Dynamic Data-Driven Adaptive Observations in Data Assimilation for Multiscale Systems

AFOSR DDDAS Program Review Meeting, Dayton OH

PI: N. Sri Namachchivaya
Ryne Beeson, Hoong Chieh Yeong,
Nicolas Perkowski* and Peter Sauer†

Department of Aerospace Engineering & Information Trust Institute
University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

September 7th, 2017

* *Applied Mathematics, Humboldt-Universität zu Berlin*
† *Electrical and Computer Engineering, University of Illinois at Urbana-Champaign*

Research Objectives

(I) Develop efficient and robust methods to produce *lower-dimensional recursive nonlinear filtering* equations driven by the observations; particle filters for the integration of observations with the simulations of large-scale complex systems.

(II) Develop an *integrated framework* that combines the ability to *dynamically steer* the measurement process, extracting useful information, with *nonlinear filtering* for inference and prediction of large scale complex systems.

Table of Contents

Motivating Problems and Characteristics

Problems are,

(i) Multiscale
(ii) Chaotic
(iii) High dimensional
(iv) Sparse observations
(v) Ability for sensor selection, placement and control - adaptive observation
(vi) Sensors correlated to environment

Motivating problems,

(a) Weather prediction and forecasting
(b) Detection and tracking of contaminants in the environment (e.g. chemical and radioactive)

Estimation and prediction in Earth (climate) system models

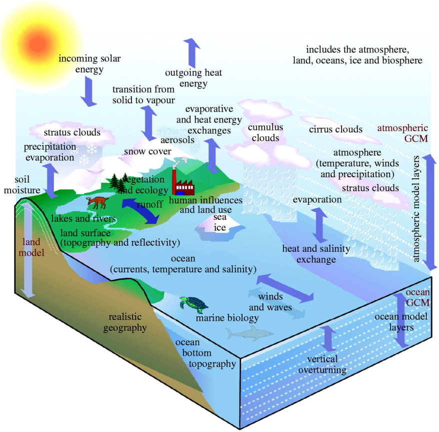


Figure: NCAR Community Climate System Model [1]

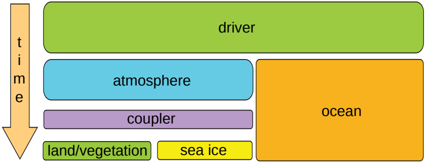


Figure: Coupling components in climate model [NCAR] [1]

Simple multiscale example: Lorenz-Maas model

Coupled equations [2], [3], [4] :

$$
\begin{aligned}
\frac{d\rho_x}{dt} &= -\rho_y L_z + (\rho_x + f'\rho_y)\rho_z - \rho_x \\
\frac{d\rho_y}{dt} &= \rho_x L_z - (f'\rho_x - \rho_y)\rho_z - \rho_y \\
&\quad + k_1 q + k_2(x - k_3\rho_y) \\
\frac{d\rho_z}{dt} &= -\rho_x^2 - \rho_y^2 - \mu\rho_z \\
\frac{\epsilon_2}{\epsilon_3}\frac{dL_z}{dt} &= -L_z - k_4 x \\
\epsilon_2\frac{dx}{dt} &= -y^2 - z^2 - ax + aF_0 \\
&\quad + \epsilon_1(k_3\rho_y - x) \\
\epsilon_2\frac{dy}{dt} &= xy - bxz - y + G \\
\epsilon_2\frac{dz}{dt} &= bxy + xz - z
\end{aligned}
$$

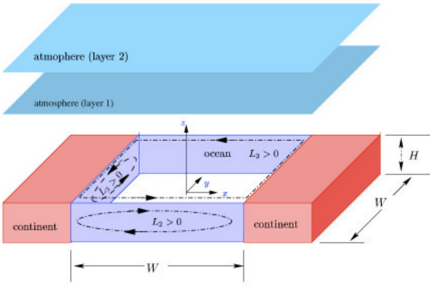


Figure: Coupled Lorenz 1984 atmosphere – Maas 2004 ocean models

Simple multiscale example: Lorenz-Maas model

Coupled equations [2], [3], [4] :

$$\frac{d}{dt}X = b(X, L_z, Z_1)$$
$$\epsilon \frac{d}{dt}L_z = -L_z - k_4 Z_1$$
$$\epsilon^2 \frac{d}{dt}Z = f_0(Z) + \epsilon f_1(X_2, Z_1),$$

where ϵ is $\mathcal{O}(10^{-2})$.

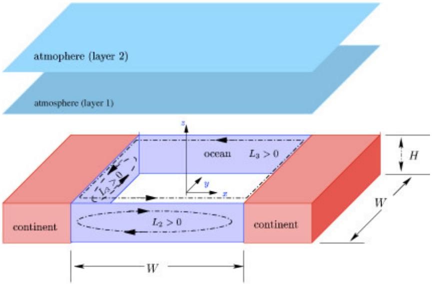
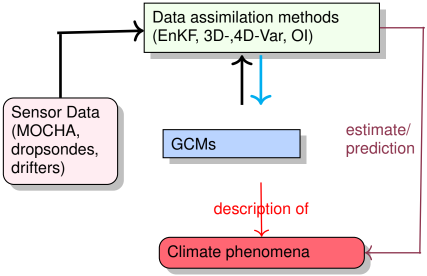


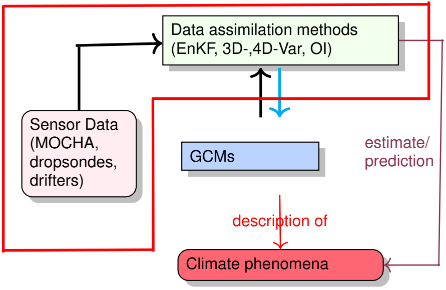
Figure: Coupled Lorenz 1984 atmosphere – Maas 2004 ocean models

Estimation and prediction in Earth (climate) system models



- Weather phenomena can be studied through estimation/prediction using GCMs.
- GCMs can be improved using data assimilation results.
- Filtering theory provides rigorous approach to quantifying probabilistic information - as opposed to methods such as 3D-Var, 4D-Var, and OI.

Estimation and prediction in Earth (climate) system models



- Weather phenomena can be studied through estimation/prediction using GCMs.
- GCMs can be improved using data assimilation results.
- Filtering theory provides rigorous approach to quantifying probabilistic information - as opposed to methods such as 3D-Var, 4D-Var, and OI.

Mobile Platforms, Adaptive Observations, and Correlated Noise

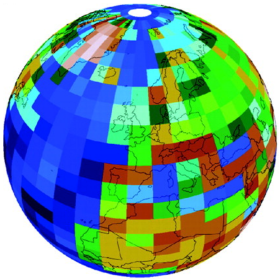


Figure: NCAR Globe, 500km Sectors

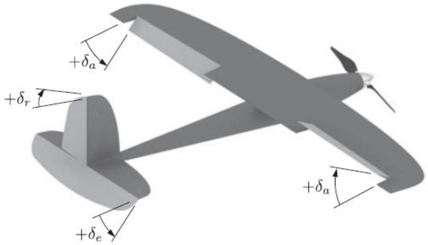


Figure: UAV Flight Controls [5]

1. Sensor Selection / Placement
2. Sensor Control
3. Mobile Platforms are Embedded in Signal Environment → Correlated Noise
4. Require Efficient and Robust Filtering Methods for Multiscale Correlated Case

Table of Contents

Multiscale Correlated Noise Problem Setup

Let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geqslant 0}, \mathbb{Q})$ be a probability space upon which the following SDEs are defined:

$$dX_t^\epsilon = \left[b(X_t^\epsilon, Z_t^\epsilon) + \frac{1}{\epsilon} b_1(X_t^\epsilon, Z_t^\epsilon)\right] dt + \sigma(X_t^\epsilon, Z_t^\epsilon) dW_t \qquad X_0^\epsilon = x$$

$$dZ_t^\epsilon = \frac{1}{\epsilon^2} f(X_t^\epsilon, Z_t^\epsilon) dt + \frac{1}{\epsilon} g(X_t^\epsilon, Z_t^\epsilon) dV_t \qquad Z_0^\epsilon = z$$

$$dY_t^\epsilon = h(X_t^\epsilon, Z_t^\epsilon) dt + \alpha dW_t + \beta dV_t + \gamma dU_t \qquad Y_0^\epsilon = 0$$

$$= h(X_t^\epsilon, Z_t^\epsilon) dt + dB_t$$

1. W_t, V_t and U_t are independent standard BM under $\mathbb{Q}$
2. $0 < \epsilon \ll 1$ is the time-scale separation
3. w.l.o.g., let $\alpha^2 + \beta^2 + \gamma^2 = 1$ and define the standard BM

$$B_t \equiv \alpha W_t + \beta V_t + \gamma U_t$$

The Nonlinear Filter

The objective in filtering theory is to obtain a solution for the *normalized conditional measure* - the filter,

Normalized Conditional Measure

$$\pi_t^\epsilon(\varphi(X_t^\epsilon, Z_t^\epsilon)) \equiv \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right],$$

where $\varphi(X_t^\epsilon, Z_t^\epsilon)$ is an integrable function and

$$\mathcal{Y}_t^\epsilon \equiv \sigma(\{Y_s^\epsilon - Y_0^\epsilon \,|\, s \in [0, t]\}).$$

1. Density equivalent of π_t^ϵ satisfies a high dimensional SPDE: "Curse of Dimensionality".
2. If $\varphi = \varphi(X_t^\epsilon)$ and $X_t^\epsilon \Rightarrow X_t^0$ as $\epsilon \to 0$, does there exists $\pi_t^\epsilon \to \pi_t^0$?
3. Proof and insight will enable improvement of nonlinear filtering algorithms for multiscale correlated noise case.
4. *$X_t^\epsilon \Rightarrow X_t^0$ does not imply $\pi_t^\epsilon \to \pi_t^0$*

The Nonlinear Filter

The objective in filtering theory is to obtain a solution for the *normalized conditional measure* - the filter,

Normalized Conditional Measure

$$\pi_t^\epsilon(\varphi(X_t^\epsilon, Z_t^\epsilon)) \equiv \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right],$$

where $\varphi(X_t^\epsilon, Z_t^\epsilon)$ is an integrable function and

$$\mathcal{Y}_t^\epsilon \equiv \sigma(\{Y_s^\epsilon - Y_0^\epsilon \,|\, s \in [0, t]\}).$$

1. Density equivalent of π_t^ϵ satisfies a high dimensional SPDE: "Curse of Dimensionality".
2. If $\varphi = \varphi(X_t^\epsilon)$ and $X_t^\epsilon \Rightarrow X_t^0$ as $\epsilon \to 0$, does there exists $\pi_t^\epsilon \to \pi_t^0$?
3. Proof and insight will enable improvement of nonlinear filtering algorithms for multiscale correlated noise case.
4. *$X_t^\epsilon \Rightarrow X_t^0$ does not imply $\pi_t^\epsilon \to \pi_t^0$*

The Nonlinear Filter

The objective in filtering theory is to obtain a solution for the *normalized conditional measure* - the filter,

Normalized Conditional Measure

$$\pi_t^\epsilon(\varphi(X_t^\epsilon, Z_t^\epsilon)) \equiv \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right],$$

where $\varphi(X_t^\epsilon, Z_t^\epsilon)$ is an integrable function and

$$\mathcal{Y}_t^\epsilon \equiv \sigma(\{Y_s^\epsilon - Y_0^\epsilon \,|\, s \in [0, t]\}).$$

1. Density equivalent of π_t^ϵ satisfies a high dimensional SPDE: “Curse of Dimensionality”.
2. If $\varphi = \varphi(X_t^\epsilon)$ and $X_t^\epsilon \Rightarrow X_t^0$ as $\epsilon \to 0$, does there exists $\pi_t^\epsilon \to \pi_t^0$?
3. Proof and insight will enable improvement of nonlinear filtering algorithms for multiscale correlated noise case.
4. *$X_t^\epsilon \Rightarrow X_t^0$ does not imply $\pi_t^\epsilon \to \pi_t^0$*

The Nonlinear Filter

The objective in filtering theory is to obtain a solution for the *normalized conditional measure* - the filter,

Normalized Conditional Measure

$$\pi_t^\epsilon(\varphi(X_t^\epsilon, Z_t^\epsilon)) \equiv \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \mid \mathcal{Y}_t^\epsilon\right],$$

where $\varphi(X_t^\epsilon, Z_t^\epsilon)$ is an integrable function and

$$\mathcal{Y}_t^\epsilon \equiv \sigma(\{Y_s^\epsilon - Y_0^\epsilon \mid s \in [0, t]\}).$$

1. Density equivalent of π_t^ϵ satisfies a high dimensional SPDE: "Curse of Dimensionality".
2. If $\varphi = \varphi(X_t^\epsilon)$ and $X_t^\epsilon \Rightarrow X_t^0$ as $\epsilon \to 0$, does there exists $\pi_t^\epsilon \to \pi_t^0$?
3. Proof and insight will enable improvement of nonlinear filtering algorithms for multiscale correlated noise case.
4. *$X_t^\epsilon \Rightarrow X_t^0$ does not imply $\pi_t^\epsilon \to \pi_t^0$*

Mathematical Tools and Proof of Convergence Approach

1. Introduce an unnormalized conditional measure

Unnormalized Conditional Measure ρ_t^ϵ

$$\frac{\rho_t^\epsilon(\varphi)}{\rho_t^\epsilon(1)} = \frac{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\varphi(X_t^\epsilon, Z_t^\epsilon)\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]}{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]} = \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right] = \pi_t^\epsilon(\varphi)$$

2. Introduce function valued dual process, v^ϵ, satisfying a BSPDE
3. Ansatz v^0, ρ^0, π^0
4. Asymptotic expansion of $v^\epsilon = v^0 + \psi + R$; ψ the corrector, and R the remainder
5. Utilize homogenization estimates [6], estimates for BSPDE [7] and Feynman-Kac representations with FBDSDE [8] to prove convergence,

Dual Convergence Implies Filter Convergence

$$\mathbb{E}\left[\left|\rho_T^{\epsilon,x}(\varphi) - \rho_T^0(\varphi)\right|^p\right] \leqslant \int_{\mathbb{R}^2} \mathbb{E}\left[\left|v_0^\epsilon(x,z) - v_0^0(x)\right|^p\right] \mathbb{Q}_{X_0^\epsilon, Z_0^\epsilon}(dx, dz)$$

Mathematical Tools and Proof of Convergence Approach

1. Introduce an unnormalized conditional measure

Unnormalized Conditional Measure ρ_t^ϵ

$$\frac{\rho_t^\epsilon(\varphi)}{\rho_t^\epsilon(1)} = \frac{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\varphi(X_t^\epsilon, Z_t^\epsilon)\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]}{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]} = \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \mid \mathcal{Y}_t^\epsilon\right] = \pi_t^\epsilon(\varphi)$$

2. Introduce function valued dual process, v^ϵ, satisfying a BSPDE
3. Ansatz v^0, ρ^0, π^0
4. Asymptotic expansion of $v^\epsilon = v^0 + \psi + R$; ψ the corrector, and R the remainder
5. Utilize homogenization estimates [6], estimates for BSPDE [7] and Feynman-Kac representations with FBDSDE [8] to prove convergence,

Dual Convergence Implies Filter Convergence

$$\mathbb{E}\left[\left|\rho_T^{\epsilon,x}(\varphi) - \rho_T^0(\varphi)\right|^p\right] \leqslant \int_{\mathbb{R}^2} \mathbb{E}\left[\left|v_0^\epsilon(x,z) - v_0^0(x)\right|^p\right] \mathbb{Q}_{X_0^\epsilon, Z_0^\epsilon}(dx, dz)$$

Mathematical Tools and Proof of Convergence Approach

1. Introduce an unnormalized conditional measure

Unnormalized Conditional Measure ρ_t^ϵ

$$\frac{\rho_t^\epsilon(\varphi)}{\rho_t^\epsilon(1)} = \frac{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\varphi(X_t^\epsilon, Z_t^\epsilon)\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]}{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]} = \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right] = \pi_t^\epsilon(\varphi)$$

2. Introduce function valued dual process, v^ϵ, satisfying a BSPDE
3. Ansatz v^0, ρ^0, π^0
4. Asymptotic expansion of $v^\epsilon = v^0 + \psi + R$; ψ the corrector, and R the remainder
5. Utilize homogenization estimates [6], estimates for BSPDE [7] and Feynman-Kac representations with FBDSDE [8] to prove convergence,

Dual Convergence Implies Filter Convergence

$$\mathbb{E}\left[\left|\rho_T^{\epsilon,x}(\varphi) - \rho_T^0(\varphi)\right|^p\right] \leqslant \int_{\mathbb{R}^2} \mathbb{E}\left[\left|v_0^\epsilon(x,z) - v_0^0(x)\right|^p\right] \mathbb{Q}_{X_0^\epsilon, Z_0^\epsilon}(dx, dz)$$

Mathematical Tools and Proof of Convergence Approach

1. Introduce an unnormalized conditional measure

Unnormalized Conditional Measure ρ_t^ϵ

$$\frac{\rho_t^\epsilon(\varphi)}{\rho_t^\epsilon(1)} = \frac{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\varphi(X_t^\epsilon, Z_t^\epsilon)\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]}{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]} = \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \mid \mathcal{Y}_t^\epsilon\right] = \pi_t^\epsilon(\varphi)$$

2. Introduce function valued dual process, v^ϵ, satisfying a BSPDE
3. Ansatz v^0, ρ^0, π^0
4. Asymptotic expansion of $v^\epsilon = v^0 + \psi + R$; ψ the corrector, and R the remainder
5. Utilize homogenization estimates [6], estimates for BSPDE [7] and Feynman-Kac representations with FBDSDE [8] to prove convergence,

Dual Convergence Implies Filter Convergence

$$\mathbb{E}\left[\left|\rho_T^{\epsilon,x}(\varphi) - \rho_T^0(\varphi)\right|^p\right] \leqslant \int_{\mathbb{R}^2} \mathbb{E}\left[\left|v_0^\epsilon(x,z) - v_0^0(x)\right|^p\right] \mathbb{Q}_{X_0^\epsilon, Z_0^\epsilon}(dx, dz)$$

Mathematical Tools and Proof of Convergence Approach

1. Introduce an unnormalized conditional measure

Unnormalized Conditional Measure ρ_t^ϵ

$$\frac{\rho_t^\epsilon(\varphi)}{\rho_t^\epsilon(1)} = \frac{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\varphi(X_t^\epsilon, Z_t^\epsilon)\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]}{\mathbb{E}_{\mathbb{P}^\epsilon}\left[\widetilde{D}_t^\epsilon \,\middle|\, \mathcal{Y}_t^\epsilon\right]} = \mathbb{E}_{\mathbb{Q}}\left[\varphi(X_t^\epsilon, Z_t^\epsilon) \,|\, \mathcal{Y}_t^\epsilon\right] = \pi_t^\epsilon(\varphi)$$

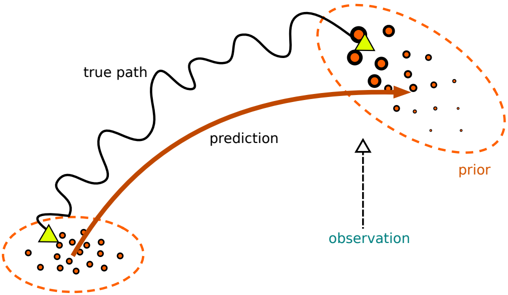
2. Introduce function valued dual process, v^ϵ, satisfying a BSPDE
3. Ansatz v^0, ρ^0, π^0
4. Asymptotic expansion of $v^\epsilon = v^0 + \psi + R$; ψ the corrector, and R the remainder
5. Utilize homogenization estimates [6], estimates for BSPDE [7] and Feynman-Kac representations with FBDSDE [8] to prove convergence,

Dual Convergence Implies Filter Convergence

$$\mathbb{E}\left[\left|\rho_T^{\epsilon,x}(\varphi) - \rho_T^0(\varphi)\right|^p\right] \leqslant \int_{\mathbb{R}^2} \mathbb{E}\left[\left|v_0^\epsilon(x,z) - v_0^0(x)\right|^p\right] \mathbb{Q}_{X_0^\epsilon, Z_0^\epsilon}(dx, dz)$$

Table of Contents

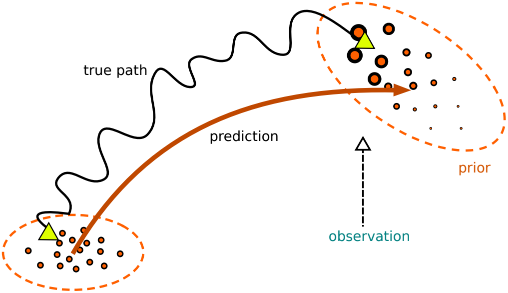
Particle Filters - Discrete Time [11]



Continuous signal, discrete observations:

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad \text{and} \quad Y_{t_k} = h(X_{t_k}) + B_{t_k}$$

Particle Filters - Discrete Time [11]

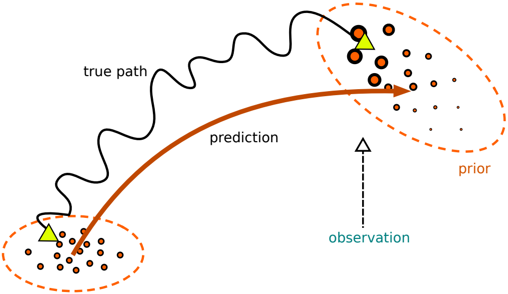


1. $\{x^i \in \mathbb{R}^m\}_{i=1}^{N_s}$, an ensemble of particles.
2. $\{w^i\}$, normalized weights: $\sum_{i=1}^{N_s} w^i = 1$.

Approximation of posterior distribution at time t_k

$$\pi_k(x|y_{0:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x - x_k^i)$$

Particle Filters - Discrete Time [11]

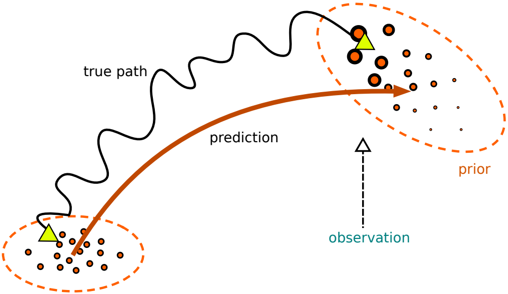


Sequential Importance Sampling - SIS

$$\pi_k(x|y_{0:k}) \propto \psi(x), \quad x_k^i \sim q(x), \quad \text{then} \quad w_k^i \propto \frac{\psi(x)}{q(x)}$$

- ψ, can be evaluated
- q, easy to draw samples from

Particle Filters - Discrete Time [11]



Weights Update

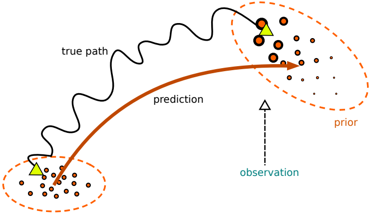
$$w_{k+1}^i \propto \frac{u(y_{k+1}|x_{k+1}^i)u(x_{k+1}^i|x_k^i)}{q(x_{k+1}^i|x_k^i)} w_k^i$$

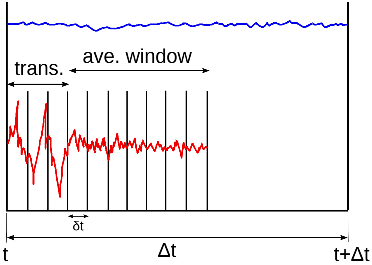
Typically (for simplicity) choose: $q(x_{k+1}^i|x_k^i) = u(x_{k+1}^i|x_k^i)$

Nudged Particle Filter

Choose $q(x_{k+1}^i|x_k^i)$ in an intelligent, but flexible hands-off manner

Particle Filters - Discrete Time [11]





Heterogenous Multiscale Method (HMM) for Homogenized Hybrid PF (HHPF),

$$dX_t^0 = \bar{b}(X_t^0)dt + \bar{\sigma}(X_t^0)dW_t \quad \text{and} \quad Y_{t_k} = \bar{h}(X_{t_k}^0) + B_{t_k}$$

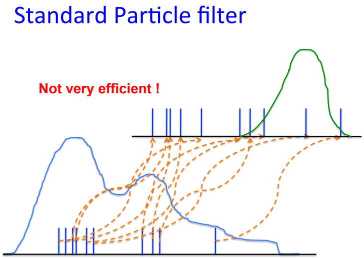
Doeblin Condition

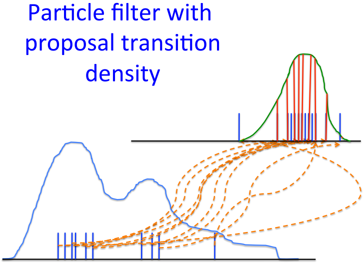
For every fixed x, the solution Z_t^x of

$$dZ_t^x = f(x, Z_t^x)dt + g(x, Z_t^x)dV_t$$

is ergodic and converges rapidly to its unique stationary distribution μ^x.

Nudging of particles





Continuous signal, discrete observations:

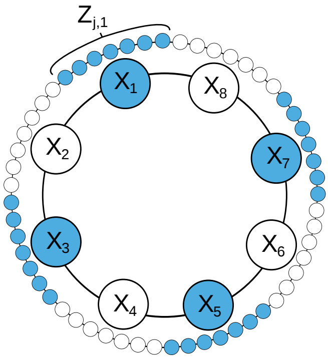
$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad \text{and} \quad Y_{t_k} = h(X_{t_k}) + B_{t_k}$$

Nudge particles:

$$d\widehat{X}_t^i = \left(b(\widehat{X}_t^i) + u_t^i\right) dt + \sigma(\widehat{X}_t^i)dW_t, \qquad t \in (t_k, t_{k+1}).$$

Multiscale Lorenz '96 Model [13], [14]

1. Mid-Latitude Atmospheric Dynamics
2. Linear Dissipation
3. External Forcing F
4. Quadratic Advection-Like Terms (Conserve Total Energy)
5. Chaotic for a wide range of F, h_x, h_z

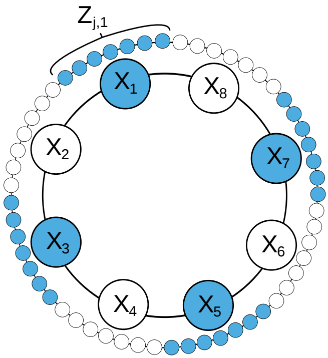


$$dX_t^k = (X_t^{k-1}(X_t^{k+1} - X_t^{k-2}) - X_t^k + F + \frac{h_x}{J}\sum_{j=1}^{J} Z_t^{k,j})dt \qquad k = 1, \ldots, K,$$

$$dZ_t^{k,j} = \frac{1}{\varepsilon}\left(Z_t^{k,j+1}(Z_t^{k,j-1} - Z_t^{k,j+2}) - Z_t^{k,j} + h_z X_t^k\right)dt \qquad j = 1, \ldots, J.$$

Multiscale Lorenz '96 Model [13], [14]

1. Mid-Latitude Atmospheric Dynamics
2. Linear Dissipation
3. External Forcing F
4. Quadratic Advection-Like Terms (Conserve Total Energy)
5. Chaotic for a wide range of F, h_x, h_z



$$dX_t^k = (X_t^{k-1}(X_t^{k+1} - X_t^{k-2}) - X_t^k + F + \frac{h_x}{J}\sum_{j=1}^{J} Z_t^{k,j})dt + \sigma_x dW_t^k, \qquad k = 1, \dots, K,$$

$$dZ_t^{k,j} = \frac{1}{\varepsilon}\left(Z_t^{k,j+1}(Z_t^{k,j-1} - Z_t^{k,j+2}) - Z_t^{k,j} + h_z X_t^k\right) dt + \frac{1}{\sqrt{\varepsilon}}\sigma_z dV_t^j, \qquad j = 1, \dots, J.$$

Nudged HHPF ($HHPF_c$) on Lorenz '96: 36 slow, 360 fast

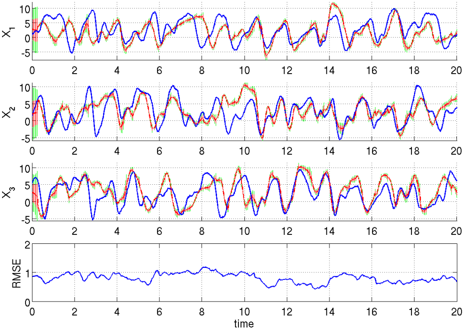


Figure: PF, Observations every 36 hours

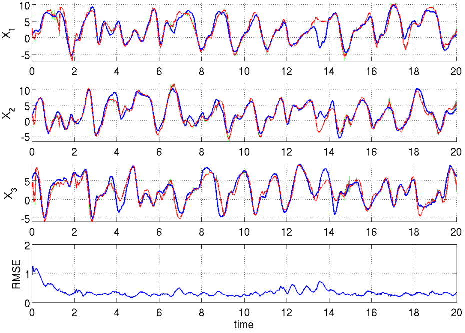


Figure: Nudged HHPF, Observations every 72 hours

Legend:

Truth (Top 3 Plots), Error (Bottom Plot); Filter mean with 1 std and 2 std

Table of Contents

Information Theory Motivation and Sensor Placement / Control

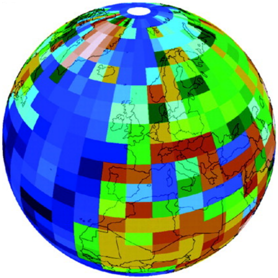


Figure: NCAR Globe, 500km Sectors

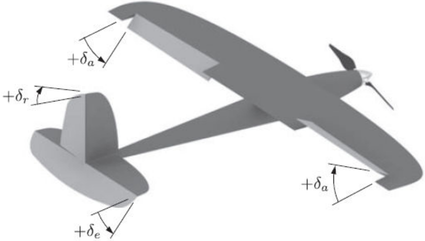


Figure: UAV Flight Controls [5]

- Improved posterior yields better prior for next observation cycle (e.g. prediction or forecasting)
- Information theory provides general tool for describing improvement in knowledge (uncertain) of random variables
- A useful 'metric' is Kullback-Leibler divergence - for filtering, expectation of 'distance' between posterior and prior over all possible observations

Basic Tools in Information Theory

Shannon Entropy

Shannon entropy, an **absolute entropy**,

$$H(X) \equiv -\int_{\mathcal{X}} p(x) \log p(x) \mu(dx)$$

quantifies the information content of a random variable. It can be interpreted as how much uncertainty there is about the random variable.

Entropy of Normal Random Variable

If $X \sim \mathcal{N}(\nu, \Sigma)$, then

$$H(X) = \log((2\pi e)^d |\Sigma|)/2,$$

where $|\cdot|$ will denote the determinant and d is the dimension of $\Sigma \in \mathbb{R}^{d \times d}$.

Conditional Entropy

$$H(X|Y) \equiv -\int_{\mathcal{X} \times \mathcal{Y}} p(x,y) \log p(x|y) \mu(dx, dy)$$

Maximization of Kullback-Leibler Divergence

Definition (Kullback-Leibler Divergence (D_{KL}))

A **relative entropy** that quantifies the 'distance' between two densities. Given densities p and q, their KL divergence is defined as:

$$D_{KL}(p||q) \equiv \int_{\mathcal{X}} p(x) \log\left(p(x)/q(x)\right) \mu(dx)$$

If p is the actual density for a random variable X, then $D_{KL}(p||q)$ can be interpreted as the loss of information due to using q instead of p.

Discrete Time Objective Functional

$$\begin{aligned} J(u_k|y_{0:k-1}) &= \int_{\mathcal{Y}} D_{KL}(p(x_k|y_{0:k};u_k)\,||\,p(x_k|y_{0:k-1};u_k))p(y_k|y_{0:k-1};u_k)dy_k \\ &= \vdots \\ &= H\big|_{y_{0:k-1}}(X_k) - H\big|_{y_{0:k-1}}(X_k|Y_k;u_k) \end{aligned}$$

Maximization of Kullback-Leibler Divergence

Definition (Kullback-Leibler Divergence (D_{KL}))

A **relative entropy** that quantifies the 'distance' between two densities. Given densities p and q, their KL divergence is defined as:

$$D_{KL}(p||q) \equiv \int_{\mathcal{X}} p(x) \log\left(p(x)/q(x)\right) \mu(dx)$$

If p is the actual density for a random variable X, then $D_{KL}(p||q)$ can be interpreted as the loss of information due to using q instead of p.

Discrete Time Objective Functional

$$J(u_k|y_{0:k-1}) = \int_{\mathcal{Y}} D_{KL}(p(x_k|y_{0:k};u_k)\,||\,p(x_k|y_{0:k-1};u_k))p(y_k|y_{0:k-1};u_k)dy_k$$

$$= \quad \vdots$$

$$= H\big|_{y_{0:k-1}}(X_k) - H\big|_{y_{0:k-1}}(X_k|Y_k;u_k)$$

Maximization of Kullback-Leibler Divergence

Definition (Kullback-Leibler Divergence (D_{KL}))

A **relative entropy** that quantifies the 'distance' between two densities. Given densities p and q, their KL divergence is defined as:

$$D_{KL}(p||q) \equiv \int_{\mathcal{X}} p(x) \log \left(p(x)/q(x) \right) \mu(dx)$$

If p is the actual density for a random variable X, then $D_{KL}(p||q)$ can be interpreted as the loss of information due to using q instead of p.

Discrete Time Objective Functional

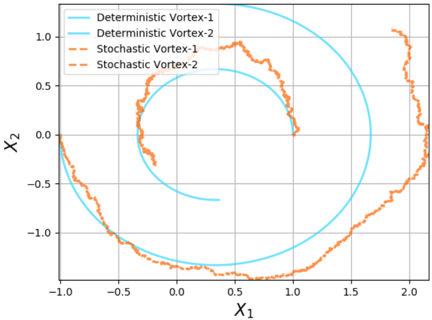
$$\begin{aligned} J(u_k|y_{0:k-1}) &= \int_{\mathcal{Y}} D_{KL}(p(x_k|y_{0:k}; u_k) \,||\, p(x_k|y_{0:k-1}; u_k)) p(y_k|y_{0:k-1}; u_k) dy_k \\ &= \ \vdots \\ &= H\big|_{y_{0:k-1}}(X_k) - H\big|_{y_{0:k-1}}(X_k|Y_k; u_k) \end{aligned}$$

Table of Contents

nVortex Flowfield Model

1. Deterministic vortex dynamics simulates the Euler equations
2. The first random point vortex method to simulate viscous incompressible flow was introduced in [15]

$$\mathbb{J} \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$



$$dX_{i,t} = \frac{1}{2\pi} \sum_{k=1}^{n} \frac{\Gamma_k}{\|X_{k,t} - X_{i,t}\|_2^2} \mathbb{J}(X_{k,t} - X_{i,t})dt + \sqrt{\sigma_x} dW_{i,t}, \qquad X_{i,0} = x \in \mathbb{R}^2,$$

A Controllable Tracer for Adaptive Observations

Assume that tracer is controllable,

$$dX_{i,t} = f_i(X_t) + b_i(u_{i,t}) + \sqrt{\sigma_x} dB_{i,t} \quad \text{and} \quad u_{i,t} \in \mathcal{U}$$
$$Y_{t_k} = h(X_{t_k}) + B_{t_k}$$

where $\mathcal{U}$ is some admissible control set.

- How one might control the tracer so as to best improve the filtering process?
- One approach, formulation of an optimal control problem; specifically in terms of information theoretic quantities.

PF Implementation and Result

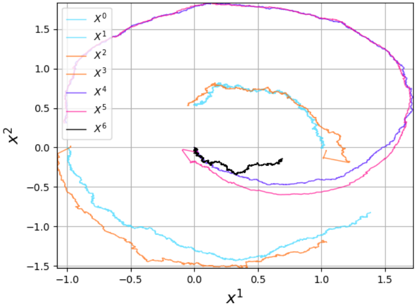


Figure: PF with no control

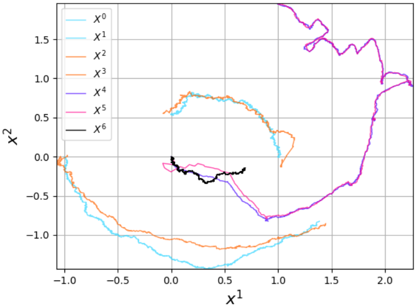


Figure: Controlled PF with RHC over 10 observation steps

PF Implementation and Result

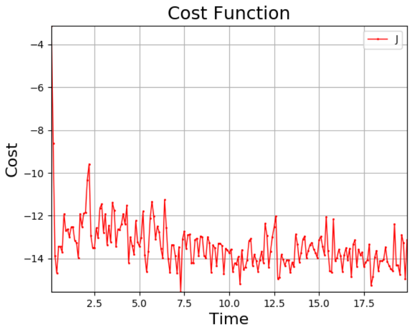


Figure: PF, no control - posterior entropy

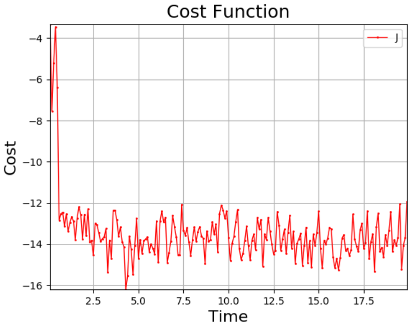


Figure: PF, control with RHC over 10 observation steps - posterior entropy

* Cost Function shown is: $-H(X|Y)$.

PF Implementation and Result

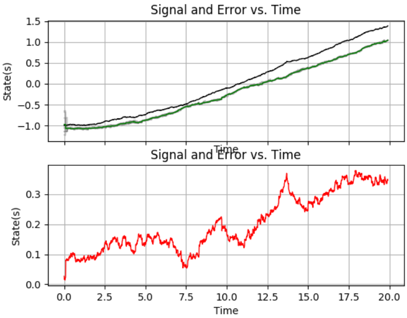


Figure: PF, no control - Vortex-1 x-state

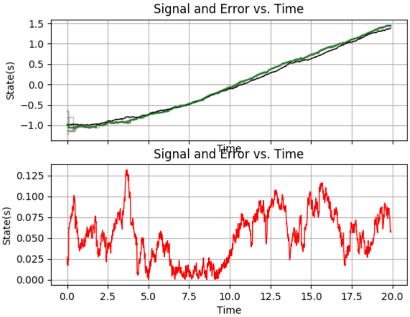


Figure: PF, control with RHC 10 observation steps - Vortex-1 x-state

Figure: x-state shown in top figure, while RMSE shown in bottom

PF Implementation and Result

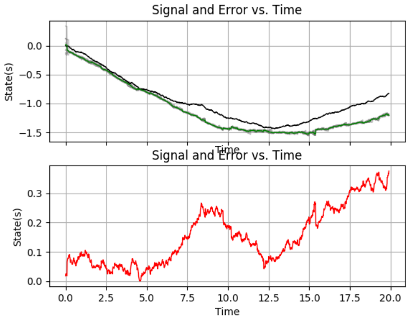


Figure: PF, no control - Vortex-1 y-state

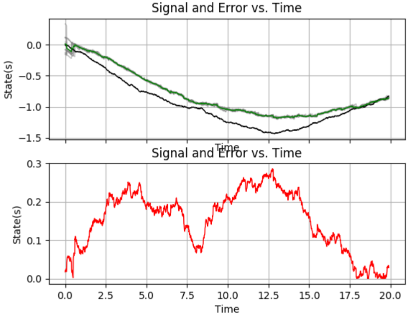


Figure: PF, control with RHC 10 observation steps - Vortex-1 y-state

Figure: x-state shown in top figure, while RMSE shown in bottom

Objectives:

(I) Develop an *integrated framework* that combines the ability to *dynamically steer* the measurement process, extracting useful information, with *nonlinear filtering* for inference and prediction of large scale complex systems.

(II) Develop efficient and robust methods to produce *lower-dimensional recursive nonlinear filtering* equations driven by the observations; particle filters for the integration of observations with the simulations of large-scale complex systems.

Presented:

(i) Use of powerful mathematical techniques - homogenization, SPDE, FBDSDE - to derive convergence results of correlated filter in multiscale problems as well as provide future mechanisms for extension of nudging particle method and information flow for the multiscale correlated noise case.

(ii) Introduced framework by which to breakdown adaptive observation problem into hierarchy of sensor selection / placement and sensor control problems.

(iii) Described preliminary algorithms using information theoretic cost functionals to drive the sensor placement and control problems - demonstrations on test bed problems.

Journal Articles:

- *Namachchivaya, N. Sri; Random dynamical systems: addressing uncertainty, nonlinearity and predictability; Meccanica, (51, 2975-2995); 2016*
  https://link.springer.com/article/10.1007%2Fs11012-016-0570-4
- *Lingala, N., Namachchivaya, N. Sri, et al.; Random perturbations of a periodically driven nonlinear oscillator: escape from a resonance zone; Nonlinearity, (30, 4, 1376); 2017*
  http://iopscience.iop.org/article/10.1088/1361-6544/aa5dc7/meta
- *Yeong, H. C., et al. Particle Filters with Nudging in Multiscale Chaotic Systems: with Application to the Lorenz-96 Atmospheric Model; Submitted to ZAMM, Journal of Applied Mathematics and Mechanics*
- *Beeson, R., et al., Dynamic Data-Driven Adaptive Observations in a Vortex Flowfield; In Preparation to European Journal of Applied Mathematics*

Conference Proceedings:

- *Beeson, R., et al., Dynamic Data-Driven Adaptive Observations in a Vortex Flowfield; 9th European Nonlinear Dynamics Conference; Budapest Hungary; June 2017*
- *Yeong, H.C., et al. Particle Filters with Nudging in Multiscale Chaotic Systems: with Application to the Lorenz-96 Atmospheric Model; Budapest Hungary; June 2017*
- *Lingala, N., Namachchivaya, N. Sri, et al.; Random perturbations of a periodically driven nonlinear oscillator: escape from a resonance zone; SIAM Conference on Dynamical Systems 2017, Snowbird Utah*

[WBC09] W. M. Washington, L. Buja, and A. Craig. "The computational future for climate and Earth system models: on the path to petaflop and beyond". In: *Phil. Trans. Royal Soc. A: Mathematical, Physical and Engineering Sciences* 367.1890 (2009), pp. 833–846.

[Lor84] E. N. Lorenz. "Irregularity: A Fundamental Property of the Atmosphere". In: *Tellus* 36A (1984), pp. 98–110.

[Maa04] Leo R. M. Maas. "Theory of basin scale dynamics of a stratified rotating fluid". In: *Surveys in Geophysics* 25 (2004), pp. 249–279.

[Van03] Lennart Van Veen. "Overturning and wind driven circulation in a low-order ocean-atmosphere model". In: *Dynamics of Atmospheres and Oceans* 37 (2003), pp. 197–221.

[EA15] Ahmed EA. "Modeling of a Small Unmanned Aerial Vehicle". In: 9 (Apr. 2015), pp. 413–421.

[PV03] E. Pardoux and A. Y. Veretennikov. "On Poisson equation and diffusion approximation 2". In: *Ann. Probab.* 31.3 (2003), pp. 1166–1192.

[Roz90] Boris L. Rozovskii. *Stochastic Evolution System: Linear Theory and Aplications to Non-linear Filtering*. Dordrecht: Kluwer Academic Publishers, 1990, p. 315.

[PP94] Etienne Pardoux and Shige Peng. "Backward Doubly Stochastic Differential Equations and Systems of Quasilinear SPDEs". In: *Probab. Theory Related Fields* 98 (1994), pp. 209–227.

[Zak69] M. Zakai. "On the optimal filtering of diffusion processes". In: *Probability Theory and Related Fields* 11.3 (1969).

[Par79] Etienne Pardoux. "Stochastic partial differential equations and filtering of diffusion processes". In: *Stochastics* 3 (1979), pp. 127–167.

[GSS93] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation". In: *IEEE Proceedings F* 140.2 (1993), pp. 107–113.

[Fle82] Wendell H. Fleming. “PLogarithmic transformations and stochastic control”. In: *Advances in Filtering and Optimal Stochastic Control: Proceedings of the IFIP-WG 7/1 Working Conference Cocoyoc, Mexico, February 1–6, 1982*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1982, pp. 131–141. ISBN: 978-3-540-39517-1. DOI: 10.1007/BFb0004532. URL: http://dx.doi.org/10.1007/BFb0004532.

[Lor95] E. N. Lorenz. *Predictability: A problem partly solved*. 1995.

[MTV01] A.J. Majda, I. Timofeyev, and E. Vanden-Eijnden. “A Mathematical Framework for Stochastic Climate Models”. In: *Comm. Pure Appl. Math.* 54 (2001), pp. 891–974.

[Cho73] A. J. Chorin. “Numerical study of slightly viscous flow”. In: *Journal of Fluid Mechanics* 57.4 (1973), pp. 785–796.

[CT06] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. NJ, USA: Wiley Series in Telecommunications and Signal Processing, John Wiley & Sons Inc., 2006.