Lecture 2: Introduction to Data Assimilation - Methods

Claudia Schillings

CEMRACS Data Assimilation and Reduced Modeling for High Dimensional Problems

Outline

Mathematical Formulation of the Problem

We assume a model of the unknown z in the form of

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad n \in \mathbb{N} \\ z_0 &\sim \mathcal{N}(m_0, C_0) \end{aligned}$$

with $\Psi \in \mathcal{C}(\mathbb{R}^{n_z}, \mathbb{R}^{n_z})$, $\zeta = (\zeta)_n$ an iid sequence with $\zeta_0 \sim \mathcal{N}(0, \Sigma)$, $\Sigma > 0$, z_0 and ζ are assumed to be independent.

There is a true trajectory of z that produces noisy observations

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad n \in \mathbb{N}$$

with $H \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_y})$ and $\eta = (\eta)_n$ an iid sequence, independent of (z_0, ζ) with $\eta_1 \sim \mathcal{N}(0, \Gamma)$, $\Gamma > 0$.

The aim of **data assimilation** is to characterize the conditional distribution of z_n given the observations.

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \dots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n \,, & j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) & \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \dots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1} \,, \qquad j \in \mathfrak{N} \,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma) \,.$$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \ldots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, & j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) & \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \ldots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

- **Prior**

$$\mathbb{P}(z_0, \ldots, z_N) = \prod_{n=0}^{N-1} \mathbb{P}(z_{n+1}|z_n)\mathbb{P}(z_0)$$

with

$$\mathbb{P}(z_0) \propto \exp(\frac{1}{2}|z_0 - m_0|^2_{C_0})$$

and

$$\mathbb{P}(z_{n+1}|z_n) \propto \exp(\frac{1}{2}|z_{n+1} - \Psi(z_n)|^2_{\Sigma})\,.$$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \dots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) \qquad \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \dots, N\}$

$$y_{n+1} = Hz_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

- **Prior**

$$\mathbb{P}(z_0, \dots, z_N) \propto \exp(-\Theta(z_0, \dots, z_n))$$

with

$$\Theta(z_0, \dots, z_N) = \frac{1}{2}|z_0 - m_0|^2_{C_0} + \sum_{n=0}^{N-1} \frac{1}{2}|z_{n+1} - \Psi(z_n)|^2_{\Sigma}\,.$$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \ldots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) \qquad \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \ldots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

- **Likelihood**

$$\mathbb{P}(y_1, \ldots, y_N | z_0, \ldots, z_N) = \prod_{n=0}^{N-1} \mathbb{P}(y_{n+1} | z_0, \ldots, z_N)$$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \dots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) \qquad \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \dots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

- **Likelihood**

$$\begin{aligned} \mathbb{P}(y_1, \dots, y_N | z_0, \dots, z_N) &= \prod_{n=0}^{N-1} \mathbb{P}(y_{n+1} | z_0, \dots, z_N) \\ &= \prod_{n=0}^{N-1} \mathbb{P}(y_{n+1} | z_{n+1}) \end{aligned}$$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \ldots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) \qquad \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \ldots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

- **Likelihood**

$$\begin{aligned} \mathbb{P}(y_1, \ldots, y_N | z_0, \ldots, z_N) &= \prod_{n=0}^{N-1} \mathbb{P}(y_{n+1} | z_0, \ldots, z_N) \\ &\propto \exp(-\Phi(z_0, \ldots, z_n; y_1, \ldots, y_n)) \end{aligned}$$

with $\quad \Phi(z_0, \ldots, z_n; y_1, \ldots, y_n) = \sum_{n=0}^{N-1} \frac{1}{2} |y_{n+1} - H z_{n+1}|_\Gamma^2$

Smoothing Problem

Find the signal z on a discrete time interval $\mathfrak{N}_0 = \{0, \ldots, N\}$

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad & j \in \mathfrak{N}_0 \\ z_0 &\sim \mathcal{N}(m_0, C_0) \qquad & \zeta_0 \sim \mathcal{N}(0, \Sigma) \end{aligned}$$

from given data y on the discrete time interval $\mathfrak{N} = \{1, \ldots, N\}$

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad j \in \mathfrak{N}\,, \qquad \eta_1 \sim \mathcal{N}(0, \Gamma)\,.$$

Bayes' Theorem

The posterior smoothing distribution on $z_0, \ldots, z_n | y_1, \ldots, y_n$ is given by

$$\mathbb{P}(z_0, \ldots, z_n | y_1, \ldots, y_n) \propto \exp(-\Phi(z_0, \ldots, z_n; y_1, \ldots, y_n) - \Theta(z_0, \ldots, z_n))\,.$$

Filtering Problem

Find the pdf $\mathbb{P}(z_n|y_1,\dots y_n)$ associated with the probability measure on the random variable $z_n|y_1,\dots y_n$, i.e. sequentially update the pdf $\mathbb{P}(z_n|y_1,\dots y_n)$ as n is incremented.

Update $\mathbb{P}(z_{n+1}|y_1,\dots y_{n+1})$ from $\mathbb{P}(z_n|y_1,\dots y_n)$ via

prediction $\mathbb{P}(z_n|y_1,\dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1,\dots y_n)$ and

analysis $\mathbb{P}(z_{n+1}|y_1,\dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1,\dots y_{n+1})$.

Filtering Problem

Find the pdf $\mathbb{P}(z_n|y_1, \dots y_n)$ associated with the probability measure on the random variable $z_n|y_1, \dots y_n$, i.e. sequentially update the pdf $\mathbb{P}(z_n|y_1, \dots y_n)$ as n is incremented.

Update $\mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$ from $\mathbb{P}(z_n|y_1, \dots y_n)$ via

prediction $\quad \mathbb{P}(z_n|y_1, \dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1, \dots y_n)$ and

analysis $\quad \mathbb{P}(z_{n+1}|y_1, \dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$.

Filtering Problem

Update $\mathbb{P}(z_{n+1}|y_1,\dots y_{n+1})$ from $\mathbb{P}(z_n|y_1,\dots y_n)$ via

prediction $\quad \mathbb{P}(z_n|y_1,\dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1,\dots y_n)$ and

analysis $\quad \mathbb{P}(z_{n+1}|y_1,\dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1,\dots y_{n+1})$.

- **Prediction**

$$
\begin{aligned}
\mathbb{P}(z_{n+1}|y_1,\dots y_n) &= \int_{\mathbb{R}^{n_z}} \mathbb{P}(z_{n+1}|y_1,\dots y_n, z_n)\mathbb{P}(z_n|y_1,\dots y_n)\mathrm{d}z_n \\
&= \int_{\mathbb{R}^{n_z}} \mathbb{P}(z_{n+1}|z_n)\mathbb{P}(z_n|y_1,\dots y_n)\mathrm{d}z_n \,.
\end{aligned}
$$

Filtering Problem

Update $\mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$ from $\mathbb{P}(z_n|y_1, \dots y_n)$ via

prediction $\mathbb{P}(z_n|y_1, \dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1, \dots y_n)$ and

analysis $\mathbb{P}(z_{n+1}|y_1, \dots y_n) \mapsto \mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$.

- **Analysis**

$$
\begin{aligned}
\mathbb{P}(z_{n+1}|y_1, \dots y_{n+1}) &= \mathbb{P}(z_{n+1}|y_1, \dots, y_n, y_{n+1}) \\
&= \frac{\mathbb{P}(y_{n+1}|z_{n+1}, y_1, \dots, y_n)\mathbb{P}(z_{n+1}|y_1, \dots, y_n)}{\mathbb{P}(y_{n+1}|y_1, \dots, y_n)} \\
&= \frac{\mathbb{P}(y_{n+1}|z_{n+1})\mathbb{P}(z_{n+1}|y_1, \dots, y_n)}{\mathbb{P}(y_{n+1}|y_1, \dots, y_n)}.
\end{aligned}
$$

Particle Filter

Sequential Importance resampling (SIR) filter, Bootstrap filter

1: Set $n = 0$ and $\mu_0^J = \mu_0$.

2: Draw J independent realizations $z_n^{(j)}$ from μ_n^J and set $w_n^{(j)} = 1/J$ for $j = 1, \ldots, J$.

3: Define $\mu_n^J = \sum_{j=1}^J w_n^{(j)} \delta_{z_n^{(j)}}$.

4: Forecast ensemble: Draw $\hat{z}_{n+1}^{(j)} \sim p(z_n^{(j)}, \cdot)$ with kernel $p(z_n, z_{n+1}) = \mathbb{P}(z_{n+1}|z_n)$.

5: Define $g_n(z_{n+1}) \propto \mathbb{P}(y_{n+1}|z_{n+1})$ and compute

$$w_{n+1}^{(j)} = \tilde{w}_{n+1}^{(j)} / (\sum_{j=1}^J \tilde{w}_{n+1}^{(j)}), \qquad \tilde{w}_{n+1}^{(j)} = g_n(\hat{z}_{n+1}^{(j)}) w_n^{(j)}, \quad j = 1, \ldots, J.$$

6: Analysis ensemble: Set $\mu_{n+1}^J = \sum_{j=1}^J w_{n+1}^{(j)} \delta_{\hat{z}_{n+1}^{(j)}}$.

7: $n \leftarrow n + 1$, goto 2.

Particle Filter

Convergence

Assume g is bounded from below and above, i.e. $\kappa \leq g_n(z) \leq \kappa^{-1}$ for $\kappa \in (0,1], z \in \mathbb{R}^{n_z}$.

For all $n \geq 0$, there exists a constant C, independent of J such that for any $\phi \in B(\mathbb{R}^{n_z})$

$$\mathbb{E}[(\mu_n^J(\phi) - \mu_n(\phi))^2] \leq C \frac{\|\phi\|^2}{J}\,.$$

See e.g. D. Crisan and A. Doucet **2002** ***A survey of convergence results on particle filtering methods for practitioners*** *IEEE Transactions on Signal Processing* **50** for a convergence proof.

Particle Filter

Convergence

Assume g is bounded from below and above, i.e. $\kappa \leq g_n(z) \leq \kappa^{-1}$ for $\kappa \in (0,1], z \in \mathbb{R}^{n_z}$.

For all $n \geq 0$, there exists a constant C, independent of J such that for any $\phi \in B(\mathbb{R}^{n_z})$

$$\mathbb{E}[(\mu_n^J(\phi) - \mu_n(\phi))^2] \leq C\frac{\|\phi\|^2}{J}\,.$$

- The rate of convergence is independent of the state dimension n_z, i.e. particle methods can circumvent the curse of dimensionality.
- The constant C depends on the state dimension n_z in general. For the standard setting, the number of particles must increase exponentially as problem sizes increases to avoid degenerary.
  T. Bengtsson, P. Bickel and B. Li **2008** ***Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems*** *IMS Collections* **2**
  C. Snyder, T. Bengtsson, P. Bickel and J. Anderson **2008** ***Obstacles to high-dimensional particle filtering*** *Monthly Wea. Rev.* **136**

Example

Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \dots, 10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \dots, 10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

Example

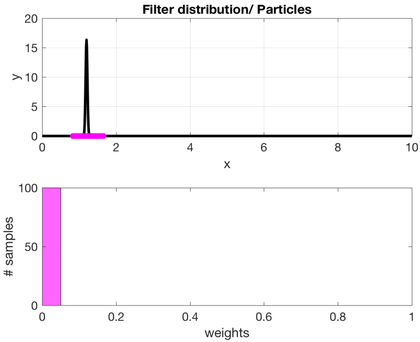
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \dots, 10\}, \qquad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \dots, 10\}, \qquad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 1}$



Example

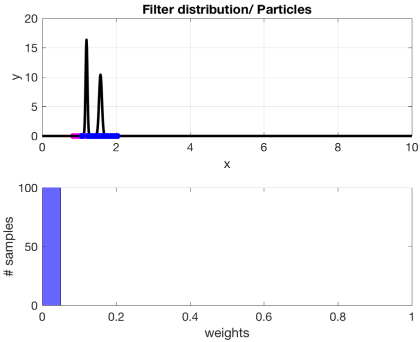
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\dots,10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\dots,10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 2}$



Example

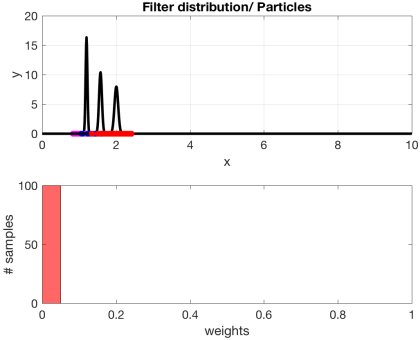
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\ldots,10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\ldots,10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 3}$



Example

Evolution model

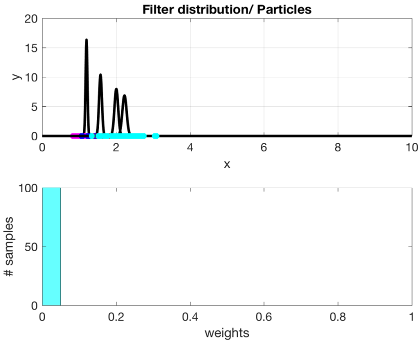
$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \dots, 10\}, \qquad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$

$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \dots, 10\}, \qquad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 4}$



Example

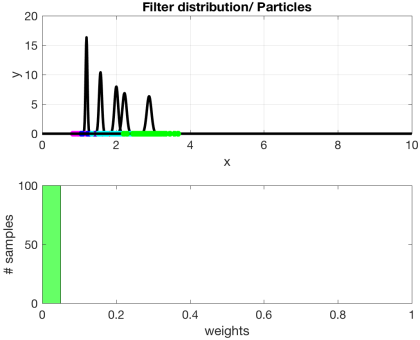
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \dots, 10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \dots, 10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 5}$



Example

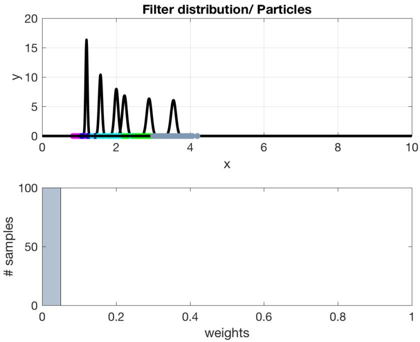
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\dots,10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d), \quad \zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\dots,10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 6}$



Example

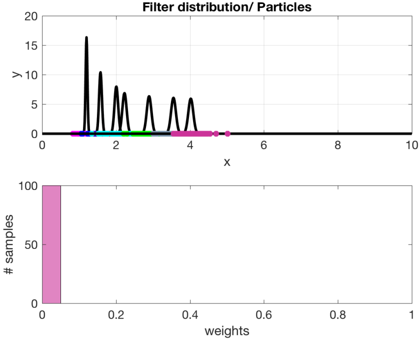
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,,\quad j \in \{0,\ldots,10\},\quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,,\quad j \in \{1,\ldots,10\},\quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 7}$



Example

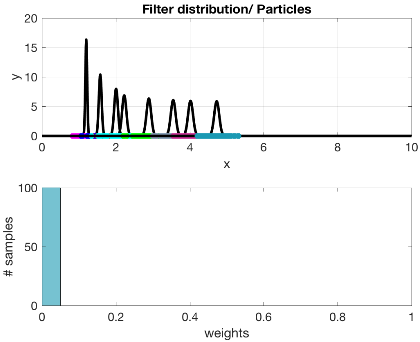
Evolution model

$$z_{n+1} = 1.2 \; I_d \; z_n + \zeta_n \,, \quad j \in \{0, \dots, 10\}, \quad z_0 \sim \mathcal{N}(1, 0.01 \; I_d), \quad \zeta_0 \sim \mathcal{N}(0, 0.01 \; I_d)\,.$$

Observation model

$$y_{n+1} = I_d \; z_{n+1} + \eta_{n+1} \,, \quad j \in \{1, \dots, 10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1 \; I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 8}$



Example

Evolution model

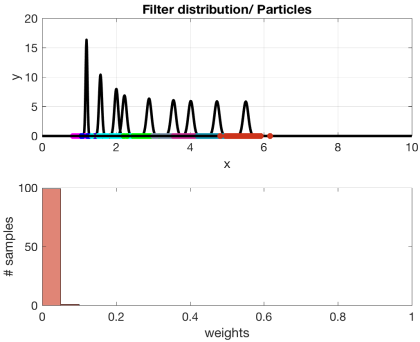
$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \dots, 10\}, \qquad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$

$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \dots, 10\}, \qquad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 9}$



Example

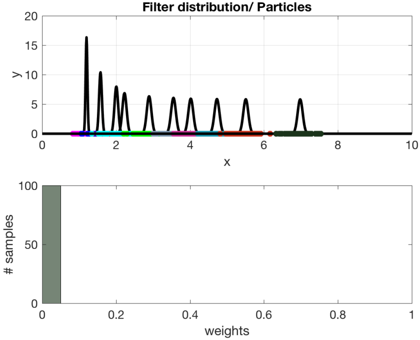
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\dots,10\}, \qquad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\dots,10\}, \qquad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 100, N = 10}$



Example

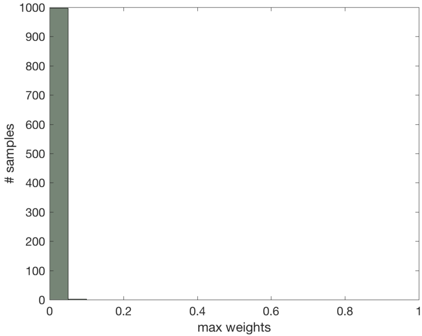
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0, \ldots, 10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1, \ldots, 10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 1, J = 1000, N = 10}$
1000 SIR runs



Example

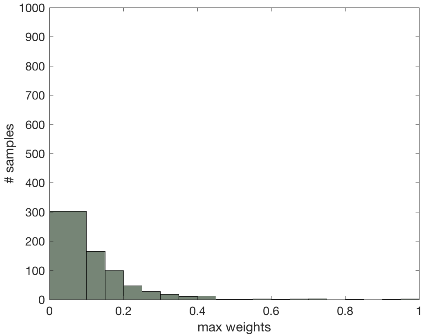
Evolution model

$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\dots,10\}, \quad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$
$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\dots,10\}, \quad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$d = 10$, $J = 1000$, $N = 1$
1000 SIR runs



Example

Evolution model

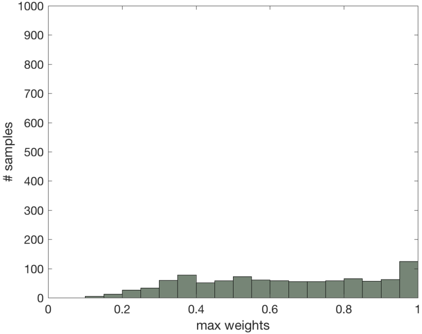
$$z_{n+1} = 1.2\ I_d\ z_n + \zeta_n\,, \quad j \in \{0,\ldots,10\}, \qquad z_0 \sim \mathcal{N}(1, 0.01\ I_d),$$

$$\zeta_0 \sim \mathcal{N}(0, 0.01\ I_d)\,.$$

Observation model

$$y_{n+1} = I_d\ z_{n+1} + \eta_{n+1}\,, \quad j \in \{1,\ldots,10\}, \qquad \eta_1 \sim \mathcal{N}(0, 0.1\ I_d)\,.$$

$\mathbf{d = 50, J = 1000, N = 1}$
1000 SIR runs



Extensions

- High-Dimensional Problems
  A. Beskos, D. Crisan, A. Jasra 2014 *On the stability of SMC methods in high dimensions* The Annals of Applied Probability **24**
  P. Rebeschini and R. van Handel 2015 *Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability* **25**
- Arnaud Doucet's SMC and Particle Filters Resources
  https://www.stats.ox.ac.uk/ doucet/smc_resources.html

Ensemble Kalman Filter

1: Set $n = 0$. Draw J independent realizations $z_n^{(j)}$ from μ_0.

2: Forecast ensemble: Set $\hat{z}_{n+1}^{(j)} = \Psi(z_n^{(j)}) + \zeta_n^{(j)}$ for $j = 1, \dots, J$. Use the ensemble $(\hat{z}_{n+1}^{(j)})_{j=1}^J$ to define the empirical mean and covariance

$$\hat{m}_{n+1} = \frac{1}{J} \sum_{j=1}^{J} \hat{z}_{n+1}^{(j)} \quad \text{and} \quad \hat{C}_{m+1} = \frac{1}{J-1} \sum_{j=1}^{J} (\hat{z}_{n+1}^{(j)} - \hat{m}_{n+1}) \otimes (\hat{z}_{n+1}^{(j)} - \hat{m}_{n+1})$$

.

3: Kalman update formulas

$$m_{n+1} = \hat{m}_{n+1} + K_{n+1}(y_{n+1} - H\hat{m}_{n+1}) \qquad C_{n+1} = \hat{C}_{n+1} - K_{n+1} H \hat{C}_{n+1}$$

with $K_{n+1} = \hat{C}_{n+1} H^\top (H \hat{C}_{n+1} H^\top + \Gamma)^{-1}$.

4: Define $(z_{n+1}^{(j)})_{j=1}^J$ by a linear transformation D with $z_{n+1}^{(j)} = \sum_{i=1}^J \hat{z}_{n+1}^{(i)} d_{ij}$ such that

$$\frac{1}{J} \sum_{j=1}^{J} z_{n+1}^{(j)} = m_{n+1} \quad \text{and} \quad \frac{1}{J-1} \sum_{j=1}^{J} (z_{n+1}^{(j)} - m_{n+1}) \otimes (z_{n+1}^{(j)} - m_{n+1}) = C_{n+1} \, .$$

5: $n \leftarrow n + 1$, go to 2.

Ensemble Kalman Filter

EnKF with perturbed observations

$$z_{n+1}^{(j)} = \sum_{i=1}^{J} \hat{z}_{n+1}^{(i)} d_{ij}$$

with observations $y_{n+1}^{(j)} = y_{n+1} + \eta_{n+1}^{(j)}$, $\eta_{n+1}^{(j)} \sim N(0,\Gamma)$ and
$d_{ij} = \delta_{ij} - \frac{1}{J-1}(\hat{z}_{n+1}^{(j)} - \hat{m}_{n+1})^\top H^\top (H\hat{C}_{n+1}H^\top + \Gamma)^{-1}(Hz_{n+1}^{(j)} - y_{n+1}^{(j)})$.

Ensemble square root filter (ESFR)

$$z_{n+1}^{(j)} = \sum_{i=1}^{J} \hat{z}_{n+1}^{(i)} d_{ij}$$

with $d_{ij} = w_i - \frac{1}{J} + s_{ij}$, where $\hat{C}_{n+1} = \frac{1}{J-1}P_{n+1}P_{n+1}^\top$,
$S = (s_{ij})_{i,j} = (I + \frac{1}{J-1}(HP_{n+1})^\top \Gamma^{-1} HP_{n+1})^{-\frac{1}{2}}$ and
$w = \frac{1}{J}1 - \frac{1}{J-1}S^2 P_{n+1}^\top H^\top \Gamma^{-1}(H\hat{m}_{n+1} - y_{n+1})$.

Ensemble Kalman Filter

EnKF with perturbed observations

$$z_{n+1}^{(j)} = \sum_{i=1}^{J} \hat{z}_{n+1}^{(i)} d_{ij}$$

with observations $y_{n+1}^{(j)} = y_{n+1} + \eta_{n+1}^{(j)}$, $\eta_{n+1}^{(j)} \sim N(0, \Gamma)$ and
$d_{ij} = \delta_{ij} - \frac{1}{J-1}(\hat{z}_{n+1}^{(j)} - \hat{m}_{n+1})^\top H^\top (H\hat{C}_{n+1}H^\top + \Gamma)^{-1}(Hz_{n+1}^{(j)} - y_{n+1}^{(j)})$.

Ensemble square root filter (ESFR)

$$z_{n+1}^{(j)} = \sum_{i=1}^{J} \hat{z}_{n+1}^{(i)} d_{ij}$$

with $d_{ij} = w_i - \frac{1}{J} + s_{ij}$, where $\hat{C}_{n+1} = \frac{1}{J-1} P_{n+1} P_{n+1}^\top$,
$S = (s_{ij})_{i,j} = (I + \frac{1}{J-1}(HP_{n+1})^\top \Gamma^{-1} HP_{n+1})^{-\frac{1}{2}}$ and
$w = \frac{1}{J}1 - \frac{1}{J-1} S^2 P_{n+1}^\top H^\top \Gamma^{-1}(H\hat{m}_{n+1} - y_{n+1})$.

- The ensemble parameter estimate lies in the linear span of the initial ensemble [23].
- In the linear case, the EnKF estimate converges in the limit $J \to \infty$ to the solution of the regularised least-squares problem [24, 31]. In the nonlinear setting, convergence to the mean-field Kalman filter is proven in [30].
- Ernst et al. [21] showed that the EnKF is not consistent with the Bayesian perspective in the nonlinear setting, but can be interpreted as a point estimator of the unknown parameters.
- Kelly et al. [28, 29, 42, 41] presented an analysis of the long-time behavior and ergodicity of the ensemble Kalman filter with arbitrary ensemble size establishing time uniform bounds to control the filter divergence and ensuring in addition the existence of an invariant measure.
- Long term stability and accuracy is established for ensemble Kalman-Bucy filters applied to continuous-time filtering problems [20, 44].
- Higher order updates by polynomial chaos expansion can be found in [34].

Connection to inverse problems

Find the unknown data $u \in X$ from noisy observations

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

Bridging Sequence

Introduction of an artificial discrete time dynamical system which maps the prior μ_0 into the posterior μ. The effective variance is amplified by $N = 1/h$ at each step, compensating for the redundant, repeated use of the data.

Analysis of Ensemble Kalman Inversion

Assumption: The forward operator is linear , i.e. $\mathcal{G} = A \in \mathcal{L}(X, \mathbb{R}^{n_y})$.

EnKF with perturbed observations

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

where $W^{(1)}, \dots, W^{(J)}$ are pairwise cylindrical Wiener processes and y denotes the noisy observational data.

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Strongly convergent discretization scheme .

Analysis of Ensemble Kalman Inversion

Assumption: The forward operator is linear , i.e. $\mathcal{G} = A \in \mathcal{L}(X, \mathbb{R}^{n_y})$.

EnKF with perturbed observations

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

where $W^{(1)}, \dots, W^{(J)}$ are pairwise cylindrical Wiener processes and y denotes the noisy observational data.

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Strongly convergent discretization scheme .

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $\quad C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\quad \overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Limiting SDE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Noise-free Case

Limiting ODE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Noise-free Case

Limiting ODE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger - u^{(j)})\,\mathrm{d}t\,,$$

or equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}u^{(j)} = -C(u)D_u\Phi(u^{(j)}; y)$$

with potential $\Phi(u; y) = \frac{1}{2}\|\Gamma^{-\frac{1}{2}}(y - Au)\|^2$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \ldots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^J$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0, \infty); \mathcal{X}_0)$ for $j = 1, \ldots, J$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \ldots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^J$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0, \infty); \mathcal{X}_0)$ for $j = 1, \ldots, J$.

Sketch of Proof

Quantities

$$e^{(j)} = u^{(j)} - \overline{u}\,, \qquad r^{(j)} = u^{(j)} - u^\dagger\,,$$

$$E_{lj} = \langle Ae^{(l)}, Ae^{(j)} \rangle_\Gamma\,, \qquad R_{lj} = \langle Ar^{(l)}, Ar^{(j)} \rangle_\Gamma\,, \qquad F_{lj} = \langle Ar^{(l)}, Ae^{(j)} \rangle_\Gamma\,.$$

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \ldots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^J$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0,\infty); \mathcal{X}_0)$ for $j = 1, \ldots, J$.

Sketch of Proof

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{(j)} = -\frac{1}{J} \sum_{k=1}^{J} E_{jk} e^{(k)}, \quad \frac{\mathrm{d}}{\mathrm{d}t} r^{(j)} = -\frac{1}{J} \sum_{k=1}^{J} F_{jk} r^{(k)}, \qquad j = 1, \ldots, J$$

$$\frac{\mathrm{d}}{\mathrm{d}t} E = -\frac{2}{J} E^2, \qquad \frac{\mathrm{d}}{\mathrm{d}t} R = -\frac{2}{J} F F^\top, \qquad \frac{\mathrm{d}}{\mathrm{d}t} F = -\frac{2}{J} F E$$

Global existence of $\mathbf{E}$, $\mathbf{R}$ and $\mathbf{F}$ $\Rightarrow$ global existence of $\mathbf{r}$ and $\mathbf{e}$

Long-time Behaviour (Linear Case)

(b) Ensemble Collapse

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \dots, J$.

Then, the matrix valued quantity $E(t)$ converges to 0 for $t \to \infty$ and, indeed $\|E(t)\| = \mathcal{O}(Jt^{-1})$.

Long-time Behaviour (Linear Case)

(b) Ensemble Collapse

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \ldots, J$.

Then, the matrix valued quantity $E(t)$ converges to 0 for $t \to \infty$ and, indeed $\|E(t)\| = \mathcal{O}(Jt^{-1})$.

The rate of convergence of E and F is algebraic with a constant growing with larger ensemble size J.

Long-time Behaviour (Linear Case)

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and the forward operator A is one-to-one. Let $Y^\parallel$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^J$ and let $Y^\perp$ denote the orthogonal complement of $Y^\parallel$ in $\mathcal{Y}$ with respect to the inner product $\langle \cdot, \cdot \rangle_\Gamma$ and assume that the initial ensemble members are chosen so that $Y^\parallel$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

Then $Ar^{(j)}(t)$ may be decomposed uniquely as

$$Ar_\parallel^{(j)}(t) + Ar_\perp^{(j)}(t) \quad \text{with } Ar_\parallel^{(j)} \in Y^\parallel \text{ and } Ar_\perp^{(j)} \in Y^\perp.$$

Furthermore $Ar_\parallel^{(j)}(t) \to 0$ as $t \to \infty$ and $Ar_\perp^{(j)}(t) = Ar_\perp^{(j)}(0) = Ar_\perp^{(1)}$.

Long-time Behaviour (Linear Case)

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and the forward operator A is one-to-one. Let $Y^\parallel$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^J$ and let $Y^\perp$ denote the orthogonal complement of $Y^\parallel$ in $\mathcal{Y}$ with respect to the inner product $\langle \cdot, \cdot \rangle_\Gamma$ and assume that the initial ensemble members are chosen so that $Y^\parallel$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

Then $Ar^{(j)}(t)$ may be decomposed uniquely as

$$Ar_\parallel^{(j)}(t) + Ar_\perp^{(j)}(t) \quad \text{with } Ar_\parallel^{(j)} \in Y^\parallel \text{ and } Ar_\perp^{(j)} \in Y^\perp.$$

Furthermore $Ar_\parallel^{(j)}(t) \to 0$ as $t \to \infty$ and $Ar_\perp^{(j)}(t) = Ar_\perp^{(j)}(0) = Ar_\perp^{(1)}$.

Adaptive choice of the initial ensemble to ensure convergence of the residuals.

Long-time Behaviour (Linear Case)

Idea of Proof

Subspace property

$$Ae^{(j)}(t) = \sum_{k=1}^{J} \ell_{jk}(t) Ae^{(k)}(0)$$

where the matrix $L = \{\ell_{jk}\}$ is invertible.

Decomposition of the residual

$$Ar^{(j)}(t) = \sum_{k=1}^{J} \alpha_k Ae^{(k)}(t) + Ar_{\perp}^{(1)}$$

Convergence of the residuals

Boundedness of the coefficient vector

$$|\alpha(t)|^2 \leq \frac{\lambda_0^{(J)}}{\lambda_0^{\min}} |\alpha(0)|^2$$

gives convergence of the residuals.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

- **No Gaussian prior assumption.**
- **Convergence result opens up the perspective to use the EnKF as a linear solver in case of a boundedly invertible forward operator.**
- **In the finite dimensional setting, the results can be used to characterise the parameter space informed by the data.**

Continuous Time Limit

EnKF with perturbed observations

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

where $W^{(1)}, \dots, W^{(J)}$ are pairwise independent cylindrical Wiener processes and y denotes the noisy observational data.

Continuous Time Limit

EnKF with perturbed observations

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger - u^{(j)})\,\mathrm{d}t\,,$$

or equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}u^{(j)} = -C(u)D_u\Phi(u^{(j)};y)$$

with potential $\Phi(u;y) = \frac{1}{2}\|\Gamma^{-\frac{1}{2}}(y - Au)\|^2$.

CIS and Stuart A M 2017 Analysis of the ensemble Kalman filter for inverse problems SINUM.
CIS and Stuart A M 2017 Convergence analysis of ensemble Kalman inversion: the linear, noisy case Applicable Analysis.

Long-time Behaviour (Linear Case)

Continuous time limit of the EnKF

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Well-posedness of the EnKF inversion

Let $\mathcal{S}$ be the linear span of $\{u_0^{(j)}\}_{j=1}^J$, then $u_t^{(j)} \in \mathcal{S}$ for all $(t,j) \in [0,\infty) \times \{1,\ldots,J\}$ almost surely.

Well-posedness of the EnKF inversion

Let $\mathcal{S}$ be the linear span of $\{u_0^{(j)}\}_{j=1}^J$, then $u_t^{(j)} \in \mathcal{S}$ for all $(t,j) \in [0,\infty) \times \{1,\dots,J\}$ almost surely.

Transformation to SDE in finite dimensional space

- Assume that the initial ensemble $(u_0^{(j)})_{j\in\{1,\dots,J\}}$ is linearly independent almost surely.
- Transformation of the original SDE to

$$\mathrm{d}u_t^{(j)} = C(u_t)A^*\Gamma^{-1}(y - Au_t^{(j)})\,\mathrm{d}t + C(u_t)A^*\Gamma^{-1/2}\,\mathrm{d}W_t^{(j)}.$$

with linear operator $A : \mathbb{R}^J \to \mathbb{R}^K$.

Well-posedness of the EnKF inversion

(a) Global Existence of Solutions [Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j\in\{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps $u_0^{(j)} : \Omega \to X$ which are linearly independent almost surely.

Then for all $T \geq 0$ there exists a unique strong solution $(u_t)_{t\in[0,T]}$ (up to $\mathbb{P}$-indistinguishability) of the set of coupled SDEs

$$\mathrm{d}u_t^{(j)} = C(u_t)A^*\Gamma^{-1}(y - Au_t^{(j)})\,\mathrm{d}t + C(u_t)A^*\Gamma^{-1/2}\,\mathrm{d}W_t^{(j)}\,. \tag{1}$$

Well-posedness of the EnKF inversion

(a) Global Existence of Solutions [Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j\in\{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps $u_0^{(j)} : \Omega \to X$ which are linearly independent almost surely.

Then for all $T \geq 0$ there exists a unique strong solution $(u_t)_{t\in[0,T]}$ (up to $\mathbb{P}$-indistinguishability) of the set of coupled SDEs

$$\mathrm{d}u_t^{(j)} = C(u_t)A^*\Gamma^{-1}(y - Au_t^{(j)})\,\mathrm{d}t + C(u_t)A^*\Gamma^{-1/2}\,\mathrm{d}W_t^{(j)}\,. \tag{1}$$

Sketch of Proof

- Local weak monotonicity
- Local weak coercivity
- Existence of a stochastic Lyapunov function $V \in C^2(X;\mathcal{R}_+)$ such that for some $c > 0$

$$LV(x) := V_x(x)F(x) + \frac{1}{2}\,\mathrm{trace}(G^T(x)V_{xx}(x)G(x)) \leq cV(x),$$

$$\inf_{|x|>R} V(x) \to \infty \text{ as } R \to \infty\,.$$

Quantification of the Ensemble Collapse

Quantities of Interest

$$e^{(j)} = u^{(j)} - \overline{u}, \quad \mathfrak{e}^{(j)} := \Gamma^{-\frac{1}{2}} A e^{(j)}$$

(b) Ensemble Collapse [Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j\in\{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps with $C_0 = \mathbb{E}[\frac{1}{J}\sum_{j=1}^{J}|\mathfrak{e}_0^{(j)}|^2] < \infty$.

Then, the ensemble collapse is quantified by

$$\mathbb{E}[\frac{1}{J}\sum_{j=1}^{J}|\mathfrak{e}_t^{(j)}|^2] \leq \frac{1}{\frac{J+1}{J^2}t + \frac{1}{C_0}}.$$

Quantification of the Ensemble Collapse

Quantities of Interest

$$e^{(j)} = u^{(j)} - \overline{u}, \quad \mathfrak{e}^{(j)} := \Gamma^{-\frac{1}{2}} A e^{(j)}$$

(b) Ensemble Collapse (Parameter Space)

[Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps with $C_0 = \mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |\mathfrak{e}_0^{(j)}|^2] < \infty$.

Further assume that the linear operator A is one-to-one.

Then, the ensemble collapse is quantified by

$$\mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |e_t^{(j)}|^2] \leq \frac{1}{\sigma_{\min}} \frac{1}{\frac{J+1}{J^2} t + \frac{1}{C_0}},$$

where $\sigma_{\min}$ is the smallest eigenvalue of $A^* \Gamma^{-1} A$.

Quantification of the Ensemble Collapse

Quantities of Interest

$$e^{(j)} = u^{(j)} - \overline{u}, \quad \mathfrak{e}^{(j)} := \Gamma^{-\frac{1}{2}} A e^{(j)}$$

(c) Almost Sure Ensemble Collapse [Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps and $\gamma : \mathcal{R}_+ \to \mathcal{R}_+$ a positive, monotonically increasing and differentiable function such that $\int_0^\infty \frac{\gamma'(s)^2}{\gamma(s)} \, \mathrm{d}s < \infty$.

Then the trivial solution of

$$\mathrm{d}\mathfrak{e}_t^{(j)} = -C(\mathfrak{e}_t)\mathfrak{e}_t^{(j)} \mathrm{d}t + C(\mathfrak{e}_t) \mathrm{d}(W_t^{(j)} - \overline{W}_t)$$

is almost surely asymptotically stable with rate function $\rho(t) = (\gamma(t))^{-\frac{1}{2}}$. In particular, $(\mathfrak{e}_t^{(j)})_{j=1,\dots,J}$ converges to zero almost surely as $t \to \infty$.

Quantification of the Ensemble Collapse

Quantities of Interest

$$e^{(j)} = u^{(j)} - \overline{u}, \quad \mathfrak{e}^{(j)} := \Gamma^{-\frac{1}{2}} A e^{(j)}$$

(c) Almost Sure Ensemble Collapse [Blömker, CIS, Wacker, Weissmann 18]

Let $u_0 = (u_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps and $\gamma : \mathcal{R}_+ \to \mathcal{R}_+$ a positive, monotonically increasing and differentiable function such that $\int_0^\infty \frac{\gamma'(s)^2}{\gamma(s)} \, \mathrm{d}s < \infty$.

Then the trivial solution of

$$\mathrm{d}\mathfrak{e}_t^{(j)} = -C(\mathfrak{e}_t)\mathfrak{e}_t^{(j)} \mathrm{d}t + C(\mathfrak{e}_t)\mathrm{d}(W_t^{(j)} - \overline{W}_t)$$

is almost surely asymptotically stable with rate function $\rho(t) = (\gamma(t))^{-\frac{1}{2}}$. In particular, $(\mathfrak{e}_t^{(j)})_{j=1,\dots,J}$ converges to zero almost surely as $t \to \infty$.

Under the assumption that A is one-to-one, the result holds true in the parameter space too.

Convergence of the Residuals

Quantities of Interest

$$r^{(j)} = u^{(j)} - u^{\dagger}, \quad \mathfrak{r}^{(j)} := \Gamma^{-\frac{1}{2}} A r^{(j)}$$

(c) Convergence of the Residuals [Blömker, CIS, Wacker, Weissmann 18]

Let y be the image of a truth $u^{\dagger} \in \mathcal{X}$ under A and $u_0 = (u_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps $u_0^{(j)} : \Omega \to \mathcal{X}$ such that $\mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |\mathfrak{r}_0^{(j)}|^2] < \infty$.

Then

$$\mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |\mathfrak{r}_t^{(j)}|^2]^{\frac{1}{2}}$$

is monotonically decreasing.

Convergence of the Residuals

Quantities of Interest

$$r^{(j)} = u^{(j)} - u^{\dagger}, \quad \mathfrak{r}^{(j)} := \Gamma^{-\frac{1}{2}} A r^{(j)}$$

(c) Convergence of the Residuals with Variance Inflation

[Blömker, CIS, Wacker, Weissmann 18]

Assume that y is the image of a truth $u^{\dagger} \in \mathcal{X}$ under A and let $\mathfrak{r}_0 = (\mathfrak{r}_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps such that $\mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |\mathfrak{r}_0^{(j)}|^2] < \infty$, $B \in \mathcal{L}(\mathcal{R}^K, \mathcal{R}^K)$ a positive definite operator and $(\mathfrak{r}_t^{(j)})_{t \geq 0, j=1,\dots,J}$ the solution of

$$\mathrm{d}\mathfrak{r}_t^{(j)} = -(C(\mathfrak{r}_t) + \frac{1}{t^{\alpha} + R} B)\mathfrak{r}_t^{(j)}\,\mathrm{d}t + C(\mathfrak{r}_t)\,\mathrm{d}W_t^{(j)}, \quad \alpha \in (0,1), R > 0.$$

Then it holds true that

$$\lim_{t \to \infty} \mathbb{E}[\frac{1}{J}\sum_{j=1}^{J} |\mathfrak{r}_t^{(j)}|^2] = 0.$$

Convergence of the Residuals

Quantities of Interest

$$r^{(j)} = u^{(j)} - u^\dagger, \quad \mathfrak{r}^{(j)} := \Gamma^{-\frac{1}{2}} A r^{(j)}$$

(c) Convergence of the Residuals with Variance Inflation

[Blömker, CIS, Wacker, Weissmann 18]

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and let $\mathfrak{r}_0 = (\mathfrak{r}_0^{(j)})_{j \in \{1,\dots,J\}}$ be $\mathcal{F}_0$-measurable maps, $B \in \mathcal{L}(\mathcal{R}^K, \mathcal{R}^K)$ a positive definite operator and $(\mathfrak{r}_t^{(j)})_{t \geq 0, j=1,\dots,J}$ the solution of

$$\mathrm{d}\mathfrak{r}_t^{(j)} = -(C(\mathfrak{r}_t) + \frac{1}{t^\alpha + R} B)\mathfrak{r}_t^{(j)}\,\mathrm{d}t + C(\mathfrak{r}_t)\,\mathrm{d}W_t^{(j)}, \quad \alpha \in (0,1), R > 0.$$

Then the solution is almost surely asymptotically stable with rate function $\rho(t) = t^{-\frac{\beta}{2}}$ for all $\beta \in (0, 1-\alpha)$.

In particular, $(\mathfrak{r}_t^{(j)})_{j=1,\dots,J}$ converges to zero almost surely as $t \to \infty$.

Convergence of the Residuals

Variance Inflation in the Parameter Space

Let $y \in A\mathcal{S}$ with $A\mathcal{S} = \operatorname{span}\{Au_0^{(1)}, \ldots, Au^{(J)}\}$.

$$\mathrm{d}u_t^{(j)} = \left(C(u_t) + \frac{1}{t^\alpha + R} B\right) A^* \Gamma^{-1} (y - Au_t^{(j)})\, \mathrm{d}t + C(u_t) A^* \Gamma^{-\frac{1}{2}}\, \mathrm{d}W_t^{(j)} \quad \text{(VI)}$$

$j = 1, \ldots, J$, for B positive definite, $R > 0$ and $\alpha \in (0, 1)$

Convergence of the Residuals

Variance Inflation in the Parameter Space

Let $y \in A\mathcal{S}$ with $A\mathcal{S} = \mathrm{span}\{Au_0^{(1)}, \dots, Au^{(J)}\}$.

$$\mathrm{d}u_t^{(j)} = (C(u_t) + \frac{1}{t^\alpha + R} B) A^* \Gamma^{-1}(y - Au_t^{(j)})\,\mathrm{d}t + C(u_t) A^* \Gamma^{-\frac{1}{2}}\,\mathrm{d}W_t^{(j)} \qquad \text{(VI)}$$

$j = 1, \dots, J$, for B positive definite, $R > 0$ and $\alpha \in (0, 1)$

Let $y \in A\mathcal{S}$ and assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and let $(u_t^{(j)})_{t \geq 0, j=1,\dots,J}$ be the solution of (VI). Then,

- $\lim_{t \to \infty} \mathbb{E}[\frac{1}{J} \sum_{j=1}^{J} |\mathfrak{e}_t^{(j)}|^2] = 0.$
- $\lim_{t \to \infty} \mathbb{E}[\frac{1}{J} \sum_{j=1}^{J} |\mathfrak{r}_t^{(j)}|^2] = 0.$
- $(\mathfrak{r}_t^{(j)})_{t \geq 0}$ converges almost surely to zero with rate function $\rho(t) = t^{-\frac{\beta}{2}}$ for all $\beta \in (0, 1 - \alpha)$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

- **No Gaussian prior assumption (in the case of the EnKF with perturbed observations).**
- **The discretization via the ensemble particles and properties of the forward operator allow to transfer the results to the parameter space informed by the data.**

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0,\pi)\,,\ p = 0 \quad \text{in } \partial D\,,$$

where

$A = \mathcal{O} \circ L^{-1}$ with $L = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + id$ and $D(L) = H^2(D) \cap H_0^1(D)$
$\mathcal{O} : X \mapsto \mathbb{R}^K$, equispaced observation points in D with spacing $\tau_N^{\mathcal{O}} = 2^{-N_K}$ at
$x_k = \frac{k}{2^{N_K}},\ k = 1, \ldots, 2^{N_K} - 1$, $o_k(\cdot) = \delta(\cdot - x_k)$ with $K = 2^{N_K} - 1$.

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0, \pi)\,, \; p = 0 \quad \text{in } \partial D\ .$$

The goal of computation is to recover the unknown data $u^\dagger$ from observations

$$y \;\; = \;\; \mathcal{O}L^{-1}u^\dagger = Au^\dagger\ .$$

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0,\pi)\,,\ p = 0 \quad \text{in } \partial D\ .$$

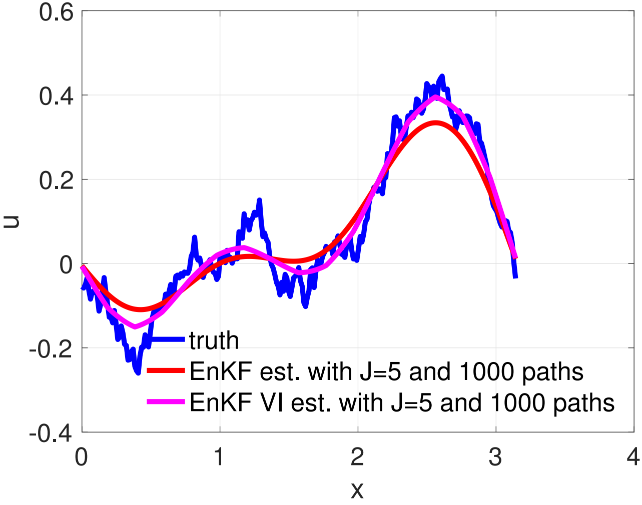
The goal of computation is to recover the unknown data $u^\dagger$ from observations

$$y \;\;=\;\; \mathcal{O}L^{-1}u^\dagger = Au^\dagger\ .$$

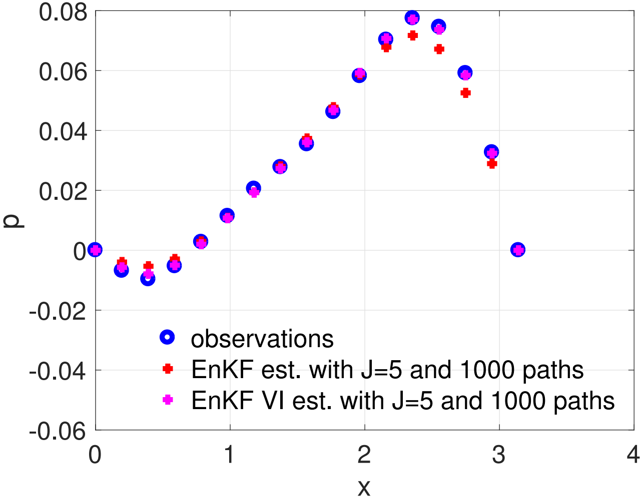
Computational Setting

- Noise-free case, $\Gamma = I$.
- $u \sim \mathcal{N}(0, C)$ with $C = \beta(A - id)^{-1}$ and with $\beta = 10$.
- Finite element method using continuous, piecewise linear ansatz functions on a uniform mesh with meshwidth $h = 2^{-8}$ (the spatial discretisation leads to a discretisation of u, i.e. $u \in \mathbb{R}^{2^8-1}$).
- The space $\mathcal{A} = \operatorname{span}\{u_0^{(j)}\}_{j=1}^J$ is chosen based on the KL expansion of $C = \beta(A - id)^{-1}$.

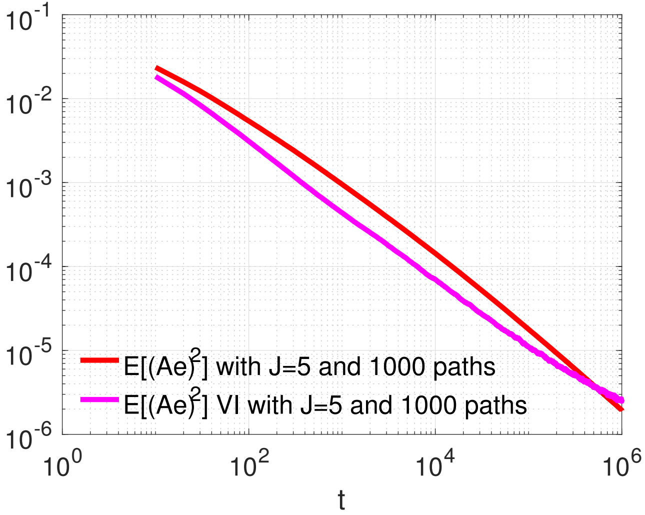
Numerical Experiments (Linear Case)



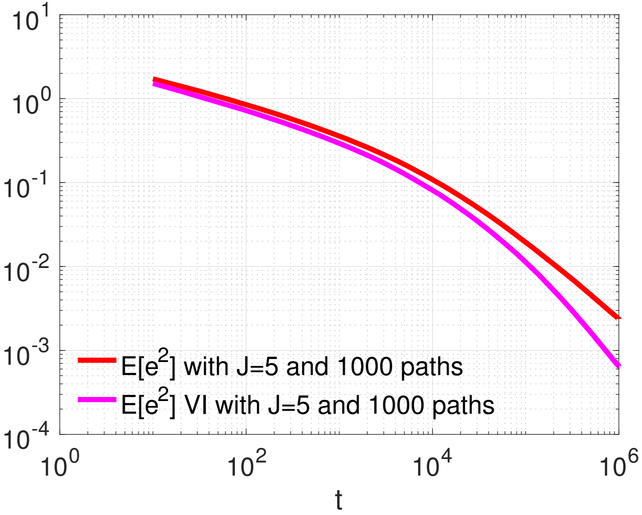
Numerical Experiments (Linear Case)



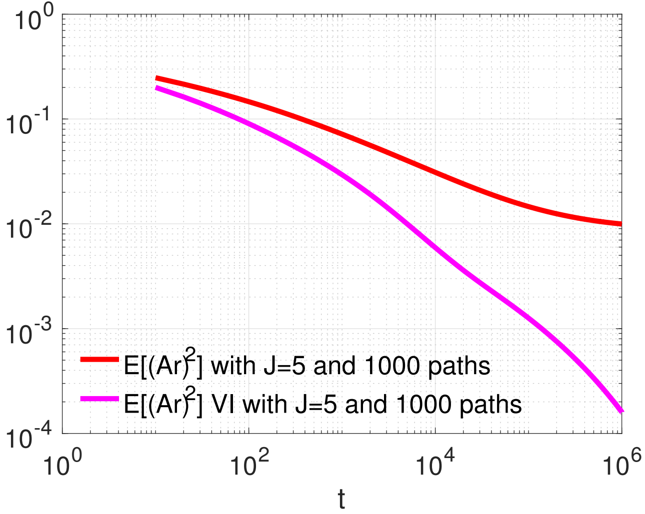
Numerical Experiments (Ensemble Collapse)



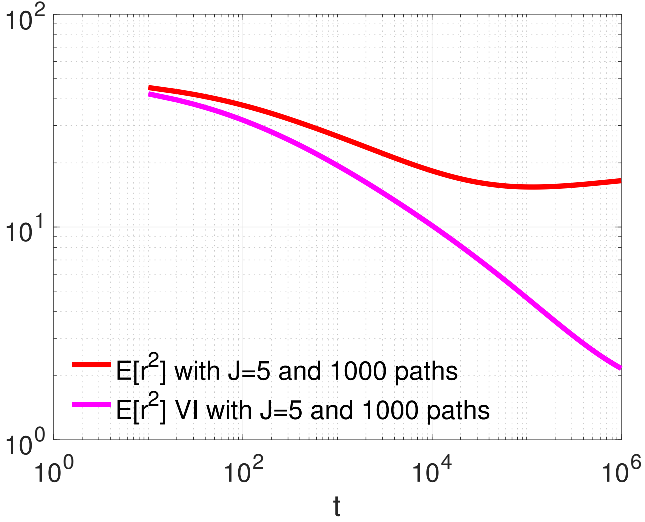
Numerical Experiments (Ensemble Collapse)



Numerical Experiments (Convergence of the Residuals)



Numerical Experiments (Convergence of the Residuals)



Extensions

- Variance inflation, Localization
  G. Evensen 2006 *Data Assimilation: The Ensemble Kalman Filter Springer*
  E. Kalnay 2003 *Atmospheric Modeling, Data Assimilation and Predictability Cambridge*
- Multilevel strategies
  A. Chernov, H. Hoel, K. Law, F. Nobile and R. Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models*
- Regularization
  N. Chada, A. Stuart, X. Tong 2019 *Tikhonov regularization within Ensemble Kalman Inversion*
- Subsampling strategies
  M. Hanu, J. Latz, ClS 2021(in preparation)
- Ensemble tranform filters
- Hybrid Methods ...

Summary

- Basic concepts of smoothing, filtering.
- (Ensemble) Kalman filter.
- Particle filter.

Main references

Kody Law, Andrew Stuart and Konstantinos Zygalakis, Data Assimilation: A Mathematical Introduction, Springer, 2015

Sebastian Reich and Colin Cotter, Probabilistic Forecasting and Bayesian Data Assimilation, Cambridge University Press, 2015

References I

M. Asch, M. Bocquet and M. Nodet 2017 *Data Assimilation. Methods, Algorithms, and Applications SIAM*.

A. Bain and D. Crisan 2009 *Fundamentals of Stochastic Filtering Springer*

A. Doucet, N. de Freitas and N. Gordon 2001 *Sequential Monte Carlo Methods in Practice Springer*

G. Evensen 2006 *Data Assimilation: The Ensemble Kalman Filter Springer*

E. Kalnay 2003 *Atmospheric Modeling, Data Assimilation and Predictability Cambridge*

P. J. van Leeuwen, Y. Cheng and S. Reich 2015 *Nonlinear Data Assimilation Springer*

K. Law, A. M. Stuart and K. Zygalakis 2015 *Data Assimilation: A Mathematical Introduction Springer*

A. Majda and J. Harlim 2011 *Mathematical Strategies for Real Time Filtering of Turbulent Signals in Complex Systems Cambridge*

P. del Moral 2004 *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications Springer*

S. Reich and C. Cotter 2015 *Probabilistic Forecasting and Bayesian Data Assimilation Cambridge*

References II

S. Agapiou, M. Burger, M. Dashti and T. Helin 2018 *Sparsity-promoting and edge-preserving maximum a posteriori estimators in non-parametric Bayesian inverse problems arXiv:1705.03286*.

T. Bengtsson, P. Bickel and B. Li 2008 *Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems IMS Collections* **2**

K. Bergemann and S. Reich 2010 *A localization technique for ensemble Kalman filters Q.J.R. Meteorol. Soc.* **136** pp. 701–707.

A. Beskos, D. Crisan, A. Jasra 2014 *On the stability of SMC methods in high dimensions* The Annals of Applied Probability **24**.

M. Bocquet and P. Sakov 2014 *An iterative ensemble Kalman smoother Q.J.R. Meteorol. Soc.* **140** 682.

D. Blömker, C. Schillings and P. Wacker 2017 *A strongly convergent numerical scheme from ensemble Kalman inversion arXiv:1703.06767*.

A. Chernov, H. Hoel, K. Law, F. Nobile and R. Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models* arXiv:1710.07282.

D. Crisan and A. Doucet 2002 *A survey of convergence results on particle filtering methods for practitioners IEEE Transactions on Signal Processing* **50**.

References III

M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss 2013 *MAP estimators and posterior consistency in Bayesian nonparametric inverse problems Inverse Problems* **29**.

P. Del Moral and J. Tugaut 2016 *On the stability and the uniform propagation of chaos properties of ensemble Kalman-Bucy filters arXiv:1605.09329*.

O. G. Ernst, B. Sprungk, and H. Starkloff 2015 *Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems*, *http://arxiv.org/abs/1504.03529*.

G. Evensen 2003 *The ensemble Kalman filter: Theoretical formulation and practical implementation Ocean dynamics* **53**.

M. A. Iglesias, K. J. H. Law and A. M. Stuart 2013 *Ensemble Kalman methods for inverse problems Inverse Problems* **29** 045001.

F. Le Gland, V. Monbet and V.-D. Tran 2011 *Large sample asymptotics for the ensemble Kalman filter Dan Crisan, Boris Rozovskii. The Oxford Handbook of Nonlinear Filtering, Oxford University Press*, 598-631.

J. Kaipio and E. Somersalo 2005 *Statistical and Computational Inverse Problems Springer*.

R. Kalman and R. Bucy 1961 *A new approach to linear filtering and prediction problems Journal of Basic Engineering* **82**.

References IV

- D. T. B. Kelly, K. J. H. Law and A. M. Stuart 2014 *Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time Nonlinearity* **27** 2579-2603.
- D. Kelly, K. Law, and A. M. Stuart 2015 *Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time Nonlinearity* **27** p. 2579.
- D. Kelly, A. J. Majda, and X. T. Tong 2015 *Concrete ensemble Kalman filters with rigorous catastrophic filter divergence Proceedings of the National Academy of Sciences* **112** pp. 10589–10594.
- K. Law, H. Tembine, and R. Tempone 2016 *Deterministic mean-field ensemble Kalman filter SIAM J. Sci. Comput* **38** pp. A1251–A1279.
- J. Mandel, L. Cobb, and J. D. Beezley 2011 *On the convergence of the ensemble Kalman filter Applications of Mathematics* **56** pp. 533–541.
- Y. M. Marzouk, T. Moselhy, M. Parno, A. Spantini 2016 *An introduction to sampling via measure transport In ?Handbook of Uncertainty Quantification.? R. Ghanem, D. Higdon, and H. Owhadi, editors* Springer.
- P. Rebeschini and R. van Handel 2015 *Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability* **25**.
- B. Rosic, A. Litvinenko, O. Pajonk, H.G. Matthies 2017 *Direct Bayesian Update of Polynomial Chaos Representations*, Preprint TU Braunschweig.

References V

C. Schillings and A. M. Stuart 2017 *Analysis of the Ensemble Kalman Filter for Inverse Problems SIAM J Numerical Analysis* 55(3), 1264-1290.

C. Schillings and A. M. Stuart 2017 *Convergence analysis of ensemble Kalman inversion: the linear, noisy case Applicable Analysis*.

C. Snyder, T. Bengtsson, P. Bickel and J. Anderson 2008 *Obstacles to high-dimensional particle filtering Monthly Wea. Rev.* **136**.

A. S. Stordal, and A. H. Elsheikh 2015 *Iterative ensemble smoothers in the annealed importance sampling framework Advances in Water Resources* **86**.

A. M. Stuart 2010 *Inverse problems: a Bayesian approach Acta Numerica* **19**.

A. Tarantola 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation SIAM*.

X. T. Tong, A. J. Majda, and D. Kelly 2015 *Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation arXiv:1507.08319*.

X. T. Tong, A. J. Majda, and D. Kelly 2016 *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29, p. 657.

C. Villani 2009 *Optimal Transportation: Old and New Springer*

References VI

J. de Wiljes, S. Reich, and W. Stannat 2016 *Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise*, arXiv:1612.06065.

G. Evensen 2003 *The ensemble Kalman filter: Theoretical formulation and practical implementation Ocean dynamics* **53** pp. 343–367.

R. Kalman and R. Bucy 2016 *A new approach to linear filtering and prediction problems Journal of Basic Engineering* **82** 95-108.