The Mean-Field Ensemble Kalman Filter

Andrew Stuart

Computing and Mathematical Sciences
California Institute of Technology

DoD, NSF, ONR

⋆ ⋆ ⋆

Leçons Jacques-Louis Lions Lectures
Sorbonne Université, Paris

December 12th–14th 2023

Collaborators

- Edo Calvello (Caltech)
- Sebastian Reich (Potsdam)

Reference: E. Calvello, S. Reich, and S

Ensemble Kalman Methods: a Mean-Field Perspective. arXiv:2209.11371 (2022).

- José Carrillo (Oxford)
- Franca Hoffmann (Caltech)
- Urbain Vaes (CERMICS)

Reference: J. A. Carrillo, F. Hoffmann, S, and U. Vaes

The Ensemble Kalman Filter in the Near-Gaussian Setting. arXiv:2212.13239 (2022).

Overview

Mean-Field Optimization Perspective

Optimization: Albers, Blancquart, Levine, Seylabi and S [1] (2022)

Mean-Field: Calvello, Reich and S [3] (2022)

Kalman Filter (Navigation)

State Space Model

$$\text{Dynamics Model:} \quad v_{n+1} = Mv_n + \xi_n, \quad n \in \mathbb{Z}^+$$

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

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

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

- Rudolph Kalman [19] (1960).
- $\approx 43,500$ citations (Google Scholar 12/23).
- Apollo 11.
- The Algorithm:
- $Y_n^\dagger = \{y_\ell^\dagger\}_{\ell=1}^n$.
- $v_n^\dagger | Y_n^\dagger \sim \mathrm{N}(m_n, C_n)$.
- $(m_n, C_n) \mapsto (m_{n+1}, C_{n+1})$.

Kalman Filter

Sequential Optimization Viewpoint

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

- $|\cdot|$ Euclidean norm on any $\mathbf{R}^r$ and induced matrix norm.
- $|\cdot|_A = |A^{-\frac{1}{2}} \cdot|$ for $A > 0$ spd.
- 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

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

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

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

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

- Andrew Lorenc [23] (1986).
- $\approx 2,000$ citations (Google Scholar 12/23).
- Introduced in UK Met Office.
- The Algorithm:
- $\{v_n\} \mapsto \{v_{n+1}\}$.
- Given $Y_n^\dagger$ want $v_n \approx v_n^\dagger$. (Again $Y_n^\dagger = \{y_\ell^\dagger\}_{\ell=1}^n$.)

3DVAR

Sequential Optimization Viewpoint

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

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

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

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

Ensemble Kalman Filter

State Space Model

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

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

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

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

- Geir Evensen [11] (1994).
- $\approx 6,000$ citations (Google Scholar 12/23).
- Originally ocean dynamics; now weather.
- $\mu_n := \mathrm{Law}(v_n^\dagger | Y_n^\dagger)$. (Here $Y_n^\dagger = \{y_\ell^\dagger\}_{\ell=1}^n$.)
- Mean-Field Algorithm:
- $(v_n, \mu_n^{EK}) \mapsto (v_{n+1}, \mu_{n+1}^{EK})$. $\quad \mu_n^{EK} := \mathrm{Law}(v_n)$.
- When is this approximation valid: $\mu_n^{EK} \approx \mu_n$?

Mean-Field Ensemble Kalman Filter

Sequential Optimization Viewpoint

$$\begin{aligned}
\text{Predict:} &\quad \widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \quad n \in \mathbb{Z}^+ \\
\text{Model/Data Compromise:} &\quad J_n(v) = \frac{1}{2}|v - \widehat{v}_{n+1}|^2_{\widehat{C}_{n+1}} + \frac{1}{2}|y^\dagger_{n+1} + \eta_{n+1} - Hv|^2_\Gamma \\
\text{Optimize:} &\quad v_{n+1} = \mathrm{argmin}_v \; J_n(v).
\end{aligned}$$

In What Sense Is This A Mean-Field Model?

- $\mu_n^{EK} := \mathrm{Law}(v_n)$.
- $\widehat{C}_{n+1}$ is the covariance under $\widehat{\mu}^{EK}_{n+1} := \mathrm{Law}(\widehat{v}_{n+1})$.
- $|\cdot|$ Euclidean norm on $\mathbf{R}^r$ and induced matrix norm (any r).
- $|\cdot|_A = |A^{-\frac{1}{2}} \cdot|$ for $A > 0$ spd.

Mean-Field Ensemble Kalman Filter

Sequential Optimization Viewpoint

$$\begin{aligned}
\text{Predict:} \quad & \widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \quad n \in \mathbb{Z}^+ \\
\text{Model/Data Compromise:} \quad & J_n(v) = \frac{1}{2}|v - \widehat{v}_{n+1}|^2_{\widehat{C}_{n+1}} + \frac{1}{2}|y^\dagger_{n+1} + \eta_{n+1} - Hv|^2_\Gamma \\
\text{Optimize:} \quad & v_{n+1} = \mathrm{argmin}_v \; J_n(v).
\end{aligned}$$

In What Sense Is This A Mean-Field Model?

- $\mu_n^{EK} := \mathrm{Law}(v_n)$.
- $\widehat{C}_{n+1}$ is the covariance under $\widehat{\mu}_{n+1}^{EK} := \mathrm{Law}(\widehat{v}_{n+1})$.
- $|\cdot|$ Euclidean norm on $\mathbf{R}^r$ and induced matrix norm (any r).
- $|\cdot|_A = |A^{-\frac{1}{2}} \cdot|$ for $A > 0$ spd.
-
- In practice: use $j \in \{1, \ldots, J\}$, J number of ensemble members.
- Use resulting ensemble $\widehat{v}_{n+1}^{(j)}$ to estimate $\widehat{C}_{n+1}$.

Summary Of Optimization Perspective

Nudging

$$\begin{aligned}
\text{Prediction:} \quad & \widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \\
\text{Analysis:} \quad & v_{n+1} = \widehat{v}_{n+1} + K(y^\dagger_{n+1} - H\widehat{v}_{n+1}) + K\eta_{n+1}, \\
\text{3DVAR:} \quad & K \quad \text{constant}, \quad \text{no noise}, \\
\text{EnKF:} \quad & K = K(\widehat{\mu}^{EK}_{n+1}), \quad \widehat{\mu}^{EK}_{n+1} = \mathrm{Law}(\widehat{v}_{n+1}).
\end{aligned}$$

Two Goals

$$\begin{aligned}
\text{Control (3DVAR, EnKF):} \quad & |v_n - v^\dagger_n| \ll 1, \quad \text{next slide.} \\
\text{UQ (EnKF):} \quad & \mu^{EK}_n \approx \mu_n = \mathrm{Law}(v^\dagger_n | Y^\dagger_n), \quad \text{rest of talk.}
\end{aligned}$$

3DVAR and Small Noise

Synchronization and Lorenz '63 Pecora and Carroll [24] (1990)

Synchronization and Navier-Stokes Hayden, Olson and Titi [15] (2011)

Theorem Law, Shukla and S [21] (2012)

Assume synchronization and small noise $\mathcal{O}(\epsilon)$ in truth, no noise in filter. Consider 3DVAR with $K = \gamma H^\star$ and $|\gamma - 1| \le 1$. Then

$$\limsup_{n\to\infty} \mathbb{E}\left|v_n - v_n^\dagger\right|^2 \le C\epsilon^2.$$

Corollary Sanz-Alonso and S [29] (2015)

Assume synchronization and small noise $\mathcal{O}(\epsilon)$ in truth, no noise in filter. The true filtering distribution $\mu_n = \mathrm{Law}(v_n^\dagger | Y_n^\dagger)$ satisfies

$$\limsup_{n\to\infty} \mathbb{E}\left|\mathbb{E}^{v\sim\mu_n} v - v_n^\dagger\right|^2 \le C\epsilon^2.$$

Probabilistic Perspective: True Filter

Filtering: Doucet, de Freitas and Gordon [10] (2004)

Unconditioned Dynamics

State-Data Viewpoint (Nonlinear)

$$\textbf{State:} \quad v_{n+1} = \Psi(v_n) + \xi_n, \qquad \xi_n \sim \mathrm{N}(0, \Sigma), \text{ i.i.d. },$$
$$\textbf{Data:} \quad y_{n+1} = h(v_{n+1}) + \eta_{n+1}, \qquad \eta_{n+1} \sim \mathrm{N}(0, \Gamma), \text{ i.i.d. }.$$

$$v_0 \sim \mathrm{N}(m_0, C_0), \quad v_0 \perp\!\!\!\perp \{\xi_n\}_{n\in\mathbb{N}} \perp\!\!\!\perp \{\eta_{n+1}\}_{n\in\mathbb{N}}$$

Probability Viewpoint (Linear)

$$v_n \sim \pi_n, \quad (v_n, y_n) \sim \mathfrak{r}_n,$$
$$\pi_{n+1} = P\pi_n,$$
$$\mathfrak{r}_{n+1} = Q\pi_{n+1}$$

Key Linear Operators on $\mathcal{P}$

Definition of $\mathcal{P}$, $\mathcal{G}$

- $\mathcal{P}(\mathbf{R}^r)$: all probability measures on $\mathbf{R}^r$.
- $\mathcal{G}(\mathbf{R}^r)$: all Gaussian probability measures on $\mathbf{R}^r$.

Definition of P

$P: \mathcal{P}(\mathbf{R}^d) \to \mathcal{P}(\mathbf{R}^d)$ is the linear operator:

$$P\pi(u) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \int \exp\left(-\frac{1}{2}|u - \Psi(v)|_\Sigma^2\right) \pi(v)\,\mathrm{d}v.$$

Definition of Q

$Q: \mathcal{P}(\mathbf{R}^d) \to \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K)$ is the linear operator:

$$Q\pi(u, y) = \frac{1}{\sqrt{(2\pi)^K \det \Gamma}} \exp\left(-\frac{1}{2}|y - h(u)|_\Gamma^2\right) \pi(u).$$

Key Nonlinear Operator on $\mathcal{P}$

Probability Viewpoint (Nonlinear)

$$Y_n^\dagger = \{y_\ell^\dagger\}_{\ell=1}^n, \quad v_n | Y_n^\dagger \sim \mu_n.$$

$$\begin{aligned}
\widehat{\mu}_{n+1} &= P\mu_n, \quad v_{n+1}|Y_n^\dagger \sim \widehat{\mu}_{n+1} \\
\rho_{n+1} &= Q\widehat{\mu}_{n+1}, \quad (v_{n+1}, y_{n+1})|Y_n^\dagger \sim \rho_{n+1} \\
\mu_{n+1} &= B(\rho_{n+1}; y_{n+1}^\dagger), \quad \text{conditioning.}
\end{aligned}$$

Conditioning (Nonlinear)

$B(\bullet; y^\dagger)\colon \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K) \to \mathcal{P}(\mathbf{R}^d)$ describes conditioning on observation $y = y^\dagger$:

$$B(\rho; y^\dagger)(u) = \frac{\rho(u, y^\dagger)}{\int_{\mathbf{R}^d} \rho(u, y^\dagger)\,\mathrm{d}u}.$$

The True Filter

Sequential Interleaving of Prediction and Bayes Theorem

$P\mu_n$ is prior prediction; $L(\bullet; y^\dagger) := B(\bullet; y^\dagger) \circ Q$ maps prior to posterior:

$$\mu_{n+1} = B(QP\mu_n; y^\dagger_{n+1}),$$
$$\mu_{n+1} = L(P\mu_n; y^\dagger_{n+1}).$$

Particle Filter Doucet [10] (2015)

$S^J : \mathcal{P}(\mathbf{R}^r) \times \Omega \to \mathcal{P}(\mathbf{R}^r)$ is empirical approximation operator:

$$S^J\mu = \frac{1}{J}\sum_{j=1}^{J} \delta_{v_j}, \quad v_j \sim \mu \text{ i.i.d.}.$$

S^J : is thus a random approximation of the identity operator on $\mathcal{P}(\mathbf{R}^r)$.

$$\mu^{PF}_{n+1} = L(S^J P\mu^{PF}_n; y^\dagger_{n+1}).$$

Particle Filter Convergence

Theorem Del Moral [7] (1997), Del Moral and Guionnet [9] (2001)

$$\sup_{0 \leq n \leq N} d(\mu_n, \mu_n^{PF}) \leq \frac{C}{\sqrt{J}}.$$

Comments on Proof Rebschini and Van Handel [25] (2015),

- Metric $d(\cdot, \cdot)$ on random probability measures:
- $d(\mu, \nu)^2 = \sup_{|f| \leq 1} \mathbb{E}\left|\mu(f) - \nu(f)\right|^2.$
- Reduces to TV between deterministic measures.
- Consistency + Stability Implies Convergence.
- Consistency: $d(S^J\mu, \mu) \leq \frac{1}{\sqrt{J}}$.
- Stability: P, L Lipschitz in $d(\cdot, \cdot)$.
- Suffers from weight collapse.

Weights

Particle Filter (Weight Collapse)

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)}, \quad v_n^{(j)} \sim \mu_n^{PF},$$
$$\ell_{n+1}^{(j)} = \exp\left(-\frac{1}{2}\big|y_{n+1}^{\dagger} - h\big(\widehat{v}_{n+1}^{(j)}\big)\big|_{\Gamma}^{2}\right),$$
$$\mu_{n+1}^{PF} = \sum_{j=1}^{J} w_{n+1}^{(j)} \delta_{\widehat{v}_{n+1}^{(j)}}, \quad w_{n+1}^{(j)} = \ell_{n+1}^{(j)} \Big/ \Big(\sum_{m=1}^{J} \ell_{n+1}^{(m)}\Big).$$

Ensemble Kalman Filter (No Weight Collapse!)

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)}, \quad v_n^{(j)} \sim \mu_n^{EK},$$
$$\widehat{y}_{n+1}^{(j)} = h\big(\widehat{v}_{n+1}^{(j)}\big) + \eta_{n+1}^{(j)},$$
$$v_{n+1}^{(j)} = \widehat{v}_{n+1}^{(j)} + \mathcal{C}^{vy}\big(\rho_{n+1}^{EK,J}\big)\mathcal{C}^{yy}\big(\rho_{n+1}^{EK,J}\big)^{-1}\big(y_{n+1}^{\dagger} - \widehat{y}_{n+1}^{(j)}\big),$$
$$\mu_{n+1}^{EK} = \frac{1}{J}\sum_{j=1}^{J} \delta_{v_{n+1}^{(j)}}.$$

Probabilistic Perspective: Ensemble Kalman Filter

Mean-Field: Calvello, Reich and S [3] (2022)

The Mean Field Ensemble Kalman Filter

Comparison With True Filter

$$\mu_{n+1}^{EK} = T(QP\mu_n^{EK}; y_{n+1}^\dagger),$$
$$\mu_{n+1} = B(QP\mu_n; y_{n+1}^\dagger).$$

Observations About T

- Choose T to recover mean-field EnKF;
- T defined through pushforward;
- Key is to understand when $T \approx B$.
- $T \equiv B$ on $\mathcal{G}(\mathbb{R}^d \times \mathbb{R}^K)$.

Approximate Conditioning

Block Form Of State-Data Covariance

Write covariance under $\rho \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K)$ as:

$$\mathrm{cov}_\rho = \begin{pmatrix} \mathcal{C}^{vv}(\rho) & \mathcal{C}^{vy}(\rho) \\ \mathcal{C}^{vy}(\rho)^\top & \mathcal{C}^{yy}(\rho) \end{pmatrix}.$$

Key Nonlinear Operator on $\mathcal{P}$ ($T(\bullet; y^\dagger) \equiv B(\bullet; y^\dagger)$ for Gaussian inputs)

$T(\bullet; y^\dagger) \colon \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K) \to \mathcal{P}(\mathbf{R}^d)$ approximates conditioning of ρ on $y = y^\dagger$:

$$\begin{aligned} \mathfrak{T}(\bullet, \bullet; \rho, y^\dagger) &\colon \mathbf{R}^d \times \mathbf{R}^K \to \mathbf{R}^d; \\ (v, y) &\mapsto v + \mathcal{C}^{vy}(\rho)\mathcal{C}^{yy}(\rho)^{-1}(y^\dagger - y), \\ T(\rho; y^\dagger) &= \big(\mathfrak{T}(\bullet, \bullet; \rho, y^\dagger)\big)_\sharp \rho, \\ \mu^{EK}_{n+1} &= T(QP\mu^{EK}_n; y^\dagger_{n+1}). \end{aligned}$$

Mean Field EnKF & Maps on Probability Measures

State-Data Space Picture

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n,$$
$$\widehat{y}_{n+1} = h(\widehat{v}_{n+1}) + \eta_{n+1},$$
$$v_{n+1} = \widehat{v}_{n+1} + \mathcal{C}^{vy}\big(\rho^{EK}_{n+1}\big)\mathcal{C}^{yy}\big(\rho^{EK}_{n+1}\big)^{-1}\big(y^{\dagger}_{n+1} - \widehat{y}_{n+1}\big).$$

Remarks

- Recovers mean-field EnKF in nudging form.
- Use equal weight particle approximation to implement.

Main Theorem

Relating The True and Ensemble Kalman Filters

Main Theorem: Carrillo, Hoffmann, S and Vaes [4] (2022)

The Mean Field Ensemble Kalman Filter

Comparison With True Filter

$$\mu_{n+1}^{EK} = T(QP\mu_n^{EK}; y_{n+1}^{\dagger}),$$
$$\mu_{n+1} = B(QP\mu_n; y_{n+1}^{\dagger}),$$

Remarks

- When is $T(\bullet; y^{\dagger}) \approx B(\bullet; y^{\dagger})$?
- A form of Consistency.
- Try Consistency + Stability Implies Convergence.

Gaussian Projection

Best Gaussian Approximation in KL

$$G : \mathcal{P} \to \mathcal{G},$$
$$G\pi = \mathit{argmin}_{\mathfrak{p} \in \mathcal{G}} \; d_{\mathrm{KL}}(\pi \| \mathfrak{p}).$$

Best Gaussian Approximation in KL

$$G\pi = \mathsf{N}(\mathrm{mean}_\pi, \mathrm{cov}_\pi).$$

The Mean Field Ensemble Kalman Filter

Comparison With True Filter

$$\mu_{n+1}^{EK} = T(QP\mu_n^{EK}; y_{n+1}^\dagger),$$
$$\mu_{n+1} = B(QP\mu_n; y_{n+1}^\dagger).$$

Key Fact

$$T(G\rho; y^\dagger) = B(G\rho; y^\dagger) \quad \forall (\rho, y^\dagger) \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^{EK}) \times \mathbf{R}^{EK}.$$

Pushforward beyond the Gaussian setting (continuous time): Yang, Mehta and Meyn [33] (2013)

Pushforward beyond the Gaussian setting (discrete time): Spantini, Baptista and Marzouk [32] (2022)

Closness of Exact Filter and EnKF

Weighted TV Metric

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

$$d_g(\mu_1, \mu_2) = \sup_{|f| \le g} \big|\mu_1[f] - \mu_2[f]\big|, \quad \mu[f] = \int f(u)\mu(du).$$

Definition

Measure of how close true filter $\{\mu_n\}$ is to being Gaussian:

$$\varepsilon := \sup{}_{0 \le n \le N} \; d_g(GQP\mu_n, QP\mu_n).$$

Closness of Exact Filter and EnKF

Weighted TV Metric

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

$$d_g(\mu_1, \mu_2) = \sup_{|f| \le g} \left|\mu_1[f] - \mu_2[f]\right|, \quad \mu[f] = \int f(u)\mu(du).$$

Definition

Measure of how close true filter $\{\mu_n\}$ is to being Gaussian:

$$\varepsilon := \sup_{0 \le n \le N} d_g(GQP\mu_n, QP\mu_n).$$

Theorem Carrillo, Hoffmann, S and Vaes [4] (2022)

Let $\mu_0^{EK} = \mu_0$ and assume that $\|\Psi\|_{L^\infty}, \|h\|_{L^\infty}$ and $|h|_{C^{0,1}}$ are bounded by r.
Then there is $C := C(N, r) > 0$:

$$\sup_{0 \le n \le N} d_g(\mu_n, \mu_n^{EK}) \le C\varepsilon.$$

Closness of Exact Filter and EnKF

Assumptions C

- Data $Y_j^\dagger$ lies in set

$$B_y := \left\{ Y^\dagger \in \mathbf{R}^{KJ} : \max_{0 \le j \le J} |y_j^\dagger| \le \kappa_y \right\}.$$

- $\Psi_0 \colon \mathbf{R}^d \to \mathbf{R}^d$ and $\mathbf{h}_0 \colon \mathbf{R}^d \to \mathbf{R}^K$ are constant functions and denote by $B_{\Psi,h}(r)$ the set (Ψ, h) satisfying $\Psi \in B_{L^\infty}(\Psi_0, r)$, $h \in B_{L^\infty}(h_0, r)$.

Corollary Carrillo, Hoffmann, S and Vaes [4] (2022)

Let Assumptions T of Theorem hold and Assumptions C. Then for any $\epsilon > 0$ there is $\delta > 0$ such that

$$\sup_{Y^\dagger \in B_y} \sup_{(\Psi,h) \in B_{\Psi,h}(\delta)} \sup_{0 \le n \le N} d_g(\mu_n, \mu_n^{EK}) \le \epsilon.$$

Proof of Theorem (Lipschitz Estimates)

Linear Maps P, Q

The maps P, Q are globally Lipschitz on $\mathcal{P}(\mathbf{R}^d)$ in d_g.

Proof of Theorem (Stability Estimate I)

Conditioning is not Lipschitz stable. However, if Ψ is bounded:

Nonlinear Conditioning Map $B^{y^\dagger}$

The maps $B^{y^\dagger}(\bullet) := B(\bullet; y^\dagger)$ satisfy:

$$\forall \mu \in \mathcal{P}(\mathbf{R}^d)$$

$$d_g\left(B^{y^\dagger}(GQP\mu), B^{y^\dagger}(QP\mu)\right) \leq \ell_B \, d_g(GQP\mu, QP\mu)\,.$$

Proof of Theorem (Stability Estimate II)

Let $\mathcal{P}_R$ denote the following subset of probability measures

$$\mathcal{P}_R(\mathbf{R}^r) = \left\{\mu \in \mathcal{P}(\mathbf{R}^r) : \max\left\{|\mathrm{mean}(\mu)|, |\mathrm{cov}(\mu)|^{\frac{1}{2}}, |\mathrm{cov}(\mu)|^{-\frac{1}{2}}\right\} \leq R\right\}.$$

Using linearity of $\mathfrak{T}$, which defines nonlinear map $T^{y^\dagger}$:

Approximate Nonlinear Conditioning Map $T^{y^\dagger}$

The maps $T^{y^\dagger}(\bullet) := T(\bullet; y^\dagger)$ satisfy, using Ψ bounded,

$$\forall(\mu, \rho) \in \mathcal{P}(\mathbf{R}^d) \times \mathcal{P}_R(\mathbf{R}^d \times \mathbf{R}^K),$$
$$d_g(T^{y^\dagger}(QP\mu), T^{y^\dagger}(\rho)) \leq \ell_T(R)\, d_g(QP\mu, \rho),$$

Proof of Theorem

Since $T^{y^\dagger_{n+1}}(G\bullet) = B^{y^\dagger_{n+1}}(G\bullet)$ we have

$$
\begin{aligned}
d_g(\mu^{EK}_{n+1}, \mu_{n+1}) &= d_g\left(T^{y^\dagger_{n+1}}(QP\mu^{EK}_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\leq d_g\left(T^{y^\dagger_{n+1}}(QP\mu^{EK}_n), T^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\qquad + d_g\left(T^{y^\dagger_{n+1}}(QP\mu_n), T^{y^\dagger_{n+1}}(GQP\mu_n)\right) \\
&\qquad\qquad + d_g\left(T^{y^\dagger_{n+1}}(GQP\mu_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\leq \ell_T(R)\, d_g\left(QP\mu^{EK}_n, QP\mu_n\right) \\
&\qquad + \ell_T(R)\, d_g\left(QP\mu_n, GQP\mu_n)\right) \\
&\qquad\qquad + d_g\left(B^{y^\dagger_{n+1}}(GQP\mu_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\leq c d_g(\mu^{EK}_n, \mu_n) + (\ell_T(R) + \ell_B)\, \varepsilon.
\end{aligned}
$$

Proof of Theorem

Since $T^{y^\dagger_{n+1}}(G\bullet) = B^{y^\dagger_{n+1}}(G\bullet)$ we have

$$
\begin{aligned}
d_g(\mu^{EK}_{n+1}, \mu_{n+1}) &= d_g\left(T^{y^\dagger_{n+1}}(QP\mu^{EK}_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\le d_g\left(T^{y^\dagger_{n+1}}(QP\mu^{EK}_n), T^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\quad + d_g\left(T^{y^\dagger_{n+1}}(QP\mu_n), T^{y^\dagger_{n+1}}(GQP\mu_n)\right) \\
&\quad\quad + d_g\left(T^{y^\dagger_{n+1}}(GQP\mu_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\le \ell_T(R)\, d_g\left(QP\mu^{EK}_n, QP\mu_n\right) \\
&\quad + \ell_T(R)\, d_g\left(QP\mu_n, GQP\mu_n)\right) \\
&\quad\quad + d_g\left(B^{y^\dagger_{n+1}}(GQP\mu_n), B^{y^\dagger_{n+1}}(QP\mu_n)\right) \\
&\le c d_g(\mu^{EK}_n, \mu_n) + (\ell_T(R) + \ell_B)\,\varepsilon.
\end{aligned}
$$

Closing

Conclusions

- Introduced in 1960 by Rudolph Kalman.
- Basic algorithm generalized: 3DVAR, Ensemble Kalman (EK).
- EK methods:
  - developing as a general methodology for state estimation;
  - developing as a general methodology for inverse problems.
- EK methods applied in numerous fields:
  - weather forecasting;
  - oceanography;
  - hydrology, subsurface flow;
  - medical imaging, machine learning $\cdots$.
- Analysis in its infancy:
  - accuracy of 3DVAR (State Estimation) – last decade.
  - accuracy of EK (UQ) – end of last year.
- Many open mathematical questions: great field to enter!

References I

[1] D. J. Albers, P.-A. Blancquart, M. E. Levine, E. E. Seylabi, and A. Stuart.
Ensemble Kalman methods with constraints.
Inverse Problems, 35(9):095007, 2019.

[2] M. Asch, M. Bocquet, and M. Nodet.
Data assimilation, volume 11 of *Fundamentals of Algorithms*.
Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016.
Methods, algorithms, and applications.

[3] E. Calvello, S. Reich, and A. M. Stuart.
Ensemble Kalman methods: a mean-field perspective.
Acta Numerica, arXiv 2209.11371, 2025.

[4] J. Carrillo, F. Hoffmann, A. Stuart, and U. Vaes.
The ensemble Kalman filter in the near Gaussian setting.
arXiv preprint, 2212.13239, 2022.

References II

[5] J. Carrillo and U. Vaes.
Wasserstein stability estimates for covariance-preconditioned fokker–planck equations.
Nonlinearity, 34(4):2275, 2021.

[6] M. Dashti and A. M. Stuart.
The Bayesian approach to inverse problems.
In *Handbook of uncertainty quantification. Vol. 1, 2, 3*, pages 311–428. Springer, Cham, 2017.

[7] P. Del Moral.
Nonlinear filtering: interacting particle resolution.
Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 325(6):653–658, 1997.

[8] P. Del Moral, A. Doucet, and A. Jasra.
Sequential Monte Carlo samplers.
J. R. Stat. Soc. Ser. B Stat. Methodol., 68(3):411–436, 2006.

References III

[9] P. Del Moral and A. Guionnet.
On the stability of interacting processes with applications to filtering and genetic algorithms.
In *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, volume 37, pages 155–194. Elsevier, 2001.

[10] A. Doucet, N. De Freitas, N. J. Gordon, et al.
Sequential Monte Carlo methods in practice, volume 1.
Springer, 2001.

[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.

References IV

[12] G. Evensen.
Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics.
J. Geophys. Res. Oceans, 99(C5):10143–10162, 1994.

[13] A. Garbuno-Inigo, F. Hoffmann, W. Li, and A. M. Stuart.
Interacting Langevin diffusions: gradient structure and ensemble Kalman sampler.
SIAM J. Appl. Dyn. Syst., 19(1):412–441, 2020.

[14] A. Garbuno-Inigo, N. Nüsken, and S. Reich.
Affine invariant interacting Langevin dynamics for Bayesian inference.
SIAM J. Appl. Dyn. Syst., 19(3):1633–1658, 2020.

References V

[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] D. Z. Huang, J. Huang, S. Reich, and A. M. Stuart.
Efficient derivative-free bayesian inference for large-scale inverse problems.
Inverse Problems, 38(12):125006, 2022.

[17] D. Z. Huang, T. Schneider, and A. M. Stuart.
Iterated kalman methodology for inverse problems.
Journal of Computational Physics, 463:111262, 2022.

[18] M. A. Iglesias, K. J. H. Law, and A. M. Stuart.
Ensemble Kalman methods for inverse problems.
Inverse Problems, 29(4):045001, 20, 2013.

References VI

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

[20] R. E. Kalman.
A new approach to linear filtering and prediction problems.
Trans. ASME Ser. D. J. Basic Engrg., 82(1):35–45, 1960.

[21] K. Law, A. Shukla, and A. Stuart.
Analysis of the 3dvar filter for thepartially observed Lorenz '63 model.
Discrete and Continuous Dynamical Systems, 34(3):1061–1078, 2013.

[22] K. Law, A. Stuart, and K. Zygalakis.
Data Assimilation.
Texts in Applied Mathematics 62. Springer, 2015.

References VII

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

[24] L. M. Pecora and T. L. Carroll.
Synchronization in chaotic systems.
Physical review letters, 64(8):821, 1990.

[25] P. Rebeschini and R. van Handel.
Can local particle filters beat the curse of dimensionality?
Ann. Appl. Probab., 25(5):2809–2866, 2015.

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

References VIII

[27] S. Reich and C. Cotter.
Probabilistic Forecasting and Bayesian Data Assimilation.
Cambridge University Press, New York, 2015.

[28] D. Sanz-Alonso, A. Stuart, and A. Taeb.
Inverse Problems and Data Assimilation, volume 107.
Cambridge University Press, arXiv 1810.06191, 2023.

[29] D. Sanz-Alonso and A. M. Stuart.
Long-time asymptotics of the filtering distribution for partially observed chaotic dynamical systems.
SIAM/ASA Journal on Uncertainty Quantification, 3(1):1200–1220, 2015.

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

References IX

[31] C. Schillings and A. M. Stuart.
Convergence analysis of ensemble Kalman inversion: the linear, noisy case.
Appl. Anal., 97(1):107–123, 2018.

[32] A. Spantini, R. Baptista, and Y. Marzouk.
Coupling techniques for nonlinear ensemble filtering.
SIAM Review, 64(4):921–953, 2022.

[33] T. Yang, P. G. Mehta, and S. P. Meyn.
Feedback particle filter.
IEEE Trans. Automat. Control, 58(10):2465–2480, 2013.

Citations

- Data Assimilation [2, 22, 27, 28]
- Particle Filter [7, 25]
- Kalman Filter [20]
- Ensemble Kalman Filter (EnKF) [12, 3]
- Inverse Problems [6, 28]
- Transport and Sequential Monte Carlo [8]
- Transport and EnKF [26]
- Ensemble Kalman Inversion (EKI) [18]
- EKI and regularized optimization [17]
- EKI and sampling [16]
- EKI and Approximate Gradient Descent [30, 31]
- EKI and Approximate Langevin Dynamics (EKS)[13, 14, 5]