A Quadrature Technique for Efficient Kalman Filtering With Model-Space Covariance Localization

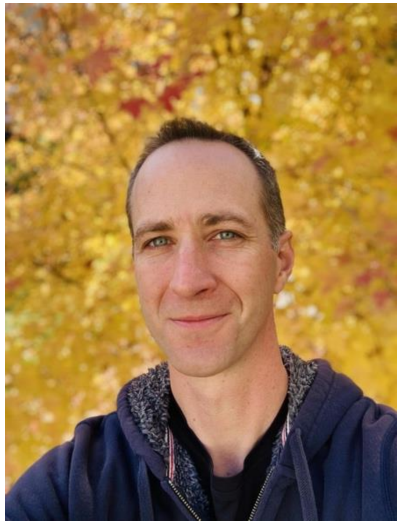
Robin Armstrong[1], Ian Grooms[2], and Chris Snyder[3]

2025 SIAM Conference on Applications of Dynamical Systems – CP25, "Data Driven Methods II"

May 15th, 2025

1. PhD Candidate, Cornell University, Center for Applied Mathematics
2. Associate Professor, CU Boulder, Department of Applied Mathematics
3. Senior Scientist, NSF-NCAR, Mesoscale and Microscale Meteorology Laboratory

Joint Work With:



Ian Grooms (CU Boulder)

Chris Snyder (NSF-NCAR MMM)

Data Assimilation

- … is a Bayesian approach for estimating an unknown system state based on noisy, partial observations.

- For example: improving atmospheric forecasts using satellite radiance measurements.
  - Data is high-dimensional, $\approx \mathcal{O}(10^6)$.
  - State is even higher-dimensional, $\approx \mathcal{O}(10^9)$.

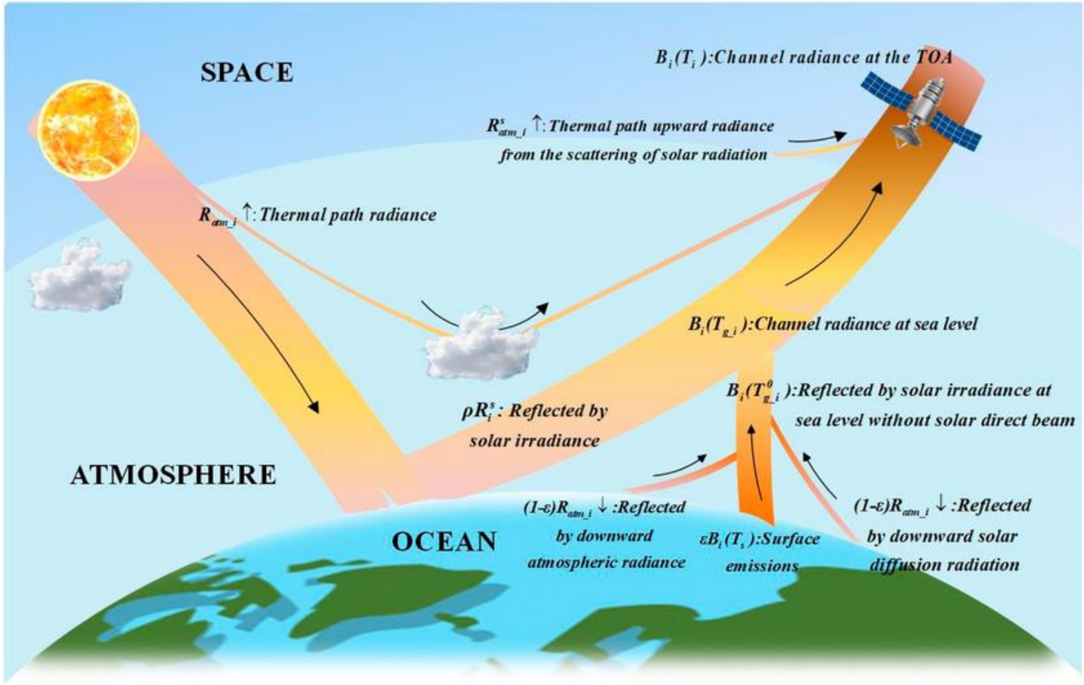
- First ingredient: a "forecast distribution" (prior) on the system state:

$$p(x) = P(X = x).$$

- Second ingredient: an observation model,

$$Y = h(X) + \xi.$$

- Goal: use Bayes' rule to find $P(X = x \mid Y = y)$.



Source: https://www.researchgate.net/figure/Radiance-received-by-the-satellite-sensor_fig4_375018928

Ensemble Kalman Filters (EnKFs)

- Kalman filter assumes a Gaussian "forecast" (prior) in joint state+observable space:

$$\begin{bmatrix} X_f \\ h(X_f) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_h \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_{xx} & \mathbf{\Sigma}_{xh} \\ \mathbf{\Sigma}_{xh}^{\mathrm{T}} & \mathbf{\Sigma}_{hh} \end{bmatrix} \right),$$

  and assumes Gaussian obs. errors, $\xi \sim \mathcal{N}(0, \mathbf{R})$.

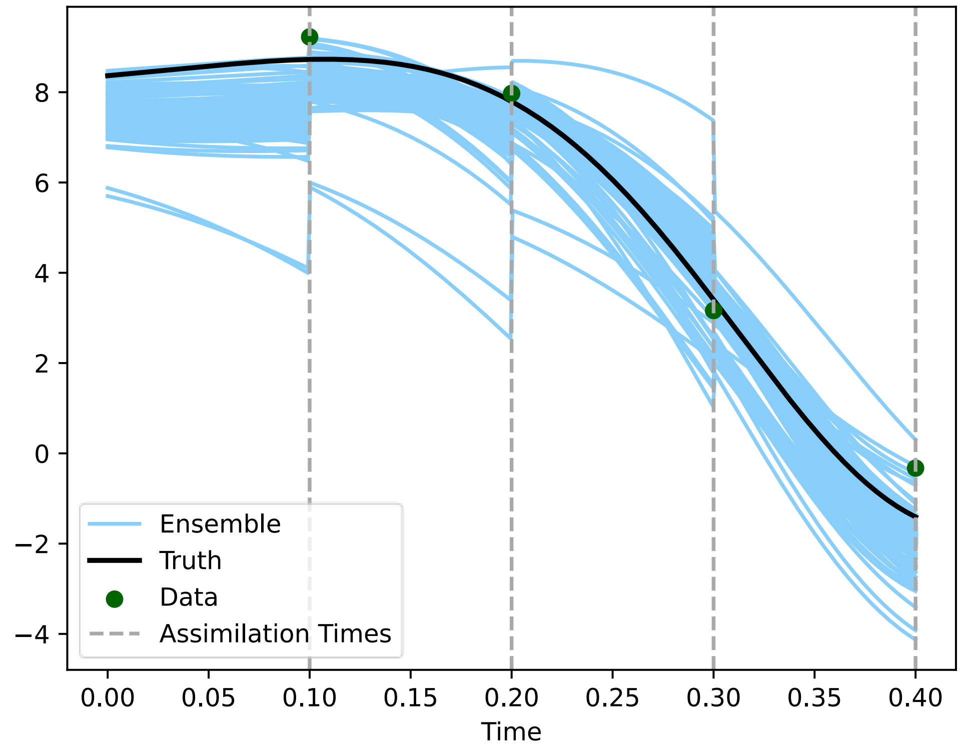
- "Analysis" (posterior) given $Y = y$ becomes Gaussian: $X_a \sim \mathcal{N}(\mu_a, \Sigma_a)$ with

$$\mu_a = \mu_x + \mathbf{K}(y - \mu_h), \qquad \mathbf{\Sigma}_a = \mathbf{\Sigma}_{xx} - \mathbf{K}\mathbf{\Sigma}_{xh}^{\mathrm{T}},$$

  where $\mathbf{K} = \mathbf{\Sigma}_{xh}(\mathbf{R} + \mathbf{\Sigma}_{hh})^{-1}$ = "Kalman gain matrix."

- Ensemble Kalman filters (EnKFs) transform samples of X_f into samples of X_a:

$$f(X_f, y) \sim X_a$$



Stochastic vs Square-Root Filters

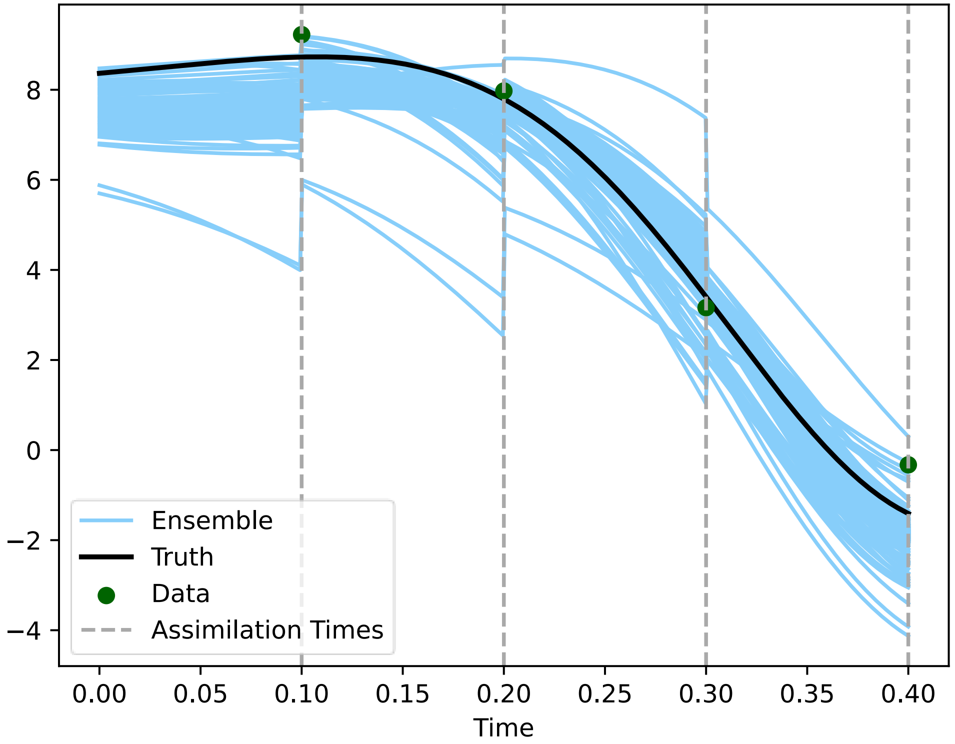
- First papers on EnKF (Evensen, 1994) updated each sample as if it were the mean.

$$\hat{X}_a = X_f + \mathbf{K}\left(y - h(X_f)\right).$$

This undercounts obs. error: Cov$[\hat{X}_a] \prec \boldsymbol{\Sigma}_a$.

- Stochastic EnKF (Houtekamer & Mitchell, 1998) adds random noise to the observation.

$$X_a = X_f + \mathbf{K}\big(y - h(X_f) - \hat{\xi}\big), \qquad \hat{\xi} \sim \mathcal{N}(0, \mathbf{R}).$$



- Square-root filter (Whitaker & Hamill, 2002) update the means and perturbations separately.

$$\mu_a = \mu_x + \mathbf{K}(y - \mu_h), \qquad \mathbf{K} = \boldsymbol{\Sigma}_{xh}(\mathbf{R} + \boldsymbol{\Sigma}_{hh})^{-1}$$

$$X_a - \mu_a = X_f - \mu_x - \mathbf{G}\big(h(X_f) - \mu_h\big), \qquad \mathbf{G} = \boldsymbol{\Sigma}_{xh}\left(\mathbf{R} + \boldsymbol{\Sigma}_{hh} + \mathbf{R}(\mathbf{I} + \mathbf{R}^{-1}\boldsymbol{\Sigma}_{hh})^{1/2}\right)^{-1}$$

Covariance Localization

- The <u>empirical ensemble covariance</u>,

$$\boldsymbol{\Sigma}_{xx} = \mathbf{Z}_x \mathbf{Z}_x^{\mathrm{T}},$$

  is low-rank (good for efficiency) but noisy (bad for accuracy).

- The <u>localized ensemble covariance</u>,

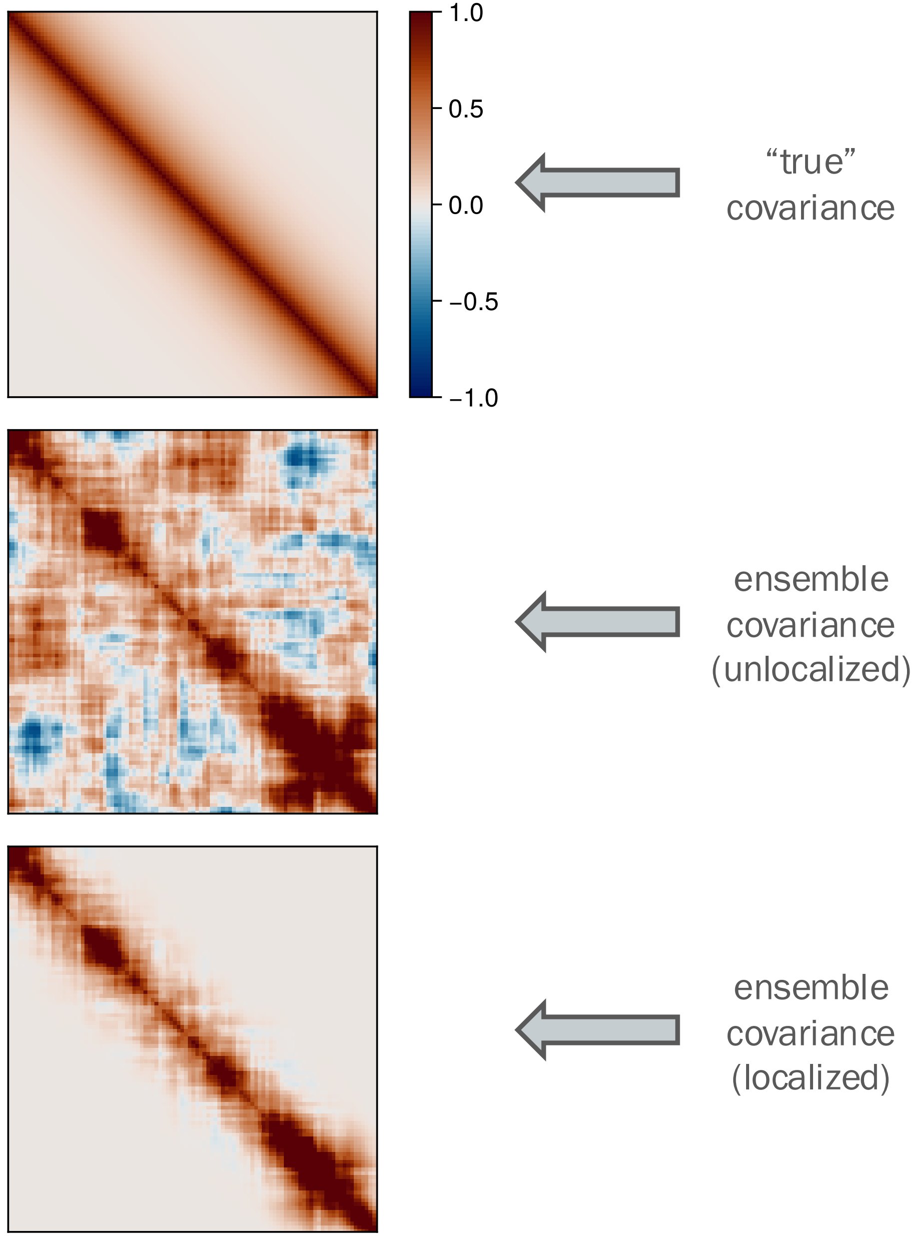
$$\boldsymbol{\Sigma}_{xx} = \mathbf{L} \circ (\mathbf{Z}_x \mathbf{Z}_x^{T})$$

  is less noisy but is <u>full-rank</u> (difficult to compute with).

- <u>Operator access</u> (Farchi and Bocquet, FAMS, 2019) lets us run Krylov methods and linear solves:

$$\boldsymbol{\Sigma}_{xx} u = \sum_i z^{(i)} \circ \mathbf{L}(z^{(i)} \circ u) \,.$$

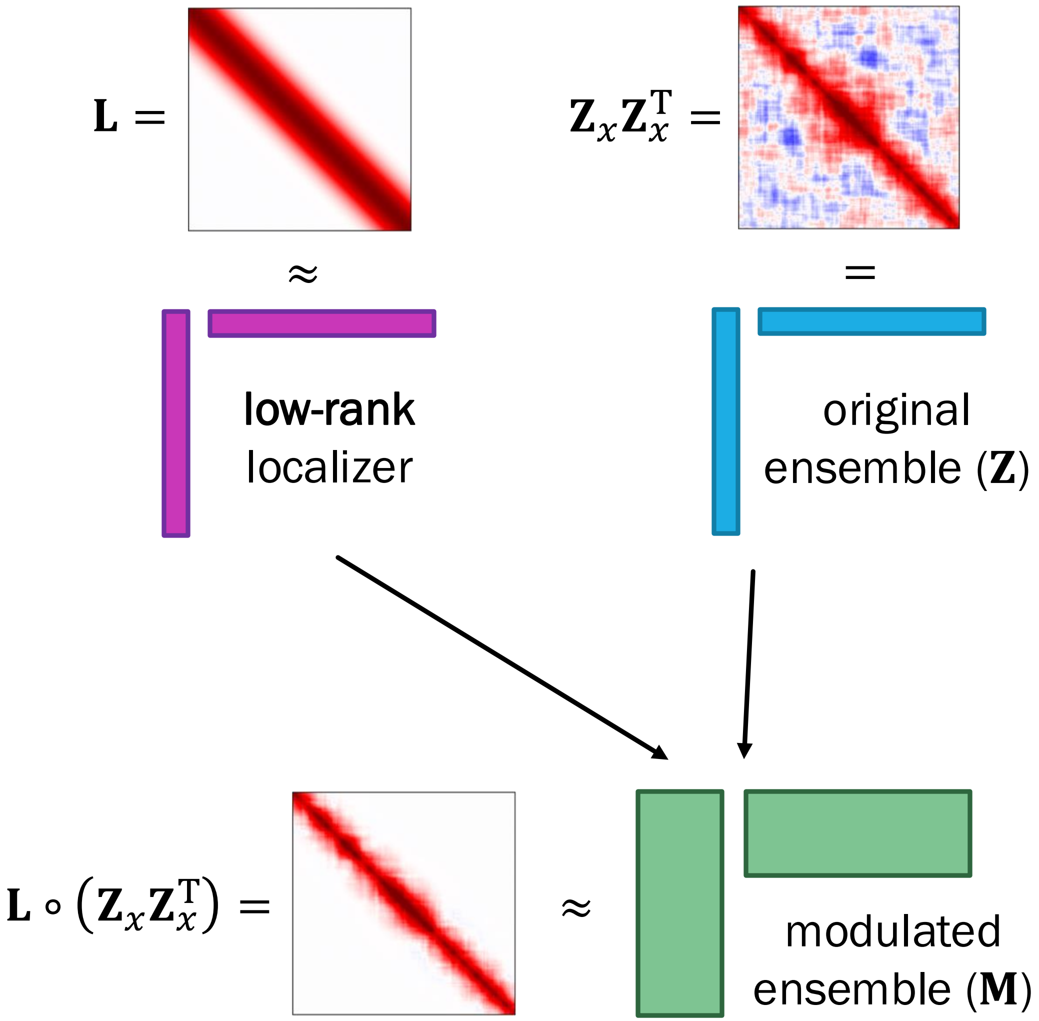
- **How to handle the square-root?**



Ensemble Modulation

- Constructs a larger modulated ensemble whose empirical covariance retains some localization.
- Essentially a low-rank factorization of $\mathbf{L} \circ (\mathbf{Z}_x \mathbf{Z}_x^{\mathrm{T}})$.
- Constructed using a spectral decomposition of $\mathbf{L}$ (Bishop et al., 2017) or using a randomized SVD (Farchi & Bocquet, 2019).
- Only works well if $\mathbf{L} \circ (\mathbf{Z}_x \mathbf{Z}_x^{\mathrm{T}})$ has fast singular value decay!
- Measure of complexity:

$$k = \frac{\text{modulated ensemble size}}{\text{original ensemble size}}$$



A Modulation-Free Perturbation Update

- Stochastic EnKF:

$$\begin{aligned} \mu_a &= \mu_x + \mathbf{K}(y - \mu_h) \\ X_a - \mu_a &= X_f - \mu_x - \mathbf{K}\big(h(X_f) + \xi - \mu_h\big), \qquad \xi \sim \mathcal{N}(0, \mathbf{R}). \end{aligned}$$

- **Key idea 1**: deterministically inflate $\mathbf{R}$.

Theorem (A. and Grooms, 2025). Let $\mathbf{K}(s) = \mathbf{\Sigma}_{xh}((s+1)\mathbf{R} + \mathbf{\Sigma}_{hh})^{-1}$, and let $p(s) = [\pi\sqrt{s}(s+1)]^{-1}$ for $s \in (0, \infty)$. If

$$\begin{aligned} \mu_a &= \mu_x + \mathbf{K}(y - \mu_h), \\ X_a - \mu_a &= X_f - \mu_x - \int_0^\infty \mathbf{K}(s)\big(h(X_f) - \mu_h\big)\, p(s) ds\,, \end{aligned}$$

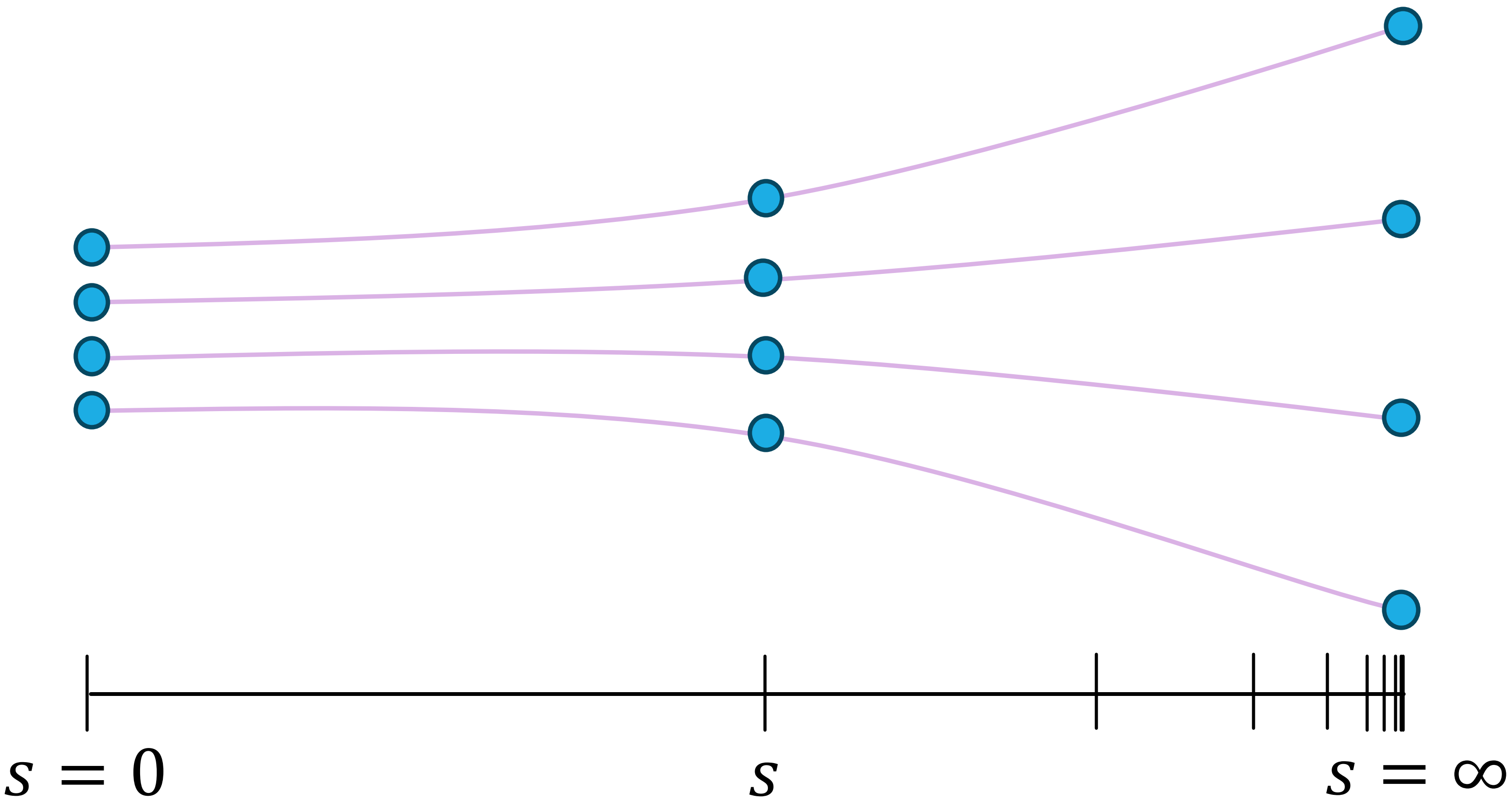
then $\mathbb{E}[X_a]$ and $\mathrm{Cov}[X_a]$ satisfy the Kalman filter equations.

- **Key idea 2**: use quadrature to discretize the integral.

$$\mathbf{K}(s) = \mathbf{\Sigma}_{xh}\big((s+1)\mathbf{R} + \mathbf{\Sigma}_{hh}\big)^{-1}$$

$$\mathbf{K}(0) = \mathbf{K}$$

un-perturbed EnKF; analysis is under-dispersed.

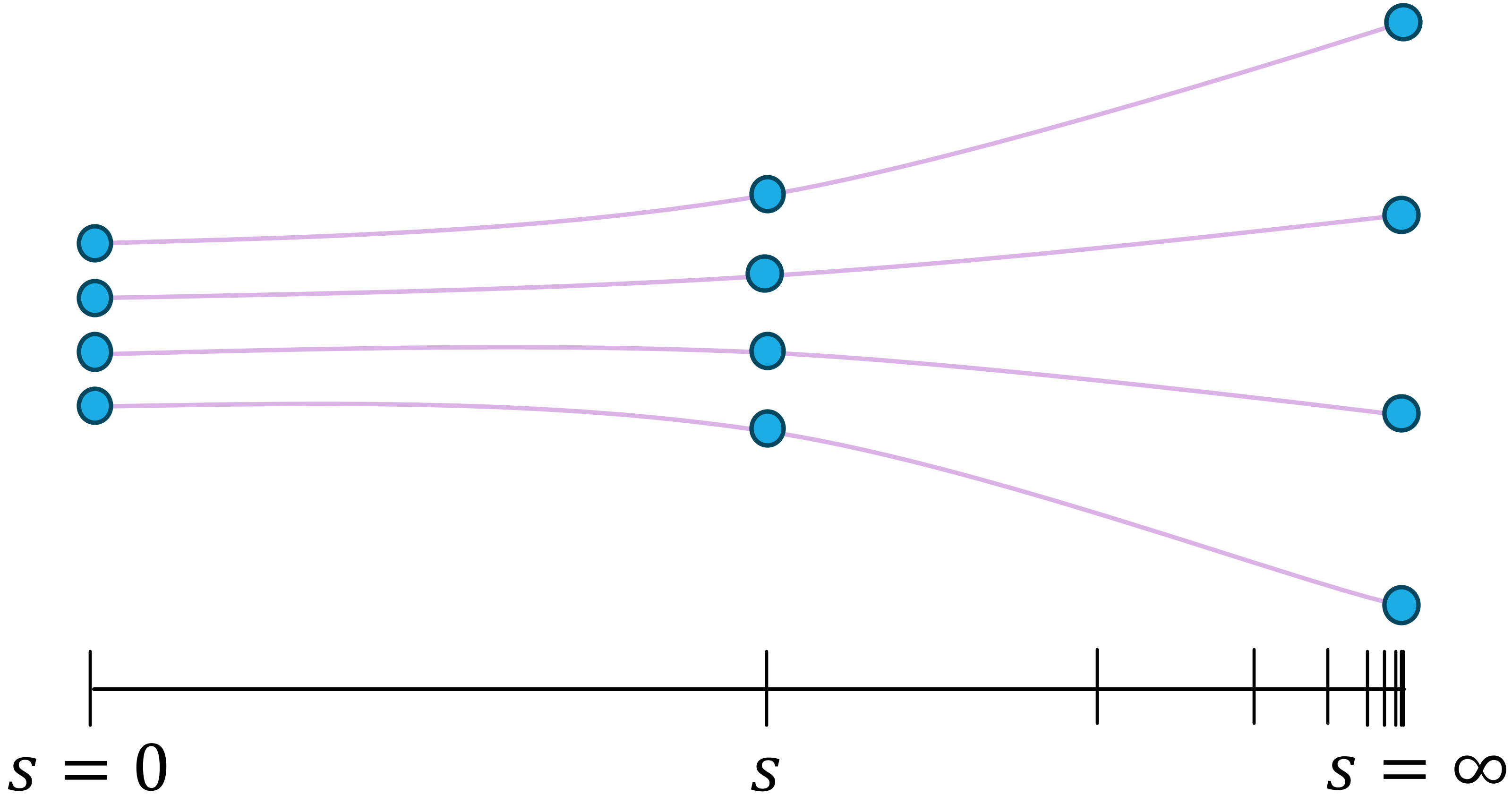


$$\mathbf{K}(\infty) = \mathbf{0}$$

data has no effect; analysis is over-dispersed

$$\mathbf{K}(s) = \mathbf{\Sigma}_{xh}\left((s+1)\mathbf{R} + \mathbf{\Sigma}_{hh}\right)^{-1}$$

$$X_a - \mu_a = X_f - \mu_x - \int_0^\infty \mathbf{K}(s)\left(h\left(X_f\right) - \mu_h\right) p(s)ds$$



$$\mathbf{K}(0) = \mathbf{K}$$

un-perturbed EnKF; analysis is under-dispersed.

$$\mathbf{K}(\infty) = \mathbf{0}$$

data has no effect; analysis is over-dispersed

Reweighting The Integral

- We want to evaluate

$$\int_0^\infty \mathbf{K}(s)\big(h(X_f) - \mu_h\big)\, p(s)ds\,,$$

where $p(s) = [\pi\sqrt{s}(s+1)]^{-1}$. Very slowly decaying!

Theorem (A. and Grooms, 2025). Define $\mathcal{D} \subseteq \mathbb{R}_+$ and $\hat{r}, \hat{s} : \mathcal{D} \to \mathbb{R}_+$ such that

$$\frac{1}{\sqrt{c+1}} = \int_{t\in\mathcal{D}} \frac{\hat{r}(t)}{c + \hat{s}(t) + 1}\, dt$$

for all c in some open set containing $\{0\} \cup \lambda(\mathbf{R}^{-1/2}\boldsymbol{\Sigma}_{hh}\mathbf{R}^{-1/2})$. Then,

$$\int_0^\infty \mathbf{K}(s)\big(h(X_f) - \mu_h\big)\, p(s)ds = \int_{t\in\mathcal{D}} \mathbf{K}(\hat{s}(t))\,\big(h(X_f) - \mu_h\big)\, \hat{p}(t)dt,$$

where $\hat{p}(t) = \hat{r}(t)(\hat{s}(t) + 1)^{-1}$ (note that $\int \hat{p}(t)\, dt = 1$).

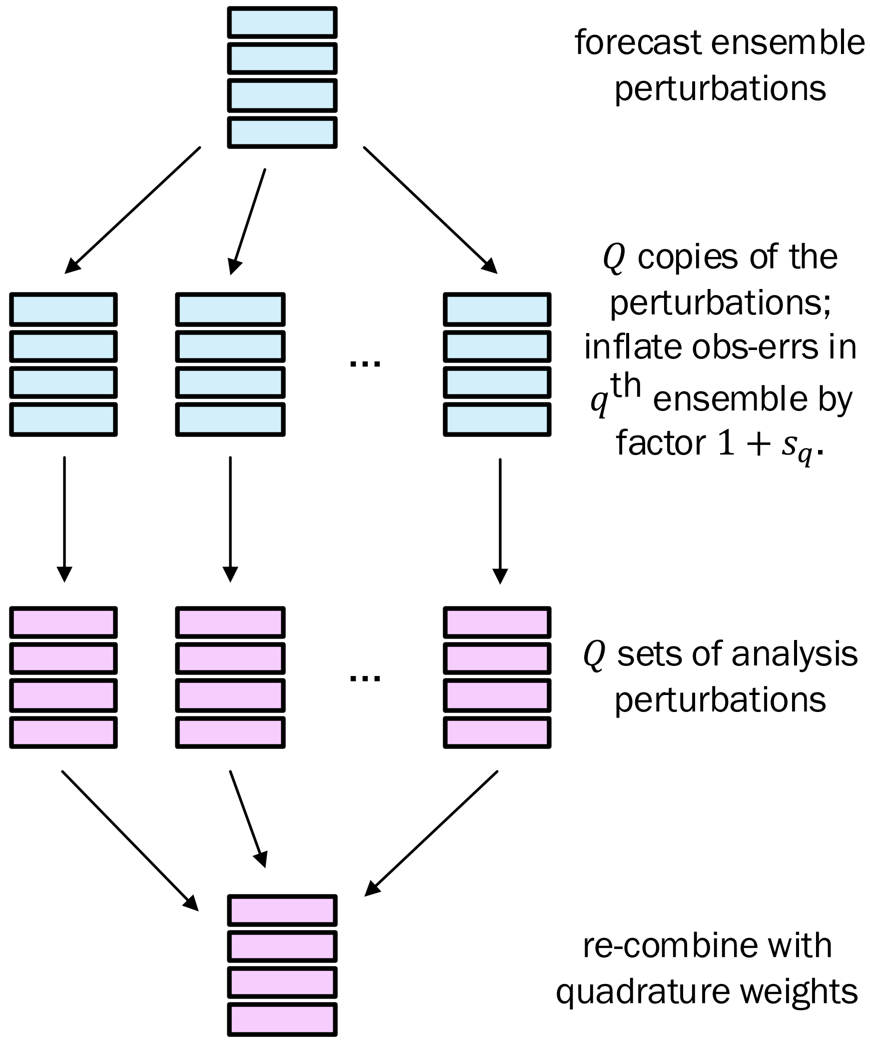
A Modulation-Free Localized Square-Root Filter

- Quadrature approximation:

$$\int_{t\in\mathcal{D}} \mathbf{K}(\hat{s}(t))\,\big(h(X_f)-\mu_h\big)\,\hat{p}(t)dt \approx \sum_{q=1}^{Q} p_q \mathbf{K}(s_q)\,\big(h(X_f)-\mu_h\big)$$

- 1. for $i = 1, \ldots, m$ # both loops parallelize
- 2. $z_a^{(i)} \leftarrow x_f^{(i)} - \mu_x$
- 3. for $q = 1, \ldots, Q$
- 4. $z_a^{(i)} \leftarrow z_a^{(i)} - p_q \mathbf{K}(s_q)\left(h\left(x_f^{(i)}\right) - \mu_h\right)$ # PGC
- 5. end
- 6. end

- Measure of complexity: $k = \text{ number of quadrature points}$



A Modulation-Free Localized Square-Root Filter

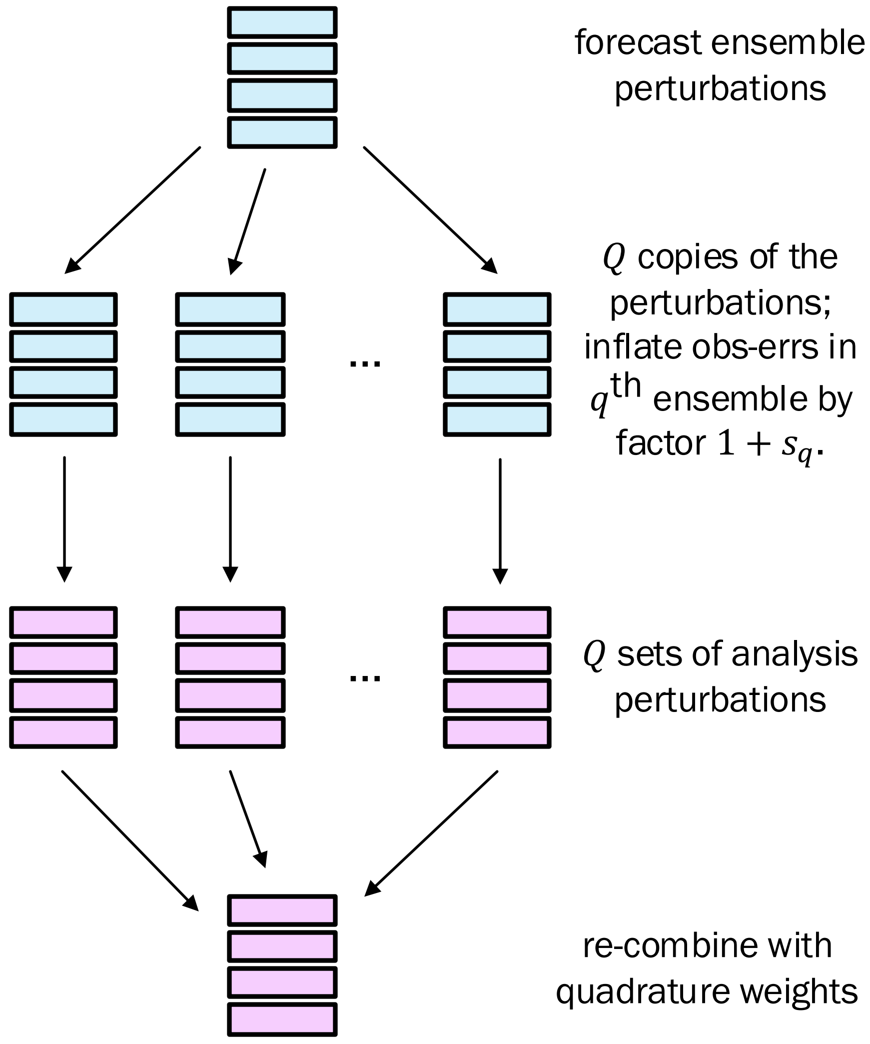
- Quadrature approximation:

$$\int_{t\in\mathcal{D}} \mathbf{K}(\hat{s}(t))\,\left(h(X_f)-\mu_h\right)\hat{p}(t)dt \approx \sum_{q=1}^{Q} p_q\mathbf{K}(s_q)\,\left(h(X_f)-\mu_h\right)$$

<u>InFo-ESRF (Integral-Form Ensemble Square-Root Filter)</u>

- 1. for $i = 1, \ldots, m$ # both loops parallelize
- 2. $z_a^{(i)} \leftarrow x_f^{(i)} - \mu_x$
- 3. for $q = 1, \ldots, Q$
- 4. $z_a^{(i)} \leftarrow z_a^{(i)} - p_q\mathbf{K}(s_q)\left(h\left(x_f^{(i)}\right) - \mu_h\right)$ # PGC
- 5. end
- 6. end

- Measure of complexity: $k = \text{ number of quadrature points}$

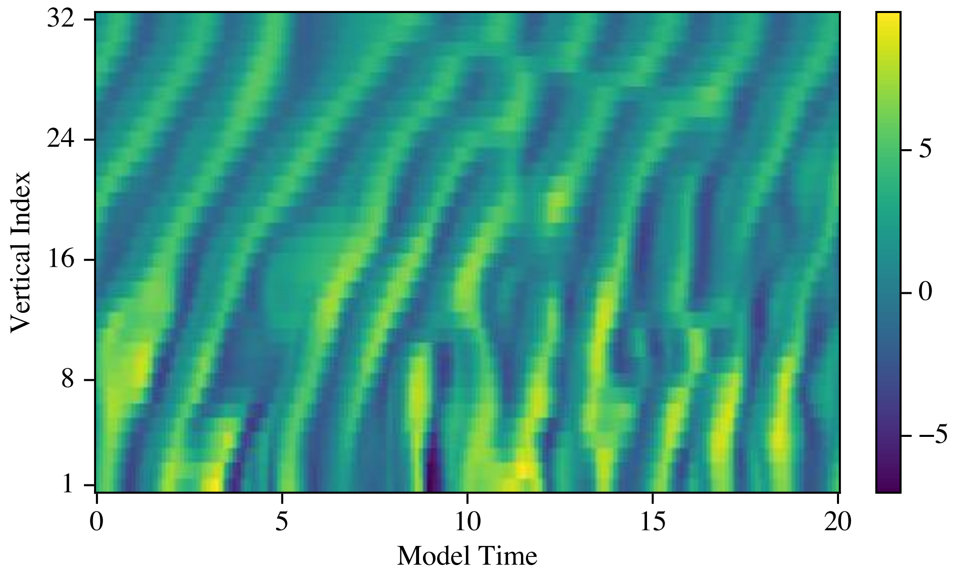


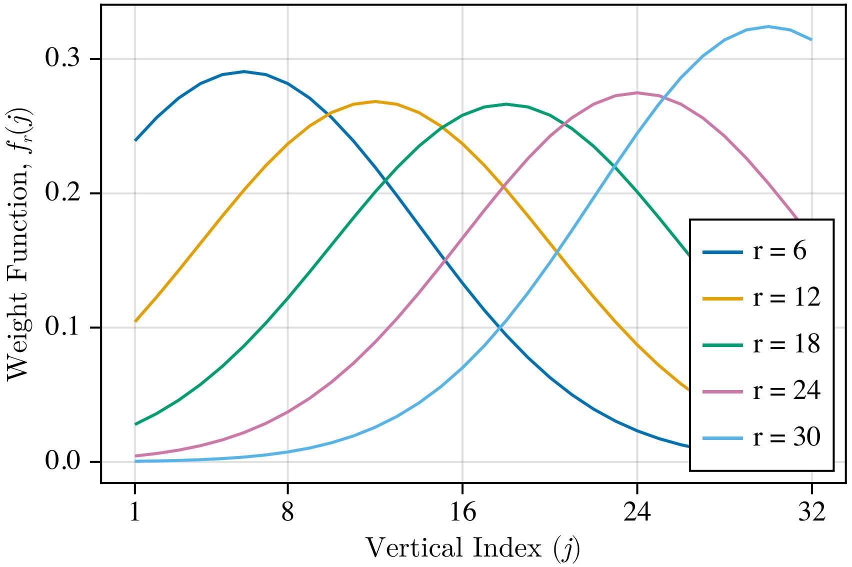
Test Case 1: Multi-Layer Lorenz System

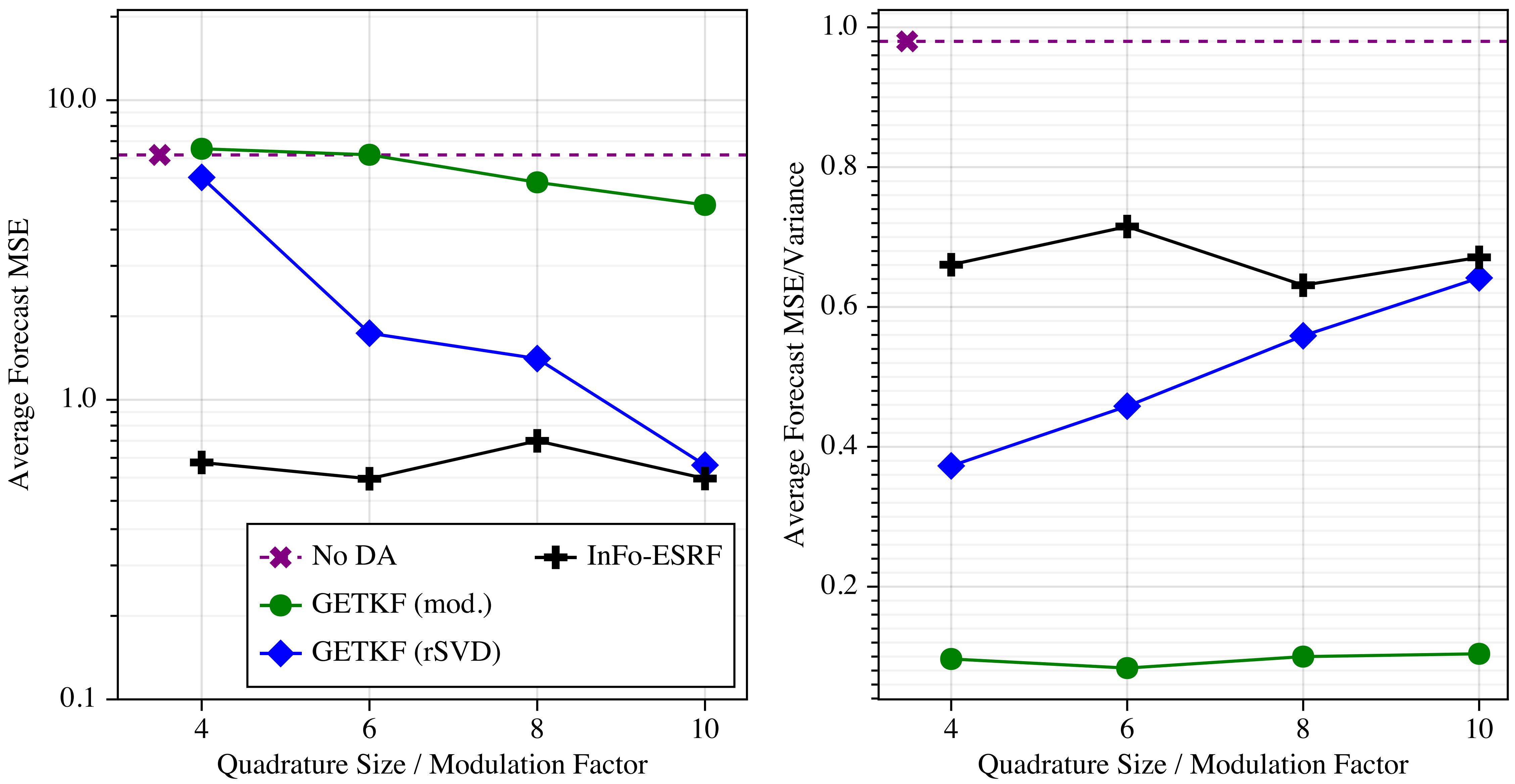
- **Model and Observing System:**
  - 32 coupled layers of 40-variable L'96 systems.
  - Forcing decreases linearly from 8 (bottom) to 4 (top).
  - "Satellite-like" measurements of 5 weighted vertical sums for every 5th column.
  - Gaussian i.i.d. noise with variance ≈ 1% of climatological variance.
  - From Farchi and Bocquet, FAMS (2019).

- **Localization:** Gaspari-Cohn model-space localization in both the horizontal and vertical. Tuned for optimal accuracy under "brute force" ESRF (no approximation of square-root).

- **Experiment:** 5000 forecast-assimilation cycles ($\Delta t = 0.05$), average MSE and spread measured over last 4000.



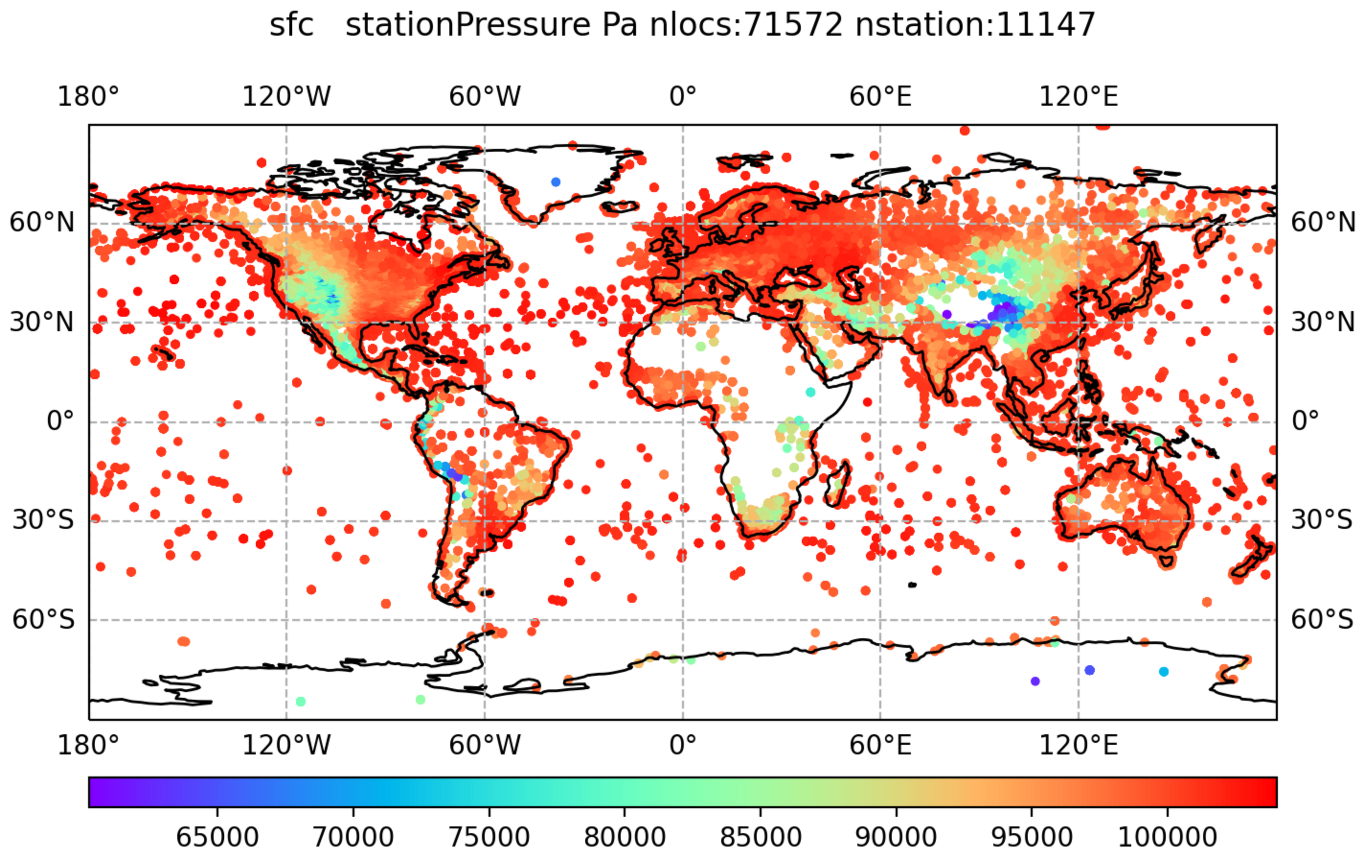




Average Forecast MSE
10.0
1.0
0.1
4
6
8
10
Quadrature Size / Modulation Factor
No DA
InFo-ESRF
GETKF (mod.)
GETKF (rSVD)
Average Forecast MSE/Variance
1.0
0.8
0.6
0.4
0.2
4
6
8
10
Quadrature Size / Modulation Factor

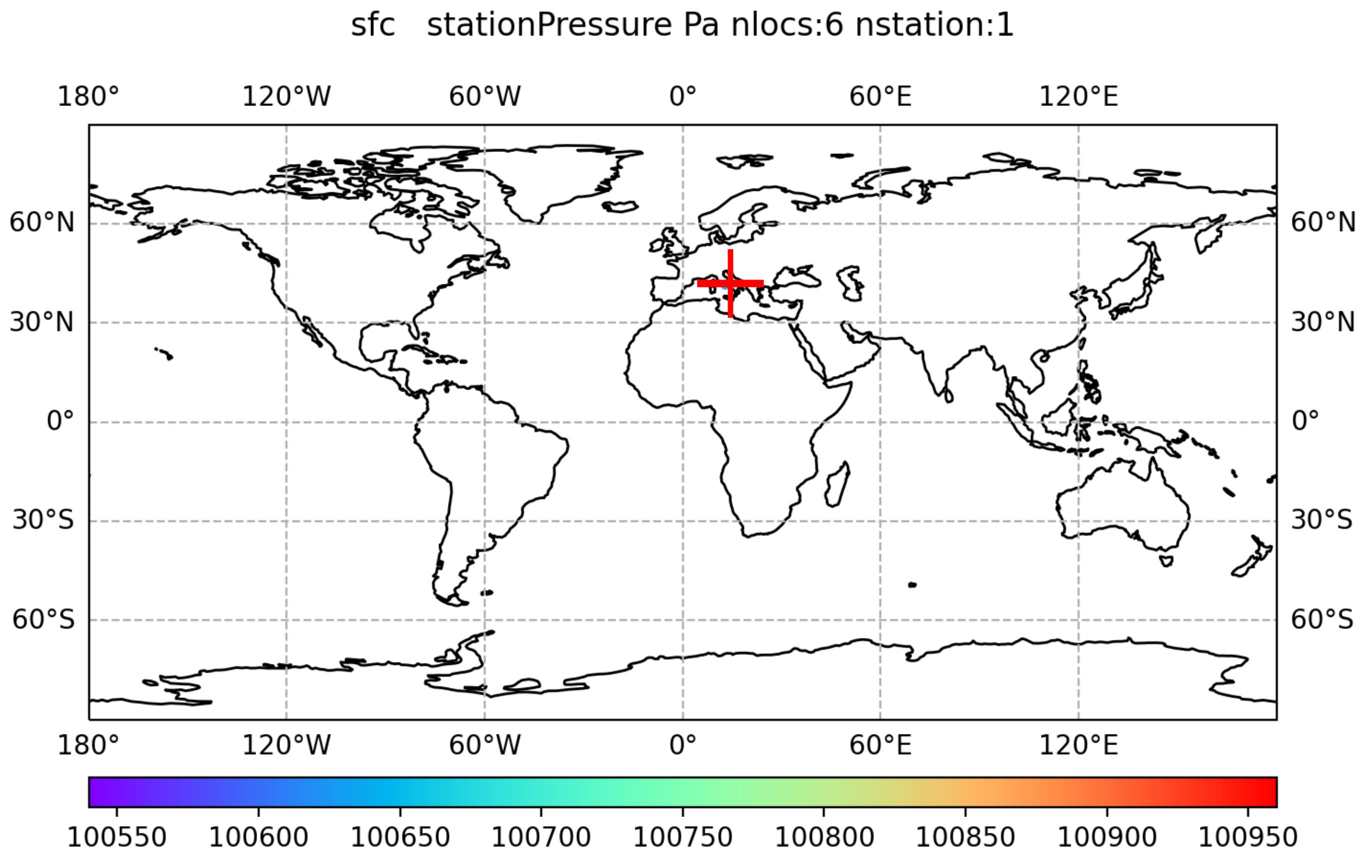
Test Case 2: MPAS-JEDI

- **Background Ensemble:** 20-members from GEFS analysis.
- **Background Covariance:** SABER ensemble covariance model, BUMP-NICAS localization.
- **Observations:** a <u>single</u> surface pressure measurement over Italy.
- **Implementation:** Kalman gain matrices applied using MPAS-JEDI 3D-EnVar executables (<u>one</u> outer loop).

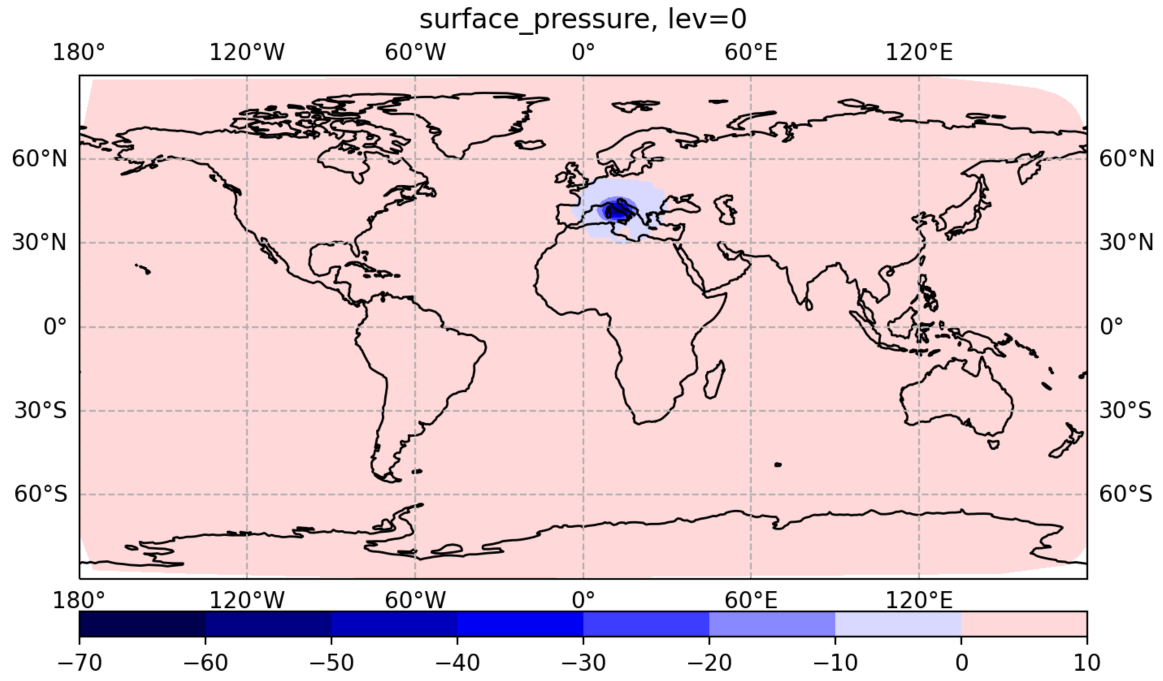


Test Case 2: MPAS-JEDI

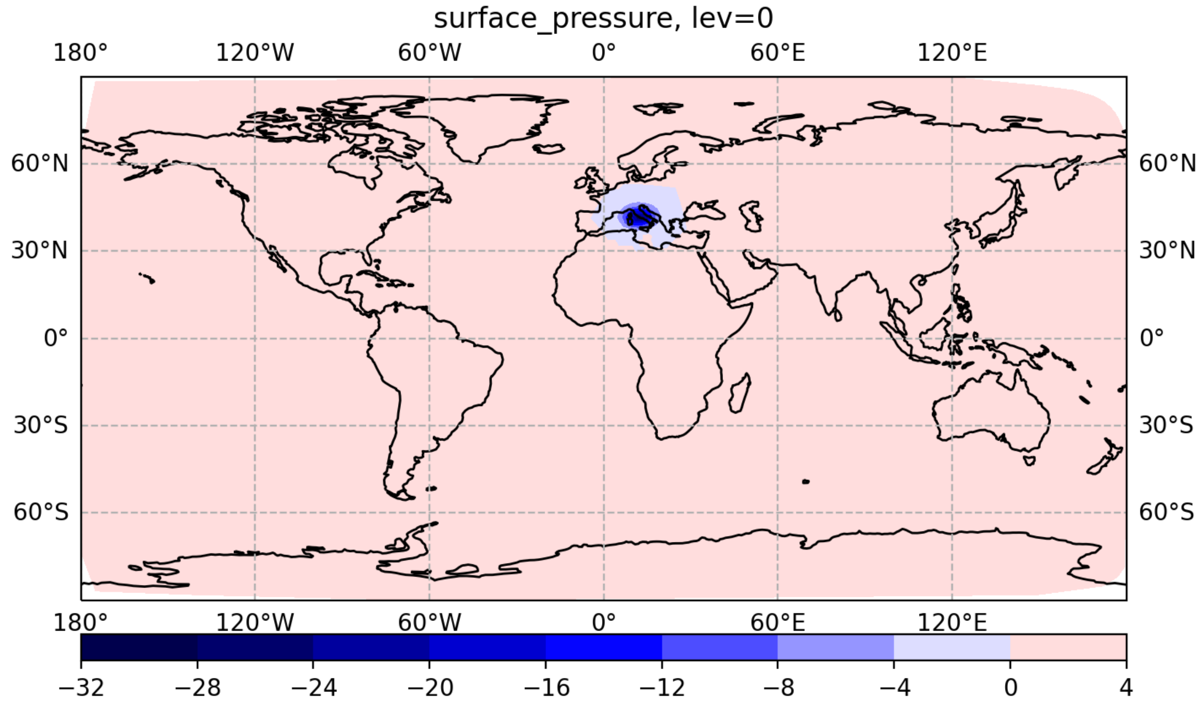
- **Background Ensemble**: 20-members from GEFS analysis.
- **Background Covariance**: SABER ensemble covariance model, BUMP-NICAS localization.
- **Observations**: a <u>single</u> surface pressure measurement over Italy.
- **Implementation**: Kalman gain matrices applied using MPAS-JEDI 3D-EnVar executables (<u>one</u> outer loop).



MPAS-JEDI Results: Changes in Ensemble Variance



EDA/3D-EnVar (perturbed obs)



InFo-ESRF (no obs perturbations)

Conclusion

- **We have demonstrated:** *an ensemble square-root filter which updates perturbations by discretizing an integral form of the Kalman filter update equations. This lets us avoid evaluating a matrix square-root, eliminating the need to approximate the forecast covariance with modulation.*

A. and Grooms, “Data Assimilation With An Integral-Form Ensemble Square-Root Filter,” arXiv:2503.00253