The EnKF-N and inflation

Patrick N. Raanes[1,2,*], Marc Bocquet[3]

[1]Nansen Environmental and Remote Sensing Center, [2]Mathematical Institute, University of Oxford, [*]patrick.raanes@nersc.no
[3]CEREA, Joint laboratory École des Ponts ParisTech and EDF R&D, Université Paris-Est

Constitutive assumptions

The ensemble is an iid sample from the same *Gaussian* distribution as the truth, i.e. $\boldsymbol{x}, \boldsymbol{x}_n \sim \mathcal{N}(\boldsymbol{b}, \mathbf{B})$. However, the true moments, $\boldsymbol{b}$ and $\mathbf{B}$ are *unknown*. Then, by marginalization,

$$p(\boldsymbol{x}|\mathbf{E}) = \int p(\boldsymbol{x}|\boldsymbol{b}, \mathbf{B}, \cancel{\mathbf{E}})\, p(\boldsymbol{b}, \mathbf{B}|\mathbf{E})\, d\boldsymbol{b}\, d\mathbf{B}\,.$$

The EnKF-N effective prior

With $p(\boldsymbol{b}, \mathbf{B}) \propto |\mathbf{B}|^{-(m+1)/2}$, the effective prior is

$$p(\boldsymbol{w}|\mathbf{E}) = \left(\varepsilon_N + \|\boldsymbol{w}\|^2_{\mathbf{I}_N}\right)^{-(N+g)/2}$$
$$\propto \boldsymbol{t}_N(\boldsymbol{w}|g; 0, \varepsilon_N \mathbf{I}_N)\,,$$

where $\varepsilon_N = 1 + \frac{1}{N}$, and g is the rank of the nullspace of $\mathbf{A}$. The posterior is $p(\boldsymbol{w}|\mathbf{E}, \boldsymbol{y}) \propto \exp -\frac{1}{2}\left(\|\boldsymbol{y} - h(\boldsymbol{x}(\boldsymbol{w}))\|^2_{\mathbf{R}} + (N+g)\log(\varepsilon_N + \|\boldsymbol{w}\|^2_{\mathbf{I}_N})\right)$

The standard EnKF prior

Denote $\mathbf{E} = [\boldsymbol{x}_1, \ \dots \ \boldsymbol{x}_n, \ \dots \ \boldsymbol{x}_N]$ and let $\mathbf{A}$ be the corresponding matrix of anomalies. Parameterize $\boldsymbol{x}$ by ensemble weights, $\boldsymbol{w}$:

$$\boldsymbol{x}(\boldsymbol{w}) = \bar{\boldsymbol{x}} + \mathbf{A}\boldsymbol{w}$$
$$\boldsymbol{y} - h(\boldsymbol{x}(\boldsymbol{w})) = \boldsymbol{y} - h(\bar{\boldsymbol{x}} + \mathbf{A}\boldsymbol{w}) \approx \bar{\boldsymbol{\delta}} - \mathbf{Y}\boldsymbol{w}$$

Then, $p(\boldsymbol{x}|\boldsymbol{b}{=}\bar{\boldsymbol{x}}, \mathbf{B}{=}\bar{\mathbf{B}}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{w}\|^2_{\mathbf{I}_N}\right)$.

The EnKF's intrinsic bias

Define $\bar{\mathbf{B}} = \frac{1}{N-1}\sum_{n=1}^N (\boldsymbol{x}_n - \bar{\boldsymbol{x}})(\boldsymbol{x}_n - \bar{\boldsymbol{x}})^{\mathsf{T}}$. Then

$$\mathbb{E}(\bar{\mathbf{B}}) = \mathbf{B}\,,$$

(with expectation over all $\boldsymbol{x}_n$). But, with $\bar{\mathbf{P}}^a = [\mathbf{I} - \bar{\mathbf{K}}\mathbf{H}]\bar{\mathbf{B}}$, and the same expectation,

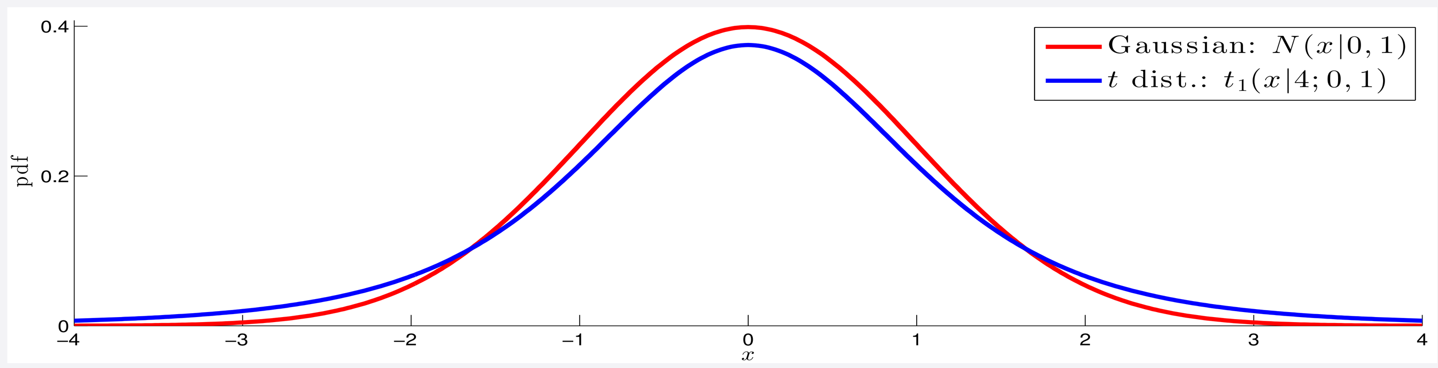
$$\mathbb{E}(\mathrm{tr}(\bar{\mathbf{P}}^a)) < \mathrm{tr}(\mathbf{P}^a)\,,$$

and idem. for $\bar{\mathbf{K}}$.

The t distribution

Note the thick tails.

$$\mathcal{N}(x|0,1) \propto e^{-\frac{1}{2}x^2}$$
$$\boldsymbol{t}_1(x|1;0,1) \propto \frac{1}{1+x^2}$$



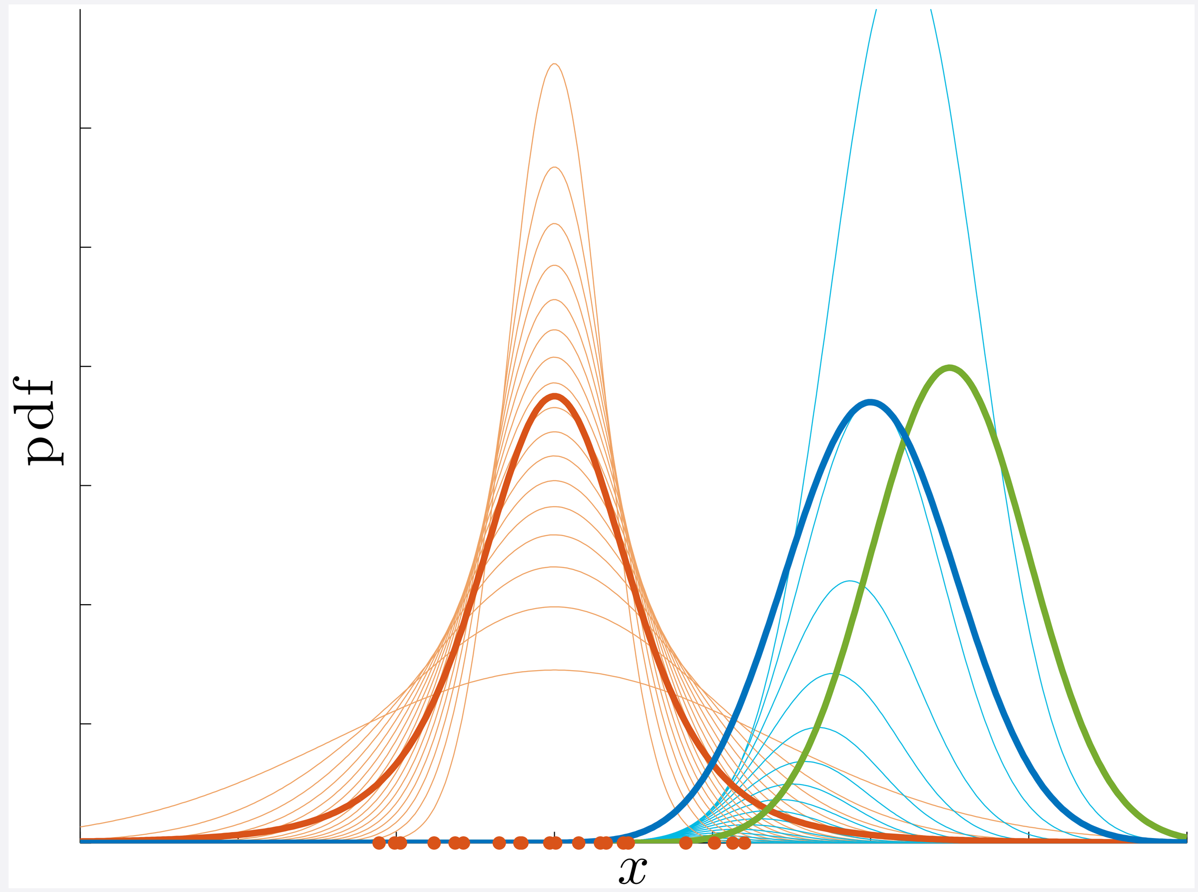
Scale mixture re-derivation

The dual EnKF-N may also be derived as a *scalar*, scale mixture. Let $\lambda^2 = \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2_{\bar{\mathbf{B}}} / \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2_{\mathbf{B}}$. Then

$$p(\boldsymbol{w}|\mathbf{E}, \boldsymbol{y}) = \int p(\boldsymbol{w}, \lambda^2|\mathbf{E}, \boldsymbol{y})\, d\lambda^2$$
$$\propto \int \mathcal{N}(\bar{\boldsymbol{\delta}}|\mathbf{Y}\boldsymbol{w}, \mathbf{R})\, \mathcal{N}(\boldsymbol{w}/\lambda|0, \tfrac{\varepsilon_N}{N-1}\mathbf{I}_N)\, \chi^{-2}(\lambda^2|N-1)\, \lambda^{-g}\, d\lambda$$
$$\propto \int \exp -\frac{1}{2}\left(\|\mathbf{Y}\boldsymbol{w} - \bar{\boldsymbol{\delta}}\|^2_{\mathbf{R}} + \frac{N-1}{\varepsilon_N \lambda^2}\|\boldsymbol{w}\|^2_{\mathbf{I}_N} + \frac{N-1}{\lambda^2} + (N+g)\log\lambda^2\right) d\lambda$$
$$\approx c \exp -\frac{1}{2}\Big(\|\boldsymbol{w} - \boldsymbol{w}_\star(\zeta_\star)\|^2_{\mathbf{G}_\star(\zeta_\star)} + \underbrace{\|\bar{\boldsymbol{\delta}}\|^2_{\mathbf{Y}\mathbf{Y}^{\mathsf{T}}/\zeta_\star + \mathbf{R}} + \varepsilon_N\zeta_\star - (N+g)\log\zeta_\star}_{D(\zeta_\star)}\Big)\,,$$

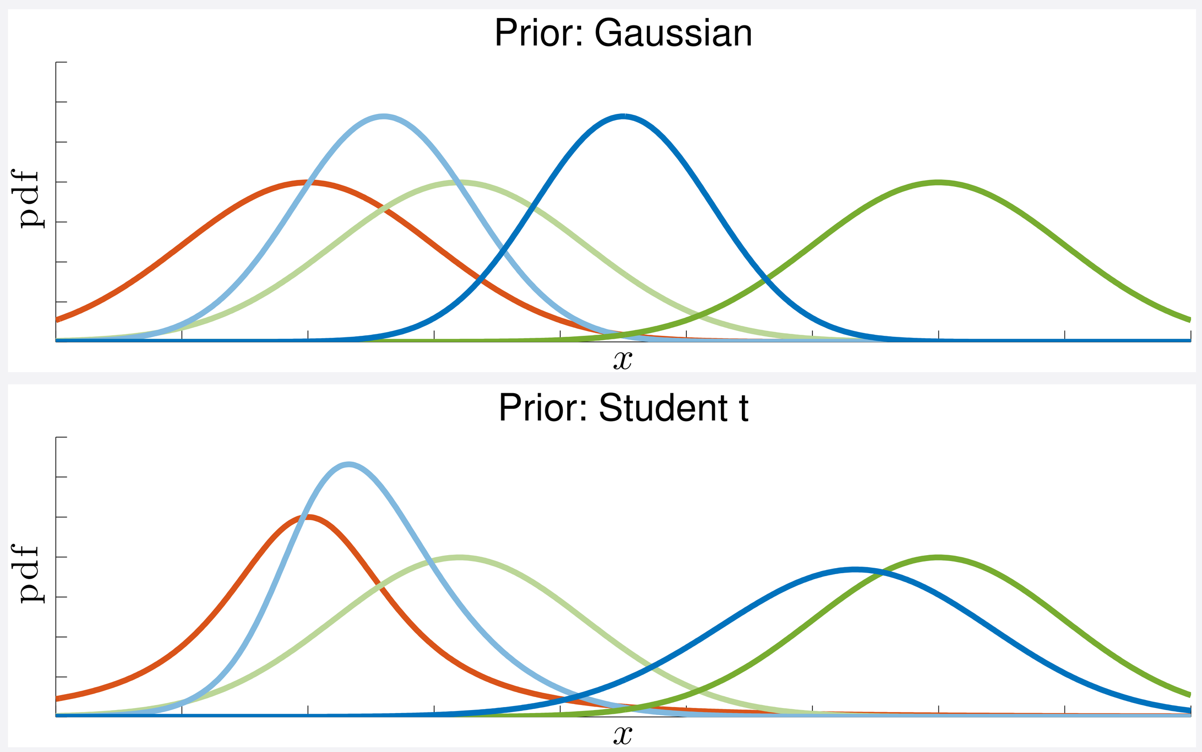
where $\zeta = \frac{N-1}{\varepsilon_N \lambda^2}$.

Scale (and loc.) mixtures



The effective (thick lines) prior and posteriors are the averages of the "candidate" (thin lines) prior and posteriors.

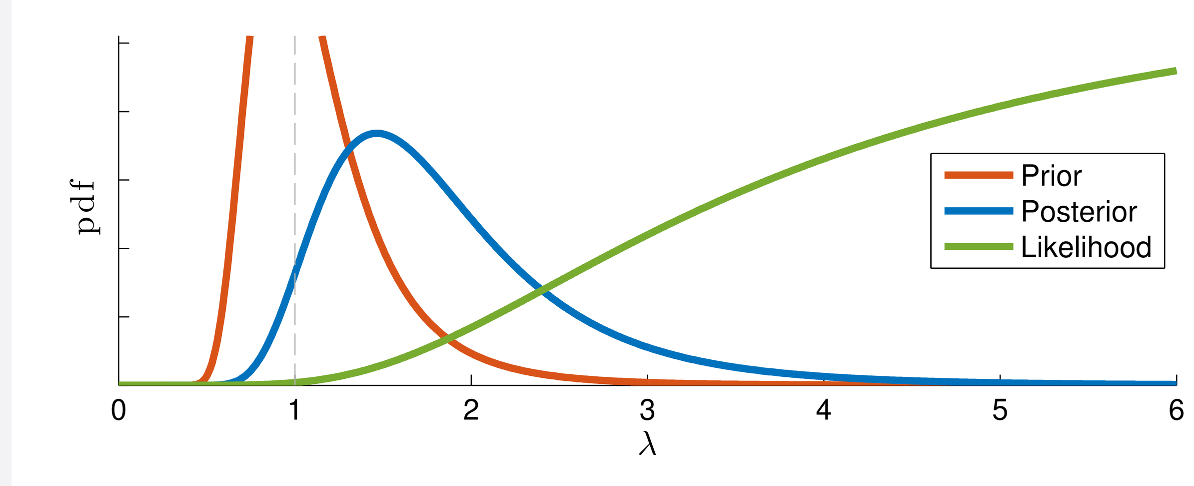
Posterior's behaviour



Summary

- Less dogmatic assumptions $\Longrightarrow$ EnKF-N
  - Posterior variance depends on innovation
  - Better than "unbiased"
  - Sequential feedback
  - Careful about parameterization and implicit assumptions
- Primal perspective: Student t prior
  - Future: include localization
- Dual perspective: scale mixture
  - Adaptive inflation
  - Future: estimate model error inflation

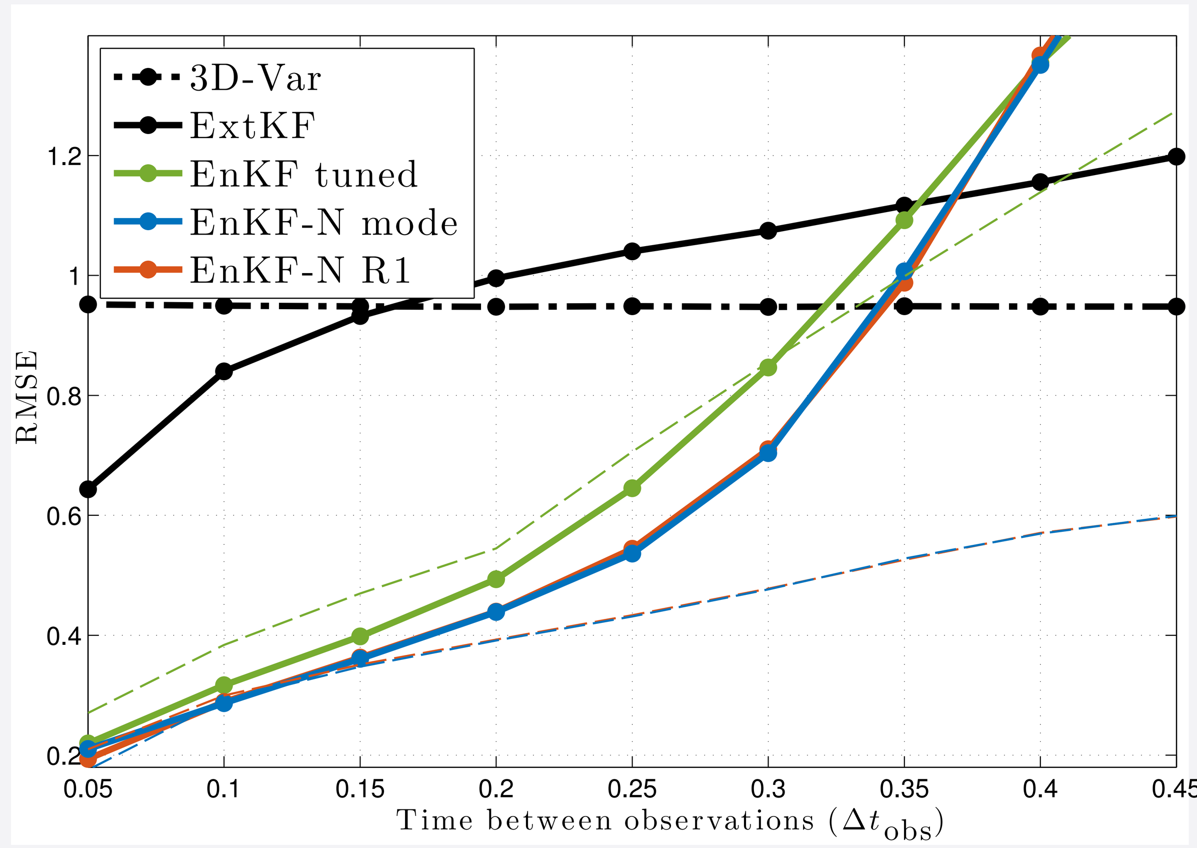
Distributions of λ^2



Prior: $p(\lambda^2|\mathbf{E}) = \chi^{-2}(\lambda^2|N-1)$
Lklhd: $p(\boldsymbol{y}, \boldsymbol{w}_\star|\mathbf{E}, \lambda^2) = \exp\left(-\frac{1}{2}\|\bar{\boldsymbol{\delta}}\|^2_{\mathbf{Y}\mathbf{Y}^{\mathsf{T}}/\zeta + \mathbf{R}}\right)$
Post: $p(\boldsymbol{w}_\star, \zeta|\mathbf{E}, \boldsymbol{y}) = \exp\left(-\frac{1}{2}D(\zeta)\right)$

Benchmarks

Lorenz-96, $N = 20$, no model noise, $\mathbf{R} = \mathbf{I}_{40}$, no localization.



References

[1] Patrick Nima Raanes. *Improvements to Ensemble Methods for Data Assimilation in the Geosciences*. PhD thesis, University of Oxford, Mathematical Institute, 2015. Chapter 6.

[2] Marc Bocquet, Patrick N. Raanes, and Alexis Hannart. Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation. *Nonlinear Processes in Geophysics*, 22(6):645–662, 2015.

Acknowledgements

Statoil