Lagrangian Data Assimilation and Manifold Detection for a Point-Vortex Model

Mid-year Progress Report

David Darmon, AMSC
Kayo Ide, AOSC, IPST, CSCAMM, ESSIC
Today’s Presentation

Background

Database Generation

The EKF, Validation and Testing

The EnKF, Validation and Testing

Moving Forward
Background
Data Assimilation
Iterative process

True State
Observation

Forecast
Analysis
Database Generation
Database Generation

True System Dynamics

\[ d\mathbf{x}^t = M(\mathbf{x}^t, t) dt + d\eta^t \]

- \( x^t \) – system state
- \( \eta^t \) – Brownian motion
- \( M(\cdot, \cdot) \) – deterministic evolution operator
Database Generation

Deterministic Dynamics

Point-Vortices

\[ \frac{d x_i}{d t} = -\frac{1}{2\pi} \sum_{i' = 1, i' \neq i}^{N_V} \frac{\Gamma_{i'}(y_i - y_{i'})}{l_{i'i}^2} \]

\[ \frac{d y_i}{d t} = \frac{1}{2\pi} \sum_{i' = 1, i' \neq i}^{N_V} \frac{\Gamma_{i'}(x_i - x_{i'})}{l_{i'i}^2} \]

\[ l_{i'i}^2 = (x_i - x_{i'})^2 + (y_i - y_{i'})^2 \]
Database Generation

Deterministic Dynamics

Drifters

\[
\frac{d\xi_i}{dt} = -\frac{1}{2\pi} \sum_{j=1}^{N_v} \frac{\Gamma_j (\eta_i - y_j)}{l_{ij}^2}
\]

\[
\frac{d\eta_i}{dt} = \frac{1}{2\pi} \sum_{j=1}^{N_v} \frac{\Gamma_j (\xi_i - x_j)}{l_{ij}^2}
\]

\[
l_{ij}^2 = (\xi_i - x_j)^2 + (\eta_i - y_j)^2
\]
Database Generation

Scalar vs. Vector Case

Stochastic Differential Equation

\[dX(t) = f(X(t), t)dt + g(X(t), t)dW(t)\]
\[X(0) = X_0, 0 \leq t \leq T\]

\[dx^t(t) = M(x^t(t), t)dt + d\eta^t(t)\]
\[x^t(0) = x_0^t, 0 \leq t \leq T\]
Database Generation
Numerical Solution
Runge-Kutta

\[ h(X_j, \tau_j) = \left[ f(X_j, \tau_j) \right. - \frac{1}{2} g(X_j, \tau_j) \frac{\partial g(X_j, \tau_j)}{\partial X_j} \left. \right] \Delta \tau + g(X_j, \tau_j) (W(\tau_j) - W(\tau_{j-1})) \]
Database Generation

Numerical Solution

Runge-Kutta

\[
K_1 = h(X_j, \tau_j)
\]

\[
K_2 = h(X_j + \frac{1}{2}K_1, \tau_j + \frac{1}{2}\Delta\tau)
\]

\[
K_3 = h(X_j + \frac{1}{2}K_2, \tau_j + \frac{1}{2}\Delta\tau)
\]

\[
K_4 = h(X_j + K_3, \tau_j + \Delta\tau)
\]

\[
X_{j+1} = X_j + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)
\]
Database Generation

Numerical Solution

Runge-Kutta

\[
K_1 = h(X_j, \tau_j)
\]

\[
K_2 = h(X_j + \frac{1}{2}K_1, \tau_j + \frac{1}{2}\Delta\tau)
\]

\[
K_3 = h(X_j + \frac{1}{2}K_2, \tau_j + \frac{1}{2}\Delta\tau)
\]

\[
K_4 = h(X_j + K_3, \tau_j + \Delta\tau)
\]

\[
X_{j+1} = X_j + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)
\]
Database Generation

Simple Scalar Case

Validation

\[ dX(t) = \lambda X(t)\,dt + \mu X(t)\,dW(t) \]

\[ X(0) = X_0 \]

\[ X(t) = X_0 \exp \left[ \left( \lambda - \frac{1}{2} \mu^2 \right) t + \mu W(t) \right] \]
True solution

RK solution
Database Generation

True Dynamics

Stochastic Differential Equation

\[ d\mathbf{x}^t = M(\mathbf{x}^t, t)dt + d\eta^t \]

\[ d\mathbf{X} = \mathbf{f}(\mathbf{X}, t)dt + \mathbf{g}(\mathbf{X}, t)dW(t) \]

\[ \mathbf{f}(\mathbf{x}^t, t) = M(\mathbf{x}^t, t) \]

\[ \mathbf{g}(\mathbf{x}^t, t) = \sqrt{2}Q^{1/2} \]
Vortex 1
Vortex 2
Drifter

\[ \sigma = 0.02 \]

\[ Q = 2\sigma^2 I \]
Vortex 1
Vortex 2
Drifter

\[ \sigma = 0.02 \]
\[ \rho = 0.02 \]

\[ Q = 2\sigma^2 I \]
\[ R = \rho^2 I \]
Data Assimilation

True State → Observation

Forecast → Analysis

Feedback loop
Data Assimilation

Approach 1:
EKF
Approach 1

Extended Kalman Filter (EKF), Forecast

\[
\frac{d}{dt} x^f = M(x^f, t)
\]

\[
\frac{d}{dt} P^f = M(t)P^f + P^f M^T(t) + Q
\]

\(Q\) – covariance matrix from SDE
\(M(t) = J[M(x, t)]\big|_{x=x^f}\) - Jacobian of \(M\)
Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

\[ \mathbf{x}_1 \quad - \quad \text{Reference Trajectory} \]
\[ \mathbf{x}_2 \quad - \quad \text{Perturbed Trajectory} \]
Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

\[ x_1 \quad - \quad \text{Reference Trajectory} \]
\[ x_2 \quad - \quad \text{Perturbed Trajectory} \]

\[ x_1(0) = x^0 \]
\[ x_2(0) = x^0 + y^0 \]
Approach 1

Extended Kalman Filter (EKF)
Validation - Jacobian via Tangent Linear Model

\[ x_1 \quad - \quad \text{Reference Trajectory} \]
\[ x_2 \quad - \quad \text{Perturbed Trajectory} \]

\[ x_1(0) = x^0 \]
\[ x_2(0) = x^0 + y^0 \]

Tangent Vector (small)
Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

\[
\frac{d\mathbf{y}}{dt} \bigg|_{x=x(t)} = \left( \frac{\partial M(x)}{\partial x} \right)_{x=x(t)} \mathbf{y} = \mathbf{M}(x(t))\mathbf{y}
\]
Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

Evolve reference and perturbed trajectories forward using deterministic ODE

Evolve tangent vector forward using Tangent Linear Model

Compare $y_{\text{perturbed}}$ and $y_{\text{TL}}$
Tangent Linear
RK
Approach 1

Extended Kalman Filter (EKF), Analysis

\[
\begin{align*}
    x_k^a &= x^f(t_k) + K_k(y_k^o - h_k(x^f(t_k))) \\
    P_k^a &= (I - K_kH_k)P^f(t_k) \\
    K_k &= P^f(t_k)H_k^T(H_kP^f(t_k)H_k^T + R_k^o)^{-1}
\end{align*}
\]

R^o – covariance matrix from observation

H_k = J[h(x, t_k)]|_{x=x^f} - Jacobian of h_k
Approach 1

Extended Kalman Filter (EKF)

Validation - No Noise

\[
\begin{align*}
\mathbf{x}_k^a &= \mathbf{x}_k^f(t_k) + \mathbf{K}_k (\mathbf{y}_k^o - h_k(\mathbf{x}_k^f(t_k))) \\
\mathbf{P}_k^a &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^f(t_k) \\
\mathbf{K}_k &= \mathbf{P}_k^f(t_k) \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f(t_k) \mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}
\end{align*}
\]
Approach 1

Extended Kalman Filter (EKF)

Validation - Full Observation, Imperfect Data

Observe vortices and drifter under observational noise
Vortex 1
Approach 1

Extended Kalman Filter (EKF)

Results - Partial Observation, Imperfect Data

Observe only drifter under observational noise
Data Assimilation

Approach 1: EnKF
Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

\[ \{ x_i^f \}_{i=1}^N \]

Forecast: Evolve particles forward using SDE

Analysis: Perform Kalman-like analysis
Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

$$\{x^f_i\}_{i=1}^N$$

Forecast: Evolve particles forward using SDE-ODE

Analysis: Perform Kalman-like analysis
Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

\[ \{x_i^f\}_{i=1}^N \]

Forecast: Evolve particles forward using ODE

Analysis: Perform Kalman-like analysis
Approach 2a

EnKF with Observational Perturbations

\[ y_{k,i}^o = y_k^o + \epsilon_i, \quad i = 1, \ldots, n \]
\[ \epsilon_i \sim N(0, R^o) \]
Approach 2a

EnKF with Observational Perturbations

\[ y_{k,i}^o = y_k^o + \epsilon_i, \quad i = 1, \ldots, n \]
\[ \epsilon_i \sim N(0, R^o) \]

\[ \{x_i^f\}_{i=1}^N \quad \text{Forecast ensemble} \]
\[ \{y_k^o\}_{i=1}^N \quad \text{Observation ensemble} \]
Approach 2a

EnKF with Observational Perturbations
Kalman-like update

\[
x_{k,i}^a = x_{i}^f(t_k) + K_k(y_{k,i}^o - h_k(x_{i}^f(t_k)))
\]
\[
K_k = P_e^f(t_k)H_k^T(\cancel{H_kP_e^f(t_k)H_k^T} + R_k^o)^{-1}
\]
Approach 2a

EnKF with Observational Perturbations
Kalman-like update

\[ \mathbf{x}_{k,i}^a = \mathbf{x}_{i}^f(t_k) + \mathbf{K}_k(\mathbf{y}_{k,i}^o - h_k(\mathbf{x}_{i}^f(t_k))) \]

\[ \mathbf{K}_k = \mathbf{P}_e^f(t_k) \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_e^f(t_k) \mathbf{H}_k^T + \mathbf{R}_k^o)^{-1} \]

Sample Covariance Matrix

\[ \mathbf{P}_e^f(t_k) = \frac{1}{N - 1} \sum_{i=1}^{N} (\mathbf{x}_{i}^f - \bar{\mathbf{x}}^f)(\mathbf{x}_{i}^f - \bar{\mathbf{x}}^f)^T \]
Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

\[ \{ x_i^f \}_{i=1}^N \]

Forecast: Evolve particles forward using SDE-ODE

Analysis: Perform Kalman-like analysis
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

\[ x_k^a = x^f(t_k) + K_k(y_k^o - h_k(x^f(t_k))) \]
\[ P_k^a = (I - K_k H_k) P_k^f(t_k) \]
\[ K_k = P_k^f(t_k) H_k^T (H_k P_k^f(t_k) H_k^T + R_k^o)^{-1} \]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

\[
x^a_k = x^f(t_k) + K_k(y^o_k - h_k(x^f(t_k)))
\]
\[
P^a_k = (I - K_kH_k)P^f(t_k)
\]
\[
K_k = P^f(t_k)H^T_k (H_kP^f(t_k)H^T_k + R^o_k)^{-1}
\]

Use sample estimates:

\[
\bar{x}^a_k = \bar{x}^f(t_k) + K_{e,k}(y^o_k - h_k(\bar{x}^f(t_k)))
\]
\[
P^a_{e,k} = (I - K_{e,k}H_k)P^e_{e}(t_k)
\]
\[
K_{e,k} = P^e_{e}(t_k)H^T_k (H_kP^e_{e}(t_k)H^T_k + R^o_k)^{-1}
\]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

\[ \mathbf{X}^f = \begin{bmatrix} x_1^f - \bar{x} & x_2^f - \bar{x} & \ldots & x_n^f - \bar{x} \end{bmatrix} \]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

\[ X^f = \begin{bmatrix} x_1^f - \bar{x} & x_2^f - \bar{x} & \ldots & x_n^f - \bar{x} \end{bmatrix} \]

\[ \frac{1}{N-1} X^f (X^f)^T = P_e^f \]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

\[
X^f = \begin{bmatrix}
x_1^f - \bar{x} \\
x_2^f - \bar{x} \\
\vdots \\
x_n^f - \bar{x}
\end{bmatrix}
\]

\[
\frac{1}{N-1} X^f (X^f)^T = P_e^f
\]

\[
X^a = X^f W
\]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

\[ D = WW^T \]

\( W \) is the matrix square root of \( D \)

Choose

\[ D = (I + (X^f)^T H^T (R^o)^{-1} H X^f)^{-1} \]
Approach 2b

Ensemble Transform Kalman Filter (ETKF)
Analysis Ensemble

\[ X^a = X^f W \]

\[ \tilde{x}^a = \tilde{x}^f + X^f D(X^f)^T H^T (R^o)^{-1} \left( y^o - H\tilde{x}^f \right) \]
Approach 2

Ensemble Kalman Filter (EnKF)

Validation - Full Observation, Imperfect Data

Observe vortices and drifter under observational noise

N = 6 ensemble members
Vortex I
$\|X_{true} - X_{filter}\|_2$ vs Time (s)
Approach 2

Ensemble Kalman Filter (EnKF)

Results - Partial Observation, Imperfect Data

Observe only drifter under observational noise

N = 6 ensemble members
$\|X_{\text{true}} - X_{\text{filter}}\|_2^2$ vs Time (s)
Data Assimilation

Testing
Testing, Phase I

Failure statistic: time to failure

Compare EKF, perturbed observations
EnKF and ETKF
Testing, Phase I

Failure statistic: time to failure

Generate $M = 500$ instances of the SDE at each of $L = 4$ drifter locations

Start vortex 1 at $(0, 1)$
Start vortex 2 at $(0, -1)$

Add observational noise according to the model
Testing, Phase I

Failure statistic: time to failure

Record when the distance between the analyzed state and the true state for either vortex is greater than 1
Testing, Phase I

Failure statistic: time to failure

<table>
<thead>
<tr>
<th></th>
<th>EKF</th>
<th>EnKF, Pert</th>
<th>ETKF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.23</td>
<td>25.64</td>
<td>25.27</td>
</tr>
<tr>
<td>Stdev</td>
<td>9.53</td>
<td>17.17</td>
<td>17.24</td>
</tr>
<tr>
<td>Fraction Completed</td>
<td>0.024</td>
<td>0.094</td>
<td>0.084</td>
</tr>
</tbody>
</table>
Moving Forward

Improvements to EnKF

Covariance Inflation

Localization

Local Ensemble Transform Kalman Filter (LETKF)
Moving Forward
Implement Particle Filter

Approximate pdf by an ensemble of weighted particles:

\[ \left\{ x_i^f(t) \right\}_{i=1}^{N} \quad \left\{ w_{i,k} \right\}_{i=1}^{N} \]

Forecast: Evolve particles forward using SDE

Analysis: Perform full Bayesian update at analysis
Moving Forward

Phase II - Manifold Detection for Observing System Design
Timeline

Phase I

– Produce database: now through mid-October
– Develop extended Kalman Filter: now through mid-October
– Develop ensemble Kalman Filter: mid-October through mid-November
– Develop particle filter: mid-November through end of January
– Validation and testing of three filters (serial): Beginning in mid-October, complete by February
Timeline

Phase I

- Produce database: now through mid-October
- Develop extended Kalman Filter: now through mid-October
- Develop ensemble Kalman Filter: mid-October through mid-November
- Develop particle filter: mid-November through end of January
- Validation and testing of three filters (serial): Beginning in mid-October, complete by February
(Some) References


Questions???