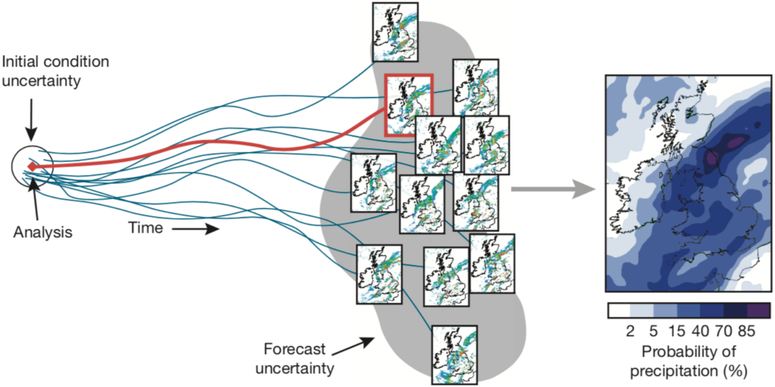
# Introduction to data assimilation

Sebastian Reich (`www.sfb1294.de`)

Universität Potsdam/ University of Reading

CIRM2018, October 26, 2018

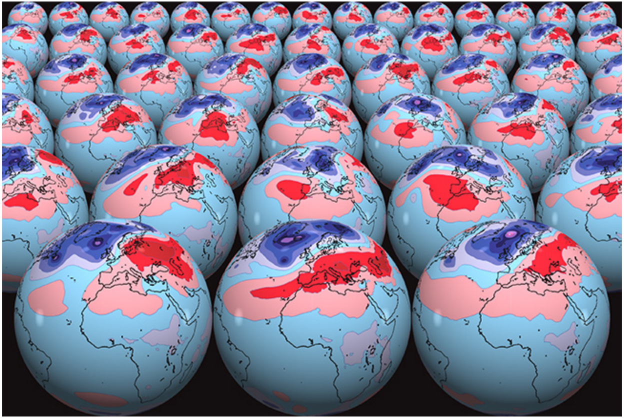
# Ensemble prediction I



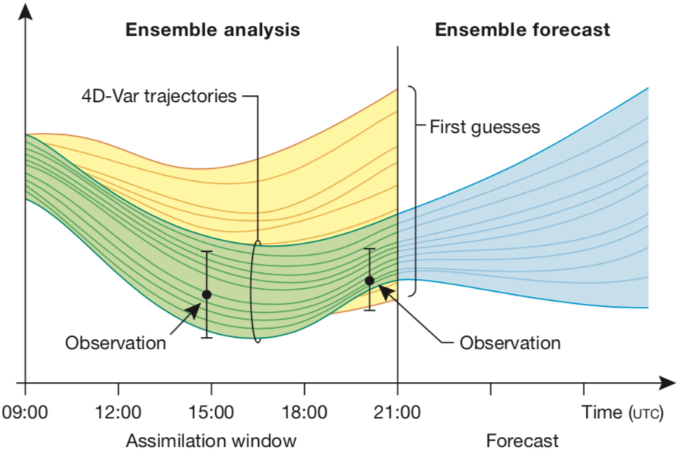

Source: The quiet revolution of numerical weather prediction, Nature, 2015

**Ensemble prediction system** with $M$ members:

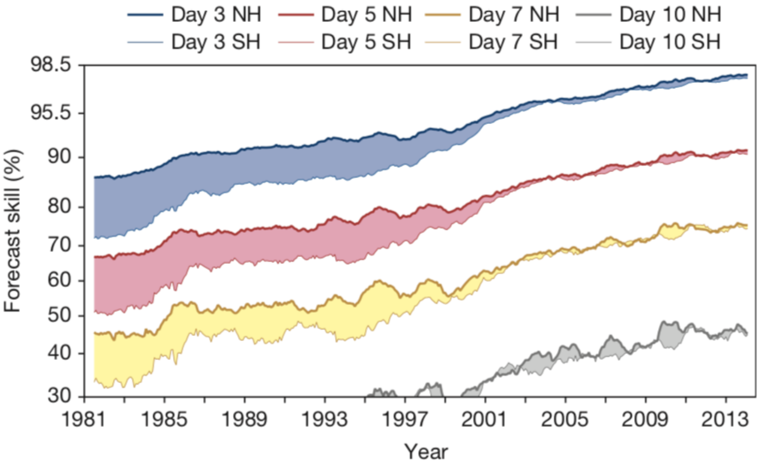
$$\frac{\mathrm{d}}{\mathrm{d}t} Z_t^i = f(Z_t^i), \qquad Z_0^i \sim \pi_0, \qquad i = 1, \ldots, M.$$



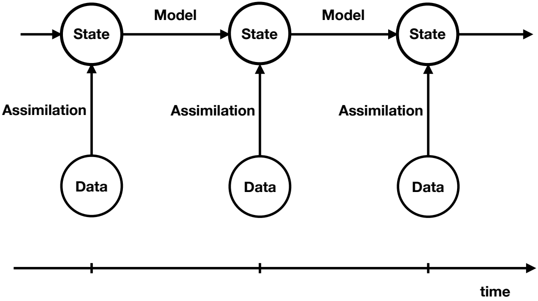
Source: ECMWF




Source: The quiet revolution of numerical weather prediction, Nature, 2015




Source: The quiet revolution of numerical weather prediction, Nature, 2015




**Model**:

$$dZ_t = f(Z_t)dt + \gamma^{1/2}dW_t$$

**Data**: $y_{t_n}$ at discrete times $t_n$

**Discrete–time observations**:

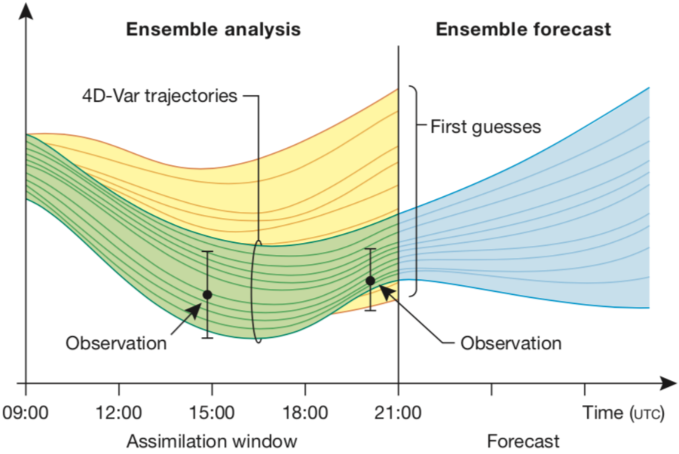
$$y_{t_n} = h(Z_{t_n}) + R^{1/2}\,\Xi_{t_n}, \qquad n = 1,\ldots,N\,.$$

**Likelihood function**:

$$L(z_{[0,T]}) := \exp\left(-\frac{1}{2}\sum_n \|y_{t_n} - h(z_{t_n})\|_R^2\right).$$

**Bayes**:

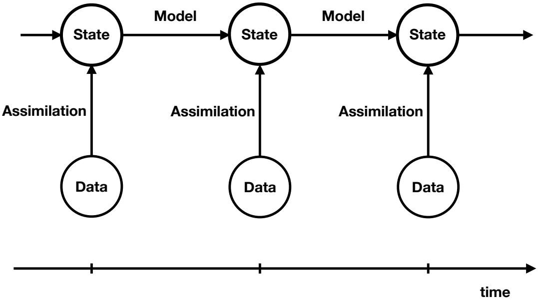
$$\frac{\mathrm{d}\widehat{\mathbb{P}}_{[0,T]}}{\mathrm{d}\mathbb{P}_{[0,T]}}(Z_{[0,T]}) := \frac{L(Z_{[0,T]})}{\mathbb{P}_{[0,T]}[L]}\,.$$




**prior** $\mathbb{P}_{[9:21]}$ in yellow; **posterior** $\widehat{\mathbb{P}}_{[9:21]}$ in green

# References


- Daley, R., Atmospheric Data Analysis, Cambridge University Press, 1991.
- Bain, A. and Crisan, D., Fundamentals of Stochastic Filtering, Springer, 2009.
- SR, Cotter, J., Probabilistic Forecasting and Bayesian Data Assimilation, Cambridge University Press, 2015.
- Law, K. et al., Data Assimilation – A Mathematical Introduction, Springer, 2015.
- P. Bauer et al, The quiet revolution of numerical weather prediction, Nature, 2015




**Notation**:

Prediction/Forecast: $\pi(\widehat{z}_{n+1} \,|\, y_{1:n})$

Filtering/Analysis: $\pi(z_{n+1} \,|\, y_{1:n+1})$

Classic **bootstrap particle filter** (sequential MC):

1) propagate $M$ particles $z_n^i$ under the model dynamics to yield **forecasts** $\hat{z}_{n+1}^i$, $i = 1, \ldots, M$.

2) data likelihood assigns **importance weights** $w_{n+1}^i$ to each propagated particle $\hat{z}_{n+1}^i$

3) **resample** to obtain equally weighted particles $z_{n+1}^i$ from the filtering distribution at $t_{n+1}$.

Classic **bootstrap particle filter** (sequential MC):

1) propagate $M$ particles $z_n^i$ under the model dynamics to yield **forecasts** $\hat{z}_{n+1}^i$, $i = 1, \ldots, M$.

2) data likelihood assigns **importance weights** $w_{n+1}^i$ to each propagated particle $\hat{z}_{n+1}^i$

3) **resample** to obtain equally weighted particles $z_{n+1}^i$ from the filtering distribution at $t_{n+1}$.

This approach is **not applicable** to geophysical/meteorological applications because of:

i) small sample sizes $M \sim 10^2$

ii) high dimensionality of models (discretised PDEs) $D \sim 10^7$

iii) medium number of observations $N \sim 10^5$

Remedies used in data assimilation:

1) Formulate assimilation of data as an **optimisation problem** (4DVar, 3DVar)

Remedies used in data assimilation:

1) Formulate assimilation of data as an **optimisation problem** (4DVar, 3DVar)
2) Replace importance–resampling step by low variance **ensemble transformations**

$$z^j_{n+1} = \sum_{i=1}^{M} \widehat{z}^i_{n+1}\, t^*_{ij}$$

# Sequential processing of data III

Remedies used in data assimilation:

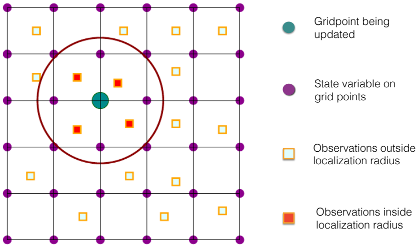
1) Formulate assimilation of data as an **optimisation problem** (4DVar, 3DVar)
2) Replace importance–resampling step by low variance **ensemble transformations**

$$z^j_{n+1} = \sum_{i=1}^{M} \hat{z}^i_{n+1}\, t^*_{ij}$$

3) **Localise** assimilation of data

$$z^j_{n+1}(\mathbf{x}) = \sum_{i=1}^{M} \hat{z}^i_{n+1}(\mathbf{x})\, t^*_{ij}(\mathbf{x})$$

$\mathbf{x} \in \mathbb{R}^3$ grid point being updated.



# Ensemble Kalman filter (EnKF)

**Forward observation model**: $y_{n+1}^{\mathrm{obs}} = h(\widehat{z}_{n+1}) + R^{1/2}\Xi_{n+1}$

**Best linear unbiased estimator** (BLUE): $z_{n+1} = A y_{n+1}^{\mathrm{obs}} + b$

**Forward observation model**: $$y^{\text{obs}}_{n+1} = h(\widehat{z}_{n+1}) + R^{1/2}\Xi_{n+1}$$

**Best linear unbiased estimator** (BLUE): $$z_{n+1} = A y^{\text{obs}}_{n+1} + b$$

**Empirical BLUE/EnKF**:

$$A = PR^{-1}$$
$$b^j = \widehat{z}^j_{n+1} - A\left(h(\widehat{z}^j_{n+1}) + R^{1/2}\widehat{\Xi}^j_{n+1}\right)$$

with

$$P = \frac{1}{M-1}\sum_{i=1}^{M} \widehat{z}^i_{n+1}(h(\widehat{z}^i_{n+1}) - \overline{h}_{n+1})^T, \qquad \widehat{\Xi}^i_{n+1} \sim \mathrm{N}(0, I)\,.$$

**Forward observation model**: $$y^{\text{obs}}_{n+1} = h(\widehat{z}_{n+1}) + R^{1/2}\Xi_{n+1}$$

**Best linear unbiased estimator** (BLUE): $$z_{n+1} = Ay^{\text{obs}}_{n+1} + b$$

**Empirical BLUE/EnKF**:

$$A = PR^{-1}$$
$$b^j = \widehat{z}^j_{n+1} - A\left(h(\widehat{z}^j_{n+1}) + R^{1/2}\widehat{\Xi}^j_{n+1}\right)$$

with

$$P = \frac{1}{M-1}\sum_{i=1}^{M}\widehat{z}^i_{n+1}(h(\widehat{z}^i_{n+1}) - \overline{h}_{n+1})^T, \qquad \widehat{\Xi}^i_{n+1} \sim \mathrm{N}(0,I)\,.$$

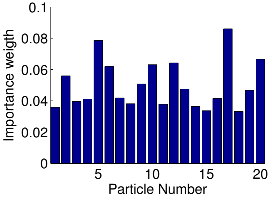
EnKF induced **particle transformation**:

$$z^j_{n+1} = \sum_{i=1}^{M}\widehat{z}^i_{n+1}t^*_{ij}\,.$$

**Importance weights**: $$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\widehat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$

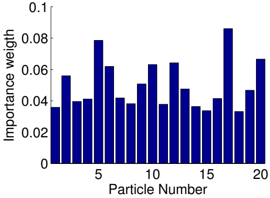
**Importance weights**:

$$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\widehat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$



**Importance weights**:

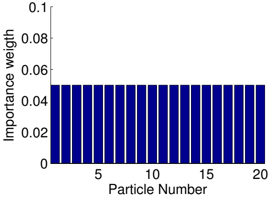
$$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\hat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$




$\Rightarrow$

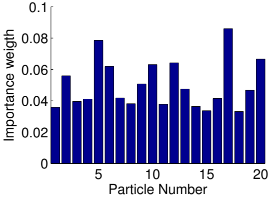
**Optimal Coupling**
$T^*$

$\Rightarrow$



**Importance weights**:

$$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\widehat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$

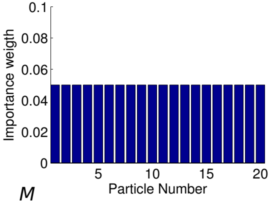



$\Rightarrow$

**Optimal Coupling**

$T*$

$\Rightarrow$




$T*$ minimiser of

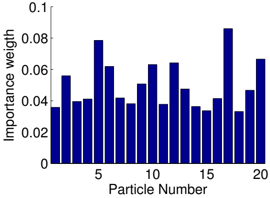
$$J(\{t_{ij}\}) = \sum_{i,j=1}^{M} t_{ij}\|\widehat{z}^i_{n+1} - \widehat{z}^j_{n+1}\|^2$$

subject to:

$$t_{ij} \geq 0, \qquad \sum_i t_{ij} = 1, \qquad \sum_j t_{ij} = w_i M$$

**Importance weights**:

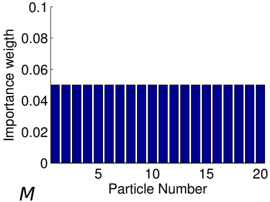
$$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\widehat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$




$\Rightarrow$

**Optimal Coupling**
$T^*$

$\Rightarrow$




$T^*$ minimiser of

$$J(\{t_{ij}\}) = \sum_{i,j=1}^{M} t_{ij}\|\widehat{z}^i_{n+1} - \widehat{z}^j_{n+1}\|^2$$

subject to:

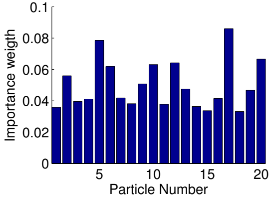
$$t_{ij} \geq 0, \qquad \sum_i t_{ij} = 1, \qquad \sum_j t_{ij} = w_i M$$

$T^*$ induces linear transformation:

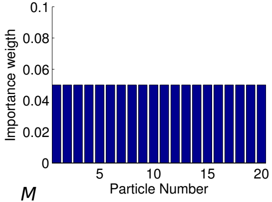
$$z^j_{n+1} = \sum_{i=1}^{M} \widehat{z}^i_{n+1} t^*_{ij}$$

**Importance weights**:

$$w^i_{n+1} \propto \exp\left(-\frac{1}{2}\|h(\widehat{z}^i_{n+1}) - y_{n+1})\|^2_R\right)$$




$\Rightarrow$ **Optimal Coupling** $T^*$ $\Rightarrow$




$T^*$ minimiser of

$$J(\{t_{ij}\}) = \sum_{i,j=1}^{M} t_{ij}\|\widehat{z}^i_{n+1} - \widehat{z}^j_{n+1}\|^2$$

subject to:

$$t_{ij} \geq 0, \qquad \sum_i t_{ij} = 1, \qquad \sum_j t_{ij} = w_i M$$

$T^*$ induces linear transformation:

$$z^j_{n+1} = \sum_{i=1}^{M} \widehat{z}^i_{n+1} t^*_{ij}$$

Lorenz-63 model, first component observed infrequently ($\Delta t = 0.12$) and with large measurement noise ($R = 8$):

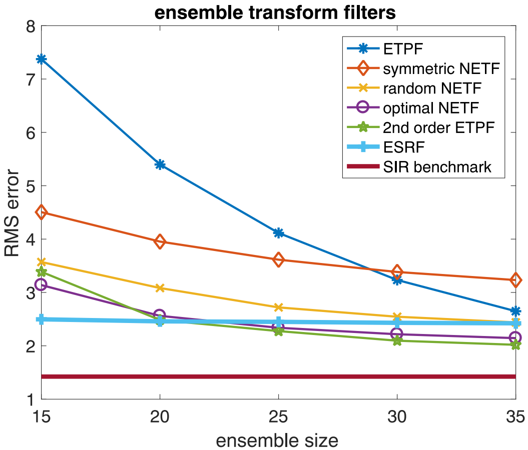



Figure: RMSEs for various second-order accurate LETFs compared to the ETPF, the ESRF, and the SIR PF as a function of the sample size, $M$.

Hybrid filter: $\mathbf{P} := \mathbf{P}_{\mathrm{ESRF}}(\alpha)\,\mathbf{P}_{\mathrm{ETPF}}(1-\alpha)$.

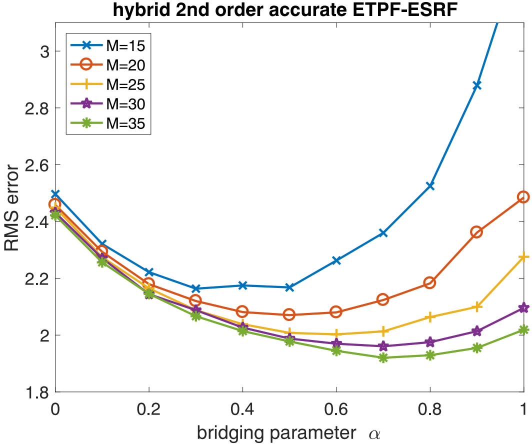



Figure: RMSEs for hybrid ESRF ($\alpha = 0$) and 2nd-order corrected NETF/ETPF ($\alpha = 1$) as a function of the sample size, $M$.

# References


- Evensen, G., Data assimilation – the ensemble Kalman filter, Springer, 2009.
- SR, A nonparametric ensemble transform method for Bayesian inference, SIAM J. Sci. Comput., 2013.
- SR, Cotter, J., Probabilistic Forecasting and Bayesian Data Assimilation, Cambridge University Press, 2015.
- Tödter J., Ahrens, B., A second–order exact ensemble square root filter for nonlinear data assimilation, Month. Weather Rev., 2015.
- Chustagulprom, N. et al., A hybrid ensemble transform particle filter for nonlinear and spatially extended systems, SIAM/ASA J. UQ, 2016.
- Acevedo, W., et al., Second–order accurate ensemble transform particle filters, SIAM J. Sci. Comput., 2017.
- Fearnhead, P., Künsch, H.R., Particle filters and data assimilation, Annual Review of Statistics and Its Application, 2018.

**Observation model**:

$$\mathrm{d}y_t = h(Z_t)\mathrm{d}t + R^{1/2}\mathrm{d}V_t$$

**Observation model**:

$$\mathrm{d}y_t = h(Z_t)\mathrm{d}t + R^{1/2}\mathrm{d}V_t$$

**Linear SDEs and observations**:

$$\mathrm{d}Z_t = (AZ_t + b)\,\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t, \qquad \mathrm{d}y_t = HZ_t\mathrm{d}t + R^{1/2}\mathrm{d}V_t\,.$$

**Observation model**:

$$\mathrm{d}y_t = h(Z_t)\mathrm{d}t + R^{1/2}\mathrm{d}V_t$$

**Linear SDEs and observations**:

$$\mathrm{d}Z_t = (AZ_t + b)\,\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t, \qquad \mathrm{d}y_t = HZ_t\mathrm{d}t + R^{1/2}\mathrm{d}V_t\,.$$

**Kalman-Bucy equations** for the mean

$$\mathrm{d}\mu_t = (A\mu_t + b)\mathrm{d}t - K_t\mathrm{d}I_t, \qquad \mathrm{d}I_t := (H\mu_t\mathrm{d}t - \mathrm{d}y_t)$$

and the covariance matrix

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t = AP_t + P_tA^T + \gamma I - K_tHP_t\,.$$

with Kalman gain matrix $K_t = P_tH^TR^{-1}$.

Controlled mean–field SDE:

$$\mathrm{d}Z_t = (AZ_t + b)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t + du_t(Z_t), \qquad du_t := -K_t \mathrm{d}\mathcal{I}_t$$

with innovation $\mathcal{I}_t$

$$\mathrm{d}\mathcal{I}_t = \frac{1}{2}(HZ_t + H\mu_t)\mathrm{d}t - \mathrm{d}y_t$$

or

$$\mathrm{d}\mathcal{I}_t = HZ_t\mathrm{d}t + R^{1/2}\mathrm{d}U_t - \mathrm{d}y_t$$

and Kalman gain $K_t = P_t H^T R^{-1}$.

Controlled mean–field SDE:

$$\mathrm{d}Z_t = (AZ_t + b)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t + du_t(Z_t), \qquad du_t := -K_t \mathrm{d}\mathcal{I}_t$$

with innovation $\mathcal{I}_t$

$$\mathrm{d}\mathcal{I}_t = \frac{1}{2}(HZ_t + H\mu_t)\mathrm{d}t - \mathrm{d}y_t$$

or

$$\mathrm{d}\mathcal{I}_t = HZ_t\mathrm{d}t + R^{1/2}\mathrm{d}U_t - \mathrm{d}y_t$$

and Kalman gain $K_t = P_t H^T R^{-1}$.

Extension to nonlinear SDEs:

**Ensemble Kalman-Bucy filter**(EnKBF)

$$\mathrm{d}tZ_t = f(Z_t)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t - K_t \mathrm{d}\mathcal{I}_t\,.$$

**Feedback particle filter**:

$$\mathrm{d}Z_t = f(Z_t)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t - K_t \circ \mathrm{d}\mathcal{I}_t$$

with Kalman gain

$$K_t := \nabla_z \psi_t R^{-1}, \qquad \nabla_z \cdot (\pi_t \nabla_z \psi_t) = -\pi_t (h - \bar{h})$$

and innovations

$$\mathrm{d}\mathcal{I}_t = \frac{1}{2}(h(Z_t) + \bar{h}_t)\mathrm{d}t - \mathrm{d}y_t$$

or

$$\mathrm{d}\mathcal{I}_t = h(Z_t)\mathrm{d}t + R^{1/2}\mathrm{d}U_t - \mathrm{d}y_t\,.$$

**Particle approximation** to elliptic PDE:

$$\nabla_z \psi_t(z_t^j) \approx \sum_{i=1}^{M} z_t^i \, a_t^{ij}$$

with $z_t^i \sim \pi_t$, $i = 1, \ldots, M$ and coefficients $a_t^{ij} \in \mathbb{R}$ appropriately defined.

**Particle approximation** to elliptic PDE:

$$\nabla_z \psi_t(z_t^j) \approx \sum_{i=1}^{M} z_t^i \, a_t^{ij}$$

with $z_t^i \sim \pi_t$, $i = 1, \ldots, M$ and coefficients $a_t^{ij} \in \mathbb{R}$ appropriately defined.

**Interacting particle system**:

$$\mathrm{d}z_t^j = f(z_t^j)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t^j - \left(\sum_{i=1}^{M} z_t^i \, a_t^{ij}\right) R^{-1} \circ \mathrm{d}\mathcal{I}_t^j$$

with innovation

$$\mathrm{d}\mathcal{I}_t^j = h(z_t^j)\mathrm{d}t + R^{1/2}\mathrm{d}U_t^j - \mathrm{d}y_t \, .$$

**Parameter–dependent model**:

$$\mathrm{d}Z_t = f(Z_t, \theta)\mathrm{d}t + \gamma^{1/2}\mathrm{d}W_t - K_t \circ \mathrm{d}\mathcal{I}_t$$

**Marginal likelihood** (evidence of the data $\mathcal{Y}_t = y_{[0,t]}$ under $\theta$):

$$\mathrm{d}\pi_t(\mathcal{Y}_t|\theta) = \pi_t(\mathcal{Y}_t|\theta)\,\bar{h}_t R^{-1}\mathrm{d}y_t\,.$$

**Parameter–dependent model**:

$$dZ_t = f(Z_t, \theta)dt + \gamma^{1/2}dW_t - K_t \circ d\mathcal{I}_t$$

**Marginal likelihood** (evidence of the data $\mathcal{Y}_t = y_{[0,t]}$ under $\theta$):

$$d\pi_t(\mathcal{Y}_t|\theta) = \pi_t(\mathcal{Y}_t|\theta)\,\bar{h}_t R^{-1} dy_t\,.$$

**Example.**

SPDE model:

$$\dot{u}_t = \theta\,\Delta u_t + \dot{W}_t$$

$\dot{W}_t$ space–time white noise.

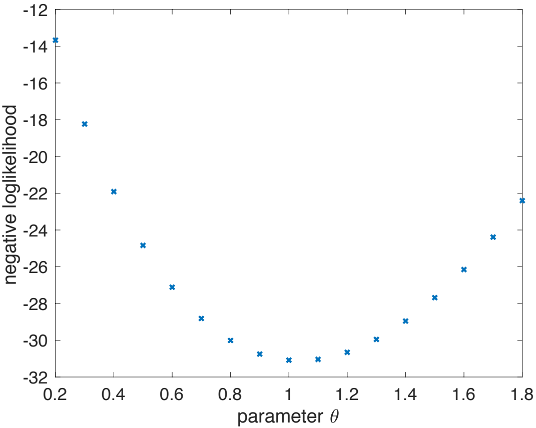
Observation:

$$dy_t = u_t(x_0)\,dt + R^{1/2}dV_t$$

for fixed spatial location $x_0$.

# Parameter estimation II

$D = 80$ grid points, state estimation with EnKBF, $M = 20$ particles, and localisation, observation interval $[0, 20]$, measurement error $R = 0.001$



# References


- Bergemann, K, SR, An ensemble Kalman–Bucy filter for continuous–time data assimilation, Meteorolog. Zeitschrift, 2012.
- Yang, T. et al., Feedback particle filter, IEEE Trans. Autom. Control, 2013.
- Taghvaei, A. et al., Kalman filter and its modern extensions for the continuous–time nonlinear filtering problem, J. Dyn. Sys., Meas., and Control, 2018.
- de Wiljes, et al., Long–time stability and accuracy of the ensemble Kalman–Bucy filter for fully observed processes and small measurement noise, SIAM J. Appl. Dyn. Sys., 2018.
- SR, Data assimilation, Acta Numerica, 2019, preliminary version on arXiv 1807.08351.

## Summary

- Data assimilation has become a workhorse for day–to–day weather forecasting
- Key are a "smart" choice of biased and robust approximations to the sequential filtering problem
- Data assimilation has applications to many areas outside weather prediction such as medicine, geosciences, biology, etc.
- Current topics of research: combined state–parameter estimation for PDEs, theoretical understanding of approximations such as linear transformations & localisation
- Continuous–in–time DA naturally leads to an attractive unifying framework in the form of mean–field SDEs.

# The End

"Miss Peterson, may I go home? I can't assimilate any more data today."

(source: J.B. Handelsman, New Yorker, 05/31/1969)

# Collaborators

- Walter Acevedo
- Kay Bergemann
- Yuan Cheng
- Nawinda Chustagulprom
- Colin Cotter
- Jana de Wiljes
- Prashant Mehta
- Wilhelm Stannat
- Amari Taghvaei