The role of additive and multiplicative noise in filtering complex dynamical systems[1]

John Harlim[2]

Department of Mathematics and Department of Meteorology
The Pennsylvania State University

Joint work with: Georg A. Gottwald[3]
School of Mathematics and Statistics
University of Sydney

October 7, 2013

Motivations:

- In practice, it is very difficult NOT to UNDERESTIMATE the error covariance of filtered solutions, especially when the filtering problem is solved with imperfect model and/or nonlinear.

Motivations:

- In practice, it is very difficult NOT to UNDERESTIMATE the error covariance of filtered solutions, especially when the filtering problem is solved with imperfect model and/or nonlinear.
- In general, it is difficult to compute the covariance explicitly even when the dimensionality of the problem is relatively low.

Motivations:

- In practice, it is very difficult NOT to UNDERESTIMATE the error covariance of filtered solutions, especially when the filtering problem is solved with imperfect model and/or nonlinear.
- In general, it is difficult to compute the covariance explicitly even when the dimensionality of the problem is relatively low.
- One attractive method that is scalable for high-dimensional problem is ensemble Kalman filter. When ensemble size is too small, the covariance is underestimated, it is common to empirically inflate the covariance via:

$$R_m \leftarrow (1+r)R_m \quad \text{(Anderson \& Anderson 1999)}$$
$$R_m \leftarrow R_m + Q \quad \text{(Kalnay et al. 2007)}$$

Motivations:

- In practice, it is very difficult NOT to UNDERESTIMATE the error covariance of filtered solutions, especially when the filtering problem is solved with imperfect model and/or nonlinear.

- In general, it is difficult to compute the covariance explicitly even when the dimensionality of the problem is relatively low.

- One attractive method that is scalable for high-dimensional problem is ensemble Kalman filter. When ensemble size is too small, the covariance is underestimated, it is common to empirically inflate the covariance via:

$$R_m \leftarrow (1+r)R_m \quad \text{(Anderson \& Anderson 1999)}$$
$$R_m \leftarrow R_m + Q \quad \text{(Kalnay et al. 2007)}$$

- There are also adaptive schemes to estimate inflation factors (Anderson 2007, 2009, Li et al. 2009, Miyoshi 2011).

► Our goal is to mathematically understand the covariance inflation in the context of filtering multiscale systems with moderate scale gap in the presence of model errors through reduced dynamics from subgrid-scale parameterizations.

- Our goal is to mathematically understand the covariance inflation in the context of filtering multiscale systems with moderate scale gap in the presence of model errors through reduced dynamics from subgrid-scale parameterizations.
- In particular, we will compute the effective "statistical" inflation for analytically solvable linear and nonlinear filtering test problems with one order higher perturbation expansion than the usual averaging limit.

Goal of this talk

- Our goal is to mathematically understand the covariance inflation in the context of filtering multiscale systems with moderate scale gap in the presence of model errors through reduced dynamics from subgrid-scale parameterizations.

- In particular, we will compute the effective "statistical" inflation for analytically solvable linear and nonlinear filtering test problems with one order higher perturbation expansion than the usual averaging limit.

- We will formulate these statistical inflations with a dynamically consistent reduced stochastic model. For linear problem, we will find a naturally arise additive noise correction. For nonlinear problem, we will find a naturally arise multiplicative noise correction in addition to additive noise.

Consider

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

We assume:

- $\sigma_x, \sigma_y \neq 0$ and the deterministic dynamics are stable to assure the existence of a unique joint invariant density $\rho_\infty(x, y)$.
- Furthermore we require $\tilde{a} = a_{11} - a_{12}a_{22}^{-1}a_{21} < 0$ and for fixed x, the fast dynamics is ergodic.

Consider

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

We assume:

- $\sigma_x, \sigma_y \neq 0$ and the deterministic dynamics are stable to assure the existence of a unique joint invariant density $\rho_\infty(x, y)$.

- Furthermore we require $\tilde{a} = a_{11} - a_{12}a_{22}^{-1}a_{21} < 0$ and for fixed x, the fast dynamics is ergodic.

Then for a finite time, the solutions of the slow dynamics converge to the solutions of

$$dX = \tilde{a}Xdt + \sigma_x dW_x.$$

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

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

Stochastic invariant manifold theory (Fenichel 1979).

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:

$$\mathbb{E}(\tilde{X}_t) = e^{\tilde{a}t}\mathbb{E}(\tilde{X}_0)$$

$$Var(\tilde{X}_t) = Var(\tilde{X}_0)e^{2\tilde{a}t} - \frac{\sigma_x^2}{2\tilde{a}}(1 - e^{2\tilde{a}t}) - \epsilon\frac{a_{12}^2\sigma_y^2}{2\tilde{a}a_{22}^2}(1 - e^{2\tilde{a}t})$$

$$= Var(x_0)e^{2\tilde{a}t} + Q_0 + \epsilon Q_1.$$

Theorem

*Consider the full linear multi-scale system with the assumptions.
Let x^ϵ be solutions of the full multi-scale system and $\tilde{X}$ be the
solution of the reduced equation,*

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

*corresponding to the same realizations W_x, W_y. Let $x^\epsilon(0) = \tilde{X}(0)$.
Then the error $e(t) = x^\epsilon(t) - \tilde{X}(t)$ is bounded for finite time T by*

$$\mathbb{E}\Big(\sup_{0 \le t \le T} |e(t)|^2\Big) \le c\epsilon^2.$$

Theorem

Consider the full linear multi-scale system with the assumptions. Let x^ϵ be solutions of the full multi-scale system and $\tilde{X}$ be the solution of the reduced equation,

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

corresponding to the same realizations W_x, W_y. Let $x^\epsilon(0) = \tilde{X}(0)$. Then the error $e(t) = x^\epsilon(t) - \tilde{X}(t)$ is bounded for finite time T by

$$\mathbb{E}\Big(\sup_{0 \le t \le T} |e(t)|^2\Big) \le c\epsilon^2.$$

Remark: This result extends that in (Zhang 2011) which provides a linear convergence rate for the classical averaged model,

$$dX = \tilde{a}X\,dt + \sigma_x dW_x.$$
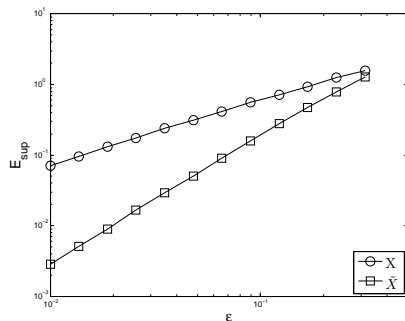
Pathwise error convergence rate.

Figure : Convergence of the full model and the reduced models on a time interval $0 \leq t \leq T = 250$. Linear regression of the two data sets reveals a slope of 0.9 and 1.8 for the error of X and $\tilde{X}$, respectively.

Implication on data assimilation

We consider approximating the following filtering problem,

$$dx = (a_{11}x + a_{12}y)\, dt + \sigma_x dW_x,$$
$$dy = \frac{1}{\epsilon}(a_{21}x + a_{22}y)\, dt + \frac{\sigma_y}{\sqrt{\epsilon}} dW_y,$$
$$z_m = x(t_m) + \varepsilon_m^o, \quad \varepsilon_m^o \sim \mathcal{N}(0, r^o),$$

with reduced stochastic filters:

- ▶ RSF (classical reduced stochastic filters)

$$dX = \tilde{a}X\, dt + \sigma_x dW_x,$$
$$z_m = X(t_m) + \varepsilon_m^o.$$

- ▶ RSFA (reduced stochastic filters with additive covariance inflation)

$$d\tilde{X} = \tilde{a}\tilde{X}\, dt + \sigma_x dW_x - \sqrt{\epsilon}\sigma_y \frac{a_{12}}{a_{22}} dW_y,$$
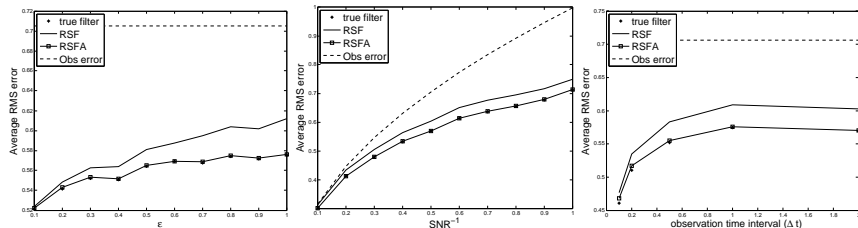$$z_m = \tilde{X}(t_m) + \varepsilon_m^o.$$

Figure : Filter accuracy: Average RMS errors as functions of ϵ for $\Delta t = 1$ and $SNR^{-1} = 0.5$ (left panel); as functions of SNR^{-1} for $\epsilon = 1$ and $\Delta t = 1$ (middle panel); as functions of Δt for $\epsilon = 1$ and $SNR^{-1} = 0.5$ (right panel).

Nonlinear SPEKF model

SPEKF "Stochastic Parameterized Extended Kalman Filter" model (GHM 2010) is defined as follows,

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

with $\hat{\lambda} = \hat{\gamma} - \mathrm{i}\omega$ and $\lambda_b = \gamma_b - \mathrm{i}\omega_b$.

Nonlinear SPEKF model

SPEKF "Stochastic Parameterized Extended Kalman Filter" model (GHM 2010) is defined as follows,

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

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

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

with $\hat{\lambda} = \hat{\gamma} - i\omega$ and $\lambda_b = \gamma_b - i\omega_b$.
Here we consider temporal scales of t/ϵ for $\tilde{b}, \tilde{\gamma}$.

Nonlinear SPEKF model

SPEKF "Stochastic Parameterized Extended Kalman Filter" model (GHM 2010) is defined as follows,

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

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

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

with $\hat{\lambda} = \hat{\gamma} - \mathrm{i}\omega$ and $\lambda_b = \gamma_b - \mathrm{i}\omega_b$.
Here we consider temporal scales of t/ϵ for $\tilde{b}, \tilde{\gamma}$.

Remark: SPEKF model has explicit statistical non-Gaussian solutions.

Solutions of SPEKF in three regimes

Recent studies by (Branicki, Gershgorin, and Majda 2012) suggest that the system above can reproduce signals in various turbulent regimes for $\epsilon = 1$:
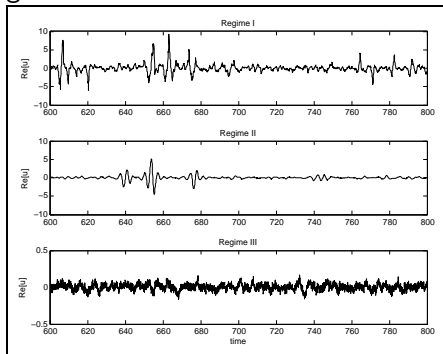
I. Turbulent transfer energy regime: $\tilde{\gamma}$ decays faster than u.

II. Dissipative range: u and $\tilde{\gamma}$ have comparable decaying time scales.

III. Laminar mode: u decays faster than $\tilde{\gamma}$.

Theorem

Assume that $f(t)$ and the statistics up to order-4th moments are bounded; $\tilde{\gamma}$ decays sufficiently faster than u. Let u^ϵ be a solution of the SPEKF model and U be a solution of

$$\frac{dU}{dt} = -\hat{\lambda}U + \hat{b} + f(t) + \sigma_u \dot{W}_u + \sqrt{\epsilon}\frac{\sigma_b}{\lambda_b}\dot{W}_b - \sqrt{\epsilon}\frac{\sigma_\gamma}{d_\gamma}U \circ \dot{W}_\gamma,$$

where $\hat{\lambda} = \hat{\gamma} - i\omega$, corresponding to the same realizations W_u, W_b, W_γ. Let $u^\epsilon(0) = U(0)$. Then, the error $e(t) = u^\epsilon(t) - U(t)$ is bounded for finite time T.

Theorem

Assume that $f(t)$ and the statistics up to order-4th moments are bounded; $\tilde{\gamma}$ decays sufficiently faster than u. Let u^ϵ be a solution of the SPEKF model and U be a solution of

$$\frac{dU}{dt} = -\hat{\lambda}U + \hat{b} + f(t) + \sigma_u \dot{W}_u + \sqrt{\epsilon}\frac{\sigma_b}{\lambda_b}\dot{W}_b - \sqrt{\epsilon}\frac{\sigma_\gamma}{d_\gamma}U \circ \dot{W}_\gamma,$$

where $\hat{\lambda} = \hat{\gamma} - i\omega$, corresponding to the same realizations W_u, W_b, W_γ. Let $u^\epsilon(0) = U(0)$. Then, the error $e(t) = u^\epsilon(t) - U(t)$ is bounded for finite time T.

Remark:

- ▶ The multiplicative noise in the Stratonovich sense is crucial for correcting the mean through a drift term.

Filtering $v_m = u(t_m) + \varepsilon_m^o$, where $\varepsilon_m^o \sim \mathcal{N}(0, r^o)$,

1. The true filter (GHM:10, MH Ch 13).
2. Reduced stochastic filter (RSF). This approach uses prior model

$$\frac{dU}{dt} = -\hat{\lambda}U + \hat{b} + f(t) + \sigma_u \dot{W}_u,$$

3. Reduced stochastic filter with additive inflation (RFSA).

$$\frac{dU}{dt} = -\hat{\lambda}U + \hat{b} + f(t) + \sigma_u \dot{W}_u + \sqrt{\epsilon}\frac{\sigma_b}{\lambda_b}\dot{W}_b,$$

4. Reduced stochastic filter with combined, additive and multiplicative, inflations (RSFC).

$$\frac{dU}{dt} = -\hat{\lambda}U + \hat{b} + f(t) + \sigma_u \dot{W}_u + \sqrt{\epsilon}\frac{\sigma_b}{\lambda_b}\dot{W}_b - \sqrt{\epsilon}\frac{\sigma_\gamma}{d_\gamma}U \circ \dot{W}_\gamma,$$
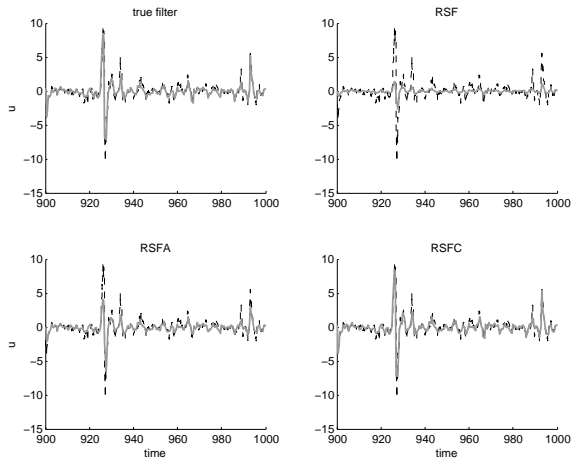
Figure : Filtered posterior estimates (grey) compared to the truth (black dashes) for regime I with $SNR^{-1} = 0.5$.
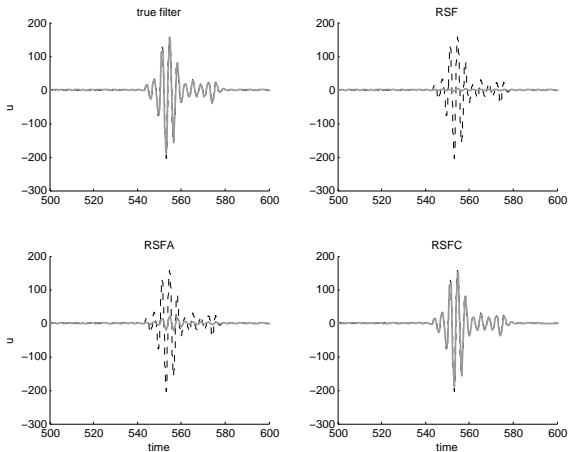
Figure : Filtered posterior estimates (grey) compared to the truth (black dashes) for regime II with $SNR^{-1} = 0.5$.

Summary:

- We presented a study of filtering partially observed multi-scale systems with reduced stochastic models obtained from a systematic closure on the unresolved fast processes.

- Here, we were not only showing convergence of solutions in the limit of large time scale separation, but we also tackled the question of how the stochasticity induced by the unresolved scales can enhance the filtering skill, and how their diffusive behaviour can be translated into effective inflation with naturally arise additive and multiplicative noises.

- The main message of this work here is that reduced stochastic models can be viewed as dynamically consistent way to introduce covariance inflation as well as mean correction, guiding the filtering process.