# The Mean Field Ensemble Kalman Filter: Near-Gaussian Setting

SIAM UQ24: Mathematical Foundations of EnK Methods

Urbain Vaes
urbain.vaes@inria.fr

MATHERIALS – Inria Paris & CERMICS – École des Ponts ParisTech

May 13, 2024

# Collaborators

José Carrillo

Mathematical Institute

Franca Hoffmann

UNIVERSITÄT BONN

Hausdorff Center for Mathematics

Andrew Stuart

Department of Computing + Mathematical Sciences

**References**:

- The Mean Field Ensemble Kalman Filter: Near-Gaussian Setting, arXiv preprint, 2022;
- Statistical Accuracy of Approximate Filtering Methods, arXiv preprint, 2024.

# Outline



**Notation**: Probability measures and densities, Operators

# The discrete-time filtering problem

## State dynamics and observations

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

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

Independence assumption:

$$v_0 \perp\!\!\!\perp \{\xi_n\} \perp\!\!\!\perp \{\eta_n\}$$

Initial state: $v_0 \sim \mathsf{N}(m_0, C_0)$.

**Notations**:

- $\{v_n\}_{n \in [\![0,N]\!]}$ 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.
- $Y_n^\dagger = \{y_1^\dagger, \dots, y_n^\dagger\}$ is a given realization of the data up to time $n$.

**Goal**: Approximate sequentially $\mathrm{Law}(v_n | Y_n^\dagger)$.

# 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

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

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

## Definition of Q

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

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

# Unconditioned Dynamics

## State dynamics and observations

$$\textbf{Stochastic dynamics:} \qquad v_{n+1} = \Psi(v_n) + \xi_n, \qquad \xi_n \sim \mathsf{N}(0, \Sigma),$$
$$\textbf{Data model:} \qquad y_{n+1} = h(v_{n+1}) + \eta_{n+1}, \qquad \eta_{n+1} \sim \mathsf{N}(0, \Gamma).$$

Independence assumption:

$$v_0 \perp\!\!\!\perp \{\xi_n\} \perp\!\!\!\perp \{\eta_n\}$$

Initial state: $v_0 \sim \mathsf{N}(m_0, C_0)$.

## Evolution of unconditioned dynamics

Let $v_n \sim \pi_n$ and $(v_n, y_n) \sim \chi_n$. Then

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

## Key Nonlinear Operator on $\mathcal{P}$: conditioning

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

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

## Probability Viewpoint (Nonlinear)

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

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

Schematically,

$$\mathrm{Law}(v_n|Y_n^\dagger) \xrightarrow{\mathsf{P}} \mathrm{Law}(v_{n+1}|Y_n^\dagger) \xrightarrow{\mathsf{B}\circ\mathsf{Q}} \mathrm{Law}(v_{n+1}|Y_{n+1}^\dagger).$$

# The True Filter

## Sequential Interleaving of Prediction and Bayes Theorem

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

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

## Particle Filter[1]

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

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

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

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

[1] A. Doucet, N. de Freitas, and N. Gordon, editors. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.

## Convergence of the particle filter[2][3]

$$\sup_{0\leqslant n\leqslant N} d\left(\mu_n, \mu_n^{\mathrm{PF}}\right) \leqslant \frac{C}{\sqrt{J}}, \qquad d(\mu,\nu)^2 := \sup_{|f|\leqslant 1} \mathbf{E}\big(\mu[f]-\nu[f]\big)^2,$$

## Comments on proof[4][5]

- Metric $d(\bullet,\bullet)$ on random probability measures:
- Reduces to TV between deterministic measures.
- Consistency + Stability Implies Convergence.
- Consistency: $d(\mathsf{S}^J\mu,\mu) \leqslant \frac{1}{\sqrt{J}}$.
- Stability: $\mathsf{P}$, $\mathsf{L}$ Lipschitz in $d(\bullet,\bullet)$.
- Suffers from weight collapse.


[2] P. Del Moral. C. R. Acad. Sci. Paris Sér. I Math., 1997.
[3] P. Del Moral and A. Guionnet. Ann. Inst. H. Poincaré Probab. Statist., 2001.
[4] P. Rebeschini and R. van Handel. Ann. Appl. Probab., 2015.
[5] D. Sanz-Alonso, A. Stuart, and A. Taeb. Cambridge University Press, 2023.

# Convergence of the Particle Filter (2/2)

**True filtering distribution and particle filter**:

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

## Sketch of proof

**Consistency.** Monte Carlo error

$$\forall \mu \in \mathcal{P}(\mathbf{R}^d), \qquad d(\mathsf{S}^J\mu, \mu) \leqslant \frac{1}{\sqrt{J}}.$$

**Stability.** Under appropriate assumptions,

$$\forall (\mu,\nu) \in \mathcal{P}(\mathbf{R}^d) \times \mathcal{P}(\mathbf{R}^d), \qquad d(\mathsf{P}\mu, \mathsf{P}\nu) \leqslant d(\mu,\nu), \qquad d(\mathsf{L}\mu, \mathsf{L}\nu) \leqslant \ell_L d(\mu,\nu).$$

**Main argument.**

$$\begin{aligned} d(\mu_{n+1}, \mu^{\mathrm{PF}}_{n+1}) &\leqslant \ell_L d\left(\mathsf{P}\mu_n, \mathsf{S}^J\mathsf{P}\mu^{\mathrm{PF}}_n\right) \\ &\leqslant \ell_L d\left(\mathsf{P}\mu_n, \mathsf{P}\mu^{\mathrm{PF}}_n\right) + \ell_L d\left(\mathsf{P}\mu^{\mathrm{PF}}_n, \mathsf{S}^J\mathsf{P}\mu^{\mathrm{PF}}_n\right) \\ &\leqslant \ell_L d\left(\mu_n, \mu^{\mathrm{PF}}_n\right) + \frac{\ell_L}{J}. \end{aligned}$$

## Particle Filter (Weight Collapse)

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

## Ensemble Kalman Filter (No Weight Collapse!)

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

# Outline

# The ensemble Kalman filter from a mean field perspective[6] (1/2)

## Block Form Of State-Data Covariance

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

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

## Mean field ensemble Kalman filter

$$\begin{aligned}
\widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, & \xi_n &\sim \mathsf{N}(0, \Sigma), \\
\widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, & \eta_{n+1} &\sim \mathsf{N}(0, \Gamma). \\
v_{n+1} &= \widehat{v}_{n+1} + \mathcal{C}^{vy}\big(\rho^{\mathrm{EK}}_{n+1}\big)\mathcal{C}^{yy}\big(\rho^{\mathrm{EK}}_{n+1}\big)^{-1}\big(y^\dagger_{n+1} - \widehat{y}_{n+1}\big), \\
(\widehat{v}_{n+1}, \widehat{y}_{n+1}) &\sim \rho^{\mathrm{EK}}_{n+1}.
\end{aligned}$$

Approximate filtering distribution $\mu^{\mathrm{EK}}_n = \mathrm{Law}(v_n)$.

[6] E. Calvello, S. Reich, and A. M. Stuart. Acta Numerica, 2025.

# Evolution equation for the approximate filtering distribution

## Key Nonlinear Operator on $\mathcal{P}$

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

$$\begin{aligned} &\mathfrak{T}(\bullet, \bullet; \rho, y^\dagger)\colon \mathbf{R}^d \times \mathbf{R}^K \to \mathbf{R}^d; \\ &\qquad (v, y) \mapsto v + \mathcal{C}^{vy}(\rho)\mathcal{C}^{yy}(\rho)^{-1}\bigl(y^\dagger - y\bigr), \\ &\mathsf{T}(\rho; y^\dagger) = \bigl(\mathfrak{T}(\bullet, \bullet; \rho, y^\dagger)\bigr)_\sharp \rho. \end{aligned}$$

With this notation:

$$\begin{aligned} \mu_{n+1} &= \mathsf{B}(\mathsf{QP}\mu_n; y^\dagger_{n+1}) \\ \mu^{\mathrm{EK}}_{n+1} &= \mathsf{T}(\mathsf{QP}\mu^{\mathrm{EK}}_n; y^\dagger_{n+1}), \qquad \mu^{\mathrm{EK}}_0 = \mu_0. \end{aligned}$$

**Key fact**: $\mathsf{T}(\bullet; y^\dagger) \equiv \mathsf{B}(\bullet; y^\dagger)$ for Gaussian inputs.

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

# Towards an error estimate for mean field EnKF

## Best Gaussian Approximation in KL

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

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

## Weighted TV Metric

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

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

## Definition

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

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

## Theorem[7]

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

$$\sup_{0 \leqslant n \leqslant N} d_g\left(\mu_n, \mu_n^{\mathrm{EK}}\right) \leqslant C\varepsilon.$$

[7] J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. arXiv preprint, 2022.

# Closeness of exact filter and EnKF (2/2)

## Assumptions

- Data $Y_j^\dagger$ lies in set

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

- $\Psi_0 \colon \mathbf{R}^d \to \mathbf{R}^d$ and $h_0 \colon \mathbf{R}^d \to \mathbf{R}^K$ are constant functions.

Denote by $B_{\Psi,h}(r)$ the set $(\Psi, h)$ satisfying $\Psi \in B_{L^\infty}(\Psi_0, r)$, $h \in B_{L^\infty}(h_0, r)$.

## Corollary[8]

Suppose that the assumptions of the previous theorem and the assumption above are satisfied. Then for any $\epsilon > 0$ there is $\delta > 0$ such that

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

[8] J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. arXiv preprint, 2022.

### Linear Maps P, Q

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

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

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

$$\forall \mu \in \mathcal{P}(\mathbf{R}^d) \quad d_g\big(\mathsf{B}^{y^\dagger}(\mathsf{GQP}\mu), \mathsf{B}^{y^\dagger}(\mathsf{QP}\mu)\big) \leqslant \ell_B \, d_g(\mathsf{GQP}\mu, \mathsf{QP}\mu)\,.$$

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

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

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

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

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

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

**Moment bounds**

Assume that $\|\Psi\|_{L^\infty}, \|h\|_{L^\infty}$ and $\Sigma, \Gamma \succ 0$. Then there is $R$ such that

$$\mathrm{Im}(\mathsf{QP}) \in \mathcal{P}_R(\mathbf{R}^{d+K})$$

# Proof of Theorem

**Strategy:** Consistency + Stability Implies Convergence

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

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

# Outline

# The Gaussian projected filter[9]

The Gaussian projected filter is defined by the iteration

$$\mu_{n+1}^{\mathrm{GF}} = \mathsf{B}(\mathsf{GQP}\mu_n^{\mathrm{GF}}, y_{n+1}^{\dagger}).$$

**Remarks:**

- corresponds to mean field version of unscented Kalman filter.
- This evolution may be rewritten in the form

$$\mu_{n+1}^{\mathrm{GF}} = \mathsf{GT}(\mathsf{QP}\mu_n^{\mathrm{GF}}, y_{n+1}^{\dagger}). \tag{3}$$

- Reproduces the exact filtering distribution in linear Gaussian setting.

[9] E. Calvello, S. Reich, and A. M. Stuart. Acta Numerica, 2025.

# Closeness of exact filter and Gaussian projected filter

## Theorem[10]

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

$$\sup_{0 \leqslant n \leqslant N} d_g\left(\mu_n, \mu_n^{\mathrm{EK}}\right) \leqslant C\varepsilon, \qquad \varepsilon := \sup_{0 \leqslant n \leqslant N} d_g(\mathsf{GQP}\mu_n, \mathsf{QP}\mu_n).$$

## Corollary

In addition to the assumptions of the theorem, suppose that

- the data is bounded in norm by $\kappa_y$,
- $\Psi_0, h_0$ are constant functions.

Then for any $\epsilon > 0$ there is $\delta > 0$ such that

$$\sup_{Y^\dagger \in B_y} \sup_{(\Psi, h) \in B_{\Psi,h}(\delta)} \left( \sup_{0 \leqslant n \leqslant N} d_g(\mu_n, \mu_n^{\mathrm{GF}}) \right) \leqslant \epsilon.$$

[10] J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. arXiv preprint, 2022.

# Additional ingredients for the proof

### Linear Map G

The map G is locally Lipschitz on $\mathcal{P}_R(\mathbf{R}^d)$ in $d_g$.

$$\forall(\mu,\nu) \in \mathcal{P}_R(\mathbf{R}^r) \times \mathcal{P}_R(\mathbf{R}^r), \qquad d_g(\mathsf{G}\mu, \mathsf{G}\nu) \leqslant \ell_{\mathsf{G}}(R)\, d_g(\mu,\nu).$$

### Nonlinear map B

The map B is locally Lipschitz on $\mathcal{G}_R(\mathbf{R}^d)$ in $d_g$.

$$\forall(\mu,\nu) \in \mathcal{G}_R(\mathbf{R}^r) \times \mathcal{G}_R(\mathbf{R}^r), \qquad d_g\Big(\mathsf{B}(\mu; y^\dagger), \mathsf{B}(\nu, y^\dagger)\Big) \leqslant \ell_{\mathsf{B}}(R)\, d_g(\mu,\nu).$$

# Proof of the theorem

With $\mathsf{B}^{y^\dagger} = \mathsf{B}(\bullet, y^\dagger)$, we have

$$
\begin{aligned}
d_g(\mu_{n+1}^{\mathrm{GF}}, \mu_{n+1}) &= d_g\left(\mathsf{B}^{y_{n+1}^\dagger}(\mathsf{GQP}\mu_n^{\mathrm{GF}}), \mathsf{B}^{y_{n+1}^\dagger}(\mathsf{QP}\mu_n)\right) \\
&\leqslant d_g\left(\mathsf{B}^{y_{n+1}^\dagger}(\mathsf{GQP}\mu_n^{\mathrm{GF}}), \mathsf{B}^{y_{n+1}^\dagger}(\mathsf{GQP}\mu_n)\right) \\
&\qquad + d_g\left(\mathsf{B}^{y_{n+1}^\dagger}(\mathsf{GQP}\mu_n), \mathsf{B}^{y_{n+1}^\dagger}(\mathsf{QP}\mu_n)\right) \\
&\leqslant \ell_T(B)\, d_g\left(\mathsf{GQP}\mu_n^{\mathrm{GF}}, \mathsf{GQP}\mu_n\right) \\
&\qquad + \ell_B(R)\, d_g\left(\mathsf{GQP}\mu_n, \mathsf{QP}\mu_n\right) \\
&\leqslant c d_g(\mu_n^{\mathrm{GF}}, \mu_n) + \ell_B(R)\varepsilon.
\end{aligned}
$$

# Outline

# Perspectives for future work

- Control growth of error with $N$;
- Extend to continuous-time setting;
- Extend to unbounded setting (in progress);
- Extend to particle approximations[11];
- Relax assumptions of non-zero noises;
- Extend to other transport maps[12].

Thank you for your attention!

[11] F. Le Gland, V. Monbet, and V.-D. Tran. In The Oxford handbook of nonlinear filtering. Oxford Univ. Press, Oxford, 2011.

[12] E. Calvello, S. Reich, and A. M. Stuart. Acta Numerica, 2025.

# Perspectives for future work

- Control growth of error with $N$;
- Extend to continuous-time setting;
- Extend to unbounded setting (in progress);
- Extend to particle approximations[11];
- Relax assumptions of non-zero noises;
- Extend to other transport maps[12].

# Thank you for your attention!

[11] F. Le Gland, V. Monbet, and V.-D. Tran. In The Oxford handbook of nonlinear filtering. Oxford Univ. Press, Oxford, 2011.

[12] E. Calvello, S. Reich, and A. M. Stuart. Acta Numerica, 2025.

# Some references


E. Calvello, S. Reich, and A. M. Stuart. Ensemble Kalman Methods: A Mean Field Perspective. Acta Numerica, 2025.

J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. The Mean Field Ensemble Kalman Filter: Near-Gaussian Setting. arXiv preprint, 2212.13239, 2022.

P. Del Moral. Nonlinear filtering: interacting particle resolution. C. R. Acad. Sci. Paris Sér. I Math., 325(6):653–658, 1997.

P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. Ann. Inst. H. Poincaré Probab. Statist., 37(2):155–194, 2001.

A. Doucet, N. de Freitas, and N. Gordon, editors. Sequential Monte Carlo methods in practice. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.

F. Le Gland, V. Monbet, and V.-D. Tran. Large sample asymptotics for the ensemble Kalman filter. In The Oxford handbook of nonlinear filtering, pages 598–631. Oxford Univ. Press, Oxford, 2011.

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. Stuart, and A. Taeb. Inverse Problems and Data Assimilation. Volume 107. Cambridge University Press, 2023.