Linear Theory for Filtering Nonlinear Multiscale Systems with Model Error

Tyrus Berry[1]

Department of Mathematics
The Pennsylvania State University

Collaborator: John Harlim
The Pennsylvania State University

Dec 18, 2013

[1]Supports: Office of Naval Research and the National Science Foundation

Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

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

Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

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

Model Error from Neglected Scales:

- Model for slow time variables x are known (f_1 and σ_x).
- Observation only depends on x and is known (h).
- Fast variables y are unknown and unobserved. (f_2 and σ_y).

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:

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

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:

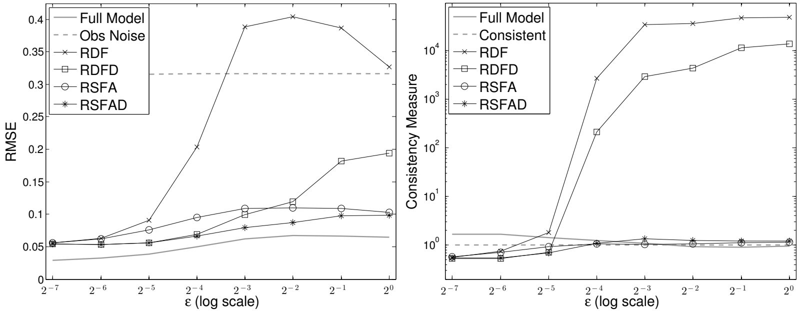
$$\begin{aligned}\frac{dx_i}{dt} &= x_{i-1}(x_{i+1} - x_{i-2}) - ax_i + F \\ &\quad + \Big(- \alpha x_i + \sum_{j=1}^{N} \sigma_{ij}\dot{W}_j + \sum_{j=1}^{N} \beta_{ij} \circ x_j \dot{V}_j \Big)\end{aligned}$$

Details of the Simulation

- $M = 9$ slow variables, $N = 8$ 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 α and σ are fit from the data.

- We measure the performance of the mean estimate (RMSE).
- We use a measure called *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 ($\beta = 0$)

Motivation for the Reduced Model

Recall: We compensate for the model error with linear damping and additive and multiplicative stochastic forcing.

Motivation for the Reduced Model

Recall: We compensate for the model error with linear damping and additive and multiplicative stochastic forcing.

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

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

Motivation for the Reduced Model

Recall: We compensate for the model error with linear damping and additive and multiplicative stochastic forcing.

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

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

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

Motivation for the Reduced Model

Recall: We compensate for the model error with linear damping and additive and multiplicative stochastic forcing.

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

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

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 $\mathcal{O}(\sqrt{\epsilon})$ closure.

Understanding covariance inflation

Gottwald & Harlim made the following $\mathcal{O}(\epsilon)$ closure rigorous.

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

Rewrite the fast equation as follows

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

and substitute it to the slow equation and ignore the $\mathcal{O}(\epsilon)$-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).

New approach: Asymptotic expansion of the *filter* (not the model).

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

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

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 ϵ we have:

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

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 ϵ 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 ϵ 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,

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

Find parameters $\{\alpha, \sigma\}$ such that $\tilde{s} = \hat{s_{11}} + \mathcal{O}(\epsilon^2)$!

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*

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

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

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

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,

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

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

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

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,

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

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

Optimal Reduced Stochastic Filter

Theorem (Existence and Uniqueness, BH2013)

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

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

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* ϵ^2.

Optimal Reduced Stochastic Filter

Theorem (Existence and Uniqueness, BH2013)

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

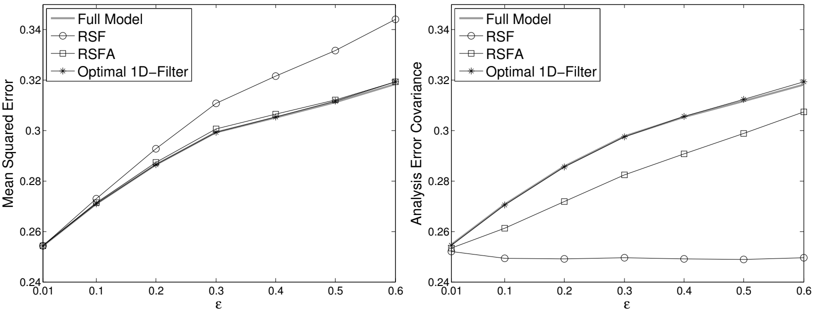
$$d\tilde{X} = (\tilde{a} - \epsilon\tilde{a}\hat{a})\tilde{X}\,dt + \sigma_x(1 - \epsilon\hat{a})dW_x - \sqrt{\epsilon}\sigma_y\frac{a_{12}}{a_{22}}dW_y, \qquad (2)$$

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* ϵ^2.

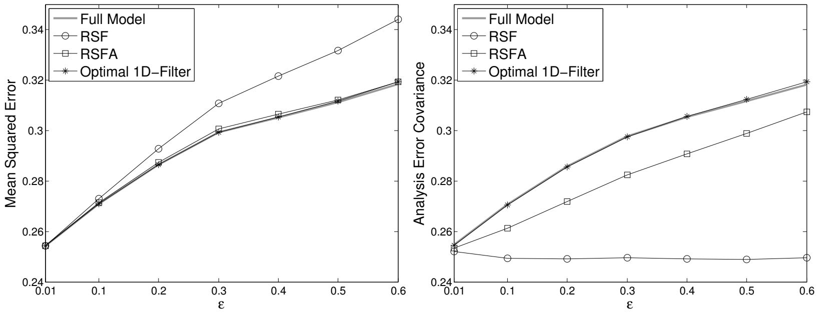
Remark: So, if $\{\tilde{x}, \tilde{s}\}$ are the solutions of the reduced filter in (2) and $\{\hat{x}, \hat{s}_{11}\}$ are the solutions of the perfect model, there exists tim-independent constants C_1, C_2, such that

$$|\hat{s}_{11}(t) - \tilde{s}(t)| \leq C_1\epsilon^2,$$

$$\mathbb{E}(|\hat{x}(t) - \tilde{x}(t)|^2) \leq C_2\epsilon^4.$$



Full Model
RSF
RSFA
Optimal 1D−Filter
Mean Squared Error
ε
0.24
0.26
0.28
0.3
0.32
0.34
0.01
0.1
0.2
0.3
0.4
0.5
0.6
Full Model
RSF
RSFA
Optimal 1D−Filter
Analysis Error Covariance
ε



Remarks:

- Notice that for optimal 1D-filter, the MSE (left) approximately equal to the Covariance estimate (right). This is what we called **consistent** filter: the actual error of the mean estimate matches the filtered covariance estimates.
- Optimal solutions are always consistent, but consistent solutions are not necessarily optimality.

Summary (Theory):

- Model error creates inconsistency in the filtered statistical estimates.

Summary (Theory):

- 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 (Theory):

- 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 (Theory):

- 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 (Theory):

- 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 (Theory):

- 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.

Summary (Practical):

Based on these results, we propose the ansatz used above,

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

as a stochastic parameterization for model error.

Summary (Practical):

Based on these results, we propose the ansatz used above,

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

as a stochastic parameterization for model error.

Many practical questions remain:

- Should α be a matrix (spatial dependent)?
- How can we estimate α, β, σ efficiently from data?
- In particular, we currently set $\beta = 0$ since we do not have an estimation procedure available.
- Is it feasible to make these parameters state dependent?

References:

- T. Berry & J. Harlim, "Linear theory for filtering nonlinear multiscale systems with model error", submitted.
- G.A. Gottwald & J. Harlim, "The role of additive and multiplicative noises in filtering complex dynamical systems", Proc R. Soc. A 469(2155), 2013.
- T. Berry & T. Sauer, "Adaptive ensemble Kalman filtering of nonlinear systems", Tellus A 65:20331, 2013.
- J. Harlim, A. Mahdi, & A.J. Majda, "An ensemble kalman filter for statistical estimation of physics constrained nonlinear regression models", J. Comput. Phys. 257A: 782-812, 2014.
- J. Harlim, "Data assimilation with model error from unresolved scales", submitted.

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,

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

We say that a filter is consistent if $\mathcal{C} = 1$ 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,

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

We say that a filter is consistent if $\mathcal{C} = 1$ 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.

Returning to Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

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

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

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

Returning to Nonlinear Filtering Problems

Consider the following prototype continuous-time filtering problem,

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

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

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

Practical issues:

- We have no access to p for high-dimensional nonlinear problems.
- Nonlinearity causes the covariance solutions to depend on higher-order moments and to not equilibrate.

Nonlinear Test model

Consider [Gershgorin, Harlim, Majda 2010]:

$$\frac{du}{dt} = -(\tilde{\gamma} + \lambda_u)u + \hat{b} + \tilde{b} + f(t) + \sigma_u \dot{W}_u,$$

$$\frac{d\tilde{b}}{dt} = -\frac{\lambda_b}{\epsilon}\tilde{b} + \frac{\sigma_b}{\sqrt{\epsilon}}\dot{W}_b,$$

$$\frac{d\tilde{\gamma}}{dt} = -\frac{\lambda_\gamma}{\epsilon}\tilde{\gamma} + \frac{\sigma_\gamma}{\sqrt{\epsilon}}\dot{W}_\gamma,$$

Nonlinear Test model

Consider [Gershgorin, Harlim, Majda 2010]:

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

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]:

$$\frac{du}{dt} = -(\tilde{\gamma} + \lambda_u)u + \hat{b} + \tilde{b} + f(t) + \sigma_u \dot{W}_u,$$

$$\frac{d\tilde{b}}{dt} = -\frac{\lambda_b}{\epsilon}\tilde{b} + \frac{\sigma_b}{\sqrt{\epsilon}}\dot{W}_b,$$

$$\frac{d\tilde{\gamma}}{dt} = -\frac{\lambda_\gamma}{\epsilon}\tilde{\gamma} + \frac{\sigma_\gamma}{\sqrt{\epsilon}}\dot{W}_\gamma,$$

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.

Numerical Solutions for the nonlinear test filtering problems in a regime that mimics dissipative range