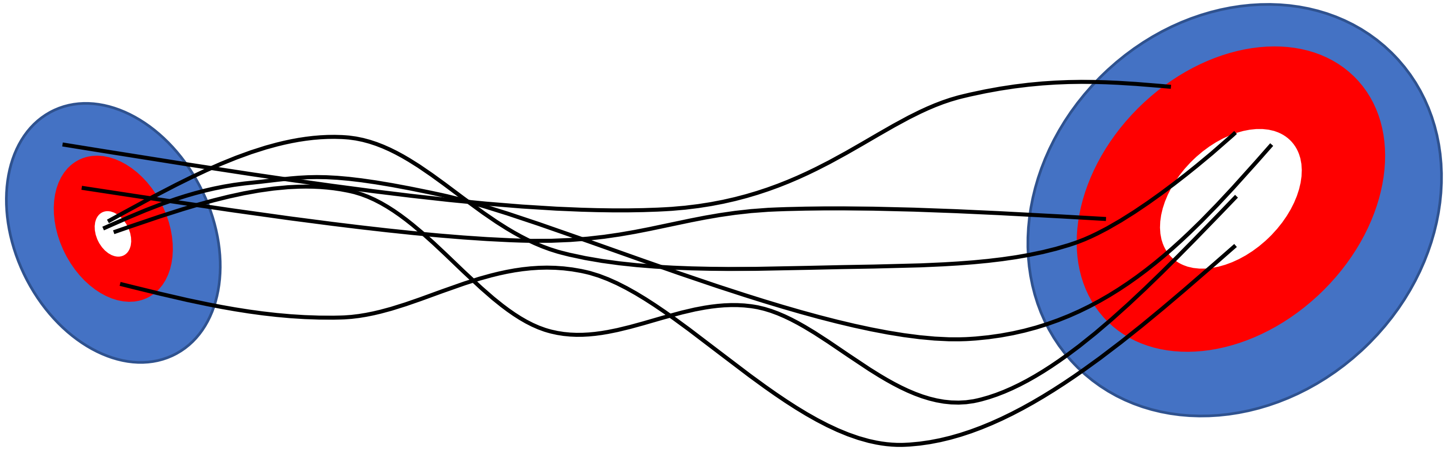
The Ensemble Kalman filter



Part II: Practicalities

Ivo Pasmans (Based on notes by Alison Fowler and Ross Bannister)

Data-assimilation training course. 7th-10th May, University of Reading

How large should the ensemble be?

- Lyapunov exponents are the ``averages'' of singular values of tangent linear model.

$$
\begin{aligned}
M_{t+1\leftarrow t}(\mathbf{x}_t) &\approx M_{t+1\leftarrow t}(\mathbf{x}_t^b) + \mathbf{M}_{t+1\leftarrow t}(\mathbf{x}_t - \mathbf{x}_t^b) \\
\mathbf{M}_{t+1\leftarrow t} &= \mathbf{U}_t \Lambda_t \mathbf{V}_t^{\mathrm{T}} \\
\lambda_i &= \limsup_{t\to\infty} \frac{1}{t} \log(\Lambda_t)_{ii}
\end{aligned}
$$

perturbation

- $\lambda_i < 0$ are associated with stable modes and will disappear over time.
- $\lambda_i \approx 0$ are associated (pseudo)-modes.
- $\lambda_i > 0$ are associated with unstable modes. These need to be constrained by observations.

- Without model error *N* > unstable modes, if model error are added more ensemble members are needed.

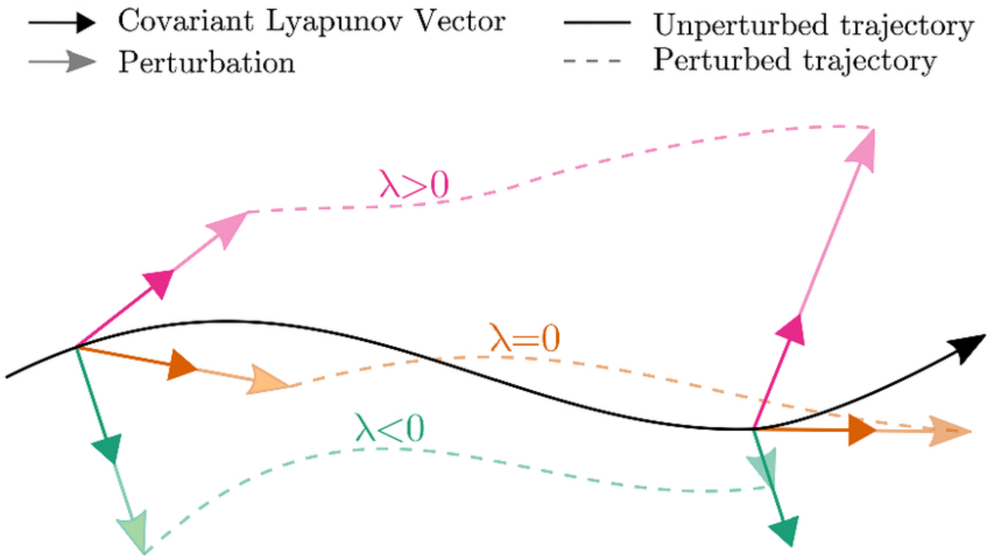


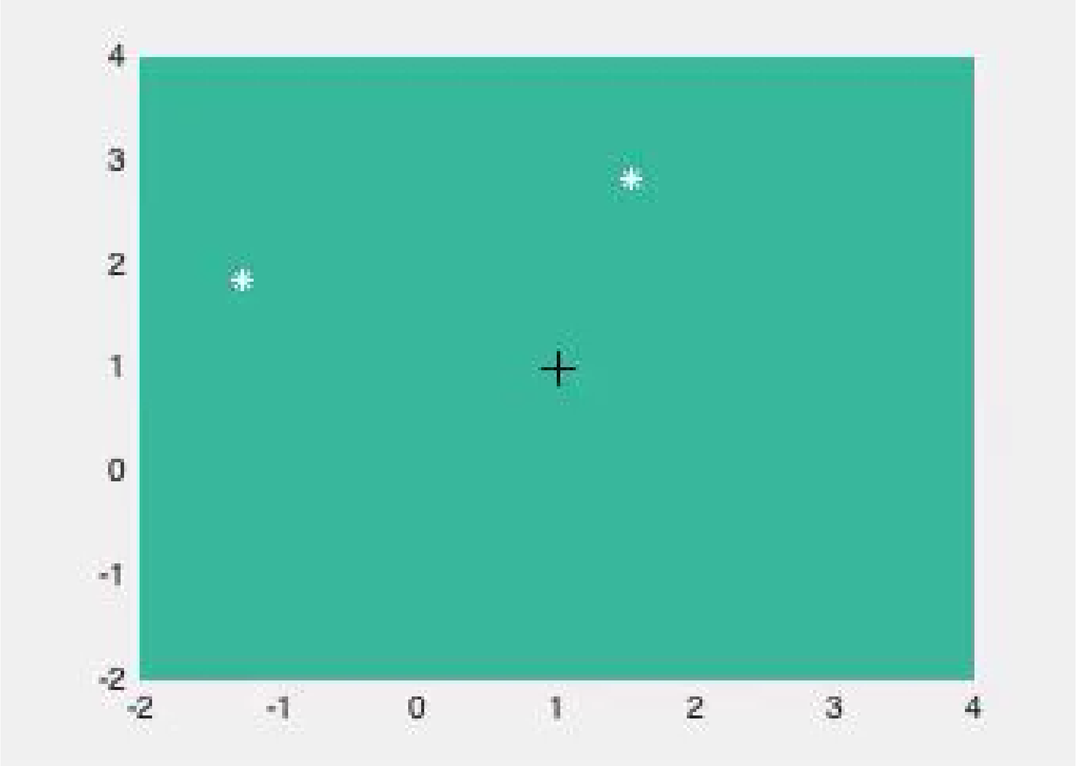
Image: Huhn, F., & Magri, L. (2020). Optimisation of chaotically perturbed acoustic limit cycles. *Nonlinear Dynamics*, *100*, 1641-1657.

How large should the ensemble be?

- Lyapunov spectrum is generally unknown.
- In practice run as large of an ensemble as practically possible.
- Use statistical tricks to compensate for limited ensemble size.
  - Inflation
  - Covariance localisation

Limited ensemble size creates sampling error

- The ensemble Kalman Filter theory assumes that the ensemble is large enough to give an accurate estimate of the sample mean and covariance, $\bar{\mathbf{x}}$ and $\mathbf{P}$.
- Even for a two variable model a large sample size is needed to accurately estimate the mean and covariance:



Example:
True distribution $\mathbf{x} \sim N([1, 1], \mathbf{I}_2)$

Consequence of sampling error

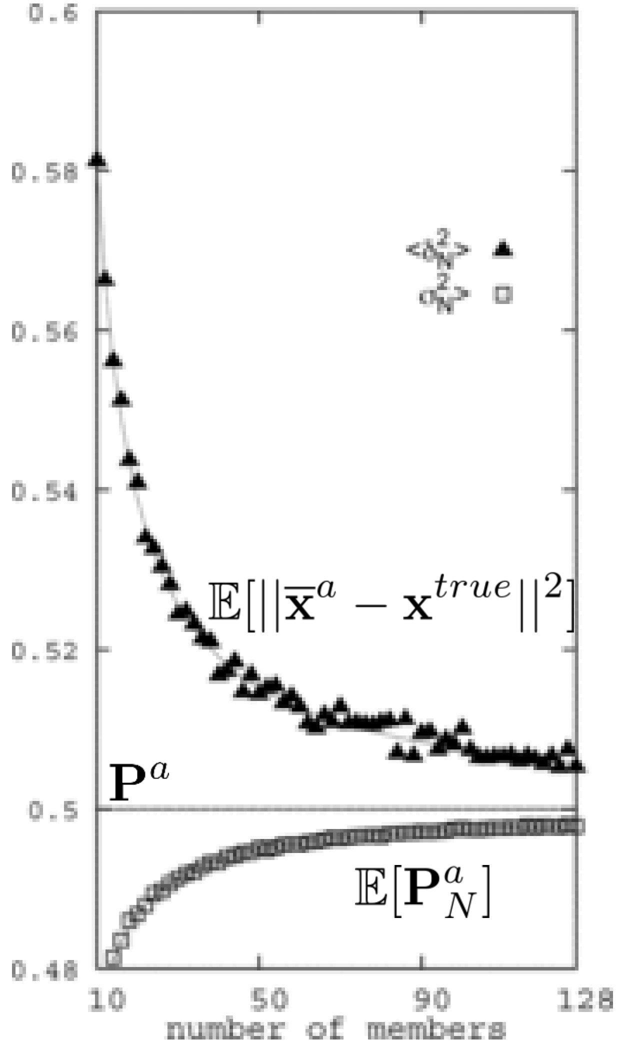
- Analysis covariance depends nonlinear on forecast covariance

$$\mathbf{P}^a = (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{P}^f = \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}}+\mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f$$

- Inserting erroneous forecast covariance $\mathbf{P}^f_N = \mathbf{P}^f + \epsilon,\ \frac{||\epsilon||}{||\mathbf{P}^f||} \ll 1$ it can be shown that

$$\begin{aligned}
(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}}+\mathbf{R}) &\overset{def}{=} \mathbf{G} \\
(\mathbf{H}\mathbf{P}^f_N\mathbf{H}^{\mathrm{T}}+\mathbf{R})^{-1} &= (\mathbf{G}-\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}})^{-1} = (\mathbf{I}-\mathbf{G}^{-1}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{G}^{-1} \\
&= \mathbf{G}^{-1} + \mathbf{G}^{-1}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1} + \mathbf{G}^{-1}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1} + \mathcal{O}(||\epsilon||^3/||\mathbf{P}^f||^3) \\
(\mathbf{I}-\mathbf{K}_N\mathbf{H})\mathbf{P}^f_N &= (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{P}^f + \mathcal{O}(\epsilon) - \epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{H}\epsilon - \mathbf{K}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}(\mathbf{P}^f)^{-1}\epsilon - [\mathbf{K}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}(\mathbf{P}^f)^{-1}\epsilon]^{\mathrm{T}} \\
&- \mathbf{K}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{H}\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}} + \mathcal{O}(||\epsilon||^3/||\mathbf{P}^f||^3) \\
\mathbb{E}[\mathbf{P}^a_N] &= \mathbf{P}^a - (\mathbf{I}+\mathbf{K}\mathbf{H})\mathbb{E}[\epsilon\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{H}\epsilon](\mathbf{I}+\mathbf{K}\mathbf{H})^{\mathrm{T}} + \mathcal{O}(||\epsilon||^3/||\mathbf{P}^f||^3)
\end{aligned}$$

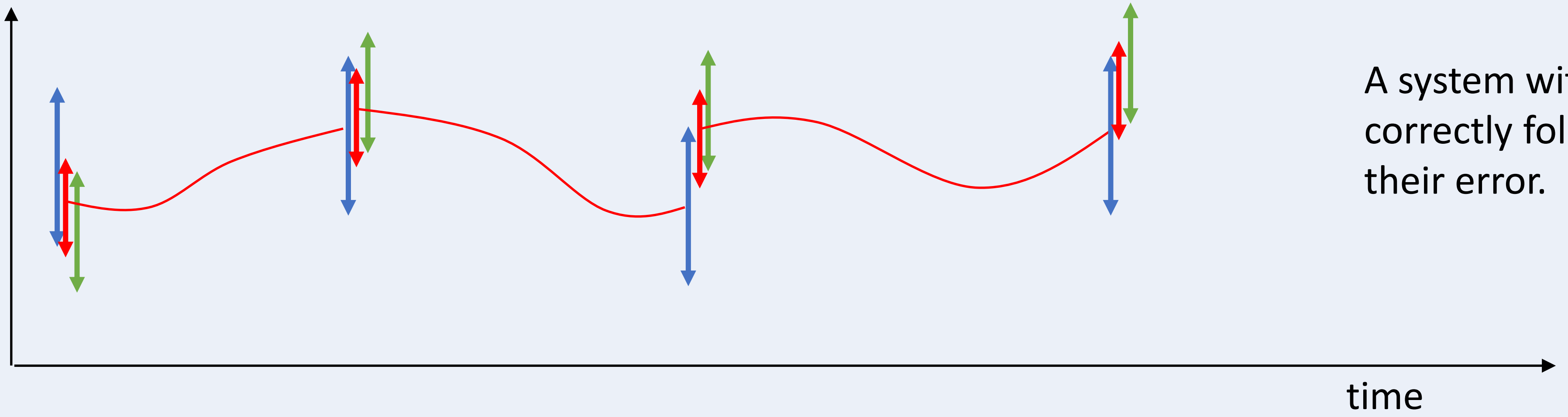
Variance bias is function of ensemble size



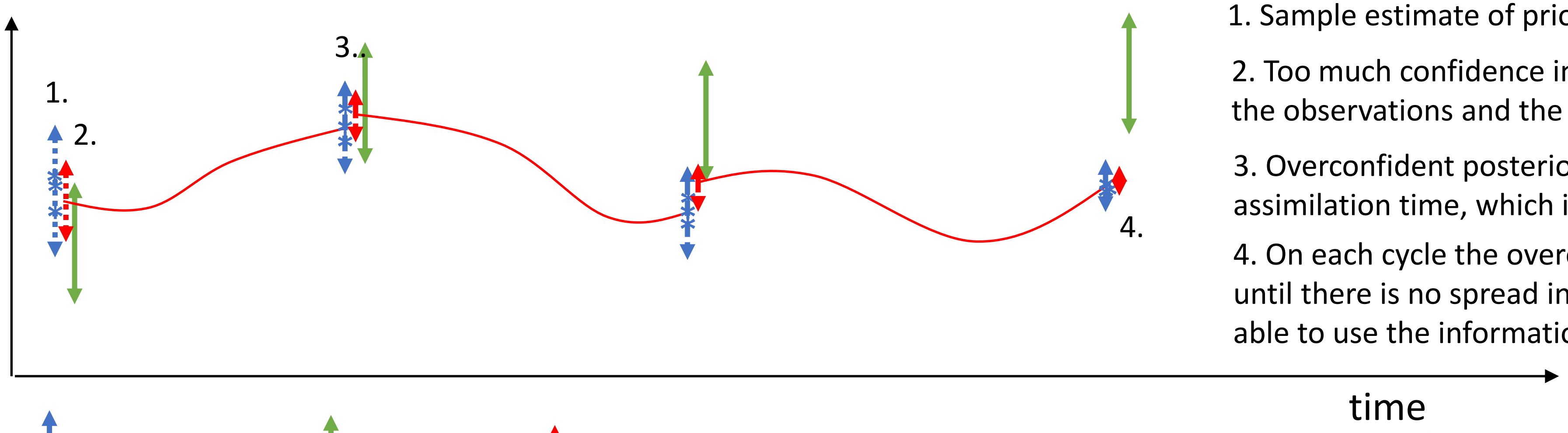
- In limit $N \to \infty$ the mean squared error and ensemble variance match on average.
- For small ensembles, the ensemble spread underestimates the analysis uncertainty.
- Smaller ensembles result in less error removal.

Image: Sacher, W., & Bartello, P. (2008). Sampling errors in ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, *136*(8), 3035-3049.

Sampling error results in filter divergence



A system without sampling error can correctly follow the observations within their error.



1. Sample estimate of prior spread is too small

2. Too much confidence in the prior means the analysis underfits the observations and the posterior spread is too small

3. Overconfident posterior leads to an overconfident prior at the next assimilation time, which is exasperated further by under sampling.

4. On each cycle the overconfidence is propagated and worsened until there is no spread in the prior and the analysis is no longer able to use the information in the observations.

prior spread | likelihood spread | posterior spread | * Sample from prior | Sample est. of prior spread | Sample est. of posterior spread

Methods to mitigate effect sampling error

- ETKF-N (Bocquet & Sakov, 2012).
- **Relaxation**
- **Inflation**

Bocquet, M., & Sakov, P. (2012). Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlinear Processes in Geophysics*, *19*(3), 383-399.

Relaxation

- Relaxation to prior perturbation (Zhang & Sun, 2004):

$$\mathbf{x}'^{a,(n)} \leftarrow (1-\alpha)\mathbf{x}'^{a,(n)} + \alpha\mathbf{x}'^{f,(n)} \text{ with } \mathbf{x}'^{(n)} = \mathbf{x}^{(n)} - \frac{1}{N}\sum_{m=1}^{N}\mathbf{x}^{(m)}$$

  - Does not work with rotation in ETKF as there is not a one-to-one relation between forecast and analysis members.
  - When using $\alpha = \frac{1}{2}$ the stochastic EnKF can be used without stochastic error perturbations. This is known as DEnKF (Sakov & Oke, 2008).

- Relaxation to prior spread (Whitaker & Hamill, 2012):

$$\mathbf{x}_i'^{a,(n)} \leftarrow \mathbf{x}_i'^{a,(n)}\left(\alpha\frac{\sigma_i^b - \sigma_i^a}{\sigma_i^a} + 1\right)$$

Zhang F, Snyder C, Sun J. 2004. Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Weather Rev.* 132: 1238–1253.
Sakov, P., & Oke, P. R. (2008). A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A: Dynamic Meteorology and Oceanography*, *60*(2), 361-371.
Whitaker, J. S., & Hamill, T. M. (2012). Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, *140*(9), 3078-3089.

Ensemble inflation

Ways to inflate

- Additive inflation (Mitchell & Houtekamer, 2000):
  - At each model time step add a random perturbation using similar ideas to representing model error given in the last lecture

$$\mathbf{x}^{f,(n)}(t+1) = \mathbf{x}^{f,(n)}(t) + \eta^{(n)}(t) \text{ with } \eta^{(n)}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$$

- Prior multiplicative inflation (Anderson & Anderson, 1999):

$$\mathbf{P}_N^f \leftarrow (1+\alpha^2)\mathbf{P}_N^f \text{ with } \alpha > 0 \qquad \mathbf{X}'^f \leftarrow \sqrt{1+\alpha^2}\mathbf{X}'^f \text{ with } \alpha > 0$$

- Posteriori multiplicative inflation:

$$\mathbf{P}_N^a \leftarrow (1+\alpha^2)\mathbf{P}_N^a \text{ with } \alpha > 0 \qquad \mathbf{X}'^a \leftarrow \sqrt{1+\alpha^2}\mathbf{X}'^a \text{ with } \alpha > 0$$

Mitchell, H. L., & Houtekamer, P. L. (2000). An adaptive ensemble Kalman filter. *Monthly Weather Review*, *128*(2), 416-433.
Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly weather review*, *127*(12), 2741-2758.

Tuning the inflation factor: rank histogram

Method 1: rank histograms (Hamill, T., 2001):

- For the ensemble to be reliable it is assumed that it is sampling the same distribution as the truth.
- Rank histogram: histogram of

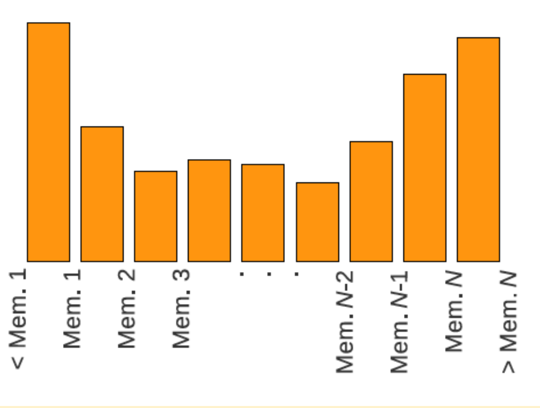
$$\text{rank}_i = \sum_{n=1}^{N} \text{heaviside}(\mathbf{y}_i + \xi_i - \mathbf{H}_i \mathbf{x}^{(n)}) \text{ with } \xi_i \sim \mathcal{N}(0, \mathbf{R}_{ii})$$

Example:
Observation: 2.45+0.06=2.51
Ensemble predictions: **1.23, 1.45, 2.32**, 2.56, 3.00
so rank is 3

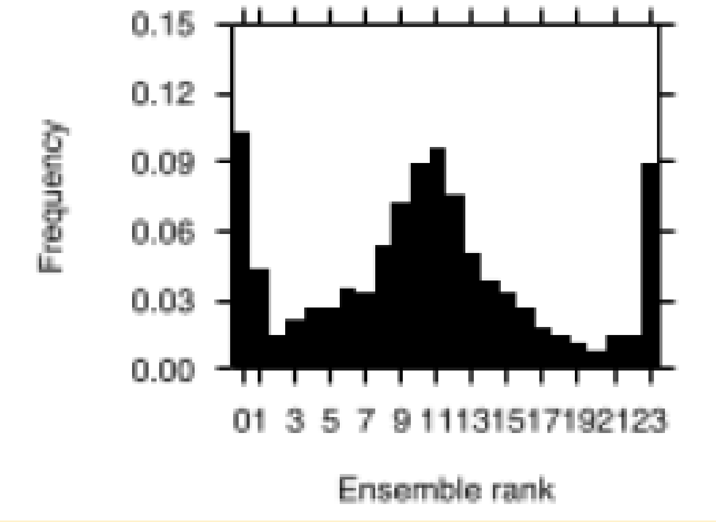
- Find factor α that makes rank histogram as uniform as possible.

Interpretation:

- Concave shape- the ensemble is under spread →
- Convex shaped- the ensemble is overspread
- Flat- the ensemble is correctly spread
- Asymmetric- the ensemble is biased



Rank histograms for surface precipitation rate rate. From Migliorini et al. (2011).



Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, *129*(3), 550-560.
Migliorini, S., Dixon, M., Bannister, R., & Ballard, S. (2011). Ensemble prediction for nowcasting with a convection-permitting model—I: description of the system and the impact of radar-derived surface precipitation rates. *Tellus A: Dynamic Meteorology and Oceanography*, *63*(3), 468-496.

Tuning the inflation factor: covariance matching

Method 2: Covariance matching

- Desroziers relations

$$\begin{aligned}\mathbb{E}[\mathbf{d}^{ob}(\mathbf{d}^{ob})^{\mathrm{T}}] &= \mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}\\ \mathbb{E}[\mathbf{d}^{ab}(\mathbf{d}^{ob})^{\mathrm{T}}] &= \mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}}\end{aligned}$$

$$\begin{aligned}\mathbf{d}^{ob} &= \mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^f\\ \mathbf{d}^{ab} &= \mathbf{H}\overline{\mathbf{x}}^a - \mathbf{H}\overline{\mathbf{x}}^f\\ \mathbf{d}^{oa} &= \mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^a\end{aligned}$$

- Inserting $\mathbf{P}^f \approx (1+\alpha^2)\mathbf{P}^f_N = \frac{1+\alpha^2}{N-1}\sum_{n=1}^{N}\mathbf{x}'^{f,(n)}(\mathbf{x}'^{f,(n)})^{\mathrm{T}}$

  and taking the trace gives that the inflation factor can be found by solving

$$(1+\alpha^2)\sum_i[(\mathbf{H}\mathbf{P}^f_N\mathbf{H}^{\mathrm{T}})_{ii} + \mathbf{R}_{ii}] = \sum_i \mathbb{E}[\mathbf{d}^{ob}(\mathbf{d}^{ob})^{\mathrm{T}}]_{ii} \approx \sum_i (\mathbf{d}^{ob}_i)^2$$

or

$$(1+\alpha^2)\sum_i(\mathbf{H}\mathbf{P}^f_N\mathbf{H}^{\mathrm{T}})_{ii} = \sum_i \mathbb{E}[\mathbf{d}^{ab}(\mathbf{d}^{ob})^{\mathrm{T}}]_{ii} \approx \sum_i \mathbf{d}^{ob}_i\mathbf{d}^{ab}_i$$

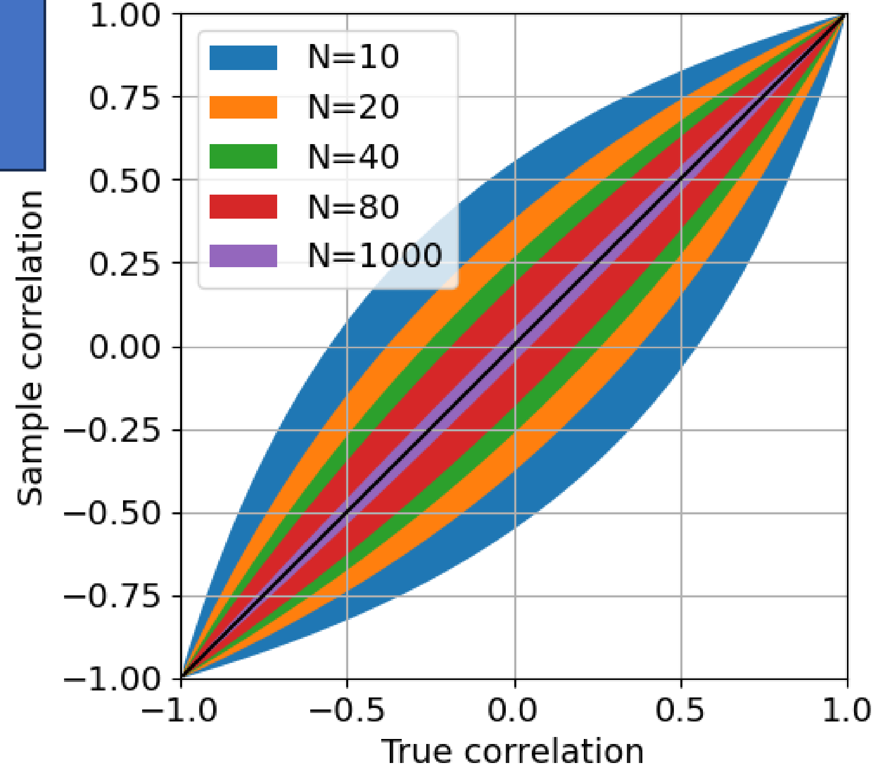
Kotsuki, S., Ota, Y., & Miyoshi, T. (2017). Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere. *Quarterly Journal of the Royal Meteorological Society*, *143*(705), 2001-2015.

Desroziers, G., Berre, L., Chapnik, B., & Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *131*(613), 3385-3396.

Spurious correlations

- $\mathbf{x}^a - \mathbf{x}^f = \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}$ shows that DA corrections lie in image $\mathbf{P}^f$. If $\mathbf{P}^f = \frac{1}{N-1}\mathbf{X}'^f(\mathbf{X}'^f)^{\mathrm{T}}$ the dimension of the image is *N-1*, limiting possible shape corrections.
- The **correlation will be subject to sampling error**. Implying that observations can influence regions and variables that they shouldn't.

90% confidence interval sample correlation



True covariance

Sample *N=16* covariance

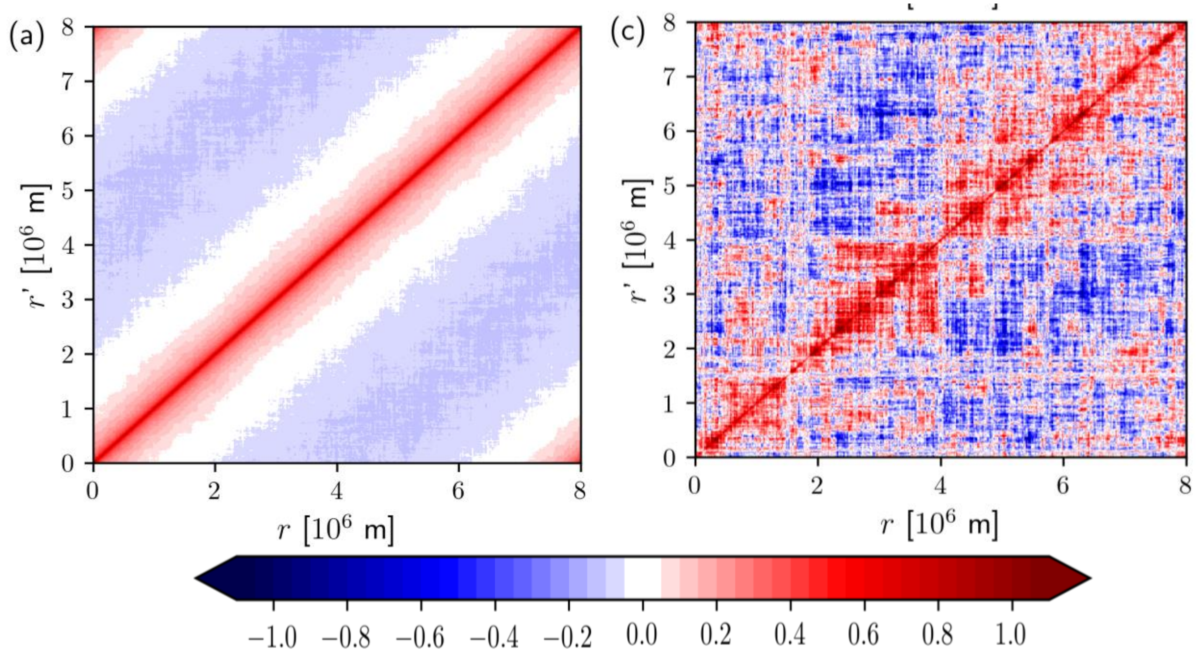


Image: Pasmans, I., Chen, Y., Carrassi, A., & Jones, C. K. (2023). Tailoring data assimilation to discontinuous Galerkin models. *arXiv preprint arXiv:2305.02950*.

Localisation

- Localisation removes spurious correlations and increases rank of the covariance.
- Different localisation techniques:
  - Scale-dependent localisation.
  - Adaptive covariance localisation.
    - Optimal localisation (Ménétrier & Auligné, 2015).
    - ECORAP (Bishop & Hodyss, 2009).
  - **Non-adaptive covariance localisation.**
  - **Domain localisation.**

Bishop, C., & Hodyss, D. (2009). Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus A: Dynamic Meteorology and Oceanography*, *61*(1), 84-96.
Ménétrier, B., Auligné, T. (2015). Optimized localization and hybridization to filter ensemble-based covariances. *Monthly Weather Review*, *143*(10), 3931-3947.

Covariance localisation

- Suppress long –distance correlations as non-physical.

$$\mathbf{P}^f = \mathbf{S}(\mathbf{S}^{-1}\mathbf{P}_N^f\mathbf{S}^{-1} \circ \mathbf{L})\mathbf{S}$$

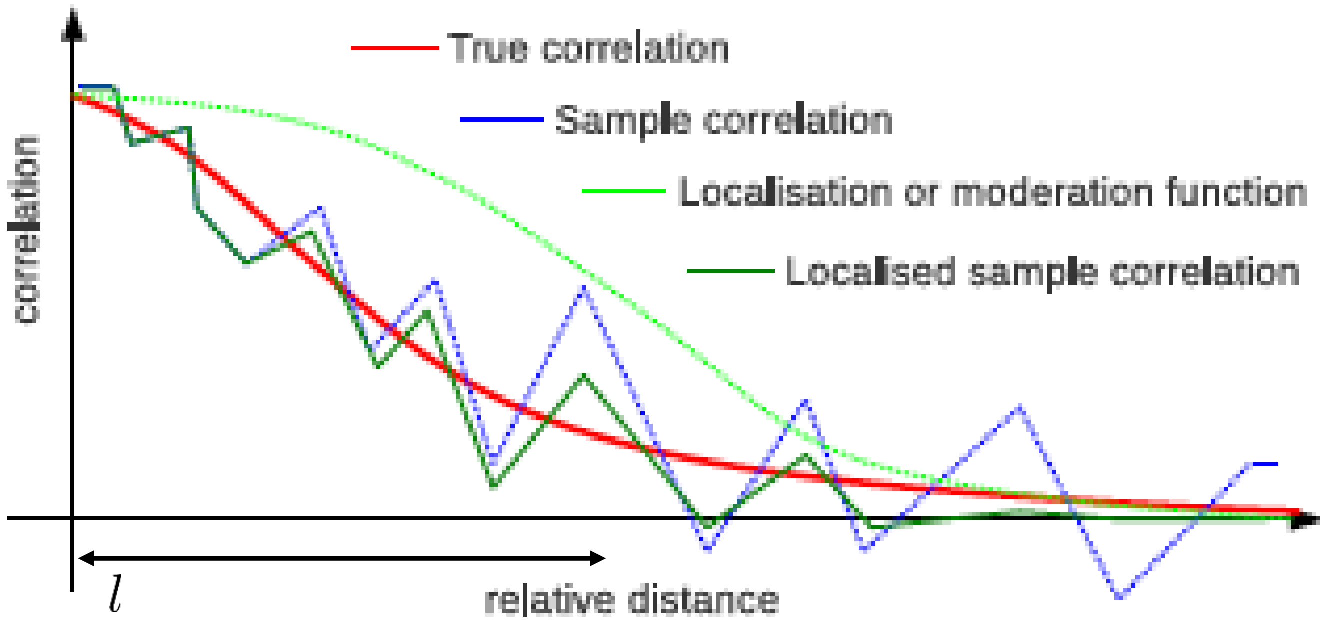
$$\mathbf{S}_{ij} = \frac{\delta_{ij}}{N-1}\sqrt{\sum_{n=1}^{N} \mathbf{x'}_i^{f,(n)}\mathbf{x'}_i^{f,(n)}} \text{ (standard deviation)}$$

$$((\mathbf{S}^{-1}\mathbf{P}_N^f\mathbf{S}^{-1}) \circ \mathbf{L})_{ij} = (\mathbf{S}^{-1}\mathbf{P}_N^f\mathbf{S}^{-1})_{ij}\mathbf{L}_{ij} \text{ (Schur product)}$$

- Entries of L far from diagonal tapper to zero, e.g.

$$\mathbf{L}_{ij} = e^{-\frac{||\mathbf{r}_i - \mathbf{r}_j||^2}{2l^2}}$$

- Pick localisation length scale *l* based on calibration or physical length scales (e.g. internal Rossby radius of deformation).

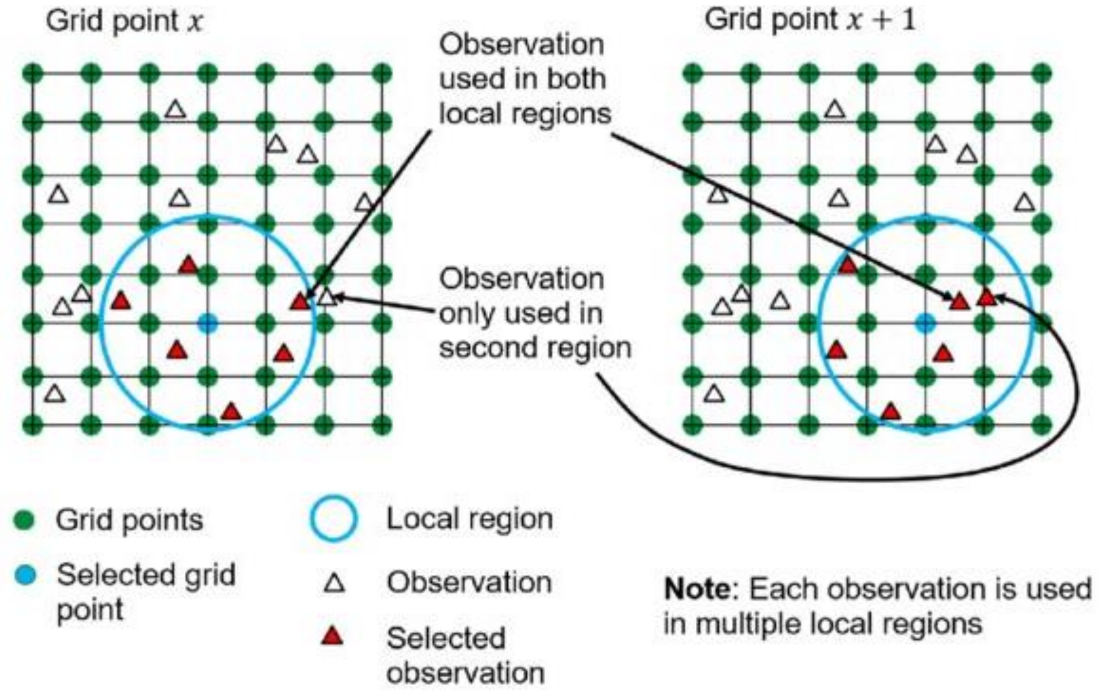


Domain localisation

- Domain localisation is used by the local ensemble transform Kalman filter (LETKF) (Hunt et al., 2007).
- LETKF procedure:

Select grid point → Select observations in vicinity → Calculate correction using ETKF → Apply correction to selected grid point only

- To avoid discontinuous DA corrections observational error variances are gradually inflated for observations far from selected grid point *i*, e.g. $\mathbf{R}_{jj} \leftarrow \mathbf{R}_{jj} e^{\frac{||\mathbf{r}_i - \mathbf{r}_j||^2}{2l^2}}$



Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1-2), 112-126.
Image: Elvidge, S., & Angling, M. J. (2019). Using the local ensemble transform Kalman filter for upper atmospheric modelling. *Journal of Space Weather and Space Climate*, *9*, A30.

Localisation comparison

	Covariance localisation	Domain localisation (LETKF)
Covariance modified	Background error covariance	Observation error covariance
Compatible solver	Iterative	Analytic
Suitable for observation types	All	Point observations
Parallelizable	Difficult	Trivial
Optimal length scale	Larger than in domain localisation	Shorter than in covariance localisation

Summary

- Ensemble data assimilation relies on a sample estimate of the mean and covariance of forecast distribution. This allows it to provide a flow-dependent estimate of the forecast uncertainty.
- If the ensemble size is much smaller than the size of the state then sampling error becomes an issue
  - Biases
  - Analysis increments lie in the subspace of the ensemble
  - Filter divergence
  - Spurious correlations
- To make ensemble DA practical need
  - Ensemble inflation
  - Localisation
  - …Hybrid methods

Further reading

- Anderson JL, Anderson SL. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev. 127: 2741–2758.
- Bannister et al. 2011: Ensemble prediction for nowcasting with a convection-permitting model - II: forecast error statistics, Tellus 63A, 497-512 .
- Corazza et al.. 2003. Use of the breeding technique to estimate the structure of the analysis 'errors of the day'. Nonlinear Processes Geophys. 10: 233–243.
- Greybush, 2011: Balance and ensemble Kalman filter localisation techniques. Mon. Wea. Rev., 139, 511-522.
- Hamill, 2001: Interpretation of Rank Histograms for Verifying Ensemble Forecasts. Mon. Wea. Rev., 129,550-560.
- Hamill, 2006: Ensemble-based atmospheric data assimilation. In *Predictability of Weather and Climate*, Palmer T, Hagedorn R (eds). Cambridge University Press: Cambridge; 124– 156.
- Houtekamer and Zhang (2016) Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation . *Mon. Weather Rev*., **144**, 4489–4532
- Houtekamer and Mitchell 1998: Data assimilation using an ensemble Kalman Filter technique, Mon. Wea. Rev. 126, 796-811.
- Houtekamer and Mitchell 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev., 129, 123–137
- Hunt et al. 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D, 230, 112–126,
- Kotsuki et al. 2017: Adaptive covariance relaxation methods for ensemble data assimilation: experiments in the real atmosphere. Q.J.R.Meteorol.Soc.143: 2001 – 2015.
- Migliorini et al. 2011: Ensemble prediction for nowcasting with a convection-permitting model - I: description of the system and the impact of radar-derived surface precipitation rates, Tellus 63A, 468-496.
- Mitchell and Houtekamer 2000. An adaptive ensemble Kalman filter. Mon. Weather Rev. 128: 416–433.
- Miyoshi et al. 2014: The 10,240-member ensemble Kalman filtering with an intermediate AGCM. Geophys. Res. Lett., 41, 5264-5271.
- Whitaker and Hamill 2012. Evaluating methods to account for system errors in ensemble data assimilation. Mon. Weather Rev. 140: 3078–3089.
- Zhang et al. 2004. Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. Mon. Weather Rev. 132: 1238–1253.