Lecture 11: Ensemble Square-Root Filters (ETKF and EAKF)

Introduction to Data Assimilation

November 20. 2025

Contents

1 Nonlinear Filtering, EKF, and EnKF: A Brief Recap

1.1 Canonical nonlinear filtering problem

We consider the discrete-time nonlinear state–space model

$$\boldsymbol{u}_{m+1} = \boldsymbol{f}(\boldsymbol{u}_m) + \boldsymbol{\sigma}_{m+1}, \tag{1}$$
$$\boldsymbol{v}_{m+1} = \boldsymbol{g}(\boldsymbol{u}_{m+1}) + \boldsymbol{\sigma}^o_{m+1}, \tag{2}$$

where

- $\boldsymbol{u}_m \in \mathbb{R}^N$ is the state (signal),
- $\boldsymbol{v}_m \in \mathbb{R}^M$ is the observation,
- $\boldsymbol{f}$ is a (possibly highly) nonlinear model map,
- $\boldsymbol{g}$ is a (possibly nonlinear) observation operator.

We assume

$$\mathbb{E}[\boldsymbol{\sigma}_m] = 0, \quad \mathbb{E}[\boldsymbol{\sigma}_m \boldsymbol{\sigma}_m^\top] = \boldsymbol{R}, \quad \mathbb{E}[\boldsymbol{\sigma}_m^o] = 0, \quad \mathbb{E}[\boldsymbol{\sigma}_m^o (\boldsymbol{\sigma}_m^o)^\top] = \boldsymbol{R}^o.$$

The filtering problem is to estimate the posterior mean and covariance of $\boldsymbol{u}_{m+1}$ given all observations up to time $m+1$,

$$\mathcal{F}_{m+1} := \{\boldsymbol{v}_j : j \le m+1\}.$$

1.2 Extended Kalman filter (EKF) recap

The extended Kalman filter approximates the nonlinear operators by their first-order Taylor expansions about the current mean states. Assume $\boldsymbol{f} \in C^1(\mathbb{R}^N)$, $\boldsymbol{g} \in C^1(\mathbb{R}^N)$. Linearize as

$$\boldsymbol{f}(\boldsymbol{u}) \approx \boldsymbol{f}(\bar{\boldsymbol{u}}_{m\,|\,m}) + \boldsymbol{F}_m(\boldsymbol{u} - \bar{\boldsymbol{u}}_{m\,|\,m}), \qquad \boldsymbol{F}_m = \nabla \boldsymbol{f}(\boldsymbol{u})\big|_{\bar{\boldsymbol{u}}_{m\,|\,m}}, \tag{3}$$

$$\boldsymbol{g}(\boldsymbol{u}) \approx \boldsymbol{g}(\bar{\boldsymbol{u}}_{m+1\,|\,m}) + \boldsymbol{G}_m(\boldsymbol{u} - \bar{\boldsymbol{u}}_{m+1\,|\,m}), \qquad \boldsymbol{G}_m = \nabla \boldsymbol{g}(\boldsymbol{u})\big|_{\bar{\boldsymbol{u}}_{m+1\,|\,m}}. \tag{4}$$

Substituting into (1)–(2) gives the linearized system

$$\boldsymbol{u}_{m+1} = \boldsymbol{f}(\bar{\boldsymbol{u}}_{m\,|\,m}) + \boldsymbol{F}_m(\boldsymbol{u}_m - \bar{\boldsymbol{u}}_{m\,|\,m}) + \boldsymbol{\sigma}_{m+1}, \tag{5}$$

$$\boldsymbol{v}_{m+1} = \boldsymbol{g}(\bar{\boldsymbol{u}}_{m+1\,|\,m}) + \boldsymbol{G}_m(\boldsymbol{u}_{m+1} - \bar{\boldsymbol{u}}_{m+1\,|\,m}) + \boldsymbol{\sigma}^o_{m+1}. \tag{6}$$

Forecast step. Define posterior mean and covariance at time m as $\bar{\boldsymbol{u}}_{m\,|\,m}$, $\boldsymbol{R}_{m\,|\,m}$. Then

$$\bar{\boldsymbol{u}}_{m+1\,|\,m} = \boldsymbol{f}(\bar{\boldsymbol{u}}_{m\,|\,m}), \tag{7}$$

$$\boldsymbol{R}_{m+1\,|\,m} = \boldsymbol{F}_m \, \boldsymbol{R}_{m\,|\,m} \, \boldsymbol{F}_m^\top + \boldsymbol{R}. \tag{8}$$

Analysis step. The EKF is obtained by minimizing the quadratic cost functional

$$J(\boldsymbol{u}) = \|\boldsymbol{u} - \bar{\boldsymbol{u}}_{m+1\,|\,m}\|^2_{\boldsymbol{R}^{-1}_{m+1\,|\,m}} + \|\boldsymbol{v}_{m+1} - \boldsymbol{g}(\boldsymbol{u})\|^2_{(\boldsymbol{R}^o)^{-1}} \tag{9}$$

$$\approx \|\boldsymbol{u} - \bar{\boldsymbol{u}}_{m+1\,|\,m}\|^2_{\boldsymbol{R}^{-1}_{m+1\,|\,m}} + \|\boldsymbol{v}_{m+1} - \boldsymbol{g}(\bar{\boldsymbol{u}}_{m+1\,|\,m}) - \boldsymbol{G}_m(\boldsymbol{u} - \bar{\boldsymbol{u}}_{m+1\,|\,m})\|^2_{(\boldsymbol{R}^o)^{-1}}. \tag{10}$$

The minimizer satisfies

$$\bar{\boldsymbol{u}}_{m+1\,|\,m+1} = \bar{\boldsymbol{u}}_{m+1\,|\,m} + \boldsymbol{K}_{m+1}(\boldsymbol{v}_{m+1} - \boldsymbol{g}(\bar{\boldsymbol{u}}_{m+1\,|\,m})), \tag{11}$$

$$\boldsymbol{R}_{m+1\,|\,m+1} = (\boldsymbol{I} - \boldsymbol{K}_{m+1}\boldsymbol{G}_m)\, \boldsymbol{R}_{m+1\,|\,m}, \tag{12}$$

with Kalman gain

$$\boldsymbol{K}_{m+1} = \big(\boldsymbol{R}^{-1}_{m+1\,|\,m} + \boldsymbol{G}_m^\top (\boldsymbol{R}^o)^{-1} \boldsymbol{G}_m\big)^{-1} \boldsymbol{G}_m^\top (\boldsymbol{R}^o)^{-1} \tag{13}$$

$$= \boldsymbol{R}_{m+1\,|\,m}\, \boldsymbol{G}_m^\top \big(\boldsymbol{G}_m \boldsymbol{R}_{m+1\,|\,m} \boldsymbol{G}_m^\top + \boldsymbol{R}^o\big)^{-1}. \tag{14}$$

1.3 Ensemble Kalman filter (EnKF) recap

The EnKF replaces the exact covariances by ensemble covariances. Let $\{\boldsymbol{u}_{m+1\,|\,m}^{(k)}\}_{k=1}^{K}$ be the forecast ensemble at time $m+1$, and define the forecast ensemble mean and perturbation matrix

$$\bar{\boldsymbol{u}}_{m+1\,|\,m} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{u}_{m+1\,|\,m}^{(k)}, \tag{15}$$

$$\boldsymbol{U}_{m+1\,|\,m} = \left[\boldsymbol{u}_{m+1\,|\,m}^{(1)} - \bar{\boldsymbol{u}}_{m+1\,|\,m}, \ldots, \boldsymbol{u}_{m+1\,|\,m}^{(K)} - \bar{\boldsymbol{u}}_{m+1\,|\,m}\right] \in \mathbb{R}^{N\times K}. \tag{16}$$

The ensemble covariance approximation is

$$\boldsymbol{R}_{m+1\,|\,m} \approx \frac{1}{K-1}\boldsymbol{U}_{m+1\,|\,m}\boldsymbol{U}_{m+1\,|\,m}^{\top}. \tag{17}$$

For linear $\boldsymbol{G}$, the Kalman gain can be written as

$$\boldsymbol{K}_{m+1} = \frac{1}{K-1}\boldsymbol{U}\,\boldsymbol{V}^{\top}\left(\frac{1}{K-1}\boldsymbol{V}\boldsymbol{V}^{\top} + \boldsymbol{R}^{o}\right)^{-1}, \tag{18}$$

where

$$\boldsymbol{V} = \left[\boldsymbol{g}(\boldsymbol{u}_{m+1\,|\,m}^{(1)}) - \bar{\boldsymbol{v}}, \ldots, \boldsymbol{g}(\boldsymbol{u}_{m+1\,|\,m}^{(K)}) - \bar{\boldsymbol{v}}\right] \in \mathbb{R}^{M\times K}, \tag{19}$$

$$\bar{\boldsymbol{v}} = \boldsymbol{g}(\bar{\boldsymbol{u}}_{m+1\,|\,m}). \tag{20}$$

In the classical stochastic EnKF, the analysis ensemble is updated as

$$\boldsymbol{u}_{m+1\,|\,m+1}^{(k)} = \boldsymbol{u}_{m+1\,|\,m}^{(k)} + \boldsymbol{K}_{m+1}\big(\boldsymbol{v}_{m+1}^{(k)} - \boldsymbol{g}(\boldsymbol{u}_{m+1\,|\,m}^{(k)})\big), \tag{21}$$

$$\boldsymbol{v}_{m+1}^{(k)} = \boldsymbol{v}_{m+1} + \boldsymbol{\eta}^{(k)}, \qquad \boldsymbol{\eta}^{(k)} \sim \mathcal{N}(0, \boldsymbol{R}^{o}). \tag{22}$$

The noise perturbations ensure the analysis ensemble covariance matches the Kalman posterior covariance in the large-K limit, but are also a source of sampling error for small ensembles; this motivates deterministic *ensemble square-root* filters.

2 Ensemble Square-Root Filters (EnSRF)

The idea of ensemble square-root filters is to:

1. Update the *mean* by the Kalman formula (no observation perturbations);
2. Transform the forecast ensemble perturbations deterministically so that the resulting analysis perturbations have the correct posterior covariance.

We denote the forecast perturbation matrix at time $m+1$ by

$$\boldsymbol{U} := \boldsymbol{U}_{m+1\,|\,m} \in \mathbb{R}^{N\times K}, \qquad \bar{\boldsymbol{u}}_{f} := \bar{\boldsymbol{u}}_{m+1\,|\,m}, \qquad \bar{\boldsymbol{u}}_{a} := \bar{\boldsymbol{u}}_{m+1\,|\,m+1}.$$

We seek a transformation matrix $T \in \mathbb{R}^{K\times K}$ such that

$$\boldsymbol{U}_{a} = \boldsymbol{U}T, \qquad \frac{1}{K-1}\boldsymbol{U}_{a}\boldsymbol{U}_{a}^{\top} = \boldsymbol{R}_{m+1\,|\,m+1},$$

where $\boldsymbol{R}_{m+1\,|\,m+1}$ is the Kalman posterior covariance.

There are two main strategies covered in this lecture:

- the *ensemble transform Kalman filter* (ETKF), which works in the ensemble space ($K \times K$ matrices);
- the *ensemble adjustment Kalman filter* (EAKF), which works by transforming in state space ($N \times N$ matrices).

2.1 ETKF: Derivation in ensemble space

We start from the linear Kalman update with a linear observation operator $\boldsymbol{G}$:

$$\bar{\boldsymbol{u}}_a = \bar{\boldsymbol{u}}_f + \boldsymbol{K}_{m+1}(\boldsymbol{v}_{m+1} - \boldsymbol{G}\,\bar{\boldsymbol{u}}_f), \tag{23}$$
$$\boldsymbol{R}_a = (\boldsymbol{R}_f^{-1} + \boldsymbol{G}^\top(\boldsymbol{R}^o)^{-1}\boldsymbol{G})^{-1}, \tag{24}$$

where $\boldsymbol{R}_f = \boldsymbol{R}_{m+1\,|\,m}$, $\boldsymbol{R}_a = \boldsymbol{R}_{m+1\,|\,m+1}$. For the ensemble approximation, we write

$$\boldsymbol{R}_f \approx \frac{1}{K-1}\boldsymbol{U}\boldsymbol{U}^\top.$$

Let us define the projected anomalies

$$\boldsymbol{Y} := \boldsymbol{G}\boldsymbol{U} \in \mathbb{R}^{M\times K}.$$

Then the Kalman gain can be written in ensemble form as

$$\boldsymbol{K}_{m+1} = \boldsymbol{R}_f\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{R}_f\boldsymbol{G}^\top + \boldsymbol{R}^o)^{-1} \tag{25}$$
$$\approx \frac{1}{K-1}\boldsymbol{U}\boldsymbol{Y}^\top\Big(\frac{1}{K-1}\boldsymbol{Y}\boldsymbol{Y}^\top + \boldsymbol{R}^o\Big)^{-1}. \tag{26}$$

Posterior covariance in ensemble space. Using the identity

$$\boldsymbol{R}_a = \boldsymbol{R}_f - \boldsymbol{K}_{m+1}\,\boldsymbol{G}\,\boldsymbol{R}_f,$$

and substituting (26), we can write

$$\boldsymbol{R}_a \approx \frac{1}{K-1}\boldsymbol{U}\boldsymbol{U}^\top - \frac{1}{K-1}\boldsymbol{U}\boldsymbol{Y}^\top\Big(\frac{1}{K-1}\boldsymbol{Y}\boldsymbol{Y}^\top + \boldsymbol{R}^o\Big)^{-1}\boldsymbol{G}\,\frac{1}{K-1}\boldsymbol{U}\boldsymbol{U}^\top. \tag{27}$$

A standard matrix identity with $A = (K-1)^{-1/2}\boldsymbol{Y}$ gives

$$A^\top(AA^\top + \boldsymbol{R}^o)^{-1} = (\boldsymbol{I} + A^\top\boldsymbol{R}^{o-1}A)^{-1}A^\top\boldsymbol{R}^{o-1}.$$

After some algebra (mirroring the derivation in the book and writing everything in terms of $\boldsymbol{Y}$), one arrives at

$$\boldsymbol{R}_a = \frac{1}{K-1}\boldsymbol{U}\Big(\boldsymbol{I} + \boldsymbol{B}\Big)^{-1}\boldsymbol{U}^\top, \tag{28}$$
$$\boldsymbol{B} := \frac{1}{K-1}\,\boldsymbol{Y}^\top\boldsymbol{R}^{o-1}\boldsymbol{Y} \in \mathbb{R}^{K\times K}. \tag{29}$$

Thus the posterior covariance is a *quadratic form in ensemble space.*

Defining the transform matrix. We want $\boldsymbol{U}_a = \boldsymbol{U}T$ such that

$$\frac{1}{K-1}\boldsymbol{U}_a\boldsymbol{U}_a^\top = \frac{1}{K-1}\boldsymbol{U}TT^\top\boldsymbol{U}^\top = \frac{1}{K-1}\boldsymbol{U}(\boldsymbol{I}+\boldsymbol{B})^{-1}\boldsymbol{U}^\top.$$

Therefore we require

$$TT^\top = (\boldsymbol{I}+\boldsymbol{B})^{-1}. \tag{30}$$

Let

$$\boldsymbol{I}+\boldsymbol{B} = X\Lambda X^\top, \qquad X^\top X = XX^\top = \boldsymbol{I},$$

be an eigenvalue decomposition with diagonal $\Lambda = \mathrm{diag}(\lambda_i)$. Then one valid square-root is

$$T = X\Lambda^{-1/2}X^\top, \tag{31}$$

which is symmetric and is the choice known as the *symmetric ETKF*.

2.1.1 ETKF algorithm (Hunt-type efficient formulation)

Given:

- prior ensemble $\{\boldsymbol{u}_{m+1\,|\,m}^{(k)}\}_{k=1}^K$,
- observation $\boldsymbol{v}_{m+1}$,
- observation operator $\boldsymbol{g}$ (linear or nonlinear),
- observation error covariance $\boldsymbol{R}^o$,
- inflation factor $r \geq 0$.

1. **Compute forecast ensemble mean and anomalies.**

$$\bar{\boldsymbol{u}}_f = \frac{1}{K}\sum_{k=1}^K \boldsymbol{u}_{m+1\,|\,m}^{(k)}, \qquad \boldsymbol{U} = [\boldsymbol{u}_{m+1\,|\,m}^{(k)} - \bar{\boldsymbol{u}}_f].$$

2. **Form observation anomalies.**

$$\bar{\boldsymbol{v}}_f = \boldsymbol{g}(\bar{\boldsymbol{u}}_f), \qquad \boldsymbol{V} = [\boldsymbol{g}(\boldsymbol{u}_{m+1\,|\,m}^{(k)}) - \bar{\boldsymbol{v}}_f].$$

3. **Apply multiplicative inflation (optional).**

$$\boldsymbol{U} \leftarrow \sqrt{1+r}\,\boldsymbol{U}, \qquad \boldsymbol{V} \leftarrow \sqrt{1+r}\,\boldsymbol{V}.$$

4. **Compute the $K \times K$ matrix**

$$J = (K-1)\,\boldsymbol{I} + \boldsymbol{V}^\top \boldsymbol{R}^{o-1}\boldsymbol{V} \in \mathbb{R}^{K\times K}.$$

   Compute its eigenvalue decomposition

$$J = X\Lambda X^\top.$$

5. **Update the mean.** Solve the linear system

$$Jx = \boldsymbol{V}^\top \boldsymbol{R}^{o-1}(\boldsymbol{v}_{m+1} - \bar{\boldsymbol{v}}_f)$$

for $x \in \mathbb{R}^K$, and set

$$\bar{\boldsymbol{u}}_a = \bar{\boldsymbol{u}}_f + \boldsymbol{U}x.$$

6. **Compute the transform matrix.** Using (31),

$$T = X\Lambda^{-1/2}X^\top.$$

7. **Transform anomalies and form the analysis ensemble.**

$$\boldsymbol{U}_a = \boldsymbol{U}T.$$

The posterior ensemble is

$$\boldsymbol{u}^{(k)}_{m+1\,|\,m+1} = \bar{\boldsymbol{u}}_a + (\boldsymbol{U}_a)_{\cdot,k}, \qquad k = 1, \dots, K,$$

where $(\boldsymbol{U}_a)_{\cdot,k}$ denotes the k-th column of $\boldsymbol{U}_a$.

2.2 EAKF: Derivation via state-space adjustment

The ensemble adjustment Kalman filter constructs a state-space "adjustment" matrix $A \in \mathbb{R}^{N\times N}$ such that

$$\boldsymbol{U}_a = A\boldsymbol{U}, \tag{32}$$

$$\boldsymbol{R}_a = A\boldsymbol{R}_f A^\top. \tag{33}$$

Again take $\boldsymbol{R}_f \approx (K-1)^{-1}\boldsymbol{U}\boldsymbol{U}^\top$. For the linear observation operator $\boldsymbol{G}$, the Kalman posterior covariance is

$$\boldsymbol{R}_a = (\boldsymbol{R}_f^{-1} + \boldsymbol{G}^\top \boldsymbol{R}^{o-1}\boldsymbol{G})^{-1}.$$

Whitening the state covariance. Let the eigenvalue decomposition of $\boldsymbol{R}_f$ be

$$\boldsymbol{R}_f = F\Sigma^2 F^\top, \tag{34}$$

where $F \in \mathbb{R}^{N\times N}$ is orthonormal and Σ^2 is diagonal with positive entries. Then

$$\Sigma^{-1}F^\top \boldsymbol{R}_f F\Sigma^{-1} = \boldsymbol{I}.$$

Introduce the "whitened" coordinates

$$\boldsymbol{y} = \Sigma^{-1}F^\top \boldsymbol{u}.$$

In these coordinates, the prior covariance is the identity.

Diagonalizing the observation information. Consider the matrix

$$\Sigma F^\top \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{G} F \Sigma.$$

Let its eigenvalue decomposition be

$$X^\top \Sigma F^\top \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{G} F \Sigma X = D, \tag{35}$$

where X is orthonormal and D is diagonal (non-negative entries).

Combining (24), (34), and (35), one can show that

$$\boldsymbol{R}_a = F \Sigma X (\boldsymbol{I} + D)^{-1} X^\top \Sigma F^\top. \tag{36}$$

This suggests defining the adjustment matrix

$$A = F \Sigma X (\boldsymbol{I} + D)^{-1/2} \Sigma^{-1} F^\top. \tag{37}$$

Then

$$A \boldsymbol{R}_f A^\top = \boldsymbol{R}_a,$$

so using $\boldsymbol{U}_a = A\boldsymbol{U}$ gives the correct posterior covariance.

2.2.1 EAKF algorithm

Given:

- prior ensemble $\{\boldsymbol{u}^{(k)}_{m+1\,|\,m}\}_{k=1}^{K}$,
- observation $\boldsymbol{v}_{m+1}$,
- *linear* observation operator $\boldsymbol{G}$,
- observation error covariance $\boldsymbol{R}^o$,
- inflation factor $r \geq 0$.

1. **Compute forecast ensemble mean and anomalies.**

$$\bar{\boldsymbol{u}}_f = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{u}^{(k)}_{m+1\,|\,m}, \qquad \boldsymbol{U} = [\boldsymbol{u}^{(k)}_{m+1\,|\,m} - \bar{\boldsymbol{u}}_f].$$

2. **Approximate forecast covariance with inflation.**

$$\boldsymbol{U} \leftarrow \sqrt{1+r}\,\boldsymbol{U}, \qquad \boldsymbol{R}_f \approx \frac{1}{K-1} \boldsymbol{U}\boldsymbol{U}^\top.$$

3. **Eigenvalue decomposition of $\boldsymbol{R}_f$.** Compute

$$\boldsymbol{R}_f \approx F \Sigma^2 F^\top,$$

where F is orthonormal and Σ^2 diagonal.

4. **Eigenvalue decomposition of the observation information.** Compute

$$\Sigma F^\top \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{G} F \Sigma = X D X^\top,$$

with orthonormal X and diagonal D.

5. **Form the adjustment matrix.** Using (37),

$$A = F \Sigma X (\boldsymbol{I} + D)^{-1/2} \Sigma^{-1} F^\top.$$

6. **Transform ensemble anomalies.**

$$\boldsymbol{U}_a = A \boldsymbol{U}.$$

7. **Update the mean.** The Kalman posterior mean satisfies

$$\bar{\boldsymbol{u}}_a = (\boldsymbol{R}_f^{-1} + \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{G})^{-1} \big(\boldsymbol{R}_f^{-1} \bar{\boldsymbol{u}}_f + \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{v}_{m+1}\big).$$

In practice, one solves the linear system

$$L \bar{\boldsymbol{u}}_a = y, \tag{38}$$

where

$$L = \boldsymbol{R}_f^{-1} + \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{G}, \tag{39}$$

$$y = \boldsymbol{R}_f^{-1} \bar{\boldsymbol{u}}_f + \boldsymbol{G}^\top \boldsymbol{R}^{o-1} \boldsymbol{v}_{m+1}. \tag{40}$$

8. **Form the analysis ensemble.** As before,

$$\boldsymbol{u}_{m+1\,|\,m+1}^{(k)} = \bar{\boldsymbol{u}}_a + (\boldsymbol{U}_a)_{\cdot,k}, \qquad k = 1, \ldots, K.$$

2.3 Comparison: ETKF vs EAKF

A schematic comparison:

Aspect	ETKF	EAKF
Space of transform	Ensemble space ($K \times K$)	State space ($N \times N$)
Posterior anomalies	$\boldsymbol{U}_a = \boldsymbol{U}T$	$\boldsymbol{U}_a = A\boldsymbol{U}$
Main decompositions	Eigen/SVD of $K \times K$ matrix J	Eigens of two $N \times N$ matrices
Computational cost	Favourable when $K \ll N$	Potentially expensive when N large
Observation operator	Nonlinear $\boldsymbol{g}$ handled via $\boldsymbol{V}$	Requires linear $\boldsymbol{G}$ in basic form
Random perturbation	None (deterministic square-root)	None (deterministic adjustment)
Typical use	NWP/ocn DA with small K	NWP, often with localization

Both ETKF and EAKF:

- avoid perturbing observations (reducing sampling noise),
- enforce the correct posterior covariance (in the Kalman sense),
- need inflation and localization for small ensembles in high dimension.

3 Remarks, Stories, and Intuition

From ocean forecasts to modern DA workhorses

The ensemble Kalman filter (EnKF) was originally proposed by Evensen in the early 1990s for ocean forecasting. The key idea was almost "physics-driven Monte Carlo": run many copies of the model, adjust them with observations, and let the ensemble statistics estimate covariances on the fly.

A decade later, operational weather centres wanted to keep the ensemble idea, but the added noise in the stochastic EnKF was painful when the ensemble size was small (say $K = 100$ for systems with $N \sim 10^7$). This motivated deterministic square-root filters:

- Bishop's ETKF: "just" transform anomalies in ensemble space so their covariance matches the Kalman posterior.
- Anderson's EAKF: work directly with the state covariance, adjust the ensemble so the covariance shrinks in just the right directions.

Physical picture: reshuffling parallel universes

A useful picture:

- Each ensemble member is a "parallel universe" weather (or ocean) realization.
- The forecast step pushes these universes forward with the nonlinear PDEs.
- The observation tells you: "some of these universes are too wet here, too warm there".
- ETKF/EAKF do not kill or resample universes; instead they *reshuffle* them: pull them closer together in directions where observations are informative, and let them remain spread where there is little information.

The square-root transform is exactly that reshuffling in a multivariate Gaussian world.

Why two flavours? ETKF vs EAKF

- ETKF is often preferred when $K \ll N$: all heavy linear algebra is in $K \times K$ space, and localization can be applied in physical space.
- EAKF gives a very clean link to the analytic Kalman formula in state space, and is conceptually close to "rotating and shrinking" the covariance ellipsoid where observations are informative.
- In practice, many operational systems use ETKF-type ideas with localisation and hybridization (mixing ensemble and climatological covariances).

Some points

Informal comments:

- *"Nonlinear filters are like students before an exam: EKF reviews only the first-order Taylor series; EnKF/ETKF/EAKF try to capture the spread of all possible answers."*
- *"Square-root filters are covariance control freaks: they guarantee the covariance is exactly what the Kalman formula says it should be, without adding extra noise."*
- *"If you remember one thing: ETKF = transform in ensemble space, EAKF = adjust in state space. Same Kalman target, different coordinates."*

These notes are intended to be read together with the detailed derivations and numerical experiments in Chapter 9 of Majda–Harlim ("Filtering Complex Turbulent Systems").

References

- A. J. Majda and J. Harlim, *Filtering Complex Turbulent Systems*, Cambridge University Press, 2012, Chapter 9.
- G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods", J. Geophys. Res., 1994.
- C. H. Bishop et al., "Adaptive sampling with the ensemble transform Kalman filter", Mon. Weather Rev., 2001.
- J. L. Anderson, "An ensemble adjustment Kalman filter for data assimilation", Mon. Weather Rev., 2001.