $x(y,t_o)$. The weighting matrices P_o,Q,R,S are bounded, continuous, symmetric, symmetric in their spatial arguments, and positive definite. The inverse kernel P_o^{-1} is defined by

$$\int_\Omega P_o^{-1}(y,y_2)P_o(y_2,y_1)d\Omega_2 = \int_\Omega P_o(y,y_2)P_o^{-1}(y_2,y_1)d\Omega_2 = \delta(y - y_1) \qquad (2.7)$$

T_1 and T_2 are symmetric bound positive definite matrices. These facts imply that a minimum for J exists [18] subject to the constraints of equations (2.1-5).

It is important to emphasize that no model for the uncertainty in either $\underline{x}$ or the various noises are used. Weighted-least squares estimation theory replaces modeling and optimality arguments by the intuitive judgment that a "reasonable" estimate of $\underline{x}$ would be obtained by minimizing a weighted functional (2.6) of all the residual errors.

2.2. The Euler-Lagrange Equations

In this section we reformulate the problem of minimizing the functional (2.6) to facilitate application of variational techniques. Then the use of the calculus of variations to minimize (2.6) leads to the derivation of the Euler-Lagrange equations for the estimation problems. These equations result in a two-point boundary value problem (TPBVP). These equations form the basis for the work of the next few sections wherin solutions to the various estimation problems are obtained from this TPBVP by converting it into an initial value problem.

The problem is now recast so that the theory of optimal control can be applied. We desire to minimize J with respect to $S_t = \{x(y,t), q(t), r(t) : y \varepsilon \bar{\Omega},\ t \varepsilon [t_o, t_f]\}$ and $\{u_i : i = 1,..,4\}$ where

$$\begin{aligned} J = \frac{1}{2}\int_{t_o}^{t_f}\Big\{ &\int_\Omega\int_\Omega (z(y,t) - h(x,y,t))^T Q(y,y_1,t)(z(y,t) - h(x,y_1,t)) \\ &d\Omega d\Omega_1 + \int_\Omega\int_\Omega u_1^T(y,t)R(y,y_1,t)u_1(y_1,t)d\Omega d\Omega_1 + \int_{\partial\Omega}\int_{\partial\Omega} u_2^T(y,t) \\ &S(y,y_1,t)u_2(y_1,t)d(\partial\Omega)d(\partial\Omega_1) + ||u_3(t)||^2_{T_1(t)} \\ &+ ||u_4(t)||^2_{T_2(t)} \Big\} dt + \int_\Omega\int_\Omega (x(y,t_o) - x_o(y))^T P_o^{-1}(y,y_1) \\ &(x(y_1,t_o) - x_o(y_1))d\Omega d\Omega_1 \equiv \int_{t_o}^{t_f} \Phi(S_t,u_i,y,t)dt + \theta(x,t_o) \end{aligned} \qquad (2.8)$$

subject to the constraints

$$\frac{\partial x(y,t)}{\partial t} = N(x,q,t) + u\ (y,t)\ ,\ E\left\{x(y,t_o)\right\} = x_o(y),\ y \in \Omega$$

$$\frac{dq}{dt} = N_q(t,q) + u_3(t)$$

$$N_b(x,r,y,t) = u_2(y,t) \quad y \in \partial\ \Omega \tag{2.9}$$

$$\frac{dr}{dt} = N_r(r,t) + u_4(t)$$

The weighting matrices P_o, Q, R, S, T_1, T_2 are as defined in Sec. 2.1 previously. The optimal control problem is recognized to be a Bolza problem in the calculus of variations.

Defining the Hamiltonian $H(u_i,\lambda_i,y,t\colon\ i=1,3,4)$ as

$$H = \Phi(S_t,u_i,y,t) + \int_\Omega \lambda_1^T(y,t)(N(x,q,y,t) + u_1(y,t))d\Omega$$

$$+ \lambda_3^T(N_q^T + u_3) + \lambda_4^T(t)\ (N_r^T + u_4(t)) \tag{2.10}$$

where Φ is as defined in (2.8) and $\lambda_1,\lambda_4,\lambda_3$ are adjoint variables, we obtain the extended functional

$$J = \theta(x,t_o) + \int_{t_o}^{t_f}\left\{H(S_t,u_i,y,t) - \int_\Omega \lambda_1^T(y,t)\left(\frac{\partial x}{\partial t}\right)d\Omega\right.$$

$$\left. - \lambda_3^T(t)\dot{q}(t) - \lambda_4^T(t)r(t)\right\}dt \tag{2.11}$$

Integrating by parts, this yields

$$J = \theta(\hat{x},t_o) + \int_{t_o}^{t_f}\left(H + <\dot{\lambda}_3,\hat{q}> + <\dot{\lambda}_4,\hat{r}> + \int_{\Omega}<\frac{\partial\lambda_1}{\partial t},\hat{x}>d\Omega\right)dt$$

$$- \left\{ \int_{\Omega}<\lambda_1,\hat{x}>d\Omega + <\lambda_3,\hat{q}> + <\lambda_4,\hat{r}>\right\}\Bigg|_{t_o}^{t_f} \tag{2.12}$$

where < , > denotes the scalar product.

A necessary condition for a minimum is that the first variation of J vanish for arbitrary variations δS_t and δu_i, $i = 1..,4$ [20]. The first variation is given by

$$\delta J = \int_{\Omega}<\frac{\delta\theta(\hat{x},t_o)}{\delta x},\delta\hat{x}>\ d\Omega + \int_{t_o}^{t_f}\left\{\int_{\Omega}<\frac{\delta H}{\delta\hat{x}},\delta\hat{x}>d\Omega + <\frac{\partial H}{\partial\hat{q}},\delta\hat{q}> + <\frac{\partial H}{\partial\hat{r}},\delta\hat{r}>\right.$$

$$\left. + \int_{\partial\Omega}<\frac{\delta H}{\delta\hat{x}},\ \delta\hat{x}>d(\partial\Omega)\right\} dt + \int_{t_o}^{t_f}\left\{\int_{\Omega}<\frac{\delta H}{\delta\hat{u}_1},\delta\hat{u}_1>d\Omega + \int_{\partial\Omega}<\frac{\delta H}{\delta\hat{u}_2},\delta\hat{u}_2>d(\partial\Omega)\right.$$

$$\left. + <\frac{\partial H}{\partial\hat{u}_3},\delta\hat{u}_3> + <\frac{\partial H}{\partial u_4},\delta\hat{u}_4>\right\}dt - \left\{\int_{\Omega}<\lambda_1,\delta\hat{x}>d\Omega - <\lambda_3,\delta\hat{q}> - <\lambda_4,\delta\hat{r}\right\}\Bigg|_{t_o}^{t_f}$$

$$+ \int_{t_o}^{t_f}\left\{<\dot{\lambda}_3,\delta\hat{q}> + <\dot{\lambda}_4,\delta\hat{r}> + \int_{\Omega}<\frac{\partial\lambda_1}{\partial t},\delta\hat{x}>d\Omega\right\}dt \tag{2.13}$$

where δ( )/δ( ) denotes the functional derivative [21].

The first variation δJ shall vanish if

$$\frac{\delta H}{\delta\hat{x}} = -\frac{\partial\lambda_1}{\partial t} \quad , \quad \frac{\delta H}{\delta\hat{u}_1} = 0 \qquad y \in \Omega$$

$$\frac{\delta H}{\delta\hat{x}} = 0 \quad , \quad \frac{\delta H}{\delta\hat{u}_2} = 0 \qquad y \in \partial\ \Omega \tag{2.14}$$

$$\dot{\lambda}_3 = \frac{-\partial H}{\partial\hat{q}} \quad , \quad \dot{\lambda}_4 = \frac{-\partial H}{\partial\hat{r}} \quad , \quad \frac{\partial H}{\partial\hat{u}_3} = 0 \quad , \quad \frac{\partial H}{\partial\hat{u}_4} = 0$$

together with the terminal conditions

$$\begin{aligned}
&\lambda_1(y,t_f) = 0 \\
&\lambda_1(y,t_o) = -\int_\Omega P_o^{-1}(y,y_1)(\hat{x}_o(y_1,t_o) - x_o(y_1))d\Omega \\
&\lambda_3(t_f) = 0 = \lambda_3(t_o) \\
&\lambda_4(t_f) = 0 = \lambda_4(t_o)
\end{aligned} \tag{2.15}$$

The control $\hat{u}_1$ can be obtained from (2.14) to yield

$$\hat{u}_1(y,t) = -\int_\Omega R^{-1}(y,y_1,t)\ (y_1,t)d\Omega \tag{2.16}$$

To obtain $\delta H/\delta\hat{x}$; $y \in \Omega$, and $\delta H/\delta\hat{x}$; $y \in \partial\,\Omega$, we write the full expression for δH:

$$\delta H = \int_\Omega <\frac{\delta\Phi}{\delta\hat{x}},\delta x>d\Omega + \int_{\partial\Omega}<\frac{\delta\Phi}{\delta\hat{x}},\delta\hat{x}>d(\partial\Omega) + \int_\Omega<\lambda_1,(N)_{\hat{x}}\delta\hat{x}>d\Omega$$

where $(N)_{\hat{x}}$ denotes the Frechet (strong) derivative of the operator N[22]. Using the compact Green identity this becomes

$$\delta H = \int_\Omega< \frac{\delta\Phi}{\delta\hat{x}} + (N)^*_{\hat{x}}\lambda_1\ ,\delta\hat{x}>d\Omega + \int_{\partial\Omega}< \frac{\delta\Phi}{\delta\hat{x}} + (N_b)^*_{\hat{x}}\lambda_1\ ,\delta\hat{x}>d(\partial\Omega) \tag{2.17}$$

Hence
$$\frac{\delta H}{\delta\hat{x}} = -\frac{\partial\lambda_1}{\partial t} = \frac{\delta\Phi}{\delta\hat{x}} + (N)^*_{\hat{x}}\lambda_1\ , \qquad y \in \Omega \tag{2.18}$$

along with the boundary condition

$$\frac{\delta H}{\delta \hat{x}} = 0 = (N_b)^*_{\hat{x}}\lambda_1 , \qquad y \in \partial\ \Omega \tag{2.19}$$

since $\delta\Phi/\delta\hat{x}$ is zero for $y \in \partial\ \Omega$. In eqns. (2.17)-(2.19), $(N)^*_{\hat{x}}$ denotes the adjoint of the linear operator $(N)_{\hat{x}}$.

Eqns. (2.14) and (2.18) along with (2.10) yield

$$\frac{-\partial\lambda_1}{\partial t} = (N)^*_{\hat{x}}\lambda_1 + \int_\Omega \left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(y,y_1,t)(z(y_1,t) - h(\hat{x},y_1,t))d\Omega_1, \ y\varepsilon\Omega \tag{2.20}$$

$$u_2(y,t) = - \int_{\partial\Omega} S^{-1}(y,y_1,t)K(y_1,t)\lambda_1(y_1,t)d(\partial\Omega_1) \ \ y \in \partial\Omega \tag{2.21}$$

$$u_3(t) = - T_1^{-1}(t)\lambda_3(t) \tag{2.22}$$

$$u_4(t) = - T_2^{-1}(t)\lambda_4(t) \tag{2.23}$$

$$- \dot{\lambda}_3(t) = \int_\Omega (N)^T_{\hat{q}}\lambda_1(y,t)d\Omega + (N_{\hat{q}})^T_{\hat{q}}\lambda_3(t) \tag{2.24}$$

$$- \dot{\lambda}_4(t) = \int_{\partial\Omega} (N_b)^T_{\hat{r}}K\lambda_1(y,t)d(\partial\Omega) + (N_{\hat{r}})^T_{\hat{r}}\lambda_4(t) \tag{2.25}$$

The form of $K(y,t)$ and the boundary conditions for $\lambda_1(y,t)$ are obtained by assuming that the linear operators $(N)_x$ and $(N_b)_x$ have the form [23]

$$\begin{aligned} (N)_x[\cdot] &= \sum_{i,j=1}^{n} A_{ij}(y,t)\frac{\partial^2[\cdot]}{\partial y_i\,\partial y_j} + \sum_{i=1}^{n} B_i(y,t)\frac{\partial[\cdot]}{\partial y_j} + C(y,t) \\ (N_b)_x[\cdot] &= A_{b_1}(y,t)[\cdot] + A_{b_2}(y,t)\frac{\partial[\cdot]}{\partial\sigma} \end{aligned} \tag{2.26}$$

where $\partial/\partial\sigma$ is the derivative w.r.t. the conormal of Ω relative to the operator $(N_b)_x$. With this formalism, one can show [23]

$$K(y,t) = (A_{b_2}^{-1})^T(y,t)A_\sigma^T(y,t) \tag{2.27}$$

where

$$A_\sigma(y,t) = \left\{ \sum_{j=1}^{n} \left[\sum_{i=1}^{n} A_{ij} \cos(n,x_i) \right]^2 \right\}^{1/2}$$

with n the interior normal of Ω. Furthermore, with this formalism, the boundary condition for $\lambda_1(y,t)$, eqn. (2.19), becomes, with the help of Green's identity [23]

$$A_\sigma^T \frac{\partial \lambda_1}{\partial \sigma} + A_{b_1}^T A_{b_2}^{-1} A_\sigma^T \lambda_1 - \sum_{i=1}^{n} D_i^T \cos)n,y_i)\lambda_1(y,t) = (N_b)_x^* \lambda_1$$

$$= 0 \qquad y \in \partial\,\Omega \tag{2.28}$$

where

$$D_i = B_i - \sum_{j=1}^{n} \frac{\partial A_{ij}}{\partial x_j}$$

The boundary value problem can now be written down as

$$\frac{\partial \hat{x}(y,t)}{\partial t} = N(\hat{x},\hat{q},y,t) - \int_\Omega R^{-1}(y,y_1,t)\lambda_1(y,t)\,d\Omega \quad y \in \Omega \tag{2.29}$$

$$\dot{\hat{q}}(t) = N_{\hat{q}}(\hat{q},t) - T_1^{-1}(t)\lambda_3(t) \tag{2.30}$$

$$\dot{\hat{r}}(t) = N_{\hat{r}}(\hat{r},t) - T_2^{-1}(t)\lambda_4(t) \tag{2.31}$$

along with the adjoint equations (2.20), (2.24), (2.25) and boundary equations (2.21) and (2.28) with terminal conditions given by (2.15).

In essence, any estimation algorithm comprises of estimator equations, adjoint equations and boundary conditions tabulated in Table 1.

Table 1. General Estimation Algorithm

Estimator	Estimator Eqn.	Adjoint Eqn.	B.C.	I.C.
$\hat{x}(y,t)$	(2.29)	(2.20)	(2.21),(2.28)	2.15a
$\hat{q}(t)$	(2.30)	(2.24)		2.15b
$\hat{r}(t)$	(2.31)	(2.25)		2.15c

The above equations are used to derive the filtering algorithm in Appendix A, and the smoothing solution in the next few sections.

2.3. The Smoothing Problem

Much attention has been devoted to the linear and nonlinear filtering algorithms [1]-[6], [8]-[10] for distributed systems, with some attention to the prediction problem [3], [9]. The smoothing problem has received considerably less attention despite the fact that considerable reduction in error, above that obtained by the filtering solution, is possible. Previous approaches to solving the linear smoothing problem have included orthogonal projections [3], Bayes' approach for white Gaussian disturbances [10]. The only algorithms for nonlinear plants existing are linearized Kalman algorithms [10], obtained by linearizing the plant about a nominal trajectory. Purely nonlinear estimators for arbitrary disturbances do not exist.

The results obtained in this section are extensions on previous attempts to get smoothing results [10] and form a major contribution of this thesis. We shall develop solutions to the various smoothing problems outlined in Sec. 2.1 which avoid the direct solution of the TPBVP posed in Sec. 2.2. The approach used is invariant embedding, which converts the BVP into a more general class of problems. To proceed with the invariant embedding solution for smoothing, we need the filtering solution. The filtering solution has already been developed for similar models by Lamont and Kumar [8], Hwang, et al. [18]. Its derivation is given in Appendix A because of the frequency with which it is referred to.

2.3A. Fixed-Interval Smoothing

Often it is necessary that better estimates than the crude filtered estimate be obtained for data interpretation purposes. If in case there is no need for immediate data interpretation, the filtered estimate can be stored in its entirity over the entire observation period, and we can refine this estimate using the fixed interval smoothing routine. The fixed interval smoothing problem was outlined in Sec. 2.1. The estimate resulting is in general the best obtainable for any smoothing routine because knowledge of the filtered estimate over the entire observation interval is available. For linear systems least covariance of error also results. The disadvantage is the amount of computer memory storage required. As the fixed interval smoothing problem involves data over the entire observation interval, it is clear that the algorithmic computation must be done off-line. We first assume that we have obtained the filtered estimate $\{\hat{x}(y,t),\hat{q}(t),\hat{r}(t)\}$ and associated covariances from the algorithm of Appendix A. Then we wish to solve the TPBVP posed by eqns. (2.20), (2.24), (2.25) and (2.29)-(2.31) along with associated boundary conditions, backward in time for all $t \in [t_o, t_f]$ with initial conditions

$$\hat{x}(y,t_f|t_f) = \hat{x}(y,t_f),\ \hat{q}(t_f|t_f) = \hat{q}(t_f),\ \hat{r}(t_f|t_f) = \hat{r}(t_f) \tag{2.32}$$

$$\lambda_1(y,t_f|t_f) = \lambda_1(y,t_f),\ \lambda_3(t_f|t_f) = \lambda_3(t_f),\ \lambda_4(t_f|t_f) = \lambda_4(t_f) \tag{2.33}$$

where $\hat{x}(y,t|t_f)$, $\hat{q}(t|t_f)$, $\hat{r}(t|t_f)$ denote smoothed estimates. In terms of our new notation for the smoothing solution and the initial conditions at at the final time t_f, we can write the TPBVP as

$$\frac{\partial \hat{x}(y,t|t_f)}{\partial t} = N(\hat{x}(y,t|t_f),y,\hat{q}(t|t_f),t) - \int_\Omega R^{-1}(y,y_1,t)\lambda_1(y,t|t_f)d\Omega_1 \tag{2.34}$$

$$\dot{\hat{q}}(t|t_f) = N_{\hat{q}}(\hat{r}(t|t_f),t) - T_1^{-1}(t)\lambda_3(t|t_f) \tag{2.35}$$

$$\dot{\hat{r}}(t|t_f) = N_{\hat{r}}(\hat{r}(t|t_f),t) - T_2^{-1}(t)\lambda_4(t|t_f) \tag{2.36}$$

with initial conditions given by (2.32) and boundary condition

$$[N_b]_{\hat{x}}\ \hat{x}(y,t|t_f) = 0 \qquad y \in \partial\,\Omega \tag{2.37}$$

The adjoint equations are

$$\frac{\partial \lambda_1(y,t|t_f)}{\partial t} = -[N]^*_{\hat{x}(y,t|t_f)}\lambda_1(y,t|t_f) + \int_\Omega \left[\frac{\partial h}{\partial \hat{x}(y,t|t_f)}\right]^T Q(z - h(x(y_1,t|t_f),\ y_1,t|t_f))d\Omega_1 \tag{2.38}$$

$$-\dot{\lambda}_3(t|t_f) = \int_\Omega [N]_{\hat{q}(t|t_f)}\lambda_1(y,t|t_f)d\Omega + [N_{\hat{q}}]_{\hat{q}(t|t_f)}\lambda_3(t|t_f) \qquad (2.39)$$

$$-\dot{\lambda}_4(t|t_f) = \int_{\partial\Omega}[N_b]^T_{\hat{r}}K\lambda_1(y,t|t_f)d(\partial\Omega) + [N_{\hat{r}}]_{\hat{r}(t|t_f)}\lambda_4(t|t_f) \qquad (2.40)$$

with boundary conditions

$$[N_b]_{\hat{x}(y,t|t_f)}\lambda_1(y,t|t_f) = 0 \qquad y \in \partial\,\Omega \qquad (2.41)$$

Eqns. (2.32-41) now constitute the fixed interval smoothing TPBVP. As is apparent from (2.32) and (2.33) the filtered estimate $\{\hat{x}(y,t_f), \hat{q}(t_f), \hat{r}(t_f)\}$ and the terminal adjoints $\lambda_1(y,t_f)$, $\lambda_3(t_f)$, $\lambda_4(t_f)$ constitute a set of boundary conditions at the terminal stage which convert our TPBVP to an initial value problem which may be backswept to obtain a fixed interval smoothing estimate. To avoid stability problems associated with reverse time solution of canonic equations, it is preferable to make a linear approximation for the λ's that expresses the dependence of the fixed-interval smoothing solution on the filtered solution. For linear systems (cf. Chapter III) this would be equivalent to a Riccatti transformation of the adjoint variable. From eqn. (2.15) we see that an appropriate choice is

$$\lambda_1(y,t|t_f) = -\int_\Omega P^{-1}_{11}(y,y_1,t)(\tilde{x}(y,t|t_f))d\Omega - P^{-1}_{13}(y,t)\tilde{q}(t|t_f)$$

$$- P^{-1}_{14}(y,t)\tilde{r}(t|t_f) \qquad (2.42a)$$

$$\lambda_3(t/t_f) = - \int_\Omega P^{-1}_{31}(y,t)\tilde{x}(y,t|t_f)d\Omega - P^{-1}_{33}(t)\tilde{q}(t|t_f)$$

$$- P^{-1}_{34}\tilde{r}(t|t_f) \qquad (2.42b)$$

$$\lambda_4(t|t_f) = - \int_\Omega P^{-1}_{41}(y,t)\tilde{x}(y,t|t_f)d\Omega - P^{-1}_{43}(t)\tilde{q}(t|t_f)$$

$$- P^{-1}_{44}(t)\tilde{r}(t|t_f) \qquad (2.42c)$$

where

$$\tilde{x}(y,t|t_f) = - \hat{x}(y,t) + \hat{x}(y,t|t_f) \qquad (2.43)$$

with similar expressions holding for $\tilde{q}(t|t_f)$ and $\tilde{r}(t|t_f)$. Substituting (2.42a,b,c) into (2.34), (2.35), (2.36) respectively yields for the smoothed estimate

$$\frac{\partial\hat{x}(y,t|t_f)}{\partial t} = N(\hat{x}(y,t|t_f),\ \hat{q}(t|t_f),\ y,t) + \int_\Omega\int_\Omega R^{-1}(y,y_1,t)$$

$$P^{-1}_{11}(y_1,y_2,t)\tilde{x}(y_2,t|t_f)d\Omega_1 d\Omega_2 + \int_\Omega R^{-1}(y,y_1,t)P^{-1}_{13}(y_1,t)$$

$$\tilde{q}(t|t_f)d\Omega_1 + \int_\Omega R^{-1}(y,y_1,t)P^{-1}_{14}(y_1,t)\tilde{r}(t|t_f)d\Omega \qquad (2.44)$$

$$\dot{q}(t|t_f) = N_{\hat{q}}(\hat{q}(t|t_f),t)) + T^{-1}_1(t)\Big\{\int_\Omega P^{-1}_{31}(y,t)\tilde{x}(y,t|t_f)d\Omega + P^{-1}_{33}(t)\tilde{q}$$

$$+ P^{-1}_{34}(t)\tilde{r}(t|t_f)\Big\} \qquad (2.45)$$

$$\dot{\hat{r}}(t|t_f) = N_{\hat{r}}(\hat{r}(t|t_f),t) + T^{-1}_2(t)\Big\{\int_\Omega P^{-1}_{41}(y,t)\tilde{x}(y,t|t_f)d\Omega$$

$$+ P_{43}^{-1}(t)\tilde{q}(t|t_f) + P^{-1}(t)\tilde{r}(t|t_f)\Big\} \tag{2.46}$$

with boundary conditions

$$[N_b]_{\hat{x}}\, \hat{x}(y,t|t_f) = 0 \qquad y \in \partial\, \Omega \tag{2.47}$$

Eqns. (2.44)-(2.47) with terminal conditions given by (2.42) along with the filtering algorithm of Appendix A constitutes the smoothing algorithm.

<u>Analysis of Smoothing Structure</u>: An examination of the smoothed estimator eqns. (2.44)-(2.47) reveals that the estimator equations are independent of the smoothing error covariances, being dependent only upon the filtering error covariances. To obtain a measure of the smoothing error covariance for our very general model is rather cumbersome. However, for a somewhat simplified model the mathematics can be kept within manageable proportions. We make the simplifying assumption that the only disturbances acting on the system are distributed in nature. No lumped disturbances are assumed. This has the effect of eliminating the q and r terms from our system model. Concomitantly, the smoothed estimate becomes (cf. eqn. (2.44))

$$\frac{\partial \hat{x}(y,t|t_f)}{\partial t} = N(\hat{x}(y,t|t_f),y,t) + \int_\Omega \int_\Omega R^{-1} P_{11}^{-1} \tilde{x}\; d\Omega_1 d\Omega_2 \tag{2.48}$$

with boundary conditions given by (2.47). Define the smoothing error covariance in the usual manner, that is ,

$$P_{11}(y,y_1,t|t_f) = E\,\{\delta x(y,t|t_f)\delta x^T(y_1,t|t_f)\} \tag{2.49}$$

where P_{11} is a symmetric, positive definite matrix. With this definition, the covariance may now be calculated, as for the lumped case, following Bryson and Frazier [26], to be compared vis-a-vis the filtering error covariance.

To perform this calculation, linearize eqn. (2.48) along a nominal trajectory with its boundary condition (2.47). This involves terms in $\hat{x}(y,t)$ and $x(y,t)$, the filtered estimate and the noisy states, too. So linearize the filtered estimate and the noisy TPBVP also, along with their boundary conditions. This gives rise to a linear TPBVP akin to the one for lumped systems treated by Bryson and Frazier [26]. (See Sec. 3 on fixed interval smoothing, too). The solution of this linear TPBVP is easy to write in terms of its Green's function, to yield $\delta\hat{x}(y,t|t_f)$. Substituting this result into (2.49) and differentiating yields

$$\frac{\partial P_{11}(y,y_1,t|t_f)}{\partial t} = MP_{11}(y,y_1,t|t_f) + P_{11}(y,y_1,t|t_f)M^T - R^{-1}(y,y_1,t) \tag{2.50}$$

with boundary and final conditions

$$[N_b]_{\hat{x}}\, P_{11}(y,y_1,t|t_f) = 0 \qquad y \,\varepsilon\, \partial\,\Omega, y_1 \varepsilon \overline{\Omega} \tag{2.51}$$

$$P_{11}(y,y_1,t_f|t_f) = P_{11}(y,y_1,t_f) \tag{2.52}$$

To summarize, the fixed interval smoothing estimate is obtained using a Riccatti transformation on the adjoint variable. It involves processing in reverse time after forward time computation of the filtered estimate has been completed, and hence is necessarily offline. The estimate is independent of the smoothing error covariance, being dependent on the filtering error covariance only. The smoothing error covariance has been obtained by linearizing the estimate about a nominal trajectory and solving the resulting BVP. Results in this section are extensions of previous results [10]. The estimator is nonlinear, while the only previous result existing [10] describes a linearized estimator.

The fixed-interval smoothing estimate can be tabulated thus (after processing the filtering algorithm of Appendix A),

Table 2. Fixed Interval Smoothing Algorithm

Estimator	Estimator Eqn.	Boundary Condition	Terminal Conditions
$\hat{x}(y,t\|t_f)$	(2.44)	(2.47)	(2.42a)
$\hat{q}(t\|t_f)$	(2.45)		(2.42b)
$\hat{r}(t\|t_f)$	(2.46)		(2.42c)

2.3b. Fixed Point Smoothing

The fixed point smoothing problem was outlined in Sec. 2.1. Here we require the smoothed estimate at only one point t_1 within the observation interval $[t_o,t_f]$, denoted by $S_{t_1|t}$. As for the case of fixed interval smoothing, at the point t_1, the filtered estimate $\{\hat{x}(y,t_1)$, $\hat{q}(t_1)$, $\hat{r}(t_1)\}$ constitutes boundary conditions for the smoothed estimate.

Likewise for the adjoint variables arising in the TPBVP formulated in Sec. 2.2. Thus we may write at the time t_1

$$\hat{x}(y,t_1|t) = \hat{x}(y,t_1),\ \hat{q}(t_1|t) = \hat{q}(t_1),\ \hat{r}(t_1|t) = \hat{r}(t_1) \tag{2.53}$$

$$\lambda_1(y,t_1|t) = \lambda_1(y,t_1),\ \lambda_3(t_1|t) = \lambda_3(t_1),\ \lambda_4(t_1|t) = \lambda_4(t_1) \tag{2.54}$$

As for the filtering solution, t_f is considered a running variable, and a 'fixed stage' invariant embedding solution [13] to the TPBVP of Sec. 2.2 obtained.

As in Sec. 2.3a, we denote the general class of solutions to the BVP by

$$\begin{aligned} x(y,t_1|t_f) &= r_1(C_1,\ C_3,\ C_4,\ y,\ t_1,\ t_f) \\ q(t_1|t_f) &= r_3(C_1,\ C_3,\ C_4,\ t_1,\ t_f) \\ r(t_1|t_f) &= r_4(C_1,\ C_3,\ C_4,\ t_1,\ t_f) \end{aligned} \tag{2.55}$$

where C_1, C_3, C_4 are as defined in eqn. (A.1). The desired solution obviously is $r_1(0,0,0,y,t_1,t_f)$, $r_3(0,0,0,t_1,t_f)$, $r_4(0,0,0,t_1,t_f)$ since the C's represent perturbations about the optimum. Also, it is apparent from (2.53) that a final condition on (2.55) at time t_1 is

$$\begin{aligned} r_1(C_1,\ C_3,\ C_4,\ y,\ t_1,\ t_1) &= S_1(C_1,\ C_3,\ C_4,\ y,\ t_1) \\ r_3(C_1,\ C_3,\ C_4,\ t_1,\ t_1) &= S_3(C_1,\ C_3,\ C_4,\ t_1) \\ r_4(C_1,\ C_3,\ C_4,\ t_1,\ t_1) &= S_4(C_1,\ C_3,\ C_4,\ t_1) \end{aligned} \tag{2.56}$$

where S_i, $i = 1,3,4$ represents the filtering solution obtained by solving the "running-time" invariant embedding equations of Appendix A.

Expanding both sides of (2.55) in a functional Taylor series expansion and proceeding in the same manner as in Appendix A, yields in the limit as $\Delta t \to 0$, the 'fixed stage' invariant embedding equations

$$\frac{\partial r_i}{\partial t_f} - \int_\Omega \frac{\delta r_i}{\delta C_1} \frac{\delta H}{\delta r_1} \, d\Omega - \frac{\partial r_i}{\partial C_3} \frac{\partial H}{\partial r_3} - \frac{\partial r_i}{\partial C_4} \frac{\partial H}{\partial r_4} = 0 \qquad i = 1,3,4 \quad (2.57)$$

Equations (2.57) alone with the 'running-time' invariant embedding equations (A.6) constitute an IVP. Consequently, we can consider only r_1, r_3, r_4 since the s_i, $i = 1,3,4$ generate the filtered estimate of Appendix A. To a first order approximation, assume solutions of the form,

$$r_1(C_1,C_3,C_4,y,t_1,t_f) = \hat{x}(y,t_1|t_f) - \int_\Omega P_{11}(y,y_1,t_1,t_f)C_1(y_1) \, d\Omega_1$$

$$- P_{13}(y,t_1,t_f)C_3 - P_{14}(y,t_1, t_f) \qquad (2.58)$$

$$r_3(C_1,C_3,C_4,t_1,t_f) = \hat{q}(t_1|t_f) - \int_\Omega P_{31}(y,t_1,t_f)C_1(y) \, d\Omega$$

$$- P_{33}(t_1,t_f)C_3 - P_{34}(t_1,t_f)C_4 \qquad (2.59)$$

$$r_4(C_2,C_3,C_4,t_1,t_f) = \hat{r}(t_1|t_f) - \int_\Omega P_{41}(y,t_1,t_f)C_1(y) d\Omega$$

$$-P_{43}(t_1,t_f)C_3 - P_{44}(t_1,t_f)C_4 \qquad (2.60)$$

with terminal conditions

$$\begin{aligned}\hat{x}(y,t_1|t_1) &= \hat{x}(y,t_1)\\ \hat{q}(t_1|t_1) &= \hat{q}(t_1)\\ \hat{r}(t_1|t_1) &= \hat{r}(t_1)\end{aligned} \tag{2.61}$$

$$\begin{aligned}P_{ij}(t_1,t_1) &= P_{ij}(t_1) \qquad i,j = 3,4\\ P_{31}(y,t_1,t_1) &= P_{31}(y,t_1)\\ P_{13}(y,t_1,t_1) &= P_{13}(y,t_1)\\ P_{14}(y,t_1,t_1) &= P_{14}(y,t_1)\\ P_{41}(y,t_1|t_1) &= P_{41}(y,t_1)\\ P_{11}(y,y_1,t_1|t_1) &= P_{11}(y,y_1,t_1)\end{aligned} \tag{2.62}$$

Substituting (2.58-2.60) into each of the equations (2.57) respectively yields after carrying out a Taylor series expansion as in the Appendix A,

$$\frac{\partial\hat{x}(y,t_1|t_f)}{\partial t_f} = \int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}(y,t_f)}\right)^T Q[z - h(\hat{x}(y_1,t_f)y_1,t_f)]d\Omega_1 d\Omega \tag{2.63}$$

$$\frac{dq(t_1|t_f)}{dt_f} = \int_\Omega\int_\Omega P_{31}(y,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}(y,t_f)}\right)^T Q[z - h(\hat{x}(y_1,t_f),y_1,t_f)]d\Omega_1 d\Omega \tag{2.64}$$

$$\frac{d\hat{r}(t_1|t_1)}{dt_f} = \int_\Omega\int_\Omega P_{41}(y,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}(y,t_f)}\right)^T$$

$$Q[z - h(\hat{x}(y_1,t_f),y_1,t_f)]d\Omega_1 d\Omega \qquad (2.65)$$

with boundary conditions

$$[N_b]_{\hat{x}(y,t_f)} \; \hat{x}(y,t_1|t_f) = 0 \qquad y \; \varepsilon \; \partial\,\Omega \qquad (2.66)$$

The moment algorithms are obtained by comparing coefficients of first order terms in C's in the Taylor series expansion. This yields

$$\frac{\partial P_{11}(y,y_1,t_1,t_f)}{\partial t_f} = P_{11}(y,y_1,t_1,t_f)[N]^T_{\hat{x}(y_1,t_f)} + P_{13}(y,t_1,t_f)[N]^T_{\hat{q}(t_f)}$$

$$+ \int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)\left[\left(\frac{\partial h}{\partial\hat{x}(y_1,t_f)}\right)^T Q(z-h(\hat{x},y_1,t_f))\right]_{\hat{x}}$$

$$P_{11}(y_3,y_1,t_f)\; d\Omega_2 d\Omega_3 \qquad (2.67)$$

$$\frac{\partial P_{13}(y,t_1,t_f)}{\partial t_f} = P_{13}(y,t_1,t_f)[N_{\hat{q}}]^T_{\hat{q}(t_f)} + \int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)$$

$$\left[\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q(z-h(\hat{x}(y_1,t_f))\right]_{\hat{x}} P_{13}(y_2,t_f)d\Omega_1 d\Omega_2 \qquad (2.68)$$

$$\frac{\partial P_{14}(y,t_1,t_f)}{\partial t_f} = P_{14}(y,t_1,t_f)[N_{\hat{r}}]^T_{\hat{r}(t_f)} + \int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)$$

$$\left[\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q(z-h(\hat{x}(y_2,t_f),y_2,t_f))\right]_{\hat{x}} P_{14}(y_2,t_f)d\Omega_1 d\Omega_2 \qquad (2.69)$$

$$\frac{\partial P_{31}(y,t_1,t_f)}{\partial t_f} = P_{31}(y,t_1,t_f)[N]^T_{\hat{x}(y_1,t_f)} + P_{33}(t_1,t_f)[N]^T_{\hat{q}(t_f)} + \int_\Omega \int_\Omega P_{31}(y,t_1,t_f)\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z-h(\hat{x}(y_1,t_f),y_2,t_f))\right]_{\hat{x}} P_{11}(y_1,y_2,t_f)\,d\Omega_1\,d\Omega_2 \tag{2.70}$$

$$\frac{\partial P_{33}(y,t_1,t_f)}{\partial t_f} = P_{33}(t_1,t_f)[N_{\hat{q}}]^T_{\hat{q}(t_f)} + \int_\Omega \int_\Omega P_{31}(y,t_1,t_f)\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z-h)\right]_{\hat{x}} P_{13}(y_1,\ t_f)\,d\Omega_1\,d\Omega \tag{2.71}$$

$$\frac{\partial P_{41}(y,t_1,t_f)}{\partial t_f} = P_{41}(y,t_1,t_f)[N]^T_{\hat{x}} + P_{43}(t_1,t_f)\ [N]^T_{\hat{q}} + \int_\Omega \int_\Omega P_{41}(y,t_1,t_f)\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z-h(\hat{x}(y_1,t_f),y_1,t_f))\right]_{\hat{x}} P_{11}(y_1,y_2,t_f)\,d\Omega_1\,d\Omega_2 \tag{2.72}$$

$$\frac{\partial P_{43}(t_1,t_f)}{\partial t_f} = P_{43}(t_1,t_f)\ [N_{\hat{q}}]^T_{\hat{q}(t_f)} + \int_\Omega \int_\Omega P_{41}(y,t_1,t_f)\left[\left(\frac{\partial h}{\partial x}\right)^T Q(z-h)\right]_{\hat{x}} P_{13}(y_1,t_f)\,d\Omega\,d\Omega_1 \tag{2.73}$$

$$\frac{\partial P_{44}(t_1,t_f)}{\partial t_f} = P_{44}(t_1,t_f)[N_{\hat{r}}]^T_{\hat{r}(t_f)} + \int_\Omega \int_\Omega P_{41}(y,t_1,t_f)\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z-h)\right]_{\hat{x}} P_{14}(y_1,t_f)\,d\Omega\,d\Omega_1 \tag{2.74}$$

where $[\cdot]_{\hat{x}}$ denotes differentiation w.r.t. the filtered estimate $\hat{x}(y,t_f)$. The boundary conditions are given by

$$P_{11}:\ [N_b]^*_{\hat{x}}P_{11}(y,y_1,t_1,t_f) + P_{14}(y,y_1,t_1,t_f)[N_b]_{\hat{r}(t_f)}K(y_1,t_f) = 0$$

$$y,\ y_1 \varepsilon\ \partial\,\Omega \qquad (2.75)$$

$$P_{13}:\ [N_b]_{\hat{x}}P_{13}(y,t_1,t_f) + [N_b]_{\hat{r}(t_f)}P_{43}(t_1,t_f) = 0 \quad y\ \varepsilon\ \partial\,\Omega \qquad (2.76)$$

$$P_{14}:\ [N_b]_{\hat{x}}P_{14}(y,t_1,t_f) + [N_b]_{\hat{r}(t_f)}P_{44}(t_1,t_f) = 0 \quad y\ \varepsilon\ \partial\,\Omega \qquad (2.77)$$

$$P_{31}:\ P_{31}(y,t_1,t_f)[N_b]^T_{\hat{x}} + P_{34}(t_1,t_f[N_b]_{\hat{r}} = 0 \qquad y\ \varepsilon\ \partial\ \Omega \qquad (2.78)$$

$$P_{41}:\ P_{41}(y,t_1,t_f)[N_b]^T_{\hat{x}} + P_{44}(t_1,t_f)[N_b]^T_{\hat{r}(t_f)} \qquad y\ \varepsilon\ \partial\ \Omega \qquad (2.79)$$

The fixed point smoothing algorithm can now be summarized. Along with this algorithm, the filtering algorithm Appendix A has to be processed. As before, the fixed point smoothing estimate is independent of the various error covariances. The fixed point algorithm covariances consists of processing the filtering algorithm forward in time along with the algorithm in Table 3. The approximate error covariances can be computed by the linearization procedure of Sec. 2.3b. Define the error covariances as

$$P_{11}(y,y_1,t_1|t_f) = \varepsilon\{\delta\hat{x}(y,t_1|t_f)\delta\hat{x}^T(y_1,t_1|t_f)\}$$

$$P_{13}(y,t_1|t_f) = \varepsilon\{\delta\hat{x}(y,t_1|t_f)\delta\hat{q}^T(t_1|t_f)\}$$

Table 3. Fixed Point Smoothing Algorithm

		Equation	I.C.	B.C.
Estimate	$\hat{x}(y,t_1 \mid t_f)$	(2.63)	(2.61)	(2.66)
	$\hat{q}(t_1 \mid t_f)$	(2.64)	(2.61)	none
	$\hat{r}(t_1 \mid t_f)$	(2.65)	(2.61)	none
Moments	$P_{11}(y,y_1,t_1,t_f)$	(2.67)	(2.62)	(2.75)
	$P_{13}(y,t_1,t_f)$	(2.68)	(2.62)	(2.76)
	$P_{14}(y,t_1,t_f)$	(2.69)	(2.62)	(2.77)
	$P_{31}(y,t_1,t_f)$	(2.70)	(2.62)	(2.78)
	$P_{33}(t_1,t_f)$	(2.71)	(2.62)	none
	$P_{41}(t_1,t_f)$	(2.72)	(2.62)	none
	$P_{43}(y,t_1,t_f)$	(2.73)	(2.62)	(2.79)
	$P_{44}(t_1,t_f)$	(2.74)	(2.62)	none

$$\begin{aligned}
P_{14}(y,t_1|t_f) &= \varepsilon\{\delta\hat{x}(y,t_1|t_f)\delta\hat{r}^T(t_1|t_f)\} \\
P_{33}(t_1|t_f) &= \varepsilon\{\delta\hat{q}(t_1|t_f)\delta\hat{q}^T(t_1|t_f)\} \\
P_{34}(t_1|t_f) &= \varepsilon\{\delta\hat{q}(t_1|t_f)\delta\hat{r}^T(t_1|t_f)\} \\
P_{44}(t_1|t_f) &= \varepsilon\{\delta\hat{r}(t_1|t_f)\delta\hat{r}^T(t_1|t_f)\}
\end{aligned} \tag{2.80}$$

where the P's are symmetric, positive definite matrix functions. Carrying out the linearization procedure yields

$$\frac{\partial P_{11}(y,y_1,t_1|t_f)}{\partial t_f} = -\int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q(y_2,y_3,t)\left(\frac{\partial h}{\partial\hat{x}}\right) P_{11}(y_3,y_1,t_1,t_f)\,d\Omega_2\,d\Omega_3 \tag{2.81}$$

$$\frac{\partial P_{13}(y,t_1|t_f)}{\partial t_f} = -\int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q\left(\frac{\partial h}{\partial\hat{x}}\right) P_{13}(y_2,t_1,t_f)\,d\Omega_1\,d\Omega_2 \tag{2.82}$$

$$\frac{\partial P_{14}(y,t_1|t_f)}{\partial t_f} = -\int_\Omega\int_\Omega P_{11}(y,y_1,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q\left(\frac{\partial h}{\partial\hat{x}}\right) P_{14}(y_2,t_1,t_f)\,d\Omega_1\,d\Omega_2 \tag{2.83}$$

$$\frac{\partial P_{33}(t_1|t_f)}{\partial t_f} = -\int_\Omega\int_\Omega P_{31}(y,t_1,t_f)\left(\frac{\partial h}{\partial\hat{x}}\right)^T Q\left(\frac{\partial h}{\partial\hat{x}}\right) P_{13}(y_1,t_1,t_f)\,d\Omega_1\,d\Omega \tag{2.84}$$

$$\frac{\partial P_{34}(t_1|t_f)}{\partial t_f} = - \int_\Omega\int_\Omega P_{31}(y,t_1,t_f)\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q\left(\frac{\partial h}{\partial \hat{x}}\right) P_{14}(y_1,t_1,t_f)d\Omega d\Omega_1 \tag{2.85}$$

$$\frac{\partial P_{33}(t_1|t_f)}{\partial t_f} = - \int_\Omega\int_\Omega P_{41}(y,t_1,t_f)\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q\left(\frac{\partial h}{\partial \hat{x}}\right) P_{14}(y_1,t_1,t_f)d\Omega d\Omega_1 \tag{2.86}$$

with boundary conditions

$$[N_b]_{\hat{x}}\, P_{11}(y,y_1,t_1|t_f) = 0 \quad y \in \partial\,\Omega, \quad y_1 \in \bar{\Omega} \tag{2.87}$$

$$[N_b]_{\hat{x}}\, P_{13}(y,t_1|t_f) = 0 \quad y \in \partial\,\Omega \tag{2.88}$$

$$[N_b]_{\hat{x}}\, P_{14}(y,t_1|t_f) = 0 \quad y \in \partial\,\Omega \tag{2.89}$$

and initial conditions given by the filtering error covariances at the time t_1, i.e.,

$$P_{11}(y,y_1,t_1|t_1) = P_{11}(y,y_1,t_1)$$

$$P_{13}(y,t_1|t_1) = P_{13}(y,t_1), \; P_{14}(y,t_1|t_1) = P_{14}(y,t_1)$$

$$P_{33}(t_1|t_1) = P_{33}(t_1) \;, \; P_{34}(t_1|t_1) = P_{34}(t_1)$$

$$P_{44}(t_1|t_1) = P_{44}(t_1) \tag{2.90}$$

To summarize, the fixed point smoothing algorithm has been obtained from the 'fixed-time' invariant embedding equation using a Taylor series expansion about the filtered trajectory. The procedure consists of processing the filtered estimate alone upto time t_1, the point at which the smoothed estimate is desired, and then concurrently processing the filtering and smoothing algorithms. The smoothing algorithm is independent of its associated error covariances. The error covariances can be approximately computed by linearizing the TPBVP and solving it using Green's functions. The results thus obtained are extension of previous results and consider a far more general model than the one considered by Tzafestas [10], which takes into account only distributed Gaussian disturbances. Furthermore, these algorithms are nonlinear, while the ones in ref. [10] are linearized Kalman structures.

2.3c. Fixed-Lag Smoothing

In Sec. 2.3a we obtained a fixed-interval smoothing algorithm that yielded an approximate solution to the nonlinear TPBVP arising in Sec. 2.2. The algorithm must necessarily be processed off-line, and it is necessary to store the filtered estimate and the associated covariance matrix. Fixed-lag smoothing may be performed on-line, and storage requirements are less than for fixed interval smoothing. However, storage requirements exceed that for fixed point smoothing.

The fixed-interval smoothing problem was outlined in Sec. 2.2. In estimate S_t at time t, we use observations upto an interval $t + \tau$, where $t + \tau < t_f$. In general, the estimate over the required interval will not be as good as fixed-interval smoothing, but will result in estimates better than the filtered estimate in that additional observation information has been utilized. If the algorithms were exact, as they would be for linear systems, the fixed-lag smoothing error covariance would lie between the filtering covariance of error and fixed interval smoothing error covariance.

The procedure for solving the fixed lag smoothing algorithm is very akin to solving the TPBVP associated with the fixed point smoothing estimate. Denoting by T the time delay we denote the general class of solutions to the BVP by

$$\begin{aligned} x(y,t|t+T) &= r_1(C_1,C_3,C_4,y,t,t+T) \\ q(t|t+T) &= r_3(C_1,C_3,C_4,y,t,t+T) \\ r(t|t+T) &= r_4(C_1,C_3,C_4,y,t,t+T) \end{aligned} \tag{2.91}$$

where C_1, C_3, C_4 are as defined in eqn. (A.1). With the correspondence

$$t_1 \rightarrow t$$

$$t_f \rightarrow t+T$$

the fixed lag smoothing invariant embedding equations are (2.57) along with (A.6) with boundary conditions (2.56). Let us assume solutions to this initial value problem of the form (cf. eqns. (2.58)-(2.60))

$$\begin{aligned} r_1(C_1,C_3,C_4,y,t,t+T) &= \hat{x}(y,t|t+T) - \int_\Omega P_{11}(y,y_1,t,t+T)C_1(y_1)d\Omega_1 \\ &\quad - P_{13}(y,t,t+T)C_3 - P_{14}(y,t,t+T)C_4 \end{aligned} \tag{2.92}$$

$$\begin{aligned} r_3(C_1,C_3,C_4,y,t,t+T) &= \hat{q}(t|t+T) - \int_\Omega P_{31}(y,t,t+T)C_1(y)d\Omega \\ &\quad - P_{33}(t,t+T)C_3 - P_{34}(t,t+T)C_4 \end{aligned} \tag{2.93}$$

$$r_4(C_1,C_3,C_4,t,t+T) = \hat{r}(t|t+T) - \int_\Omega P_{41}(y,t,t+T)C_1(y)d\Omega$$
$$- P_{43}(t,t+T)C_3 - P_{44}(t,t+T)C_4 \qquad (2.94)$$

with terminal conditions $S_{t_o|t_o+T} = \{\hat{x}(y,t_o|t_o+T), \hat{q}(t_o|t_o+T), \hat{r}(t_o|t_o+T)\}$ obtained from the fixed point smoothing estimate. The quasi moments (the P terms) also have their terminal condition governed by the corresponding quasi-moments of the fixed point smoothing algorithm. Following the analogous procedure to Sec. 2.3b of substituting (2.92-94) into each of the equations (2.57) respectively and carrying out a Taylor series expansion yields

$$\frac{\partial\hat{x}(y,t|t+T)}{\partial t} = N(\hat{x}(y,t|t+T), \hat{q}(t+T),y,t) + \int_\Omega R^{-1}P_{13}^{-1}(y_1,t)\tilde{q}(t|t+T)$$
$$d\Omega_1 + \int_\Omega R^{-1}P_{14}^{-1}\tilde{r}(t|t+T)d\Omega_1 + \int_\Omega\int_\Omega R^{-1}P_{11}^{-1}(y_1,y_2,t)$$
$$\tilde{x}(y_2,t|t+T)d\Omega_1 d\Omega_2 + \int_\Omega\int_\Omega P_{11}(y,y_1,t,t+T)\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T$$
$$Q(y_1,y_2,t+T)(z - h(\tilde{x}(y_2,t|t+T),y_2,t+T)\ d\Omega_1 d\Omega_2$$
$$(2.95)$$

$$\dot{\hat{q}}(t|t+T) = N_q(\hat{q}(t|t+T),t) + T_1^{-1}(t) \int_\Omega P_{31}^{-1}(y,t)\tilde{x}(y,t|t+T)d\Omega$$
$$+ P_{33}^{-1}(t)\tilde{q}(t|t+T) + P_{34}^{-1}(t)\tilde{r}(t|t+T) +$$
$$+ \int_\Omega\int_\Omega P_{31}(y,t,t+T)\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T Q(y,y_1,t+T)$$
$$(z - h(\hat{x}(y_1,t+T),y_1,t+T))d\Omega d\Omega_1 \qquad (2.96)$$

$$\dot{r}(t|t+T) = N_r(\hat{r}(t|t+T),t) + T_2^{-1}(t)\int_\Omega P_{41}^{-1}(y,t)\tilde{x}(y,t|t+T)d\Omega$$

$$+ P_{43}^{-1}(t)\tilde{q}(t|t+T) + P_{44}^{-1}(t)\tilde{r}(t|t+T)$$

$$+ \int_\Omega\int_\Omega P_{41}(y,t,t+T)\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T Q(y,y_1,t+T)$$

$$(z - h(\hat{x}(y_1,t+T),y_1,t+T)d\Omega d\Omega_1 \qquad ((2.97)$$

with boundary conditions

$$[N_b]_{\hat{x}(y,t+T)}\hat{x}(y,t|t+T) = 0 \qquad y \in \partial\,\Omega \qquad (2.98)$$

where

$$\tilde{x}(y,t|t+T) = \hat{x}(y,t|t+T) - \hat{x}(y,t) \qquad (2.99)$$

and identical expressions hold for $\tilde{q}$, $\tilde{r}$.

The quasi-moment algorithms are obtained by comparing coefficients of first order terms in a Taylor series expansion. This yields

$$\frac{\partial P_{11}(y,y_1,t,t+T)}{\partial t} = \int_\Omega\int_\Omega P_{11}(y,y_2,t,t+T)\left[\left(\frac{\partial h}{\partial\hat{x}(y_2,t+T)}\right)^T Q(z-h)\right]_{\hat{x}(y_3,t+T)}$$

$$P_{11}(y_3,y_1,t+T)d\Omega_2 d\Omega_3 + P_{11}(y,y_1,t,t+T)$$

$$[N]^T_{\hat{x}(y_1,t+T)} + [N]^T_{\hat{x}(y,t+T)}P_{11}(y,y_1,t,t+T)$$

$$+ P_{13}(y,t,t+T)[N]^T_{\hat{q}(t+T)} + \int_\Omega\int_\Omega R^{-1}(y,y_2,t)$$

$$\begin{aligned} &P_{11}^{-1}(y_2,y_3,t)P_{11}(y_3,y_1,t,t+T)\,d\Omega_2 d\Omega_3 + [N]_{\hat{q}(t+T)} \\ &P_{31}(y,t,t+T) \end{aligned} \tag{2.100}$$

$$\begin{aligned} \frac{\partial P_{13}(y,t,t+T)}{\partial t} = {} & P_{13}(y,t,t+T)[N_q]_{\hat{q}(t+T)}^{T} + [N]_{\hat{x}}P_{13}(y_1,t,t+T) \\ & + \int_\Omega\int_\Omega P_{11}(y,y_1,t,t+T)\left[\left(\frac{\partial h}{\partial\hat{x}(t+T,y_1)}\right)^{T} Q(z - h(\hat{x}(y_2,t+T), \right. \\ & \left. y_2,t+T))\right]_{\hat{x}} P_{13}(y_2,t+T)\,d\Omega_1 d\Omega_1 + \int_\Omega R^{-1}(y,y_2,t) \\ & P_{11}^{-1}(y_1,y_2,t)P_{13}(y_2,t,t+T)\,d\Omega_1 \end{aligned} \tag{2.101}$$

$$\begin{aligned} \frac{\partial P_{14}(y,t,t+T)}{\partial t} = {} & P_{14}(y,t,t+T)\;[N_r]_{\hat{r}(t+T)}^{T} + [N]_{\hat{x}}\,P_{14}(y_1,t,t+T) \\ & + \int_\Omega\int_\Omega P_{11}(y,y_1,t,t+T)\left[\left(\frac{\partial h}{\partial\hat{x}(y_1,t+T)}\right)^{T} Q(z - h(\hat{x}(y_2,t+T), \right. \\ & \left. y_2,t+T))\right]_{\hat{x}} P_{14}(y_2,t+T)\,d\Omega_1 d\Omega_2 + \int_\Omega R^{-1}(y,y_1,t) \\ & P_{11}^{-1}(y_1,y_2,t)P_{14}(y_2,t,t+T)\,d\Omega_1 \end{aligned} \tag{2.102}$$

$$\begin{aligned} \frac{\partial P_{33}(y,t,t+T)}{\partial t} = {} & P_{33}(y,t,t+T)\;[N]_{\hat{x}(y,t+T)}^{T} + [N_q]_{\hat{q}(t+T)}P_{33}(y,t,t+T) \\ & + \int_\Omega\int_\Omega P_{31}(y,t,t+T)\left[\left(\frac{\partial h}{\partial\hat{x}(t+T,y)}\right)^{T} Q(z - h(\hat{x}(y_1,t+T), \right. \end{aligned}$$

$$y_1, t + T))\Bigg]_{\hat{x}} P_{11}(y_1,y_2,t + T)d\Omega_1\, d\Omega_2 + \int_\Omega T_1^{-1}(t)$$

$$P_{31}^{-1}(y,t)P_{11}(y,y_1,t,t + T)d\Omega_1 \qquad (2.103)$$

$$\frac{\partial P_{33}(t,t+T)}{\partial t} = P_{33}(t,t + T)\ [N_q]^T_{\hat{q}(t+T)} + [N_q]_{\hat{q}(t+T)} P_{33}(t,t + T)$$

$$+ T_1^{-1}P_{33}^{-1}(t)P_{33}(t,t + T) + \int_\Omega\int_\Omega P_{31}(y,t,t + T)\left[\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T\right.$$

$$\left. Q(z - h(\hat{x}(y_1,t + T),\ y_1,T + t))\right]_{\hat{x}} P_{13}(y,t + T)d\Omega d\Omega_1$$

$$(2.104)$$

$$\frac{\partial P_{41}(y,t,t+T)}{\partial t} = P_{41}(y,t,t + T)\ [N]^T_{\hat{x}} + (N_r]_{\hat{r}(t+T)} P_{41}(y,t,t + T)$$

$$+ \int_\Omega\int_\Omega P_{41}(y,t,t + T)\left[\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T Q(z - h(\hat{x}(y_1,t + T),\right.$$

$$\left. y_1,t + T))\right]_{\hat{x}} P_{11}(y_1,y_2,t + T)d\Omega_1 d\Omega_2 + \int_\Omega T_2^{-1}(t)$$

$$P_{41}^{-1}(y,t)P_{11}(y_1,t,t + T)d\Omega_1 \qquad (2.105)$$

$$\frac{\partial P_{43}(t,t+T)}{\partial t} = P_{43}(t,t + T)\ [N_q]^T_{\hat{q}(t+T)} + [N_q]_{\hat{q}} P_{43}(t,t + T)$$

$$+ \int_\Omega\int_\Omega P_{41}(y,t,t + T)\left[\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T Q(z - h(\hat{x}(y_1,t + T),\right.$$

$$\left. y_1,t + T))\right]_{\hat{x}} P_{13}(y_1,t + T)d\Omega d\Omega_1 + T_2^{-1}(t)P_{43}^{-1}(t)P_{33}(t,t+T)$$

$$(2.106)$$

$$\frac{\partial P_{44}(t,t+T)}{\partial t} = P_{44}(t,t+T)\,[N_r]^T_{\hat{r}(t+T)} + [N_r]_{\hat{r}}P_{44}(t,t+T) + T^{-1}(t)$$

$$P_{44}(t)P_{44}(t,t+T) + \int_\Omega\int_\Omega P_{41}(y,t,t+T)\left[\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T\right.$$

$$\left. Q(z-h)\right]_{\hat{x}} P_{14}(y_1,t+T)\,d\Omega\, d\Omega_1 \qquad (2.107)$$

with initial conditions determined from corresponding terms in the fixed point smoothing algorithm, and boundary conditions

$$P_{11} : [N_b]^*_{\hat{x}}P_{11}(y,y_1,t,t+T) + P_{14}(y,y_1,t,t+T)[N_b]_{\hat{r}(t+T)}K(y_1,t_f) = 0$$

$$y,\ y_1 \in \partial\,\Omega \qquad (2.108)$$

$$P_{13} : [N_b]_{\hat{x}(y,t+T)}P_{13}(y,t,t+T) + [N_b]_{\hat{r}(t+T)}P_{43}(t,t+T) = 0$$

$$y \in \partial\Omega \qquad (2.109)$$

$$P_{14} : [N_b]_{\hat{x}(y,t+T)}P_{14}(y,t,t+T) + [N_b]_{\hat{r}(t+T)}P_{44}(t,t+T) = 0$$

$$y \in \partial\Omega \qquad (2.110)$$

$$P_{31} : P_{31}(y,t,t+T)\ [N_b]^T_{\hat{x}(y,t+T)} + P_{34}(t,t+T)[N_b]_{\hat{r}(t+T)} \qquad (2.111)$$

$$P_{41} : P_{41}(y,t,t+T)[N_b]^T_{\hat{x}(y,t+T)} + P_{44}(t,t+T)[N_b]^T_{\hat{r}(t+T)} \qquad (2.112)$$

The fixed lag smoothing algorithm can now be summarized.

Table 4. Fixed Lag Smoothing Algorithm

		Equation	I.C.	B.C.
Estimates				
	$\hat{x}(y,t \mid t+T)$	(2.95)	From fixed point smoothing algorithm	(2.98)
	$\hat{q}(t \mid t+T)$	(2.96)	"	none
	$\hat{r}(t \mid t+T)$	(2.97)	"	none
Moments	$P_{11}(y,y_1,t,t+T)$	(2.100)	"	(2.108)
	$P_{13}(y,t,t+T)$	(2.101)	"	(2.109)
	$P_{14}(y,t,t+T)$	(2.102)	"	(2.110)
	$P_{31}(y,t,t+T)$	(2.103)	"	(2.111)
	$P_{33}(t,t+T)$	(2.104)	"	none
	$P_{41}(y,t,t+T)$	(2.105)	"	(2.112)
	$P_{43}(t,t+T)$	(2.106)	"	none
	$P_{44}(t,t+T)$	(2.107)	"	none

To recapitulate, the fixed lag smoothing algorithm has been obtained from the 'fixed time' invariant embedding equations. It is seen that the estimates are independent of their error covariances. An approximate computation of the error covariances can be made by linearizing the TPBVP and solving the resulting linear system using the technique described in Sec. 2.3b. We define the error covariances thus

$$P_{11}(y,y\ ,t|t + T) = \varepsilon\{\delta\hat{x}(y,t|t + T)\delta\hat{x}^T(y_1,t|t + T)\}$$

$$P_{13}(y,t|t + T) = \varepsilon\{\delta\hat{x}(y,t|t + T)\delta\hat{q}^T(t|t + T)\}$$

$$P_{14}(y,t|t + T) = \varepsilon\{\delta\hat{x}(y,t|t + T)\delta\hat{r}^T(t|t + T)\}$$

$$P_{33}(t|t + T) = \varepsilon\{\delta\hat{q}(t|t + T)\delta\hat{r}^T(t|t + T)\}$$

$$P_{34}(t|t + T) = \varepsilon\{\delta\hat{q}(t|t + T)\delta\hat{r}^T(t|t + T)\}$$

$$P_{35}(t|t + T) = \varepsilon\{\delta\hat{r}(t|t + T)\delta\hat{r}^T(t|t + T)\}$$

where the P's are symmetric, positive definite matrix functions. Carry ing out the linearization procedure yields

$$\frac{\partial P_{11}(y,y_1,t,t+T)}{\partial t} = - \int_\Omega\int_\Omega P_{11}(y,y_2,t,t + T)\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right)^T Q(y_2,y_3,t)$$

$$\left(\frac{\partial h}{\partial\hat{x}(y,t+T)}\right) P_{11}(y_3,y_1,t,t + T)d\Omega_2 d\Omega_3 - R^{-1}(y,y_1,t)$$

$$+ [N]_{\hat{x}(y,t)} P_{11}(y,y_1,t\ \ t + T) + P_{11}(y,y_1,t|t + T)$$

$$[N]^{T}_{\hat{x}(y_1,t)} + [N]_{\hat{q}(t)} P_{31}(y,t|t+T)$$

$$+ P_{13}(y,t|t+T)[N]^{T}_{\hat{q}(t)} + \Xi_1 P_{11}(y,y_1,t|t+T)$$

$$+ P_{11}(y,y_1,t|t+T)\Xi_1^{T} \qquad (2.114)$$

where

$$\Xi_1 P_{11}(y,y_1,t|t+T) = \int_\Omega\int_\Omega R^{-1}(y,y_2,t)P_{11}^{-1}(y_2,y_3,t)P_{11}(y_3,y_1,t|t+T)d\Omega_2 d\Omega_3$$

$$\frac{\partial P_{31}(y,t|t+T)}{\partial t} = P_{31}(y,t|t+T)$$

$$[N_q]^{T}_{\hat{q}(t)} + [N]_{\hat{x}(y,t)} P_{13}(y,t|t+T)$$

$$- \int_\Omega\int_\Omega P_{11}(y,y_2,t,t+T)\left(\frac{\partial h}{\partial \hat{x}(y_2,t+T)}\right)^{T} Q\left(\frac{\partial h}{\partial \hat{x}(y_3,t+T)}\right)$$

$$P_{13}(y_3,t,t+T)d\Omega_2 d\Omega_3 + \Xi_2 P_{13}(y,t|t+T)$$

$$+ P_{13}(y,t\ t+T)\Xi_2^{T} \qquad (2.115)$$

where

$$\Xi_2 P_{13}(y,t|t+T) = \int_\Omega\int_\Omega R^{-1}(y,y_2,t)P_{11}^{-1}(y_2,y_3,t)P_{13}(y_3,t|t+T)d\Omega_2 d\Omega_3$$

$$\frac{\partial P_{14}(y,t|t+T)}{\partial t} = P_{14}(y,t|t+T)\ [N_r]^{T}_{\hat{r}(t)}$$

$$+ [N]_{\hat{x}(y,t)} P_{14}(y,t|t+T) - \int_\Omega\int_\Omega P_{11}(y,y_2,t,\ t+T)$$

$$\left(\frac{\partial h}{\partial \hat{x}(y_2,t+T)}\right)^{T} Q\left(\frac{\partial h}{\partial \hat{x}(y_3,t+T)}\right)P_{14}(y_3,t,t+T)d\Omega_2 d\Omega_3$$

$$+ \Xi_3 P_{14}(y,t|t+T) + P_{14}(y,t|t+T)\Xi_3^T \tag{2.116}$$

where

$$\Xi_3 P_{14}(y,t|t+T) = \int_\Omega\int_\Omega R^{-1}P_{11}^{-1}(y_2,y_3,t)P_{14}(y_3,t|t+T)\,d\Omega_2\,d\Omega_3$$

$$\begin{aligned}\frac{\partial P_{31}(y,t|t+T)}{\partial t} &= P_{31}(y,t|t+T)\ [N]^T_{\hat{x}(t)} \\ &\quad + [N_q]_{\hat{q}(t)}P_{31}(y,t|t+T) - \int_\Omega\int_\Omega P_{31}(y,t,t+T) \\ &\quad \left[\frac{\partial h}{\partial\hat{x}(y,t+T)}\right]^T Q\left[\frac{\partial h}{\partial\hat{x}(y_1,t+T)}\right] P_{11}(y_1,y_2,t,t+T)\,d\Omega_1\,d\Omega_2 \\ &\quad + \Xi_4 P_{31}(y,t|t+T) + P_{31}(y,t|t+T)\Xi_4^T\end{aligned} \tag{2.117}$$

where

$$\Xi_4 P_{31}(y,t|t+T) = T_1^{-1}(t)\int_\Omega (P_{31}^{-1}(y,t)P_{11}(y,y_1,t,t+T)\,d\Omega_1$$

$$\begin{aligned}\frac{\partial P_{33}(t|t+T)}{\partial t} &= -\int_\Omega\int_\Omega P_{31}(y,t,t+T)\left[\frac{\partial h}{\partial\hat{x}(y,t+T)}\right]^T Q\left[\frac{\partial h}{\partial\hat{x}(y_1,t+T)}\right] \\ &\quad P_{13}(y_1,t,t+T)\,d\Omega_1\,d\Omega_2 - T_1^{-1}(t) + \Xi_5 P_{33}(t|t+T) \\ &\quad + P_{33}(t|t+T)\Xi_5^T\end{aligned} \tag{2.118}$$

where

$$\Xi_5 P_{33}(t|t+T) = [N_q]_{\hat{q}(t)}P_{33}(t|t+T) + T_1^{-1}(t)P_{33}^{-1}(t)P_{33}(t|t+T)$$

$$\frac{\partial P_{43}(t|t+T)}{\partial t} = - \int_{\Omega}\int_{\Omega} P_{41}(y,t,t+T) \left(\frac{\partial h}{\partial \hat{x}(y,t+T)}\right)^T Q\left(\frac{\partial h}{\partial \hat{x}(y_1,t+T)}\right) P_{13}(y_1,t+T)\,d\Omega d\Omega_1 + \Xi_6 P_{43}(t|t+T) P_{43}(t|t+T)\,\Xi_6{}^T \tag{2.119}$$

where

$$\Xi_6 P_{43}(t|t+T) = [N_q]_{\hat{q}(t)} + T_2^{-1}(t) P_{43}^{-1}(t) P_{33}(t|t+T)$$

$$\frac{\partial P_{44}(t|t+T)}{\partial t} = - \int_{\Omega} P_{41}(y,t,t+T)\left(\frac{\partial h}{\partial \hat{x}(y,t+T)}\right)^T Q\left(\frac{\partial h}{\partial \hat{x}(y_1,t+T)}\right) P_{14}(y_1,t,t+T)\,d\Omega_1 - T_2^{-1}(t) + \Xi_7 P_{44}(t|t+T) + P_{44}(t|t+T)\,\Xi_7{}^T \tag{2.120}$$

where

$$\Xi_7 P_{44}(t|t+T) = [N_r]_{\hat{r}(t)} P_{44}(t|t+T) + T_2^{-1}(t) P_{44}^{-1}(t) P_{44}(t|t+T)$$

with initial conditions obtained from the fixed point smoothing covariance algorithms, and boundary conditions

$$[N_b]_{\hat{x}(y,t)} P_{11}(y,y_1,t|t+T) = 0 \qquad y \,\epsilon\, \partial\,\Omega,\ y_1 \,\epsilon\, \bar{\Omega} \tag{2.121}$$

$$[N_b]_{\hat{x}} P_{13}(y,y_1,t|t+T) = 0 \qquad y \,\epsilon\, \partial\,\Omega \tag{2.122}$$

$$[N_b]_{\hat{x}} P_{14}(y,t|t+T) = 0 \qquad y \in \partial\Omega \qquad (2.123)$$

To recapitulate, the fixed lag smoothing algorithm has been obtained from the extended 'fixed-time' invariant embedding equation using a Taylor series expansion about the filtered trajectory. The procedure consists of processing the fixed point smoothing algorithm from the time t_o to time $t_o + T$, to obtain initial conditions for the fixed lag smoothing algorithm, and then simultaneously processing the fixed lag smoothing and filtering algorithms. The smoothing algorithm is independent of the associated error covariances. The error covariances can be approximately computed by linearizing the TPBVP about a nominal trajectory and then solving it, using Green's functions. The results thus obtained are extensions of previous results by Tzafestas [10] in that a more general model is considered for the system, with both volume and boundary disturbance. Furthermore the resulting algorithm is nonlinear in nature instead of extended Kalman.

2.4. Algorithms for Linear Systems

Until now, we have considered rather general second order operators with linear Gateaux derivatives for describing our system equation (2.1). We now show how the results obtained in the previous sections can be applied to linear systems. The results obtained in this section shall display the generality of the results obtained in previous sections.

For simplicity, we shall consider only the fixed interval smoothing algorithm for linear distributed systems alone. The other linear algorithms are obtained using an identical approach. With this assumption let the linear system be

$$\frac{\partial x(y,t)}{\partial t} = L_y x(y,t) + w(y,t) \qquad y \varepsilon \Omega \tag{2.124}$$

with boundary condition

$$N_b x(y,t) = 0 \qquad y \varepsilon \partial \Omega \tag{2.125}$$

where L_y and N_b have the elliptic form

$$L_y(\cdot) = \sum_{i,j=1}^{n} A_{ij}(y,t) \frac{\partial(\cdot)}{\partial y_i \partial y_j} + \sum_{i=1}^{n} B_i(y,t) \frac{\partial(\cdot)}{\partial y_i} + C(y,t)(\cdot)$$

$$N_b(\cdot) = A_{b_1}(\cdot) + A_{b_2}(y,t) \frac{\partial(\cdot)}{\partial\sigma}$$

where $\partial/\partial\sigma$ denotes partial differentiation along the conormal of $\partial\Omega$ relative to $N_b(\cdot)$.

The estimation criterion relative to the linear observation

$$z(y,t) = H(y,t)x(y,t) + v(y,t) \qquad y \varepsilon \Omega \tag{2.126}$$

is

$$J = \frac{1}{2} \int_{t_o}^{t_f} \int_\Omega \int_\Omega (z(y,t) - H(y,t)x(y,t))^T Q(y,y_1,t)(z(y_1,t) - H(y_1,t)$$

$$x(y_1,t) d\Omega d\Omega_1 + \int_\Omega \int_\Omega w^T(y,t) R(y,y_1,t) w(y_1,t) d\Omega d\Omega_1 \, dt \tag{2.127}$$

where Q and R are bounded, symmetric, symmetric in their spatial arguments and continuous.

The canonical equations then yield the estimate (cf. eqn. 2.74)

$$\frac{\partial \hat{x}(y,t|t_f)}{\partial t} = L_y\, x(y,t|t_f) + \int_\Omega \int_\Omega R^{-1}(y,y_1,t)P^{-1}(y_1,y_2,t)$$

$$(\tilde{x}(y_2,t|t_f)d\Omega_1 d\Omega_2 \tag{2.128}$$

with boundary conditions

$$[N_b]\hat{x}(y,t) = 0 \qquad y \in \partial\,\Omega$$

From this we see that the linear estimate is obtained from the nonlinear estimate by recognizing the following equivalence between linear and nonlinear systems

Nonlinear		Linear
$[N]_{\hat{x}}\ \hat{x}(y,t)$	$\longleftrightarrow$	$L_y\ \hat{x}(y,t)$
$[N_{\hat{q}}]_{\hat{q}}\ \hat{q}(t)$	$\longleftrightarrow$	$L_q\ \hat{q}(t)$
$[N_r]_{\hat{r}}\ \hat{r}(t)$	$\longleftrightarrow$	$L_r\ \hat{r}(t)$

because of the fact that the derivatives $[N]_{\hat{x}}$ $[N_q]_{\hat{q}}$ $[N_r]_{\hat{r}}$ are assumed linear (i.e., they are Frechet in nature). Similarly the equivalence $\frac{\partial h}{\partial \hat{x}} \rightarrow H(y,t)\hat{x}(y,t)$ holds.

With the above noted equivalences, the linear distributed parameter algorithms result. They are not summarized here as they follow easily from the nonlinear algorithms using the above relations.

Also, we notice that as the domain $\Omega \subset E^n$ shrinks to a point in E^n, that is, the state equation is given by

$$\frac{dx(t)}{dt} = \mathcal{L}x(y,t) + w(t) \quad , \qquad \mathcal{L} \text{ a bounded matrix}$$

the estimate reduces to the smoothed estimate for the finite dimensional case [12]

$$\hat{\dot{x}}(t|t_f) = \mathcal{L}\hat{x}(y,t|t_f) + R^{-1}P^{-1}(t)\tilde{x}(t|t_f) \qquad (2.219)$$

This points to the validity of our derivation. In fact, for all the nonlinear algorithms given earlier, as the domain Ω approaches a point, the algorithms collapse to corresponding finite dimensional algorithms [12], verifying their validity.

2.5. Summary

In this chapter, we have developed general filtering and smoothing algorithms for mixed distributed and lumped parameter systems with volume and boundary disturbances. Only volume measurements were considered. The estimation criterion was assumed to be a minimization of a positive definite quadratic form using classical least squares. The TPBVP arising from minimizing this quadratic functional was converted into an initial value problem using invariant embedding. To obviate introduction of noisy measurements into the smoothed estimate, recourse was made to results from linear estimation theory and the fixed interval smoothing estimate assumed a linear functional about the filtered trajectory. The other smoothed estimates were obtained from simultaneous solution of the 'fixed stage' and 'running stage' invariant embedding equations. It was found that all smoothing estimates were independent of their associated error covariances. These covariances were computed by linearizing the TPBVP and solving the resulting system using a Green's function approach.

Furthermore, linear algorithms are shown to be a subclass of the more general nonlinear algorithms. It is also shown that as the domain of estimation reduces to a point in Euclidean n-space, the algorithm collapse to the corresponding lumped parameter algorithms. This serves to prove the validity of these algorithms.

The important problem of numerical algorithms shall be considered in Chapter IV to demonstrate the feasibility of the previously developed algorithms.

CHAPTER III

INNOVATIONS APPROACH TO DISTRIBUTED ESTIMATION

In the previous chapter we considered the problem of estimating the states of quasilinear second order systems with additive dynamic and measurement noises. No assumption was made on the statistics of the noises, and the problem was treated as an optimal control problem in an infinite dimensional space. The optimal estimate was assumed to minimize a positive definite quadratic form, and the resulting two-point boundary value problem (TPBVP) converted into an initial value problem (IVP). This IVP was then solved approximately using a Taylor series expansion.

In case the noise statistics are known, however, a number of different approaches to the estimation problem can be taken, like the maximum a posteriori approach [10], maximum-likelihood ratio approach [9], etc. All these methods end up with an attempt to maximize a positive definite function, called the likelihood function. However, for noise densities other than Gaussian, the likelihood function is difficult to obtain, if not impossible, which severely restricts the applicability of the above methods. To obviate these difficulties, Kailath [15] developed the innovations approach to least squares estimation. This method, which is an extension of Bode and Shannon's [24] derivation of the stationary Wiener-filter, can be used with success for a large class of white and colored noise processes. Furthermore, unlike the methods of Chapter II, exact quantitative error calculations can be made, using this method.

For the analysis in this chapter, we consider linear second order systems with independent increment white additive dynamic and measurement noises. This is the DP analogue of the class of lumped linear systems considered by Kailath and Frost [15], [16]. The objective function is considered to be a minimum variance estimator objective function. The minimum variance estimator is obtained using innovations [15] and with the proper identification, reduces to the invariant embedding estimate of the previous chapter.

3.1. System Model and Problem Formulation

Consider the class of noisy linear second order DP systems described on an open bounded domain $\Omega \in E^n$, with piecewise smooth boundary $\partial\Omega$, by

$$\frac{\partial x(y,t)}{\partial t} = L_y x(y,t) + G(y,t)w(y,t) \;;\quad y \in \Omega,\ t \geq t_o \tag{3.1}$$

with boundary conditions

$$N_b x(y,t) = 0 \qquad y \in \partial\,\Omega \tag{3.2}$$

and measurement process described by*

$$z(y,t) = H(y,t)\ x(y,t\ ) + v(y,t) \tag{3.3}$$

where $x(y,t)$: $y \in \Omega \cup \partial\,\Omega$ is the n-dimensional state of the system, $w(y,t)$: $y \in \Omega$ is the independent increment additive white dynamic noise,

---

*The system $\{ L_y, H\}$ is assumed to be observable (see Appendix C).

$v(y,t)$: $y \in \Omega$ is the independent increment additive white measurement noise, $H(y,t)$, $G(y,t)$ are bounded matrix functions over Ω. L_y is a spatial matrix operator of the form

$$L_y[\cdot] = \sum_{i,j=1}^{n} a_{ij}(y) \frac{\partial^2 [\ ]}{\partial y_i \, \partial y_j} + \sum_{i=1}^{n} b_i(y) \frac{\partial [\ ]}{\partial y_i} + C(y)[\ ]; y \in \Omega \qquad (3.4)$$

and N_b is a first order boundary operator of the form

$$N_b[\ ] = \sum_{i=1}^{n} d_i(y) \frac{\partial [\ ]}{\partial n} + f(y)[\ ] \ ; \ y \in \partial \Omega \qquad (3.5)$$

n being the outward normal. The noises are zero mean, white in time with covariance matrices defined by

$$\begin{aligned} E\left\{w(y,t)w^T(y',t')\right\} &= \psi_w(y,y',t)\delta(t-t') \\ E\left\{v(y,t)v^T(y',t')\right\} &= \psi_v(y,y',t)\delta(t-t') \qquad (3.6) \\ E\left\{v(y,t)w^T(y',t')\right\} &= 0 \end{aligned}$$

It is assumed that in the absence of noise, system (3.1) with b. c. given by (3.2) is well-posed in the sense of Hadamard [17], i.e., a regular solution exists, is unique, and evolves continuously with the data.

The estimation problems to be considered here are the same as outlined in Sec. 2.2 of the previous chapter, viz. (i) filtering,

(ii) fixed interval smoothing, (iii) fixed point smoothing, (iv) fixed lag smoothing.

The estimation criterion for any admissible estimate $S_t = \{x(y,t): y \in \Omega,\ t\in[t_o,t_f]\}$ is the minimization of the mean square error functional,

$$E \langle x(y,t_1) - \hat{x}(y,t_1|t_2), \quad x(y,t_1) - \hat{x}(y,t_1|t_2)\rangle \tag{3.7}$$

The best estimate $\hat{x}(y,t_1|t_2)$ of $x(y,t_1)$ which minimizes (3.7) given the initial estimate $\hat{x}(y,t_o|t_o)$ and its covariance matrix

$$V_{\tilde{x}}(y,y_1,t_o) = E\ \{\tilde{x}(y,t_o|t_o)\ \tilde{x}^T(y_1,\ t_o|t_o)\} \tag{3.8}$$

is called the minimum variance estimator [13]. The estimation error is

$$\tilde{x}(y,t_1|t_2) = x(y,t_1) - \hat{x}(y,t_1|t_2) \tag{3.9}$$

Define by $I(y,t)$ the innovation

$$I(y,t) = z(y,t) - H(y,t)\ \hat{x}\ (y,t) \tag{3.10}$$

where $I(y,t)$ has the same covariance as that of the noise $v(y,t)$ [3], i.e.,

$$E\ \{I(y,t)\ I^T\ (y_1,t_1)\} = \psi_v(y,y_1,t)\ \delta\ (t - t_1) \tag{3.11}$$

Using the above property of the innovation process, namely it is white with the same covariance as the measurement noise, we shall derive expressions for the estimator $\hat{x}(y,t_1|t_2)$ in the following section, by assuming that $\hat{x}(y,t_1|t_2)$ is a linear transformation on the innovation process such that the variance of the error is minimized.

3.2 The Innovations Approach

Innovations, as developed by Kailath and Frost [15], [16], applied to linear least squares estimation problems consists basically of whitening the observation sequence and then treating the simpler whitened problem. In effect it is an extension, to linear time varying systems, of of Bode and Shannon's approach to the derivation of the classical stationary Wiener filter [24].

Recently Atre and Lamba [25] appear to have taken an approach similar to the one defined in this section. In this research a composite solution of entire estimation problem is considered, of which the filtering result of Atre and Lamba [25] is a fallout.

The basic problem in the use of the innovations approach is to obtain by means of a causal, invertible linear operation on the observation $z(S,T)$ a white noise process referred to as the innovation $I(S,T)$ (eqn. (3.10)). In the general problem, the determination of a suitable transformation may be a deterrent to the use of the approach. However, in linear distributed systems, as pointed out in the previous section, the innovation process is white [3] for zero mean uncorrelated white dynamic and process noises. Since $z(y,t)$ can be obtained from $I(y,t)$ as

$$z(y,t) = I(y,t) + H(y,t)\,\hat{x}(y,t)\ ; \quad y \in \Omega \tag{3.12}$$

we see that $I(y,t)$ and $z(y,t)$ are related by a causal, invertible linear transformation, and hence $I(y,t)$ is an innovation process for this problem. In other words, any estimate based on $z(y,t)$ can also be written in terms of $I(y,t)$ so that $z(y,t)$ and $I(y,t)$ are "equivalent."

From the orthogonal projection lemma for distributed systems [3], it is known that the estimation error $\tilde{x}(y,t_1|t_2)$ and the observation sequence $z(s,\tau) : s \in \Omega,\ t_o \leq \tau \leq t$ are orthogonal. That is,

$$E \left\{z(s,\tau)\ \tilde{x}^T(y,t_1|t_2)\right\} = 0 \tag{3.13}$$

Because of the equivalence of $z(s,\tau)$ and $I(s,\tau)$, we may then express the estimate as

$$\hat{x}(y,t_1|t_2) = \int_{t_o}^{t_2} \int_{\Omega} \xi(y,t_1\ ;\ s,\tau) I(s,\tau)\ ds\ d\tau \tag{3.14}$$

where $\xi(y,t_1\ ;\ s,\tau)$ is chosen such that the trace of the variance of the error

$$\mathrm{tr}(\mathrm{var}\ (x(y,t) - \hat{x}(y,t)))$$

is a minimum. This requires, from the orthogonal projection lemma [3] (cf. (3.13)),

$$E\ \{\tilde{x}(y,t_1|t_2) I^T(s,\tau)\} = 0\ ;\ y,\ s \in \Omega \tag{3.15}$$

Substituting from (3.14) into (3.15) we get

$$E\{\tilde{x}(y,t_1|t_2)\ I^T(s,\tau)\} = E\{(x(y,t_1) - \hat{x}(y,t_1|t_2))\ I^T(s,\tau)\}$$

$$= E\ \{x(y,t_1)\ I^T(s,\tau)\} - \int_{t_o}^{t_2}\int_{\Omega} \xi(y,t_1\ ;\ s_1,\tau_1)\ E\ \{I(s_1,\tau_1)$$

$$I^T(s,\tau)\ ds_1 d\tau_1 = 0 \tag{3.16}$$

such that using (3.11) we get

$$E\ \{x(y,t_1)\ I^T(y',\tau')\} = \int_{\Omega} \xi(y,t_1\ ;\ s,\tau')\psi_V(s,y',\tau')\ ds \tag{3.17}$$

as the Wiener-Hopf equation governing the system. This is the infinite dimensional analogue of the finite dimensional Wiener-Hopf equation [15]. Using (3.17) we can write our estimate (3.14) as

$$\hat{x}(y,t_1|t_2) = \int_{t_o}^{t_2}\int_{\Omega}\int_{\Omega} E\ \{x(y,t_1)\ I^T(y',\tau)\}\psi_V^{-1}(y',s,\tau)I(s,\tau)dy\ ds\ dt \tag{3.18}$$

This is a principal result of this section, and represents the solution for filtering ($t_2 = t_1$), smoothing ($t_2 > t_1$) as well as prediction ($t_2 < t_1$). The only remaining step is to find an expression for $E\ \{x(y,t_1)\ I^T(y,\tau)\}$ which we do in the next few subsections for the various estimation problems outlined in Sec. 2.2.

3.2A. Solution to the Linear Filtering Problem

The filtering problem was outlined in Sec. 2.2. As previously pointed out this problem has also been recently treated by Atre and Lamba [25], whose results parallel the ones obtained in this subsection.

In terms of the notation of the previous subsection, with $t_1 = t_2 = t$, we can write the filtered estimate

$$\hat{x}(y,t) = \int_{t_0}^{t} \int_{\Omega} \int_{\Omega} E\{x(y,t) I^T(y',\tau)\} \psi_v^{-1}(y',s,\tau) I(s,\tau)\, dy'\, ds\, d\tau \tag{3.19}$$

Recognizing

$$E\{x(y,t) I^T(y',\tau)\} = V_{\tilde{x}}(y,y',t) H^T(y',\tau) \quad \tau \geq t \tag{3.20}$$

we get, on differentiating (3.19) w.r.t. t and substituting from (3.1), the filtered estimate,

$$\frac{\partial \hat{x}(y,t)}{\partial t} = L_y \hat{x}(y,t) + \int_{\Omega} \int_{\Omega} V_{\hat{x}}(y,y',t) H^T(y',t) \psi_v^{-1}(y',y'',t)$$

$$[z(y'',t) - H\hat{x}(y'',t)]\, dy'\, dy'' \tag{3.21a}$$

with boundary condition $\qquad N_b\, \hat{x}(y,t) = 0, \ y \in \partial\Omega \qquad$ (3.21b)

The expression for the error covariance kernel $V_{\tilde{x}}(y,y',t)$ is obtained after differentiating the error equation (3.9) w.r.t. t to get

$$\frac{\partial \tilde{x}(y,t)}{\partial t} = L_y \tilde{x} + B(y,t) w(y,t) - \int_{\Omega} \int_{\Omega} V_{\tilde{x}}(y,y',t) H^T(y',t) \psi_v^{-1}(y', y'',t)$$

$$[H(y'',t) x(y'',t) + v(y'',t)]\, dy'\, dy'' \tag{3.22}$$

with homogeneous boundary condition

$$N_b\tilde{x}(y,t) = 0 \qquad y \in \partial\,\Omega \tag{3.23}$$

and initial condition

$$\hat{x}(y,t_o) = x_o(y) \tag{3.24}$$

where

$$V_{\tilde{x}}(y,y',t) = E\,\{\tilde{x}(y,t)\ \tilde{x}^{\,T}(y',t)\} \tag{3.25}$$

Eqns. (3.22-24) constitute a linear Cauchy problem whose solution can easily be written in terms of the system's Green function [3] as

$$\tilde{x}(y,t) = \int_\Omega \Xi\,(y,t;S,t_o)x_o(S)\ ds + \int_{t_o}^{t}\int_\Omega \Xi\,(y,t;S,\tau)\Big[B(S,\tau)w(S,\tau) - \int_\Omega\int_\Omega V_{\tilde{x}}(S,y',\tau)H^T(y',\tau)\psi_v^{-1}(y',y'',\tau)v(y'',\tau)\ dy'dy''\Big]dS\ d\tau \tag{3.26}$$

where the Green's matrix $\Xi\,(y,t;S,\tau)$ satisfied the linear system

$$\frac{\partial\,\Xi\,(y,t;S,\tau)}{\partial t} = N_y\,\Xi\,(y,t;S,\tau)\ ; \quad y,\ S \in \Omega \tag{3.27a}$$

with

$$\Xi\,(y,t;S,t) = \delta\,(y - S)I \quad \forall t \tag{3.27b}$$

and

$$N_y(\cdot) = L_y(\cdot) - \int_\Omega \int_\Omega V_{\tilde{x}}(y,y',t)H^T(y',t)\psi_v^{-1}(y',y'',t)(\cdot)\ dy'dy'' \qquad (3.27c)$$

Substituting (3.26) into (3.25) and differentiating yields for the error covariance

$$\frac{\partial V_{\tilde{x}}(y,y_1,t)}{\partial t} = L_y V_{\tilde{x}}(y,y_1,t) + V_{\tilde{x}}(y,y_1,t)L_{y_1}^T + G(y,t)\psi_w(y,y_1,t)G^T(y_1,t)$$
$$-\int_\Omega \int_\Omega V_{\tilde{x}}(y_1,y_2,t)H^T(y_2,t)\psi_v^{-1}(y_2,y_3,t)V_{\tilde{x}}(y_3,y_1,t)\ dy_2dy_3 \qquad (3.28)$$

with boundary condition

$$N_b V_{\tilde{x}}(y,\ y_1,t) = 0 \qquad y\varepsilon\partial\Omega, y_1\varepsilon\bar{\Omega} \qquad (3.29)$$

Eqns. (3.21) and (3.28-29) constitute the minimum variance estimator and its associated error covariance for the filtering problem.

The results obtained here are identical to the filtering algorithm for white Gaussian disturbances obtained by Tzafestas [10] using a Bayesian approach and the complicated limit argument of Meditch [12]. The model assumed here does not restrict the noises to be Gaussian, and the derivation is much simplier too. If the noises were assumed white in time and space the filter reduces to the one obtained by Tzafestas and Nightingale [3], Thau [4] using orthogonal projections. The technqiues used here allow the derivation of the filter in a much simpler manner than in the previous references. Also, with the identification

$$Q \leftrightarrow \psi_v^{-1}$$

$$R \leftrightarrow \psi_w^{-1}$$

and assuming the model of Chapter II is purely distributed and linear, this filter is identical to the invariant embedding filter of the last chapter and to the one derived by Lamont [8].

Table 5. The Filtering Algorithm

Estimate	Equation	Initial Condition	Boundary Condition
$\hat{x}(y,t)$	(3.21a)	(3.24)	(3.21b)
Variance $V_{\tilde{x}}$	(3.28)	(3.8)	(3.29a)

3.2B. The Fixed Interval Smoothing Problem

In this subsection and the others following we consider the problem of obtaining smoothed estimates. The importance of smoothing has already been pointed before and need not be emphasized again. The results obtained in these subsections are extensions beyond Atre and Lamba's [25] work on the filtering problem, and emcompass their work as a particular case of our general solution.

The smoothing problem was defined in Sec. 2.2. As for the filtered estimate, the starting point for our estimate is eqn. (3.18), viz.

$$\hat{x}(y,t_1|t_2) = \hat{x}(y,t_1) + \int_{t_1}^{t_2}\int_{\Omega}\int_{\Omega} E\{x(y,t_1)I^T(y',\tau)\}\psi_v^{-1}(y',y'',\tau)I(y'',\tau)dy'dy''d\tau \tag{3.30}$$

From (3.20) we have

$$E\{x(y,t_1)I^T(y',\tau)\} = V_{\tilde{x}}(y,t_1\ ;\ y',\tau)H^T(y',\tau) \tag{3.31}$$

where $V_{\tilde{x}}(y,t_1\ ;\ y',\tau)$ is the smoothing error covariance described by

$$V_{\tilde{x}}(y,t\ ;\ s,\tau) = \mathrm{Cov}\left\{\tilde{x}(y,t),\tilde{x}(s,\tau)\right\} \tag{3.32}$$

Substituting (3.30) into (3.32) and recognizing that $\tilde{x}(y,t_1)$ and $v(y,t_1)$ are uncorrelated for $t_1 > t$, we get, using (3.26-27)

$$V_{\tilde{x}}(y,t\ ;s,\tau) = \int_\Omega V_{\tilde{x}}(y,r,t)\,\Xi^T(s,\tau;\ r,t)dr \tag{3.33}$$

Substitute (3.33) and (3.31) into (3.30) to get for the smoothed estimate

$$\hat{x}(y,t_1|t_2) = \hat{x}(y,t_1) + \int_\Omega V_{\tilde{x}}(y,s,t_1)\lambda(s,t_1,t_2)\ ds \tag{3.34}$$

where

$$\lambda(s,t_1,t_2) = \int_{t_1}^{t_2}\int_\Omega\int_\Omega \Xi^T(y,\tau;s,t_1)H^T(y,\tau)\psi_v^{-1}(y,r,\tau)I(r,\tau)dy\,dr\,d\tau \tag{3.35}$$

Differentiating (3.34) with respect to t_1 and using Green's identity yields for the fixed interval smoothing estimate

$$\frac{\partial\hat{x}(y,t_1|t_2)}{\partial t_1} = L_y\hat{x}(y,t_1|t_2) + \int_\Omega\int_\Omega G(y,t_1)\psi_w(y,y_1,t)G^T(y_1,t_1)$$

$$V_{\tilde{x}}^{-1}(y_1,y_2,t_1)(\hat{x}(y_2,t_1|t_2) - \hat{x}(y_2,t_1))dy_1dy_2 \tag{3.36}$$

with boundary conditions

$$N_b\hat{x}(y,t_1|t_2) = 0 \qquad y \in \partial\Omega \tag{3.37}$$

and terminal condition

$$\hat{x}(y,t_1|t_2) = \hat{x}_o(y,t_2), \quad \text{the filtered estimate} \tag{3.38}$$

An examination of the estimator equation (3.36) shows that the estimate depends only on the filtering error covariance, being independent of the smoothing error covariance. It would, however, be desirable to have a quantitative measure of the effect of smoothing on the estimate error variance. To this effect, substituting from (3.34) into the smoothing error expression (3.9) yields

$$\tilde{x}(y,t_1|t_2) = \tilde{x}(y,t_1) - \int_\Omega V_{\tilde{x}}(y,y',t_1)\lambda(y',t_1,t_2)\, dy' \tag{3.39}$$

The smoothing error variance is therefore

$$V_{\tilde{x}}(y,y',t_1|t_2) = E\left\{x(y,t_1|t_2)x^T(y',t_1|t_2)\right\} \tag{3.40}$$

which on substituting from (3.53) yields

$$V_{\tilde{x}}(y,y',t_1|t_2) = V_{\tilde{x}}(y,y',t_1) - \int_{t_1}^{t_2}\int_\Omega\int_\Omega\int_\Omega\int V_{\tilde{x}}(y,y_2,t_1)\,\Xi^T(y_3,\tau;y_2,t_1)$$

$$H^T(y_3,\tau)\psi_v^{-1}(y_3,y_4,\tau)H(y_4,\tau)\,\Xi(y_5,t_1;y_4,\tau)$$

$$V_{\tilde{x}}(y_5,y',t_1)\; dy_2\,dy_3\,dy_4\,dy_5\,d\tau \tag{3.40a}$$

This is a very appealing result, since it indicates exactly the amount by which the smoothing of future data reduces the variance of the error in the filtering estimate of $x(y,t_1)$. In this form, however, the answer is computationally not very tractable since the system's Green's function must be evaluated. To have a more tractable solution, we differentiate with respect to t_1 and substitute from eqn. (3.27) to get

$$\frac{\partial V_{\tilde{x}}(y,y',t_1|t_2)}{\partial t_1} = (L_y + \Pi_y)V_{\tilde{x}}(y,y',t_1|t_2) + V_{\tilde{x}}(L_{y'} + \Pi_{y'})^T$$

$$- G(y,t)\psi_w(y,y',t)G^T(y',t) \tag{3.41}$$

with boundary conditions

$$N_b V_{\tilde{x}}(y,y',t_1|t_2) = 0 \qquad y\varepsilon\partial\Omega, y'\varepsilon\bar{\Omega} \tag{3.42}$$

and terminal condition

$$V_{\tilde{x}}(y,y',t_2|t_2) = V_{\tilde{x}}(y,y',t_2) \tag{3.43}$$

where

$$\Pi_y(\cdot) = \int_\Omega\int_\Omega G(y,t)\psi_w(y,y_1,t)G^T(y_1, t)V_{\tilde{x}}^{-1}(y_1,y_2,t)(\cdot)dy_1dy_2 \tag{3.44}$$

To summarize, we have obtained the fixed interval smoothing algorithm for linear DP systems (with t_2 fixed) using innovations. The algorithm consists of processing the filtering algorithm of Sec. 3.2A forward in time upto t_2 and then processing the smoothing estimator, eqns. (3.36-37), backward in time from t_2 to time t_o. The terminal condition is determined by the filtering estimate at time t_2. The smoothing error covariance has been shown to be an improvement over the filtering error covariance. As before, a backward sweep must be made in order to compute the smoothing error covariance.

The algorithm in this subsection is identical to the one obtained by Tzafestas [10] for linear DP systems with white Gaussian disturbances using a rather cumbersome limit technique. The ease of derivation of our algorithm is apparent from the simple mathematics involved, and moreover, no restriction on noise statistics are imposed. In this sense, it is more general than the one in ref. [10]. Moreover, this algorithm involves fewer equations than the 'two filter' algorithms proposed by Tzafestas and Nightingale [3], [9], which saves on storage and computer execution time.

The algorithm can be tabulated as shown in Table 6. (The filtering algorithm of the previous section needs to be processed before the implementation of this algorithm.)

Table 6. Fixed Interval Smoothing Algorithm

	Equation	B.C.	I.C.
Estimate $\hat{x}$	(3.36)	(3.37)	(3.38)
Variance $V_{\tilde{x}}$	(3.41)	(3.42)	(3.43)

3.2C. Fixed Point Smoothing Estimation

In Sec. 3.2B on fixed interval smoothing, we considered smoothing with $t_1 = t$ and t_2 fixed. In fixed point smoothing we have $t_1 = t_*$ fixed and $t_2 = t$, the running variable. With this identification, our general smoothing estimator expression, eqn. (3.30, becomes

$$\hat{x}(y,t_*|t) = \hat{x}(y,t_*) + \int_{t_*}^{t}\int_\Omega\int_\Omega\int_\Omega V_{\tilde{x}}(y,y_1,t_*)\, \Xi^{T}(y_2,\tau;y_1,t_*) H^T(y_2,\tau)\psi_v^{-1}(y_2,y_3,\tau)I(y_3,\tau)dy_1dy_2dy_3d\tau \qquad (3.44)$$

Differentiating this expression with respect to t yields the fixed point smoothing estimator

$$\frac{\partial\hat{x}(y,t_*|t)}{\partial t} = \int_\Omega\int_\Omega V_{\tilde{x}}(y_1,t;y,t_*)H^T(y_1,t)\psi_v^{-1}(y_1,y_2,t)(z - H(y_2,t)\hat{x}(y_2,t))\ dy_1dy_2 \qquad (3.45)$$

using (3.33). The moment algorithm is obtained by differentiating (3.33) to yield

$$\frac{\partial V_{\tilde{x}}(y_1,t;y,t_*)}{\partial t} = V_{\tilde{x}}(y_1,t;y,t_*)\ N^T_{y_1} \qquad (3.46)$$

with initial conditions

$$\hat{x}(y,t_*|t_*) = \hat{x}(y,t_*) \qquad (3.47a)$$

$$V_{\tilde{x}}(y,t_*;y',t_*) = V_{\tilde{x}}(y,y',t_*) \qquad (3.47b)$$

and boundary conditions

$$N_b\ \hat{x}(y,t_*|t) = 0 \qquad y\varepsilon\partial\Omega \qquad (3.48a)$$

$$N_b\ V_{\tilde{x}}(y_1,t\ ;\ y,t_*) = 0 \qquad y\varepsilon\partial\Omega,\ y_1\varepsilon\bar{\Omega} \qquad (3.48b)$$

As yet we have not obtained an expression for the smoothing error covariance. This can be obtained by differentiating (3.4) w.r.t. $t_2 = t$ to yield

$$\frac{\partial V_{\tilde{x}}(y,y',t_*|t)}{\partial t} = -\int_\Omega\int_\Omega V_{\tilde{x}}(y_1,t\ ;\ y,t_*)H^T(y_1,t)\psi_v^{-1}(y_1,y_2,t)$$

$$H(y_2,t)V_{\tilde{x}}(y',t_*\ ;\ y_2,t)\ dy_1dy_2 \qquad (3.49)$$

with initial condition

$$V_{\tilde{x}}(y,y',t_*|t_*) = V_{\tilde{x}}(y,y',t_*) \qquad (3.50)$$

and boundary condition

$$N_bV_{\tilde{x}}(y,y';\ t_*|t) = 0 \quad y\varepsilon\partial\Omega,\ y'\varepsilon\bar{\Omega} \qquad (3.51)$$

To summarize, the fixed point smoothing estimate is obtained by processing the filtering algorithms upto t_* and the processing the smoothing algorithm eqns. (3.45-48). The smoothed estimate is independent of the smoothing error covariance.

The results in this section are extensions of the previous work of Tzafestas, who obtained a somewhat identical algorithm for Gaussian disturbances [10] using very complicated limiting procedures. The derivation in this section is rather simple and extends to a much broader class of disturbances than Gaussian noise. These algorithms are identical to those derived in Chapter II for linear purely distributed systems if the correspondence noted at the end of Sec. 3.2A is observed.

Table 7. Fixed Point Smoothing Algorithm

	Equation	B.C.	I.C.
Estimate	(3.45)	(3.48a)	(3.47a)
Moment	(3.46)	(3.48b)	(3.47b)
Covariance	(3.49)	(3.51)	(3.50)

The filtering algorithm of Sec. 3.2A needs to be implemented with this algorithm.

3.2D. The Fixed Lag Smoothing Problem

The fixed lag smoothing problem was outlined in Sec. 2.2. The analysis over here is an extension of the general smoothing analysis of Sec. 3.2A, with $t_1 = t$ and $t_2 = t + T$, T being the fixed delay.

Our starting point, as before, is eqn. (3.30), the smoothed estimate, and eqn. (3.31). Differentiate (3.31) with respect to t $(= t_1)$ to get

$$\frac{\partial\lambda(y',t\,|\,t+T)}{\partial t} = -\ [N_y{}^{*}]\lambda - \int_\Omega H^{T}(y',t)\psi_v^{-1}(y',y'',t)I(y'',t)\ dy''$$

$$+ \int_\Omega\int_\Omega \Xi^{T}(y'',t+T;y',t)H^{T}(y'',t+T)\psi_v^{-1}(y'',y,t+T)$$

$$I(y,t+T)\ dy'dy'' \tag{3.52}$$

Differentiate (3.30) w.r.t. $t(=t_1)$ and substitute from (3.52), (3.22), (3.28) to get the fixed lag smoothed estimate.

$$\frac{\partial\hat{x}(y,t\,|\,t+T)}{\partial t} = L_y x(y,t\,|\,t+T) + \int_\Omega\int_\Omega G(y,t)\psi_w(y,y_1,t)G^{T}(y_1,t)$$

$$V_{\tilde{x}}^{-1}(y_1,y_2,t)(\hat{x}(y_2,t\,|\,t+T) - \hat{x}(y_2,t))\ dy_1dy_2$$

$$+ \int_\Omega\int_\Omega\int_\Omega V_{\tilde{x}}(y,y',t)\,\Xi^{T}(y_1,t+T;y',t)H^{T}(y_1,t+T)$$

$$\psi_v^{-1}(y_1,y_2,t+T)I(y_2,t+T)\ dy'dy_1dy_2 \tag{3.53}$$

with initial conditions obtained from processing the fixed point algorithm from t_o to $t_o + T$. The error kernel matrix is obtained from (3.27) and is

$$\frac{\partial\,\Xi\,(y,\ t+T\ ;\ y',t)}{\partial t} = N_y\ \Xi\ (y,t+T\ ;\ y',t) - \Xi\ (y,t+T;y',t)N_y^{T} \tag{3.54}$$

the boundary conditions are

$$N_b\hat{x}(y,t\,|\,t+T) = 0 \qquad y\varepsilon\partial\Omega \tag{3.55a}$$

$$N_b \, \Xi \, (y, t + T \; ; \; y', t) = 0 \qquad (3.55b)$$

An examination of the estimator equation (3.53) reveals that it is independent of the error covariance. An expression for the smoothing error can be obtained, as in the other smoothing cases, by differentiating (3.9) with respect to $t(= t_1)$ to get

$$\frac{\partial V_{\tilde{x}}(y, y', t | t + T)}{\partial t} = (L_y + \Pi_y) V_{\tilde{x}}(y, y', t | t + T) + V_{\tilde{x}}(y, y', t | t + T)$$

$$(L_{y'} + \Pi_{y'})^T - \int_\Omega \int_\Omega \int_\Omega V_{\tilde{x}}(y, y_1, t) \, \Xi^T(y_2, t + T; \; y_1, t) H^T(y_2, t + T)$$

$$\psi_v^{-1}(y_2, y_3, t + T) H(y_3, t + T) \, \Xi (y_3, t + T; y', t) dy_1 \, dy_2 dy_3 - G \psi_w G^T \qquad (3.56)$$

with initial conditions obtained from the fixed point smoothing algorithm and boundary conditions

$$N_b \, V_{\tilde{x}}(y, y', t | t + T) = 0 \quad y \varepsilon \partial\Omega, \; y' \varepsilon \bar{\Omega} \qquad (3.57)$$

To summarize, the fixed lag algorithm proceeds by processing the filtering algorithm along with eqns. (3.53-57). The initial conditions are obtained by processing the fixed point smoothing algorithm upto the time interval $t_o + T$.

These algorithms apply to a more general class of noise statistics than those considered by Tzafestas [10] which applies only to Gaussian distributions. The derivation is far more simpler, too. For the purely distributed case, these algorithms are identical to the ones derived in

Chapter II using invariant embedding.

Table 8. Fixed Lag Smoothing Algorithm

	Equation	B.C.	I.C.
Estimate	(3.53)	(3.55a)	From fixed point smoothing algorithm
Moment	(3.54)	(3.55b)	"
Variance	(3.56)	(3.57)	"

The fixed point smoothing algorithm must be implemented alongside the above tabulated algorithm, along with the filtering algorithm.

3.3. Summary

In this chapter, we have developed general filtering and smoothing algorithms for second order linear DP systems with independent increment volume disturbances in the dynamics and measurement process. The estimation criterion was considered to be the minimization of the least squares estimation error which results in a minimum variance estimator. The approach taken is the innovations approach. The innovation for the class of systems considered here is white, which facilitates in the derivation.

The procedure consists of expressing the minimum variance estimator as a linear transformation on the innovations process. This results in an infinite dimensional Wiener-Hopf equation, which represents the solution for all the estimation problems, viz. filtering, smoothing and prediction. Solution of the Wiener-Hopf equation is facilitated by the fact that the innovations process is white, as previously noted.

The error covariance is obtained by differentiating the error to yield a linear TPBVP whose solution is obtained using a Green's function approach. The resulting expression clearly indicate the superiority of smoothing vis-á-vis filtering in that the error covariance is reduced.

The implementation of the various estimation algorithms has already been tabulated in Tables 5 and 8. The smoothing algorithms either require a parallel implementation of the filtering algorithm(fixed point and fixed lag smoothing) or serial implementation after the filtering algorithm (fixed interval smoothing). Furthermore, as pointed out in Section 3.2A, these algorithms are identical to the purely linear distributed least squares estimation algorithms of Chapter II if the following correspondence between the noise covariances of this chapter and the weighting matrices of Chapter II are observed.

Least Squares		Innovations
Q	$\longleftrightarrow$	ψ_v^{-1}
R	$\longleftrightarrow$	ψ_w^{-1}

CHAPTER IV

NUMERICAL IMPLEMENTATION OF REPRESENTATIVE EXAMPLES

In the previous chapters we concentrated upon developing estimation algorithms for a class of linear and nonlinear DP systems using various techniques. No mention was made of the computational feasibility of the developed algorithms. In this chapter we shall examine this aspect of the previously developed algorithms, and examine various methods of solving the PDE's on a digital computer, since the object of any algorithm is to execute it on a digital machine.

Three considerations are important in any numerical algortihm being executed on a digital computer, viz. accuracy, speed of execution and storage requirements. In our estimation scheme a measure of the accuracy, ignoring computer round-off and truncation, is the error covariance speed of execution and computer memory requirements are functions of the finite difference approach used to solve the estimator integro-partial differential equation. We shall examine these requirements in the following sections for different finite difference methods.

While there are numerous schemes for solving partial differential equations, for instance finite difference techniques, collocation schemes and spline methods, we shall concentrate only on finite difference schemes in this dissertation. The purpose of this chapter is not to exhaustively review solution techniques, but only to explore one of the alternate schemes, specifically the finite difference method, for solving two point boundary value problems. The application of various methods for solving distributed estimation problems has received little attention, and except for the work of Seinfeld et al. [14] using a discrete space continuous time approach, no other study appears to have been made.

It should also be pointed out, as done before in Chapter III, that the approaches of Chapters II and II result in essentially identical algorithms. With this in mind, we shall treat examples using the algorithms of Chapter II alone in our numerical experimentation.

4.1 Numerical Experiments

In this section we consider the application of the algorithms of Chapter II (and concomitantly, as mentioned previously, Chapter III) to representative engineering problems. Two schemes of discretization shall be studied, viz.:

1. Discrete Space-Discrete Time (DSDT) in which the linear and nonlinear integro-partial differential equations are discretized in both their spatial and temporal co-ordinates.

2. Discrete Space Continuous Time (DSCT) in which the system model partial differential equations are discretized in their spatial co-ordinate only. This results in estimator equations as derived by Seinfeld et al. [14]. The resulting system of ordinary differential equations are then solved using any of the numerous techniques available for solving such equations.

The first problem considered is the estimation of mass flow x(y,t) in a chemical reactor. The chemical reactor is a cylinderical tube packed with catalyst. The reactants flow through the packed bed, react on the surface of the catalyst, and the products continue through the bed. Dispersion in the radial and axial directions occurs because of the flow around the particles. We consider the axial dispersion problem here alone, and compare the results of DSDT simulation of the equations of Chapter II with the DSCT estimator algorithms of Seinfeld, et al. [14].

Consider an isothermal situation with a second order reactor [27]. The mass balance is, in dimensionless form

$$\frac{\partial x}{\partial t} = \frac{\partial x^2}{\partial y^2} - \mathrm{Pe}\,\frac{\partial x}{\partial y} - R(\mathrm{Pe})x^2 + w(y,t) \tag{4.1}$$

where Pe is the axial Peclet number VL/D, R is the reaction rate group, V the axial velocity, L the length of the reactor and D the axial dispersion coefficient. $w(y,t)$ is the additive noise of unknown statistics. The boundary condtions are

$$\frac{\partial x}{\partial y} = 0 \qquad \text{at} \qquad y = 1, \quad \text{all } t \tag{4.2a}$$

$$\frac{\partial x}{\partial y} + \mathrm{Pe}(1 - x) = 0 \qquad y = 0, \quad \text{all } t \tag{4.2b}$$

with initial condition

$$x(y,0) = 0 \qquad \text{all } y \tag{4.3}$$

Measurements of the form

$$z(y,t) = x(y,t) + v(y,t) : \quad y \in (0,1) \tag{4.4}$$

are assumed.

The coefficients Pe and R were chosen as $Pe = 10$, $R = 1/2$. The filtering algorithm (Appendix A) was solved along with the discrete estimation algorithm of Seinfeld, et al. [14]. The basis for comparison that was made is shown in Table 9 and the actual equations are given in Appendix B.

Table 9. Filtering Algorithm

	This Research	Seinfeld, et al. [14]
(a) No. of Space Intervals	8	8
(b) Time Step (Δt)	.01 sec.	.01 sec.
(c) Method Used	DSDT approach using Crank-Nicholson method to solve estimator equation and alternating direction implicit technique to solve the error covariance equations.	Adams-Moulton predictor-corrector approach to solve both the estimator and error covariance ordinary differential equations, using a step size $\Delta t = .01$ sec.
(d) Noise Statistics	Independent, zero mean with covariances $\delta(t - t')/\Delta t$ (because of time discretization) and uniform distribution.	Independent, zero mean, with covariances $\delta(t)$ and a uniform distribution.
(e) Initial Condition		
(i) Estimator	$\hat{x}(y,0) = .00$	$\hat{x}_i(0) = .00,\ i = 1, ..8$
(ii) Error Covariance	$P(y,y_1,0) = 2.50$	$P_i(t) = 2.50$

Results of the simulation using the two approaches are shown in Figures 1 and 2, which indicate the superiority of the DSDT (this research) approach to the DSCT approach in that error covariance and estimator error is reduced considerably. Besides the reduction in error covariance, a considerable amount of computer execution time is saved. Computer execution time alone is considered as a factor as it precludes idiocyncracies of programming taken into account by the compilation time. For this problem we have

Table 10. Filtering Computational Requirements.

	Current Research	Seinfeld
Dynamic Storage	~2500 words	3300 words
Execution Time	25 sec.	170 sec.

Results for fixed interval and fixed point smoothing at $t_1 = 0.5$ sec. are also shown in Figures 1 and 2. The error covariance Figure 2 clearly points out the superiority of smoothing over filtering. As before the smoothing algorithms were solved using 8 space intervals and the DSDT approach listed in Table 9. The procedure is recapitulated in Table 11.

Results for storage requirements and computer execution time for both smoothing and filtering are given in Table 12. Smoothing, as is to be expected, is slower than filtering.

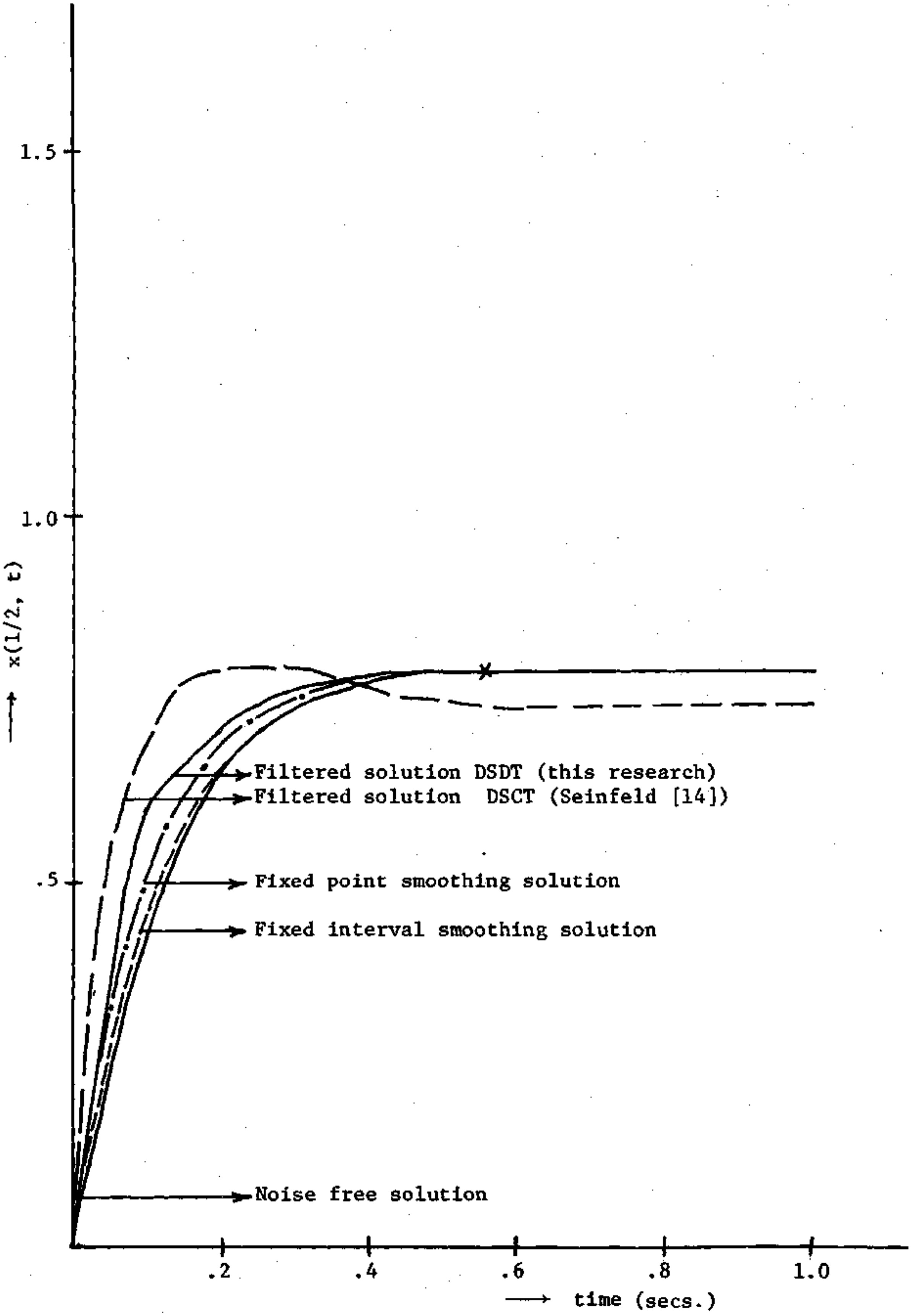

Figure 1. Chemical Reactor Kinetics-State Estimate.

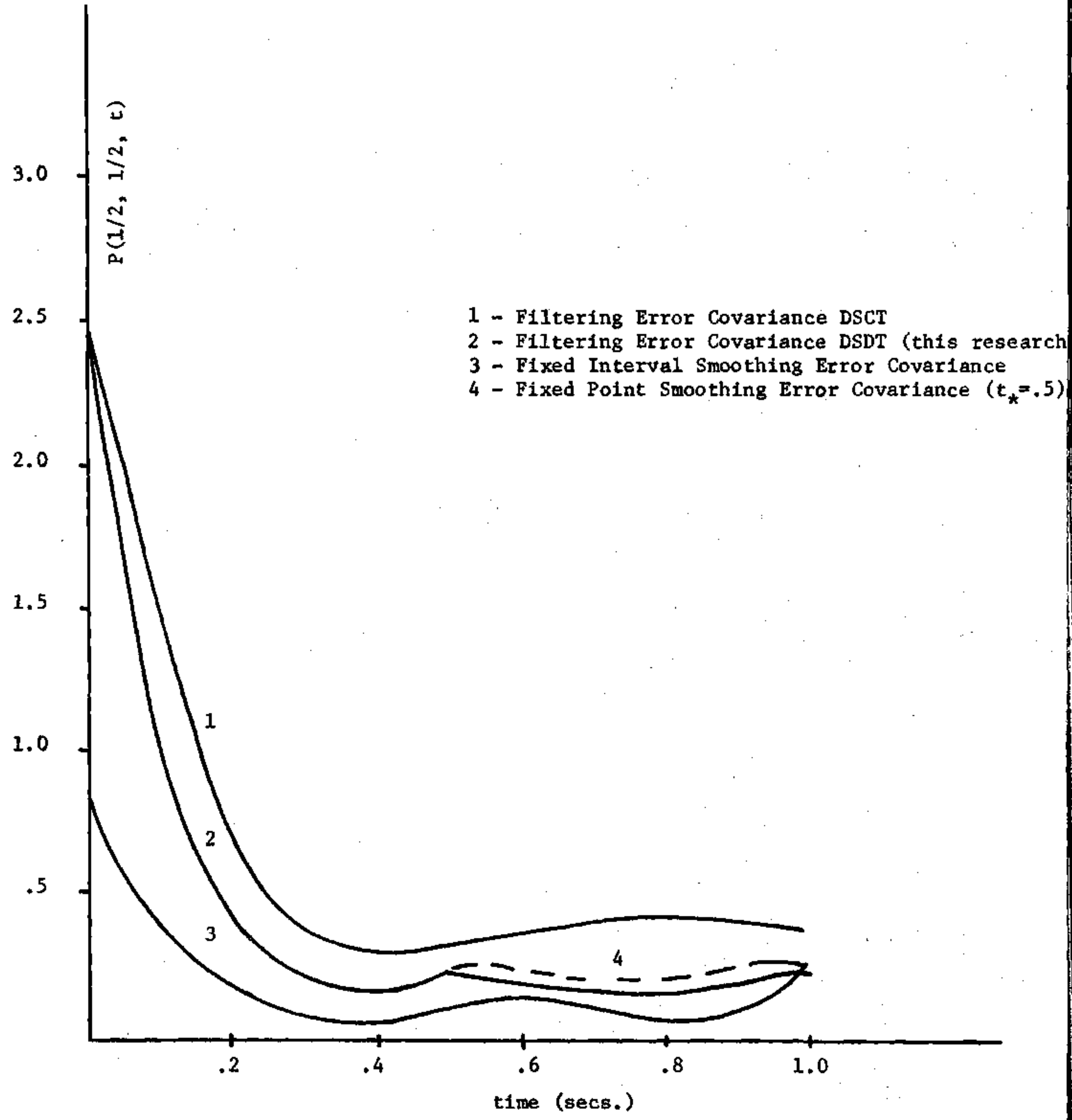

Figure 2. Chemical Reactor Kinetics - Error Covariances

Table 11. Smoothing Algorithm

	Smoothing	Filtering
(a) No. of Space Intervals	8	8
(b) Time Step Size	.01 sec.	.01 sec.
(c) Method Used	DSDT as outlined in Table 9.	DSDT as outlined in Table 9.
(d) Error Covariances	As in Table 9 for the DSDT approach.	As in Table 9.
(e) Initial Conditions		
(i) Error Covariance	$P(y,y_1,0) = 2.50$	For fixed interval smoothing: $P_1(y,y_1|1) = P(y,y_1,1)$ the filtering error covariance. For Fixed Point Smooth. $P_2(y,y_1,.5/1) = P_1(y,y_1,.5/1)$, the filtering error covariance
(ii) Estimator	$\hat{x}(y,0) = .00$	Fixed Interval Smoothing: $\hat{x}(y,1/1) = \hat{x}(y,1)$, the filtered estimate. Fixed Point Smoothing: $\hat{x}_2(y,.5/1) = \hat{x}_1(y,.5/1)$, the filtered estimate.

Table 12. Smoothing Computational Requirements.

	Filtering DSDT	Smoothing Fixed Pt.	Fixed Int.	Filtering DSCT(Seinfeld)
Storage	~2500 words	2700	3100	3300
Execution Time	25 sec.	32	45	170

In all of the above experimentation the weighting coefficients for the dynamic and measurement residuals in the performance index were both chosen to be constants equal to one. These coefficients being kept constant, next the estimation algorithm was implemented for various values of dynamic and measurement noises. A plot of the fixed interval smoothing error covariance for fixed value of the dynamic noise covariance (= $\delta(t - t')$ is shown in Figure 3. It seen that as the measurement noise covariance is increased, because it is being weighted less and less in the performance index, the error covariance increases. For very small values of the measurement noise covariance, it is relatively heavily weighted, and the dynamic noise predominates. However, the magnitude of the dynamic noise is kept constant, as is its weightage, so that for measurement noise covariance of upto $1.\delta(t - t')$, the error covariance is almost constant. This graphically demonstrates the need of selection of proper weightage to keep the error covariance within bounds.

Example 2. The next example is taken from the field of nuclear reactor kinetics [28]. Consider the equation

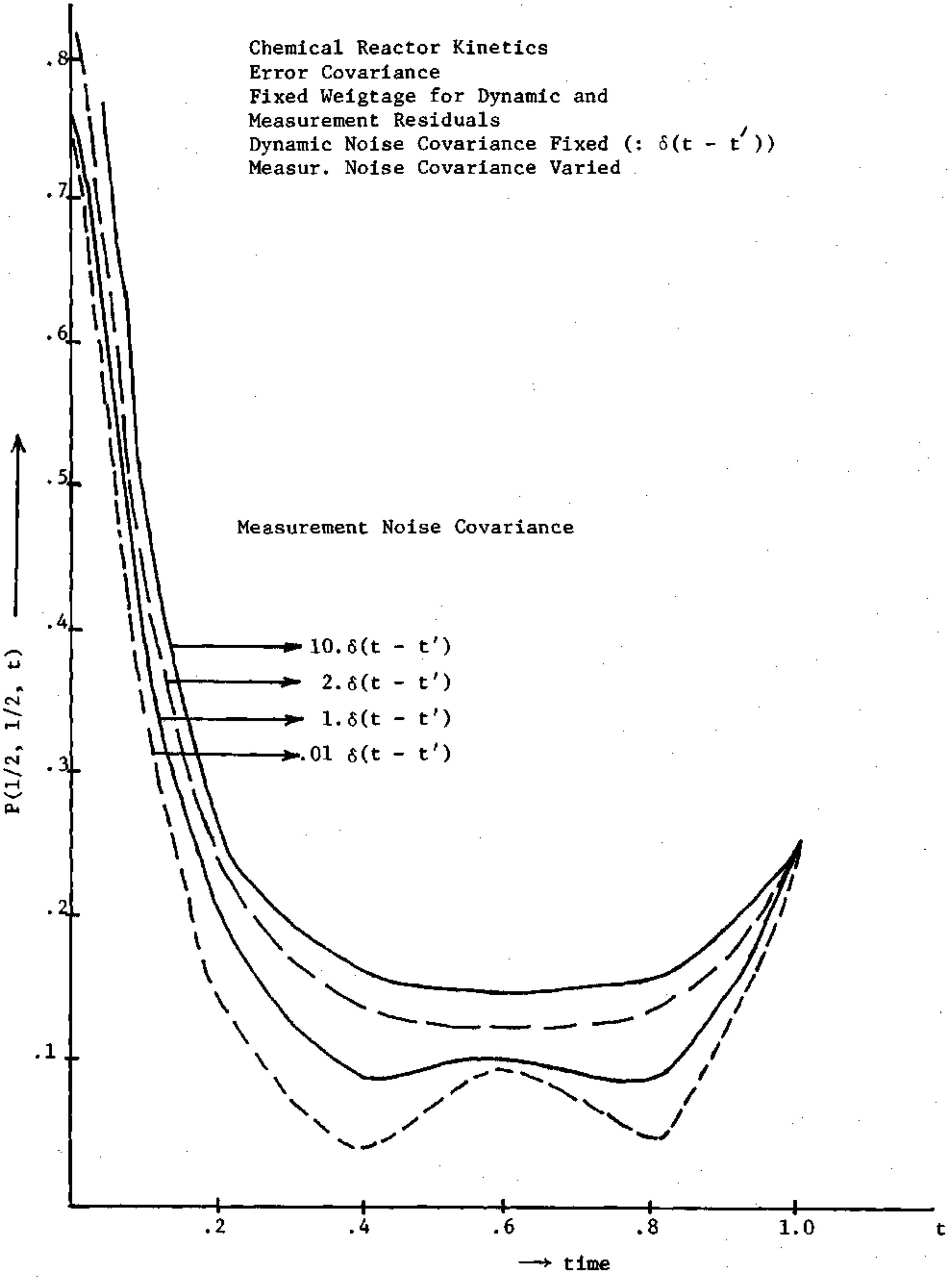

Figure 3. Chemical Reactor Kinetics–Error Covariances for Varied Measur. Noise Covariances.

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial y^2} + P^2(y)x - \delta y^2 + w(y,t) \qquad y \in (0,\Pi) \tag{4.5}$$

which represents the dynamics of a space dependent reactor with no delayed-neutron effects and a general prompt feedback [28]. The initial and boundary conditions are

$$x(y,0) = f(y) > 0, \qquad x(0,t) = x(\Pi,t) = 0 \tag{4.6}$$

The space dependent buckling $P^2(x)$ gives the initial multiplication of the neutron distribution; P^2 is positive and piecewise continuous, and in all cases it is sufficiently large so that at time zero the neutron distribution is increasing with time and not in a steady state. δ is the constant negative feedback coefficient $(\delta > 0)$. $w(y,t)$ is the additive noise of unknown statistics. Interior measurements of the form

$$z(y,t) = x(y,t) + v(y,t) \tag{4.7}$$

were assumed, where v is additive noise of unknown statistics.

As in the previous example, both the discrete (Seinfeld [14]) and continuous filtering and smoothing algorithms were applied. The basis of comparison was exactly the same as in Tables 9 and 10, except that instead of white independent, uniform noises, Gaussian white independent noises with zero mean and unity covariances were used. The numerical solution was obtained for a homogeneous reactor with

$$P^2 = 10, \quad \delta = 1$$

$$f(y) = .1 \sin y$$

The DSDT filtering, fixed interval smoothing and fixed point smoothing ($t_1 = .5$ sec.) were obtained for $t \in [0,2.]$ secs. The DSCT equations, with the same number of interior points as before (i.e., 8) were also simulated for the filtering algorithm using identical probability density functions as before. Results of this simulation are displayed in Figures 4 and 5. Besides the reduction in error covariance, a considerable amount of storage and execution time reduction is achieved. For this problem

Table 13. Computational Requirements.

	Filtering (This Research)	Fixed Int. Smooth.	Fixed Pt. Smooth.	Filtering (Seinfeld)
Storage (words)	~2500	3200	2700	3200
Execution Time (secs.)	26	50	35	175

4.2. Discussion and Comparison of Results

As reported before, the two examples considered in this section were simulated on a UNIVAC 1108 computer. The costs and other important performance factors resulting from the simulation are listed in Tables 9-13 and the results are elaborated upon below.

For purposes of comparison, each example in evaluated using the same initial conditions for the estimate and error covariance for both the DSDT (this research) and DSCT approaches. Furthermore, to ensure proper comparison, the space and time intervals were kept the same.

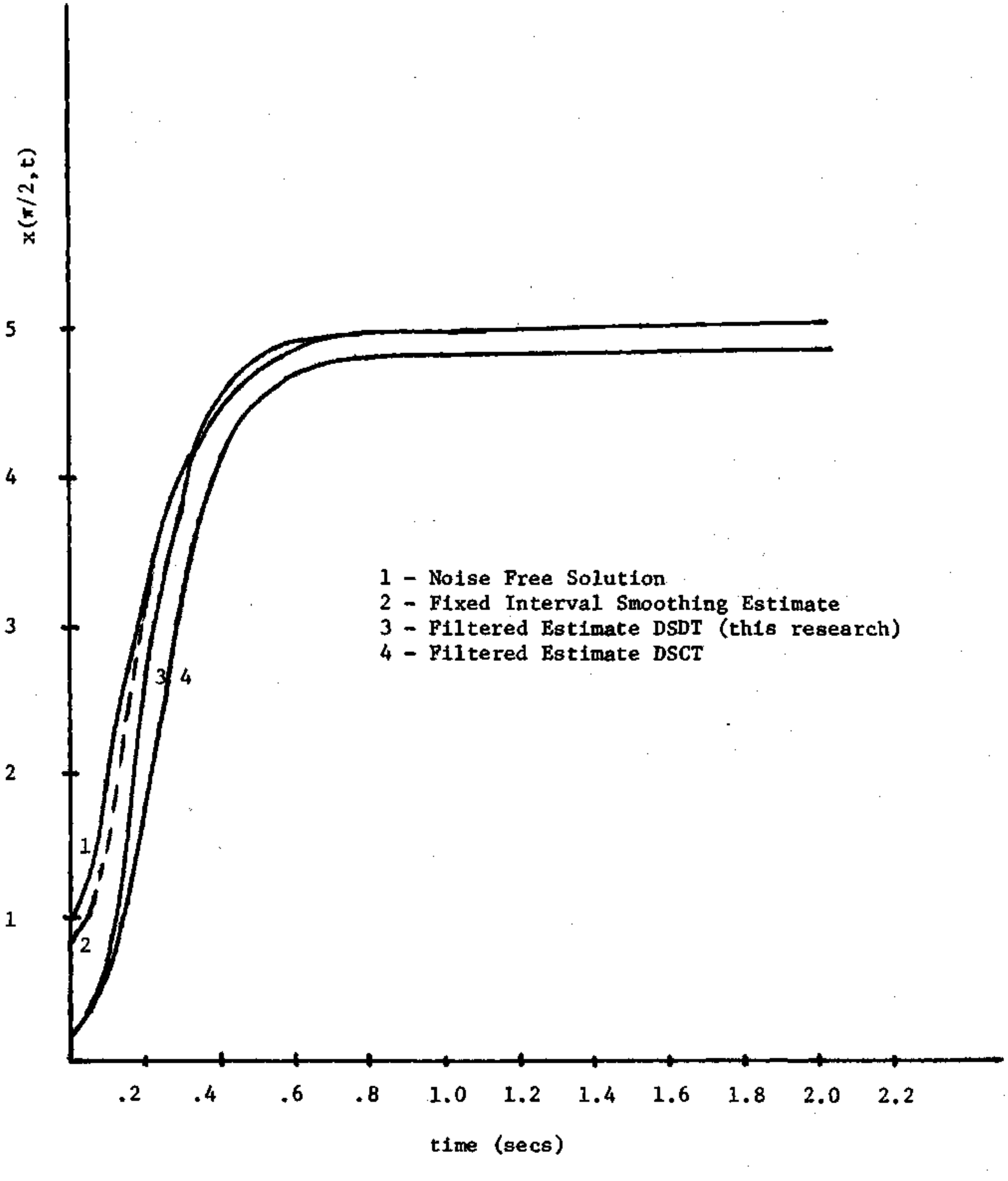

Figure 4. State Estimate for Nuclear Reactor Kinetics

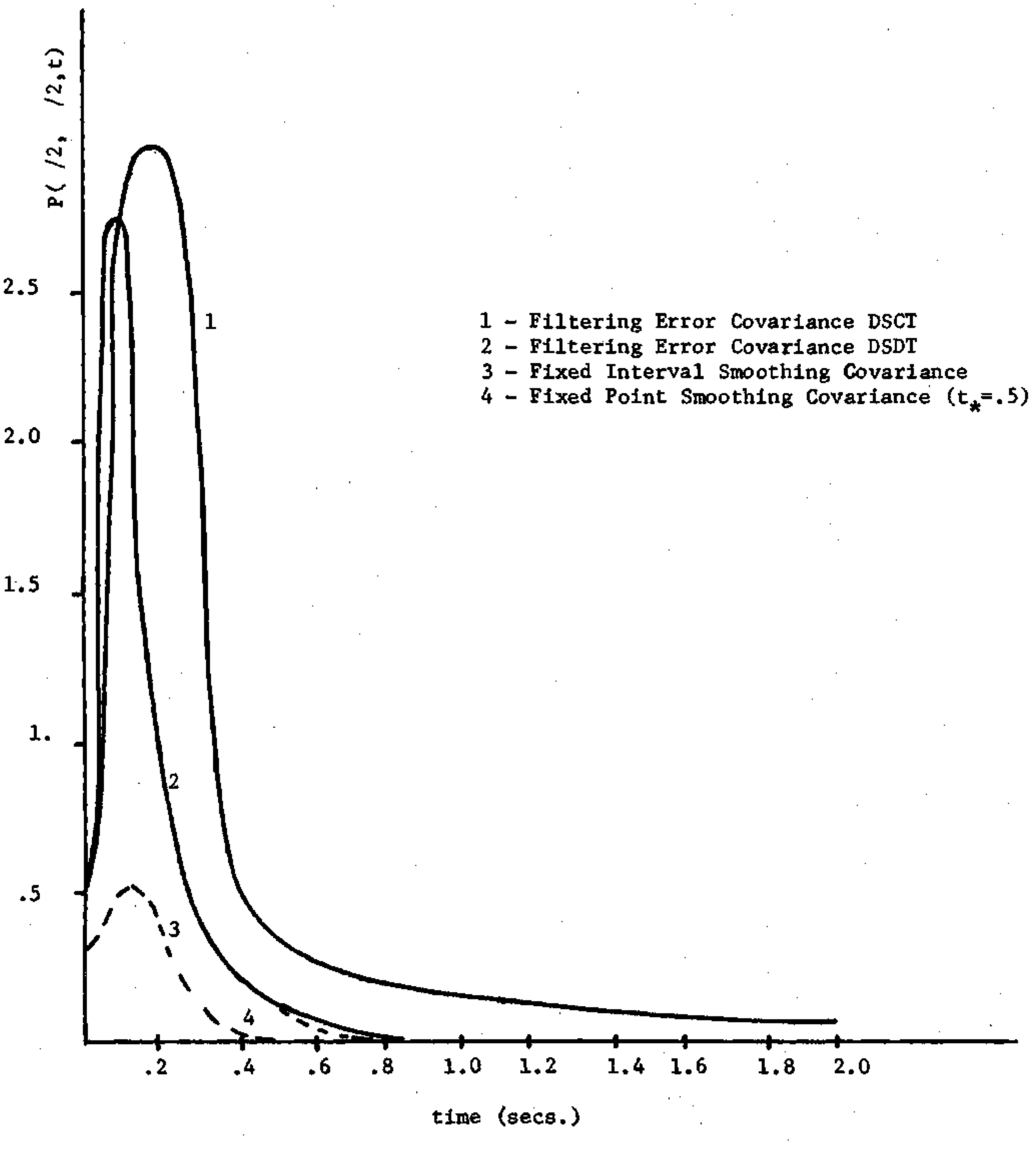

Figure 5. Nuclear Reactor Kinetics - Estimator Error Covariances

All these factors are tabulated in Tables 9 and 10. To account for the time discretization in the DSDT approach, the introduced noise covariances were adjusted by a factor $1/\Delta t$, Δt being the time step size. All these factors ensured that a rationale for a proper comparison between the DSDT and DSCT approaches existed.

In the DSDT approach, the well known Crank-Nicholson method for parabolic equations [29] was used for simulating both the noisy data (eqns. (4.1) and (4.5) and the estimator equations. The error covariance equations, being in two space dimensions, was solved using the implicit alternating direction method [29].

For the DSCT approach, the discretized noisy equation along with the estimator and error variance equations were solved using the Adams-Moulton predictor-corrector method with the starting values supplied by a fourth order Runge-Kutta routine. The time steps for the differential equation solving routine were kept the same as for the DSDT approach.

Results on the performance of the two approaches are given in Tables 9-13 along with Figures 1-5. As is apparent from Tables 10, 12 and 13, the execution time for the DSDT approach is considerably (~6/7 times) less than the DSCT approach of Seinfeld [14]. This is not surprising, keeping in view the fact that in general digital computers are slow in solving differential equations vis á vis algebraic equations. This should not however detract from the efficacy of the Seinfeld algorithm. Implemented on an analogue system, a striking improvement in results would be achieved for this algorithm. Results comparable to our filtering algorithm shall be obtained on a hybrid system with the Seinfeld algorithm.

So far we have demonstrated the efficacy of the DSDT approach over the DSCT approach. This is not a property of our estimation algorithm, rather a property of the numerical algorithm used to implement the estimation algorithm. It has been shown that the DSDT approach can play a significant part in reducing computational time, as well as lessening storage requirements, for distributed estimation problems.

The smoothing algorithms of Chapter II were implemented using the DSDT approach alone. The same space-time intervals were observed for smoothing as were done for filtering to facilitate comparison. Results of the smoothed trajectories are shown in Figures 1 and 4 while the smoothing error covariance is displayed in Figures 2 and 5. While these results are qualitative in nature alone, we can deduce from them that smoothing does indeed improve the system performance, in that error covariance for smoothing is less than for the filtering algorithm (Figures 2 and 5), and also that the smoothed trajectory follows the noise free estimate much more closely than the filtered trajectory (Figures 1 and 4). Among the different kinds of smoothing, fixed interval smoothing gave the best results using the above two criteria. However, as is indicated by Tables 12 and 13, this was at the expense of increased storage and computer execution time vis-a-vis filtering. Fixed-point smoothing gave results intermediate between fixed interval smoothing and the filtered estimates.

A quantitative estimate of the improvement rendered by smoothing for nonlinear systems is difficult to obtain. An approximate expression for the examples considered in this chapter may be obtained from the smoothing error variance, eqn. 3.40a after linearizing the nonlinear systems considered (eqns. (4.1) and (4.5). Eqn. (3.40a) repeated here is

$$V_{\hat{x}}(y,y',t_1|t_2) = V_{\hat{x}}(y,y',t_1) - \int_{t_1}^{t_2}\int_\Omega\int_\Omega\int_\Omega V_{\hat{x}}(y,y_2,t)\ \Xi^T(y_3,\tau;y_2,t)$$

$$\left(\frac{\partial h}{\partial \tau}\right)^T R\left(\frac{\partial h}{\partial \tau}\right)\ \Xi\ (y_5,t_1,y_4;\tau)\ V_{\tilde{x}}(y_5,y',t_1)$$

$$dy_2 dy_3 dy_4 dy_5 d\tau \tag{4.8}$$

where the Green's function matrix (4.8) Ξ satisfies the linearized system (3.27) with N replaced by its functional derivative. (4.8) indicates that the smoothing error variance $V_{\tilde{x}}(y,y',t_1|t_2)$ is always less than the filtering error variance $V_{\tilde{x}}(y,y_2,t)$. However, this is not an easy expression to evaluate because of the multiple integrations involved. Furthermore, this is a coarse approximation. So our safest guide is a qualitative rather than a quantitative examination of the reduction in error variance achieved.

4.3. Summary

To summary, in this chapter we have considered application of the various estimation algorithms developed in other chapters of this thesis to representative engineering problems arising in chemical and nuclear reactor theory. It was found that the algorithms of this research result in a considerable amount of saving in computer memory and execution time over the previously existing algorithms, when executed numerically. However, it was found that the equation giving a quantitative relationship between the filtering and smoothing error covariances was rather difficult to solve numerically, and hence a qualitative judgment factor was used. For the examples considered in this research, smoothing did indeed improve system performance, in that it approximated more closely the noise free trajectory than the filtered estimate, and also resulted in lesser error covariance than the filtered estimate. This was, however, achieved at the cost of increasing execution time and storage requirements.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

This dissertation presents the development of estimation algorithms for a class of linear and nonlinear second order distributed systems. The salient features of the algorithms developed are that they (1) are on-line, except for the fixed-interval smoothing algorithm, which is necessarily off-line, (2) are substantiated by theoretical analysis, (3) require fewer partial differential equations to be solved than previously existing algorithms. For instance, the 'two-filter' version of the fixed interval smoothing estimator [3] employs six partial differential equations, whereas the filter, using essentially an identical model, as developed in Chapter III uses only three equations.

The general class of systems for which estimation algorithms are developed using a least squares approach are described in Chapter II. No prior knowledge of noise statistics is assumed. Basically, the noisy plant is described by a vector partial differential equation

$$x_t = f(x,y,x_y,x_{yy},t,g) + w(y,t) \tag{5.1}$$

where w(y,t) is the distributed noise of unknown statistics. g is a parameter described by the ordinary differential equation

$$\dot{g}(t) = f(t,g) + \xi(t) \tag{5.2}$$

$\xi_1(t)$ is lumped additive noise of unknown statistics. Furthermore, homogeneous boundary conditions of the form

$$f(x,y,x_y,t,\xi_2) = 0 \quad , \quad y \in \Omega \tag{5.3}$$

are assumed, where $\xi_2(t)$ is a noisy term whose behavior is described by

$$r(t) = f(r,t) + \xi_3(t) \tag{5.4}$$

Such a general system governs a large class of distributed systems, involving, for instance, heat exchangers.

Furthermore, in the absence of the parameters q and r, this lumped-distributed system reduces to a more conventional distributed system, either parabolic or hyperbolic, whose behavior is governed by

$$x_t = f(x,y,x_y,x_{yy},t) + w(y,t) \qquad y \in \Omega \tag{5.5}$$

with homogeneous boundary conditions

$$g(x,y,t,x_y) = 0 \qquad y \in \partial\Omega \tag{5.6}$$

Such systems arise extensively in the physical world.

The solution to the estimation problem is obtained by constraining the estimator to minimize a positive definite quadratic form. This gives rise to a two-point boundary value problem. The boundary value problem is converted to an initial value problem using invariant embedding. In the case of filtering a 'running-stage' invariant embedding technique is employed to obtain the estimator equation. In the case of fixed point and fixed lag smoothing, 'fixed-stage' invariant embedding is employed. The Ricatti transformation is used to affect the fixed-interval smoothing estimator.

Such a solution to the estimation problem is significant, not only for purely distributed systems, but also for the lumped-distributed systems which arise so frequently in heat exchange problems, for instance. The nonlinear and linear smoothing problem considered in Chapter II enables one to take advantage of the considerable reduction in error covariance achieved using smoothing vis á vis filtering. A quantitative examination of simulation results in Chapter IV brought out the improvement rendered possible. Furthermore, these smoothing algorithms, as pointed out earlier, are online.

In Chapter III, a more restricted class of linear second order distributed systems are considered, to take advantage of the well developed adjoint formalism, using Green's theorem, available for such systems. The system equations are

$$x_t = L(y, x_y, x_{yy}, x, t) + w(y,t) \qquad y \in \Omega \tag{5.7}$$

where L is a linear second order vector operator, along with the first order homogeneous boundary conditions

$$N_b(x, y, x_y, t) = 0 \qquad y \in \partial\,\Omega \tag{5.8}$$

and linear measurements

$$Z(y,t) = H(y,t)\ x(y,t) + v(y,t) \tag{5.9}$$

The noises w and v are assumed zero mean, independent increment white processes. Using the innovations approach, minimum variance estimates for the filtering and smoothing problem are obtained. This has the advantage of again having the smoothing algorithms online. An exact analytical expression for the reduction in error variance using smoothing vis á vis filtering is available (Eqn. (3.41)) and demonstrates the feasibility of smoothing over filtering.

With proper identification the results for linear systems, using the techniques of Chapters II and III are identical. It is shown that the same algorithms result when the weighting matrices for the dynamic and observation residuals using least squares are chosen as the inverse of the dynamic and measurement noise covariances. Furthermore, as the domain of estimation reduces to a point in Euclidean n-space, corresponding lumped estimation algorithms result.

The results are applied to representative problems arising in various engineering disciplines. The numerical problem posed by these algorithms is treated in Chapter IV. One of the representative methods of solving partial differential equations, viz. the finite difference method is examined from two different viewpoints. In the first approach, akin to the one used by Seinfeld [14], the discretization is done only in space. The resulting set of ordinary differential equations are then solved using any of the conventional methods for solving these equations.

In the second approach, discretization in both space and time is affected. The resulting set of algebraic equations are then solved using the Thomas algorithm [29]. The results of the two approaches taken are compared for accuracy, computer execution time and storage requirements. It is seen that for the class of problems considered here and for the approach taken here, namely the finite difference method, discretization in both space and time is far superior to the other method, discretization in space alone. This comparison is based on the criteria established previously, viz. accuracy, execution time and storage requirements.

Recommendations for Further Study

As pointed out in Chapter I, a rigorous extension of Fokker-Planck techniques to the infinite-dimensional estimation problem needs to be investigated. Recently Balakrishnan [1] has initiated the study of Itô processes in abstract spaces. These methods have been successfully exploited by Bensoussan [30], Lions [31], etc., to derive Kalman type algorithms for unbounded linear operators. However, as yet the application of Fokker-Planck techniques to unbounded (linear and nonlinear) operators has not been initiated rigorously. This study would be very profitable and give a great deal of insight to the general estimation problem, as it would entail a study of the spectral properties of the operators to which such techniques can be applied. Thau [32] took a step in applying Fokker-Planck techniques to linear DP systems by calculating the backward Kolmogorov equation for the heat equation, in one dimension, excited by white Gaussian noise. However, the approach taken was purely heuristic and cannot be extended easily to higher order linear (or nonlinear) systems.

Another field which needs investigation is the sensitivity problem of various linear and nonlinear estimation algorithms to errors in modeling the plant, initial error covariance or the noise covariance when they are known. This would help give insight into errors that may arise due to any simplifying assumptions in the plant dynamics or noise statistics.

A field of major interest which has not been studied at all in the DP literature is the adaptive estimation problem. The techniques and the algorithms developed in this dissertation are very useful in carrying out online adaptive estimation, and can be extended to incorporate identification also. This area needs further investigation.

APPENDICES

APPENDIX A

THE INVARIANT EMBEDDING SOLUTION TO THE FILTERING PROBLEM

As pointed out in Section (2.3) a filtering algorithm is a necessary adjunct to any smoothing algorithm. The presentation here is not new and has already been presented elsewhere by Lamont [8] and Seinfeld et al. [14]. This is given over here only to complement the various smoothing algorithms derived in Chapter II.

Invariant embedding primarily consists of converting a boundary value problem (BVP) into an initial value problem by embedding the BVP in a more general class of problems. For the BVP arising in Sec. 2.2, we embed the terminal condition at time t_f, Eqn. (2.15) into a more general class of final conditions thus:

$$\begin{aligned} \lambda_1(y,t_f) &= C_1(y) \quad , \; y \in \bar{\Omega} \\ \lambda_3(t_f) &= C_3 \\ \lambda_4(t_f) &= C_4 \end{aligned} \tag{A.1}$$

Denote the general class of solutions to the BVP by

$$\begin{aligned} x(y,t_f) &= r_1(C_1,C_3,C_4,y,t_f) \\ q(t_f) &= r_3(C_1,C_3,C_4,t_f) \\ r(t_f) &= r_4(C_1,C_3,C_4,t_f) \end{aligned} \tag{A.2}$$

The solution desired is $r_1(0,0,0,y,t_f)$, $r_3(0,0,0,t_f)$, $r_4(0,0,0,t_f)$. Since

$$x(y,t_f + \Delta t) = r_1(C_1 + \Delta C_1,\ C_3 + \Delta C_3,\ C_4 + \Delta C_4,\ y, t_f + \Delta t_f) \qquad (A.3)$$

the left and right hand sides of (A.3) are expanded in a Taylor series to yield

$$r_1(C_i + \Delta C_i, y, t_f) = r_1(C_i, y, t_f) + \frac{\delta H}{\delta C_1}\Delta t = r_1(C_i, y, t_f) + \int_\Omega \frac{\delta r_1}{\delta C_1}\Delta C_1 d\Omega$$

$$+ \frac{\partial r_1}{\partial t_f}\Delta t + \frac{\partial r_1}{\partial C_3}\Delta C_3 + \frac{\partial r_1}{\partial C_4}\Delta C_4 \quad i = 1,3,4 \qquad (A.4)$$

From (2.14)

$$\Delta C_1 = -\frac{\delta H}{\delta r_1}\Delta t$$

$$\Delta C_4 = -\frac{\partial H}{\partial r_4}\Delta t \qquad (A.5)$$

$$\Delta C_3 = -\frac{\partial H}{\partial r_3}\Delta t$$

Substituting (A.5) into (A.4) yields, in the limit as $\Delta t \rightarrow 0$

$$\frac{\partial r_i}{\partial t_f} - \int_\Omega \frac{\delta r_i}{\delta C_1}\frac{\delta H}{\delta r} d\Omega - \frac{\partial r_i}{\partial r_3}\frac{\partial H}{\partial r_3} - \frac{\partial r_i}{\partial C_4}\frac{\partial H}{\partial r_1} = \frac{\delta H}{\delta C_1} \qquad (A.6)$$

$$i = 1,3,4$$

Equations (A.6) give the desired initial value problem for S_t: $\{x(y,t),\ q(t),\ r(t),\ y \in \bar{\Omega}\}$. Assume solutions of the form [8]

$$
\begin{aligned}
r_1(C_1,C_3,C_4,y,t_f) &= \hat{x}(y,t_f) - \int_\Omega P\ (y,y_1,t_f)C_1(y_1)\ d\Omega_1 \\
&\quad - P_{13}(y,t_f)C_3 - P_{14}(y,t_f)C_4 \\
r_3(C_1,C_3,C_4,t_f) &= \hat{q}(t_f) - \int_\Omega P_{31}(y,y_1,t_f)C_1(y_1)\ d\Omega_1 \\
&\quad - P_{33}(y,t_f)C_3 - P_{34}(y,t_f)C_4 \qquad \text{(A.7)} \\
r_4(C_1,C_3,C_4,t_f) &= \hat{r}(t_f) - \int_\Omega P_{41}(y,y_1.t_f)C_1(y_1)\ d\Omega_1 \\
&\quad - P_{43}(y,t_f)C_3 - P_{44}(y,t_f)C_4
\end{aligned}
$$

where $\hat{x}$, $\hat{q}$, $\hat{r}$ are the desired solutions. The P_{11} term is assumed continuous, bound, symmetric and symmetric in its arguments, too, i.e.,

$$P_{11}(y_1,y_2,t) = P_{11}{}^T(y_2,y_1,t_f)$$

and similar results hold for the other P terms, viz.

$$
\begin{aligned}
P_{13}(y,t) &= P_{31}{}^T(y,t)\ ; \qquad P_{14}(y,t) = P_{41}{}^T(y,t) \\
P_{33}(t) &= P_{33}{}^T(t)\ ; \qquad P_{44}(t) = P_{44}{}^T(t) \\
P_{43}(t) &= P_{34}{}^T(t)
\end{aligned}
$$

Substituting (2.39a) into (2.36) yields

$$\frac{\partial x(y,t_f)}{\partial t_f} - \int_\Omega \frac{\partial P_{11}}{\partial t_f} C_1 d\Omega_1 - \frac{\partial P_{13}}{\partial t_f} C_3 - \frac{\partial P_{14}}{\partial t_f} C_4 + \int_\Omega P_{11}\Big\{[N]^*_{r_1} C_1$$

$$- \int_\Omega \left(\frac{\partial h}{\partial r_1}\right)^T Q(z-h)\, d\Omega_1\Big\}\, d\Omega_1 + P_{13}\Big\{[N_q]^T_{r_3} C_3$$

$$+ \int_\Omega [N]_{r_3} C_1 d\Omega_1\Big\} + P_{14}\Big\{[N_r]_{r_4} C_4 + \int_{\partial\Omega} [N_b]_{r_1} KC_1 d(\partial\Omega)\Big\}$$

$$= N(r_1(C_1,C_3,C_4,y,t_f), r_3,y,t_f) - \int_\Omega R^{-1} C_1 d\Omega \qquad \text{(A.8)}$$

Linearize (A.8) about a nominal trajectory $r_1(0,0,0,y,t_f)$, $r_3(0,0,0,t_f)$, $r_4(0,0,0,t_f)$ and compare coefficients of like terms in $C^{(o)}$ and $C^{(1)}$ to get

$$C_1^{(o)}: \frac{\partial \hat{x}}{\partial t_f} = N(\hat{x},q,y,t_f) + \int_\Omega \int_\Omega P_{11}\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(y_1,y_2,t_f)$$

$$(z - h(\hat{x},y_1,t_f))\, d\Omega_1 d\Omega_2 \quad y \in \Omega \qquad \text{(A.9)}$$

$$C_1^{(1)}: \int_\Omega \frac{\partial P_{11}}{\partial t_f} C_1 d\Omega_1 = \int_\Omega \Big\{P_{11}[N]^*_{\hat{x}} + [N]_{\hat{x}} P_{11} + R^{-1} + P_{13}[N]_{\hat{q}} P_{31}$$

$$+ \int_\Omega \int_\Omega \left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h(x,y_3,t))\right]_{\hat{x}} P_{11} d\Omega_2 d\Omega_3\Big\}$$

$$C_1 d\Omega_4 + \int_{\partial\Omega} P_{14}[N_b]_{\hat{r}} KC_1 d(\partial\Omega_1) \qquad \text{(A.10)}$$

In (A.10) the first term on the right can be written, using Green's identity,

$$\int_{\Omega} P_{11} N^{*}_{\hat{x}} C_1 \, d\Omega_1 = \int_{\Omega} P_{11} \left[N_{\hat{x}_1}^{T} \right] C_1 \, d\Omega_1 + \int_{\partial\Omega} ([N_b]^{*}_{x} P_{11}) C_1 d(\partial\Omega_1) \qquad \text{(A.11)}$$

With these relations, identifying P_{11} as the error covariance of the estimate of x, we get

$$\frac{\partial P_{11}}{\partial t_f} = N_{\hat{x}} P_{11} + P_{11} N_{\hat{x}_1}^{T} + P_{13} [N]_{\hat{q}}^{T} + [N]_{\hat{q}} P_{31}^{T} + R^{-1}$$

$$+ \int_{\Omega} \int_{\Omega} P_{11} \left[\left(\frac{\partial h}{\partial \hat{x}} \right)^{T} Q(z - h(x, y_3, t)) \right]_{\hat{x}} P_{11} d\Omega_2 d\Omega_3 , \; y , y_1 \in \Omega \qquad \text{(A.12)}$$

where the arguments of the various terms are self evident. The boundary condition for P_{11} is obtained from (A.10) and (A.11) and is

$$[N_b]^{*}_{\hat{x}} P_{11}(y, y_1, t_f) + P_{14}(y, t_f) [N_b]_{\hat{r}} K(y_1, t_f) = 0 \qquad \text{(A.13)}$$

$$y \in \partial \Omega , y_1 \in \bar{\Omega}$$

with initial conditions

$$P_{11}(y, y_1, t_o) = P_o(y, y_1) \; ; y, \; y_1 \in \Omega \qquad \text{(A.14)}$$

The estimator equation for $\hat{x}(y, t_f)$ is given by (A.9) with boundary condition

$$[N_b]_{\hat{x}} \hat{x} = 0 \qquad y \in \partial \Omega \qquad \text{(A.15)}$$

and initial condition

$$\hat{x}(y,t_o) = x_o(y) \qquad (A.16)$$

Similarly, comparing coefficients of $C_3^{(1)}$ and $C_4^{(1)}$ we get

$$\frac{\partial P_{13}}{\partial t_f}(y,t_f) = [N]_{\hat{x}}P_{13} + [N]_{\hat{q}}P_{33} + P_{13}[N\hat{q}]_{\hat{q}}^T$$

$$+ \int_\Omega\int_\Omega P_{11}\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h)\right]_{\hat{x}} P_{13}\,d\Omega_1 d\Omega_2, \quad y\,\varepsilon\,\Omega \qquad (A.17)$$

$$\frac{\partial P_{14}}{\partial t_f} = [N]_{\hat{x}}P_{14} + [N]_{\hat{q}}P_{34} + P_{14}[N_{\hat{r}}]_{\hat{r}}^T$$

$$+ \int_\Omega\int_\Omega P_{11}\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h)\right]_{\hat{x}} P_{14}\,d\Omega_1 d\Omega_2 \qquad (A.18)$$

with boundary conditions

$$P_{13}: \quad [N_b]_{\hat{x}}P_{13} + [N_b]_{\hat{r}}P_{43} = 0 \qquad y\ \varepsilon\ \partial\,\Omega \qquad (A.19)$$

$$P_{14}: \quad [N_b]_{\hat{x}}P_{14} + [N_b]_{\hat{r}}P_{44} = 0 \qquad y\ \varepsilon\ \partial\,\Omega \qquad (A.20)$$

In an identical fashion, substituting (A.7b) and (A.7c) into (A.6b) and A.6c) respectively, comparing coefficients of like terms in the C's after a first order linearization yields for the estimator equations

$$\frac{d\hat{q}}{dt_f} = N_{\hat{q}}(t_f,\hat{q}) + \int_\Omega\int_\Omega P_{31}\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h(\hat{x},y_1,t_f))\,d\Omega_1 d\Omega_2 \qquad (A.21)$$

$$\frac{d\hat{r}}{dt_f} = N_{\hat{r}}(t_f,\hat{r}) + \int_\Omega\int_\Omega P_{41}\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h(\hat{x},y_1,t_f))\ d\Omega_1 d\Omega_2 \qquad \text{(A.22)}$$

and the error covariance equations

$$\frac{dP_{33}(t_f)}{dt_f} = [N_{\hat{q}}]_{\hat{q}}P_{33} + P_{33}[N_{\hat{q}}]^T_{\hat{q}} + T_1^{-1}(t_f)$$

$$+ \int_\Omega\int_\Omega P_{13}\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h)\right]_{\hat{x}} P_{31} d\Omega_1 d\Omega_2 \qquad \text{(A.23)}$$

$$\frac{dP_{34}(t_f)}{dt_f} = [N_{\hat{q}}]_{\hat{q}}P_{34} + P_{34}[N_{\hat{r}}]^T_{\hat{r}}$$

$$+ \int_\Omega\int_\Omega P_{31}\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h)\right]_{\hat{x}} P_{14} d\Omega_1 d\Omega_2 \qquad \text{(A.24)}$$

$$\frac{dP_{44}(t_f)}{dt_f} = [N_{\hat{r}}]_{\hat{r}}P_{44} + P_{44}[N_{\hat{r}}]^T_{\hat{r}} + T_2^{-1}(t_f)$$

$$+ \int_\Omega\int_\Omega P_{41}\left[\left(\frac{\partial h}{\partial \hat{x}}\right)^T Q(z - h)\right]_{\hat{x}} P_{14} d\Omega_1 d\Omega_2 \qquad \text{(A.25)}$$

The filter obtained is summarized in Table 14.

The results obtained in this section are not new, but a more general model than previous works by Lamont [8], Seinfeld [14] has been considered. In case the state of the system is described only by $\{x(y,t)\}$ rather than $\{x(y,t), g(t), r(t)\}$ the estimator reduces to the one obtained by Lamont [8]. Furthermore these equations may be regarded as the continuous analogue of the discrete estimator obtained by Seinfeld

Table 14. Filtering Algorithm

Estimates		Equation	Init. Condition	B.C.
	$\hat{x}(y,t_f)$	(A.9)	(A.1b)	(A.15)
	$\hat{q}(t_f)$	(A.21)	$\hat{q}(0)$	None
	$\hat{r}(t_f)$	(A.22)	$\hat{r}(0)$	None
Covariances				
	$P_{11}(y_1,y_2,t_f)$	(A.12)	P_o	(A.13)
	$P_{13}(y,t_f)$	(A.17)	$P_{13}(y,0)$	(A.19)
	$P_{14}(y,t_f)$	(A.18)	$P_{14}(y,0)$	(A.20)
	$P_{33}(t_f)$	(A.23)	$P_{33}(0)$	None
	$P_{34}(t_f)$	(A.24)	$P_{34}(0)$	None
	$P_{44}(t_f)$	(A.25)	$P_{44}(0)$	None

et al. [14]. The purpose of this derivation, as previously stated, is to to serve as a necessary adjunct for the smoothing algorithms developed in Section 2.3.

APPENDIX B

NUMERICAL IMPLEMENTATION OF THE ESTIMATION ALGORITHM

It was pointed out in Chapter IV that the estimation algorithms developed in Chapters II and II would be studied using the finite difference method. Two techniques for discretizing were outlined in that chapter--viz. discrete space discrete time (DSDT) and the discrete space continuous time (DSCT) approaches. In this appendix we examine the discretization of the estimator equations using these two approaches. The techniques used in discretizing the integro-partial differential equations are detailed briefly, and the resulting discrete equations are outlined fully.

The procedure is detailed for the chemical reactor problem considered in Chapter IV. Similar remarks apply to the other problem considered in that chapter. For convenience, we repeat the model equations.

dynamics: $$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial x^2} - 10\,\frac{\partial y}{\partial x} - 5y^2 + w(x,t) \qquad \text{(B.1)}$$

b.c.: $$\frac{\partial y}{\partial x} = 0 \qquad x = 1 \quad \forall t \qquad \text{(B.1a)}$$

b.c.: $$\frac{\partial y}{\partial x} + 10(1 - y) = 0 \qquad x = 0 \quad \forall t \qquad \text{(B.1b)}$$

i.c.: $$y(x,0) = 0 \qquad \forall\ x \qquad \text{(B.2)}$$

measurements: $$z(x,t) = y(x,t) + v(x,t) \ : \ x \in (0,1) \qquad \text{(B.3)}$$

The noisy dynamics and measurements were simulated by assuming w and v to be zero mean, white in time and independent, viz.,

$$E\{v(x,t)\} = E\{w(x,t)\} = 0$$

$$E\{v(x,t)v^T(x',t')\} = \delta(t - t')\delta(x - x')$$

$$E\{w(x,t)w^T(x',t')\} = \delta(t - t')\delta(x - x')$$

The distribution of v and w was simulated as being uniform in the interval $[-\sqrt{3},\sqrt{3}]$ which gave the required mean and variance. Choosing a performance index of the form (cf. eqn. (2.6)

$$J = \frac{1}{2}\int_o^{t_f}\int_o^1 \left\{(z(x,t) - y(x,t))^2 + (\frac{\partial y}{\partial t} - N)^2\right\} dxdt \qquad \text{(B.4)}$$

where N is the nonlinear spatial operator on the R.H.S. of (B.1) results in (cf. Appendix A, Table 1) <u>the filtering estimate u</u>

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - 10 \ \frac{\partial u}{\partial x} - 5u^2 + \int_o^1 P(x,x_1,t)(z(x_1,t) - u(x_1,t)) \ dx_1 \qquad \text{(B.5)}$$

with boundary conditions

$$\frac{\partial u}{\partial x} = 0 \qquad x = 1 \quad \forall t \qquad \text{(B.5a)}$$

$$\frac{\partial u}{\partial x} + 10(1 - u) = 0, \qquad x = 0 \quad \forall t \qquad \text{(B.5b)}$$

and initial conditions (assumed)

$$u(x,0) = .10 \quad \forall \; x \in (0,1) \tag{B.6}$$

<u>the error covariance</u>

$$\frac{\partial P}{\partial t}(x,x_1,t) = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial {x_1}^2} - 10P(u(x,t) + u(x_1,t)) - 10\left(\frac{\partial P}{\partial x} + \frac{\partial P}{\partial x_1}\right)$$

$$+ \delta(x - x_1) - \int_o^1 P(x,x_2,t)P(x_2,x_1,t)\, dx_2 \tag{B.7}$$

and boundary conditions

$$\frac{\partial P}{\partial x} = 0 \quad \text{at} \quad x = 1 \;\; \forall \; x_1,t \qquad \text{and at} \quad x_1 = 0 \;\; \forall \; x,t \tag{B.8}$$

$$\frac{\partial P}{\partial x} - 10u = 0 \qquad \left\{ \begin{array}{ll} x = 0 & x_1 = 0 \\ \forall \; x_1 \in D & \forall \; x \in D \\ \forall \; t & \forall \; t \end{array} \right. \tag{B.8a}$$

and initial conditions (assumed)

$$P(x,x_1,0) = 2.50 \tag{B.9}$$

The various smoothed estimates can also be similary written down, following the analysis of Sec. 2.3. For illustration's sake we give below the fixed interval smoothing algorithm that results.

<u>Fixed interval smoothing estimate:</u>

$$\frac{\partial u(x,t|t_f)}{\partial t} = \frac{\partial^2 u(x,t|t_f)}{\partial x^2} - 10\frac{\partial u(x,t|t_f)}{\partial x} - 5u(x,t|t_f)$$

$$+ \int_0^1 P^{-1}(x,x_1,t)\ (u(x_1,t|t_f) - u(x,t))\ dx_1 \qquad \text{(B.10)}$$

with boundary conditions

$$\frac{\partial u(x,t|t_f)}{\partial x} = 0 \qquad x = 1 \quad \forall\, t \qquad \text{(B.11a)}$$

$$\frac{\partial u(x,t|t_f)}{\partial x} + 10(1 - u(x,t|t_f)) = 0 \qquad x = 0 \quad \forall\, t \qquad \text{(B.11b)}$$

initial condition

$$u(x,t_f|t_f) = u(x,t_f), \text{ the filtered estimate} \qquad \text{(B.12)}$$

The smoothed estimate is, of course, independent of its corresponding error covariance. Similar results hold for other algorithms. We shall not discuss them further here, for brevity's sake.

As pointed out previously, the main thrust of this appendix is to obtain the discretized equations that were actually implemented on the U-1108. The noisy data was generated using the Crank-Nicholson method [29] with forward projection in time. The Crank-Nicholson procedure is a second-order correct analog of the time derivative. All the finite differences are written about the point x_i, $t_{n+1/2}$, which is halfway between the known and the unknown time levels. In Figure 6 this point is shown as a cross. Values of the dependent variable, u, are computed only at the points designated by circles. The second-order correct analog of the time derivative at x_i, $t_{n+1/2}$ is

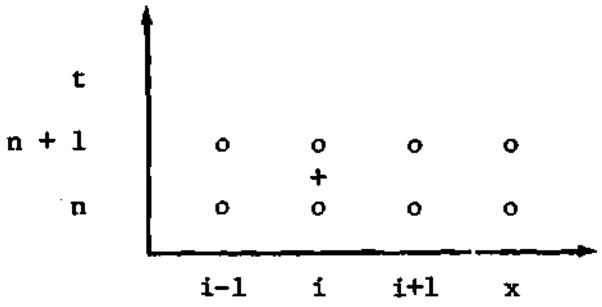

Figure 6. Center of Analogs for Crank-Nicholson Method.

$$\frac{\partial u}{\partial t}_{i,n+1/2} = \frac{u_{i,n+1} - u_{i,n}}{\Delta t} - \left(\frac{\partial^3 x}{\partial t^3}\right)_{i,n+1/2} \frac{(\Delta t)^2}{24} - \cdot \cdot$$

The second-order correct approximation for $(\partial^2 u/\partial x^2)$ is obtained by the arithmetic average of its finite difference analogs at x_i, t_n and x_i, t_{n+1}. The resulting analog is

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,n+1/2} \approx \frac{1}{2}\left(\frac{u_{i+1,n} - 2u_{i,n} + u_{i-1,n}}{(\Delta x)^2} + \frac{u_{i+1,n+1} - 2u_{i,n+1} + u_{i-1,n+1}}{(\Delta x)^2}\right)$$

For this problem define $x_i = (i - 1/2)\Delta x$, Δx and Δt being space-time intervals to get

$$\frac{u_{i,n+1} - u_{i,n}}{\Delta t} = \frac{1}{2}\left\{\delta_x^2 u_{i,n+1/2} + \delta_x^2 u_{i,n+1/2}\right\} - 5\left\{\mu_x u_{i,n} + \mu_x u_{i+1,n}\right\}$$

$$- \frac{5}{2}\left\{u_{i,n+1/2}(u_{i,n} + u_{i,n+1})\right\} + w_{i,n+1/2} \qquad \text{(B.13)}$$

where

$$\delta_x u_{i,n} = (u_{i+1/2,n} - u_{i-1/2,n})/\Delta x$$

$$\mu_x u_{i,n} = (u_{i+1,n} - u_{i-1,n})/2\Delta x$$

Incorporating the boundary conditions (B.1) we can write the difference equation for R equi-spaced intervals,

$$(1 + 5\Delta x)u_{i-1,n+1} - \left(2 + 5\Delta x^2 u_{i,n+1/2} + \frac{2\Delta x^2}{\Delta t}\right)u_{i,n+1}$$

$$+ (1 - 5\Delta x)\ u_{i+1,n+1} = -(1 + 5\Delta x)\ u_{i-1,n} + \left(2 + 5\Delta x^2 u_{i,n+1/2} - \frac{2\Delta x^2}{\Delta t}\right)u_{i,n}$$

$$+ (-1 + 5\Delta x)\ u_{i+1,n} - 2\Delta x\ w_{i,n} \qquad 2 \leq i \leq R - 1 \qquad \text{(B.14)}$$

The L.H.S. of this equation is the value of u at the new time level (n + 1) in terms of the old time level. As such it involves an unknown coefficient $u_{i,n+1/2}$. This can be obviated by projecting this coefficient forward in time thus

$$u_{i,n+1/2} = u_{i,n} + \left(\frac{\partial u}{\partial t}\right)_{i,n} \frac{\Delta t}{2} \qquad \text{(B.15a)}$$

where

$$\left(\frac{\partial u}{\partial t}\right)_{i,n} = u_{i,n} + \frac{\Delta t}{2\Delta x^2} \left\{(1 + 5\Delta x)\ u_{i-1,n} + (1 - 5\Delta x)\ u_{i+1,n}\right.$$

$$\left. -(2 + 5\Delta x^2 u_{i,n})\ u_{i,n} + \Delta x^2 w_{i,n}\right\} \qquad \text{(B.15b)}$$

Then eqn. (B.14) and (B.15) give the difference equation governing the interior behavior of the system. The boundary equation becomes

$$-\left\{5\Delta x^2 u_{1,n+1/2} + \frac{2\Delta x^2}{\Delta t} + (1 + 5\Delta x)\right\} u_{1,n+1} + (1 - 5\Delta x)\ u_{2,n+1}$$

$$= -(1 - 5\Delta x)\ u_{2,n} + \left\{5(\Delta x)^2 u_{1,n+1/2} - \frac{2\Delta x^2}{\Delta t} + (1 + 5\Delta x)\right\} u_{1,n}$$

$$- 20\Delta x - 2\Delta x^2 w_{i,n} \qquad i = 1 \qquad \text{(B.16)}$$

with $u_{1,n+1/2}$ computed from (B.15). Equations (B.14) and (B.16) comprise a set of implicit tridiagonal system of equations which is solved using the Thomas algorithm [29] for such systems.

The filter equation (B.4) was solved using the above procedure, too. The integral was approximated by Simpson's rule. The error covariance equation (B.7) was solved using the implicit alternating direction method [29].

A brief sketch of the alternating direction-implicit method is in order here. For a two space dimension problem such as posed by the covariance equation (B.7) this method involves the alternate use of two different finite difference analogs to (B.7). For the first finite difference equation the analog to $\partial^2 u/\partial x^2$ is written at the new time level, t_{n+1}, and the analog to the y-derivative, $\partial^2 u/\partial y^2$ is written at the old time level, t_n. Consequently, they are implicit in the x-direction and explicit in the y-direction. The resulting tridiagonal system can be easily solved using, for instance, the Thomas algorithm [29]. If this were used repeatedly, the method would be unstable except for restricted values of the ratio $(\Delta t/\Delta y^2)$ since the difference equation is explicit in the y-direction. So for the next time step a tridiagonal system is set up which is explicit in the x-direction and implicit in the y-direction with the same size Δt, hence the name of the method. The procedure is repeated till such time as convergence is achieved.

We can now write the discretized error covariance equation (B.7). Denote by i the index in the x-direction, by j the index in the x_1-direction. Let n denote the time index. The increment in both spatial dimensions is the same, Δx. Then we can write x-implicit equations:

$$\left(P^{n+1}_{i+1,j} - 2P^{n+1}_{i,j} + P^{n+1}_{i-1,j}\right)/\Delta x^2 + \left(P^{n}_{i,j+1} - 2P^{n}_{i,j} + P^{n}_{i-1,j}\right)/\Delta x^2$$

$$- 10\left[\left(P^{n+1}_{i+1,j} - P^{n+1}_{i-1,j}\right)/2\Delta x + \left(P^{n}_{i,j+1} - P^{n}_{i,j-1}\right)/2\Delta x\right]$$

$$- 10P^{n+1}_{i,j}\left(u^{n+1}_{i} + u^{n+1}_{j}\right) + \delta_{i,j} - \int_o^1 P^n(x_i,x)P^n(x,xj)dx$$

$$= \left(P^{n+1}_{i,j} - P^{n}_{i,j}\right)/\Delta t \qquad \text{(B.17a)}$$

x_1-implicit equations:

$$\left(P^{n+1}_{i+1,j} - 2P^{n+1}_{i,j} + P^{n+1}_{i-1,j}\right)/\Delta x^2 + \left(P^{n+2}_{i,j+1} - 2P^{n+2}_{i,j} + P^{n+2}_{i,j-1}\right)/\Delta x^2$$

$$- 10\left[\left(P^{n+1}_{i+1,j} - P^{n+1}_{i-1,j}\right)/2\Delta x + \left(P^{n+2}_{i,j+1} - P^{n+2}_{i,j-1}\right)/2\Delta x\right]$$

$$- 10P^{n+2}_{i,j}\left(u_i^{n+2} + u_j^{n+2}\right) + \delta_{ij} - \int_o^1 P^{n+1}(x_i,x)P^{n+1}(x,x_j)dx$$

$$= \left(P^{n+2}_{i,j} - P^{n+2}_{i,j}\right)/\Delta t \qquad \text{(B.17b)}$$

These equations are succesively solved along with the associated boundary conditions using the Thomas algorithm [29]. The integrals were computed by Simpson's rule.

Next the DSCT method as used by Seinfeld et al. [14] is described Conceptually, this method is rather simple. The space derivatives are approximated thus (i denotes the spatial index)

$$\partial u/\partial x = [u(i,t) - u(i - 1,t)]/\Delta x$$

$$\partial^2 u/\partial x^2 = [u(i + 1,t) - 2u(i,t) + u(i - 1,t)]/\Delta x^2$$

With this correspondence, the noisy system (B.1-2) can be discretized to yield

$$y_i = (y_{i+1} - 2y_i + y_{i-1})/\Delta x^2 - 10(y_i - y_{i-1})/\Delta x$$

$$- 5y_i + w_i(t) = f(y,t) + w_i(t) \qquad \text{(B.18)}$$

with (cf. B.la and B.lb)

$$y_o = (y_1 + 10\Delta x)/(1 + 10\Delta x) \qquad \text{(B.19a)}$$

$$y_n = y_{n-1} \qquad \text{(B.19b)}$$

The measurement process is

$$z_i(t) = y_i(t) + v_i(t) \qquad i = 2,3..n - 1 \tag{B.20}$$

The simulation was carried out by assuming the noises to be independent white processes with zero mean and unit covariance. The distribution was assumed to be uniform. Thus

$$E\ \{v_i(t)\} = E\ \{w_i(t)\} = 0$$

$$E\ \{v_i(t) w_j^T(\tau)\} = 0$$

$$E\ \{v_i(t) v_i^T(\tau)\} = E\ \{w_i(t) w_i^T(\tau)\} = \delta(t - \tau)$$

The estimate was desired to minimize the quadratic form

$$J = \frac{1}{2} \int_o^{t_f} \left[(\underline{z}(t) - \underline{y}(t))^2 + (\underline{\dot{y}} - \underline{f}(y,t))^2 \right] dt \tag{B.21}$$

This resulted in the invariant embedding equations corresponding to the ones obtained by Seinfeld [14].

Estimate: $\underline{u} = f(\underline{u},t) + P(\underline{z} - \underline{u})$ (B.22)

Error Covariance: $P = \frac{\partial f}{\partial u} P + P\left(\frac{\partial f}{\partial u}\right)^T + I - P^2$ (B.23)

with initial conditions (assumed)

$$u(t_o) = .00 \tag{B.24}$$

$$\underline{P}(t_o) = 2.5 \tag{B.25}$$

APPENDIX C

OBSERVABILITY CRITERION

There exists no general observability criterion for the class of nonlinear systems considered in Chapter II [31]. However, for linear systems as considered in Chapter III, the following Lemma [31] holds. For convenience, we repeat the unforced system model of Chapter III.

$$\frac{\partial x}{\partial t} = L_y x(y,t) \qquad \text{(C.1)}$$

$$\text{b.c.} \quad N_b x(y,t) = 0 \qquad \text{(C.2)}$$

$$\text{Measurement:} \quad z(y,t) = H(y,t)x(y,t) \qquad \text{(C.3)}$$

where L_y is an n x n matrix differential operator, $x(y,t)$ is the n vector state, and $H(y,t)$ is a bounded m x n matrix.

Lemma [31]: The system (C.1-3) is observable at t_o if and only if the linear self adjoint operator

$$M = \int_{t_o}^{t} \int_{\Omega} [H(y,t)K(y,t;y_1,t_o)]^*[H(y,t)K(y,t;y_1,t_o)]d\Omega_1 dt$$

has a bounded inverse for some $t_1 \geq t_o$. Here $K(\cdot)$ denotes the system's Green's function, and $(*)$ denotes the adjoint operation. This lemma implies that the system is observable at t_o if and only if M is positive definite and the columns of $H(\cdot)K(\cdot;y_1,t_o)$ are independent at t_o for some $t_1 \geq t_o$.

An example of a system which is not observable [31] is one in which a measurement is taken at just one point, i.e.,

$$H(\cdot) = \int_\Omega \delta(y - y_1)(\cdot)d\Omega$$

where the operator L in (C.1) is elliptic. For such a system M (eqn. (3.4)) is always unbounded. Such systems are not considered in our research.

BIBLIOGRAPHY

1. Balakrishnan, A. V., Introduction to Optimization Theory in a Hilbert Space, Springer-Verlag, Berlin, 1971.
2. Falb, P. L., "Infinite Dimensional Filtering: the Kalman-Bucy Filter in Hilbert Space," Info. and Control, Vol. 11, pp. 102-137, 1967.
3. Tzafestas, S., and Nightingale, J., "Optimal Filtering, Smoothing and Prediction in Linear Distributed Parameter Systems," Proc. I.E.E., Vol. 115, pp. 1208-1213, 1968.
4. Thau, F. E., "On Optimum Filtering for a Class of Linear Distributed Parameter Systems," ASME Journal of Basic Eng., Vol. 91, pp. 173-178, 1969.
5. Balakrishnan, A. V., and Lions, J. L., "State Estimation for Infinite Dimensional Systems," Journ. Computer and System Sciences, Vol. 1, pp. 391-403, 1967.
6. Meditch, J. S., "On State Estimation for Distributed Parameter Systems," Jour. Franklin Inst., Vol. 290, pp. 49-59, 1970.
7. Detchmendy, D. M. and Sridhar, R., "Sequential Estimation of States in Noisy Nonlinear Dynamical Systems," Jour. Basic. Eng., Vol. 88, pp. 362-368, 1966.
8. Lamont, G. B. and Kumar, K. S. P., "State Estimation in Distributed Parameter Systems Via Least Squares and Invariant Embedding," Jour. Math. Anal. Appl., Vol. 38, pp. 588-606, 1972.
9. Tzafestas, S. and Nightingale, J., "Maximum-Likelihood Approach to the Optimal Filtering of Distributed-Parameter Systems," Proc. I.E.E., Vol. 116, pp. 1085-1093, 1969.
10. Tzafestas, S., "Bayesian Approach to Distributed-Parameter Filtering and Smoothing," Intl. Jour. Control, Vol. 15, pp. 273-295, 1972.
11. Mayne, D. Q., "A Solution of the Smoothing Problem for Linear Dynamic Systems," Automatica, Vol. 3, pp. 245-255, 1966.
12. Meditch, J. S., Stochastic Optimal Linear Estimation and Control, McGraw-Hill, New York, 1970.
13. Sage, A. and Melsa, J., Estimation Theory with Applications to Communications and Control, McGraw-Hill, New York, 1971.

14. Seinfeld, J., Gavalas, G. and Hwang, M., "Nonlinear Filtering in Distributed Parameter Systems," Jour. Dynamic Systems, Measurement and Control, Vol. G93, pp. 157-163, 1971.

15. Kailath, T., "An Innovations Approach to Least-Squares Estimation, Part I: Linear Filtering in Additive White Noise," I.E.E.E. Trans. Auto. Cont., Vol. AC-13, pp. 646-655, 1968.

16. ________ and Frost, p., "An Innovations Approach to Least-Squares Estimation, Part II: Linear Smoothing in Additive White Noise," I.E.E.E. Trans. Auto. Cont., Vol. AC-13, pp. 655-660, 1968.

17. Courant, R. and Hilbert, D., "Methods of Mathematical Physics, Vol. II," Wiley-Interscience, New York, 1962.

18. Hwang, M., Seinfeld, J. and Gavalas, G., "Optimal Least-Squares Filtering and Interpolation in Distributed Parameter Systems," Jour. Math. Anal. Appl., Vol. 39, pp. 49-74, 1972.

19. Russell, D. and Lukes, D., "The Quadratic Criterion for Distributed Systems," SIAM Jour. Control, Vol. 7, pp. 75-83, 1969.

20. Gelfand, I.,and Fomin, S., "Calculus of Variations," Prentice-Hall, Englewood Cliffs, 1963.

21. Volterra, V., "Theory of Functionals," Dover, New York, 1958.

22. Vainberg, M., "Variational Methods for the Study of Nonlinear Operators," Holden-Day, San Fransisco, 1964.

23. Tzafestas, S., "Differential-Dynamic-Programming Approach to Nonlinear Distributed Parameter Control Systems," Proc. I.E.E., Vol. 116, pp. 1079-1085, 1969.

24. Bode, H. and Shannon, C., "A Simplified Derivation of Linear Least-Squares Smoothing and Prediction Theory," Proc. I.R.E., Vol. 38, pp. 417-425, 1950.

25. Atre, S. and Lamba, S., "Optimal Estimation in Distributed Processes Using the Innovations Approach," I.E.E.E. Trans. Auto. Control, Vol. AC-17, pp. 710-711, 1972.

26. Bryson, A. and Frazier, M., "Smoothing for Linear and Nonlinear Dynamic Systems," Proc. Optimum System Synthesis Conf., Sept. 11-13, 1962, Wright-Patterson AFB, Ohio.

27. Aris, R., "Elementary Chemical Reactor Analysis," Prentice-Hall, Englewood Cliffs, 1969.

28. Hetrick, D., ed.,"Dynamics of Nuclear Systems," Univ. of Arizona Press, Tucson, Ariz., 1972.

29. Ames, W., "Numerical Methods for Partial Differential Equations," Barnes and Noble, London, 1970.

30. Bensoussan, A., "Filtrage Optimale de Systemes Lineaires," Dunod, Paris, 1971.

31. Lions, J.-L., "Controle Optimale de Systemes Gouvernes par des Equations aux Derivees Partielles," Dunod, Paris, 1968.

32. Thau, F., "A Backward Equation for a Randomly Excited Diffusion Process," I.E.E.E. Trans. Auto. Control, Vol. AC-13, pp. 714-716, 1968.

33. Sahgal, R. and Webb, R. P., "Nonlinear Distributed Parameter Estimation," Proc. I.E.E.E. Conf. Decision and Control, pp. 89-93, 1972.

VITA

Raman Kumar Sahgal was born in Gujrat, Pakistan. He received the B. Tech. E.E. and M.S.E.E. from the Indian Institute of Technology, Delhi, and Carnegie-Mellon University in 1967 and 1968, respectively. He was employed by the Lockheed-Georgia Company from 1968 to 1970, when he joined the Georgia Institute of Technology to pursue studies toward the Ph.D. in Electrical Engineering.