A generalisation of the updating step in EnKF

Håkon Tjelmeland and Margrethe K. Loe
Department of Mathematical Sciences
Norwegian University of Science and Technology

EnKF Workshop 2021

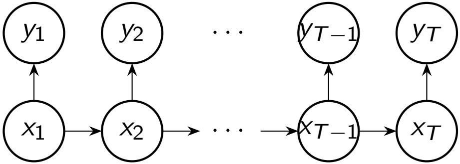
June 11th 2021

Talk outline

- ⋆ The state-space model
- ⋆ Ensemble Kalman filter (EnKF)
  - model and algorithm
  - simulation example
  - identify some issues
- ⋆ An improved model (and algorithm)
  - modified model
  - modified updating algorithm
  - simulation example (revisited)
- ⋆ Closing remarks
- ⋆ Note: We do not consider
  - variance inflation
  - localisation
  - computational efficiency

The state-space model

- State-space model



- Model components

$$
\begin{aligned}
x_1 &\sim p(x_1) \\
x_t|x_{t-1} &\sim p(x_t|x_{t-1}) \\
y_t|x_t &\sim \mathrm{N}(Hx_t, R)
\end{aligned}
$$

- Goal: Find/represent the filtering distribution

$$p(x_t|y_{1:t})$$

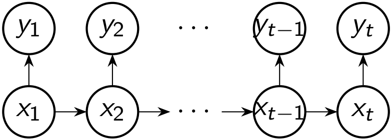
  - represent $p(x_t|y_{1:t})$ by an ensemble $\{x_t^1, \ldots, x_t^M\}$

Ensemble Kalman filter — recursive solution

- From $p(x_t|y_{1:t})$ to $p(x_{t+1}|y_{1:t+1})$ in to steps

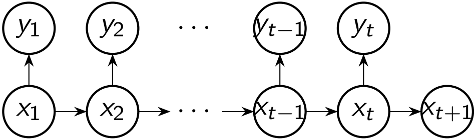
$p(x_t|y_{1:t})$

$\{x_t^1, \ldots, x_t^M\}$



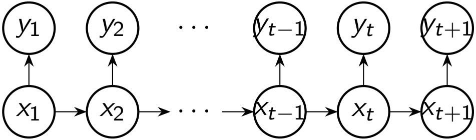
$p(x_{t+1}|y_{1:t})$

$\{\tilde{x}_{t+1}^1, \ldots, \tilde{x}_{t+1}^M\}$



$p(x_{t+1}|y_{1:t+1})$

$\{x_{t+1}^1, \ldots, x_{t+1}^M\}$

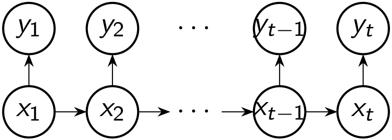


Ensemble Kalman filter — recursive solution

- ⋆ From $p(x_t|y_{1:t})$ to $p(x_{t+1}|y_{1:t+1})$ in to steps

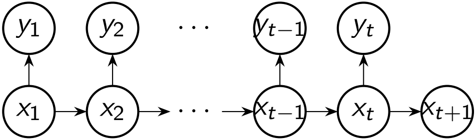
$p(x_t|y_{1:t})$

$\{x_t^1, \ldots, x_t^M\}$



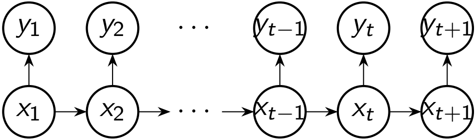
$p(x_{t+1}|y_{1:t})$

$\{\widetilde{x}_{t+1}^1, \ldots, \widetilde{x}_{t+1}^M\}$



$p(x_{t+1}|y_{1:t+1})$

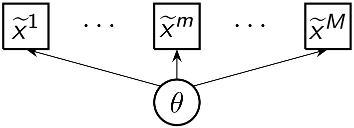
$\{x_{t+1}^1, \ldots, x_{t+1}^M\}$

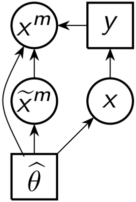


- ⋆ From now on: remove the time index notation

Ensemble Kalman filter — update step

- Assumed models when updating $\widetilde{x}^m$ to x^m (new data: y)





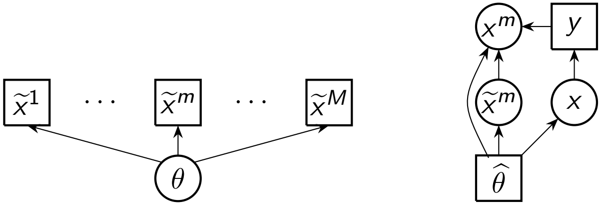
- First assumed model

$$\widetilde{x}^1, \ldots, \widetilde{x}^M | \theta \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \Sigma)$$

  - $\theta = (\mu, \Sigma)$
  - $\widehat{\theta} = (\widehat{\mu}, \widehat{\Sigma})$: empirical quantities

Ensemble Kalman filter — update step

⋆ Assumed models when updating $\widetilde{x}^m$ to x^m (new data: y)



⋆ Second assumed model

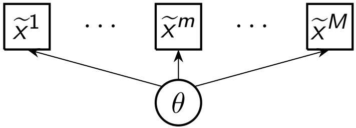
$$\widetilde{x}^m, x|\widehat{\theta} \overset{\text{iid}}{\sim} \mathrm{N}(\widehat{\mu}, \widehat{\Sigma})$$

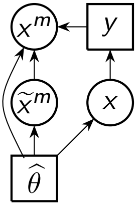
$$y|x \sim \mathrm{N}(Hx, R)$$

⋆ Require: $x^m|\widehat{\theta}, y \quad \overset{\text{d}}{=} \quad x|\widehat{\theta}, y$

Ensemble Kalman filter — update step

- Assumed models when updating $\widetilde{x}^m$ to x^m (new data: y)





- Second assumed model

$$\widetilde{x}^m, x|\widehat{\theta} \overset{\text{iid}}{\sim} \mathrm{N}(\widehat{\mu}, \widehat{\Sigma})$$

$$y|x \sim \mathrm{N}(Hx, R)$$

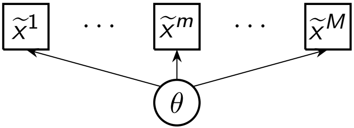
- Require: $x^m|\widehat{\theta}, y \;\overset{\text{d}}{=}\; x|\widehat{\theta}, y$
- Stochastic EnKF:

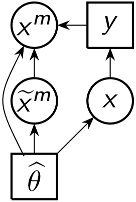
$$x^m = \widetilde{x}^m + K(y + \varepsilon - H\widetilde{x}^m)$$

where $K = \widehat{\Sigma}H^T(H\widehat{\Sigma}H^T + R)^{-1}$ and $\varepsilon \sim \mathrm{N}(0, R)$

Ensemble Kalman filter — update step

- ⋆ Assumed models when updating $\widetilde{x}^m$ to x^m (new data: y)





- ⋆ Second assumed model

$$\widetilde{x}^m, x | \widehat{\theta} \overset{\text{iid}}{\sim} \mathrm{N}(\widehat{\mu}, \widehat{\Sigma})$$

$$y | x \sim \mathrm{N}(Hx, R)$$

- ⋆ Require: $x^m | \widehat{\theta}, y \quad \overset{\text{d}}{=} \quad x | \widehat{\theta}, y$
- ⋆ Deterministic EnKF (square root filter)

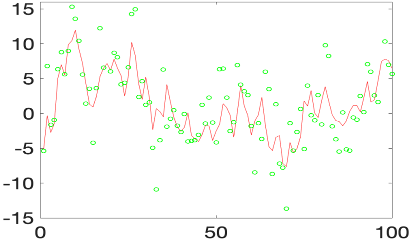
$$x^m = \widehat{\mu} + K(y - H\widehat{\mu}) + B(\widetilde{x}^m - \widetilde{\mu})$$

where $B\widehat{\Sigma}B^T = (\mathbb{I} - KH)\widetilde{\Sigma}$

Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

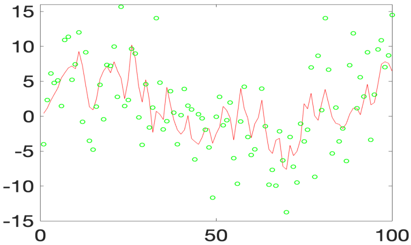
$t = 1$



Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

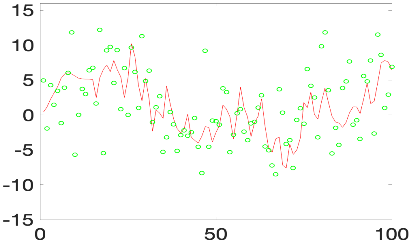
$t = 2$



Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

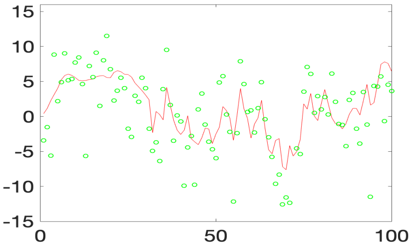
$t = 3$



Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

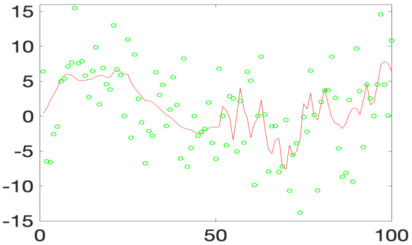
$t = 6$



Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function $(p(x_{t+1}|x_t))$ is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

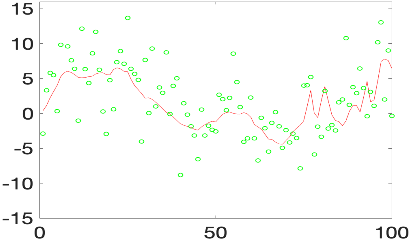
$t = 10$



Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

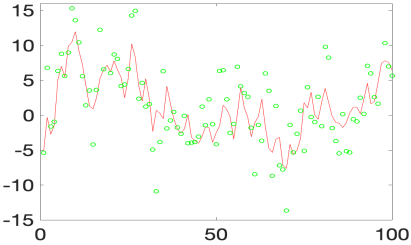
$t = 15$



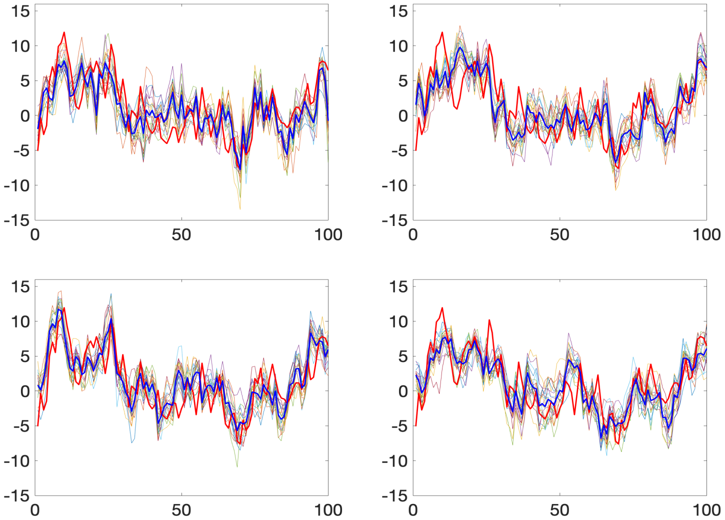
Simulation example with stochastic EnKF

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_t is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function ($p(x_{t+1}|x_t)$) is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Likelihood, $y_n|x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

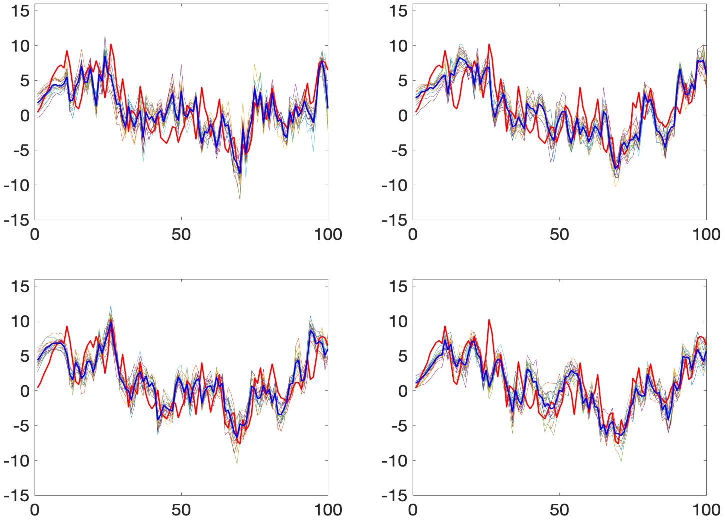
$t = 1$



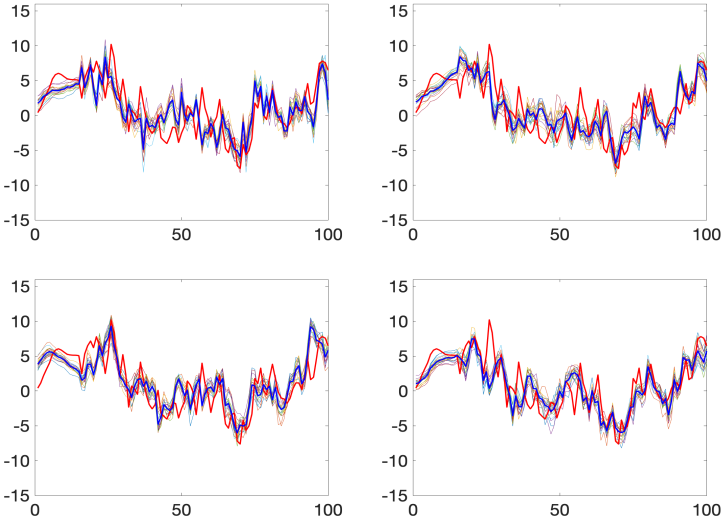
EnKF results with $M = 19$ ensemble members, time $t = 1$



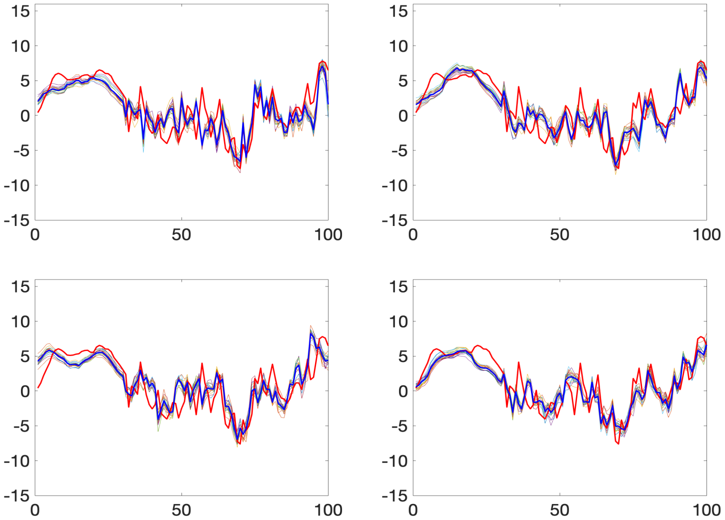
EnKF results with $M = 19$ ensemble members, time $t = 2$



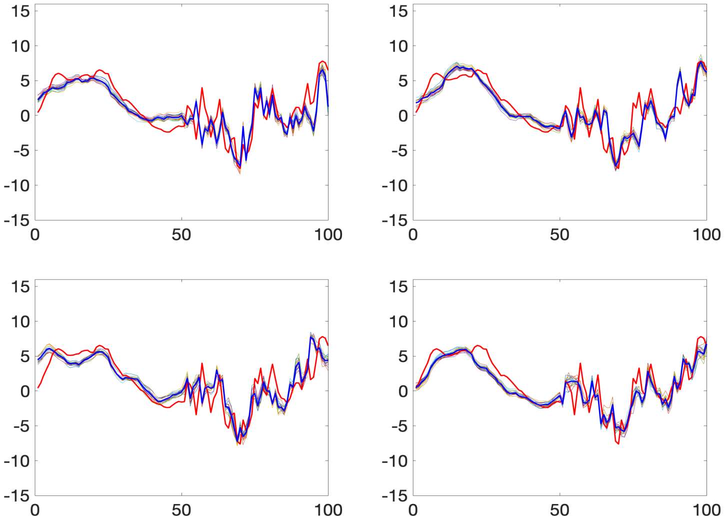
EnKF results with $M = 19$ ensemble members, time $t = 3$



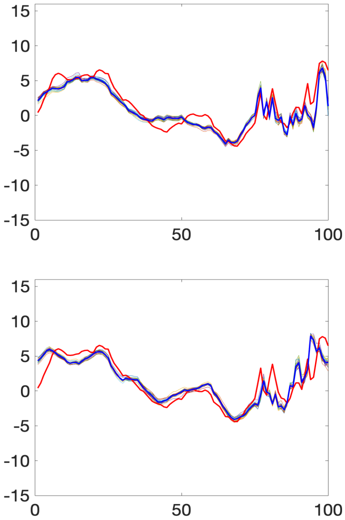
EnKF results with $M = 19$ ensemble members, time $t = 6$



EnKF results with $M = 19$ ensemble members, time $t = 10$

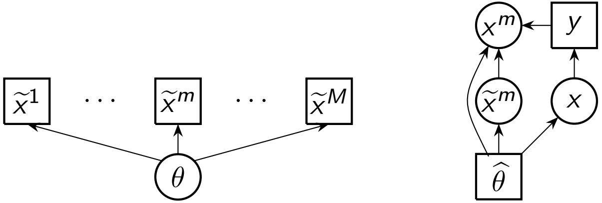


EnKF results with $M = 19$ ensemble members, time $t = 15$



Issues with the EnKF

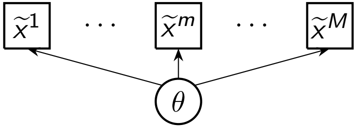
⋆ Recall: Underlying models for the EnKF updating of $\widetilde{x}^m$ to x^m

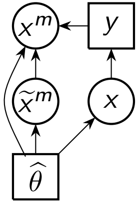


⋆ Issues:
- uncertainty in $\widehat{\theta}$ is ignored
  + Myrseth and Omre (2010), Tsyrulnikov and Rakitko (2017)
- the information in y about θ is ignored
  + discussed in Myrseth and Omre (2010)
- the information in $\widetilde{x}^m$ is used two times
  + and inconsistently?

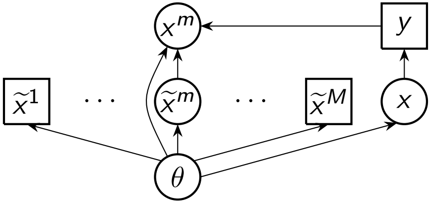
Propose new model for the updating step

⋆ Recall: Standard EnKF model for updating $\widetilde{x}^m$ to x^m



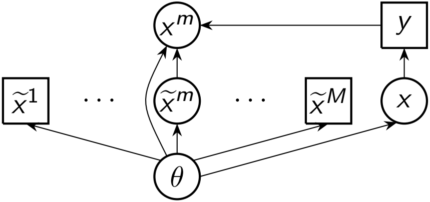


⋆ Propose to base update on a merged model



Propose new model for the updating step

- ⋆ Propose to base update on a merged model



- ⋆ Assumed model:

$$\widetilde{x}^1, \ldots, \widetilde{x}^M, x|\theta \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \Sigma)$$

$$\theta = (\mu, \Sigma) \sim \mathsf{NIW}(\mu_0, \lambda, \Psi, \nu)$$

$$y|x \sim \mathsf{N}(Hx, R)$$

- ⋆ Require: $x^m|\widetilde{x}^{-m}, y \quad \overset{\text{d}}{=} \quad x|\widetilde{x}^{-m}, y$

Class of possible of update procedures

- ⋆ Require: $x^m | \widetilde{x}^{-m}, y \quad \stackrel{\mathsf{d}}{=} \quad x | \widetilde{x}^{-m}, y$
- ⋆ Class of possible solutions:

$$\theta = (\mu, \Sigma) \sim f(\theta | \widetilde{x}^{-m}, y)$$
$$x^m = B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu) + \varepsilon$$

where

$$\varepsilon \sim \mathrm{N}(0, S) \quad \text{and} \quad S = (\mathbb{I} - KH)\Sigma - B\Sigma B^T \geq 0$$

Class of possible of update procedures

- Require: $x^m|\widetilde{x}^{-m}, y \quad \overset{\mathsf{d}}{=} \quad x|\widetilde{x}^{-m}, y$
- Class of possible solutions:

$$\theta = (\mu, \Sigma) \sim f(\theta|\widetilde{x}^{-m}, y)$$
$$x^m = B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu) + \varepsilon$$

where

$$\varepsilon \sim \mathrm{N}(0, S) \quad \text{and} \quad S = (\mathbb{I} - KH)\Sigma - B\Sigma B^T \geq 0$$

- Three special cases:
  - $S = 0$: square root filter, many possible B's
  - $B = \mathbb{I} - KH$: ensemble Kalman filter update
  - $B = 0$: sample x^m independently of $\widetilde{x}^m$

Class of possible of update procedures

- Require: $x^m|\widetilde{x}^{-m}, y \quad \stackrel{\text{d}}{=} \quad x|\widetilde{x}^{-m}, y$
- Class of possible solutions:

$$\begin{aligned} \theta &= (\mu, \Sigma) \sim f(\theta|\widetilde{x}^{-m}, y) \\ x^m &= B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu) + \varepsilon \end{aligned}$$

  where

$$\varepsilon \sim \mathrm{N}(0, S) \quad \text{and} \quad S = (\mathbb{I} - KH)\Sigma - B\Sigma B^T \geq 0$$

- Three special cases:
  - $S = 0$: square root filter, many possible B's
  - $B = \mathbb{I} - KH$: ensemble Kalman filter update
  - $B = 0$: sample x^m independently of $\widetilde{x}^m$
- What is "the best" (B, S)?

Optimality criterion

- ⋆ What is the best (B, S)?
- ⋆ If the assumed model is correct:
  - all allowed choices of (B, S) are equally good!
- ⋆ If the assumed model is wrong:
  - all allowed choices of (B, S) are wrong
  - wants an update procedure that is robust
- ⋆ Intuition: Should do small changes

$$x^m = B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu) + \varepsilon$$

- ⋆ Our optimality criterion: want to minimise

$$\mathrm{E}\left[(x^m - \widetilde{x}^m)^T (x^m - \widetilde{x}^m) \middle| \widetilde{x}^{-m}, y\right]$$

with respect to B and S, under the restriction

$$S = (\mathbb{I} - KH)\Sigma - B\Sigma B^T \geq 0$$

- ⋆ The solution can be found analytically

Optimal solution

- ⋆ Recall: Update procedure:

$$\theta = (\mu, \Sigma) \sim f(\theta | \widetilde{x}^{-m}, y)$$
$$x^m = B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu) + \varepsilon$$

where

$$\varepsilon \sim \mathrm{N}(0, S) \quad \text{and} \quad S = (\mathbb{I} - KH)\Sigma - B\Sigma B^T \geq 0$$

- ⋆ Optimal solution of (B, S):

$$S = 0 \quad \text{and} \quad B = U\Lambda^{\frac{1}{2}} F P^T D^{-\frac{1}{2}} V^T$$

where (using singular value decomposition)

$$\Sigma = VDV^T, \quad (\mathbb{I} - KH)\Sigma = U\Lambda U^T \quad \text{and} \quad \Lambda^{\frac{1}{2}} U^T \Sigma V D^{-\frac{1}{2}} = PGF^T$$

Resulting computational procedure

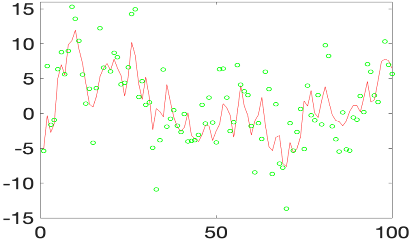
- ⋆ For $m = 1, \ldots, M$
    1. Sample $\theta = (\mu, \Sigma) \sim f(\theta | \widetilde{x}^{-m}, y)$
    2. From θ and y compute optimal weight matrix B
    3. Compute

$$x^m = B(\widetilde{x}^m - \mu) + \mu + K(y - H\mu)$$

Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

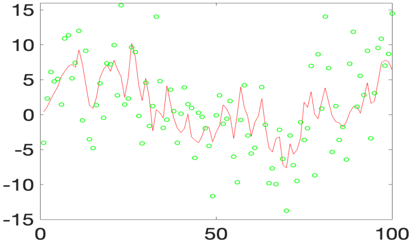
$t = 1$



Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

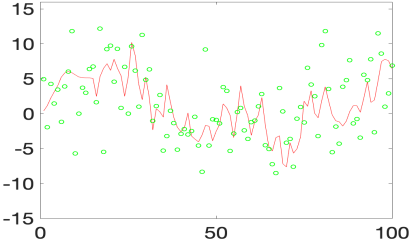
$t = 2$



Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

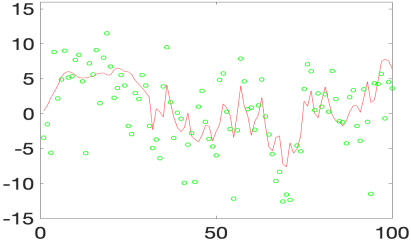
$t = 3$



Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

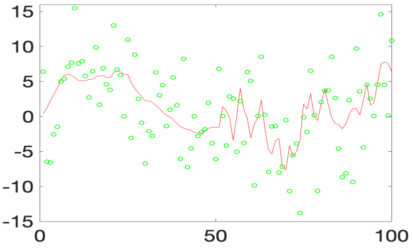
$t = 6$



Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - linear: smoothing for ten nodes
  - (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

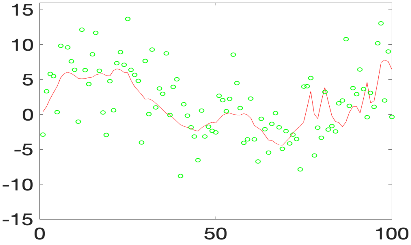
$t = 10$



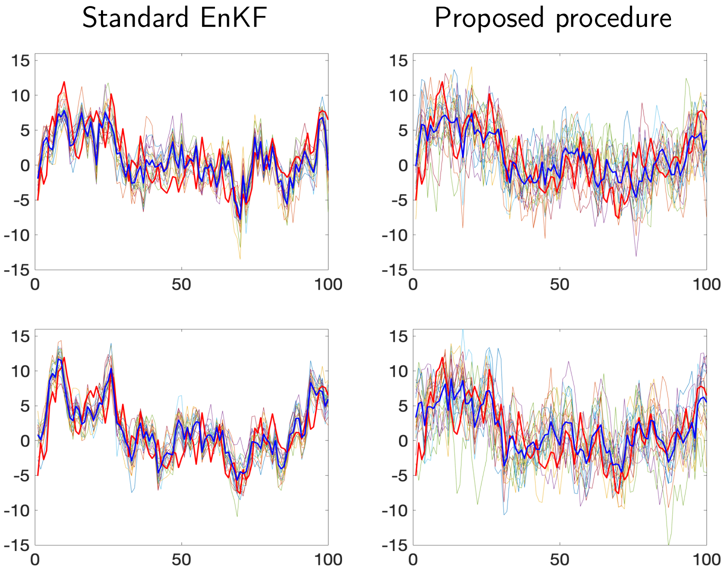
Simulation example revisited

- ⋆ Use an example introduced in Myrseth and Omre (2010)
- ⋆ State vector x_n is a vector of size 100
- ⋆ $x_1 \sim \mathrm{N}(0, \Sigma)$, exponential correlation function
- ⋆ Forward function is deterministic
  - – linear: smoothing for ten nodes
  - – (non-linear example)
- ⋆ Prior for $\theta = (\mu, \Sigma)$: vague, the same for all time steps
- ⋆ Likelihood, $y_n | x_n \sim \mathrm{N}(x_n, 20\mathbb{I})$

$t = 15$



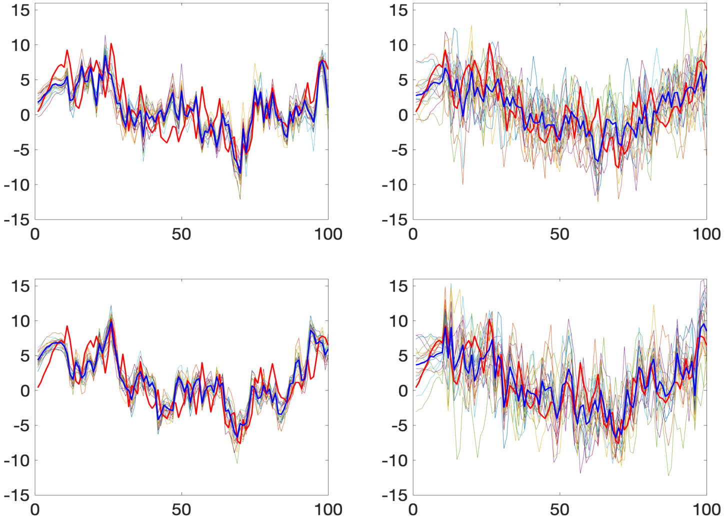
Results with $M = 19$, linear example, time $t = 1$



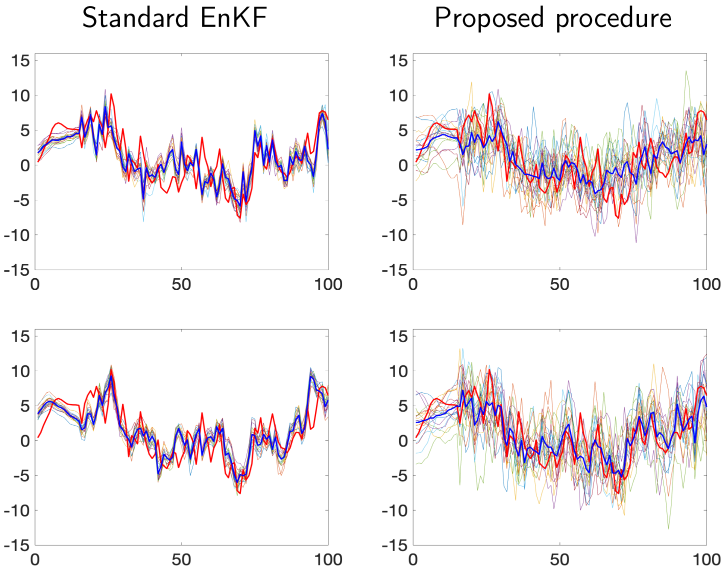
Results with $M = 19$, linear example, time $t = 2$

Standard EnKF

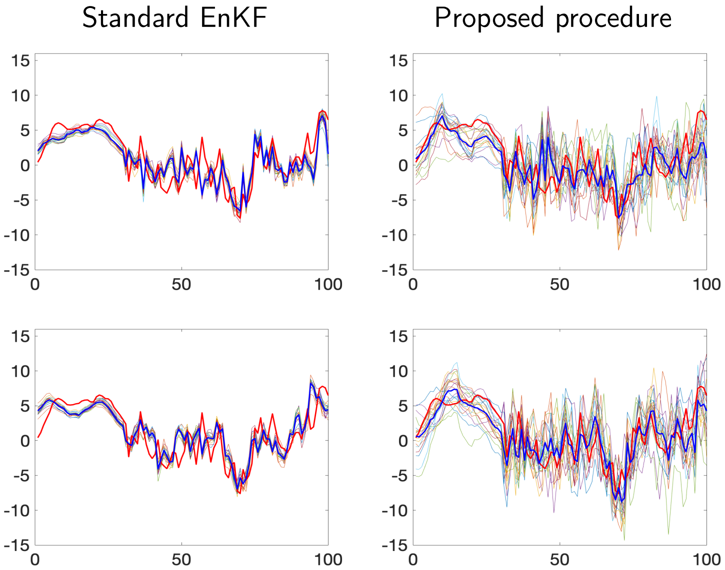
Proposed procedure



Results with $M = 19$, linear example, time $t = 3$

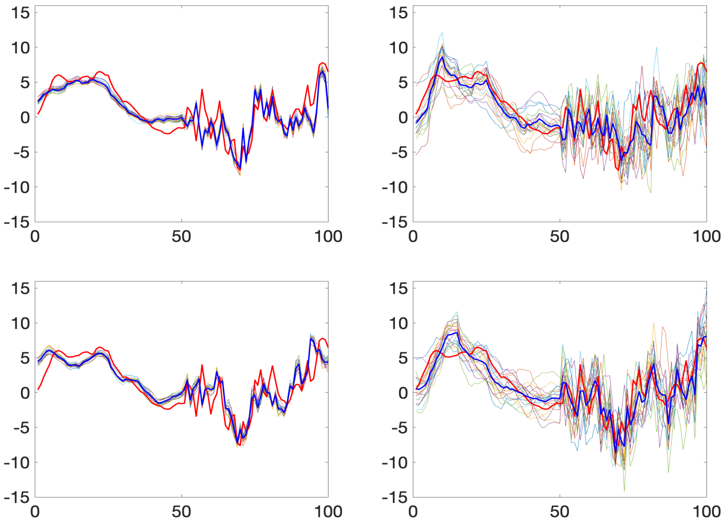


Results with $M = 19$, linear example, time $t = 6$

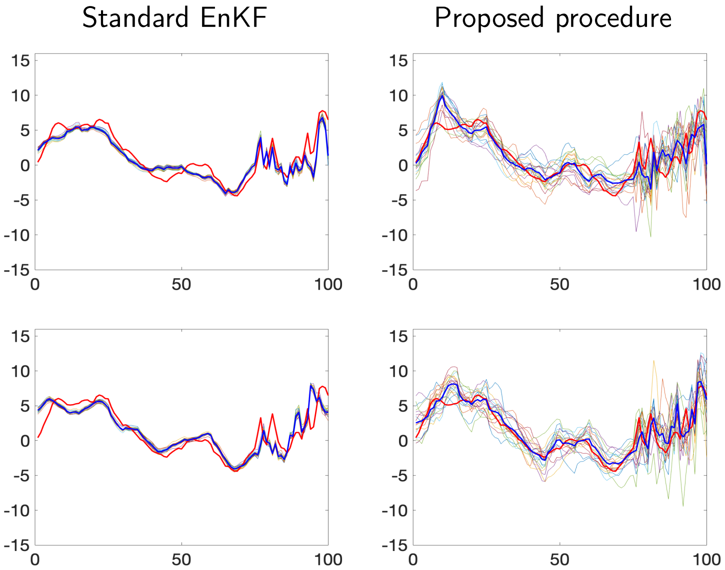


Results with $M = 19$, linear example, time $t = 10$

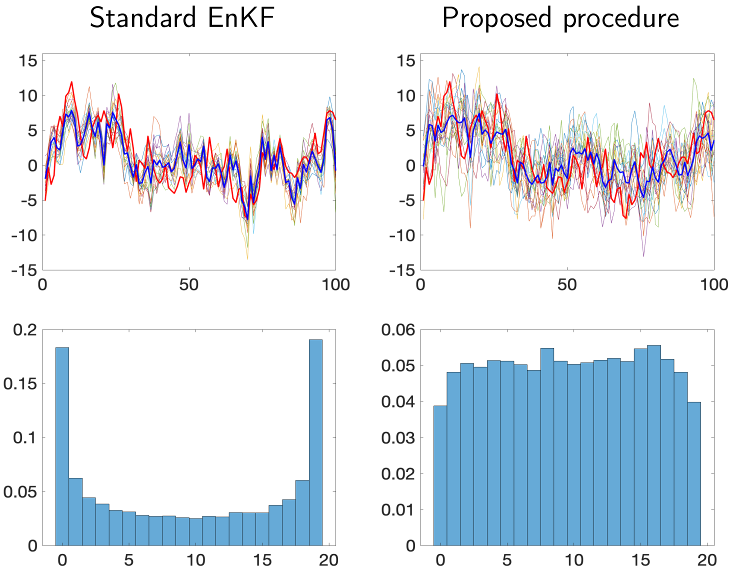
Standard EnKF | Proposed procedure



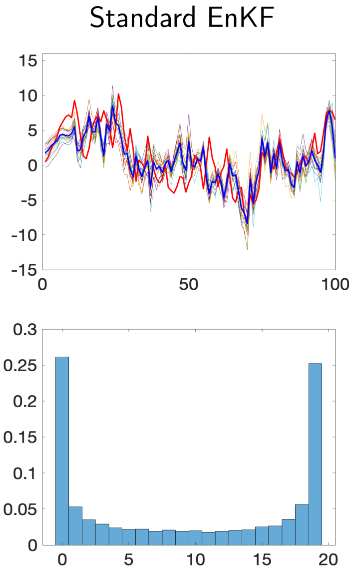
Results with $M = 19$, linear example, time $t = 15$



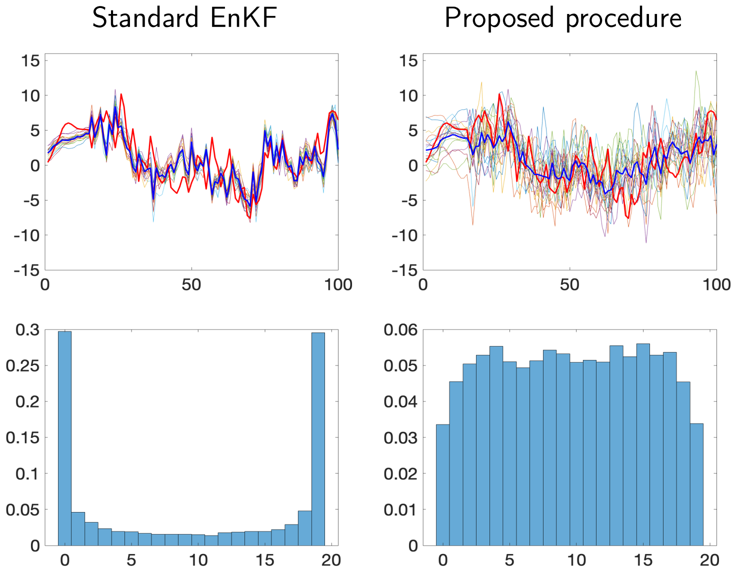
Results with $M = 19$, linear example, time $t = 1$



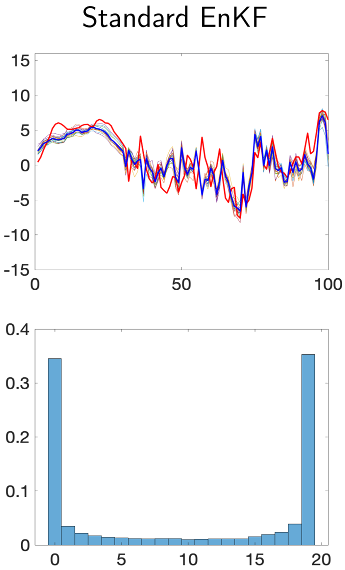
Results with $M = 19$, linear example, time $t = 2$



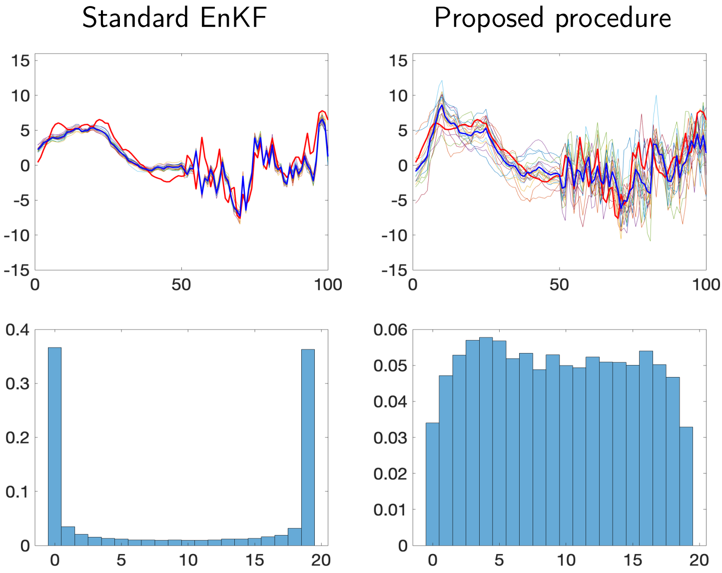
Results with $M = 19$, linear example, time $t = 3$



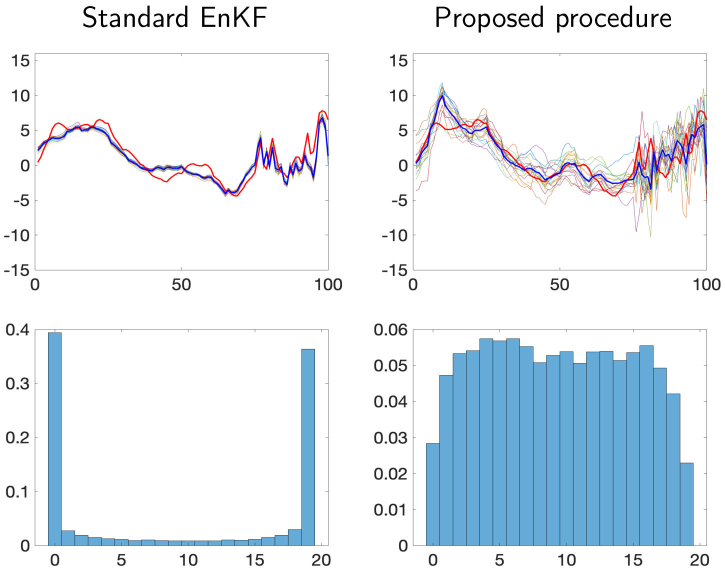
Results with $M = 19$, linear example, time $t = 6$



Results with $M = 19$, linear example, time $t = 10$



Results with $M = 19$, linear example, time $t = 15$



Closing remarks

- ⋆ Proposed procedure avoids variance underestimation
  - no need for variance inflation
- ⋆ Three changes in proposed update procedure
  - sample $\theta = (\mu, \Sigma)$
  - condition also on the new data y
  - do not condition on $\widetilde{x}^m$
- ⋆ Need more experience for non-linear models
- ⋆ The presented solution use matrices of full rank
  - not computationally feasible for high dimensional problems
  - ongoing work: formulate a sparse matrix variant of the proposed solution
- ⋆ We have also worked on generalising the idea underlying EnKF to a situation with categorical variables

References to our work

- ⋆ Loe, M.K. and Tjelmeland, H. (2021). A generalised and fully Bayesian framework for ensemble updating, `https://arxiv.org/abs/2103.14565`.
- ⋆ Loe, M.K. and Tjelmeland, H. (2020). Ensemble updating of binary state vectors by maximizing the expected number of unchanged components, *Scandinavian Journal of Statistics*, To appear, `https://doi.org/10.1111/sjos.12483`
- ⋆ Loe, M.K., Grana, D. and Tjelmeland, H. (2021). Geophysics-based fluid-facies predictions using ensemble updating of binary state vectors, *Mathematical Geosciences*, **53**, 325–347, `https://doi.org/10.1007/s11004-021-09922-4`
- ⋆ Loe, M.K. (2021). Ensemble updating for a state-space model with categorical variables, PhD thesis, NTNU, To appear.