Particle and Ensemble Kalman Filters for Nonlinear Filtering Problems

Lecture 1: Introduction to Data Assimilation

Jana de Wiljes, Claudia Schillings

Minitutorial MT7 SIAM UQ 2018

UNIVERSITÄT MANNHEIM

Outline

1. Motivation
2. Mathematical Formulation of the Problem
3. Particle Filter
4. Ensemble Kalman Filter

What is Data Assimilation?

The seamless integration of large data sets into sophisticated computational models provides one of the central challenges for the mathematical sciences in the 21st century. When the computational model is based on dynamical systems, and the data set is time ordered, the process of combining models and data is called data assimilation.

Sebastian Reich and Andrew Stuart, SIAM News 2015

source: https://www.focus.de/panorama/videos/wettervorhersage-unwettergefahr-hier-drohen-starkregen-und-hagel_id_5763860.html

Outline

1. Motivation
2. Mathematical Formulation of the Problem
3. Particle Filter
4. Ensemble Kalman Filter

What is Data Assimilation?

The seamless integration of large data sets into sophisticated computational models provides one of the central challenges for the mathematical sciences in the 21st century. When the computational model is based on dynamical systems, and the data set is time ordered, the process of combining models and data is called data assimilation.

Sebastian Reich and Andrew Stuart, SIAM News 2015

source: https://www.focus.de/panorama/videos/wettervorhersage-unwettergefahr-hier-drohen-starkregen-und-hagel_id_5763860.html

Applications of Data Assimilation

Data assimilation provides important techniques for the incorporation of data in models in various fields of science and engineering:

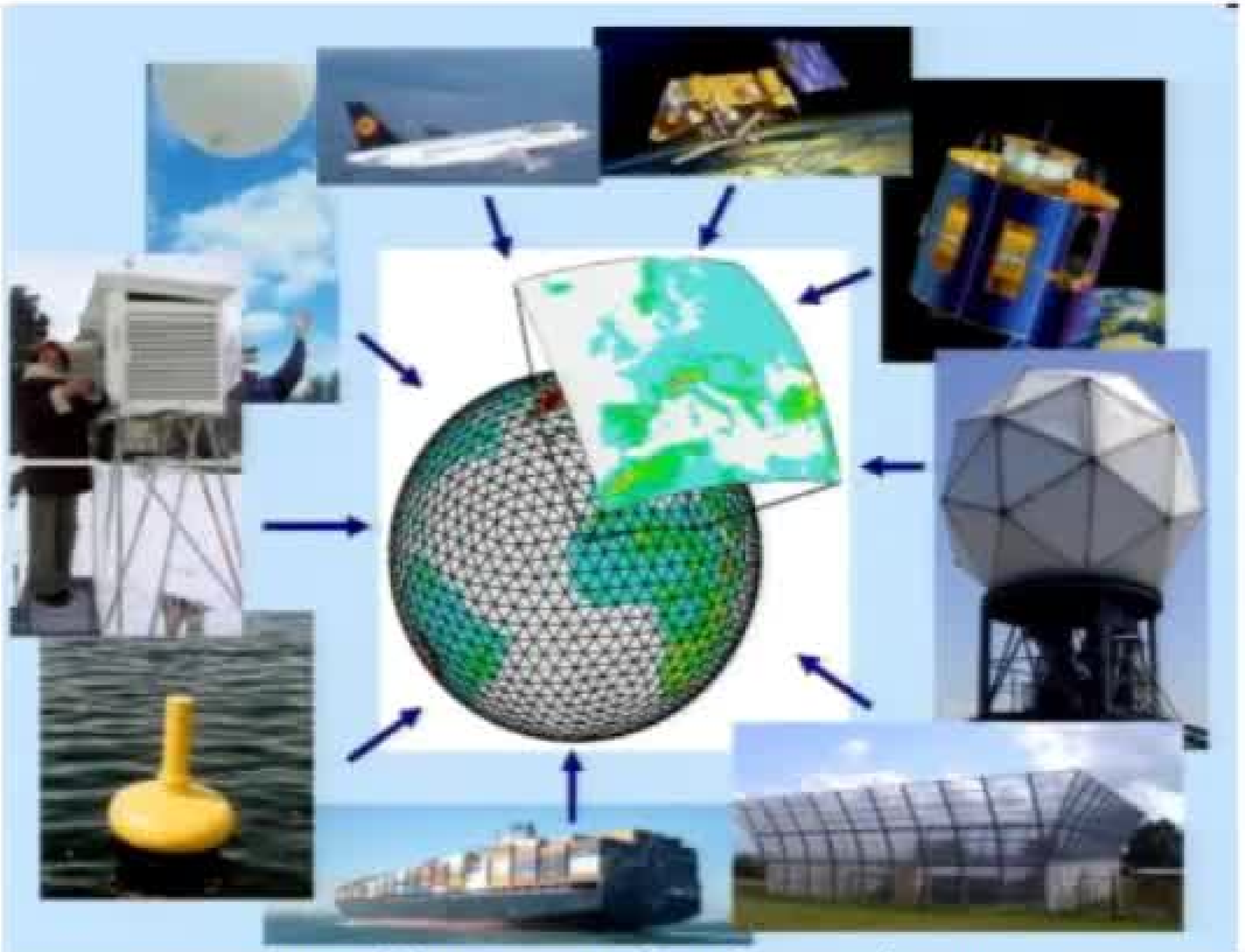
Numerical Weather Prediction

extrapolation of the present state of the atmosphere using computational models into the future

Oil reservoir modeling

Biological systems

Different observation systems



source: https://www.dwd.de/EN/research/weatherforecasting/num_modelling/02_data_assimilation/data_assimilation_node.html

Applications of Data Assimilation

Data assimilation provides important techniques for the incorporation of data in models in various fields of science and engineering:

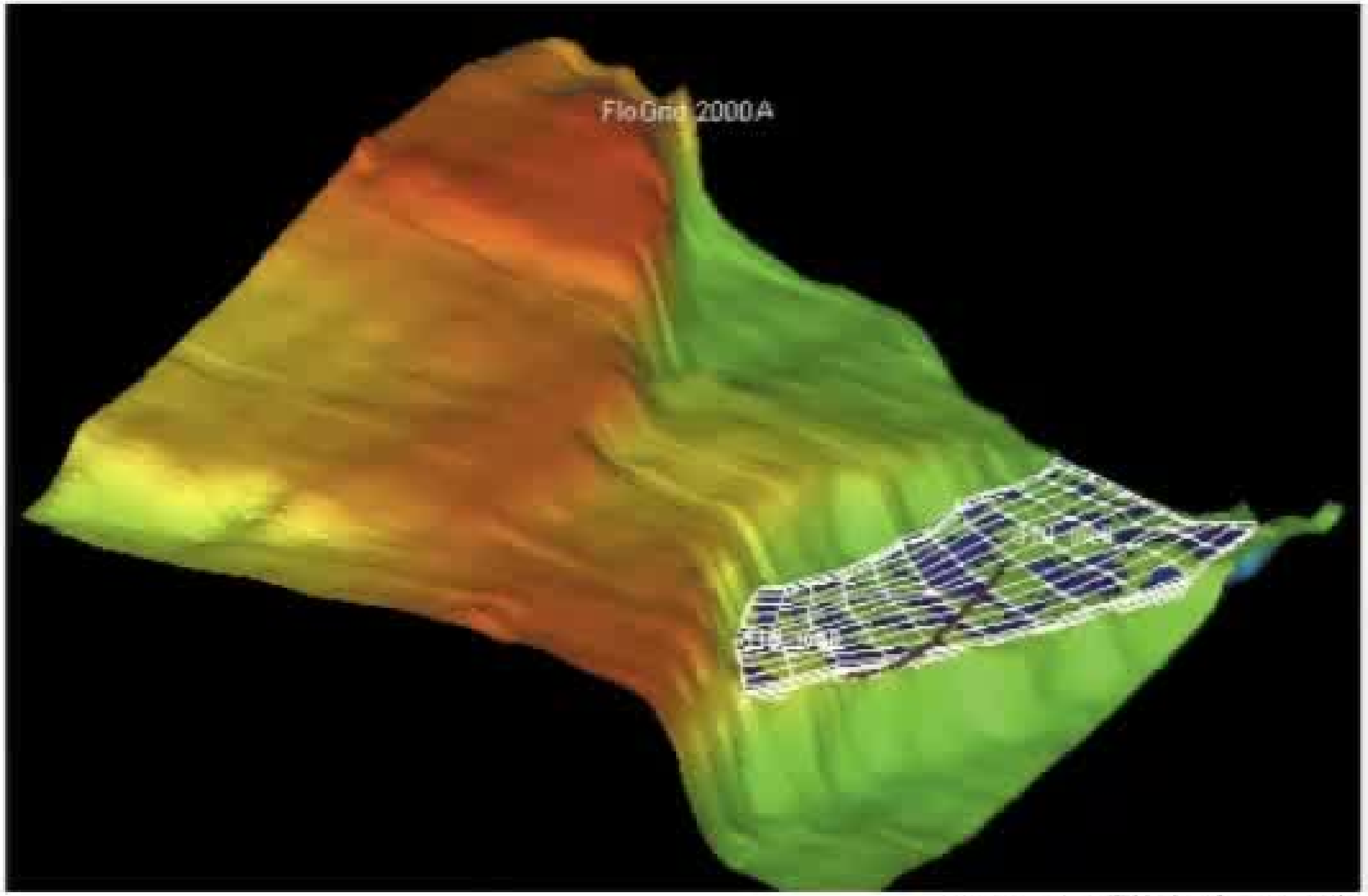
Numerical Weather Prediction

Oil reservoir modeling

improvement of the accuracy of the reservoir simlulator by data

Biological systems

Reservoir in the Gulf of Mexico



source: Christie et al.

Applications of Data Assimilation

Data assimilation provides important techniques for the incorporation of data in models in various fields of science and engineering:

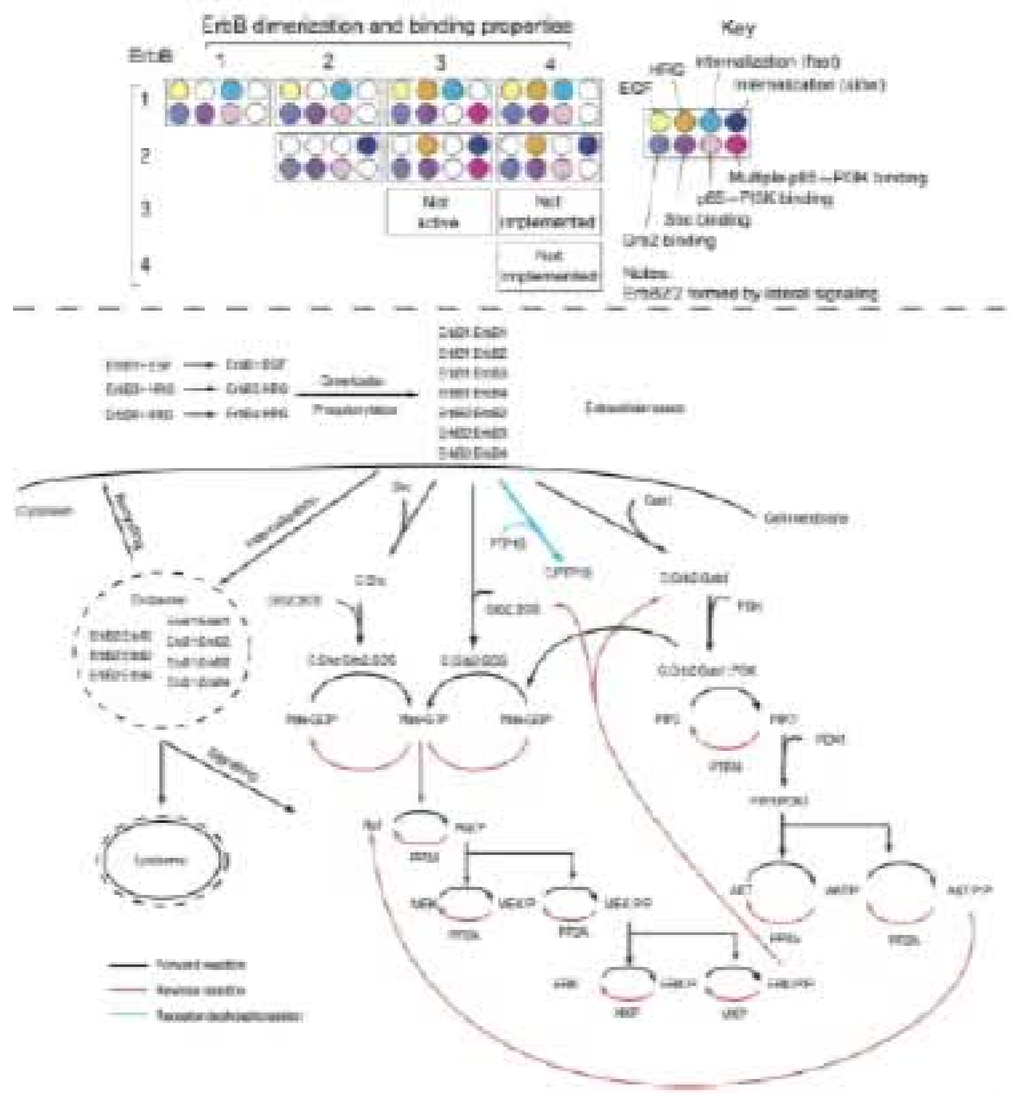
Numerical Weather Prediction

Oil reservoir modeling

Biological systems

confront biological models with measurement data

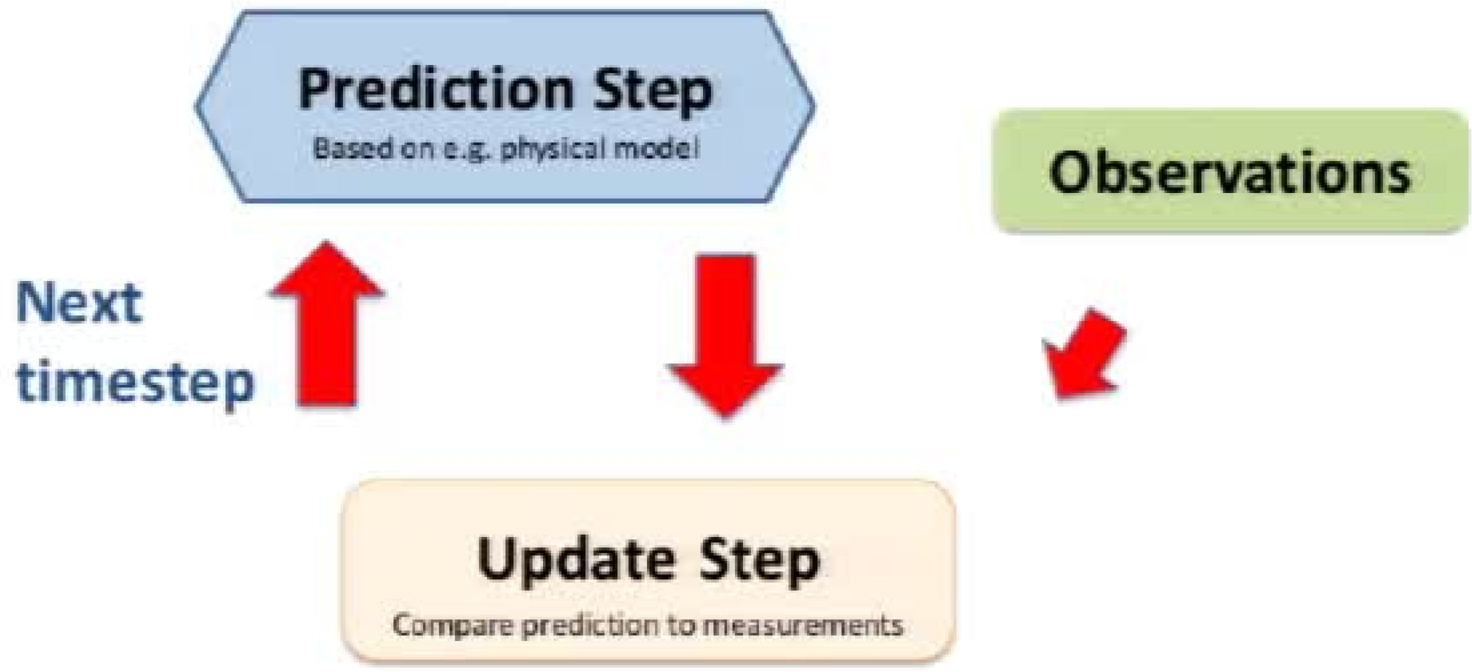
ErbB signaling pathways



source: Chen et al.

Challenges in Data Assimilation

- **High-dimensional problems.**
- **Highly nonlinear forward models.**
- **Robustness of data assimilation algorithms w.r.t. to numerical / model error.**
- **Assessment of uncertainty in the predictions.**



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 $\mathfrak{N}_0 = \{0, \ldots, N\}$

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

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

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

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

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

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

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

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 $\mathfrak{N} = \{1, \dots, N\}$

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

Bayes' Theorem

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

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

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**

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

Relation between Smoothing and Filtering

We denote by $\mathbb{P}(z_0, \ldots, z_N | y_1, \ldots, y_N)$ the smoothing distribution and by $\mathbb{P}(z_N | y_1, \ldots y_N)$ the filtering distribution at time N. Then, the marginal of the smoothing distribution on z_N coincides with the filtering distribution at time N, i.e.

$$\int \mathbb{P}(z_0, \ldots, z_N | y_1, \ldots, y_N) \mathrm{d}z_0 \ldots \mathrm{d}z_{N-1} = \mathbb{P}(z_N | y_1, \ldots y_N).$$

Note that the marginal of the smoothing distribution on z_n, $n < N$ is in general **not equal** to the filter $\mathbb{P}(z_n | y_1, \ldots y_n)$.

Linear Example

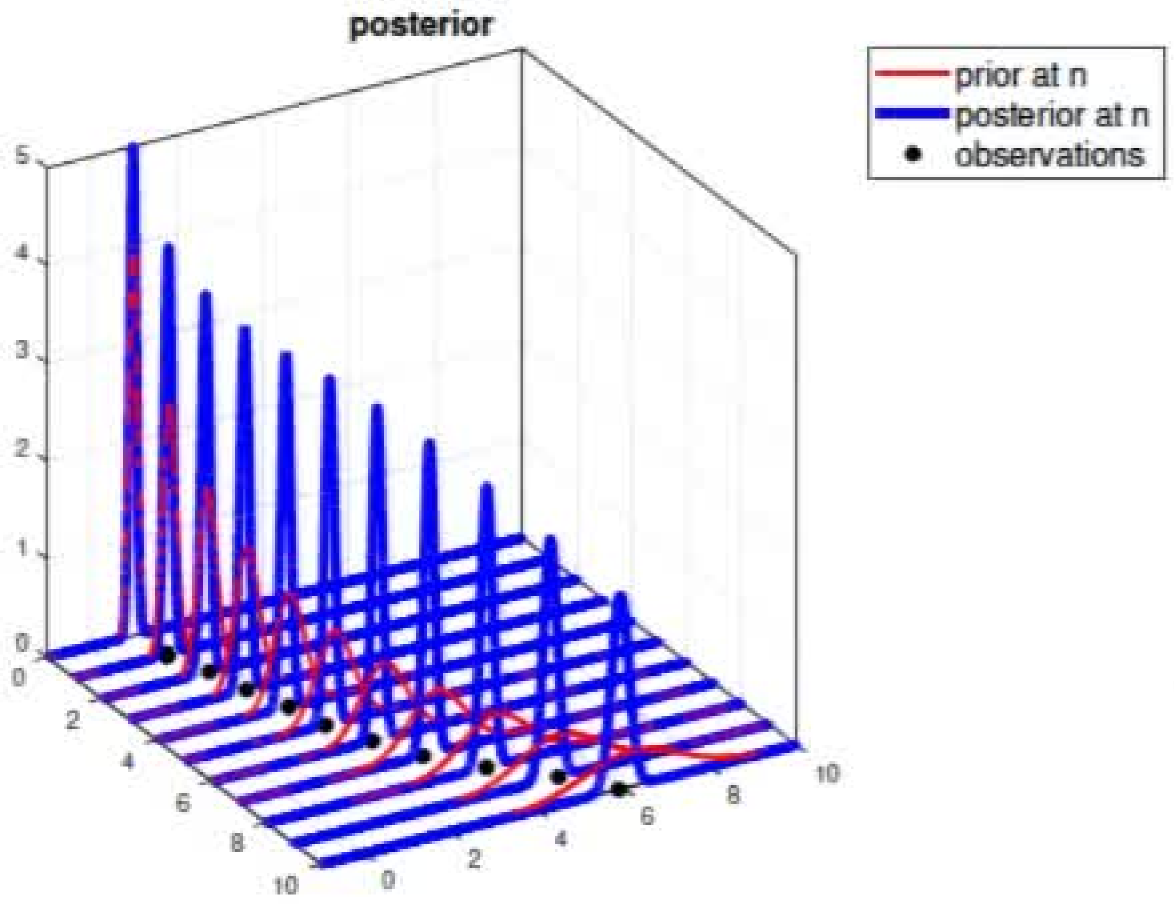
Evolution model

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

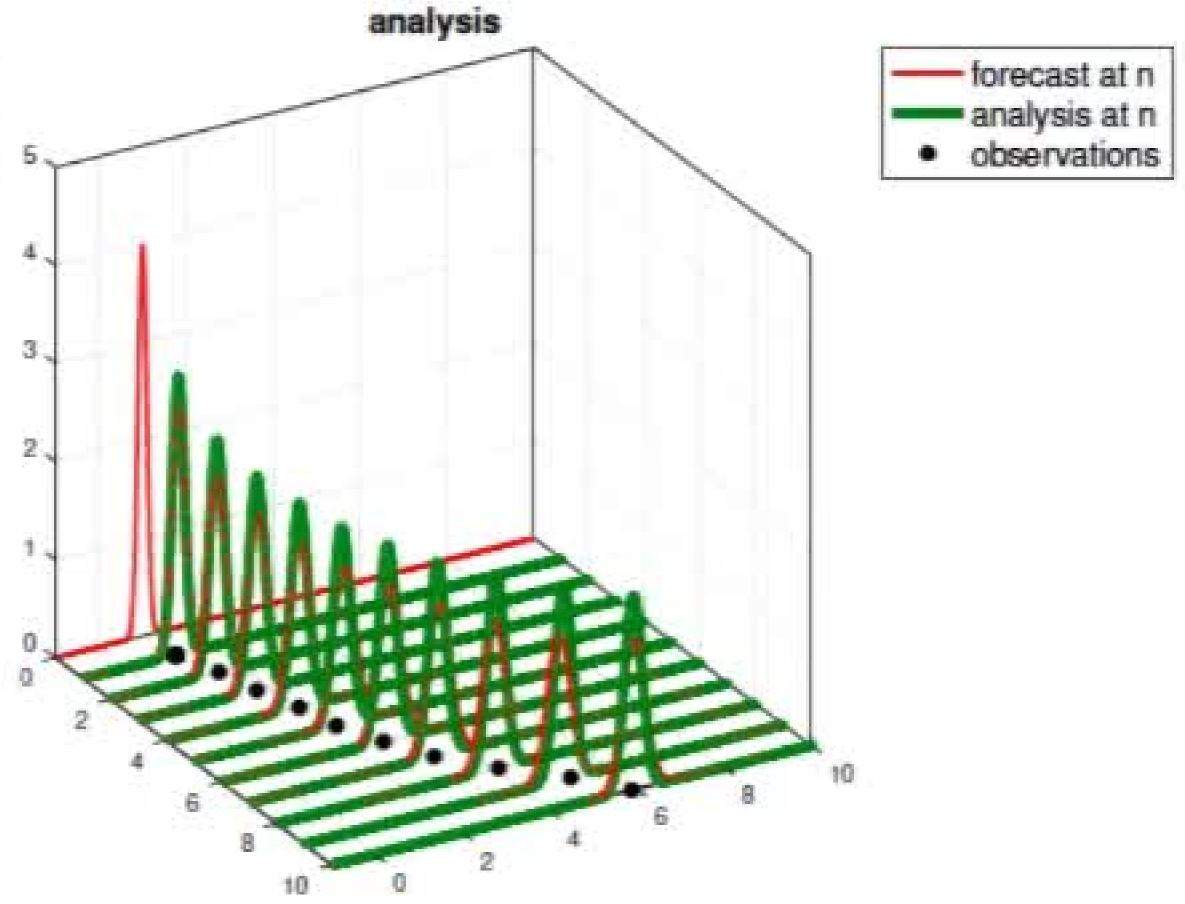
Observation model

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

Smoothing



Filtering



Linear Example

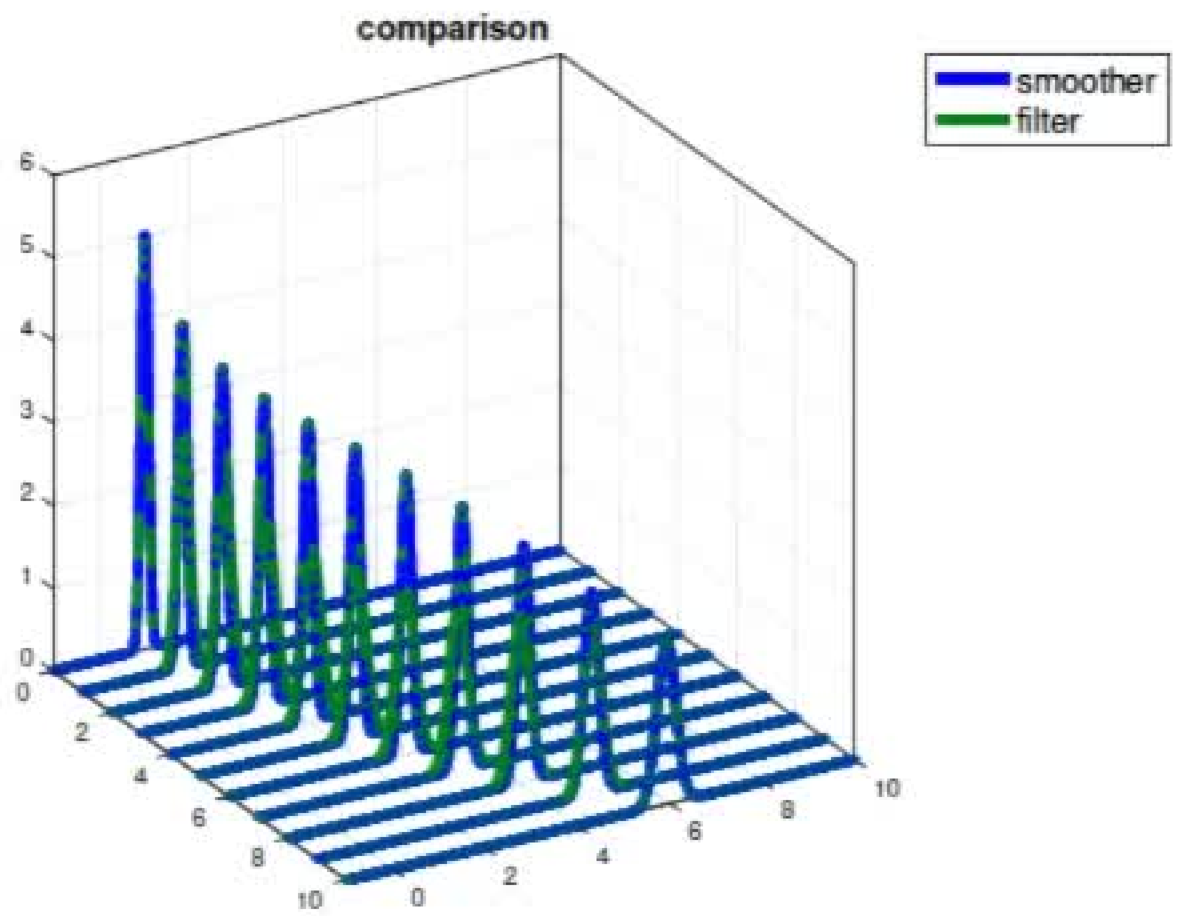
Evolution model

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

Observation model

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

Comparison Smoothing and Filtering



Kalman Filter

For a **linear model** $z_{n+1} = Az_n + \zeta_n$ with $A \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_z})$ and **linear observations** $y_{n+1} = Hz_{n+1} + \eta_{n+1}$ with $H \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_y})$, the conditional density $\mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$ is proportional to

$$\exp(-\frac{1}{2}|y_{n+1} - Hu|_\Gamma^2 - \frac{1}{2}|u - \hat{m}_{n+1}|_{\hat{C}_{n+1}}^2)$$

where $(\hat{m}_{n+1}, \hat{C}_{n+1}) = (Am_n, AC_nA^\top + \Sigma)$ is the **forecast mean and covariance**.

We have $z_{n+1}|y_1, \dots y_{n+1} \sim \mathcal{N}(m_{n+1}, C_{n+1})$ with

$$\begin{aligned} m_{n+1} &= (I - K_{n+1}H)\hat{m}_{n+1} + K_{n+1}y_{n+1} \\ C_{n+1} &= (I - K_{n+1}H)\hat{C}_{n+1} \end{aligned}$$

where $K_{n+1} = \hat{C}_{n+1}H^\top(\Gamma + H\hat{C}_{n+1}H^\top)^{-1}$ is the **Kalman gain**.

Linear Example

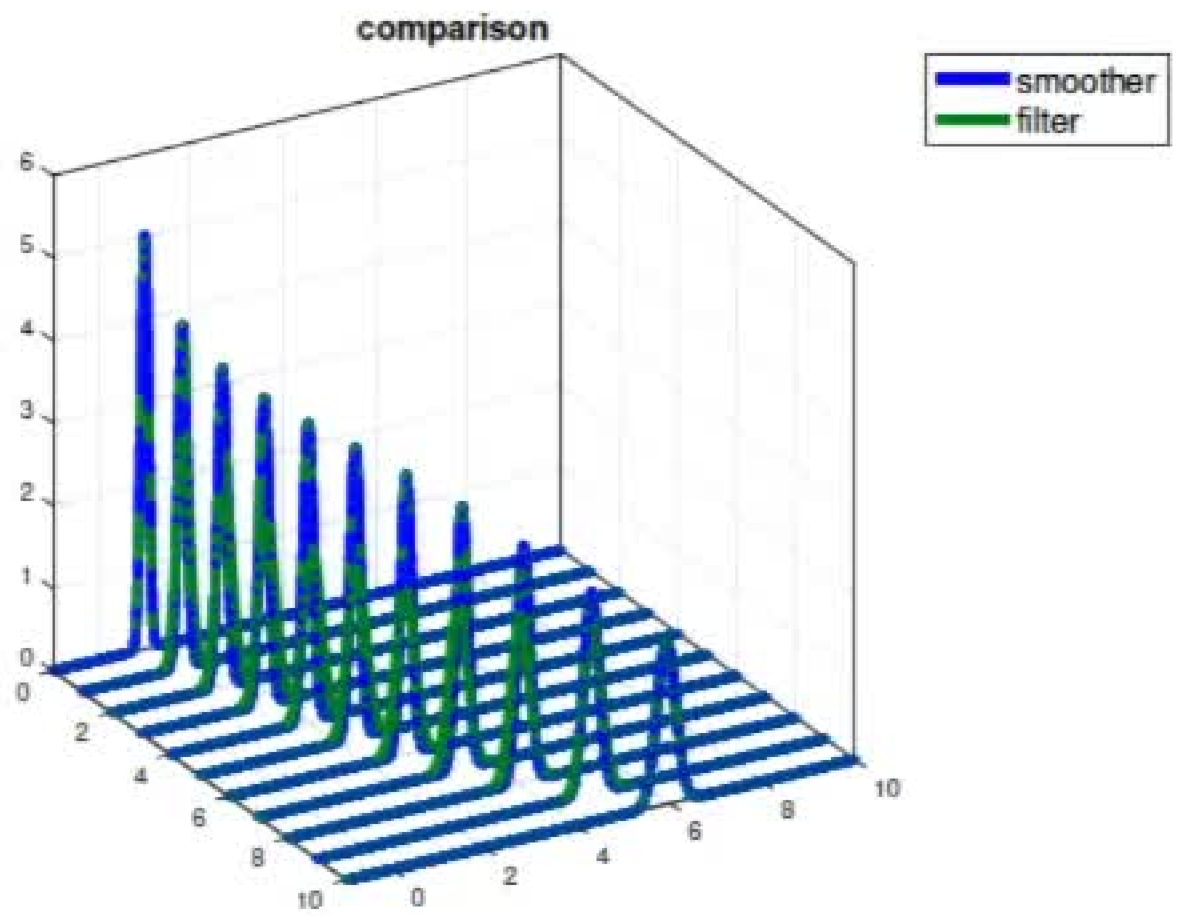
Evolution model

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

Observation model

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

Comparison Smoothing and Filtering



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

Relation between Smoothing and Filtering

We denote by $\mathbb{P}(z_0, \ldots, z_N | y_1, \ldots, y_N)$ the smoothing distribution and by $\mathbb{P}(z_N | y_1, \ldots y_N)$ the filtering distribution at time N. Then, the marginal of the smoothing distribution on z_N coincides with the filtering distribution at time N, i.e.

$$\int \mathbb{P}(z_0, \ldots, z_N | y_1, \ldots, y_N) \mathrm{d}z_0 \ldots \mathrm{d}z_{N-1} = \mathbb{P}(z_N | y_1, \ldots y_N) .$$

Note that the marginal of the smoothing distribution on z_n, $n < N$ is in general **not equal** to the filter $\mathbb{P}(z_n | y_1, \ldots y_n)$.

Linear Example

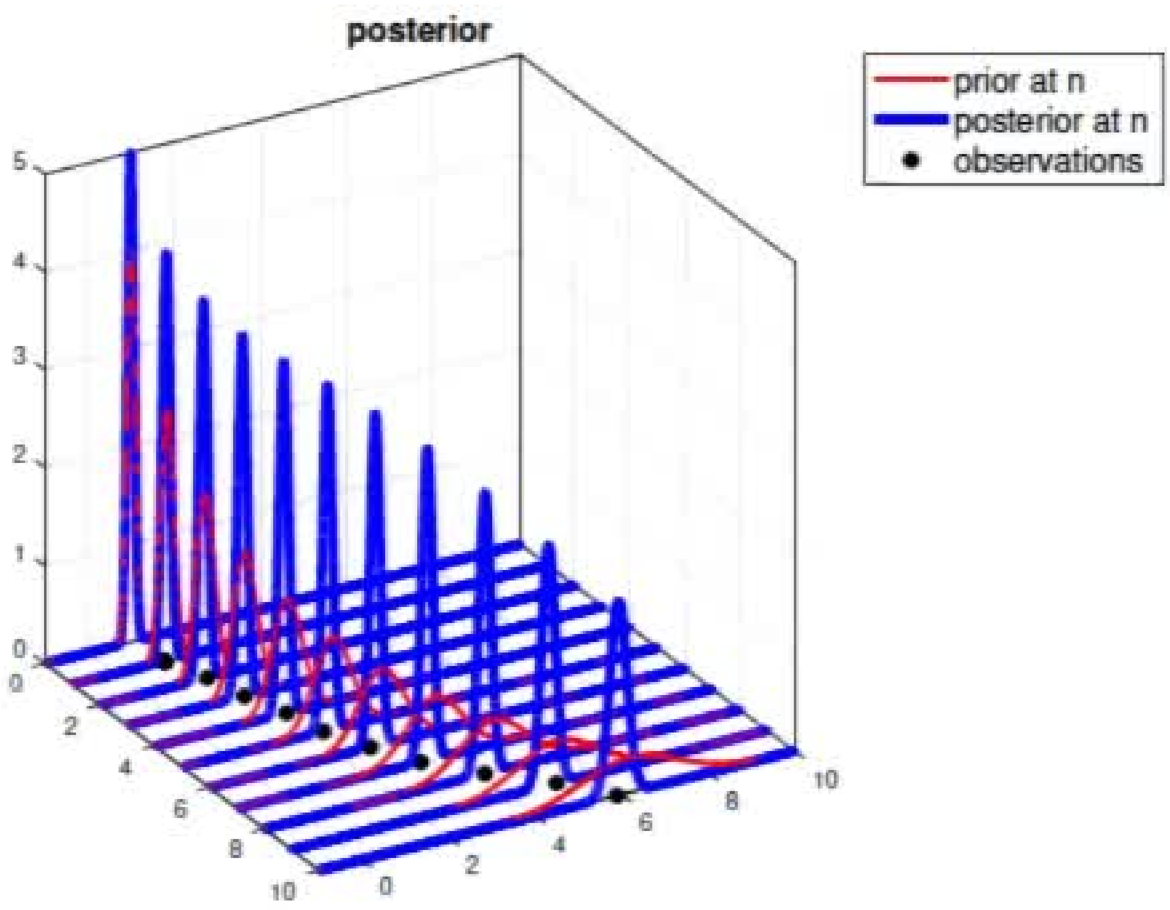
Evolution model

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

Observation model

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

Smoothing



Kalman Filter

For a **linear model** $z_{n+1} = Az_n + \zeta_n$ with $A \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_z})$ and **linear observations** $y_{n+1} = Hz_{n+1} + \eta_{n+1}$ with $H \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_y})$, the conditional density $\mathbb{P}(z_{n+1}|y_1, \dots y_{n+1})$ is proportional to

$$\exp(-\frac{1}{2}|y_{n+1} - Hu|_\Gamma^2 - \frac{1}{2}|u - \hat{m}_{n+1}|_{\hat{C}_{n+1}}^2)$$

where $(\hat{m}_{n+1}, \hat{C}_{n+1}) = (Am_n, AC_nA^\top + \Sigma)$ is the **forecast mean and covariance**.

We have $z_{n+1}|y_1, \dots y_{n+1} \sim \mathcal{N}(m_{n+1}, C_{n+1})$ with

$$\begin{aligned} m_{n+1} &= (I - K_{n+1}H)\hat{m}_{n+1} + K_{n+1}y_{n+1} \\ C_{n+1} &= (I - K_{n+1}H)\hat{C}_{n+1} \end{aligned}$$

where $K_{n+1} = \hat{C}_{n+1}H^\top(\Gamma + H\hat{C}_{n+1}H^\top)^{-1}$ is the **Kalman gain**.

Particle Filter

Let $z_n | y_1, \dots y_n \sim \mu_n$ and $z_{n+1} | y_1, \dots y_n \sim \hat{\mu}_{n+1}$.

Basic idea: Approximate μ_n and $\hat{\mu}_{n+1}$ by a J-particle Dirac measure.

$$\mu_n \approx \mu_n^J = \sum_{j=1}^{J} w_n^{(j)} \delta_{z_n^{(j)}} \qquad \hat{\mu}_{n+1} \approx \hat{\mu}_{n+1}^J = \sum_{j=1}^{J} \hat{w}_{n+1}^{(j)} \delta_{\hat{z}_{n+1}^{(j)}}$$

with $\sum_{j=1}^{J} w_n^{(j)} = \sum_{j=1}^{J} \hat{w}_{n+1}^{(j)} = 1$.

Define update rules of the particle positions and weights for the **prediction approximation**

$$\{z_n^j, w_n^{(j)}\}_{j=1}^J \mapsto \{\hat{z}_{n+1}^j, \hat{w}_{n+1}^{(j)}\}_{j=1}^J$$

and the **analysis approximation**

$$\{\hat{z}_{n+1}^j, \hat{w}_{n+1}^{(j)}\}_{j=1}^J \mapsto \{z_{n+1}^j, w_{n+1}^{(j)}\}_{j=1}^J .$$

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 $j = 1, \dots, J$.
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

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

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

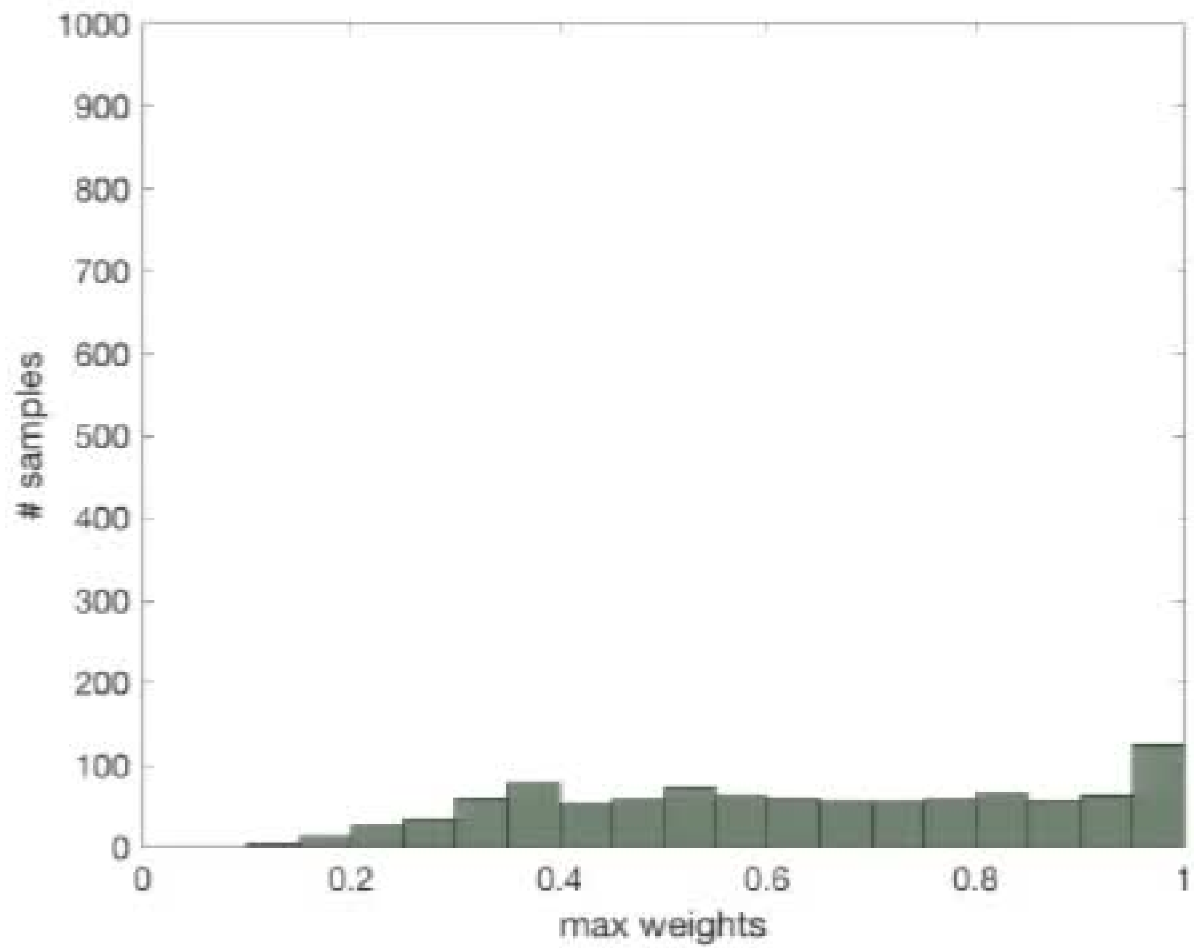
Evolution model

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

Observation model

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

$\mathbf{d = 50, J = 1000, N = 1}$
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**
- IP7 Data Assimilation and Uncertainty Quantification
  A Lagrangian Interacting Particle Perspective, Sebastian Reich
- Nonlinear Filtering and Data Assimilation in Complex Dynamical Systems MS 49, 63, 76
- Efficient Sampling Methods for Bayesian Inference in Computational Problems MS 94, 107
- Uncertainty Quantification and Data Assimilation in Earth System Modeling and Prediction MS 1, 14
- Controlled Interacting Particle Systems for Nonlinear Filtering MS 36
- Efficient Sampling Methods for Bayesian Inference in Computational Problems MS 94, 107
- Arnaud Doucet's SMC and Particle Filters Resources
  https://www.stats.ox.ac.uk/ doucet/smc_resources.html

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 .

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 startegies
  A. Chernov, H. Hoel, K. Law, F. Nobile and R. Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models*
- Ensemble tranform filters $\rightarrow$ Part II
- Hybrid Methods $\rightarrow$ Part II

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*

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 startegies
  A. Chernov, H. Hoel, K. Law, F. Nobile and R. Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models*
- Ensemble tranform filters → Part II
- Hybrid Methods → Part II

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 startegies
  A. Chernov, H. Hoel, K. Law, F. Nobile and R. Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models*
- Ensemble tranform filters → Part II
- Hybrid Methods → Part II

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 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.