Linear Theory for Filtering Nonlinear Multiscale Systems with Model Error

Tyrus Berry

Department of Mathematics
The Pennsylvania State University

Collaborator: John Harlim
The Pennsylvania State University

January 21, 2014
Problem Statement

**Linear Example:** Consider a two-dimensional system of SDEs,

\[ dx = (a_{11}x + a_{12}y) \, dt + \sigma_x dW_x, \]
\[ dy = \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y, \]
**Linear Example:** Consider a two-dimensional system of SDEs,

\[
     dx = (a_{11}x + a_{12}y) \, dt + \sigma_x \, dW_x, \\
     dy = \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} \, dW_y,
\]

Given noisy observations \( z_m = x(t_m) + \epsilon_m, \ \epsilon_m \sim \mathcal{N}(0, R) \) the filtering problem is to estimate the posterior density \( p(x(t_m), y(t_m) \mid z_1, z_2, ..., z_m) \).
Problem Statement

**Linear Example:** Consider a two-dimensional system of SDEs,

\[ dx = (a_{11}x + a_{12}y) \, dt + \sigma_x dW_x, \]
\[ dy = \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y, \]

- Given noisy observations \( z_m = x(t_m) + \epsilon_m, \; \epsilon_m \sim \mathcal{N}(0, R) \)
  the filtering problem is to estimate the posterior density \( p(x(t_m), y(t_m) \mid z_1, z_2, \ldots, z_m). \)

- Given the full model above, the Kalman filter gives the optimal posterior estimate, in the sense of minimum variance.
Problem Statement

**Linear Example:** Consider a two-dimensional system of SDEs,

\[
\begin{align*}
    dx &= (a_{11}x + a_{12}y) \, dt + \sigma_x dW_x, \\
    dy &= \frac{1}{\epsilon}(a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y,
\end{align*}
\]

Given noisy observations \( z_m = x(t_m) + \epsilon_m, \quad \epsilon_m \sim \mathcal{N}(0, R) \) the filtering problem is to estimate the posterior density

\[ p(x(t_m), y(t_m) \mid z_1, z_2, \ldots, z_m). \]

Given the full model above, the Kalman filter gives the optimal posterior estimate, in the sense of minimum variance.

Now assume the fast variable is unknown and we only have an averaged model

\[ dX = \alpha X dt + \sigma dW_x \] for the slow variable.
Problem Statement

**Linear Example:** Consider a two-dimensional system of SDEs,

\[
\begin{align*}
    dx &= (a_{11}x + a_{12}y) \, dt + \sigma_x \, dW_x, \\
    dy &= \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} \, dW_y,
\end{align*}
\]

- Given noisy observations \( z_m = x(t_m) + \epsilon_m, \ \epsilon_m \sim \mathcal{N}(0, R) \)
  the filtering problem is to estimate the posterior density 
  \( p(x(t_m), y(t_m) | z_1, z_2, ..., z_m) ). \)

- Given the full model above, the Kalman filter gives the optimal posterior estimate, in the sense of minimum variance.

- Now assume the fast variable is unknown and we only have an averaged model \( dX = \alpha X \, dt + \sigma \, dW_x \) for the slow variable.

- Can we recover \( p(x(t_m) | z_1, z_2, ..., z_m) \) as accurately as the full model?
Motivation for the Reduced Model

**Linear Example:** Consider a two-dimensional system of SDEs,

\[
\begin{align*}
    dx &= (a_{11}x + a_{12}y) \, dt + \sigma_x \, dW_x, \\
    dy &= \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} \, dW_y,
\end{align*}
\]

Want best parameters in reduced model: \(dX = \alpha X dt + \sigma dW_x\)
**Linear Example:** Consider a two-dimensional system of SDEs,

\[
\begin{align*}
    dx &= (a_{11}x + a_{12}y) \, dt + \sigma_x \, dW_x, \\
    dy &= \frac{1}{\epsilon}(a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} \, dW_y,
\end{align*}
\]

Want best parameters in reduced model: \(dX = \alpha X dt + \sigma dW_x\)

Standard approach applies averaging theory to find reduced model

\[dX = \tilde{a}X dt + \sigma_x dW_x,
\]

where \(\tilde{a} = a_{11} - a_{12}a_{22}^{-1}a_{21}\). This is an \(O(\sqrt{\epsilon})\) closure.
Motivation for the Reduced Model

**Linear Example:** Consider a two-dimensional system of SDEs,

\[
\begin{align*}
    dx &= (a_{11}x + a_{12}y) \, dt + \sigma_x dW_x, \\
    dy &= \frac{1}{\epsilon}(a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y, 
\end{align*}
\]

Want best parameters in reduced model: \( dX = \alpha X dt + \sigma dW_x \)

Standard approach applies averaging theory to find reduced model

\[
dX = \tilde{\alpha} X dt + \sigma_x dW_x, \]

where \( \tilde{\alpha} = a_{11} - a_{12}a_{22}^{-1}a_{21} \). This is an \( O(\sqrt{\epsilon}) \) closure.

The **Reduced Stochastic Filter** (RSF) uses the Kalman filter with the averaged model.
Numerical Results: Model Error with RSF.
Gottwald & Harlim made the following $O(\epsilon)$ closure rigorous.

\[ dx = (a_{11}x + a_{12}y) \, dt + \sigma_x dW_x, \]
\[ dy = \frac{1}{\epsilon} (a_{21}x + a_{22}y) \, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y, \]

Rewrite the fast equation as follows

\[ y = -\frac{a_{21}}{a_{22}} x - \sqrt{\epsilon} \frac{\sigma_x}{a_{22}} \dot{W}_y + O(\epsilon) \]

and substitute it to the slow equation and ignore the $O(\epsilon)$-term, we obtain

\[ d\tilde{X} = \tilde{a} \tilde{X} \, dt + \sigma_x dW_x - \sqrt{\epsilon} \sigma_y \frac{a_{12}}{a_{22}} dW_y. \]

**Remarks:** This closure approach is known as the stochastic invariant manifold theory (Fenichel 1979, Boxler 1989).
Numerical Results: RSF with additive covariance inflation

Improved mean estimates, but the covariance estimates are still underestimated for large $\epsilon$!
New approach: Asymptotic expansion of the *filter* (not the model).

The full model steady-state filter covariance $\hat{S}$ solves,

$$A_\epsilon \hat{S} + \hat{S} A_\epsilon^T + \hat{S} G^T R^{-1} G \hat{S} + Q_\epsilon = 0.$$
New approach: Asymptotic expansion of the filter (not the model).

The full model steady-state filter covariance $\hat{S}$ solves,

$$
A_\epsilon \hat{S} + \hat{S} A_\epsilon^\top + \hat{S} G^\top R^{-1} G \hat{S} + Q_\epsilon = 0.
$$

Solving for $\hat{s}_{11}$ and expanding in $\epsilon$ we have:

$$
- \left( \frac{1 + 2\epsilon \hat{a}}{R} \right) \hat{s}_{11}^2 + 2\tilde{a} (1 + \epsilon \hat{a}) \hat{s}_{11} + \left( \sigma_x^2 + \epsilon \sigma_y^2 \frac{a_{12}^2}{a_{22}^2} \right) + O(\epsilon^2) = 0
$$
New approach: Asymptotic expansion of the *filter* (not the model).

The full model steady-state filter covariance $\hat{S}$ solves,

$$A_\epsilon \hat{S} + \hat{S} A_\epsilon^\top + \hat{S} G^\top R^{-1} G \hat{S} + Q_\epsilon = 0.$$  

Solving for $\hat{s}_{11}$ and expanding in $\epsilon$ we have:

$$-\left( \frac{1 + 2\epsilon \hat{a}}{R} \right) \hat{s}_{11}^2 + 2\tilde{a} (1 + \epsilon \hat{a}) \hat{s}_{11} + \left( \sigma_x^2 + \epsilon \sigma_y^2 \frac{a_{12}^2}{a_{22}^2} \right) + \mathcal{O}(\epsilon^2) = 0$$

The reduced model has steady state covariance solution, $\tilde{s}$, that satisfies the 1D Riccati equation,

$$-\frac{\tilde{s}^2}{R} + 2\alpha \tilde{s} + \sigma^2 = 0.$$
New approach: Asymptotic expansion of the filter (not the model).

The full model steady-state filter covariance \( \hat{S} \) solves,

\[
A_\epsilon \hat{S} + \hat{S} A_\epsilon^\top + \hat{S} G^\top R^{-1} G \hat{S} + Q_\epsilon = 0.
\]

Solving for \( \hat{s}_{11} \) and expanding in \( \epsilon \) we have:

\[
- \left( \frac{1 + 2\epsilon \hat{a}}{R} \right) \hat{s}_{11}^2 + 2\hat{a} (1 + \epsilon \hat{a}) \hat{s}_{11} + \left( \sigma_x^2 + \epsilon \sigma_y^2 \frac{a_{12}^2}{a_{22}^2} \right) + O(\epsilon^2) = 0.
\]

The reduced model has steady state covariance solution, \( \tilde{s} \), that satisfies the 1D Riccati equation,

\[
- \frac{\tilde{s}^2}{R} + 2\alpha \tilde{s} + \sigma^2 = 0.
\]

Find parameters \( \{\alpha, \sigma\} \) such that \( \tilde{s} = s_{11} + O(\epsilon^2) \)!
Theorem (Manifold of Parameters, BH2013)

Let $\hat{s}_{11}$ be the first diagonal component of the 2D algebraic Riccati equation associated with the true filter and let $\tilde{s}$ be the solution of one-dimensional Ricatti equation associated with the reduced filter. Then

$$\lim_{\epsilon \to 0} \frac{\tilde{s} - \hat{s}_{11}}{\epsilon} = 0$$

if and only if

$$\sigma^2 = 2(\alpha - \tilde{a}(1 - \epsilon \hat{a})) \hat{s}_{11} + \sigma_x^2(1 - 2\epsilon \hat{a}) + \epsilon \sigma_y^2 \frac{a_{12}^2}{a_{22}^2} + O(\epsilon^2). \quad (1)$$
Theorem (Manifold of Parameters, BH2013)

Let $\hat{s}_{11}$ be the first diagonal component of the 2D algebraic Riccati equation associated with the true filter and let $\tilde{s}$ be the solution of one-dimensional Ricatti equation associated with the reduced filter. Then $\lim_{\epsilon \to 0} \frac{\tilde{s} - \hat{s}_{11}}{\epsilon} = 0$ if and only if

$$
\sigma^2 = 2(\alpha - \tilde{a}(1 - \epsilon\hat{a}))\hat{s}_{11} + \sigma^2_x(1 - 2\epsilon\hat{a}) + \epsilon\sigma^2_y \frac{a_{12}^2}{a_{22}^2} + O(\epsilon^2). \tag{1}
$$

Remarks: For any parameters on the manifold (1), the reduced filter mean estimate solves,

$$
d\tilde{x} = \alpha\tilde{x} \, dt + \frac{\tilde{s}}{R} (dz - \tilde{x} \, dt),
$$

while the true filter mean estimate for $x$-variable solves,

$$
d\hat{x} = GA_\epsilon(\hat{x}, \hat{y})^T \, dt + \frac{\hat{s}_{11}}{R} (dz - \hat{x} \, dt). $$
Theorem (Manifold of Parameters, BH2013)

Let $\hat{s}_{11}$ be the first diagonal component of the 2D algebraic Riccati equation associated with the true filter and let $\tilde{s}$ be the solution of one-dimensional Ricatti equation associated with the reduced filter. Then $\lim_{\epsilon \to 0} \frac{\tilde{s} - \hat{s}_{11}}{\epsilon} = 0$ if and only if

$$
\sigma^2 = 2(\alpha - \tilde{a}(1 - \epsilon \hat{a}))\hat{s}_{11} + \sigma_x^2(1 - 2\epsilon \hat{a}) + \epsilon \sigma_y^2 \frac{\tilde{a}_{12}^2}{\tilde{a}_{22}^2} + O(\epsilon^2). \quad (1)
$$

Remarks: For any parameters on the manifold (1), the reduced filter mean estimate solves,

$$
d\tilde{x} = \alpha \tilde{x} \, dt + \frac{\tilde{s}}{R} (dz - \tilde{x} \, dt),
$$

while the true filter mean estimate for $x$-variable solves,

$$
d\hat{x} = GA_{\epsilon}(\hat{x}, \hat{y})^T \, dt + \frac{\hat{s}_{11}}{R} (dz - \hat{x} \, dt).
$$

Impose consistency between the actual error covariance, $E(e^2)$, where $e \equiv \tilde{x} - x$, and $\tilde{s}$ to obtain a unique $\{\alpha, \sigma\}$ in the manifold.
Theorem (Existence and Uniqueness, BH2013)

There exists a unique optimal reduced filter given by the following prior model,

\[ d\tilde{X} = (\tilde{a} - \epsilon \tilde{a} \hat{a})\tilde{X} \, dt + \sigma_x (1 - \epsilon \hat{a})dW_x - \sqrt{\epsilon} \sigma_y \frac{a_{12}}{a_{22}} dW_y, \]

(2)

where \( \tilde{a} = a_{11} - a_{12}a_{21}a_{22}^{-1} < 0 \) and \( \hat{a} = a_{12}a_{21}a_{22}^{-2} \). The optimality is in the sense that, both the mean and covariance estimates converges uniformly to the corresponding estimates from the true filter, with convergence rate on the order of \( \epsilon^2 \).
Optimal Reduced Stochastic Filter

Theorem (Existence and Uniqueness, BH2013)

There exists a unique optimal reduced filter given by the following prior model,

\[ d\tilde{X} = (\tilde{a} - \epsilon\tilde{a}\hat{a})\tilde{X} \, dt + \sigma_x(1 - \epsilon\hat{a})dW_x - \sqrt{\epsilon}\sigma_y \frac{a_{12}}{a_{22}} dW_y, \] (2)

where \( \tilde{a} = a_{11} - a_{12}a_{21}a^{-1}_{22} < 0 \) and \( \hat{a} = a_{12}a_{21}a^{-2}_{22} \). The optimality is in the sense that, both the mean and covariance estimates converges uniformly to the corresponding estimates from the true filter, with convergence rate on the order of \( \epsilon^2 \).

Remark: So, if \( \{\tilde{x}, \tilde{s}\} \) are the solutions of the reduced filter in (2) and \( \{\hat{x}, \hat{s}_{11}\} \) are the solutions of the perfect model, there exists tim-independent constants \( C_1, C_2 \), such that

\[ |\hat{s}_{11}(t) - \tilde{s}(t)| \leq C_1\epsilon^2, \]
\[ \mathbb{E}(|\hat{x}(t) - \tilde{x}(t)|^2) \leq C_2\epsilon^4. \]
Remarks:

▶ Notice that for optimal 1D-filter, the MSE (left) approximately equal to the Covariance estimate (right). We call a filter consistent when the actual error of the mean estimate matches the filtered covariance estimate.

▶ Optimal solutions are always consistent, but consistent solutions are not necessarily optimal.
Remarks:

- Notice that for optimal 1D-filter, the MSE (left) approximately equal to the Covariance estimate (right). We call a filter **consistent** when the actual error of the mean estimate matches the filtered covariance estimates.

- Optimal solutions are always consistent, but consistent solutions are not necessarily optimal.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
- For linear problems there exists a unique reduced model for the slow variables which gives optimal mean and covariance estimates.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
- For linear problems there exists a unique reduced model for the slow variables which gives optimal mean and covariance estimates.
- Finding the reduced model requires imposing consistency on the filter mean and covariance estimates.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
- For linear problems there exists a unique reduced model for the slow variables which gives optimal mean and covariance estimates.
- Finding the reduced model requires imposing consistency on the filter mean and covariance estimates.
- The reduced model includes correction terms in the form of a linear damping and an additive stochastic forcing.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
- For linear problems there exists a unique reduced model for the slow variables which gives optimal mean and covariance estimates.
- Finding the reduced model requires imposing consistency on the filter mean and covariance estimates.
- The reduced model includes correction terms in the form of a linear damping and an additive stochastic forcing.
- For general nonlinear filtering problems, it is impractical to find the unique reduced model since it requires imposing consistency on higher-order moments.
Summary (Linear Example):

- Model error creates inconsistency in the filtered statistical estimates.
- For linear problems there exists a unique reduced model for the slow variables which gives optimal mean and covariance estimates.
- Finding the reduced model requires imposing consistency on the filter mean and covariance estimates.
- The reduced model includes correction terms in the form of a linear damping and an additive stochastic forcing.
- For general nonlinear filtering problems, it is impractical to find the unique reduced model since it requires imposing consistency on higher-order moments.
- A simple test case shows that general nonlinear problems require multiplicative noise.
Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

\[ dx = f_1(x, y; \theta) dt + \sigma_x(x, y; \theta) \, dW_x, \]

\[ dy = \frac{1}{\epsilon} f_2(x, y; \theta) dt + \frac{\sigma_y(x, y; \theta)}{\sqrt{\epsilon}} \, dW_y, \]

\[ dz = h(x) \, dt + \sqrt{R} \, dV. \]

The true filter solutions are characterized by conditional distribution \( p(x, y, t|z_\tau, 0 \leq \tau \leq t) \), which satisfies the Kushner equation (1964):

\[ dp = L^* p \, dt + p(h - \mathbb{E}[h])^\top R^{-1} (dz - \mathbb{E}[h] \, dt), \]
Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

\[\begin{align*}
    dx &= f_1(x, y; \theta) dt + \sigma_x(x, y; \theta) \, dW_x, \\
    dy &= \frac{1}{\epsilon} f_2(x, y; \theta) dt + \frac{\sigma_y(x, y; \theta)}{\sqrt{\epsilon}} \, dW_y, \\
    dz &= h(x) \, dt + \sqrt{R} \, dV.
\end{align*}\]

The true filter solutions are characterized by conditional distribution \( p(x, y, t|z_\tau, 0 \leq \tau \leq t) \), which satisfies the Kushner equation (1964):

\[dp = \mathcal{L}^* p \, dt + p(\mathbb{E}[h])^\top R^{-1} (dz - \mathbb{E}[h] \, dt),\]

Practical issues:

- We have no access to \( p \) for nonlinear problems (SPDE).
- Nonlinearity causes the covariance solutions to depend on higher-order moments and to not equilibrate.
Our strategy:

- Do not look for the unique reduced filter since it will require knowledge and consistency of all higher-order moments.
- Pick a simple nonlinear test problem and consider the first two moments of the posterior distribution, $p$.
- Apply a Gaussian closure to the evolution of these moments.
- Find parameters in a reduced model ansatz by matching the first and second moments of the filtered solutions of perfect model and the reduced models.
- Due to Gaussian closure, even the perfect model may not produce consistent statistics.
- We introduce a consistency metric to determine the performance of the covariance estimate.
Empirical Consistency measure

Definition (Consistency of Covariance)

Let \( \tilde{x}(t) \) and \( \tilde{S}(t) \) be a realization of the solution to a filtering problem for which the true signal of the realization is \( x(t) \). The consistency of the realization is defined to be,

\[
C(x, \tilde{x}, \tilde{S}) = \langle \| x - \tilde{x} \|^2 \rangle = \frac{1}{n} \langle (x(t) - \tilde{x}(t))\top \tilde{S}(t)^{-1} (x(t) - \tilde{x}(t)) \rangle.
\]

We say that a filter is consistent if \( C = 1 \) almost surely (independent of the realization).
Empirical Consistency measure

Definition (Consistency of Covariance)

Let $\tilde{x}(t)$ and $\tilde{S}(t)$ be a realization of the solution to a filtering problem for which the true signal of the realization is $x(t)$. The consistency of the realization is defined to be,

$$C(x, \tilde{x}, \tilde{S}) = \langle \|x - \tilde{x}\|^2 \rangle = \frac{1}{n} \langle (x(t) - \tilde{x}(t))\top \tilde{S}(t)^{-1}(x(t) - \tilde{x}(t)) \rangle.$$

We say that a filter is consistent if $C = 1$ almost surely (independent of the realization).

Remarks:

- Consistency does not imply accurate filter.
- A consistency filter with a good estimate of posterior mean has a good estimate of posterior covariance.
Consider [Gershgorin, Harlim, Majda 2010]:

\[
\frac{du}{dt} = -(\tilde{\gamma} + \lambda_u)u + \hat{b} + \tilde{b} + f(t) + \sigma_u \dot{\mathcal{N}}_u, \\
\frac{d\tilde{b}}{dt} = -\frac{\lambda_b}{\epsilon} \tilde{b} + \frac{\sigma_b}{\sqrt{\epsilon}} \dot{\mathcal{W}}_b, \\
\frac{d\tilde{\gamma}}{dt} = -\frac{\lambda_{\gamma}}{\epsilon} \tilde{\gamma} + \frac{\sigma_{\gamma}}{\sqrt{\epsilon}} \dot{\mathcal{W}}_{\gamma},
\]
Nonlinear Test model

Consider [Gershgorin, Harlim, Majda 2010]:

\[
\begin{align*}
\frac{du}{dt} &= -(\tilde{\gamma} + \lambda_u)u + \hat{b} + \tilde{b} + f(t) + \sigma_u \dot{W}_u, \\
\frac{d\tilde{b}}{dt} &= -\frac{\lambda_b}{\epsilon} \tilde{b} + \frac{\sigma_b}{\sqrt{\epsilon}} \dot{W}_b, \\
\frac{d\tilde{\gamma}}{dt} &= -\frac{\lambda_\gamma}{\epsilon} \tilde{\gamma} + \frac{\sigma_\gamma}{\sqrt{\epsilon}} \dot{W}_\gamma,
\end{align*}
\]

Using the same strategy as for the linear model, we perform asymptotic expansion on the solutions of the optimal filter.
Nonlinear Test model

Consider [Gershgorin, Harlim, Majda 2010]:

\[
\begin{align*}
\frac{du}{dt} &= -(\tilde{\gamma} + \lambda_u)u + \hat{b} + \tilde{b} + f(t) + \sigma_u \dot{W}_u, \\
\frac{d\tilde{b}}{dt} &= -\frac{\lambda_b}{\epsilon} \tilde{b} + \frac{\sigma_b}{\sqrt{\epsilon}} \dot{W}_b, \\
\frac{d\tilde{\gamma}}{dt} &= -\frac{\lambda_\gamma}{\epsilon} \tilde{\gamma} + \frac{\sigma_\gamma}{\sqrt{\epsilon}} \dot{W}_\gamma,
\end{align*}
\]

Using the same strategy as for the linear model, we perform asymptotic expansion on the solutions of the optimal filter.

A detailed computation proves that the optimal reduced filter requires both additive and multiplicative noise.
Continuous-time reduced SPEKF filter:

**Theorem (Existence, BH13)**

Let $\lambda_u > 0$, and observations of the full nonlinear test model,

$$dz = u \, dt + \sqrt{R} \, dV.$$  

Given identical initial statistics, $\bar{u}(0) = \hat{u}(0)$ and $\bar{S}(0) = \hat{S}(0) > 0$, the mean and covariance estimates of a stable continuous-time reduced SPEKF

$$dU = -\alpha U \, dt + \beta U \circ dW_\gamma + \sigma_1 dW_u + \sigma_2 dW_b,$$

with parameters

$$\{ \alpha = \lambda_u, \beta^2 = \frac{\epsilon \sigma^2_\gamma}{\lambda_u (\lambda_u \epsilon + \lambda_\gamma)}, \sigma^2_1 = \sigma^2_u, \sigma^2_2 = \frac{\epsilon \sigma^2_b}{2 \lambda_b (\lambda_b \epsilon + \lambda_u)} \}$$

agree with mean and covariance of a stable continuous-time SPEKF for variable $u$ uniformly, with convergence rate of order-$\epsilon$.

Furthermore, the reduced filtered solutions are also consistent, up to order-$\epsilon$. 

Remarks: Loss an order of $\epsilon$ accuracy due to multiplicative noise.
Continuous-time reduced SPEKF filter:

Theorem (Existence, BH13)

Let $\lambda_u > 0$, and observations of the full nonlinear test model, $dz = u \, dt + \sqrt{R} \, dV$. Given identical initial statistics, $\tilde{u}(0) = \hat{u}(0)$ and $\tilde{S}(0) = \hat{S}(0) > 0$, the mean and covariance estimates of a stable continuous-time reduced SPEKF

$$
dU = -\alpha U \, dt + \beta U \circ dW_\gamma + \sigma_1 dW_u + \sigma_2 dW_b,
$$

with parameters

$$
\left\{ \alpha = \lambda_u, \beta^2 = \frac{\epsilon \sigma_\gamma^2}{\lambda_\gamma(\lambda_u \epsilon + \lambda_\gamma)}, \sigma_1^2 = \sigma_u^2, \sigma_2^2 = \frac{\epsilon \sigma_b^2}{2 \lambda_b(\lambda_b + \epsilon \lambda_u)} \right\}
$$

agree with mean and covariance of a stable continuous-time SPEKF for variable $u$ uniformly, with convergence rate of order-$\epsilon$.

Furthermore, the reduced filtered solutions are also consistent, up to order-$\epsilon$.

Remarks: Loss an order of $\epsilon$ accuracy due to multiplicative noise.
Numerical solutions in the turbulent transfer energy regime with $\epsilon = 1$. 

- SPEKF
- RSF
- RSFA
- Reduced SPEKF
Numerical solutions in the turbulent transfer energy regime with $\epsilon = 1$. 

![Posterior Error Variance Graph](image)
Numerical Solutions for the nonlinear test filtering problems in a regime that mimics dissipative range

Based on these results, we propose the following ansatz,

\[
\left( -\alpha x_i + \sum_{j=1}^{N} \sigma_{ij} \dot{W}_j + \sum_{j=1}^{N} \beta_{ij} \circ x_j \dot{V}_j \right)
\]

as a stochastic parameterization for model error.
Example: Strategy for filtering with model errors

Consider the two-layer Lorenz-96 model,

\[
\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F + \frac{h_x}{M} \sum_{j=1}^{M} y_{i,j},
\]

\[
\epsilon \frac{dy_{i,j}}{dt} = y_{i,j+1}(y_{i,j-1} - y_{i,j+2}) - y_{i,j} + h_y x_i,
\]

where \( x = x(t) \in \mathbb{R}^N \) and \( y = y(t) \in \mathbb{R}^{NM} \) and the subscript \( i \) is taken modulo \( N \) and \( j \) is taken modulo \( M \).
Example: Strategy for filtering with model errors

Consider the two-layer Lorenz-96 model,

\[
\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F + \frac{h_x}{M} \sum_{j=1}^{M} y_{i,j},
\]

\[
\epsilon \frac{dy_{i,j}}{dt} = y_{i,j+1}(y_{i,j-1} - y_{i,j+2}) - y_{i,j} + h_y x_i,
\]

where \( x = x(t) \in \mathbb{R}^N \) and \( y = y(t) \in \mathbb{R}^{NM} \) and the subscript \( i \) is taken modulo \( N \) and \( j \) is taken modulo \( M \).

**Proposed Reduced Filter Model:**

\[
\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F
\]
Consider the two-layer Lorenz-96 model,

\[
\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F + \frac{h_x}{M} \sum_{j=1}^{M} y_{i,j},
\]

\[
\epsilon \frac{dy_{i,j}}{dt} = y_{i,j+1}(y_{i,j-1} - y_{i,j+2}) - y_{i,j} + h_y x_i,
\]

where \( x = x(t) \in \mathbb{R}^N \) and \( y = y(t) \in \mathbb{R}^{NM} \) and the subscript \( i \) is taken modulo \( N \) and \( j \) is taken modulo \( M \).

**Proposed Reduced Filter Model:**

\[
\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F
\]

\[
+ \left( -\alpha x_i + \sum_{j=1}^{N} \sigma_{ij} \dot{W}_j + \sum_{j=1}^{N} \beta_{ij} \circ x_j \dot{V}_j \right)
\]
Details of the Simulation

- $N = 9$ slow variables, $M = 8$ implies 72 fast variables.
- Data generated from the 81-dimensional two-layer L96 model.
- The 9 slow variables are observed with Gaussian noise.
- Ensemble Kalman Filter (EnKF) with each model.
- Parameters $\alpha$ and $\sigma$ are fit from the data.

- We measure the performance of the mean estimate (RMSE).
- We consistency to measure the accuracy of the covariance estimate.
- Consistency $> 1$ $\implies$ Underestimating covariance.
- Consistency $< 1$ $\implies$ Overestimating covariance.
Numerical results \((x \in \mathbb{R}^9, y \in \mathbb{R}^{72})\)

RDF = Reduced Deterministic Filter \((\alpha = \beta = \sigma = 0)\)
RDFD = Reduced Deterministic Filter with damping \((\beta = \sigma = 0)\)
RSFA = Reduced Stochastic Filter with additive noise \((\alpha = \beta = 0)\)
RSFAD = Reduced Stochastic Filter with damping and additive noise \((\beta = 0)\)
References:


- J. Harlim, “Data assimilation with model error from unresolved scales”, submitted.