The Ensemble Kalman Filter

Andrew Stuart

Computing and Mathematical Sciences
California Institute of Technology

DoD, NSF, ONR

⋆ ⋆ ⋆

ICIAM 2023
Tokyo, Japan

August 24th 2023

Collaborators

Trajectory Accuracy of Kalman Methods

- Kody Law (Manchester)
- Matt Levine (MIT)
- Daniel Sanz-Alonso (Chicago)

Mean Field Perspective on Kalman Methods

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

Probabilistic Accuracy of Kalman Methods

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

Overview

Kalman Filtering & Generalizations

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

Mean-Field: Calvello, Reich and S [2] (2022)

Unconditioned Dynamics

The Problem

$$\textbf{State:} \quad v^\dagger_{n+1} = \Psi(v^\dagger_n) + \xi^\dagger_n, \qquad \xi^\dagger_n \sim \mathsf{N}(0,\Sigma), \text{ i.i.d. },$$
$$\textbf{Data:} \quad y^\dagger_{n+1} = h(v^\dagger_{n+1}) + \eta^\dagger_{n+1}, \qquad \eta^\dagger_{n+1} \sim \mathsf{N}(0,\Gamma), \text{ i.i.d. }.$$

$$v^\dagger_0 \sim \mathsf{N}(m_0, C_0), \quad v^\dagger_0 \perp\!\!\!\perp \{\xi^\dagger_n\}_{n\in\mathbb{N}} \perp\!\!\!\perp \{\eta^\dagger_{n+1}\}_{n\in\mathbb{N}}$$

Goals

$Y^\dagger_n := \{y^\dagger_\ell\}_{\ell=1}^n$

- Estimate state $v^\dagger_n$ from data $Y^\dagger_n$.
- Estimate probability of state conditioned on data: $\mathbb{P}(v^\dagger_n | Y^\dagger_n)$.
- Perform estimation sequentially in n.

Kalman Filter (Navigation)

Sequential Optimization Viewpoint

$\Psi(\cdot) = M\cdot,\ h(\cdot) = H\cdot$

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

- Rudolph Kalman [12] (1960).
- $\approx 43,000$ citations (Google Scholar 8/23).
- $|\cdot|_A = |A^{-\frac{1}{2}}\cdot|$ for $A > 0$ spd.
- The Algorithm:
- $Y^\dagger_n = \{y^\dagger_\ell\}_{\ell=1}^n$.
- $v^\dagger_n | Y^\dagger_n \sim \mathrm{N}(m_n, C_n)$.
- $(m_n, C_n) \mapsto (m_{n+1}, C_{n+1})$.

3DVAR Filter (Weather Forecasting)

Sequential Optimization Viewpoint

$h(\cdot) = H\cdot$

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

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

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

- Andrew Lorenc [16] (1986).
- $\approx 2,000$ citations (Google Scholar 8/23).
- Introduced in UK Met Office.
- $\widehat{C}$ fixed.
- The Algorithm:
- $\{v_n\} \mapsto \{v_{n+1}\}$.
- When is $v_n \approx v_n^{\dagger}$?

Ensemble Kalman Filter (Oceanography)

Sequential Optimization Viewpoint

$h(\cdot) = H\cdot$

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

- Geir Evensen [9] (1994).
- $\approx 6,000$ citations (Google Scholar 8/23).
- $\widehat{C}_{n+1} = \mathrm{cov}(\widehat{v}_{n+1})$.
- The Algorithm:
- $(v_n, \mu_n^{EK}) \mapsto (v_{n+1}, \mu_{n+1}^{EK})$. $\quad \mu_n^{EK} := \mathrm{Law}(v_n)$.
- (In practice: use J ensemble members.)
- When is $\mu_n^{EK} \approx \mu_n := \mathrm{Law}(v_n^\dagger | Y_n^\dagger)$?

Summary Of Optimization Perspective

Nudging

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

Two Goals

$$
\begin{aligned}
\text{Control (3DVAR, EnKF):} \quad & |v_n - v_n^\dagger| \ll 1, \\
\text{UQ (EnKF):} \quad & \mu_n^{EK} \approx \mu_n = \text{Law}(v_n^\dagger | Y_n^\dagger).
\end{aligned}
$$

Accuracy: State Estimation

Synchronization and Lorenz '63 Pecora and Carroll [17] (1990)

Synchronization and Navier-Stokes Foias and Prodi [10] (1967)

Synchronization and Navier-Stokes Hayden, Olson and Titi [11] (2011)

Dynamics Model

The Problem

$$\frac{dv}{dt} + Av + B(v, v) = f, \quad (2a)$$

$$v(0) = v_0, \quad (2b)$$

$$\Psi(v_0) := v(\tau). \quad (2c)$$

Asssumptions

- $\exists \alpha > 0$: for all v $\langle Av, v\rangle \geq \alpha |v|^2$;
- for all v $\langle B(v, v), v\rangle = 0$;
- time-independent forcing f.

Many geophysical systems (Lorenz '63 and '96, Navier-Stokes) Temam [22] (1990)

3DVAR and Small Noise

Theorem

Assume synchronization and small noise $\mathcal{O}(\epsilon)$ in truth.
Consider 3DVAR with $K = \gamma H^\star$ and $|\gamma - 1| \leq 1$. Then

$$\limsup_{n\to\infty} \mathbb{E}\Big|v_n - v_n^\dagger\Big|^2 \leq C\epsilon^2.$$

Lorenz '63: Law, Shukla and S [14] (2013)

Lorenz '96: Law, Sanz-Alonso, Shukla and S [15] (2016)

2D Navier-Stokes: Sanz-Alonso and S [20] (2015)

EnKF and Small Noise

Theorem

Assume $H = I$ and small noise $\mathcal{O}(\epsilon)$ in truth.
Consider EnKF with variance inflation. Then

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

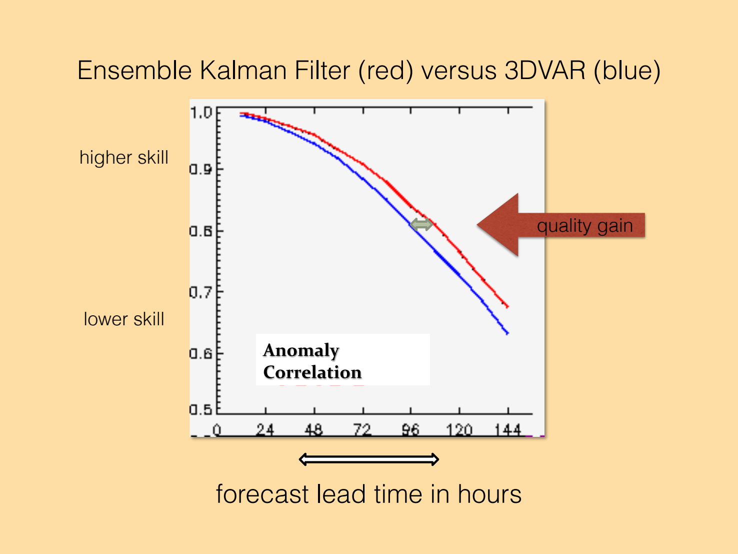
2D Navier-Stokes: Kelly, Law and S [13] (2012)

Continuous time variants: De Wiljes, Reich and Stannat [4] (2018)

Continuous time variants: Del Moral and Tugaut [7] (2018)

Impact of EnKF over 3DVAR

courtesy Roland Potthast(DWD)



Accuracy: Uncertainty Quantification

No synchronization/large noise:

Important to compare μ_n and μ_n^{EK}

Mean-Field: Calvello, Reich and S [2] (2022)

Main Theorem: Carrillo, Hoffmann, S and Vaes [3] (2022)

Unconditioned Dynamics

The Problem

$$\textbf{State:} \quad v^\dagger_{n+1} = \Psi(v^\dagger_n) + \xi^\dagger_n, \qquad \xi^\dagger_n \sim \mathsf{N}(0, \Sigma), \text{ i.i.d. },$$

$$\textbf{Data:} \quad y^\dagger_{n+1} = h(v^\dagger_{n+1}) + \eta^\dagger_{n+1}, \qquad \eta^\dagger_{n+1} \sim \mathsf{N}(0, \Gamma), \text{ i.i.d. }.$$

$$v^\dagger_0 \sim \mathsf{N}(m_0, C_0), \quad v^\dagger_0 \perp\!\!\!\perp \{\xi^\dagger_n\}_{n\in\mathbb{N}} \perp\!\!\!\perp \{\eta^\dagger_{n+1}\}_{n\in\mathbb{N}}$$

Probability Viewpoint (Linear)

$$v^\dagger_n \sim \pi_n, \quad (v^\dagger_n, y^\dagger_n) \sim \mathfrak{r}_n,$$

$$\pi_{n+1} = P\pi_n,$$

$$\mathfrak{r}_{n+1} = Q\pi_{n+1}$$

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:

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

Conditioned Dynamics (μ_n)

Probability Propagation (Nonlinear)

$$\begin{aligned} Y_n^\dagger &= \{y_\ell^\dagger\}_{\ell=1}^n, \\ v_n^\dagger | Y_n^\dagger &\sim \mu_n, \\ \mu_{n+1} &= B\big(QP\mu_n; y_{n+1}^\dagger\big). \end{aligned}$$

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 Mean Field Ensemble Kalman Filter

Comparing The True and Ensemble Kalman Filters

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

Observations About T

- Choose T to recover mean-field EnKF;
- T defined through pushforward;
- Leads to easily implementable particle algorithms;
- But key is to understand when $T \approx B$.

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

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

Optimal transport connection: Reich and Cotter [19] (2015)

Pushforward beyond the Gaussian setting (continuous time): Yang, Mehta and Meyn [23] (2013)

Pushforward beyond the Gaussian setting (discrete time): Spantini, Baptista and Marzouk [21] (2022)

Exact Filter and EnKF are Close

Weighted TV Metric

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

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

Close to Gaussian Assumption on μ_n

True filter $\{\mu_n\}$ satisfies

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

Exact Filter and EnKF are Close

Weighted TV Metric

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

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

Close to Gaussian Assumption on μ_n

True filter $\{\mu_n\}$ satisfies

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

Main Theorem Carrillo, Hoffmann, S and Vaes [3] (2022)

Let $\mu_0^{EK} = \mu_0$. Under Close to Gaussian Assumption on μ_n there is $C > 0$:

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

Closing

Conclusions: Ensemble Kalman Filtering

- Introduced in 1960 by Rudolph Kalman (linear Gaussian).
- 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] E. Calvello, S. Reich, and A. M. Stuart.
Ensemble Kalman methods: a mean-field perspective.
arXiv preprint, 2209.11371, 2022.

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

[4] J. De Wiljes, S. Reich, and W. Stannat.
Long-time stability and accuracy of the ensemble kalman–bucy filter for fully observed processes and small measurement noise.
SIAM Journal on Applied Dynamical Systems, 17(2):1152–1181, 2018.

References II

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

[6] 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.

[7] P. Del Moral and J. Tugaut.
On the stability and the uniform propagation of chaos properties of ensemble Kalman-Bucy filters.
Ann. Appl. Probab., 28(2):790–850, 2018.

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

References III

[9] 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.

[10] C. Foias and G. Prodi.

Sur le comportement global des solutions non-stationnaires des équations de navier-stokes en dimension 2.

Rendiconti del Seminario matematico della Universita di Padova, 39:1–34, 1967.

[11] 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.

[12] R. Kalman.

A new approach to linear filtering and prediction problems.

Journal of Basic Engineering, 82:35–45, 1960.

References IV

[13] 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.

[14] 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.

[15] K. J. Law, D. Sanz-Alonso, A. Shukla, and A. M. Stuart.
Filter accuracy for the Lorenz '96 model: Fixed versus adaptive observation operators.
Physica D: Nonlinear Phenomena, 325:1–13, 2016.

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

References V

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

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

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

[20] 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.

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

References VI

[22] R. Temam.
Infinite-dimensional dynamical systems in mechanics and physics, volume 68.
Springer Science & Business Media, 2012.

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

True Filter and Small Noise

Corollary (Trajectory Accuracy) Sanz-Alonso and S [20] (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

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

True Filter and UQ – Proof of Main Theorem

Lipschitz Estimates

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

True Filter and UQ – Proof of Main Theorem

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

Stability Estimate I – Nonlinear Conditioning Map $B^{y^\dagger}$

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

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

$$d_g\big(B^{y^\dagger}(GQP\mu), B^{y^\dagger}(QP\mu)\big) \le \ell_B\, d_g(GQP\mu, QP\mu)\,.$$

True Filter and UQ – Proof of Main Theorem

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

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

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

Stability Estimate II – Approximate Conditioning Map $T^{y^\dagger}$

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

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

True Filter and UQ – Convergence

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

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

True Filter and UQ – Convergence

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

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

The True and Particle Filters

Sequential Interleaving of Prediction and Bayes Theorem

$P\mu_n$ is prior prediction; $L := B \circ Q$ 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 [8] (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 [5] (1997), Del Moral and Guionnet [6] (2001)

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

Comments on Proof Rebschini and Van Handel [18] (2015),

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