Mathematical Strategies for Filtering Turbulent Dynamical Systems

By Andrew J. Majda
Morse Professor of Arts and Sciences
Center for Atmosphere/Ocean Science (CAOS)
Courant Institute of Mathematical Sciences (CIMS)
New York University

<u>Main Collaborators:</u>
John Harlim, North Carolina State University
Boris Gershgorin, CIMS, New York University (postdoc)
Marcus Grote, University of Basel

Supported by National Science Foundation, Office of Naval Research and DARPA

Modern Applied Modus Operandi

Theory: Important mathematical guidelines
Qualitative Exactly Solvable Models

Novel Algorithms:
Applications to Real Problems in Science/Engineering

General Refs for Talk: Research/Expository

A. Majda, J. Harlim, and B. Gershgorin "Mathematical Strategies for Filtering Turbulent Dynamical System" 2010, Dis. Cont. Dyn. Sys., 27, pp 441-486

Introductory Graduate Text

A. Majda and J. Harlim, "Mathematical Strategies for Real Time Filtering of Turbulent Signals in Complex Systems," Cambridge University Press (2011)

Exactly Solvable Test Models and NEKF Algorithms

Prototype Test Problems which are Nonlinear yet exactly solvable statistically for filtering multiple time scale systems

Examples: Gravity Waves, Moisture, and Large Scale Flow in Tropics or Mesoscale, Tracking hazardous pollutants in real time from partial observations

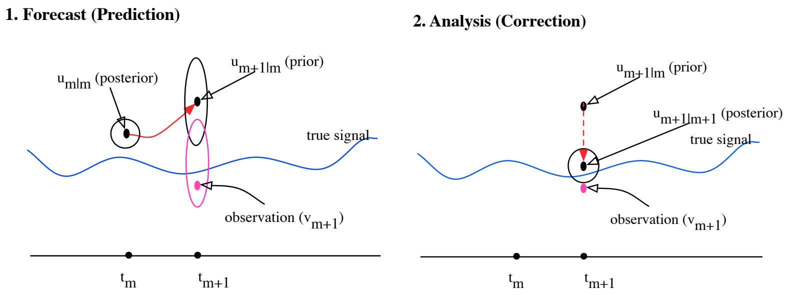
References:

1. B. Gershgorin and A. Majda, 2008, "**A nonlinear test model for filtering slow-fast systems**," *Comm. Math. Sci.*, 6, 3, pp. 611 – 649

2. B. Gershgorin and A. Majda, 2010, "**Filtering a nonlinear slow-fast system with strong fast forcing**," *Comm. Math. Sci.* 8, 1, pp. 67 – 92

3. B. Gershgorin and A. Majda, 2011, "**Filtering a statistically exactly solvable test model for turbulent tracers from partial observations**," *J. Comp. Phys*, Vol. 230, February 2011, pp 1602-1638

What is filtering?

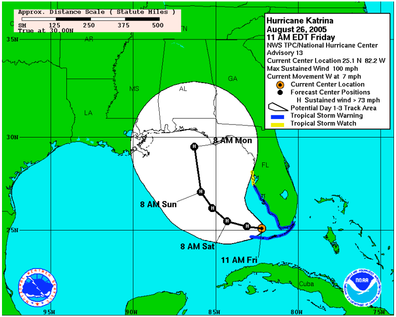


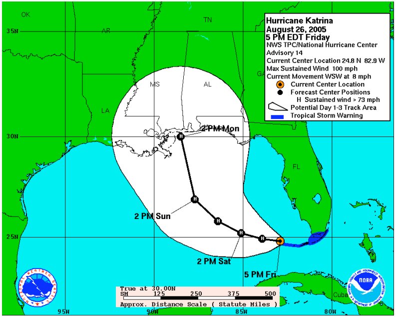
The correction step is an application of Bayesian update

$$p(u_{m+1|m+1}) \equiv p(u_{m+1|m}|v_{m+1}) \sim p(u_{m+1|m})p(v_{m+1}|u_{m+1|m})$$

Kalman filter formula produces the optimal unbiased posterior mean and covariance by assuming linear model and Gaussian observations and forecasts errors.

Example of application: predicting path of hurricane





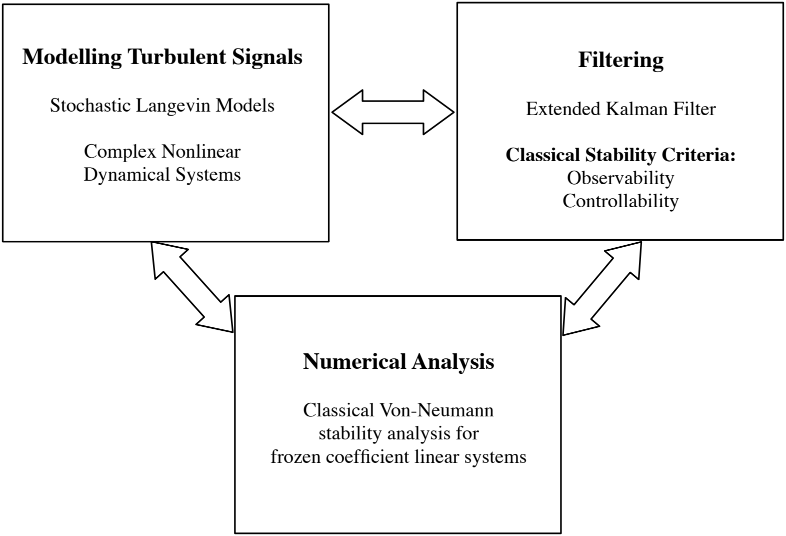
Theoretical and Computational Issues:

- Handling nonlinearity! Why not particle filter? Convergence requires ensemble size that grows exponentially with respect to the ensemble spread relative to observation errors rather than to the state dimension per se(Bengtsson, Bickel, and Li 2008).
- How to handle large system? Perhaps $N = 10^6$ state variables (e.g., 200 km resolved Global Weather Model)
- Where is the computational burden? Propagating covariance matrix of size $N \times N$ (6N minutes = 300,000 hours).
- Some successful strategies: Ensemble Kalman filters (ETKF of Bishop et al. 2001, EAKF of Anderson 2001). Each involves computing singular value decomposition (SVD).
- However, these accurate filters are not immune from "catastrophic filter divergence" (diverge beyond machine infinity) when observations are sparse, even when the true signal is a dissipative system with "absorbing ball property".

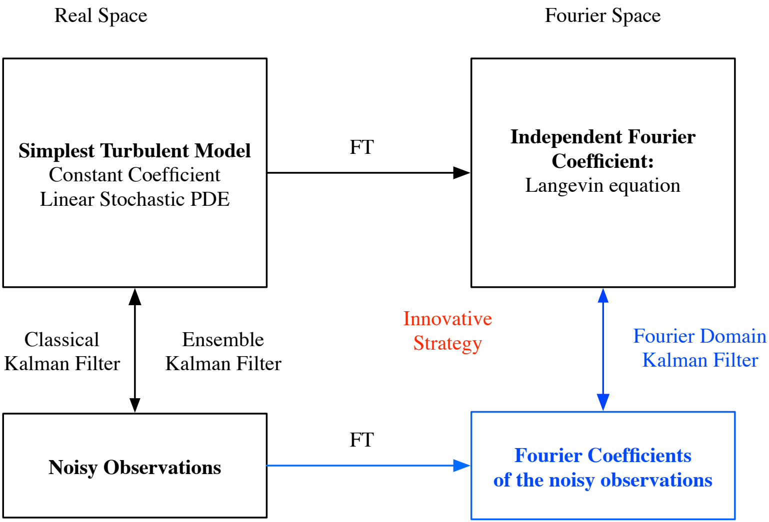
Main theoretical questions:

1. How to develop simple off-line mathematical test criteria as guidelines for filtering extremely stiff multiple space-time scale problems that often arise in filtering turbulent signals through plentiful and sparse observations?
2. For turbulent signals from nature with many scales, even with mesh refinement the model has inaccuracies from parametrization, under-resolution, etc. Can judicious model error help filtering?
3. Can new strategies and stochastic parameterizations be developed to reduce model error and improve the filtering as well as the prediction skills?

Goal: Provide math guidelines and new numerical strategies thru modern applied math paradigm

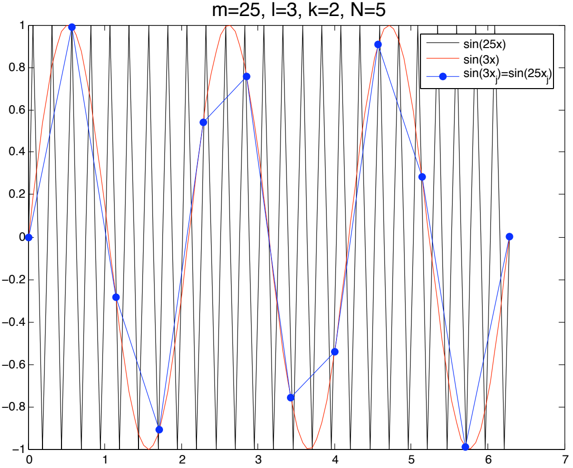


PART I: Filtering Linear Problem



How to deal with Sparse Regularly Spaced Observations?

ALIASING !!



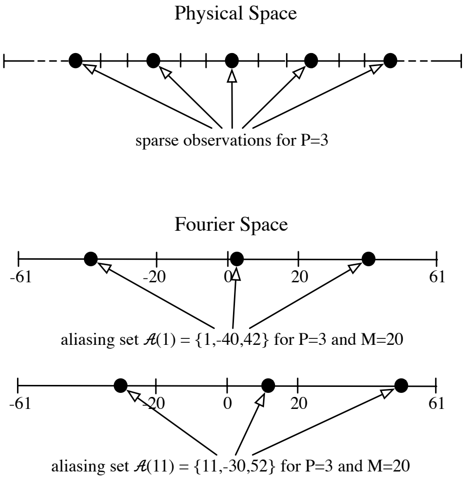
Recall Aliasing Formula:

- Fine mesh: $f(x_j) = \sum_{|k|\leq N} \hat{f}_{fine}(k)\mathrm{e}^{\mathrm{i}kx_j}$ where $x_j = jh$ and $(2N+1)h = 2\pi$.
- Coarse mesh: $f(\tilde{x}_j) = \sum_{|\ell|\leq M} \hat{f}_{coarse}(\ell)\mathrm{e}^{\mathrm{i}\ell\tilde{x}_j}$ where $\tilde{x}_j = j\tilde{h}$ and $(2M+1)\tilde{h} = 2\pi$.
- Suppose the coarse grid points $\tilde{x}_j$ coincide with the fine mesh grid points x_j at every $P = (2N+1)/(2M+1)$ fine grid points.
- Since $\mathrm{e}^{\mathrm{i}k\tilde{x}_j} = \mathrm{e}^{\mathrm{i}(\ell+q(2M+1))\tilde{x}_j} = \mathrm{e}^{\mathrm{i}\ell\tilde{x}_j}$,
- We deduce

$$\hat{f}_{coarse}(\ell) \quad = \quad \sum_{k_j\in\mathcal{A}(\ell)} \hat{f}_{fine}(k_j), \quad |\ell| \leq M,$$

where $\mathcal{A}(\ell) = \{k : |k| \leq N, k = \ell + q(2M+1), q \in \mathbb{Z}\}$

Example: 123 grid pts (61 modes) but only 41 observations (20 modes) available



Stochastically forced advection-diffusion equation

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\partial}{\partial x}u(x,t) + \bar{F}(x,t) + \mu\frac{\partial^2}{\partial x^2}u(x,t) + \sigma(x)\dot{W}(t)$$

on a periodic domain $0 \leq x \leq 2\pi$.

Observations at sparse grid points:

$$v(\tilde{x}_j, t_m) = u(\tilde{x}_j, t_m) + \sigma_m^o, \quad \tilde{x}_j = j\tilde{h}, (2M+1)\tilde{h} = 2\pi.$$

In Fourier Domain

- $d\hat{u}_k(t) = [(-\mu k^2 - \mathrm{i}k)\hat{u}_k(t) + A\mathrm{e}^{-\mathrm{i}kt}\delta_{|k|<M}]dt + \sigma_k dW_k(t)$
- Equilibrium variance: $E_k = \sigma_k^2/2\mu k^2$, decorrelation time $1/\mu k^2$.
- Observation at time t_m (apply the aliasing formula):

$$\hat{v}_{\ell,m} = \sum_{k_j \in \mathcal{A}(\ell)} \hat{u}_{k_j,m} + \hat{\sigma}_{\ell,m}^o, = G\vec{\hat{u}}_{\ell,m} + \hat{\sigma}_{\ell,m}^o$$

where $G = [1, 1, \ldots, 1]$ and $\hat{\sigma}_{\ell,m}^o \sim \mathcal{N}(0, r^o/(2M+1))$.

The Fourier Domain Kalman Filter (FDKF)
The standard Kalman filter algorithm applied to all the disjoint aliasing sets $\mathcal{A}(\ell)$ for all $0 \le \ell \le M$.
Prior Update

$$\begin{aligned} &\text{A)} \quad \bar{\hat{u}}_{k_j,m+1|m} = F_{k_j}\bar{\hat{u}}_{k_j,m|m} + \bar{F}_{k_j,m}, \quad k_j \in \mathcal{A}(\ell) \\ &\text{B)} \quad R_{\ell,m+1|m} = F_\ell R_{\ell,m|m} F_\ell^* + R_\ell, \end{aligned}$$

where

$$\begin{aligned} F_\ell(j,j) &= \mathrm{e}^{(-\mu k_j^2 + \mathrm{i}k_j)\Delta t}, \text{ and} \\ R_\ell(j,j) &= \frac{\sigma_{k_j}^2}{2\mu k_j^2}(1 - \mathrm{e}^{-\mu k_j^2 \Delta t}), \quad k_j \in \mathcal{A}(\ell), \end{aligned}$$

Posterior update:

$$\begin{aligned} &\text{D)} \quad \vec{\bar{\hat{u}}}_{\ell,m+1|m+1} = (\mathcal{I} - K_{\ell,m+1}G)\vec{\bar{\hat{u}}}_{\ell,m+1|m} + K_{\ell,m+1}\hat{v}_{\ell,m+1} \\ &\text{E)} \quad R_{\ell,m+1|m+1} = (\mathcal{I} - K_{\ell,m+1}G)R_{\ell,m+1|m}, \\ &\text{F)} \quad K_{\ell,m+1} = R_{\ell,m+1|m}G^T(GR_{\ell,m+1|m}G^T + \hat{r}^o)^{-1}. \end{aligned}$$

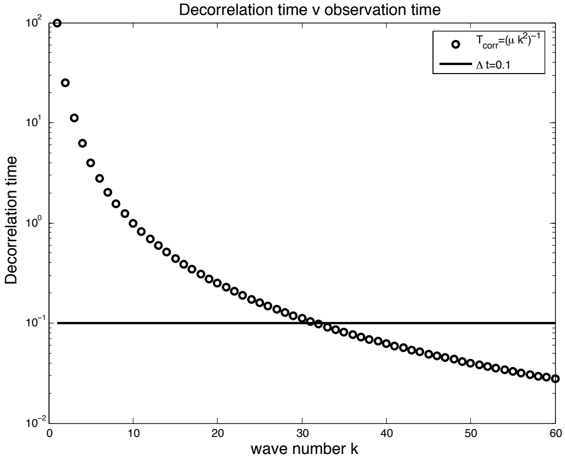
Strongly Damped Approximate Filter (SDAF):
In each aliasing set $\mathcal{A}(\ell) = \{k_1, k_2, \dots, k_P\}$, the damping strength varies enormously (e.g. $\mathcal{A}(1) = \{1, -40, 42\}$)

$$\begin{aligned} \text{Low wave number } k_1\text{:} \quad \mathrm{e}^{-\mu k_1^2 \Delta t} &= \mathcal{O}(1), \\ \text{High wave numbers } k_j\text{:} \quad \mathrm{e}^{-\mu k_j^2 \Delta t} &= \mathcal{O}(\epsilon) \ll 1, \quad 2 \le j \le P. \end{aligned}$$

SDAF algorithm approximates covariance matrix, $R_{m+1|m}$, by block diagonal covariance matrix

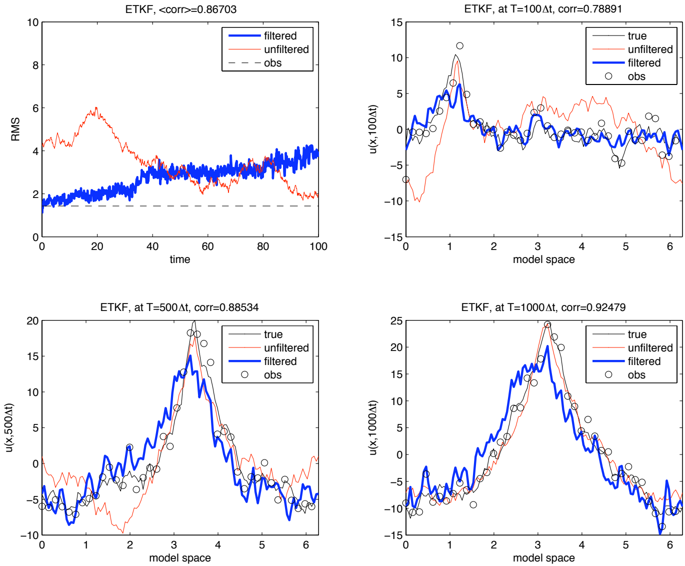
$$R_{m+1|m} = \left(\begin{array}{c|cccc} R_{k_1,m+1|m} & & & 0 & \\ \hline & R_{k_2} & & & \\ & & R_{k_3} & & \\ 0 & & & \ddots & \\ & & & & R_{k_P} \end{array} \right)$$

Decorrelation time vs observation time:

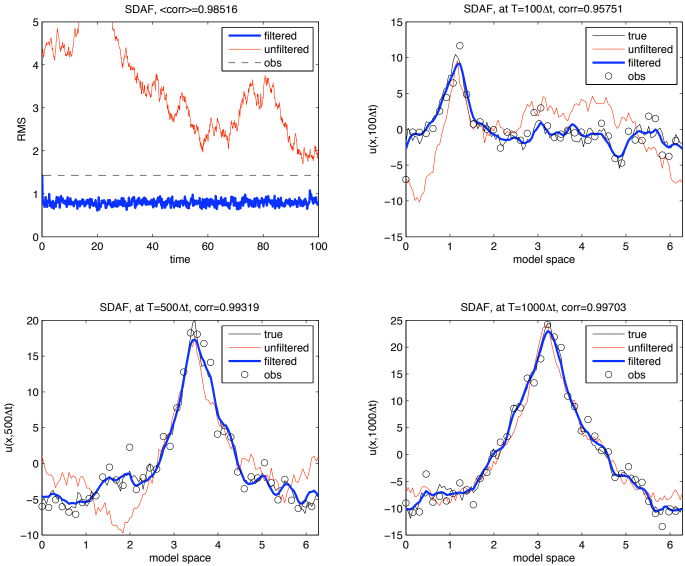


ETKF Filter Divergence (Ensemble size $= 150 > N = 123$)

Extreme event, $\Delta t_2 = 0.1$, $E_k = k^{-5/3}$, $P = 3$, $r^o = 2.05$



SDAF high skill Spontaneous development of extreme event for $\Delta t_2 = 0.1$ and $E_k = k^{-5/3}$, $P = 3$, $r^o = 2.05$



In the previous numerical simulation, the results of FDKF and SDAF yield comparable accuracy, this is true only for -5/3 spectrum and the given observation time.

We also find that when observation time is shortened and for equipartition spectrum, FDKF is better than SDAF.

Besides numerical advantage over FDKF, what is the significant of SDAF?

Can SDAF avoid singularity due to violation of observability?

- **Filter stability** (Kalman, 1960)
  Consider the following discrete time linear filtering problem

$$\begin{aligned} u_{m+1} &= F u_m + \sigma_{m+1} + \bar{F}_{m+1} \\ v_{m+1} &= G u_{m+1} + \sigma^o_{m+1} \end{aligned}$$

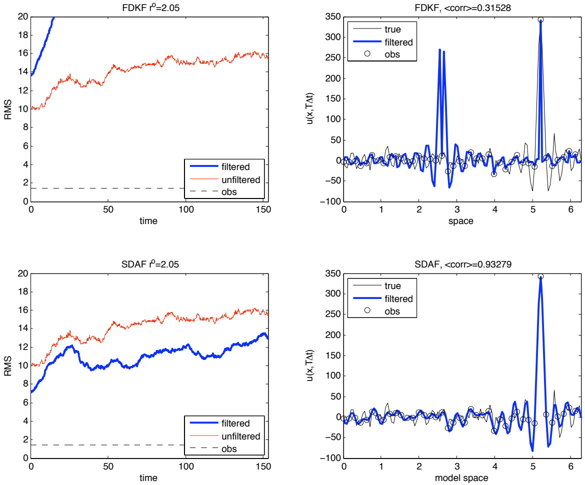
  The stability of the filter is guaranteed when $|F| < 1$; for $|F| \geq 1$ stability is satisfied if the pair of operators F, G is observable.

- **Observability:** Matrix $[G, GF, \ldots, GF^{p-1}]$ has a full rank.
- In our example, observability in each aliasing subspace with $G = [1, \ldots, 1]$ requires

$$\prod_{i \neq j} (F_{k_i} - F_{k_j}) \neq 0,$$

  where $F_{k_i} = \mathrm{e}^{-(\mu k_i^2 + \mathrm{i} k_i)\Delta t}$ for adv-diff eqn and $F_{k_i} = \mathrm{e}^{-(d + \mathrm{i} k_i)\Delta t}$ for uniformly adv eqn.

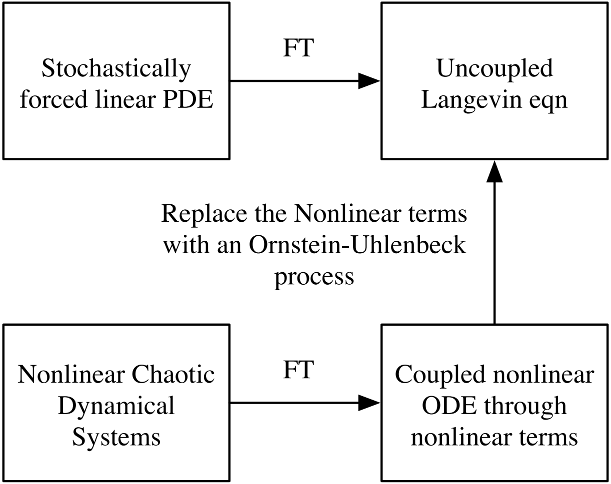
Weakly uniformly damped advection equation $E_k = 1$,
Non-observable time $\Delta t = 2\pi/(2M+1) = 0.15325$, resonantly forced signal; high skill of SDAF over FDKF



Summary of Part I:

- Given $(2M+1) = (2N+1)/P$ sparse regular observations, FDKF reduces a (2N+1)-dim filtering problem to M decoupled P-dim filtering problems, each with a scalar observation.
- In our assessment, we find that filtering sparsely observed linear problem with ensemble Kalman filter with ensemble size larger than the model dimensionality does not guaranteed convergence solution.
- FDKF suggests that ignoring the cross covariance between different aliasing sets is not only computationally advantageous but it also produces more accurate solutions.
- Intuitively, this reduced filter avoid the spurious correlations between different wave numbers.
- Observability violation can be avoided by further reduction through strong damping approximation even when the high mode in the true signal are not strongly damped.

Radical Filtering Strategy for Nonlinear System

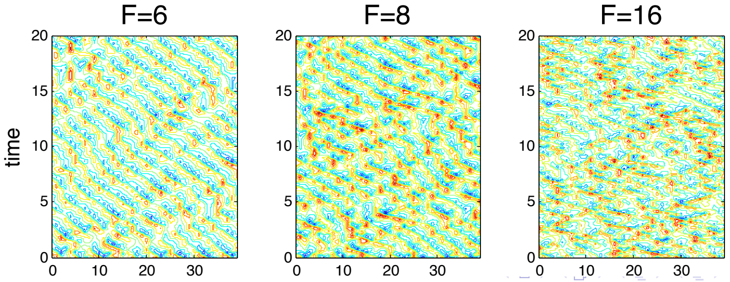


Filtering turbulent nonlinear dynamical systems

L-96 model (Lorenz 1996), 40-dim, "absorbing ball property".

$$\frac{du_j}{dt} = (u_{j+1} - u_{j-2})u_{j-1} - u_j + F, \quad j = 0, \ldots, J-1$$

	F	λ_1	N^+	KS	T_{corr}
Weakly chaotic	6	1.02	12	5.547	8.23
Strongly chaotic	8	1.74	13	10.94	6.704
Fully turbulent	16	3.945	16	27.94	5.594



References:

a) Test Models for Improving Filtering with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):1-31, 2010.

b) Improving Filtering and Prediction of Spatially Extended Turbulent Systems with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):32-57, 2010.

c) Filtering Turbulent Sparsely Observed Geophysical Flows (with A.J. Majda), Monthly Weather Review, 138(4): 1050-1083, 2010.

Review Article: Mathematical Strategies for Filtering Turbulent Dynamical Systems (with B. Gershgorin and A.J. Majda), DCDS-A: 27(2), 441-486, 2010.

Ch 13 of: Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems (with A.J. Majda), Cambridge University Press (in preparation), 2010.

Online Model Error Estimation Strategy

A simple strategy to cope with model errors for filtering with an imperfect model nonlinear dynamical system depending on parameters, b,

$$\frac{du}{dt} = F(u, b)$$

is to augment the state variable u, by the parameters λ, and adjoin an approximate dynamical equation for the parameters

$$\frac{db}{dt} = g(b).$$

In hierarchical notation, filtering this augmented system is one way of estimating

$$[u, b|v] = [v|u, b]\frac{[u|b][b]}{[v]}$$

Test model for true signal

Consider the following SDE

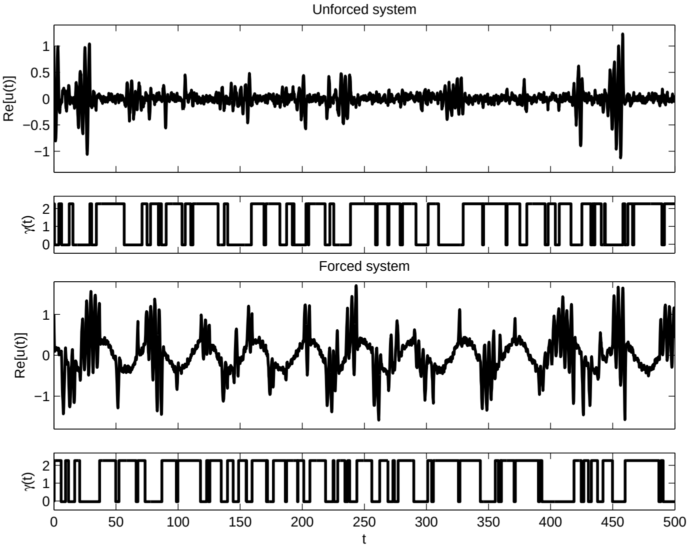
$$\frac{du(t)}{dt} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma\dot{W}(t) + f(t)$$

as a test model for filtering with model error.

To generate significant model errors as well as to mimic intermittent chaotic instability as often occurs in nature, we allow $\gamma(t)$ to switch between stable ($\gamma > 0$) and unstable ($\gamma < 0$) regimes according to a two-state Markov jump process.
Assume the following observation model:

$$v_m = u(t_m) + \sigma_m^o, \quad \sigma_m^o \sim \mathcal{N}(0, r^o). \tag{1}$$

True Signals for Unforced and Forced cases



Mean Stochastic Model

The prototype one-mode stochastic mean model

$$du(t) = \Big[(-\bar{\gamma} + \mathrm{i}\omega)u(t) + F(t)\Big]dt + \sigma dW(t)$$

where one fits the parameters using climatological statistical quantities such as the energy spectrum and correlation time.

This "poor-man" strategy is discussed in Harlim and Majda Nonlinearity 2008, Comm. Math. Sci. 2010.

Stochastic Parameterized Extended Kalman Filter:

We consider the following canonical model that accounts additive and multiplicative biases:

$$
\begin{aligned}
du(t) &= \Big[(-\gamma(t) + \mathrm{i}\omega)u(t) + F(t) + b(t)\Big]dt + \sigma dW(t) \\
db(t) &= (-\gamma_b + \mathrm{i}\omega_b)b(t)dt + \sigma_b dW_b(t) \\
d\gamma(t) &= -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t)
\end{aligned}
$$

We find stochastic parameters $\{\gamma_b, \omega_b, \sigma_b, d_\gamma, \sigma_\gamma\}$ that are robust for high filter skill beyond the MSM and in many occasions comparable to the perfectly specified filter model.

This special form has exactly solvable nonlinear solutions and moments and we do not need any linearization as in the standard EKF.

Next, we find the mean $\langle u(t) \rangle$: (Use the calculus tricks in Gershgorin-Majda 2008, 2010)

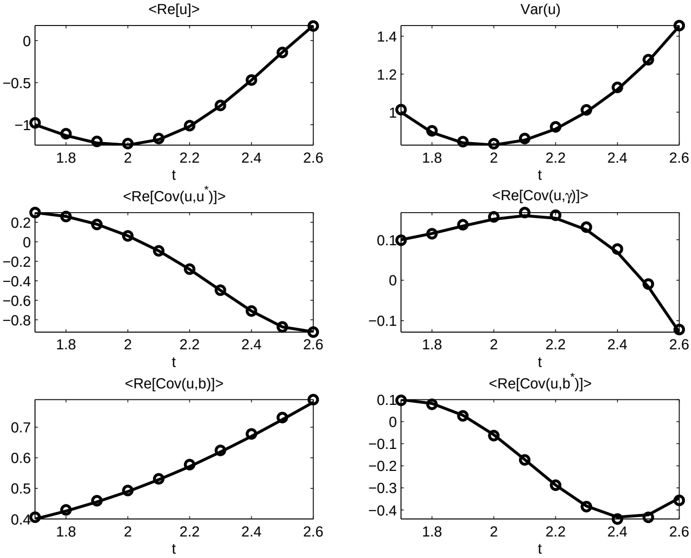
$$
\begin{aligned}
\langle u(t) \rangle &= e^{\hat{\lambda}(t-t_0)}\Big(\langle u_0 \rangle - Cov(u_0, J(t_0,t))\Big) e^{-\langle J(t_0,t)\rangle + \frac{1}{2}Var(J(t_0,t))} \\
&+ \int_{t_0}^{t} e^{\hat{\lambda}(t-s)}\Big(\hat{b} + e^{\lambda_b(s-t_0)}\Big(\langle b_0 \rangle - \hat{b} - Cov(b_0, J(s,t))\Big)\Big) \\
&\times e^{-\langle J(s,t)\rangle + \frac{1}{2}Var(J(s,t))} ds \\
&+ \int_{t_0}^{t} e^{\hat{\lambda}(t-s)} f(s) e^{-\langle J(s,t)\rangle + \frac{1}{2}Var(J(s,t))} ds \qquad (2)
\end{aligned}
$$

where

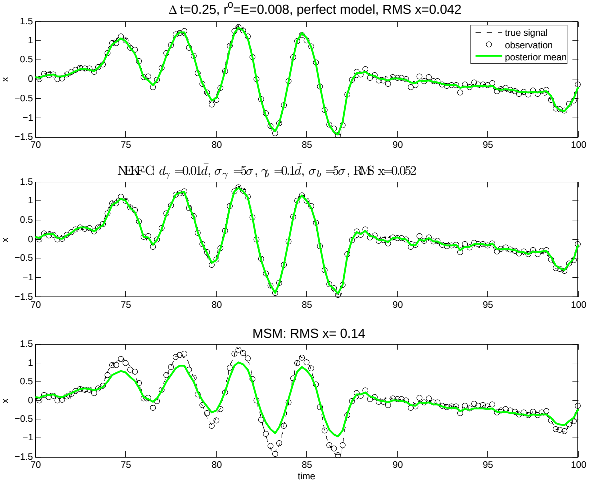
$$
\begin{aligned}
\hat{\lambda} &= -\hat{\gamma} + \mathrm{i}\omega, \\
J(s,t) &= \int_{s}^{t} (\gamma(s') - \hat{\gamma}) ds',
\end{aligned}
$$

and next the cross-covariances ...

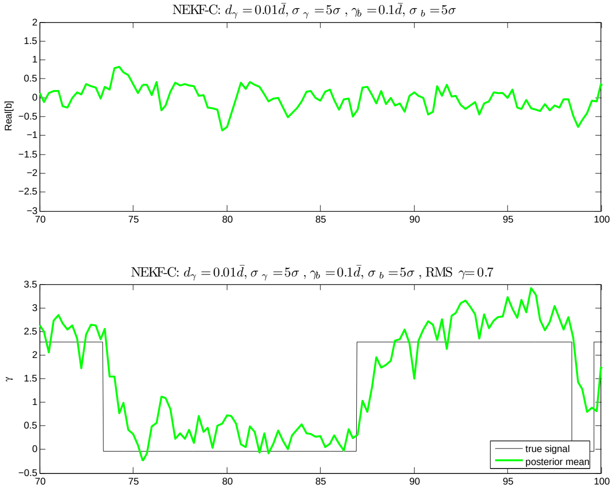
SPEKF: Checking first and second ordered statistics



One mode demonstration of the filtered solution: observed mode



One mode demonstration of the filtered solution: unobserved parameters



Canonical Spatially Extended Turbulent Systems

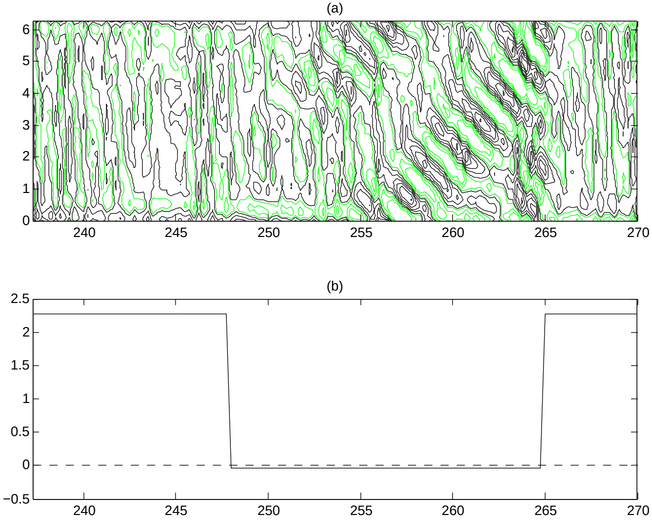
We consider a stochastic PDE with time-dependent damping Langevin equation for the first five Fourier modes, i.e.,

$$\frac{du_k(t)}{dt} = -\gamma_k(t)u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k = 1, \dots, 5,$$

and linear Langevin equation with constant damping $\bar{d}$ for modes $k > 5$,

$$\frac{du_k(t)}{dt} = -\bar{d}u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k > 5.$$

Turbulent barotropic Rossby wave equation: $\omega_k = -\beta/k, E_k = k^{-3}$



Incorrectly specified forcings:

Here, we consider a true signal with forcing given by

$$\hat{f}_k(t) = A_{f,k} \exp\Big(\mathrm{i}(\omega_{f,k} t + \phi_{f,k})\Big), \tag{3}$$

for $k = 1, \ldots, 7$ with amplitude $A_{f,k}$, frequency $\omega_{f,k}$, and phase $\phi_{f,k}$ drawn randomly from uniform distributions,

$$\begin{aligned} A_{f,k} &\sim U(0.6, 1), \\ \omega_{f,k} &\sim U(0.1, 0.4), \\ \phi_{f,k} &\sim U(0, 2\pi), \\ \hat{f}_k &= \hat{f}^*_{-k}, \end{aligned}$$

and unforced, $\hat{f}_k(t) = 0$, for modes $k > 7$. However, we do not specify this true forcing to the filter model, i.e., we use $\tilde{f}_k = 0$ for all modes.

Reduced Filter Domain Kalman Filter for regularly spaced sparse observation

We consider regularly spaced sparse observations: $(2M+1)$ observations of $(2N+1)$ model grid points. The Fourier coefficients of the observation model is given as

$$\hat{v}_{\ell,m} = \sum_{k\in\mathcal{A}(\ell)} \hat{u}_{k,m} + \hat{\sigma}^o_m,$$

where

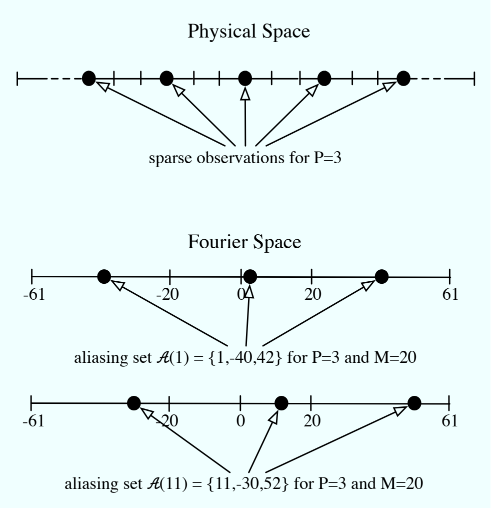
$$\mathcal{A}(\ell) = \{k | k = \ell + (2M+1)q, q \in \mathbb{Z}, |\ell| \leq N\}$$

is the aliasing set of wavenumber ℓ. (Majda-Grote PNAS 2007)

When the energy spectrum is decaying as a function of k, we can use the following reduced observation model

$$\hat{v}'_{\ell,m} \equiv \hat{v}_{\ell,m} - \sum_{k\in\mathcal{A}(\ell), k\neq\ell} \hat{u}_{k,m|m-1} = \hat{u}_{\ell,m} + \hat{\sigma}^o_m.$$

Example: 123 grid pts (61 modes) but only 41 observations (20 modes) available



Incorrectly specified forcings, observed only 15 observations of 105 grid points

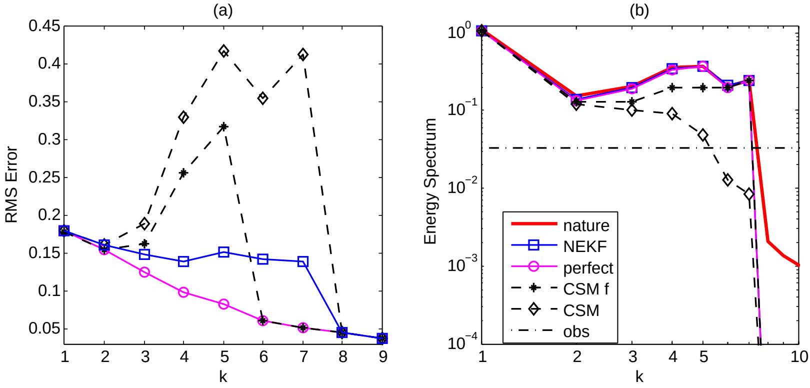


Table of RMSE for the SPDE test case with intermittent burst of instability

Forcing	unforced case	correct forcing	incorrect forcing
r^o	0.2	0.3	0.5
perfect filter	0.35	0.39	0.45
MSM	0.39	0.48	0.73
$\text{MSM}_{f=0}$	-	-	1.17
SPEKF-C	0.38	0.44	0.59
SPEKF-M	0.36	0.42	0.79
SPEKF-A	0.39	0.46	0.60

Canonical Model for Midlatitude Geophysical Flows:

The dynamical equations for the perturbed variables are:

$$\begin{aligned} \frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} + \nu\nabla^8 q_1 &= 0 \\ \frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial x} + \nu\nabla^8 q_2 + \kappa\nabla^2\psi_2 &= 0 \end{aligned}$$

where q_j is the quasi-geostrophic potential vorticity given as

$$q_j = \beta y + \nabla^2\psi_j + \frac{k_d^2}{2}(\psi_{3-j} - \psi_j)$$

with $\vec{u} = \nabla^{\perp}\psi$, $k_d = \sqrt{8}/L_d$.

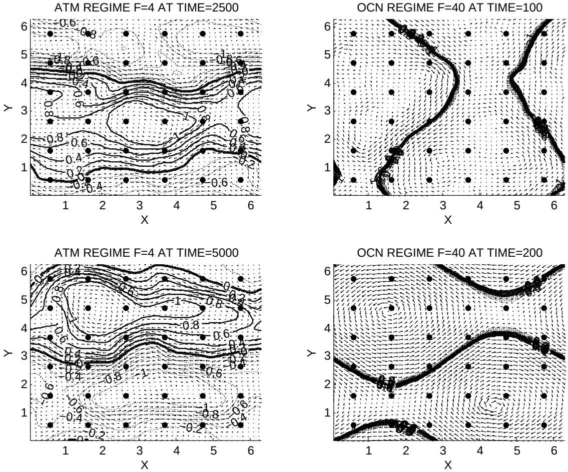
In the two-layer case, the barotropic vertical and baroclinic modes are defined as $\psi_b = (\psi_1 + \psi_2)/2$ and $\psi_c = (\psi_1 - \psi_2)/2$, respectively.

Notice that the barotropic mode dynamical equation,

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \left(J(\psi_c, q_c) + U \frac{\partial \nabla^2 \psi_c}{\partial x} - \kappa \nabla^2 \psi_c \right) = 0$$

is numerically stiff when k_d^2 is large (ocean case).

The 2-layer QG model with baroclinic instability



Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2 \psi_b$.

Poorman's stochastic models: replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes.

Discrete Fourier Transform:

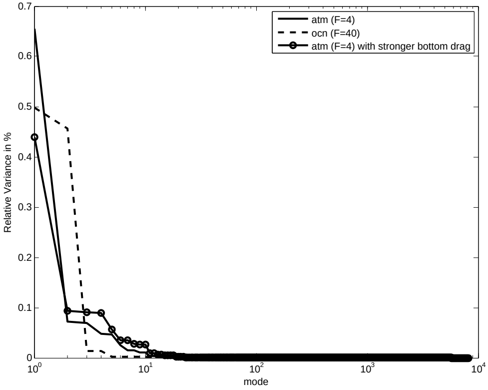
$$\psi = \sum_{k,\ell} \hat{\psi}_{k,\ell} e^{\mathrm{i}(kx+\ell y)}$$

Thus, each horizontal mode has the following form

$$d\hat{\psi}(t) = (-d + \mathrm{i}\omega)\hat{\psi}(t)dt + f(t)dt + \sigma dW(t)$$

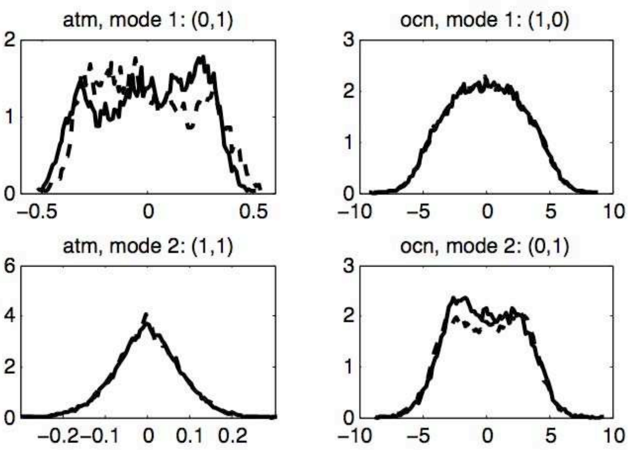
and our task is to parameterize $d, \omega, f(t), \sigma$?

Statistical Quantities: Climatological variances of the barotropic mode

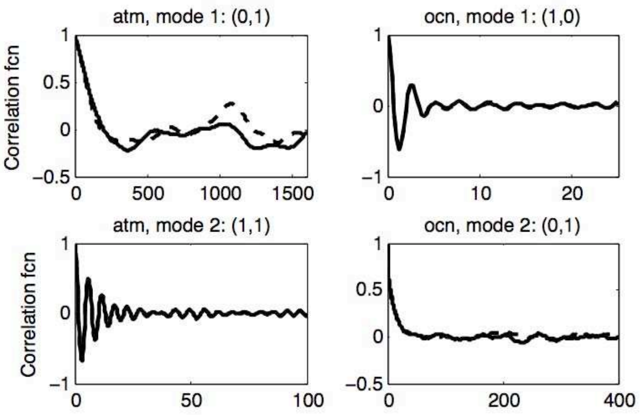


"Atmospheric" case (k_d^2 is small) and "oceanic" case (k_d^2 is large).

Statistical Quantities: Histogram "marginal pdf's"



Statistical Quantities: Correlation functions



Mean Stochastic Models: parameterize d, ω, f, σ

We set $f(t)$ to be a constant equals to the climatological mean $\langle \hat{\psi} \rangle$ (long time average).

MSM2 We use the linear dispersion ω, and we fit the damping and noise strengths to the spectrum and decorrelation time

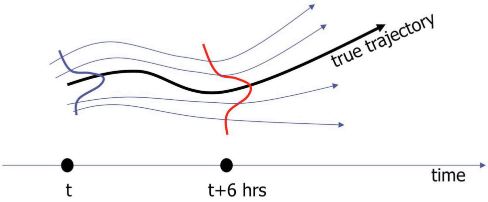
$$
\begin{aligned}
Var(\hat{\psi}) &= \frac{\sigma^2}{2d} \\
Re[T_{corr}] &\equiv \frac{1}{Var(\hat{\psi})} \int_0^\infty Re[C(\tau)] d\tau = \frac{1}{d}
\end{aligned}
$$

MSM1 Ignore the linear dispersion and solve the following

$$
\begin{aligned}
Var(\hat{\psi}) &= \frac{\sigma^2}{2d} \\
T_{corr} &\equiv \frac{1}{Var(\hat{\psi})} \int_0^\infty C(\tau) d\tau = \frac{1}{d + \mathrm{i}\omega}
\end{aligned}
$$

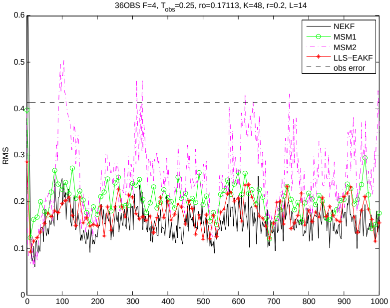
Local least squares EAKF (Anderson 2003)

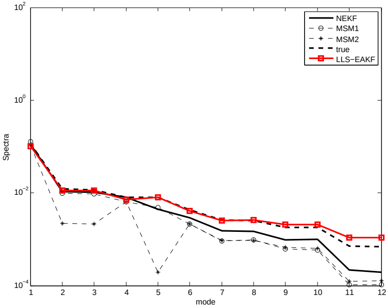
Approximate the prior error covariance matrix by ensemble covariance.

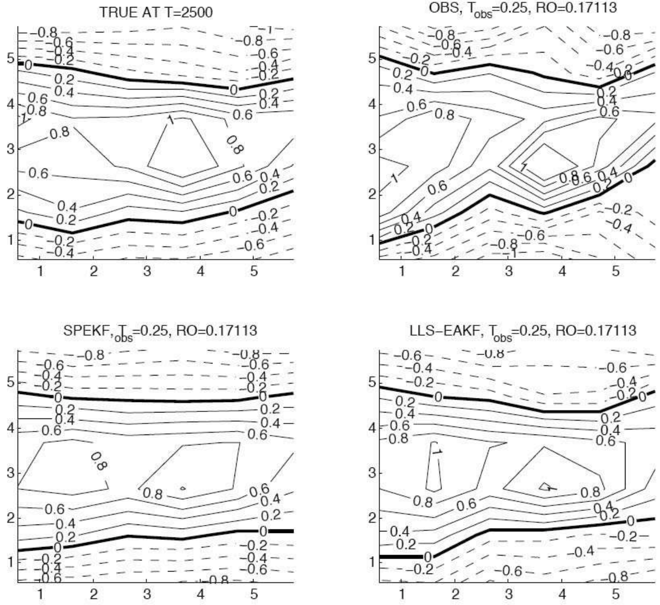


How many ensemble member? How to avoid ensemble collapse and spurious correlations due to finite ensemble size? Computationally, EAKF requires extensive tunings of ensemble size, local box size, covariance inflation, and in the ocean case, integration time step need to be reduced.

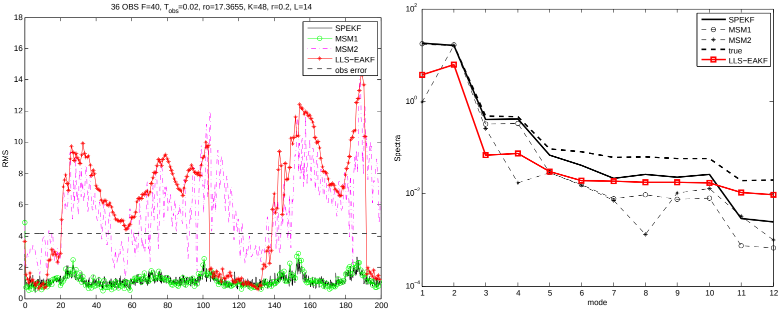
Longer deformation radius case ("atmospheric" regime).

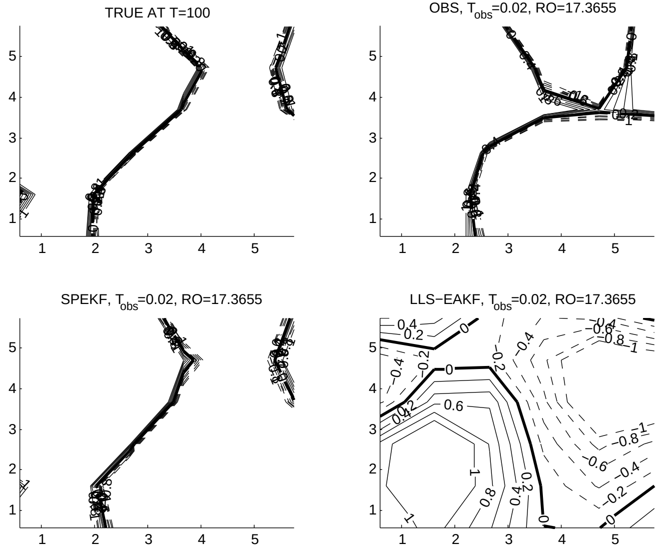





TRUE AT T=2500
OBS, T_{obs}=0.25, RO=0.17113
SPEKF, T_{obs}=0.25, RO=0.17113
LLS−EAKF, T_{obs}=0.25, RO=0.17113

Shorter deformation radius case ("oceanic" regime).

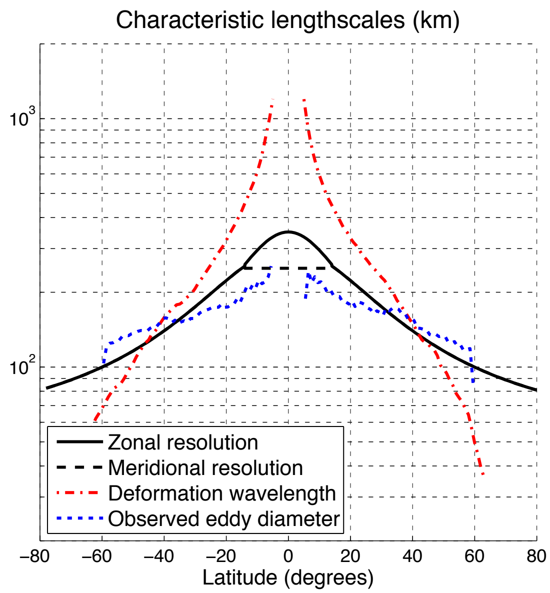



TRUE AT T=100
OBS, T_obs=0.02, RO=17.3655
SPEKF, T_obs=0.02, RO=17.3655
LLS−EAKF, T_obs=0.02, RO=17.3655

Summary:

1. MSM: We introduce reduced stochastic models through replacing the nonlinearity and baroclinic components with Ornstein-Uhlenbeck process for filtering purpose. This reduced poor man's strategy is numerically very cheap and accurate in a regime when the dynamical systems is strongly chaotic and fully turbulent.
2. SPEKF: We introduce a paradigm model for "online" learning both the additive and multiplicative biases from observations beyond the MSM. This model is analytically solvable such that NO LINEARIZATION is needed when Kalman filter formula is utilized.

Stochastic Super-resolution: Estimating turbulent heat transport in the ocean using satellite altimetry

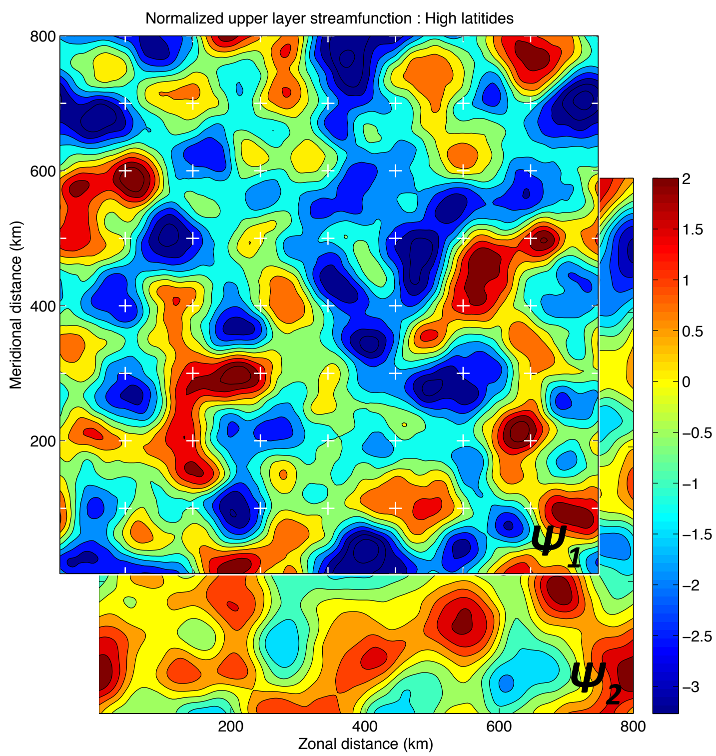


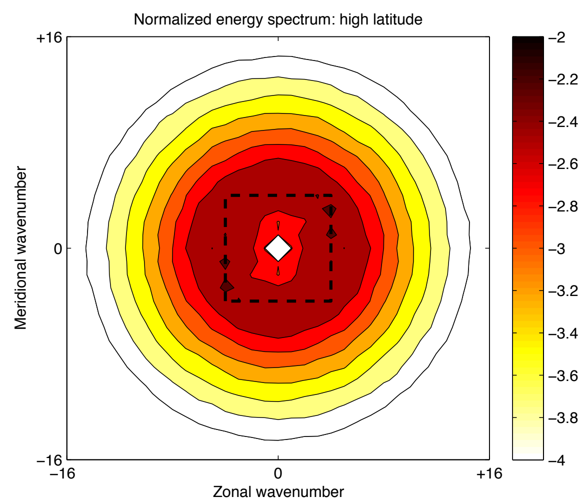
"New methods for estimating poleward eddy heat transport using satellite altimetry"

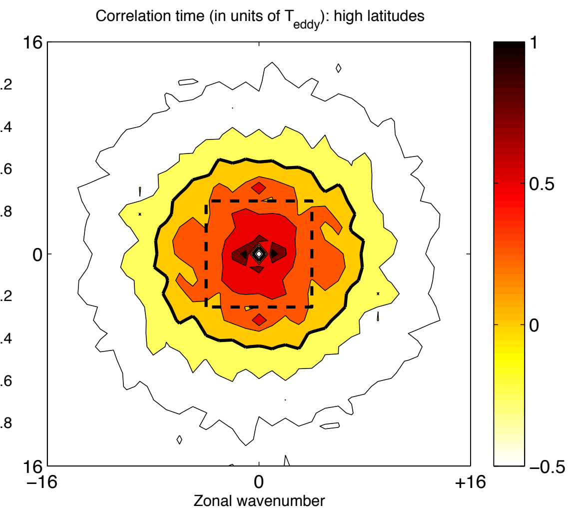
Shane R. Keating, **Andrew J. Majda** & K. Shafer Smith

J. Phys. Oceanogr. 2011 (submitted)

Estimating turbulent heat transport



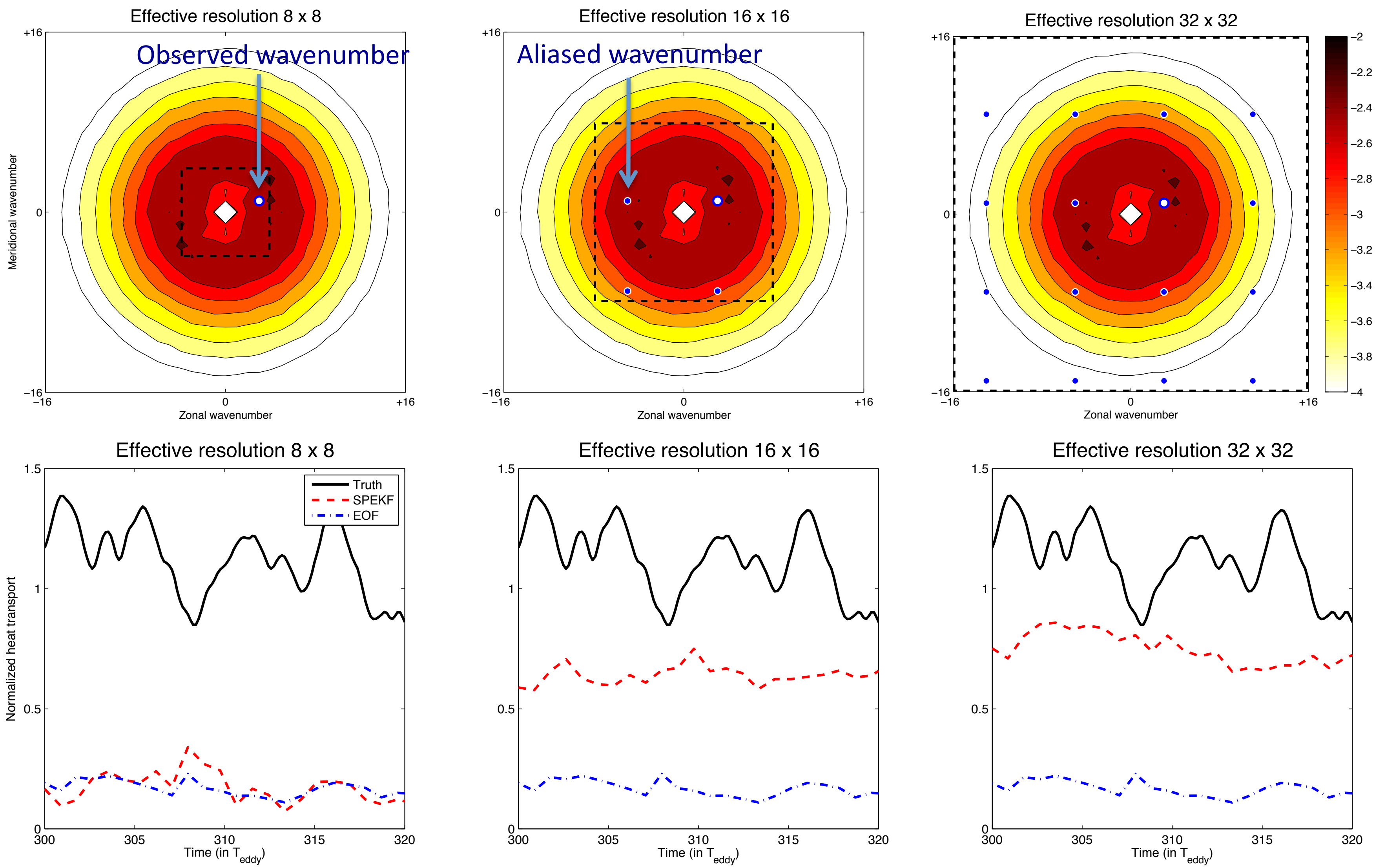




Key challenges:

- High dimensionality
- Sparseness in **vertical** space
- Sparseness in **horizontal** space
- Sparseness in **time**

Stochastic super-resolution



$$\langle v_1 \tau \rangle = -\sqrt{d_1 d_2} \langle v_1 \psi_2 \rangle$$