Blending Data And Models: Kalman Based Approaches

Andrew Stuart

Computing and Mathematical Sciences
California Institute of Technology

AFOSR, DARPA, EPSRC, NSF, ONR
Allen Philanthropies, Mission Control for Earth, Schmidt Futures

⋆⋆⋆

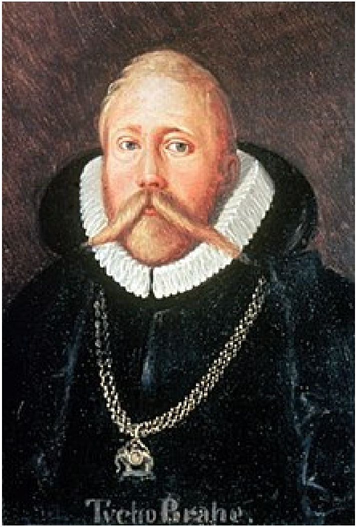
CUHK

January 22nd 2020

Overview

Historical Context: Celestial Mechanics

Brahe



Purely observational data – initially by eye.

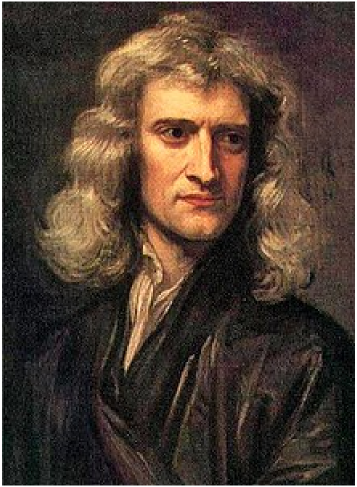
Big data c. 1600s.

Kepler

Mathematical formulae which interpolated Brahe's data.

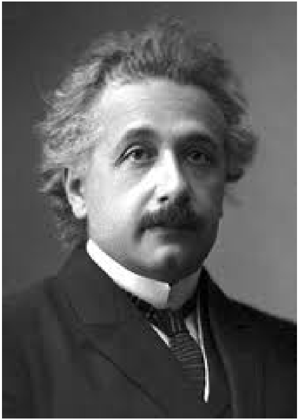
Data-driven model: Kepler's Law.

Newton



Kepler's Law rationalized through Newtonian mechanics.
Led to theory of conservation laws: extrapolation.

Einstein



Discrepancy between data and predictions of Newtonian mechanics.
Mercury perehilion; resolved by special and then general relativity.
The scientific method.

Kalman Methodology & Applications

Kalman Filter (Navigation)

State Space Model

$$\begin{aligned}
\text{Dynamics Model:} \quad & v_{n+1} = Mv_n + \xi_n, \quad n \in \mathbb{Z}^+ \\
\text{Data Model:} \quad & y_{n+1} = Hv_{n+1} + \eta_{n+1}, \quad n \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:} \quad & v_0 \sim N(m_0, C_0), \quad \xi_n \sim N(0, \Sigma), \quad \eta_n \sim N(0, \Gamma) \\
\text{Probabilistic Structure:} \quad & v_0 \perp \{\xi_n\} \perp \{\eta_n\} \text{ independent}
\end{aligned}$$

- Rudolph Kalman.
- J. Basic Engineering **82**(1960); [19].
- $\approx 35,000$ citations (Google Scholar 1/21).
- Apollo 11.
- $Y_n = \{y_\ell\}_{\ell=1}^n$.
- $v_n | Y_n \sim N(m_n, C_n)$.
- $(m_n, C_n) \mapsto (m_{n+1}, C_{n+1})$.

Kalman Filter

Sequential Optimization Perspective

$$\begin{aligned} \text{Predict:} \quad & \widehat{m}_{n+1} = M m_n, \quad n \in \mathbb{Z}^+ \\ \text{Model/Data Compromise:} \quad & J_n(m) = \frac{1}{2}|m - \widehat{m}_{n+1}|^2_{\widehat{C}_{n+1}} + \frac{1}{2}|y_{n+1} - Hm|^2_\Gamma \\ \text{Optimize:} \quad & m_{n+1} = \mathrm{argmin}_m \, J_n(m). \end{aligned}$$

- $|\cdot|_A = |A^{-\frac{1}{2}} \cdot|$ for $A > 0$. Here $|\cdot|$ is Euclidean norm.
- d the state space dimension ($m_n, v_n \in \mathbb{R}^d$).
- Updating $\widehat{C}_{n+1}$ is expensive: $\mathcal{O}(d^2)$ storage.

3DVAR Filter (Weather Forecasting)

State Space Model

$$
\begin{aligned}
\text{Dynamics Model:} \quad & v_{n+1} = \Psi(v_n) + \xi_n, \quad n \in \mathbb{Z}^+ \\
\text{Data Model:} \quad & y_{n+1} = Hv_{n+1} + \eta_{n+1}, \quad n \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:} \quad & v_0 \sim N(m_0, C_0), \quad \xi_n \sim N(0, \Sigma), \quad \eta_n \sim N(0, \Gamma) \\
\text{Probabilistic Structure:} \quad & v_0 \perp \{\xi_n\} \perp \{\eta_n\} \text{ independent}
\end{aligned}
$$

- Andrew Lorenc.
- Introduced in UK Met Office.
- See [25]
- $\{v_n\} \mapsto \{v_{n+1}\}$.

3DVAR

Sequential Optimization Perspective

$$\text{Predict:} \quad \widehat{v}_{n+1} = \Psi(v_n), \quad n \in \mathbb{Z}^+$$

$$\text{Model/Data Compromise:} \quad J_n(v) = \frac{1}{2}|v - \widehat{v}_{n+1}|^2_{\widehat{C}} + \frac{1}{2}|y_{n+1} - Hv|^2_{\Gamma}$$

$$\text{Optimize:} \quad v_{n+1} = \mathrm{argmin}_v \; J_n(v).$$

- $\widehat{C}$ is a fixed model covariance (not updated sequentially).
- $d = \mathcal{O}(10^9)$; $\mathcal{O}(d^2)$ entries of $\widehat{C}$ prohibitive in general.
- $\widehat{C}$ based on climatology + simple, computable, structure.

Ensemble Kalman Filter (Oceanography)

State Space Model

$$
\begin{aligned}
\text{Dynamics Model:} \quad & v_{n+1} = \Psi(v_n) + \xi_n, \quad n \in \mathbb{Z}^+ \\
\text{Data Model:} \quad & y_{n+1} = Hv_{n+1} + \eta_{n+1}, \quad n \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:} \quad & v_0 \sim N(m_0, C_0), \quad \xi_n \sim N(0, \Sigma), \quad \eta_n \sim N(0, \Gamma) \\
\text{Probabilistic Structure:} \quad & v_0 \perp \{\xi_n\} \perp \{\eta_n\} \text{ independent}
\end{aligned}
$$

- Geir Evensen.
- See [11].
- Motivated by extended Kalman filter; see [17, 14].
- Original paper in oceanography.
- Now used in weather forecasting centres worldwide.
- $\{v_n^{(j)}\}_{j=1}^J \mapsto \{v_{n+1}^{(j)}\}_{j=1}^J$.

Ensemble Kalman Filter

Sequential Optimization Perspective

$$\text{Predict:} \quad \widehat{v}_{n+1}^{(j)} = \Psi(v_n^{(j)}) + \xi_n^{(j)}, \quad n \in \mathbb{Z}^+$$

$$\text{Model/Data Compromise:} \quad J_n^{(j)}(v) = \frac{1}{2}|v - \widehat{v}_{n+1}^{(j)}|^2_{\widehat{C}_{n+1}} + \frac{1}{2}|y_{n+1} - Hv|^2_{\Gamma}$$

$$\text{Optimize:} \quad v_{n+1}^{(j)} = \operatorname{argmin}_v \, J_n^{(j)}(v).$$

- $j \in \{1, \ldots, J\}$, J number of ensemble members.
- $\widehat{C}_{n+1}$ is empirical covariance of the $\{\widehat{v}_{n+1}^{(k)}\}$.
- Updating $\widehat{C}_n$ requires only $\mathcal{O}(Jd)$ storage.

Ensemble Kalman Filter (Mathematical Structure)

State Space Model

Dynamics Model: $v_{n+1} = \Psi(v_n) + \xi_n, \quad n \in \mathbb{Z}^+$

Data Model: $y_{n+1} = Hv_{n+1} + \eta_{n+1}, \quad n \in \mathbb{Z}^+$

Probabilistic Structure: $v_0 \sim N(m_0, C_0), \quad \xi_n \sim N(0, \Sigma), \quad \eta_n \sim N(0, \Gamma)$

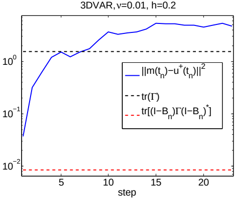
Probabilistic Structure: $v_0 \perp \{\xi_n\} \perp \{\eta_n\}$ independent

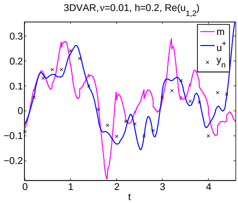
- Sebastian Reich (Potsdam)
- Continuous Time Limits:
- [4, 5, 27, 6].
- Optimal Transport Connections:
- [27, 29, 28].

Weather Forecasting

3DVAR Overcomes Butterfly Effect

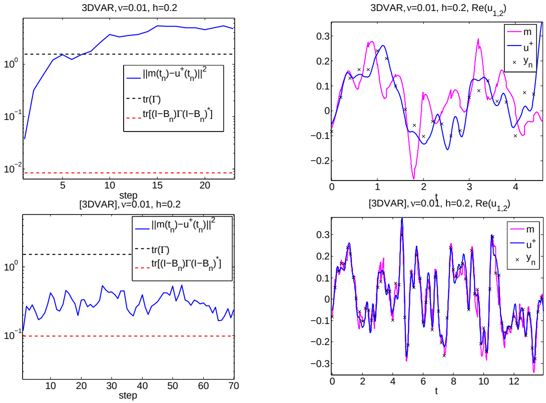
KJH Law and S [24].





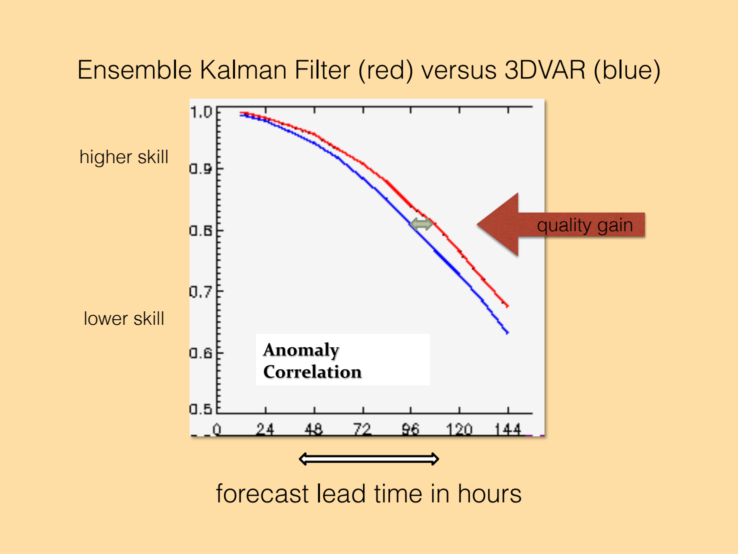
3DVAR Overcomes Butterfly Effect

KJH Law and S [24].



Impact of EnKF over 3DVAR

Courtesy of Roland Potthast (Head of Data Assimilation, DWD)



Ensemble Kalman Inversion

- Chen & Oliver [8]
- Emerick and Reynolds [9]
- Iglesias, Law and S [16]
- Ernst, Sprungk and Starkloff [10]

Inverse Problem

Problem Statement

Find u from y where $\mathrm{G} : \mathcal{U} \mapsto \mathcal{Y}$, η is noise and

$$y = \mathrm{G}(u) + \eta.$$

Main Approaches

$$\textit{Optimization} \quad \Phi(u) = \frac{1}{2}|y - G(u)|_{\Gamma}^{2} + \frac{1}{2}|u|_{\Sigma}^{2};$$

$$\textit{Probability} \quad \mathbb{P}(u|y) \propto \exp\big(-\Phi(u)\big).$$

Inverse Problem

Dynamical Formulation

$$\begin{aligned} \text{Dynamics Model:} \quad & u_{n+1} = u_n, \quad n \in \{0, \cdots, M-1\} \\ \text{Dynamics Model:} \quad & w_{n+1} = \mathsf{G}(u_n), \quad n \in \{0, \cdots, M-1\} \\ \text{Data Model:} \quad & y_{n+1} = w_{n+1} + \eta_{n+1}, \quad n\{0, \cdots, M-1\} \end{aligned}$$

- Evensen moved to Statoil.
- Dean Oliver & Al Reynolds; see [8, 9].
- $y_{n+1} = y, \quad \eta_{n+1} \sim N(0, M\Gamma)$.
- Methodology now widely used in oil industry.
- Methodology now widely used by hydrologists.

Discrete Time Kalman Inversion Algorithm

Covariances

$$C_n^{ww} = \frac{1}{J}\sum_{j=1}^{J}\left(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\right) \otimes \left(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\right), \quad \overline{w}_n = \frac{1}{J}\sum_{j=1}^{J}\mathsf{G}(u_n^{(j)}),$$

$$C_n^{uw} = \frac{1}{J}\sum_{j=1}^{J}\left(u_n^{(j)} - \overline{u}_n\right) \otimes \left(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\right), \quad \overline{u}_n = \frac{1}{J}\sum_{j=1}^{J}u_n^{(j)}.$$

Iteration $n \mapsto n+1$

$$u_{n+1}^{(j)} = u_n^{(j)} + C_n^{uw}\left(C_n^{ww} + M\Gamma\right)^{-1}\left(y - \mathsf{G}(u_n^{(j)})\right)$$

Discrete Time Kalman Inversion Algorithm

Covariances

$$C_n^{ww} = \frac{1}{J}\sum_{j=1}^{J}\big(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\big) \otimes \big(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\big), \quad \overline{w}_n = \frac{1}{J}\sum_{j=1}^{J}\mathsf{G}(u_n^{(j)}),$$

$$C_n^{uw} = \frac{1}{J}\sum_{j=1}^{J}\big(u_n^{(j)} - \overline{u}_n\big) \otimes \big(\mathsf{G}(u_n^{(j)}) - \overline{w}_n\big), \quad \overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}.$$

Iteration $n \mapsto n+1$

$$u_{n+1}^{(j)} = u_n^{(j)} + C_n^{uw}\big(C_n^{ww} + M\Gamma\big)^{-1}\big(y - \mathsf{G}(u_n^{(j)})\big)$$

Continuous Time Limit

$$u_n^{(j)} \approx u^{(j)}(t)|_{t=n/M} : \dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\Big\langle \mathsf{G}(u^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(u^{(j)}) - y\Big\rangle_{\Gamma}\,\big(u^{(k)} - \bar{u}\big)$$

Gradient Flow In Parameter Space

- Ensemble Filtering Continuous Time: Bergemann & Reich (2010a, 2010b, 2012) [4, 5, 6]
- Ensemble Filtering Continuous Time: Reich (2011) [27]
- Connection to Foais/Prodi: Titi and coworkers [15, 2]
- 3DVAR Filtering Continuous Time: Blömker, Law, S & Zygalakis (2013) [7]
- Ensemble Filtering Continuous Time: Kelly, Law & S (2015) [20]
- Ensemble Inversion Continuous Time: Schillings & S (2017) [30]
- Text: Reich & Cotter (2015) [28]
- Text: Law, S & Zygalakis (2015) [23]
- Ensemble Filtering Continuous Time: Lange & Stannat [22]
- Ensemble Square Root Filtering Continuous Time: Lange & Stannat [21]

Gradient Flow In Space Of Probability Measures

- Jordan, Kinderlehrer & Otto 1998 [18]
- Otto 2001 [26]
- Benamou & Brenier 2000 [3]
- Ambrosio, Gigli & Savare 2008 [1]
- Villani 2008 [31]
- Reich & Cotter 2013 [29]
- Garbuno-Inigo, Hoffmann, Li & Stuart 2020 [12]
- Garbuno-Inigo, Nüsken & Reich [13]

Ensemble Kalman Inversion (EKI) – Linear G See [27], [30]

EKI Is Self-Preconditioned Gradient Descent

$$\dot{u}^{(j)} = -C(u)\nabla\Phi_0(u^{(j)}), \quad \Phi_0(u) = \frac{1}{2}\left|y - G(u)\right|_\Gamma^2,$$

$$\bar{u} = \frac{1}{J}\sum_{k=1}^{J} u^{(k)}, \quad C(u) = \frac{1}{J}\sum_{k=1}^{J}\left(u^{(k)} - \bar{u}\right) \otimes \left(u^{(k)} - \bar{u}\right).$$

Theorem [30]

EKI minimizes Φ_0 over a finite dimensional subspace determined by the initial conditions $\{u^{(j)}(0)\}_{j=1}^{J}$. The rate of convergence is $\mathcal{O}(1/t)$.

Ensemble Kalman Sampling – Linear G See [12]

EKS Is Self-Preconditioned Langevin Equation

$$\dot{u}^{(j)} = -C(u)\nabla\Phi(u^{(j)}) + \sqrt{2C(u)}\dot{W}^{(j)}, \quad \Phi(u) = \frac{1}{2}|y - G(u)|_\Gamma^2 + \frac{1}{2}|u|_\Sigma^2,$$

$$\bar{u} = \frac{1}{J}\sum_{k=1}^{J} u^{(k)}, \quad C(u) = \frac{1}{J}\sum_{k=1}^{J}\Big(u^{(k)} - \bar{u}\Big) \otimes \Big(u^{(k)} - \bar{u}\Big).$$

Mean Field Limit: Nonlinear Nonlocal Fokker-Planck Eq.

$$\dot{u} = -\mathcal{C}(\rho)\nabla\Phi(u) + \sqrt{2C(\rho)}\dot{W},$$

$$\mathcal{C}(\rho) = \int (u - \bar{u}) \otimes (u - \bar{u})\rho(u,t)du, \quad \bar{u} = \int u\rho(u,t)du,$$

$$\partial_t\rho = \nabla \cdot \big(\rho\,\mathcal{C}(\rho)\,\nabla\Phi\big) + \mathcal{C}(\rho) : D^2\rho, \quad \rho(0) = \rho_0.$$

Nonlinear Nonlocal Fokker-Planck Equation See [18], [12]

Theorem [12]

The nonlinear Fokker-Planck equation may be written as

$$\partial_t \rho = \nabla \cdot \left(\rho \mathcal{C}(\rho) \nabla \frac{\delta \mathcal{E}}{\delta \rho} \right) \;, \mathcal{E}(\rho) = \int (\Phi + \ln \rho)\, \rho \, \mathrm{d}u.$$

- Gradient flow in $\mathcal{P}_+$ (probability measures) w.r.t. the (next slide) Kalman-Wasserstein metric;
- in linear setting convergence to equilibrium $\propto \exp(-\Phi)$ occurs at exponential rate $\exp(-t)$, indepedently of the linear inverse problem being solved.

Nonlinear Nonlocal Fokker-Planck Equation (Metric) [29],[12]

Kalman-Wasserstein Metric Tensor (Otto [18, 26])

Define $g_{\rho,\mathcal{C}} \; : \; T_\rho \mathcal{P}_+ \times T_\rho \mathcal{P}_+ \to \mathbb{R}$ by

$$g_{\rho,\mathcal{C}}(\sigma_1, \sigma_2) := \int_\Omega \langle \nabla \psi_1 \, , \, \mathcal{C}(\rho) \nabla \psi_2 \rangle \; \rho \, \mathrm{d}x,$$

where $\sigma_i = -\nabla \cdot (\rho \mathcal{C}(\rho) \nabla \psi_i) \in T_\rho \mathcal{P}_+$ for $i = 1, 2$.

Kalman-Wasserstein Metric (Benamou-Brenier [3])

For ρ^0, $\rho^1 \in \mathcal{P}_+$, $\mathcal{W}_\mathcal{C} : \mathcal{P}_+ \times \mathcal{P}_+ \to \mathbb{R}$

$$\mathcal{W}_\mathcal{C}(\rho^0, \rho^1)^2 := \inf_{(\rho_t, \psi_t)} \int_0^1 \int_\Omega \langle \nabla \psi_t \, , \, \mathcal{C}(\rho_t) \nabla \psi_t \rangle \; \rho_t \, \mathrm{d}x$$

$$\text{subject to} \quad \partial_t \rho_t + \nabla \cdot (\rho_t \mathcal{C}(\rho_t) \nabla \psi_t) = 0, \;\; \rho_0 = \rho^0, \;\; \rho_1 = \rho^1,$$

Closing

Conclusions: Kalman Methodologies

- Introduced in 1960 by Rudolph Kalman.
- Basic algorithm generalized in many directions.
- Applications in numerous fields:
  - Navigation;
  - Weather forecasting;
  - Oceanography;
  - Hydrology, Subsurface flow;
  - Medical imaging, Machine learning $\cdots$.
- Developing as a general methodology for state estimation.
- Developing as a general methodology for inverse problems:
  - Gradient flow structure: parameter space;
  - Gradient flow structure: probability space.
- Connections to Wasserstein gradient flows, optimal transport.
- Many open mathematical questions.

References I

[1] L. Ambrosio, N. Gigli, and G. Savaré.
Gradient flows: in metric spaces and in the space of probability measures.
Springer Science & Business Media, 2008.

[2] A. Azouani, E. Olson, and E. S. Titi.
Continuous data assimilation using general interpolant observables.
Journal of Nonlinear Science, 24(2):277–304, 2014.

[3] J.-D. Benamou and Y. Brenier.
A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem.
Numerische Mathematik, 84(3):375–393, 2000.

[4] K. Bergemann and S. Reich.
A localization technique for ensemble Kalman filters.
Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 136(648):701–707, 2010.

References II

[5] K. Bergemann and S. Reich.
A mollified ensemble Kalman filter.
Quarterly Journal of the Royal Meteorological Society, 136(651):1636–1643, 2010.

[6] K. Bergemann and S. Reich.
An ensemble Kalman-Bucy filter for continuous data assimilation.
Meteorologische Zeitschrift, 21(3):213–219, 2012.

[7] D. Blömker, K. Law, A. M. Stuart, and K. C. Zygalakis.
Accuracy and stability of the continuous-time 3DVAR filter for the navier–stokes equation.
Nonlinearity, 26(8):2193, 2013.

[8] Y. Chen and D. S. Oliver.
Ensemble randomized maximum likelihood method as an iterative ensemble smoother.
Mathematical Geosciences, 44(1):1–26, 2012.

References III

[9] A. A. Emerick and A. C. Reynolds.
Ensemble smoother with multiple data assimilation.
Computers & Geosciences, 55:3–15, 2013.

[10] O. G. Ernst, B. Sprungk, and H.-J. Starkloff.
Analysis of the ensemble and polynomial chaos kalman filters in bayesian inverse problems.
SIAM/ASA Journal on Uncertainty Quantification, 3(1):823–851, 2015.

[11] G. Evensen.
Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics.
Journal of Geophysical Research: Oceans, 99(C5):10143–10162, 1994.

[12] A. Garbuno-Inigo, F. Hoffmann, W. Li, and A. M. Stuart.
Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler.
SIAM Journal on Applied Dynamical Systems, 19(1):412–441, 2020.

References IV

[13] A. Garbuno-Inigo, N. Nüsken, and S. Reich.
Affine invariant interacting Langevin dynamics for bayesian inference.
SIAM Journal on Applied Dynamical Systems, 19(3):1633–1658, 2020.

[14] M. Ghil, S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson.
Application of estimation theory to numerical weather prediction.
Dynamic Meteorology: Data Assimilation Methods, pages 139–224, 1981.

[15] K. Hayden, E. Olson, and E. S. Titi.
Discrete data assimilation in the lorenz and 2d navier–stokes equations.
Physica D: Nonlinear Phenomena, 240(18):1416–1425, 2011.

[16] M. A. Iglesias, K. J. Law, and A. M. Stuart.
Ensemble Kalman methods for inverse problems.
Inverse Problems, 29(4):045001, 2013.

[17] A. Jazwinski.
Stochastic Processes and Filtering Theory.
Academic Press, 1970.

References V

[18] R. Jordan, D. Kinderlehrer, and F. Otto.
The variational formulation of the fokker–planck equation.
SIAM journal on mathematical analysis, 29(1):1–17, 1998.

[19] R. Kalman.
A new approach to linear filtering and prediction problems.
Journal of Basic Engineering, 82:35–45, 1960.

[20] D. T. Kelly, K. J. Law, and A. M. Stuart.
Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time.
Nonlinearity, 27(10):2579, 2014.

[21] T. Lange and W. Stannat.
On the continuous time limit of ensemble square root filters.
arXiv preprint arXiv:1910.12493, 2019.

[22] T. Lange and W. Stannat.
On the continuous time limit of the ensemble Kalman filter.
arXiv preprint arXiv:1901.05204, 2019.

References VI

[23] K. Law, A. Stuart, and K. Zygalakis.
Data assimilation.
Cham, Switzerland: Springer, 2015.

[24] K. J. Law and A. M. Stuart.
Evaluating data assimilation algorithms.
Monthly Weather Review, 140(11):3757–3782, 2012.

[25] A. C. Lorenc.
Analysis methods for numerical weather prediction.
Quarterly Journal of the Royal Meteorological Society, 112(474):1177–1194, 1986.

[26] F. Otto.
The geometry of dissipative evolution equations: the porous medium equation.
2001.

References VII

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

[28] S. Reich and C. Cotter.
Probabilistic forecasting and Bayesian data assimilation.
Cambridge University Press, 2015.

[29] S. Reich and C. J. Cotter.
Ensemble filter techniques for intermittent data assimilation.
Large Scale Inverse Problems. Computational Methods and Applications in the Earth Sciences, 13:91–134, 2013.

[30] C. Schillings and A. M. Stuart.
Analysis of the ensemble Kalman filter for inverse problems.
SIAM Journal on Numerical Analysis, 55(3):1264–1290, 2017.

[31] C. Villani.
Optimal transport: old and new, volume 338.
Springer Science & Business Media, 2008.

Ensemble Kalman Inversion (EKI)

Continuous Time Formulation

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\left\langle \mathsf{G}(u^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(u^{(j)}) - y \right\rangle_{\Gamma} \left(u^{(k)} - \bar{u}\right)$$

$$\bar{u} = \frac{1}{J}\sum_{k=1}^{J} u^{(k)}, \quad \bar{\mathsf{G}} = \frac{1}{J}\sum_{k=1}^{J} G(u^{(k)}),$$

$$C(u) = \frac{1}{J}\sum_{k=1}^{J}\left(u^{(k)} - \bar{u}\right) \otimes \left(u^{(k)} - \bar{u}\right).$$

Linear Approximation

$$(\mathsf{G}(u^{(k)}) - \bar{\mathsf{G}}) \approx D\mathsf{G}(u^{(j)})(u^{(k)} - \bar{u}).$$

Ensemble Kalman Sampling (EKS)

Continuous Time Formulation: Put EKI in a heat bath

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\left\langle \mathsf{G}(u^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(u^{(j)}) - y \right\rangle_{\Gamma} \left(u^{(k)} - \bar{u}\right) - C(u)\Sigma^{-1}u^{(j)} + \sqrt{2C(u)}\dot{W}^{(j)},$$

$$C(u) = \frac{1}{J}\sum_{k=1}^{J}\left(u^{(k)} - \bar{u}\right) \otimes \left(u^{(k)} - \bar{u}\right).$$

Gradient Structure Of NNLFP in MFL

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\rho) = -\int \rho \left|\mathcal{C}(\rho)^{\frac{1}{2}}\nabla(\Phi + \ln\rho)\right|^2 \mathrm{d}u = -g_{\rho,\mathcal{C}}(\partial_t\rho, \partial_t\rho).$$