The Mean-Field Ensemble Kalman Filter: Gaussian and Particle Approximations

Edoardo Calvello[1]

Collaborators: Franca Hoffmann [1], Sebastian Reich [2], Andrew M. Stuart [1]

[1]Department of Computing and Mathematical Sciences
California Institute of Technology

[2]Institut für Mathematik
Universität Potsdam

International Symposium on Data Assimilation
October 20th, 2023

Caltech

Table of Contents

Caltech

State Estimation and Non-linear Filtering

State-observation Model

$$
\begin{aligned}
\text{state evolution} \quad & v^\dagger_{n+1} = \Psi(v^\dagger_n) + \xi^\dagger_n, && \text{(1a)}\\
\text{observation evolution} \quad & y^\dagger_{n+1} = h(v^\dagger_{n+1}) + \eta^\dagger_{n+1}, && \text{(1b)}\\
\text{probabilistic info} \quad & v^\dagger_0 \sim \mathrm{N}(m_0, C_0), \quad \xi^\dagger_n \sim \mathrm{N}(0, \Sigma) \text{ i.i.d.}, \quad \eta^\dagger_n \sim \mathrm{N}(0, \Gamma) \text{ i.i.d.} && \text{(1c)}
\end{aligned}
$$

The Problem

From $Y^\dagger_n = \{y^\dagger_\ell\}_{1\le\ell\le n}$ and known Ψ and h

- **State Estimation:** output v_n from $Y^\dagger_n$ so that $\{v_n\}_{n\in\mathbb{Z}^+}$ estimates $\{v^\dagger_n\}_{n\in\mathbb{Z}^+}$;
- **Filtering:** estimate the distribution of random variable $v^\dagger_n | Y^\dagger_n \sim \mu_n$.

State Estimation and Non-linear Filtering

Sample Path Picture

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{2a}$$

$$\widehat{h}_{n+1} = h(\widehat{v}_{n+1}) + \eta_{n+1}, \tag{2b}$$

$$v_{n+1} = \widehat{v}_{n+1} + K_{n+1}(y^{\dagger}_{n+1} - \widehat{h}_{n+1}). \tag{2c}$$

Filtering Cycle

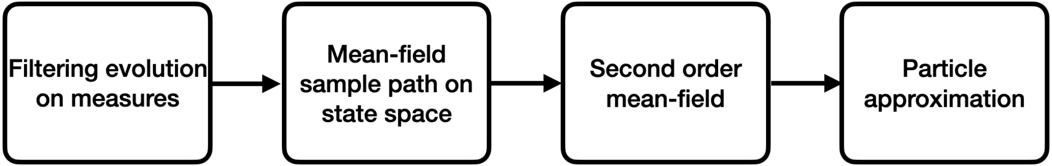
Prediction:

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n.$$

Analysis:

$$v_{n+1} = \widehat{v}_{n+1} + K_{n+1}\Big(y^{\dagger}_{n+1} - h\big(\widehat{v}_{n+1}\big) - \eta_{n+1}\Big).$$

Roadmap to Ensemble Kalman Methods



Transport perspective: Daum et al. [2010], Reich [2011], El Moselhy and Marzouk [2012], Cotter and Reich [2013], Spantini et al. [2019].

Table of Contents

Maps on Measures

Sample Paths: $v_n \sim \mu_n$

$$\begin{aligned} \widehat{v}_{n+1} &\sim \widehat{\mu}_{n+1} \\ (\widehat{v}_{n+1}, \widehat{h}_{n+1}) &\sim \nu_{n+1} \\ v_{n+1} &\sim \mu_{n+1} \end{aligned}$$

Evolution on Measures

$$\widehat{\mu}_{n+1} = P\mu_n, \tag{4a}$$

$$\nu_{n+1} = Q\widehat{\mu}_{n+1}, \tag{4b}$$

conditioning

$$\mu_{n+1} = B_n(\nu_{n+1}), \tag{4c}$$

Filtering Cycle

Prediction

$$\widehat{\mu}_{n+1} = P\mu_n, \tag{5a}$$

Application of Bayes' theorem

$$\mu_{n+1} = B_n Q(\widehat{\mu}_{n+1}). \tag{5b}$$

Caltech

Gaussian Projected Filtering

Let $\mu_n^G = \mathrm{N}(m_n, C_n)$.

Gaussian Projection

$$\widehat{\mu}_{n+1}^G = P\mu_n^G, \tag{6a}$$

$$\nu_{n+1}^G = Q\widehat{\mu}_{n+1}^G, \tag{6b}$$

$$\text{conditioning} \qquad \mu_{n+1}^G = B_n(G\nu_{n+1}^G). \tag{6c}$$

Gaussian Projected Filter

$$\text{mean evolution} \qquad m_{n+1} = \widehat{m}_{n+1} + \widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}\left(y_{n+1}^\dagger - \mathbb{E}h(\widehat{v}_{n+1})\right), \tag{7a}$$

$$\text{covariance evolution} \qquad C_{n+1} = \widehat{C}_{n+1} - \widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}\left(\widehat{C}_{n+1}^{vy}\right)^\top, \tag{7b}$$

Caltech

Table of Contents

Caltech

From Perfect to Approximate Transport

Mean-Field Sample Path on State Space

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{8a}$$
$$\widehat{y}_{n+1} = h(\widehat{v}_{n+1}) + \eta_{n+1}, \tag{8b}$$
$$v_{n+1} = T(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \nu_{n+1}, y^\dagger_{n+1}). \tag{8c}$$

Transport of Measures

$$\widehat{\mu}_{n+1} = P\mu_n, \tag{9a}$$
$$\nu_{n+1} = Q\widehat{\mu}_{n+1}, \tag{9b}$$
$$\mu_{n+1} = (T_n)^\sharp \nu_{n+1}. \tag{9c}$$

Caltech

From Perfect to Approximate Transport

Second Order Mean-Field Sample Path on State Space

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{10a}$$

$$\widehat{y}_{n+1} = h(\widehat{v}_{n+1}) + \eta_{n+1}, \tag{10b}$$

$$v_{n+1} = \widetilde{T}(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \nu_{n+1}, y^{\dagger}_{n+1}). \tag{10c}$$

Transport of Measures

$$\widehat{\mu}_{n+1} = P\mu_n, \tag{11a}$$

$$\nu_{n+1} = Q\widehat{\mu}_{n+1}, \tag{11b}$$

$$\mu_{n+1} = (\widetilde{T}_n)^{\sharp}\nu_{n+1}. \tag{11c}$$

Caltech

Table of Contents

Caltech

Approximate Transport

Second Order Transport

Seek $\widetilde{T}$ such that for any ν

$$G\big((\widetilde{T}_n)^\sharp\nu\big) = B_n(G\nu). \tag{12}$$

Second Order Mean-Field

Seek $\widetilde{T}$ in the form

$$v_{n+1} = \widetilde{T}(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \nu, y^\dagger) := A\widehat{v}_{n+1} + B\widehat{y}_{n+1} + a. \tag{13}$$

s.t. v_{n+1} has mean and covariance given by the Gaussian Projected Filter (7).

Caltech

Approximate Transport

Choice $B = 0$ and $A = C_{n+1}^{\frac{1}{2}} \widehat{C}_{n+1}^{-\frac{1}{2}}$

$$\widetilde{T}(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \nu_{n+1}, y_{n+1}^{\dagger}) := m_{n+1} + C_{n+1}^{\frac{1}{2}} \widehat{C}_{n+1}^{-\frac{1}{2}} (\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}), \tag{14}$$

Choice $B = -\widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}$ and $A = I$

$$\widetilde{T}(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \widetilde{\nu}_{n+1}, y_{n+1}^{\dagger}) := \widehat{v}_{n+1} + \widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}(y_{n+1}^{\dagger} - \widehat{y}_{n+1}), \tag{15}$$

We refer to (15) as Kalman Transport. Set $K_{n+1} = \widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}$.

Caltech

Particle Approximation

Ensemble Kalman Filter

Prediction:

$$\widehat{v}_{n+1}^{(j)} = \Psi(v_n^{(j)}) + \xi_n^{(j)}, \; n \in \mathbb{Z}^+, \tag{16a}$$

$$\widehat{y}_{n+1}^{(j)} = h(\widehat{v}_{n+1}^{(j)}) + \eta_{n+1}^{(j)}, \; n \in \mathbb{Z}^+, \tag{16b}$$

Analysis:

$$v_{n+1}^{(j)} = \widehat{v}_{n+1}^{(j)} + \color{red}{K_{n+1}}\big(y_{n+1}^{\dagger} - \widehat{y}_{n+1}^{(j)}\big), \tag{16c}$$

$$\nu_{n+1}^{\mathrm{J}} = \frac{1}{J}\sum_{j=1}^{J} \delta_{(\widehat{v}_{n+1}^{(j)}, \widehat{y}_{n+1}^{(j)})}. \tag{16d}$$

$$\widehat{C}_{n+1}^{vh} = \mathbb{E}_{n+1}^{\mathrm{J}}\Big(\big(\widehat{v}_{n+1} - \mathbb{E}_{n+1}^{\mathrm{J}}\widehat{v}_{n+1}\big) \otimes \big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathrm{J}} h(\widehat{v}_{n+1})\big)\Big), \tag{17a}$$

$$\widehat{C}_{n+1}^{hh} = \mathbb{E}_{n+1}^{\mathrm{J}}\Big(\big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathrm{J}} h(\widehat{v}_{n+1})\big) \otimes \big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathrm{J}} h(\widehat{v}_{n+1})\big)\Big) + \Gamma. \tag{17b}$$

Caltech

Table of Contents

Set-up and Assumptions

Sample Paths: Ψ and h **bounded**

$$v_{n+1} = \Psi(v_n) + \xi_n$$
$$y_{n+1} = h(v_{n+1}) + \eta_{n+1}.$$

Maps on Measures Set-up

$$\mu_{n+1} = B_n(QP\mu_n),$$
$$\mu^{MF}_{n+1} = (\widetilde{T}_n)^\sharp(QP\mu^{MF}_n).$$

Weighted Total Variation Metric

Let $g(v) = 1 + |v|^2$.

$$\mathrm{d}_g(\mu_1, \mu_2) = \sup_{|f|\le g} \big|\mu_1[f] - \mu_2[f]\big|, \qquad \mu[f] := \int f(u)\mu(du). \tag{20}$$

Caltech

Error Estimate for Mean-Field EnKF

Assumption: True Filter is Close to Gaussian

True filter $\{\mu_n\}$ satisfies

$$\sup_{0 \le n \le N} \mathrm{d}_g(GQP\mu_n, QP\mu_n) \le \epsilon. \tag{21}$$

Theorem (Carrillo et al. [2022])

Let $\mu_0^{MF} = \mu_0$. Under Assumtion, there exists $C > 0$ such that

$$\sup_{0 \le n \le N} \mathrm{d}_g\left(\mu_n, \mu_n^{MF}\right) \le C\epsilon. \tag{22}$$

Ongoing Work

Extend analysis to Ψ and h being **linear** + **bounded**.

Outlook: Error Estimate for EnKF

Goal

Under certain assumptions, obtain

$$\sup_{0\leq n\leq N} \mathrm{d}\big(\mu_n, \mu_n^{EK}\big) \leq C\epsilon. \tag{23}$$

Idea: Non-Asymptotic + Asymptotic Analysis

$$\mathrm{d}\big(\mu_n, \mu_n^{EK}\big) \leq \mathrm{d}\big(\mu_n, \mu_n^{MF}\big) + \mathrm{d}\big(\mu_n^{MF}, \mu_n^{EK}\big). \tag{24}$$

Sample path non-asymptotic analysis in Ghattas and Sanz-Alonso [2023].

Caltech

References I

J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. The ensemble kalman filter in the near-gaussian setting, 2022. URL https://arxiv.org/abs/2212.13239.

C. Cotter and S. Reich. Ensemble filter techniques for intermittent data assimilation. *Radon Ser. Comput. Appl. Math.*, 13:91–134, 2013. doi: 10.1515/9783110282269.91.

F. Daum, J. Huang, and A. Noushin. Exact particle flow for nonlinear filters. In I. Kadar, editor, *Signal Processing, Sensor Fusion, and Target Recognition XIX*, volume 7697, pages 92 – 110. International Society for Optics and Photonics, SPIE, 2010. doi: 10.1117/12.839590.

T. A. El Moselhy and Y. M. Marzouk. Bayesian inference with optimal maps. *Journal of Computational Physics*, 231:7815–7850, 2012. doi: https://doi.org/10.1016/j.jcp.2012.07.022.

O. A. Ghattas and D. Sanz-Alonso. Non-asymptotic analysis of ensemble kalman updates: Effective dimension and localization, 2023.

E. N. Lorenz. Predictability: A problem partly solved. In *Proc. Seminar on Predictability*, volume 1. ECMWF, 1996.

References II

S. Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51(1):235–249, 2011.

A. Spantini, R. Baptista, and Y. Marzouk. Coupling techniques for nonlinear ensemble filtering. *arXiv preprint arXiv:1907.00389*, 2019.

Check out our preprint!

E. Calvello, S. Reich, and A. M. Stuart. Ensemble Kalman Methods: A Mean Field Perspective. 2022. doi: 10.48550/ARXIV.2209.11371. URL https://arxiv.org/abs/2209.11371.

The Lorenz '96 Model

Lorenz '96 Singlescale [Lorenz, 1996]

Consider unknown $v \in C(\mathbb{R}^+, \mathbb{R}^L)$ satisfying the equations

$$\dot{v}_\ell = -v_{\ell-1}(v_{\ell-2} - v_{\ell+1}) - v_\ell + F + h_v m(v_\ell), \quad \ell = 1 \dots L, \tag{25a}$$

$$v_{\ell+L} = v_\ell, \quad \ell = 1 \dots L. \tag{25b}$$

Given Ψ the solution operator for (25), observations $\{y_n^\dagger\}_{n\in\mathbb{Z}^+}$ arise from the model

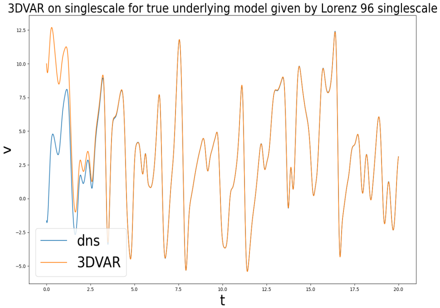
true state evolution $\quad v_{n+1}^\dagger = \Psi(v_n^\dagger) + \xi_n^\dagger, \quad$ (26a)

true observation evolution $\quad y_{n+1}^\dagger = h(v_{n+1}^\dagger) + \eta_{n+1}^\dagger, \quad$ (26b)

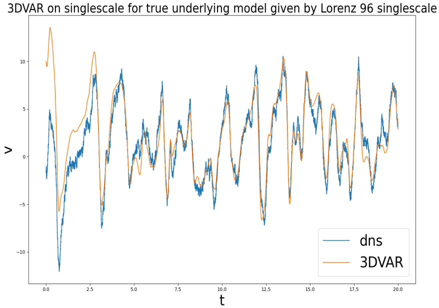
true probabilistic info $\quad \xi_n^\dagger \sim \mathrm{N}(0, \sigma^2 I)$ i.i.d., $\quad \eta_n^\dagger \sim \mathrm{N}(0, \gamma^2 I)$ i.i.d. . $\quad$ (26c)

Caltech

Lorenz ‘96: 3DVAR and Synchronization to Overcome Chaos



(a) $\sigma^2 = 10^{-3}, \gamma^2 = 10^{-3}$



(b) $\sigma^2 = 10^{-1}, \gamma^2 = 10^{-1}$

Ensemble Kalman Filter

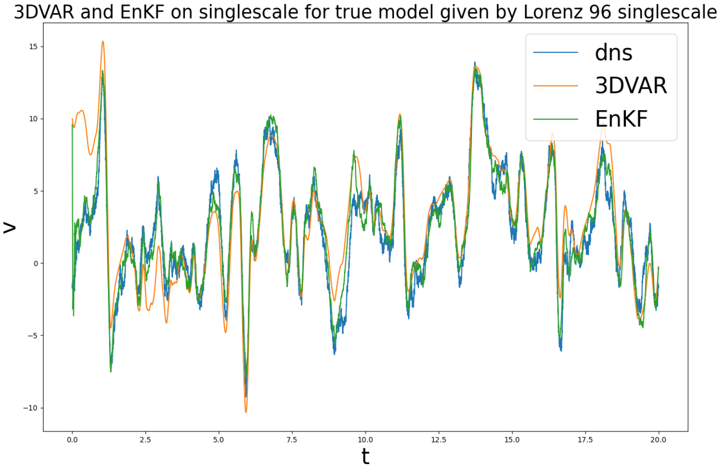


Figure: $\sigma^2 = 10^{-1}, \gamma^2 = 10^{-1}$

Approximate Transport

Lemma

Let $(\widehat{v}_{n+1}, \widehat{y}_{n+1}) \sim \nu_{n+1}$. *If* v_{n+1} *given by* (13) *has mean given by (7a) then*

$$a_{n+1} = (I - A)\mathbb{E}\widehat{v}_{n+1} + \widehat{C}^{vy}_{n+1}(\widehat{C}^{yy}_{n+1})^{-1}y^{\dagger}_{n+1} - \big(B + \widehat{C}^{vy}_{n+1}(\widehat{C}^{yy}_{n+1})^{-1}\big)\mathbb{E}\widehat{y}_{n+1}.$$

Theorem

Let $(\widehat{v}_{n+1}, \widehat{y}_{n+1}) \sim \nu_{n+1}$. *Then* v_{n+1} *defined by* (13) *has covariance (7b) if and only if*

$$F\widehat{C}^{-1}_{n+1}F^{\top} = C_{n+1} - B\big(\widehat{C}^{yy}_{n+1} - (\widehat{C}^{vy}_{n+1})^{\top}\widehat{C}^{-1}_{n+1}\widehat{C}^{vy}_{n+1}\big)B^{\top}, \tag{27}$$

where

$$F = A\widehat{C}_{n+1} + B(\widehat{C}^{vy}_{n+1})^{\top}. \tag{28}$$

Approximate Transport

Define, for positive definite $W \in \mathbb{R}^{d_v \times d_v}$,

$$I_W(A, B) = \frac{1}{2}\mathbb{E}\langle (v - \widehat{v}), W(v - \widehat{v})\rangle.$$

Theorem

$\widetilde{T}$ evaluated at an (A, B) minimizing $I_W(A, B)$ and satisfying (27) and (28) is an optimal transport in the $W-$weighted Euclidean distance. *For any positive definite W, such minimizers satisfy $B = 0$.*

Example: Optimal Transport Solution

$$B = 0, \quad S = \left(C^{\frac{1}{2}}\widehat{C}C^{\frac{1}{2}}\right)^{-\frac{1}{2}}, \quad A = C^{\frac{1}{2}}SC^{\frac{1}{2}}.$$

Caltech

Approximate Transport - Summary

Second Order Approximation

Let μ^{MF} denote the measure associated with using the mean-field map $\widetilde{T}_n$. Note

$$\mu_{n+1} = B_n(QP\mu_n), \quad \mu_0 = \mathsf{N}(m_0, C_0), \tag{29a}$$

$$\mu^G_{n+1} = B_n(GQP\mu^G_n), \quad \mu^G_0 = \mathsf{N}(m_0, C_0), \tag{29b}$$

and

$$\mu^G_{n+1} = G\big((\widetilde{T}_n)^\sharp(QP\mu^G_n)\big), \quad \mu^G_0 = \mathsf{N}(m_0, C_0), \tag{30a}$$

$$\mu^{MF}_{n+1} = (\widetilde{T}_n)^\sharp(QP\mu^{MF}_n), \quad \mu^{MF}_0 = \mathsf{N}(m_0, C_0). \tag{30b}$$

Key question: when is $\widetilde{T}^\sharp(\cdot)$ close to B?

Open Problems

- **Theory:**
  - Determine conditions, under which the true state is well-approximated by the mean or sample path of mean field models based on second order transport.
  - Determine conditions under which the filtering distribution is well-approximated by mean field models based on second order transport.
  - Derive error bounds for particle approximations.
- **Methodology:**
  - Given an ensemble of evaluations of the combined state-observation system, determine the optimal way to estimate the state from an observation sequence, or the filtering distribution.
  - Investigate the role that machine learning might play in addressing the design of algorithms.