Non-Gaussian Data Assimilation using Mixture and Kernel

Piyush Tagade
Hansjorg Seybold
Sai Ravela

Earth Systems and Signals Group
Earth Atmospheric and Planetary Sciences,
Massachusetts Institute of Technology
Environmental DDDAS

- Mapping localized atmospheric phenomena (caos.mit.edu)
Key Areas

• Theory
  – Generalized Coherent Features
  – Pattern theory for coherent fluids
  – Statistical Inference for GCF

• Methodology
  – Data-driven Nowcasting
  – DBFE

• Information theoretic UQ, Targeting, Estimation, Planning, and Control

• Application
  – Mapping localized phenomena

Field Alignment -- DA

Coalescence -- UQ

PAG modes – Model Reduction
Dynamically Biorthogonal Field Equations

- KL Expansion
- Biorthogonal Expansion

\[ u(x, t; \omega) = \bar{u}(x, t) + \sum_{i=1}^{N} \sum_{p=1}^{P} Y_p^i(t) \psi_p(\xi(\omega)) u_i(x, t) \]

- Dynamic Orthogonality: Closed form solution for mean, eigenvalues and expansion coefficients
- Faster than gPC with comparable accuracy
Inference Challenges: Nonlinearity, Dimensionality

- **System Dynamics**
  \[ \mathbf{x}_{t+\Delta t} = f(\mathbf{x}_t, \mathbf{u}_t) \]

- **Measurements**
  \[ y_{t'} = g(\hat{x}_{t'}) + v_{t'}, \]

- **Bayesian Inference**
  - **Smoothing**
    \[ p(\mathbf{x}_{0:t} | y_{1:t}) \propto p(y_{1:t} | \mathbf{x}_{0:t}) p(\mathbf{x}_{0:t}) \]
  - **Filtering**
    \[ p(\mathbf{x}_t | y_{1:t}) \propto p(y_t | \mathbf{x}_t) p(\mathbf{x}_t | y_{1:t-1}) \]
Two Approaches

• **Mutual Information Filter**
  – Non Gaussian prior and likelihood
  – Distributions oriented
  – Quadratic non-Gaussian estimation, optimization approach
  – Tractable mutual information measures

• **Mixture Ensemble Filter (MEnF)**
  – GMM prior and Gaussian likelihood, joint state-parameter estimation
  – Direct Ensemble update, compact ensemble transform
  – Fast smoother forms
  – Marginal additional cost
  – Note: not for UQ
Approach 1: Mutual Information Filter

**Mutual Information Filter (MuIF):** Non-Gaussian prior and likelihood, Maximization of Kapoor’s quadratic mutual information, equivalence with RKHS

- Shannon Entropy
  \[
  \mathcal{H}(X) = - \int f_X(x) \log(f_X(x)) \, dx
  \]

- Mutual Information
  \[
  \mathcal{J}(X; Y) = \mathcal{H}(X) - \mathcal{H}(X \mid Y)
  \]

- Inference: Maximization of Mutual Information

**Tractable?**

- We use Generalization of Shannon Entropy: Renyi Entropy
  - \(\alpha\)-Information Potential:
  \[
  \mathcal{V}_\alpha(X) = \int (f_X(x))^\alpha \, dx
  \]
  - Others: Havrda-Chardat, Behara-Chawla, Sharma-Mittal
  - Kapur establishes equivalence
  \[
  \mathcal{R}_\alpha(X) = \frac{1}{1 - \alpha} \log \mathcal{V}_\alpha(X)
  \]
**Quadratic Mutual Information**

- Special Interest:
  
- Kapoor’s result:
  - Also Cauchy-Schwartz

- Mutual Information via Information potentials

- Noparametric density estimation
  - Kernel density estimate

- Reinterpretation using RKHS

- Inference

- Filter Parameterization

- Maximization of MI: *Steepest Gradient Ascent*
Application Results: Lorenz 95

- Lorenz 95 Model

\[ \frac{dx_i}{dt} = -x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + u \]

MuIF Outperforms EnKF

Effect of Kernel Bandwidth
Approach 2: Mixture Ensemble Filter

\[ p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx} \]

\[ p(x) \sim \sum_{m=1}^{M} \alpha_m N(x; \mu_m, P_m) \]

State of the Art:
- Means and Covariances of GMMs explicitly propagated and updated (Alspach and Sorenson, 1972)
- Ensemble to propagate uncertainty. Mixture moments explicitly calculated (Sondergaard and Lermusiaux, 2013)
- Ensemble members used with balance and membership rules (Frei and Knuscht, 2013)

Key Contributions:
- Ensemble update equation via Jensen’s inequality
- Update reduces to compact ensemble transform
- Direct model-state adjustment, resampling not compulsory
- EM for mixture parameter estimation, also in reduced form
- Familiar message passing framework for smoothing O(1) – fixed lag, O(N) – fixed interval, in time
Mixture Ensemble Filter

- Joint state and parameter estimation

\[
p(x|y) \approx \frac{1}{E} \sum_{e=1}^{E} p(y|x)p(x|\theta)
\]

\[
p(\theta|y) \approx \frac{1}{E} \sum_{e=1}^{E} p(\theta|x_e)
\]

- Ensemble Update equation via Jensen’s inequality

\[
x_e^+ = x_e^- + \sum_{m=1}^{M} \omega_{em} K(P_m^-)(y - \mathcal{H}x_e^-)
\]

- Ensemble Transform \( A^+ = A^- \chi \)

- Smoother Form \( A_k^s = A_k^- \prod_{i=k}^{N} \chi^{(i)} \)
MEnF: Performance

Double Well

Lorenz - 63

90% Confidence Bound
Comparison for MEnF with EnKF and MCMC
MEnF and Coherent Structures

\[ \dot{u} + \frac{1}{2} (u^2)_x + u_{xxx} + \gamma u = \eta f(x) \]

KdV: Idealizes coherent structure of solitary waves, non-linear superposition, chaotic
Conclusions and Future Work

- Nonlinearity and Dimensionality challenges in inference for large-scale fluids
- Optimization approach to non-Gaussian estimation
- Mixture Ensemble Filter and Mutual Information Filter
- Complexity of Mutual Information Filter
  - Fast Gauss Transform, and variations from RKHS
  - Smoother forms
- Convergence rates of MEnF
- Combine Inference with UQ and Planning for Clouds and Volcanoes.