JOURNAL OF COMPUTATIONAL PHYSICS 115, 1–11 (1994)

REVIEW ARTICLE

Optimal Estimation of Dynamically Evolving Diffusivities

KURT S. RIEDEL

Courant Institute of Mathematical Sciences, New York University, New York, New York 10012-1185

Received December 9, 1992; revised August 26, 1993

The augmented, iterated Kalman smoother is applied to system identification for inverse problems in evolutionary differential equations. In the augmented smoother, the unknown, time dependent coefficients are included in the state vector and have a stochastic component. At each step in the iteration, the estimate of the time evolution of the coefficients is linear. We update the slowly varying mean temperature and conductivity by averaging the estimates of the Kalman smoother. Applications include the estimation of anomalous diffusion coefficients in turbulent fluids and the plasma rotation velocity in plasma tomography.

I. INTRODUCTION

Estimation and system identification of distributed systems of partial differential equations (PDEs) are much researched fields. However, existing research in inverse problems concentrates almost exclusively on the case where the system equations are deterministic with unknown coefficients and where only the measurements have errors [3–5, 22, 26]. In contrast, finite-dimensional estimation and control theory allows and stresses the importance of including stochastic forcing in the system evolution equations to account for model error [15].

Researchers attempt to model the effect of microscopic turbulence in plasmas and fluids with anomalous diffusion coefficients. These effective equations for fluid flow are only an approximation of the actual evolution equations, and in many cases the model error is much larger than the measurement errors. Also, the anomalous diffusion coefficients are often time dependent, while most research in inverse problems is restricted to time independent coefficients. Model error has been included in the optimal estimation schemes of numerical weather prediction [9, 12]. By extending the state space, we are able not only to estimate the state of the system, but also the coefficients.

The Kalman filter–smoother is the optimal estimator of time dependent state vectors given noisy measurements and evolution equations with stochastic forcing. The Kalman smoother minimizes a quadratic functional which includes the residual squared error to the measurements. However, the Kalman filter generalizes the standard least squares analysis by including a second term which is proportional to the square of the stochastic forcing. The estimation problem is the mathematical dual to the control problem, and we refer the reader to Ref. [13] for an excellent description of the computational aspects of fluid dynamical control.

Since inverse problems are generally analyzed off-line, the fixed interval Kalman smoother is more appropriate than the Kalman filter. In this article, we propose to analyze time dependent inverse problems with model error using a novel quasilinear extended Kalman smoother. To determine the unknown coefficients, we extend the state space to include the original state space and the coefficient space.

As a running example, we consider the evolution of the temperature with an unknown, time dependent diffusivity, $\kappa(\mathbf{x}, t)$:

$$\partial_t T = \nabla \cdot \kappa(\mathbf{x}, t)\, \nabla T + \mathbf{s}(t) + \xi_1, \tag{1.1a}$$

where $\mathbf{s}$ is the known source function. ξ_1 is a random field which stochastically forces the heat equation. The random term represents small scale errors in the model. We assume that the unknown diffusivity evolves as

$$\partial_t \tilde{\theta} = \mu_2 \Delta \tilde{\theta} + \xi_2, \tag{1.1b}$$

where $\theta = \ln(\kappa(\mathbf{x}, t))$ and $\theta \equiv \bar{\theta}(\mathbf{x}, t) + \tilde{\theta}(\mathbf{x}, t)$ with known $\bar{\theta}(\mathbf{x}, t)$.

More generally, we consider stochastic evolutionary systems of the form

$$\begin{aligned} \partial_t \psi &= \mathbf{L}(\theta, t)\psi + \mathbf{s} + \mathbf{B}\xi_1, \\ \partial_t \tilde{\theta} &= \mathbf{M}(t)\tilde{\theta} + \xi_2, \end{aligned} \tag{1.2}$$

where ψ is the state variable and $\mathbf{s}$ is the known source func-

0021-9991/94 $6.00

tion. The unknown parameter vector is $\theta = \bar{\theta} + \tilde{\theta}$, where $\bar{\theta}(\mathbf{x}, t) \equiv E[\theta(\mathbf{x}, t)]$ and $\tilde{\theta} \equiv \theta - \bar{\theta}$. $E[\cdot]$ denotes the probabalistic expectation. $\mathbf{L}(\theta, t)$ and $\mathbf{M}(t)$ are generators of *smooth evolutionary semigroups*. In our case, ψ is the temperature field, $\psi = T(\mathbf{x}, t)$. ξ_1 and ξ_2 are random fields which represent the stochastic forcing. The dimension of the noise, ξ, can be smaller than the dimension of $(T, \tilde{\theta})^*$. $\mathbf{B}$ is a bounded linear operator which propagates the noise.

Equations (1.1) and (1.2) are quasilinear stochastic differential equations (sde's) with a particular upper triangular structure. We solve the nonlinear sde's iteratively, by linearizing the random component of Eqs. (1.1)–(1.2) about our current estimate of its expected value and using a quasilinear closure for the mean field. We define $\bar{\psi}(\mathbf{x}, t) \equiv E[\psi(\mathbf{x}, t)]$ and $\tilde{\psi} \equiv \psi - \bar{\psi}$. The linearized evolution equations for the fluctuations are

$$\partial_t \tilde{\psi} = \mathbf{L}(\bar{\theta}, t)\tilde{\psi} + \mathbf{G}(\bar{\theta}, t)\tilde{\theta} + \mathbf{B}\xi_1, \tag{1.3a}$$

$$\partial_t \tilde{\theta} = \mathbf{M}(t)\tilde{\theta} + \xi_2, \tag{1.3b}$$

where $\mathbf{G}(\bar{\theta}, t) \equiv (\partial \mathbf{L}(\theta, t)\bar{\psi}/\partial\theta)(\bar{\theta})$. We neglect the nonlinear terms in the fluctuating amplitudes for the fluctuation equation. For the inverse heat conductivity problem of Eq. (1.2), Eq. (1.3) reduces to

$$\partial_t \tilde{T} = \nabla \cdot \bar{\kappa}(\mathbf{x}, t)[\nabla \tilde{T} + \tilde{\theta}\, \nabla \bar{T}] + \xi_1, \tag{1.4a}$$

$$\partial_t \tilde{\theta} = \mu_2 \Delta \tilde{\theta} + \xi_2. \tag{1.4b}$$

Equations (1.3)–(1.4) describe the stochastic component of ψ and θ. When $\bar{T}$ and $\bar{\kappa}$ are given, Eq. (1.4) is a standard distributed estimation problem as discussed in [6, 25]. When $\bar{T}$ and $\bar{\kappa}$ are unknown, one approach is to minimize the residual sum of squares as a function of $\bar{T}$ and $\bar{\kappa}$ as well as $\tilde{T}$ and $\tilde{\theta}$ with a smoothness penalty function [1, 6]. This minimization is numerically difficult. In Section V, we describe a different approach, where we modify the standard iterated Kalman filter by updating $\bar{T}$ and $\bar{\theta}$ with the slowly varying part of the estimates of $\tilde{T}$ and $\tilde{\theta}$.

In Section II and the Appendices A and B, we review the Kalman filter-smoother. In Appendix A, we give variational formulations which may be useful for more advanced time discretizations. Appendix C describes the competing penalized least squares approach and gives a simple hybrid model.

In Section III, we discuss discrete approximations to the distributed estimation problem and regularizations of them. We describe a numerical implementation of the Kalman smoother for a one-dimensional heat equation with an unknown time dependent diffusion coefficient. In our augmented Kalman smoother formulation, a number of quantities need to be specified a priori: the covariances of ξ_1 and ξ_2, the operator, $\mathbf{M}(t)$, and the strength of the smoothness penalty. In Section IV, we examine the selection of these a priori terms. Appendix E relates the a priori smoothness of $\bar{T}$ and $\bar{\kappa}$ to $\mathbf{M}$ and the covariances of ξ and ξ_2. In Section V, we update the Kalman smoother by adding the slowly varying part of the filter estimates to $\bar{T}$ and $\bar{\theta}$. In Section VI, we discuss applications to plasma physics and to fluid dynamics.

II. INFINITE DIMENSIONAL FILTERING OF DISTRIBUTED SYSTEMS

We briefly summarize the similarities and differences for estimating infinite dimensional sytems. For more rigorous treatments, we recommend [6, 10, 25]. Appendix B contains a more explicit review of the finite dimensional, discrete time case. We rewrite Eq. (1.4) using the augmented state vector, $\mathbf{u}^* = (T, \tilde{\theta})^*$:

$$\partial_t \mathbf{u}(t) = \mathbf{F}(t)\, \mathbf{u}(t) + \mathbf{B}(t)\, \xi_u(t), \tag{2.1}$$

where $\xi_u^* = (\xi_1, \xi_2)^*$, with $E[\xi_u(t)\, \xi_u^*(s)] = \mathbf{Q}(t)\, \delta(t-s)$ and

$$\mathbf{F}_t \equiv \begin{pmatrix} \nabla \cdot \bar{\kappa}(\mathbf{x}, t)\nabla & \nabla \cdot \bar{T}(\mathbf{x}, t) \\ 0 & \mu_2 \Delta \end{pmatrix}, \tag{2.2}$$

where $\bar{T}(\mathbf{x}, t) \equiv \bar{\kappa}(\mathbf{x}, t) \cdot \nabla \bar{T}$ is the mean heat flux. To properly treat the infinite dimensional case, we assume that $\mathbf{Q}_u(t)$ is a positive definite trace class covariance and that $\mathbf{F}(\bar{\theta}(t), t)$ generates a *smooth evolutionary semigroup*. The noise propagator, $\mathbf{B}$, is useful because $\mathbf{B}\mathbf{Q}_u\mathbf{B}^*$ can be semidefinite.

In this section, we assume that the measurements are continuous in time and satisfy

$$\mathbf{y}(t) = \mathbf{H}(t)\, \mathbf{u}(t) + \varepsilon(t), \tag{2.3a}$$

where $\varepsilon(t)$ is the measurement error and its covariance is $E[\varepsilon(t)\, \varepsilon^*(s)] = \mathbf{R}(t)\, \delta(t-s)$. $\mathbf{H}$ is a *bounded* linear map. In our example, $\mathbf{y}(t)$ consists of the measured temperature values at m distinct locations, r_l, with sampling times t_i, where $i = 1, \ldots, t_f$ and $l = 1, \ldots, m$. The measurements are assumed to be spatially and temporally independent:

$$T_{l,i} = T(\mathbf{x}_l, t_i) + \varepsilon_{li}, \qquad \mathrm{cov}(\varepsilon_{l,i}, \varepsilon_{k,j}) = \sigma_k^2 \delta_{l,k} \delta_{i,j}. \tag{2.3b}$$

In a basis function representation, $\mathbf{u}(\mathbf{x}, t) = \sum_k u_k(t)\, g_k(\mathbf{x})$ and $\mathbf{H}$ has the representation: $H_{l,k} = g_k(r_l)$. Equation (2.3) allows for chordal cross-section measurements since $y_l(t) = H_l(t)\, \mathbf{u}(t)$ can be rewritten as $y_l(t) = \int H_l(\mathbf{x}, t)\, u(\mathbf{x}, t)\, d\mathbf{x}$.

If the measurement times are fast relative to the characteristic evolution time, the optimal filter will be well modeled by the continuous measurement time model. The continuous measurement time filter is the limit of the discrete measurement time filter with a measurement

covariance of $\mathbf{R}_t = \mathbf{R}(t)/\Delta$, where Δ is the time between measurements. In this limit, the Kalman filter simplifies because the measurement covariance dominates the filter covariance.

We suppress the time dependence of $\mathbf{H}(t)$, $\mathbf{B}(t)$, $\mathbf{Q}(t)$, and $\mathbf{R}(t)$ in this section for notational simplicity. The Kalman filter satisfies

$$\partial_t \hat{\mathbf{u}}(t) = \mathbf{F}_t \hat{\mathbf{u}}(t) + \mathbf{P}(t)\,\mathbf{H}^*\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{u}), \tag{2.4}$$

where $\hat{\mathbf{u}}$ denotes the estimate of $\mathbf{u}$. The covariance of $\hat{\mathbf{u}}$ evolves as

$$\partial_t \mathbf{P}(t) = \mathbf{F}_t \mathbf{P}(t) + \mathbf{P}(t)\,\mathbf{F}_t^* + \mathbf{Q}_\mathbf{B} - \mathbf{P}\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}\mathbf{P}, \tag{2.5}$$

where $\mathbf{Q}_\mathbf{B} \equiv \mathbf{B}\mathbf{Q}\mathbf{B}^*$. The initial conditions are $E[\mathbf{u}](t=0) = \mathbf{u}_0$ and $\mathbf{P}(t=0) = \mathbf{P}_0$. $\mathbf{P}^{-1}(t)$ evolves as well:

$$\partial_t \mathbf{P}^{-1}(t) = -\mathbf{P}^{-1}\mathbf{F}_t - \mathbf{F}_t^*\mathbf{P}^{-1} - \mathbf{P}^{-1}\mathbf{Q}_\mathbf{B}\mathbf{P}^{-1} + \mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}. \tag{2.6}$$

For the continuous time fixed interval smoother, we integrate backwards in time the stabilized equations

$$\begin{aligned}\partial_t \hat{\mathbf{u}}(t\,|\,t_f) &= \mathbf{F}_t \hat{\mathbf{u}}(t\,|\,t_f) + \mathbf{Q}_\mathbf{B}\mathbf{P}^{-1}(t) \\ &\quad \times [\hat{\mathbf{u}}(t\,|\,t_f) - \hat{\mathbf{u}}(t\,|\,t)], \qquad (2.7)\\ \partial_t \mathbf{P}(t\,|\,t_f) &= [\mathbf{F}_t + \mathbf{Q}_\mathbf{B}\mathbf{P}^{-1}(t)]\,\mathbf{P}(t\,|\,t_f) \\ &\quad + \mathbf{P}(t\,|\,t_f)[\mathbf{F}_t + \mathbf{Q}_\mathbf{B}\mathbf{P}^{-1}(t)]^* \\ &\quad - \mathbf{Q}_\mathbf{B} - \mathbf{P}(t\,|\,t_f)\,\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}\mathbf{P}(t\,|\,t_f), \qquad (2.8)\end{aligned}$$

with the final conditions $\hat{\mathbf{u}}(t_f\,|\,t_f) = \hat{\mathbf{u}}(t_f)$, $\mathbf{P}(t_f\,|\,t_f) = \mathbf{P}(t_f)$. The backwards time integration is normally ill-conditioned. To remedy this, the Bryson–Frazier (B–F) formulation [8, p. 395] of the continuous time smoother uses the auxiliary variables, $\mathbf{g}(t)$ and $\mathbf{G}(t)$, where

$$\hat{\mathbf{u}}(t\,|\,t_f) = \hat{\mathbf{u}}(t) - \mathbf{P}(t)\,\mathbf{g}(t), \tag{2.9}$$

$$\mathbf{P}(t\,|\,t_f) = \mathbf{P}(t) + \mathbf{P}(t)\,\mathbf{G}(t)\,\mathbf{P}(t). \tag{2.10}$$

The auxiliary variables satisfy

$$\begin{aligned}\partial_t \mathbf{g}(t) &= -[\mathbf{F}_t - \mathbf{P}(t)\,\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}]^*\,\mathbf{g} \\ &\quad + \mathbf{H}^*\mathbf{R}^{-1}[\dot{\mathbf{y}}_t - \mathbf{H}\hat{\mathbf{u}}(t)], \qquad (2.11)\\ \partial_t \mathbf{G}(t) &= -[\mathbf{F}_t - \mathbf{P}(t)\,\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}]^*\,\mathbf{G} \\ &\quad - \mathbf{G}[\mathbf{F}_t - \mathbf{P}(t)\,\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}] \\ &\quad + \mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}, \qquad (2.12)\end{aligned}$$

with the final conditions $\mathbf{g}(t_f) = 0$, $\mathbf{G}(t_f) = 0$. The B–F formulation replaces $-\mathbf{F}_t$ with $\mathbf{F}_t - \mathbf{P}(t)\,\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}$, and thereby tends to stabilize the backward time integration. Although the B–F formulation is common in finite dimensional control, *we are unaware of any previous usage in distributed systems.*

III. DISCRETE APPROXIMATIONS OF DISTRIBUTED SYSTEMS

To apply the Kalman smoother to distributed systems of partial differential equations, we represent the system using a truncated set of basis functions, and discretize the time evolution of the augmented system with a stable, consistent time advance. We assume that the augmented evolution equations are well posed, and constitute a strongly continuous semigroup. The Lax equivalence theorem [24] implies that any stable, consistent discretization converges to the continuous time limit. Similarly, the covariance evolution is discretized with a stable, consistent numerical scheme. We replace Eqs. (2.7)–(2.8) with the stabler Bryson–Frazier formulation, (2.9)–(2.12).

We now examine the inverse heat diffusivity problem in the one-dimensional periodic case. As discussed in Appendix C, the inverse problem is ill-conditioned, and we regularize the Kalman smoother by adding a small higher order spatial dissipation operator (hyperdiffusion). The hyperdiffusion damps the small scale oscillations and thereby aids in convergence. Furthermore, we remain within the standard framework of Kalman filters. Section IV addresses the selection of the size of the hyperdiffusivity coefficients. Thus we replace Eq. (1.4a) with

$$\begin{aligned}\partial_t \tilde{T}(\mathbf{x}, t) &= \partial_x[\bar{\kappa}(\mathbf{x}, t)\,\partial_x \tilde{T} + \bar{\Gamma}(\mathbf{x}, t)\tilde{\theta}] \\ &\quad - \mu_1 \partial_x^4 \tilde{T} + S(\mathbf{x}, t) + w(\mathbf{x}, t), \qquad (3.1)\end{aligned}$$

where $\bar{\Gamma}(\mathbf{x}, t) \equiv \bar{\kappa}(\mathbf{x}, t)\,\partial_x \bar{T}$ is the mean heat flux. $S(\mathbf{x}, t)$ is a known source term, and $w(\mathbf{x}, t)$ is random forcing. We assume periodic boundary conditions: $T(\pi, t) = T(-\pi, t)$ and $\partial_x T(\pi, t) = \partial_x T(-\pi, t)$. We expand both the mean quantities $\bar{T}(\mathbf{x}, t)$ and $\bar{\Gamma}(\mathbf{x}, t)$ and the fluctuating quantities $\tilde{T}(\mathbf{x}, t)$ and $\tilde{\theta}(\mathbf{x}, t)$ in truncated Fourier series,

$$\begin{aligned}T(\mathbf{x}, t) &= \sum_{k=-N_T}^{N_T} T_k(t)\, e^{ikx}, \\ \bar{\Gamma}(\mathbf{x}, t) &= \sum_{k=-N_T}^{N_T} \bar{\Gamma}_k(t)\, e^{ikx}, \qquad (3.2)\\ \tilde{\theta}(\mathbf{x}, t) &= \sum_{k=-N_T}^{N_T} \tilde{\theta}_k(t)\, e^{ikx},\end{aligned}$$

where $T_k(t)$, $\bar{\Gamma}_k(t)$, and $\theta_k(t)$ are complex with $T_{-k}(t) = T_k(t)^*$, $\bar{\Gamma}_{-k}(t) = \bar{\Gamma}_k(t)^*$, and $\tilde{\theta}_{-k}(t) = \tilde{\theta}_k^*(t)$. N_T is the truncation number of the Fourier series. The nonlinear transformation between $\kappa(\mathbf{x}, t)$ and $\theta(\mathbf{x}, t)$ is performed by collocation at the spatial points, $x_k = \pi k/N_t$, $k = -N_T \cdots N_T$, using the Fourier transform.

In Fourier space, the diffusion equation becomes

$$\frac{d\tilde{T}_k}{dt}(t) = -\sum_{k'=-N_T}^{N_T} kk'\bar{\kappa}_{k-k'}\tilde{T}_{k'} - \sum_{k'=-N_T}^{N_T} k\bar{\Gamma}_{k-k'}\tilde{\theta}_{k'} - \mu_1 k^4 \tilde{T}_k(t) + S_k(t) + \xi_k(t). \tag{3.3}$$

We assume that the stochastic forcing of the different modes is statistically independent and decays algebraically: $E[\xi_k(t)\,\xi_{k'}^*(t')] = \alpha_1 |k|^{-\beta_1}\delta_{k,k'}\delta(t-t')$. Our model for the stochastic evolution of θ_k is

$$\frac{d\tilde{\theta}_k}{dt}(t) = -\mu_2 |k|^2 \tilde{\theta}_k(t) + \xi_{2,k}(t), \tag{3.4}$$

where $E[\xi_{2,k}(t)\,\xi_{2,k'}^*(t')] = \alpha_2 |k|^{-\beta'}\delta_{k,k'}\delta(t-t')$. In Section IV, we discuss the selection of the free parameters such as α_1, α_2, β, and β'. For the time discretization, the stochastic forcing is scaled as the square root of the time step size: $\xi_k \equiv dw_k/dt$ with

$$E[w_k(t)\,w_{k'}^*(t')] = \alpha_1 |k|^{-\beta_1}\delta_{k,k'}\,dt,$$
$$E[w_{2,k}(t)\,w_{2,k'}^*(t')] = \alpha_2 |k|^{-\beta_2}\delta_{k,k'}\,dt. \tag{3.5}$$

Thus the Brownian increment, w_k, is large relative to the time step. As a result, the numerical accuracy of the finite difference approximation of the sde can be much worse than the accuracy of the same scheme on deterministic differential equations [21]. Recently, higher order difference schemes have been developed, and we use the Milstein implicit second-order weak Taylor scheme [21, p. 499]. This scheme is globally second-order accurate for computing weak solutions, but it is only first-order accurate for computing strong/pathwise solutions. Since we are primarily interested in estimating the mean quantities, $\bar{T}(\mathbf{x}, t)$ and $\bar{\kappa}(\mathbf{x}, t)$, weak convergence is adequate. The implicit Milstein discretization of Eqs. (3.3)–(3.4) is

$$\tilde{\theta}_k(t+dt) = \frac{1-(dt/2)\,\mu_2 |k|^2}{1+(dt/2)\,\mu_2 |k|^2}\tilde{\theta}_k(t) + \frac{w_{2,k}(t)}{1+(dt/2)\,\mu_2 |k|^2}, \tag{3.6}$$

$$\tilde{T}_k(t+dt) = \tilde{T}_k(t) + dt\sum_{k'=-N_T}^{N_T} kk'\bar{\kappa}_{k-k'} \times\left(\frac{\tilde{T}_{k'}(t)+\tilde{T}_{k'}(t+dt)}{2}\right) + dt\sum_{k'=-N_T}^{N_T} k\bar{\Gamma}_{k-k'}\left(\frac{\tilde{\theta}_{k'}(t)+\tilde{\theta}_{k'}(t+dt)}{2}\right) + S_k\left(t+\frac{dt}{2}\right)dt + w_k(t). \tag{3.7}$$

The term, $\sum_{k'=-N_T}^{N_T} kk'\bar{\kappa}_{k-k'}\tilde{T}_{k'}(t+dt)$, makes Eq. (3.7) intractable, and therefore we expand Eq. (3.7) assuming that the average diffusion, $\bar{\kappa}_0$, is large relative to the spatial variation of $\bar{\kappa}$. Using this semi-implicit approximation [23], the temperature time advance becomes

$$\tilde{T}_k(t+dt) = f_k\tilde{T}_k(t) - \sum_{\substack{k'=-N_T \\ k'\neq k}}^{N_T} g_{k,k'}\tilde{T}_{k'} + \sum_{k'=-N_T}^{N_T} h_{k,k'}\tilde{\theta}_{k'} + \sum_{k'=-N_T}^{N_T} c_{k,k'}w_{2,k'} + S_k\left(t+\frac{dt}{2d_k}\right)dt + w_k(t)d_k, \tag{3.8}$$

where $d_k \equiv 1 + (dt/2)(\kappa_0 |k|^2 + \mu_1 |k|^4)$ and

$$f_k \equiv \frac{1-(dt/2)(\kappa_0 |k|^2+\mu_1 |k|^4)}{1+(dt/2)(\kappa_0 |k|^2+\mu_1 |k|^4)},$$
$$g_{k,k'} \equiv \frac{kk'\bar{\kappa}_{k-k'}\,dt}{2d_k}(1+f_k),$$
$$h_{k,k'} \equiv \frac{k\bar{\Gamma}_{k-k'}\,dt}{2d_k}\left(1+\frac{1-(dt/2)\,\mu_2 |k'|^2}{1+(dt/2)\,\mu_2 |k'|^2}\right),$$
$$c_{k,k'} \equiv \frac{k\bar{\Gamma}_{k-k'}\,dt}{2d_k(1+(dt/2)\,\mu_2 |k'|^2)}.$$

Note that $w_{2,k}$ appears in Eq. (3.8) due to the difference scheme and this term is included in the Kalman smoother via the propagator matrix, **B**, in Section II and Appendix B. In Eq. (3.8), only the nonlinear evolution of the nonlinear terms is first order in time while the other terms are second-order accurate. The partially implicit difference scheme adds extra dissipation.

The observational data is transformed to mode space: $y_k(t_i) = \sum_{l=1}^m y_{l,i}e^{-ikx_l}$, where $|k| \leq m/2$. $y_k(t_i)$ is complex with $y_{-k}(t_i) = y_k^*(t_i)$. When m is even, we omit the *cosine* component of the last coefficient, $k = m/2$. In mode space, the measurement error matrix is $R_{k,k'} = \sum_{l=1}^m \sigma_l^2 \exp(i(k'-k)\,x_l)$, where $0 \leq k, k' \leq m/2$. The measurement evaluation matrix is $H_{k,k'} = \sum_{l=1}^m \exp(i(k-k')\,x_l)$, where $0 \leq k \leq m/2$ and $|k'| \leq N_T$. When $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma^2$ and the measurement locations are uniformly distributed, $x_l = (2\pi l/m)$, the measurement error matrix reduces to $R_{k,k'} = m\sigma^2\delta_{k-k'}$, and the evaluation matrix reduces to $H_{k,k'} = m$ when $k - k' = 0 \bmod m$, and zero otherwise.

Equations (3.6) and (3.8) constitute a discrete dynamical system. We assume that the measurement times occur on the time scale of a single time step advance of the diffusion

equation and use the discrete measurement time filter-smoother which is described in Appendix B. The time discretized equations have now been placed in the block tridiagonal form of Eq. (B11). Our preferred numerical method to solve Eq. (B11) is sequential quadratic programming [11, 29]. Alternatively, we can solve Eq. (B12) using conjugate gradient iterations on the diagonal subblocks.

IV. MODELS FOR STOCHASTIC PARAMETER EVOLUTION

Ideally, the model error covariances, $\mathbf{Q_B}$ and $\mathbf{Q}_\theta$, are given a priori or are estimated from the residuals. We denote by $\mathbf{C}(t, s)$ the covariance of $\mathbf{u}^* = (\tilde{T}, \tilde{\theta})^*$ in the *absence of measurements*. ($\mathbf{P}(t, s)$ is the covariance of the estimates, $\hat{\mathbf{u}}$ is *conditional on the measurements*.) In practice, we often have better knowledge of $\mathbf{C}$ than $\mathbf{Q}$. Appendix E expresses $\mathbf{C}$ in terms of $\mathbf{Q}$ for time independent evolution equations.

To explicitly evaluate an approximate covariance, we neglect the spatial temporal variation in $\bar{\kappa}(\mathbf{x}, t)$ and make a large k expansion of Eq. (E4). In this case, the evolution of T_k and θ_k decouple to leading order and the quasistationary approximate covariance satisfies

$$C_{\theta,k,k}(t) \simeq \frac{Q_{\theta,k}(t)}{2\mu_2 |k|^2}, \qquad C_{T,k,k}(t) \simeq \frac{Q_{T,k,k}(t)}{2(\bar{\kappa} |k|^2 + \mu_1 |k|^4)}. \tag{4.1}$$

These balance equations for the evolution of $C_{T,k}$ and $C_{\theta,k}$ can be used to define values of α_T, α_θ, β, and β', provided that $C_{T,k}$ and $C_{\theta,k}$ are given.

To choose μ_1, we neglect $\partial_t T_k$ in Eq. (3.3) and assume that $\kappa(\mathbf{x}, t)$ is spatially independent with value κ_0. In this case, we have $\hat{T}_k(t) = S_k(t)/(\kappa_0 k^2 + \mu_1 k^4)$, which has a bias error of $-\mu_1 S_k(t)/(\kappa_0^2 + \kappa_0 \mu_1 k^2)$ and a variance of $Q_{T,k}(t)/(\kappa_0 k^2 + \mu_1 k^4)$. Thus the total expected error for it is approximately equal to

$$\sum_k \left[\frac{Q_{T,k}(t)}{(\kappa_0 k^2 + \mu_1 k)^4} + \frac{|(\mu_1 S_k(t)|^2}{(\kappa_0^2 + \kappa_0 \mu_1 k^2)^2} \right]. \tag{4.2}$$

The size of the regularizing term, $\mu_1 \Delta\Delta$, is chosen to minimize this expected error.

The initial conditions are unknown, and so we include the initial conditions in the iteration. Our initial guess is $\kappa(\mathbf{x}, t=0) = \kappa_0$ and $T_k(t=0) = S_k/\kappa_0 k^2$ with $T_0(t=0)$ chosen by dimensional considerations: $T_0 = S_{\text{rms}}/\kappa_0$, where S_{rms} is the root mean square average of $S(\mathbf{x}, t)$. We also choose a smoothness prior for $\mathbf{P}(t=0|0)$: $\mathbf{P}(t=0|0) = (S_k^0/\kappa_0 k^2)^2$, where $S_k^0 \equiv \max\{S_k(t=0), S_{\text{rms}}\}$.

V. TEMPORAL AVERAGING TO UPDATE $\bar{\mathbf{T}}(\mathbf{x}, t)$ AND $\bar{\theta}(\mathbf{x}, t)$

In Eqs. (1.3)–(1.4) we separated the temperature into a mean field, $\bar{T}(\mathbf{x}, t)$, and a fluctuating field, $\tilde{T}(\mathbf{x}, t)$, which we estimate with the Kalman smoother. In practice, $\bar{T}(\mathbf{x}, t)$ and $\bar{\theta}(\mathbf{x}, t)$ are unknown, and we use an iterative procedure to estimate them. We let $\bar{T}^{(l)}(\mathbf{x}, t)$ and $\bar{\theta}^{(l)}(\mathbf{x}, t)$ denote the lth iterate of the mean field and $\hat{\bar{T}}^{(l)}(\mathbf{x}, t)$ and $\hat{\bar{\theta}}^{(l)}(\mathbf{x}, t)$ denote the Kalman smoother estimate of $\tilde{T}(\mathbf{x}, t)$ and $\tilde{\theta}(\mathbf{x}, t)$ when the filter is linearized about $\bar{T}^{(l)}(\mathbf{x}, t)$ and $\bar{\theta}^{(l)}(\mathbf{x}, t)$. The model misspecification on the lth iterate, $\bar{T}^{(l)}(\mathbf{x}, t) - \bar{T}(\mathbf{x}, t)$ and $\bar{\theta}^{(l)}(\mathbf{x}, t) - \bar{\theta}(\mathbf{x}, t)$, is modeled by the Kalman smoother as part of the stochastic forcing in the model. Thus the Kalman smoother implicitly includes and corrects for the possibility of model misspecification in its estimator. A simple iteration scheme would be

$$\bar{T}^{(l+1)}(\mathbf{x}, t) = \bar{T}^{(l)}(\mathbf{x}, t) + \mathrm{E}[\hat{\bar{T}}^{(l)}(\mathbf{x}, t)], \tag{5.1a}$$

$$\bar{\theta}^{(l+1)}(\mathbf{x}, t) = \bar{\theta}^{(l)}(\mathbf{x}, t) + \mathrm{E}[\hat{\bar{\theta}}^{(l)}(\mathbf{x}, t)]. \tag{5.1b}$$

However, $\bar{T}^{(l+1)}(\mathbf{x}, t)$, estimated from Eq. (5.1a), is not the solution of the heat equation with the diffusion coefficient, $\bar{\kappa}^{(l+1)} \equiv \exp[\bar{\theta}^{(l+1)}(\mathbf{x}, t)]$, from Eq. (5.1b). Therefore, we replace the naive $\bar{T}$ update with a heat flux averaging update.

In our heat flux averaging implementation of the mean temperature update, we define the $(l+1)$th estimate of the total heat flux by $\hat{\Gamma}^{(l+1)} \equiv [\bar{\kappa}^{(l)} + \hat{\tilde{\kappa}}^{(l)}] \nabla[\bar{T}^{(l)} + \hat{\tilde{T}}^{(l)}]$. The mean heat flux is the expectation of the nonlinear heat flux:

$$\bar{\Gamma}^{(l)}(\mathbf{x}, t) = \mathrm{E}[\hat{\Gamma}^{(l)}(\mathbf{x}, t)]. \tag{5.2}$$

We evolve the mean temperature with the mean heat flux:

$$\partial_t \bar{T}^{(l)}(\mathbf{x}, t) = \nabla \cdot \bar{\Gamma}^{(l)}(\mathbf{x}, t) - \mu_1 \Delta\Delta \bar{T}^{(l)} + S(\mathbf{x}, t). \tag{5.3}$$

Given $\bar{\Gamma}^{(l)}$ and $\bar{T}^{(l)}$ from Eqs. (5.2)–(5.3), we solve for $\bar{\kappa}^{(l)}(\mathbf{x}, t)$ using

$$\begin{aligned} \nabla \cdot \bar{\kappa}^{(l)} \nabla \bar{T}^{(l)}(\mathbf{x}, t) &= \nabla \bar{T}^{(l)} \cdot \nabla \bar{\kappa}^{(l)}(\mathbf{x}, t) + \bar{\kappa}^{(l)}(\mathbf{x}, t) \Delta \bar{T}^{(l)} \\ &= \bar{\Gamma}^{(l)}(\mathbf{x}, t). \end{aligned} \tag{5.4}$$

Since $\bar{T}^{(l)}(\mathbf{x}, t)$ is given, Eq. (5.4) is a linear hyperbolic equation for $\bar{\kappa}^{(l)}(\mathbf{x}, t)$. Equation (5.4) is well posed, provided that $\nabla \bar{T}^{(l)}$ does not vanish in the domain and that all the $\nabla \bar{T}^{(l)}$ characteristics intersect the boundary, where boundary data for $\bar{\kappa}$ is given.

We implement the expectation operator by averaging $\nabla \cdot \hat{\Gamma}^{(l+1)}(\mathbf{x}+\mathbf{x}', t)$ with a smoothing kernel,

$$\nabla \cdot \bar{\Gamma}^{(l+1)}(\mathbf{x}, t) = \int K(\mathbf{x}', t') \nabla \cdot \hat{\Gamma}^{(l+1)}(\mathbf{x}+\mathbf{x}', t')\, dt'\, d\mathbf{x}', \tag{5.5}$$

where $K(\mathbf{x}, t)$ is a smoothing kernel with a characteristic duration of three to five autocorrelation times. In the periodic heat equation of Section III, we smooth the kth harmonic of the heat flux, $\hat{\Gamma}_k$, with a kernel of duration $1/\bar{\kappa}\,|k|^2$. The longer the kernel window, the more that the high frequency oscillations are suppressed. We average the divergence of the heat flux instead of the heat flux because the divergence of the heat flux enters into Eq. (5.3).

In Sections II–V, we have described a detailed model for using the Kalman smoother to estimate $\kappa(\mathbf{x}, t)$, including stochastic dynamics. We couple the numerical discretization of Section III with the covariance model of Section IV. Section V completes the iteration cycle by updating $\bar{T}(\mathbf{x}, t)$ and $\bar{\kappa}(\mathbf{x}, t)$. Our heat flux averaging update eliminates the nonlinear discrepency, $\mathrm{E}[\nabla \cdot \hat{\kappa} \nabla \hat{T}] - \nabla \cdot \mathrm{E}[\hat{\kappa}] \nabla \mathrm{E}[\hat{T}]$ from the updating scheme. We now sketch several more realistic inverse problems for the augmented Kalman smoother.

VI. POTENTIAL APPLICATIONS

A. *Elliptic Identification Problems*

Coefficients for elliptic equations may be identified using the augmented Kalman smoother by adding a vanishingly small time derivative to the equations [17]. Similarly, when the temperature is time dependent and the coefficients are time dependent, we can regularize the diffusivity evolution equation with a vanishingly small time derivative of the parameters, $\varepsilon \partial_t \theta$.

B. *Anomalous Transport Coefficients in Fusion Plasmas*

In magnetic fusion, high temperature plasmas are confined in diffusive equilibrium by external magnetic fields. Due to the presence of low level turbulence, the transport of heat and particles is anomalously large. We consider the important problem of estimating anomalous transport coefficients in fusion plasmas.

To good approximation, fusion plasmas are determined by the electron and ion temperatures, $T_e(r, t)$, $T_i(r, t)$, and the electron and ion densities, $n_e(r, t)$ and $n_i(r, t)$. A simple, but realistic set of power balance equations are

$$\frac{3}{2}\frac{\partial(n_e T_e)}{\partial t}(r, t) = \frac{3}{2}\nabla \cdot (\chi_e n_e \nabla T - n_e T V) - n_e T_e \nabla \cdot V + S_e(r, T, n) + n_e \frac{T_i - T_e}{\tau_{ei}}, \tag{6.1}$$

$$\frac{\partial n_e}{\partial t} = \nabla \cdot (n_e V + D_e \nabla n_e), \tag{6.2}$$

$$\frac{3}{2}\frac{\partial(n_i T_i)}{\partial t}(r, t) = \frac{3}{2}\nabla \cdot (\chi_i n_i \nabla T_i - n_e T V) - n_i T_i \nabla \cdot V + S_i(r, T, n) + n_e \frac{T_i - T_e}{\tau_{ei}}, \tag{6.3}$$

$$n_e = Z n_i. \tag{6.4}$$

Equations (6.1)–(6.4) consist of three parabolic equations for n_e, T_e, and T_i with a constraint to determine n_i. There are four unknown transport coefficients: the electron and ion diffusivities, $\chi_e(r, t)$, $\chi_i(r, t)$, the electron density diffusivity, $D(r, t)$, and the pinch velocity, $V(r, t)$. The source terms, $S_e(r, T, n)$ and $S_i(r, T, n)$, and the coupling coefficient, τ_{ei}, depend only weakly on the unknown profiles. For convenience, we neglect these nonlinearities in the Kalman smoother. Z is the change number. Similar sets of transport equations have been used for thermal control of fusion plasmas [7, 14, 18, 20, 26, 28].

Presently, two algorithms are applied to estimate the transport coefficients. When the diffusion coefficients are specified using low order parametric models, least squares can be used to estimate the unknown free parameters [26, 28]. A second nonparametric approach is to estimate the measured profiles, $n_e(r, t)$, $T_e(r, t)$, and $T_i(r, t)$, by smoothing the raw data in space and time, and then inverting Eqs. (6.1)–(6.4) for the transport coefficients [14]. The smoothing of the raw measurements will bias the estimated transport coefficients to higher values.

The augmented Kalman smoother is ideally suited to estimate the state variables, $n_e(r, t)$, $T_e(r, t)$, and $T_i(r, t)$, and the time dependent transport coefficients, given point measurements of the state variables. Furthermore, the augmented Kalman smoother produces realistic covariance estimates for the diffusion coefficient, including the sizable model error in Eqs. (6.1)–(6.4).

C. *Turbulent Closures of Fluid Equations*

Another application is to estimate the effective equations for low level fluid turbulence. The simplest variant of this problem is to assume that the fluid velocity, $\mathbf{u}(\mathbf{r}, t)$, satisfies the Navier–Stokes equations:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = \nabla p + \nu \Delta \mathbf{u} + \xi,$$
$$\nabla \cdot \mathbf{u} = 0.$$

The stochastic forcing, $\xi(\mathbf{r}, t)$, models the fine scale fluid turbulence. The viscosity, $\nu(\mathbf{r}, t)$, is anomalously large to account for the macroscopic response of the fluid to the subgrid scale turbulence. To model the evolution of $\nu(\mathbf{r}, t)$, we

assume that $\theta(\mathbf{r}, t) = \ln(\nu(\mathbf{r}, t))$ is convected and stochastically forced: $\partial_t \theta + \mathbf{u} \cdot \nabla\theta = \xi$.

We are given continuous time measurements of velocity field on a coarse grid in space. We expand the Navier–Stokes equation in the set of eigenfunctions of the laminar flow linear stability problem and apply the augmented Kalman smoother to the problem. In imposing a higher order spatial dissipation for additional numerical stability, we note that many turbulence theories actually contain a hyperviscosity.

D. *Time Dependent Tomography of Plasma Instabilities*

X-ray tomography is used to analyze the $m = 1$ instability in tokamak plasmas [16]. Presently, each time slice is inverted separately to construct a time dependent image of the plasma emissivity, $\varepsilon(\mathbf{x}, t)$. The emissivity is advected by an unknown, time dependent velocity, $\mathbf{u}(\mathbf{x}, t)$, so we postulate a stochastic model:

$$\partial_t \varepsilon + \mathbf{u} \cdot \nabla\varepsilon = \kappa_0 \Delta\varepsilon + \xi_1. \tag{6.5}$$

To apply the extended Kalman smoother to estimate both $\varepsilon(\mathbf{x}, t)$ and $\mathbf{u}(\mathbf{x}, t)$, we augment the system with

$$\begin{aligned} \partial_t \mathbf{u} &= \nu \Delta \mathbf{u} + \xi_2, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned} \tag{6.6}$$

To simplify the model, we have dropped the forces from Eq. (6.6). Both κ_0 and μ can be given by neoclassical theory or can be estimated. The advantages of this approach are: (a) the velocity field is estimated from the time dependent emissivity; (b) by using many time slices of the tomography data simultaneously, coupled by Eqs. (6.5)–(6.6), we reduce the error in the estimate of $\varepsilon(\mathbf{x}, t)$.

VII. DISCUSSION

In this article, we applied the augmented Kalman smoother to inverse problems in partial differential equations. The stochastic forcing term in the evolution equations represents the model error. This stochastic forcing reduces the chances of filter divergence. Our approach is common for finite dimensional engineering problems, but we are unaware of any previous work which uses the augmented Kalman smoother to estimate unknown coefficients in distributed systems of partial differential equations.

The nonlinear coupling term, $\hat{\hat{\kappa}} \nabla\hat{\hat{T}}$, introduces non-Gaussianity into the system. In using the extended iterated Kalman smoother, we neglect the non-Gaussian part of the estimation problem. A rigorous alternative is to consider the full nonlinear filtering problem using the Zakai equation in function space. Unfortunately, this funcion space estimation approach is computationally infeasible.

Thus we use the extended Kalman smoother with two modifications to reduce the non-Gaussianity and numerical ill-conditioning. First, we add a regularizing term, $\mu_1 \partial_x^4 T(\mathbf{x}, t)$, to damp out higher order oscillations which are below the resolution threshold and thereby aid in convergence. Second, we update the Kalman smoother by adding a temporally smoothed version of $\hat{\bar{T}}$. The kernel smoother reduces the variance of $\bar{T}$ and corresponds to the probabilistic expectation. We update $\bar{T}$ by averaging the heat flux, $\hat{\kappa} \nabla\hat{\bar{T}}$, and then solving Eq. (5.3) for $\bar{T}$.

APPENDIX A: VARIATIONAL FORMULATIONS

For completeness, we state two variational formulations of the Kalman smoother. The variational formulations are useful from a Bayesian perspective and may yield higher order discretizations. For certain variational formulations [6, 23], we need to assumed that $\mathbf{F}(\bar{\theta}(t), t)$ maps the Hilbert space, $V \subset L^2(R^n)$, into its dual space, V', such that (i) the map from V to V' is bounded in the operator norm, (ii) the map is coercive: $(\mathbf{F}(\bar{\theta}(t), t)\, \mathbf{u}(t), \mathbf{u}(t))_{L^2} + c\, \|\mathbf{u}\|_{L^2}^2 \geqslant \alpha\, \|\mathbf{u}\|_V^2$. We assume that these properties are satisfied not only at $\theta = \bar{\theta}$, but for all θ in a compact, convex region, Ω, with $\bar{\theta} \in \Omega$. For parabolic problems, V is the space of functions with square integrable derivatives and the appropriate boundary conditions and $\|u\|_V^2 \equiv \int |\nabla u|^2 + u^2\, dx^n$. For the inverse heat conductivity problem of Eq. (1.4), the hypotheses are satisfied, provided that the diffusivity remains bounded from above and below.

The constrained least squares formulation of the Kalman smoother is

$$\begin{aligned} l(\mathbf{u}, \lambda, \xi \,|\, \mathbf{y}, \mathbf{F}, \mathbf{Q}) \equiv & \int_0^{t_f} dt [\mathbf{y}_t - \mathbf{H}_t \mathbf{u}_t]^* \mathbf{R}^{-1} [\mathbf{y}_t - \mathbf{H}_t \mathbf{u}_t] \\ & + \int_0^{t_f} dt \int d\mathbf{x}\, \xi_t^* \mathbf{Q}_t^{-1} \xi_t \\ & + \int_0^{t_f} dt \int d\mathbf{x}\, \lambda_t [\dot{\mathbf{u}}_t - \mathbf{F}_t \mathbf{u}_t - \mathbf{s}_t - \xi_t], \end{aligned} \tag{A1}$$

which yields to the backward time adjoint equation for the Lagrange multiplier:

$$\partial_t \lambda(t) = -\mathbf{F}_t^* \lambda - \mathbf{H}^* \mathbf{R}^{-1} [\mathbf{y}_t - \mathbf{H}_t \hat{\mathbf{u}}(t \,|\, t_f)]. \tag{A2}$$

When $\mathbf{Q}_{B,t}$ is also positive definite and trace class, the least square functional is

$$l(\mathbf{u}\,|\,\mathbf{y},\mathbf{F},\mathbf{Q}) \equiv \int_0^{t_f} dt \Big[(\mathbf{y}_t-\mathbf{H}_t\mathbf{u}_t)^*\,\mathbf{R}^{-1}(\mathbf{y}_t-\mathbf{H}_t\mathbf{u}_t) + \int d\mathbf{x}(\dot{\mathbf{u}}_t-\mathbf{F}_t\mathbf{u}_t-\mathbf{s}_t)^*\,\mathbf{Q}_{\mathbf{B},t}^{-1}(\dot{\mathbf{u}}_t-\mathbf{F}_t\mathbf{u}_t-\mathbf{s}_t)\Big]. \tag{A3}$$

Equation (A3) is the continuous time analogy of Eq. (B10).

APPENDIX B: DISCRETE TIME KALMAN FILTER-SMOOTHERS

We review discrete Kalman filters and smoothers; more extensive presentations are given in the textbooks by Jazwinski [19] and Anderson and Moore [2].

A. *Discrete Kalman Filter-Smoothers*

We consider the discrete linear state space model,

$$\mathbf{u}_{i+1} = \Phi(i+1,i)\,\mathbf{u}_i + \mathbf{B}_i\mathbf{w}_i + \mathbf{s}_i, \tag{B1}$$

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{u}_i + \varepsilon_i, \tag{B2}$$

where $\mathbf{u}_i$ is the state vector of dimension N, $\mathbf{y}_i$ is the measurement vector of dimension m, and $\mathbf{s}_i$ is the known source vector of dimension N. $\Phi(j,i)$ is the $N\times N$ nonsingular deterministic part of the map from time i to time j. The system noise, $\mathbf{w}_i$, is assumed be an r-dimensional white Gaussian processes with covariance $\mathbf{Q}_i$. The measurement noise, ε_i, is an m-dimensional white Gaussian sequence with nondegenerate covariance $\mathbf{R}_i$. The $m\times N$ measurement evaluation matrix, $\mathbf{H}_i$, maps the state vector, $\mathbf{u}_i$, onto the deterministic part of the measurements. To simplify the notation, we define the $N\times N$ matrices, $\mathbf{F}_i \equiv \Phi(i+1,i)$, $\mathbf{Q}_{\mathbf{B},i} \equiv \mathbf{B}_i\mathbf{Q}_i\mathbf{B}_i^*$ and $\mathbf{J}_i \equiv \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i$.

The standard Kalman filter estimates the state vector, $\hat{\mathbf{u}}(i\,|\,j)$, at time i given the measurements, $\mathbf{y}_1, \ldots, \mathbf{y}_j$ up to time j by the time evolution update:

$$\hat{\mathbf{u}}(i+1\,|\,i) = \mathbf{F}_i\hat{\mathbf{u}}(i\,|\,i) + \mathbf{s}_i. \tag{B3}$$

The covariance, $\mathbf{P}(i\,|\,j)$, of the estimate, $\hat{\mathbf{u}}(i\,|\,j)$, evolves as

$$\mathbf{P}(i+1\,|\,i) = \mathbf{F}_i\mathbf{P}(i\,|\,i)\,\mathbf{F}_i^* + \mathbf{Q}_{\mathbf{B},i}. \tag{B4}$$

We assume that $\hat{\mathbf{u}}(0\,|\,0)$ are given. The measurement update is

$$\hat{\mathbf{u}}(i\,|\,i) = \hat{\mathbf{u}}(i\,|\,i-1) + \mathbf{K}_i(\mathbf{y}_i - \mathbf{H}_i\hat{\mathbf{u}}(i\,|\,i-1)), \tag{B5}$$

$$\mathbf{P}(i\,|\,i)^{-1} = \mathbf{P}(i\,|\,i-1)^{-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i, \tag{B6}$$

where $\mathbf{K}_i$ is the $N\times m$ Kalman gain:

$$\mathbf{K}_i = [\mathbf{P}(i\,|\,i-1)^{-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i]^{-1}\,\mathbf{H}_i^*\mathbf{R}_i^{-1} = \mathbf{P}(i\,|\,i)\,\mathbf{H}_i^*\mathbf{R}_i^{-1}. \tag{B7}$$

Following the Rauch–Tung–Striebel (RTS) formulation [9, Chap. 13.2], we divide the Kalman smoother into a forward Kalman filter followed by a backward smoother correction:

$$\hat{\mathbf{u}}(i\,|\,N_f) = \hat{\mathbf{u}}(i\,|\,i) + \mathbf{P}(i\,|\,i)\,\mathbf{F}_i^*\mathbf{P}^{-1}(i+1\,|\,i)[\hat{\mathbf{u}}(i+1\,|\,N_f) - \hat{\mathbf{u}}(i+1\,|\,i)], \tag{B8}$$

$$\mathbf{P}(i\,|\,N_f) = \mathbf{P}(i\,|\,i) + \mathbf{P}(i\,|\,i)\,\mathbf{F}_i^*\mathbf{P}^{-1}(i+1\,|\,i) \times [\mathbf{P}(i+1\,|\,N_f) - \mathbf{P}(i+1\,|\,i)] \times \mathbf{P}^{-1}(i+1\,|\,i)\,\mathbf{F}_i\mathbf{P}(i\,|\,i), \tag{B9}$$

where $|N_f)$ denotes the values conditional on the measurements up to time N_f.

B. *Least Squares Formulation of the Kalman Smoother*

The fixed interval Kalman smoother is the least squares and maximum likelihood estimator of $\mathbf{u}_i$, given the measurements, $\mathbf{y}_1 \cdots \mathbf{y}_{N_f}$, where the final measurement time is N_f. The Kalman smoother is a statistical estimation problem of very large dimension, NN_f.

The generalized least squares funcional, $l(\mathbf{u}\,|\,\mathbf{y},\mathbf{F},\mathbf{Q})$, for the Kalman smoother is

$$l(\mathbf{u}\,|\,\mathbf{y},\mathbf{F},\mathbf{Q}) \equiv \sum_{i=1}^{N_f} (\mathbf{y}_i-\mathbf{H}_i\mathbf{u}_i)^*\,\mathbf{R}^{-1}(\mathbf{y}_i-\mathbf{H}_i\mathbf{u}_i) + \sum_{i=1}^{N_f} (\mathbf{u}_{i+1}-\mathbf{F}_i\mathbf{u}_i-\mathbf{s}_i)^* \times \mathbf{Q}_{\mathbf{B},i}^{-1}(\mathbf{u}_{i+1}-\mathbf{F}_i\mathbf{u}_i-\mathbf{s}_i), \tag{B10}$$

where $\mathbf{Q}_{\mathbf{B},i}^{-1} \equiv (\mathbf{B}_i\mathbf{Q}_i\mathbf{B}_i^*)^{-1}$. The least squares functional has a Bayesian interpretation: $\sum_{i=1}^{N_f}(\mathbf{u}_{i+1}-\mathbf{F}_i\mathbf{u}_i-\mathbf{s}_i)^*\,\mathbf{Q}_{\mathbf{B},i}^{-1}(\mathbf{u}_{i+1}-\mathbf{F}_i\mathbf{u}_i-\mathbf{s}_i)$ is the argument of the a priori probability density of the stochastic system, and $\sum_{i=1}^{N_f}(\mathbf{y}_i-\mathbf{H}_i\mathbf{u}_i)^*\,\mathbf{R}^{-1}(\mathbf{y}_i-\mathbf{H}_i\mathbf{u}_i)$ represents the conditional probability density. Thus $l(\mathbf{u}\,|\,\mathbf{y})$ is the argument of the a posteriori probability density. In the Bayesian interpretation, $\hat{\mathbf{u}}_i$, is the weighted sum of the a priori estimate and a new independent estimate from the measurements, $\mathbf{y}_1 \cdots \mathbf{y}_{N_f}$.

In dynamic programming, $\mathbf{u}$ is estimated by minimizing $l(\mathbf{u}\,|\,\mathbf{y},\Phi,\mathbf{Q})$ directly. Differentiating $l(\mathbf{u}\,|\,\mathbf{y},\mathbf{F},\mathbf{Q})$ with respect to $\mathbf{u}_i$ yields the estimation equations for $\hat{\mathbf{u}}(i\,|\,N_f)$:

$$-\mathbf{Q}_{\mathbf{B},i-1}^{-1}\mathbf{F}_{i-1}\hat{\mathbf{u}}(i-1\,|\,N_f) + [\mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i}^{-1}\mathbf{F}_i + \mathbf{Q}_{\mathbf{B},i-1}^{-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i]\,\hat{\mathbf{u}}(i\,|\,N_f) - \mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i-1}^{-1}\hat{\mathbf{u}}(i+1\,|\,N_f) = -\mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i}^{-1}\mathbf{s}_i + \mathbf{Q}_{\mathbf{B},i-1}^{-1}\mathbf{s}_{i-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{y}_i. \tag{B11a}$$

We decompose $\mathbf{u}_i$ into a deterministic component, $\bar{\mathbf{u}}_i$,, and the stochastic component, $\tilde{\mathbf{u}}$. Equation (B11) reduces to $\bar{\mathbf{u}}_{i+1} - \mathbf{F}_i \bar{\mathbf{u}}_i - \mathbf{s}_i$, and

$$
\begin{aligned}
&-\mathbf{Q}_{\mathbf{B},i-1}^{-1}\mathbf{F}_{i-1}\tilde{\mathbf{u}}(i-1\,|\,N_f) \\
&\quad + (\mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i}^{-1}\mathbf{F}_i + \mathbf{Q}_{\mathbf{B},i-1}^{-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i) \\
&\quad \times \tilde{\mathbf{u}}(i\,|\,N_f) - \mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i-1}^{-1}\tilde{\mathbf{u}}(i+1\,|\,N_f) \\
&\quad = \mathbf{H}_i^*\mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i\bar{\mathbf{u}}). \qquad \text{(B11b)}
\end{aligned}
$$

Equation (B11b) is a coupled set of NN_f equations with a symmetric, block tridiagonal structure. The inverse of the a posteriori covariance matrix, Σ, is defined by

$$
\begin{aligned}
&\Sigma_{i,i}^{-1} \equiv (\mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i}^{-1}\mathbf{F}_i + \mathbf{Q}_{\mathbf{B},i-1}^{-1} + \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i), \\
&\Sigma_{i,i-1}^{-1} \equiv -\mathbf{Q}_{\mathbf{B},i-1}^{-1}\mathbf{F}_{i-1},\ \Sigma_{i,i+1}^{-1} \equiv -\mathbf{F}_i^*\mathbf{Q}_{\mathbf{B},i}^{-1}.
\end{aligned} \qquad \text{(B12)}
$$

The block tribanded structure enables the Kalman smoother equations to be solved using a forward sweep and then a backward sweep of the equations. *The forward–backward sweeps of the RTS smoother correspond to the standard algorithm for solving block tribanded matrices.*

If the estimated covariance, $\mathbf{P}(t)$, is small relative to $\mathbf{R}$ (more precisely, if $\mathbf{P}(i\,|\,i-1)^{-1} \gg \mathbf{H}_i^*\mathbf{R}_i^{-1}\mathbf{H}_i$ (in the discrete case)), the continuous measurement time filter may be used. If the measurement times are slower than the characteristic evolution time, the time map, $\Phi(i+1, i)$, between the ith and $(i+1)$th measurements is evaluated by the composition of many time steps.

APPENDIX C: COMBINING THE KALMAN SMOOTHER WITH SMOOTHNESS PENALTY FUNCTIONS

A number of researchers have examined parameter estimation when $\kappa(\mathbf{x})$ is an unknown, deterministic function of space ($\mathbf{Q}_T \equiv 0$ and $\mathbf{Q}_\theta \equiv 0$). To regularize the infinite dimensional minimization, either $\kappa(\mathbf{x})$ is required to be in a compact convex subset of smooth functions where $0 < \kappa_L \leqslant \kappa(\mathbf{x}) \leqslant \kappa_U$ [3, 4], or a smoothness penalty function is added. Thus the penalized least squares problem is to minimize

$$
l(\mathbf{u}, \bar{\kappa}\,|\,\mathbf{y}) \equiv \int_0^{t_f} dt(\mathbf{y}_t - \mathbf{H}\mathbf{u}_t)^* \mathbf{R}^{-1}(\mathbf{y}_t - \mathbf{H}\mathbf{u}_t) + \delta \|\bar{\theta}\|^2, \qquad \text{(C1)}
$$

over initial conditions and over $\bar{\theta} = \ln(\kappa(\mathbf{x}))$ in a convex set with $0 < \kappa_L \leqslant \kappa(\mathbf{x}) \leqslant \kappa_U$ and $\|\theta\|_s^2$ is a smoothness penalty. Equation (C1) can be interpreted as the Bayesian a posteriori likelihood given a smoothness prior. Alternatively, Eq. (C1) can be interpreted as an approximate likelihood where the penalty term induces a bias error in order to reduce the variance of the estimate by damping the high frequency modes.

The convergence of discretized approximations is proven in [4, 6, 20]. Kravaris and Seinfeld [22] give general conditions when penalized least squares functionals have an unique minimum. Recently, Aihara [1] has applied this analysis to the case of stochastic forcing in the temperature equation with $\kappa(\mathbf{x})$ an unknown, deterministic function of space.

Banks and Lamm [4] consider the case of deterministic dynamics with time dependent diffusivities. They show that the temporally and spatially discretized parameter estimates converge to the correct parameter estimates of the continuous problem as the discretization becomes finer. Alternately, the results of Aihara can be generalized to time dependent diffusivities. To ensure the existence of a unique minimizer of the loss functional, $\kappa(\mathbf{x}, t)$ needs to be restricted to a compact set or a smoothness penalty in both time and space, such as

$$
\|\bar{\theta}\|_s^2 = \int_0^{t_f} dt \int d\mathbf{x}\,[\,|\partial_t^2\bar{\theta}(\mathbf{x}, t)|^2 + |\Delta\bar{\theta}(\mathbf{x}, t)|^2], \qquad \text{(C2)}
$$

must be applied. Similarly, our augmented Kalman smoother formulation adds a penalty term of the form

$$
\int d\mathbf{x}(\vec{\bar{\theta}}_t - \mathbf{M}\tilde{\theta})^* \mathbf{Q}_\theta^{-1}(\vec{\bar{\theta}}_t - \mathbf{M}\tilde{\theta}). \qquad \text{(C3)}
$$

As $\mathbf{Q}_\theta$ increases, the importance of the most recent measurements in determining the instantaneous value of $\bar{\theta}$ increases. Thus the stochastic forcing, $\xi_{\bar{\theta}}$, reduces the tendency of the smoother to diverge.

Both Eq. (C2) and Eq. (C3) penalize against rapid, unphysical variations in the parameter estimates by adding an a priori smoothness density to the likelihood function. Our augmented Kalman filter approach has several advantages. For given $\bar{T}$ and $\bar{\kappa}$, the Kalman smoother is the exact minimizer of the likelihood function with respect to $\tilde{T}$ and $\tilde{\theta}$. In contrast, the standard formulation of the penalized likelihood approach of Eqs. (C1)–(C2) requires many iterations of a steepest descent algorithm to minimize with respect to κ [1, 19]. Furtermore, the computation of the gradient of Eq. (C1) with respect to κ is extremely complicated [1]. The computational simplicity of our formulation arises because $\tilde{\theta}$ and $\tilde{T}$ enter quadratically in our likelihood functional.

The two approaches have slightly different functions. The smoothness penalty damps all rapid variation while the prior of Eq. (C3) damps large innovations, $\mathbf{w}$. The two approaches may be combined by adding a prior distribution of $\|\tilde{T}\|^2 + \|\tilde{\theta}\|^2$ to the standard variational formulation of Eq. (A1). As an alternative to the smoothness penalty, we

prefer to add an artificial hyperdiffusivity, $\mu_4 \Delta^4 \bar{T}$, to the evolution equations. The hyperdiffusivity damps the high frequency modes, as does the smoothness penalty. The addition of a hyperdiffusivity is easy to implement numerically and lets us remain in the Kalman smoother framework. The artificial hyperdiffusivity induces a systematic bias error in exchange for reducing th ill-conditioning. Equation (4.2) quantifies this bias versus variance trade-off.

APPENDIX D: NEARLY BLOCK DIAGONAL IMPLEMENTATION OF THE KALMAN SMOOTHER

In [27], we showed that the computational cost and numerical ill-conditioning of the Kalman filter/smoother can be noticeably reduced when the stochastic system, including $\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}$ and $\mathbf{Q}$, is nearly block diagonal (NBD), and first order approximations are sufficient. In our case, the computational effort decreases by a factor of m. To apply the NBD approximation to Eqs. (3.6)–(3.8), we assume that the conductivity is nearly constant and expand $\theta(r, t) = \theta_0 + \theta_1(r, t)$.

Spatial aliasing strongly couples the Fourier modes with $k - k' = 0 \bmod m$. To preserve the block diagonal structure for $\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}$, we reorder the indices in the state vector, $\mathbf{u}$. The kth block of the reordered $\mathbf{u}$ is T_k, θ_k, $T_{k+m} \cdots$. We expand $\mathbf{H}^*\mathbf{R}^{-1}\mathbf{H}$ about a block diagonal matrix and include the corrections due to unequal variances ($\sigma_k^2 \neq \sigma_{k'}^2$), and nonuniformly distributed measurement locations ($x_l \neq 2\pi l/m$). The reordered and expanded system is then solved using the first-order approximation to the iterated Kalman smoother.

APPENDIX E: MODELS FOR STOCHASTIC PARAMETER EVOLUTION

Ideally, the model error covariances, $\mathbf{Q_B}$ and $\mathbf{Q}_\theta$, are given a priori or are estimated from the residuals. $\mathbf{P}(t, s)$ is the covariance of the estimates, $\hat{\mathbf{u}}$ *is conditional on the measurements.* We examine $\mathbf{C}(t, s)$, the covariance of $\mathbf{u}^* = (\tilde{T}, \tilde{\theta})^*$, in the *absence of measurements.* In practice, we often have a better understanding of the size and autocorrelation time of $\mathbf{C}$ then of the stochastic model parameters, $\mathbf{Q}$ and $\mathbf{M}$.

We now determine $\mathbf{C}(s, t)$ in terms of $\mathbf{F}$ and $\mathbf{Q}$ when $\mathbf{F}$ and $\mathbf{Q}$ are time independent. The covariance evolves $\partial_t \mathbf{C}(s, t) = \mathbf{FC} + \mathbf{CF}^* + \mathbf{Q}$. We assume that all of the eigenvalues of $\mathbf{F}$ have negative real parts, and therefore $\mathbf{u}$ has a stationary covariance: $\mathbf{C}(s, t) = \bar{\mathbf{C}}(s - t)$, where

$$\mathbf{C}(t, t) = \bar{\mathbf{C}}(0) \equiv \int_0^\infty \exp(t'\mathbf{F})\, \mathbf{Q} \exp(t'\mathbf{F}^*)\, dt', \tag{E1}$$

$$\mathbf{C}(s, t) = \bar{\mathbf{C}}(s - t) = \bar{\mathbf{C}}(0) \exp((t - s)\, \mathbf{F}^*) \quad \text{for} \quad s \leqslant t. \tag{E2}$$

If we use a basis where $\mathbf{F}$ is diagonal, Eqs. (E1)–(E2) reduce to

$$C_{i,j}(s, t) = \frac{e^{(t-s)\lambda_j} Q_{i,j}}{(\lambda_i + \lambda_j)} \quad \text{for} \quad s \leqslant t. \tag{E3}$$

Thus if $\mathbf{Q}$ and $\mathbf{F}$ are simultaneously diagonalizable, then the different eigenvector directions are uncoupled and evolve independently. We will generally choose $\mathbf{Q}$ to be diagonal in the basis of $\mathbf{F}$ eigenfunctions.

When the autocorrelation time of the fluctuating $\tilde{T}$ and $\tilde{\theta}$ fields is short in comparison with the characteristic time scale for $\bar{T}$ and $\bar{\kappa}$ evolution, we can treat the fluctuations as quasistationary. Since Eq. (1.4) has a block upper triangular structure, the eigenfunctions of the frozen time version of Eq. (1.4) are of two types. The first class are eigenfunctions of the form: $(T_j^1(\mathbf{x}), \theta_j^1(\mathbf{x}))^* = (T_j^1(\mathbf{x}), 0)^*$, where $T_j^1(\mathbf{x})$ is an eigenfunction of $\nabla \cdot \bar{\kappa}(\mathbf{x}, t) \nabla T_j^1(\mathbf{x}) = \lambda_j^1 T_j^1(\mathbf{x})$. The second class is based on the eigenfunctions of Eq. (1.4b), $(T_j^2(\mathbf{x}), \theta_j^2(\mathbf{x}))^*$, where $\mu_2 \Delta \theta_j^2(\mathbf{x}) = \lambda_j^2 \theta_j^2(\mathbf{x})$. Thus the eigenfunctions of the perturbed temperature operator, $\nabla \cdot \bar{\kappa}(\mathbf{x}, t) \nabla$, do not couple to $\tilde{\theta}$, but the eigenfunctions of $\mu_2 \nabla$ couple to $\tilde{T}$. Correspondingly, we divide the covariance, $\mathbf{C}$, into

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{T,T}^1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{C}_{T,T}^2 & \mathbf{C}_{T,\theta}^2 \\ 0 & \mathbf{C}_{\theta,\theta}^2 \end{pmatrix}. \tag{E4}$$

As the mode number, k, increases, the crossterm, $\mathbf{C}_{T,\theta}$, decreases relative to the diagonal terms as $1/k$. The asymptotic decoupling of the $\tilde{T}$ equation and the $\tilde{\theta}$ equation implied that the variation of $\tilde{\theta}$ becomes increasingly difficult to identify as the wavelength of $\tilde{\theta}$ variation decreases.

Thus we use the extended Kalman smoother with two modifications to reduce the non-Gaussianity and numerial ill-conditioning. First, we add a regularizing term, $\mu_1 \partial_x^4 T(\mathbf{x}, t)$, to damp out higher order oscillations which are below the resolution threshold and thereby aid in convergence. Second, we update the Kalman smoother by adding a temporally smoother version of $\hat{T}$. The kernel smoother reduces the variance of $\bar{T}$ and corresponds to the probabilistic expectation. We update $\bar{T}$ by aveaging the heat flux, $\hat{\kappa} \nabla \hat{T}$, and then solving Eq. (5.3) for $\bar{T}$.

ACKNOWLEDGMENTS

The author has benefited from discussions with G. Papanicolaou and R. Kohn. The referee's comments are gratefully acknowledged. This work was performed under U.S. Department of Energy Grant No. DE-FG02-86ER-53223.

REFERENCES

1. S. I. Aihara, *SIAM J. Control Opt.* **30**, 745 (1992).
2. B. D. O. Anderson and J. B. Moore, *Optimal Filtering* (Prentice-Hall, Englewood Cliffs, NJ, 1979).
3. A. V. Balakrishnan, "Parameter Estimation in Stochastic Differential

Systems: Theory and Applications," in *Developments in Statistics*, edited by P. R. Krishnaih (Academic Press, New York, 1978).

4. H. T. Banks and P. D. Lamm, *IEEE Trans. Automat. Control* **30**, 386 (1985).
5. J. V. Beck, B. Blackwell, and C. R. St. Clair, *Inverse Heat Conduction—Ill-posed Problems* (Wiley-Interscience, New York, 1985).
6. A. Bensoussan, *Filtrage Optimal des Systèmes Linéaires* (Dunod, Paris, 1971).
7. J. Blum, *Numerical Simulation and Optimal Control in Plasma Physics* (Wiley/Gauthiers-Villars, Chichester/New York, 1989).
8. A. E. Bryson, Jr. and Y. C. Ho, *Applied Optimal Control* (Blaisdell, New York, 1969).
9. S. E. Cohn and D. F. Parrish, *Mon. Weather Rev.* **120**, 1757 (1991).
10. R. Curtain, *SIAM Rev.* **17**, 395 (1975).
11. A. Fiacco and G. McCormick, *Nonlinear Programming: Sequentially Unconstrained Minimization Techniques* (Wiley, New York, 1968).
12. M. Ghil, *Dynamics of Oceans and Atmospheres* **13**, 171 (1989).
13. R. Glowinski, *J. Comput. Phys.* **103**, 189 (1992).
14. R. J. Goldston, D. C. McCune, H. H. Towner, *et al.*, *J. Comput. Phys.* **43**, 61 (1981).
15. G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experimental Design and Data Analysis* (Academic Press, New York, 1977).
16. R. S. Granetz and P. Smeulders, *Nuclear Fusion* **28**, 457 (1988).
17. K.-H. Hoffman and J. Spreckels, *Numer. Funct. Anal. Optim.* **7**, 157 (1984).
18. S. C. Jardin, N. Pomprey, and J. Delucia, *J. Comput. Phys.* **66**, 481 (1986).
19. A. H. Jazwinski, *Stochastic Processes and Filtering Theory* (Academic Press, New York, 1970).
20. C. E. Kessel and M. A. Firestone, *IEEE Trans. Plasma Sci.* **19**, 29 (1991).
21. P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992).
22. C. Kravaris and J. H. Seinfeld, *SIAM J. Control Opt.* **23**, 217 (1985).
23. M. Kress and K. S. Riedel, *J. Comput. Phys.* **83**, 237 (1989).
24. P. D. Lax and R. D. Richtmyer, *Commun. Pure Appl. Math* **9**, 267 (1956).
25. S. Omatu and J. H. Seinfeld, *Distributed Parameter Systems, Theory and Applications* (Oxford Sci., Oxford, 1989).
26. K. S. Riedel, E. Eberhagen, O. Gruber, *et al.*, *Nucl. Fusion* **28**, 1509 (1988).
27. K. S. Riedel, *Automatica* **29**, 779 (1993).
28. P. K. C. Wang, Identificaton problems in plasma physics, in *Lecture Notes in Control Inform. Sci.*, Vol. 1 (Springer-Verlag, Berlin, 1978), p. 424.
29. S. J. Wright, Tech. Rep. MCS-P229-0491, Argonne Natl. Labs. Argonne, IL, 1991 (unpublished).