Ensemble Kalman filtering in the near-Gaussian setting

SIAM MDS22: Frontiers in Monte Carlo Methods for Physics

Urbain Vaes
urbain.vaes@inria.fr

MATHERIALS – Inria Paris & CERMICS – École des Ponts ParisTech

September 30, 2022

Collaborators

José Carrillo

UNIVERSITY OF OXFORD

Mathematical Institute

Franca Hoffmann

UNIVERSITÄT BONN

Hausdorff Center for Mathematics

Andrew Stuart

CALIFORNIA INSTITUTE OF TECHNOLOGY 1891

Department of Computing + Mathematical Sciences

Outline

The discrete-time filtering problem

We consider the following stochastic dynamics and data model:

Dynamics and observations

$$\begin{aligned} \textbf{Stochastic dynamics:} \quad & u_{n+1} = \Psi(u_n) + \xi_n, \quad && \xi_n \sim \mathsf{N}(0, \Sigma), \\ \textbf{Data model:} \quad & y_{n+1} = h(u_{n+1}) + \eta_{n+1}, \quad && \eta_{n+1} \sim \mathsf{N}(0, \Gamma). \end{aligned}$$

Independence assumption:

$$u_0 \perp\!\!\!\perp \xi_n \perp\!\!\!\perp \eta_n$$

Initial state: $u_0 \sim \mathsf{N}(m_0, C_0)$.

Notations:

- $\{u_n\}_{n \in [\![0,N]\!]}$ is the unknown state in $\mathbf{R}^d$.
- $\{y_n\}_{n \in [\![1,N]\!]}$ are the observations in $\mathbf{R}^K$.
- $\Psi \colon \mathbf{R}^d \to \mathbf{R}^d$ and $h \colon \mathbf{R}^d \to \mathbf{R}^K$ are nonlinear operators.
- $Y_n = \{y_1^\dagger, \dots, y_n^\dagger\}$ is a given realization of the data up to time n.

Goal: Approximate sequentially the probability measure $\mathbf{P}(u_n | Y_n)$ for $n \in [\![1, N]\!]$.

Evolution of the filtering distribution

Update formula for the true filtering distribution

Let $\rho_n = \mathbf{P}(u_{n+1}|Y_{n+1})$. Then

$$\rho_{n+1} = L_n P \rho_n$$

For simplicity, we assume that all the measures have densities.

- The map $P \colon \mathcal{P}(\mathbf{R}^d) \to \mathcal{P}(\mathbf{R}^d)$ is the prediction:

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

- The map $L_n \colon \mathcal{P}(\mathbf{R}^d) \to \mathcal{P}(\mathbf{R}^d)$ is the analysis (Bayes' theorem):

$$L_n\rho(u) = \frac{\exp\left(-\frac{1}{2}\left|y_{n+1}^\dagger - h(u)\right|_\Gamma^2\right)\rho(u)}{\int_{\mathbf{R}^d} \exp\left(-\frac{1}{2}\left|y_{n+1}^\dagger - h(U)\right|_\Gamma^2\right)\rho(U)\,\mathrm{d}U}$$

Schematically,

$$\mathbf{P}(u_n|Y_n) \xrightarrow{P} \mathbf{P}(u_{n+1}|Y_n) \xrightarrow{L_n} \mathbf{P}(u_{n+1}|Y_{n+1}).$$

Example of a particle approximation for the filtering problem

Bootstrap particle filter

One of the simplest filtering methods is based on the iteration

$$\rho_{n+1}^B = L_n S^J P \rho_n^B.$$

where S^J is the sampling operator

$$S^J \rho := \frac{1}{J} \sum_{j=1}^{J} \delta_{u^{(j)}}, \qquad u^{(j)} \sim \rho \text{ i.i.d.}$$

- The filtering distribution is approximated by a sum of Dirac masses;
- The convergence of this approach can be proved[1],[2] in the metric

$$d_1(\mu, \nu) := \sup_{\|f\|_\infty \leq 1} \mathbf{E} \sqrt{\left|\mu[f] - \nu[f]\right|^2}.$$

[1] P. Rebeschini and R. van Handel. Can local particle filters beat the curse of dimensionality? Ann. Appl. Probab., 2015.

[2] D. Sanz-Alonso, A. M. Stuart, and A. Taeb. Inverse Problems and Data Assimilation with Connections to Machine Learning. arXiv preprint, 2018.

Goal of our work

The bootstrap particle filter

- converges to the true filtering distribution in the limit $J \to \infty$,
- but tends to perform poorly for high-dimensional problems[3].

In this talk, we study the ensemble Kalman filter,

- more robust in high-dimensional setting,
- but converges to the right limit as $J \to \infty$ only in the Gaussian setting.

Objective: obtain an error estimate away from the Gaussian setting.

---

[3] P. Bickel, B. Li, and T. Bengtsson. Sharp failure rates for the bootstrap particle filter in high dimensions. In Pushing the limits of contemporary statistics: contributions in honor of Jayanta K. Ghosh, Inst. Math. Stat. (IMS) Collect. 2008.

Using stochastic filtering to solve inverse problems (1/2)

Inverse problem considered

Find an unknown parameter $u \in \mathbf{R}^d$ from data $y \in \mathbf{R}^K$ where

$$y = h(u) + \eta,$$

- h is the forward operator;
- η is observational noise.

Bayesian approach:

- Probability distribution on parameter: $u \sim \rho_0$, encoding **prior knowledge**;
- Probability distribution for the noise: $\eta \sim \nu$.

Bayesian **posterior distribution** given observation $y^\dagger$:

$$\rho^y(u) \propto \rho_0(u)\, \nu\big(y^\dagger - h(u)\big) = \text{prior} \times \text{likelihood}.$$

In the Gaussian case where $\nu = \mathsf{N}(0, \Gamma)$,

$$\rho^y(u) = \exp\left(-\left(\frac{1}{2}\left|y^\dagger - h(u)\right|_\Gamma^2\right)\right)\rho_0(u).$$

Using stochastic filtering to solve inverse problems (2/2)

Connection with stochastic filtering

The posterior distribution coincides with the filtering distribution at $n = N$ for

$$\begin{aligned} u_{n+1} &= u_n, \\ y_{n+1} &= h(u_{n+1}) + \eta_{n+1}, \qquad \eta_{n+1} \sim \mathsf{N}(0, N\Gamma). \end{aligned}$$

with observations $y_n^\dagger = y^\dagger$ and initial state $u_0 \sim \rho_0$.

Indeed, in this case P is the identity map and so

$$\begin{aligned} \rho_{n+1} = L\rho_n &= \frac{\exp\left(-\frac{1}{2}\left|y^\dagger - h(u)\right|_{N\Gamma}^2\right)\rho_n(u)}{\int_{\mathbf{R}^d} \exp\left(-\frac{1}{2}\left|y^\dagger - h(U)\right|_{N\Gamma}^2\right)\rho_n(U)\,\mathrm{d}U} \\ &= \ldots = \frac{\exp\left(-\frac{1}{2}\left|y^\dagger - h(u)\right|_{N\Gamma/n}^2\right)\rho_0(u)}{\int_{\mathbf{R}^d} \exp\left(-\frac{1}{2}\left|y^\dagger - h(U)\right|_{N\Gamma/n}^2\right)\rho_0(U)\,\mathrm{d}U}. \end{aligned}$$

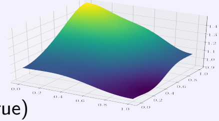
$\rightsquigarrow$ Numerical methods for filtering are useful in the context of inverse problems.

Mathematical model:

$$-\nabla \cdot \big(u(x)\nabla T(x)\big) = f(x), \qquad x \in \Omega,$$
$$T(x) = 0, \qquad x \in \partial\Omega.$$

Unknown parameter:

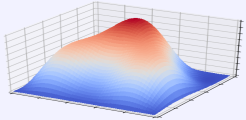
Thermal conductivity $u(x)$



(true)

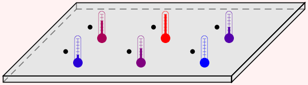
Forward problem

Solution:



Temperature field $T(x)$

Data:

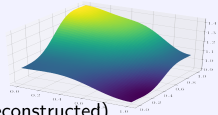


Noisy temperature measurements:

$$y = \big(T(x_1), \ldots, T(x_m)\big) + \eta.$$

Inverse problem

MAP estimator:



(reconstructed)

Outline

Analysis of a simple Gaussian projection scheme

A simple (mean field) filter based on Gaussian projections

We consider the dynamics

$$\rho_{n+1}^G = G L_n P \rho_n^G, \qquad G\rho := \mathsf{N}\big(\mathcal{M}(\rho), \mathcal{C}(\rho)\big).$$

where

$$\mathcal{M}(\rho) = \int \theta \rho(\mathrm{d}\theta), \qquad \mathcal{C}(\rho) = \int \big(\theta - \mathcal{M}(\rho)\big) \otimes \big(\theta - \mathcal{M}(\rho)\big) \rho(\mathrm{d}\theta).$$

- If Ψ and h are linear, then $\rho_n^G = \rho_n$.
- The Gaussian projection satisfies the following property:

$$G\rho = \arg\min_{\pi \in \mathcal{G}} d_{\mathrm{KL}}(\rho || \pi), \qquad \text{where } \mathcal{G} = \text{Gaussians}.$$

- **Next step**: approximation by a particle system:

$$\varrho_{n+1}^G = G L_n S^J P \varrho_n^G, \qquad S^J \rho := \frac{1}{J} \sum_{j=1}^{J} \delta_{u^{(j)}}, \qquad u^{(j)} \sim \rho \text{ i.i.d.}$$

Objective and assumptions

What we want to show

In the near-Gaussian setting, the probability measures $\{\rho_n^G\}_{n \in [\![1,N]\!]}$ are close to the true filtering distributions $\{\rho_n\}_{n \in [\![1,N]\!]}$.

Metric on probability measures: weighted total variation

$$d_g(\mu,\nu) = \sup_{|f| \leq g} \left|\mu[f] - \nu[f]\right|, \qquad g : u \to 1 + |u|^2, \qquad \rho[f] := \int f \mathrm{d}\rho.$$

Assumptions:

- (Near-Gaussian setting) There is $\varepsilon > 0$ such that

$$d_g(G\rho_n, \rho_n) \leq \varepsilon \qquad \forall n \in [\![0, N]\!].$$

- (Boundedness) There is $\kappa < \infty$ such that

$$\|\Psi\|_{L^\infty(\mathbf{R}^d)} \vee \|h\|_{L^\infty(\mathbf{R}^d)} \leq \kappa.$$

- (Non-degenerate noise) $\Sigma > 0$ and $\Gamma > 0$.

Strategy of proof

Let $\mathcal{P}_R$ denote the set of probability measures with bounded first and second moments:

$$\mathcal{P}_R(\mathbf{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbf{R}^d) : \max\left\{ |\mathcal{M}(\mu)|, \|\mathcal{C}(\mu)\|, \|\mathcal{C}(\mu)\|^{-1} \right\} \leq R \right\}.$$

Auxiliary results:

- The map P is Lipschitz over $\mathcal{P}(\mathbf{R}^d)$ for d_g.
- For any $R > 1$, the maps L_n and G are Lipschitz over $\mathcal{P}_R(\mathbf{R}^d)$ for d_g.
- Moment bounds: $P\rho_n, P\rho_n^G, L_n P\rho_n, L_n P\rho_n^G$ belong to $\mathcal{P}_{R_*}(\mathbf{R}^d)$ for some R_*.

Main theorem

Under the assumptions,

$$\forall n \in [\![0, N]\!], \qquad d_g(\rho_n^G, \rho_n) \leq \varepsilon \left(\frac{\ell^n - 1}{\ell - 1} \right).$$

Proof. Denoting by ℓ the Lipschitz constant of $G L_n P$ over $\mathcal{P}_{R_*}(\mathbf{R}^d)$, we have

$$\begin{aligned} d_g\left(\rho_{n+1}^G, \rho_{n+1}\right) &\leq d_g\left(\rho_{n+1}^G, G\rho_{n+1}\right) + d_g\left(G\rho_{n+1}, \rho_{n+1}\right) \\ &= d_g\left(G L_n P \rho_n^G, G L_n P \rho_n\right) + d_g\left(G\rho_{n+1}, \rho_{n+1}\right) \\ &\leq \ell\, d_g\left(\rho_n^G, \rho_n\right) + \varepsilon \leq \dots \end{aligned}$$

Auxiliary result 1: the prediction is globally Lipschitz[4]

Recall the definition

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

Take any $f \leq g$. Since $\Psi(v)$ is bounded by assumption,

$$\begin{aligned} b(v) := \int f(u) p(v,u)\, \mathrm{d}u &\leq \int g(u) p(v,u)\, \mathrm{d}u \\ &= \left|\Psi(v)\right|^2 + \mathrm{tr}(\Sigma) \leq \kappa^2 + \mathrm{tr}(\Sigma). \end{aligned}$$

Therefore,

$$\begin{aligned} \Bigl|P\mu[f] - P\nu[f]\Bigr| &= \left|\int \left(\int f(u) p(v,u)\, \mathrm{d}u\right) \bigl(\mu(v) - \nu(v)\bigr)\, \mathrm{d}v\right| \\ &= \Bigl|\mu[b] - \nu[b]\Bigr| \leq \Bigl(\kappa^2 + \mathrm{tr}(\Sigma)\Bigr) d_g(\mu,\nu). \end{aligned}$$

[4] P. Rebeschini and R. van Handel. Can local particle filters beat the curse of dimensionality? Ann. Appl. Probab., 2015.

Auxiliary result 2: the analysis is locally Lipschitz

Recall the definition

$$L_n\mu(u) = \frac{\exp\left(-\frac{1}{2}\left|y^\dagger_{n+1} - h(u)\right|^2_\Gamma\right)\mu(u)}{\int_{\mathbf{R}^d} \exp\left(-\frac{1}{2}\left|y^\dagger_{n+1} - h(U)\right|^2_\Gamma\right)\mu(U)\,\mathrm{d}U}$$

Since h is bounded by assumption, there exists K such that

$$\forall u \in \mathbf{R}^d, \qquad K^{-1} \leq \phi_n(u) := \exp\left(-\frac{1}{2}\left|y^\dagger_{n+1} - h(u)\right|^2_\Gamma\right) \leq K.$$

For any $f \leq g$, we have

$$\begin{aligned}
\left|L_n\mu[f] - L_n\nu[f]\right| &= \left|\frac{\mu[f\phi_n]}{\mu[\phi_n]} - \frac{\nu[f\phi_n]}{\nu[\phi_n]}\right| \\
&\leq \left|\frac{\mu[f\phi_n] - \nu[f\phi_n]}{\mu[\phi_n]}\right| + \left|\frac{\nu[f\phi_n]\big(\nu[\phi_n] - \mu[\phi_n]\big)}{\mu[\phi_n]\,\nu[\phi_n]}\right| \\
&\leq K^2 d_g(\mu,\nu) + K^4\nu[g]\,d_g(\mu,\nu) = \left(K^2 + K^4\nu[g]\right)d_g(\mu,\nu).
\end{aligned}$$

Auxiliary result 3: the Gaussian projection is locally Lipschitz

This is established in two steps:

- Control $|\mathcal{M}(\mu) - \mathcal{M}(\nu)|$ and $|\mathcal{C}(\mu) - \mathcal{C}(\nu)|$ using $d_g(\mu, \nu)$. For example:

$$
\begin{aligned}
|\mathcal{M}(\mu) - \mathcal{M}(\nu)| &= \sup_{|\mathbf{a}|=1} \left| \mathbf{a}^\mathsf{T} \big(\mathcal{M}(\mu) - \mathcal{M}(\nu)\big) \right| \\
&= \sup_{|\mathbf{a}|=1} \left| \mu[\mathbf{a}^\mathsf{T} u] - \nu[\mathbf{a}^\mathsf{T} u] \right| \leq d_g(\mu, \nu).
\end{aligned}
$$

  $\rightsquigarrow$ The weight in d_g is essential for this step!

- Show that, for any two Gaussian measures $\mu = \mathsf{N}(\mathbf{m}_1, S_1)$ and $\nu = \mathsf{N}(\mathbf{m}_2, S_2)$,

$$
d_g(\mu, \nu) \leq \sqrt{\left(\mu[g^2] + \nu[g^2]\right)} \left(3 \left\| S_2^{-1} S_1 - I_d \right\|_{\mathrm{F}} + |\mathbf{m}_1 - \mathbf{m}_2|_{S_2}\right),
$$

  This generalizes a similar result for d_1[5].

[5] L. Devroye, A. Mehrabian, and T. Reddad. The total variation distance between high-dimensional Gaussians. arXiv e-prints, 2018.

Extending the error estimate to the particle approximation

Goal: Generalize the approach to the iteration

$$\varrho_{n+1}^G = G L_n S^J P \varrho_n^G, \qquad S^J \mu := \frac{1}{J} \sum_{j=1}^{J} \delta_{u^{(j)}}, \qquad u^{(j)} \sim \mu \text{ i.i.d.}$$

Since (ϱ_n^G) are random measures, we extend the definition of d_g to the random setting:

$$d_g(\mu, \nu) := \sup_{f \le g} \mathbf{E} \sqrt{\left|\mu[f] - \nu[f]\right|^2}.$$

- The sampling operator satisfies[6]

$$d_g(\mu, S^J \mu) \le \frac{1}{\sqrt{J}} \mathbf{E}\Big(1 + |\mathcal{M}(\mu)|^2 + \operatorname{tr}\big(\mathcal{C}(\mu)\big)\Big)$$

- but the Lipschitz continuity of L_n and G is difficult to show for random measures...

$$\left|L_n\mu[f] - L_n\nu[f]\right| \le \Big(K^2 + K^4 \nu[g]\Big) \left|\mu[\phi_n f] - \nu[\phi_n f]\right|.$$

[6] D. Sanz-Alonso, A. M. Stuart, and A. Taeb. Inverse Problems and Data Assimilation with Connections to Machine Learning. arXiv preprint, 2018.

Outline

Rewriting the evolution of the true filtering distribution

Filtering problem

Stochastic dynamics: $u_{n+1} = \Psi(u_n) + \xi_n, \qquad \xi_n \sim \mathsf{N}(0, \Sigma),$

Data model: $y_{n+1} = h(u_{n+1}) + \eta_{n+1}, \qquad \eta_{n+1} \sim \mathsf{N}(0, \Gamma).$

The true filtering evolves according to

$$\rho_{n+1} = L_n P \rho_n = B^n Q P \rho_n.$$

- $Q \colon \mathcal{P}(\mathbf{R}^d) \to \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K)$

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

- $B^n \colon \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K) \to \mathcal{P}(\mathbf{R}^d)$ is the conditioning on the observation $y_{n+1} = y_{n+1}^\dagger$.

$$B^n \rho(u) = \frac{\rho(u, y_{n+1}^\dagger)}{\int \rho(u, y_{n+1}^\dagger)\, \mathrm{d}u}.$$

Schematically,

$$\mathbf{P}(u_n | Y_n) \xrightarrow{P} \mathbf{P}(u_{n+1} | Y_n) \xrightarrow{Q} \mathbf{P}\big((u_{n+1}, y_{n+1}) \,\big|\, Y_n\big) \xrightarrow{B^n} \mathbf{P}(u_{n+1} | Y_{n+1})$$

The ensemble Kalman filter from a mean field perspective[7] (1/2)

One iteration of ensemble Kalman at the mean field level

$$
\begin{aligned}
\widehat{u}_{n+1} &= \psi(u_n) + \xi_n, & \xi_n &\sim \mathsf{N}(0, \Sigma), \\
\widehat{y}_{n+1} &= h(\widehat{u}_{n+1}) + \eta_{n+1}, \\
u_{n+1} &= \widehat{u}_{n+1} + \mathcal{C}^{uy}(\widehat{\pi}_{n+1})\mathcal{C}^{yy}(\widehat{\pi}_{n+1})^{-1}\bigl(y^\dagger_{n+1} - \widehat{y}_{n+1}\bigr), & \eta_{n+1} &\sim \mathsf{N}(0, \Gamma).
\end{aligned}
$$

Here $\widehat{\pi}_{n+1} = \mathrm{Law}(\widehat{u}_{n+1}, \widehat{y}_{n+1})$ and, for $\pi \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^K)$,

$$
\mathcal{C}(\pi) = \begin{pmatrix} \mathcal{C}^{uu}(\pi) & \mathcal{C}^{uy}(\pi) \\ \mathcal{C}^{uy}(\pi)^\mathsf{T} & \mathcal{C}^{yy}(\pi) \end{pmatrix}.
$$

The third equation may be rewritten

$$
u_{n+1} = T^n(\widehat{u}_{n+1}, \widehat{y}_{n+1}; \widehat{\pi}_{n+1}),
$$

with T^n the following mean-field transport map:

$$
\begin{aligned}
T^n(\bullet, \bullet; \pi) \colon \mathbf{R}^d \times \mathbf{R}^K &\to \mathbf{R}^d; \\
(u, y) &\mapsto u + \mathcal{C}^{uy}(\pi)\mathcal{C}^{yy}(\pi)^{-1}\bigl(y^\dagger_{n+1} - y\bigr).
\end{aligned}
$$

In the following, we use the shorthand notation $T^n_\sharp \pi = \bigl(T^n(\bullet, \bullet; \pi)\bigr)_\sharp \pi$.

---

[7] E. Calvello, S. Reich, and A. M. Stuart. Ensemble Kalman Methods: A Mean Field Perspective. arXiv preprint, 2022.

The ensemble Kalman filter from a mean field perspective

One iteration of ensemble Kalman at the mean field level

$$
\begin{aligned}
\widehat{u}_{n+1} &= \Psi(u_n) + \xi_n, && \xi_n \sim \mathsf{N}(0, \Sigma), \\
\widehat{y}_{n+1} &= h(\widehat{u}_{n+1}) + \eta_{n+1}, \\
u_{n+1} &= \widehat{u}_{n+1} + \mathcal{C}^{uy}(\widehat{\pi}_{n+1})\mathcal{C}^{yy}(\widehat{\pi}_{n+1})^{-1}\big(y^\dagger_{n+1} - \widehat{y}_{n+1}\big), && \eta_{n+1} \sim \mathsf{N}(0, \Gamma).
\end{aligned}
$$

Let $\rho_n^K = \mathrm{Law}(u_n)$. Then

$$
\begin{aligned}
\rho_{n+1} &= B^n Q P \rho_n && \text{(True filtering distribution)} \\
\rho_{n+1}^K &= T_\sharp^n Q P \rho_{n+1}^K && \text{(Ensemble Kalman filtering distribution)}
\end{aligned}
$$

Schematically, for mean field ensemble Kalman:

$$
\mathrm{Law}(u_n) \xrightarrow{P} \mathrm{Law}(\widehat{u}_{n+1}) \xrightarrow{Q} \mathrm{Law}\big((\widehat{u}_{n+1}, \widehat{y}_{n+1})\big) \xrightarrow{T_\sharp^n} \mathrm{Law}(u_{n+1})
$$

Key result for the analysis: For any Gaussian ρ

$$
T_\sharp^n \rho = B^n \rho.
$$

$\rightsquigarrow$ as expected, mean field ensemble Kalman is exact in the Gaussian setting.

Estimate in the near-Gaussian setting

What we want to show

In the near-Gaussian setting, the probability measures $\{\rho_n^K\}_{n\in[\![1,N]\!]}$ are close to the true filtering distributions $\{\rho_n\}_{n\in[\![1,N]\!]}$.

Assumptions:

- (Near-Gaussian setting) There is $\varepsilon > 0$ such that

$$d_g(QP\rho_n, GQP\rho_n) \leq \varepsilon \qquad \forall n \in [\![0,N]\!].$$

- (Boundedness) There is $\kappa < \infty$ such that

$$\|\Psi\|_{L^\infty(\mathbf{R}^d)} \vee \|h\|_{L^\infty(\mathbf{R}^d)} \leq \kappa.$$

- (Lipschitz continuity) The map h is globally Lipschitz continuous.
- (Non-degenerate noise) $\Sigma > 0$ and $\Gamma > 0$.

Main theorem: There exists C_N independent of ε such that

$$\forall n \in [\![0,N]\!], \qquad d_g(\rho_n^K, \rho_n) \leq C_N\varepsilon.$$

Strategy of proof

Auxiliary results:

- The maps P, L_n and G are Lipschitz on $\mathcal{P}_R(\mathbf{R}^d)$ with constant $\ell(R)$.
- The maps B^n and $T^n_\sharp$ satisfy: $\forall(\mu,\nu) \in \mathcal{P}(\mathbf{R}^d) \times \mathcal{P}_R(\mathbf{R}^d \times \mathbf{R}^K)$,

$$d_g(B^n QP\mu, B^n\nu) \leq \ell(R)\, d_g(QP\mu, \nu)$$
$$d_g(T^n_\sharp QP\mu, T^n_\sharp\nu) \leq \ell(R)\, d_g(QP\mu, \nu),$$

  $\rightsquigarrow$ "Lipschitz" when the first argument is in the range of QP.
- Moment bounds: all the appropriate measures are in $\mathcal{P}_{R_*}$.

The main idea of the proof is to use the triangle inequality. Since

$$d_g(\rho^K_{n+1}, \rho_{n+1}) = d_g\left(T^n_\sharp QP\rho^K_n, B^n QP\rho_n\right)$$

and "$T^n_\sharp G = B_n G$" we have

$$\begin{aligned} d_g\left(T^n_\sharp QP\rho^K_n, B^n QP\rho_n\right) &\leq d_g\left(T^n_\sharp QP\rho^K_n, T^n_\sharp QP\rho_n\right) + d_g\left(T^n_\sharp QP\rho_n, T^n_\sharp GQP\rho_n\right) \\ &\quad + d_g\left(B^n GQP\rho_n, B^n QP\rho_n\right) \\ &\leq \ell(R_*)^3 d_g(\rho^K_n, \rho_n) + 2\ell(R_*)\varepsilon. \end{aligned}$$

Outline

The discrete-time filtering problem

Analysis of a simple Gaussian projection filter

Analysis of the ensemble Kalman filter in the near-Gaussian setting

Conclusions and perspectives

Perspectives and conclusions

In this presentation,

- we confined ourselves to the mean field setting;
- we analysed a simple filter based on Gaussian projections;
- we analysed the ensemble Kalman filter in the near-Gaussian setting.

Perspectives for future work:

- Obtain error estimates with a better scaling with respect to N;
- Obtain error estimates for continuous-time Gaussian filtering methods;
- Derive error bounds for the particle approximations;
- Identify settings in which the "near-Gaussian" assumption is provably correct.

Some references

- P. Bickel, B. Li, and T. Bengtsson. Sharp failure rates for the bootstrap particle filter in high dimensions. In Pushing the limits of contemporary statistics: contributions in honor of Jayanta K. Ghosh. Volume 3, Inst. Math. Stat. (IMS) Collect. 2008.
- E. Calvello, S. Reich, and A. M. Stuart. Ensemble Kalman Methods: A Mean Field Perspective. arXiv preprint, 2209.11371, September 2022.
- L. Devroye, A. Mehrabian, and T. Reddad. The total variation distance between high-dimensional Gaussians. arXiv e-prints, 1810.08693, October 2018.
- P. Rebeschini and R. van Handel. Can local particle filters beat the curse of dimensionality? Ann. Appl. Probab., 25(5):2809–2866, 2015.
- D. Sanz-Alonso, A. M. Stuart, and A. Taeb. Inverse Problems and Data Assimilation with Connections to Machine Learning. arXiv preprint, 1810.06191, October 2018.

Thank you for your attention!