On the impact of dynamics on ensemble data assimilation

Marc Bocquet,
CEREA, joint laboratory École des Ponts ParisTech and EDF R&D, Université Paris-Est, France

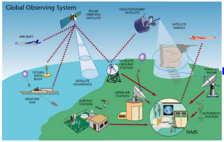
Alberto Carrassi
Nansen Environmental and Remote Sensing Center & Geophysical Institute - Uni. of Bergen, Norway

Colin Grudzien
Dept. of Mathematics and Statistics, University of Nevada in Reno, USA.

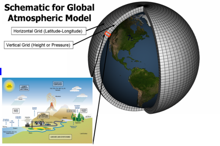
A. Apte (ICTS, India), **K. Gurumoorthy** (Amazon, India), **C.K.R.T. Jones** (University of North Carolina), **A. Fillion** (ENPC/CERFACS, France), **S.Gratton** (IRIT, France)

Outline

Data assimilation (DA) in the geosciences



Data assimilation
best combines
observations and models



An ongoing expansion from numerical weather prediction to the climate science/geosciences:

- Oceanography
- Atmospheric chemistry
- Climate prediction and assessment
- Glaciology
- Hydrology and hydraulics
- Geology
- Space weather
- and many other fields

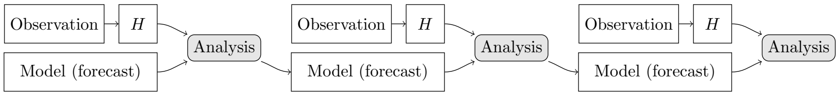
DA as used in climate/atmosphere/ocean

- In the geosciences: Dynamical numerical models are often computationally costly.

- In the geosciences: The state space and observations space are huge (up to $10^9/10^7$ for operational systems, up to $10^7/10^5$ for research systems). A big data problem with costly models to integrate.

- What for?: estimate initial state of chaotic systems for forecasting, re-analysis, parameter estimation ($\sim$ inverse modelling).

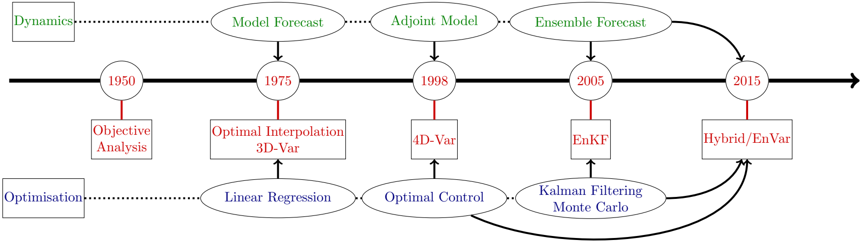
- Data assimilation for forecasting chaotic geofluids: sequential schemes



- This design is the implicit consequence of the unstable dynamics of chaotic geofluids! With this notable expection, DA schemes use models as black boxes.

Mathematical methods in DA

- Introduction of mathematical methods in operational numerical weather prediction:



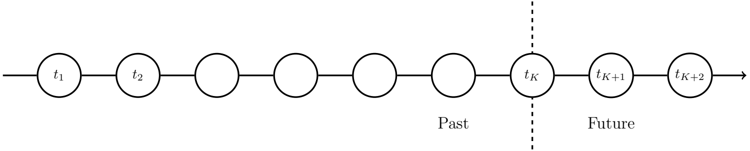
- Using increasingly complex mathematical methods and increasingly resolved high-dimensional models.

Data assimilation system

▶ Data assimilation system = observation and evolution models + statistics of the errors. Typically:

$$\mathbf{x}_k = M_{k:k-1}(\mathbf{x}_{k-1}) + \eta_k$$
$$\mathbf{y}_k = H_k(\mathbf{x}_k) + \varepsilon_k$$

with $\eta_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ and $\varepsilon_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$.



▶ Denoting $\mathbf{x}_{K:1} = \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$, $\mathbf{y}_{K:1} = \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$:

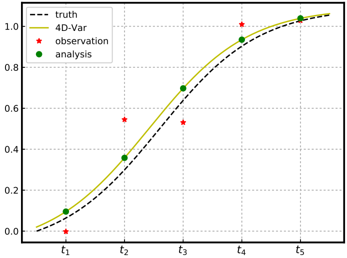
- Prediction: Estimate $\mathbf{x}_k$ for $k > K$, knowing $\mathbf{y}_{K:1}$;
- Filtering: Estimate $\mathbf{x}_K$, knowing $\mathbf{y}_{K:1}$;
- Smoothing: Estimate $\mathbf{x}_{K:1}$, knowing $\mathbf{y}_{K:1}$.

4D-Var (optimal control)

- Strongly constrained 4D-Var, i.e. assuming the model is perfect

$$J(\mathbf{x}_0) = \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}_0^{\mathbf{b}}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\sum_{k=1}^{K}\|\mathbf{y}_k - H_k(\mathbf{x}_k)\|^2_{\mathbf{R}_k^{-1}} + \sum_{k=1}^{K}\lambda_k^\top\left(\mathbf{x}_{k+1} - M_{k+1:k}(x_k)\right).$$

- Fits a model trajectory through the 4D data points.
- In high-dimensional spaces, requires $\nabla_{\mathbf{x}_0} J$ for an efficient minimisation. But $\nabla_{\mathbf{x}_0} J$ depends on the adjoint of $M_{k+1:k}$ and H_k.



- Weakly constrained 4D-Var, i.e. assuming the model is imperfect

$$J(\mathbf{x}_{K:0}) = \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}_0^{\mathbf{b}}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\sum_{k=0}^{K}\|\mathbf{y}_k - H_k(\mathbf{x}_k)\|^2_{\mathbf{R}_k^{-1}} + \frac{1}{2}\sum_{k=1}^{K}\|\mathbf{x}_k - M_{k:k-1}(\mathbf{x}_{k-1})\|^2_{\mathbf{Q}_k^{-1}}.$$

The ensemble Kalman filter (EnKF)

- Mimics the Kalman filter (KF) but replaces the forecast error covariance matrix by

$$\mathbf{P}^{\mathrm{f}} \simeq \mathbf{X}_{\mathrm{f}}\mathbf{X}_{\mathrm{f}}^{\top} \quad \text{where} \quad [\mathbf{X}_{\mathrm{f}}]_i = \frac{\mathbf{x}_{(i)} - \overline{\mathbf{x}}}{\sqrt{m-1}} \quad \text{and} \quad \overline{\mathbf{x}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{x}_{(i)}.$$

- The stochastic EnKF is the closest to traditional KF, but adds stochastic perturbations to the observations of each members [Burgers et al., 1998]:

$$\mathbf{x}^{\mathrm{a}}_{(i)} = \mathbf{x}^{\mathrm{f}}_{(i)} + \mathbf{K}\left(\mathbf{y} + \varepsilon_{(i)} - \mathbf{H}\mathbf{x}^{\mathrm{f}}_{(i)}\right).$$

- The deterministic EnKF avoids the stochasticity by updating the square root of $\mathbf{P}^{\mathrm{f}}$, i.e. $\mathbf{X}_{\mathrm{f}}$. One of the variant (ETKF, [Hunt et al., 2007]) operates the linear algebra in the space of the perturbations ($\mathbf{Y}_{\mathrm{f}} = \mathbf{H}\mathbf{X}_{\mathrm{f}}$):

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\mathbf{w}^{\mathrm{a}} \quad \text{where} \quad \mathbf{w}^{\mathrm{a}} = \left(\mathbf{I}_m + \mathbf{Y}_{\mathrm{f}}^{\top}\mathbf{R}^{-1}\mathbf{Y}_{\mathrm{f}}\right)^{-1}\mathbf{Y}_{\mathrm{f}}^{\top}\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{f}}\right),$$

The perturbations around the mean are updated via

$$\mathbf{X}_{\mathrm{a}} = \mathbf{X}_{\mathrm{f}}\left(\mathbf{I}_m + \mathbf{Y}_{\mathrm{f}}^{\top}\mathbf{R}^{-1}\mathbf{Y}_{\mathrm{f}}\right)^{-\frac{1}{2}}\mathbf{U}, \quad \text{where} \quad \mathbf{U} \in \mathrm{O}(m) \quad \text{and} \quad \mathbf{U}\mathbf{1} = \mathbf{1}.$$

The downside of the EnKF: rank-deficiency

▶ Sampling errors: replacing $\mathbf{P}^{\mathrm{f}}$ by $\mathbf{X}_{\mathrm{f}}\mathbf{X}_{\mathrm{f}}^{\top}$ is in practice rank-deficient and generates spurious correlations for distant state components. If $\mathbf{P} = \mathbf{X}_{\mathrm{f}}\mathbf{X}_{\mathrm{f}}^{\top}$ and $\mathbf{B}$ is the true error covariance matrix of a Gaussian process:

$$\mathrm{Cov}\left([\mathbf{P}]_{ii}, [\mathbf{P}]_{jj}\right) = \frac{2}{N-1}[\mathbf{B}]_{ij}^2, \qquad \mathrm{Cov}\left([\mathbf{P}]_{ij}, [\mathbf{P}]_{ij}\right) = \frac{1}{N-1}\left([\mathbf{B}]_{ij}^2 + [\mathbf{B}]_{ii}[\mathbf{B}]_{jj}\right).$$

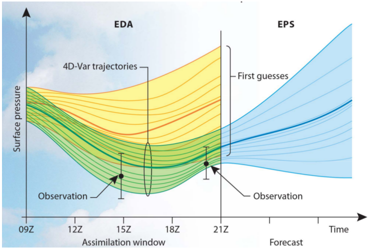
For geophysical systems, we know that most long-range correlations are dampened exponentially. Consequently, the covariances are misestimated (too low variances, too high long-range covariances) and leads to the divergence of the EnKF.
$\longrightarrow$ Practically, this is solved using two fixes: inflation and localisation.

▶ Inflation consists in inflating the covariances by a scalar in the hope to compensate for the underestimation of the error statistics [Pham et al., 1998, Anderson et al., 1999]:

$$\mathbf{x}_{(i)} \longleftarrow \mathbf{x}_{(i)} + \lambda\left(\mathbf{x}_{(i)} - \overline{\mathbf{x}}\right).$$

Hybridising ensemble and variational methods

- A collection of algorithms meant to capture the best of variational and ensemble filtering techniques:
  - Hybrid covariance schemes
  - 4D-LETKF
  - Ensemble of data assimilation (EDA)
  - 4DEnVar
  - IEnKS



- Several of these methods do not require an explicit model adjoint, which is a strong motivation in operations

- Mathematically, an EnVar method such as the IEnKS combines (in addition to avoiding the adjoint):
  - a nonlinear variational analysis (like 4D-Var),
  - a flow-dependent representation of the error (like the EnKF).

The iterative ensemble Kalman smoother (IEnKS)

- Reduced scheme in ensemble space, $\mathbf{x}_0 = \overline{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\mathbf{w}$, where $\mathbf{X}_0$ is the ensemble perturbation matrix:

$$\widetilde{J}(\mathbf{w}) = J(\overline{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\mathbf{w}) .$$

- Analysis IEnKS cost function in ensemble space:

$$\widetilde{J}(\mathbf{w}) = \frac{1}{2}\sum_{k=1}^{L} \|\mathbf{y}_k - H_k \circ M_{k:0}\left(\overline{\mathbf{x}}^{\mathrm{f}} + \mathbf{X}_{\mathrm{f}}\mathbf{w}\right)\|^2_{\beta_k \mathbf{R}_k^{-1}} + \frac{1}{2}(N-1)\|\mathbf{w}\|^2 .$$

$\{\beta_0, \beta_1, \dots, \beta_L\}$ weight the observations impact within the window.

- As a variational reduced method, one can use Gauss-Newton [Sakov et al., 2012], Levenberg-Marquardt [Bocquet & Sakov, 2012], quasi-Newton, trust region, etc., minimisation schemes.

- Perturbation update: same as the ETKF

$$\mathbf{E}_0^{\star} = \mathbf{x}_0^{\star}\mathbf{1}^{\mathrm{T}} + \sqrt{N-1}\mathbf{X}_{\mathrm{f}}\left[\nabla^2_{\mathbf{w}}\widetilde{J}\right]_{\star}^{-1/2}\mathbf{U} \quad \text{where} \quad \mathbf{U} \in \mathrm{O}(N) \quad \text{and} \quad \mathbf{U}\mathbf{1} = \mathbf{1} .$$

Recent textbooks/reviews in DA

- A.J. Majda and J. Harlim, *Filtering complex turbulent systems*, Cambridge University Press, 2012.

- S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian DataAssimilation*, Cambridge University Press, 2015.

- K. Law et al., *Data Assimilation – A Mathematical Introduction,Springer*, 2015.

- M. Asch et al., *Data Assimilation: Methods, Algorithms, and Applications*, SIAM, 2016.

- S.J. Fletcher. *Data Assimilation for the Geosciences: From Theory to Application*, Elsevier, 2017.

- A. Carrassi et al., *Data Assimilation in the Geosciences: An overview on methods, issues, and perspectives*, WIREs Climate Change, 9, e535, 2018.

Outline

Data assimilation and the dynamics

- So far, the model has (essentially) been considered as a black box.

- The atmosphere and ocean exhibit chaotic dissipative dynamics: Highly state-dependent error growth. DA must track and incorporate this flow-dependency in the quantification of the uncertainty (i.e. error covariances).

- Dissipation induces dimensional reduction: The error dynamics are confined to a subspace of much smaller dimension, $n_0 \ll m$: the unstable subspace. The existence of the underlying unstable-stable splitting of the phase space expected to have critical impact on the efficiency and accuracy of DA.

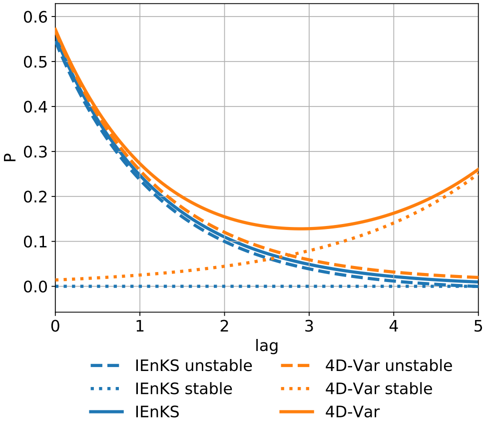
$\longrightarrow$ A set of ideas put forward and initially developed by Anna Trevisan et al. [Trevisan et al. 2004-2015; Palatella et al., 2013], and called AUS (assimilation in the unstable subspace).

Motivations

- Is there any fingerprint of the unstable subspace on the fate of the (En)KF and the (En)KS?
- Understand the interaction between DA and the dynamics.
- Can dynamical properties be used to design computationally cheap DA schemes?

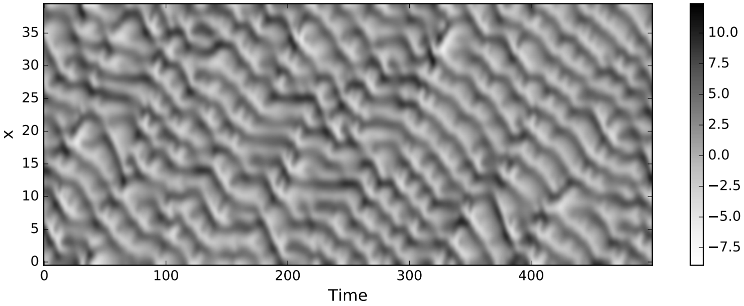
DA and the dynamics: the linear and scalar case

▶ Analytical formulae for the forecast and analysis variances can be obtained in the linear, diagonal dynamics case [Fillion et al. 2018]



▶ 4D-Var is impacted by its imperfect representations of the error stable modes as opposed to the IEnKS [Talagrand et al., 2010; Fillion et al. 2018].

Nonlinear chaotic models: the Lorenz-96 low-order model



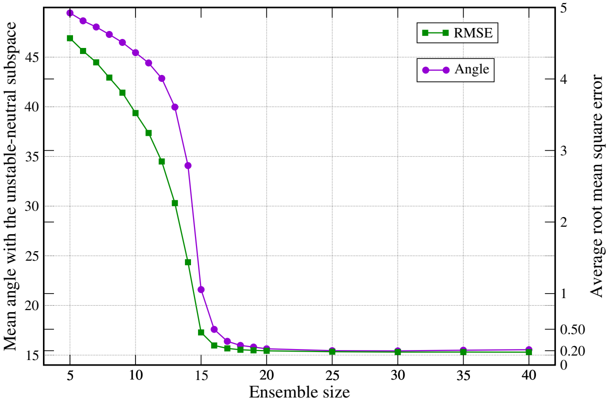
- It represents a mid-latitude zonal circle of the global atmosphere.
- Set of $M = 40$ ordinary differential equations [Lorenz and Emmanuel 1998]:

$$\frac{\mathrm{d}x_m}{\mathrm{d}t} = (x_{m+1} - x_{m-2})x_{m-1} - x_m + F\,, \tag{1}$$

where $F = 8$, and the boundary is cyclic.

- Conservative system except for a forcing term F and a dissipation term $-x_m$.
- Chaotic dynamics, 13 positive and 1 neutral Lyapunov exponents, a doubling time of about 0.42 time units.

Nonlinear chaotic model



▶ Average angle (in degrees) between a perturbation (from the ensemble) and the unstable-neutral subspace as a function of the DAW length (IEnKS, Lorenz-96), as well as the corresponding RMSE of the analysis.

Outline

Linear case: Degenerate Kalman filter equations

- Model dynamics and observation model:

$$\mathbf{x}_k = \mathbf{M}_k \mathbf{x}_{k-1} + \mathbf{w}_k, \tag{2}$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k. \tag{3}$$

The model and observation noises, $\mathbf{w}_k$ and $\mathbf{v}_k$, are assumed mutually independent, unbiased Gaussian white sequences with statistics

$$\mathrm{E}[\mathbf{v}_k \mathbf{v}_l^\top] = \delta_{k,l} \mathbf{R}_k, \quad \mathrm{E}[\mathbf{w}_k \mathbf{w}_l^\top] = \delta_{k,l} \mathbf{Q}_k, \quad \mathrm{E}[\mathbf{v}_k \mathbf{w}_l^\top] = \mathbf{0}. \tag{4}$$

- Forecast error covariance matrix $\mathbf{P}_k$ recurrence of the Kalman filter (KF)

$$\mathbf{P}_{k+1} = \mathbf{M}_{k+1} \left(\mathbf{I} + \mathbf{P}_k \mathbf{\Omega}_k\right)^{-1} \mathbf{P}_k \mathbf{M}_{k+1}^\top + \mathbf{Q}_{k+1}, \tag{5}$$

where

$$\mathbf{\Omega}_k \equiv \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \tag{6}$$

are the precision matrices and $\mathbf{P}_0$ can be of arbitrary rank.

- In the case $\mathbf{Q}_k \equiv \mathbf{0}$, it was proven that the full-rank KF $\mathbf{P}_k$ collapses onto the unstable subspace [Gurumoorthy at al. 2017].
- Still in the case $\mathbf{Q}_k \equiv \mathbf{0}$, it will be generalised in the following and for degenerate $\mathbf{P}_0$ required to connect to reduced-order methods such as the ensemble Kalman filter (EnKF) [Bocquet at al. 2017].

Result 1: Bound of the covariance free forecast

- Simple inequality in the set of the semi-definite symmetric matrices

$$\mathbf{P}_k \leq \mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^{\top} + \mathbf{\Xi}_k. \tag{7}$$

where

$$\mathbf{\Xi}_0 \equiv \mathbf{0} \quad \text{and for } k \geq 1 \quad \mathbf{\Xi}_k \equiv \sum_{l=1}^{k} \mathbf{M}_{k:l}\mathbf{Q}_l\mathbf{M}_{k:l}^{\top} \tag{8}$$

is known as the *controllability* matrix [Jazwinski, 1970].

- In the absence of model noise ($\mathbf{Q}_k \equiv \mathbf{0}$ for the rest of this talk), it reads

$$\mathbf{P}_k \leq \mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^{\top}. \tag{9}$$

Assuming the dynamics is non-singular

$$\mathrm{Im}(\mathbf{P}_k) = \mathbf{M}_{k:0}\left(\mathrm{Im}(\mathbf{P}_0)\right). \tag{10}$$

If n_0 is the dimension of the unstable-neutral subspace, it can further be shown that

$$\lim_{k\to\infty} \mathrm{rank}(\mathbf{P}_k) \leq \min\left\{\mathrm{rank}(\mathbf{P}_0), n_0\right\}. \tag{11}$$

Result 2: Collapse onto the unstable subspace

- Let σ_i^k, for $i = 1, \ldots, n$ denote the eigenvalues of $\mathbf{P}_k$ ordered as $\sigma_1^k \geq \sigma_2^k \cdots \geq \sigma_n^k$. We can show that

$$\sigma_i^k \leq \alpha_i \exp\left(2k\lambda_i^k\right) \tag{12}$$

where $k\lambda_i^k$ is a log-singular value of $\mathbf{M}_{k:0}$ and $\lim_{k\to\infty} \lambda_i^k = \lambda_i$. This gives an upper bound for all eigenvalues of $\mathbf{P}_k$ and a rate of convergence for the $n - n_0$ smallest ones.

- If $\mathbf{P}_k$ is uniformly bounded, it can further be shown that the stable subspace of the dynamics is asymptotically in the null space of $\mathbf{P}_k$, *i.e.* for any vector $\mathbf{u}_{k:0}$ in the stable subspace

$$\lim_{k\to\infty} \|\mathbf{P}_k \mathbf{u}_{k:0}\| = 0. \tag{13}$$

Result 3: Explicit dependence of $\mathbf{P}_k$ on $\mathbf{P}_0$

▶ Using either analytic continuation or the symplectic symmetry of the linear representation of covariances, we have proven that

$$\mathbf{P}_k = \mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^\top\left(\mathbf{I}+\mathbf{\Gamma}_k\mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^\top\right)^{-1}. \tag{14}$$

where

$$\mathbf{\Gamma}_k \equiv \sum_{l=0}^{k-1}\mathbf{M}_{k:l}^{-\top}\mathbf{\Omega}_l\mathbf{M}_{k:l}^{-1}. \tag{15}$$

▶ An alternative is

$$\mathbf{P}_k = \mathbf{M}_{k:0}\mathbf{P}_0\left[\mathbf{I}+\mathbf{\Theta}_k\mathbf{P}_0\right]^{-1}\mathbf{M}_{k:0}^\top \tag{16}$$

where

$$\mathbf{\Theta}_k \equiv \mathbf{M}_{k:0}^\top\mathbf{\Gamma}_k\mathbf{M}_{k:0} = \sum_{l=0}^{k-1}\mathbf{M}_{l:0}^\top\mathbf{\Omega}_l\mathbf{M}_{l:0}. \tag{17}$$

is the *information* matrix, directly related to the observability of the DA system.

Result 4: Asymptotics of $\mathbf{P}_k$

- Questions: Under which conditions does $\mathbf{P}_k$ forget about $\mathbf{P}_0 = \mathbf{X}_0\mathbf{X}_0^\top$? Can we analytically compute its asymptotics?

- We proposed a sufficient set of conditions

  - **Condition 1:** Assume the forward Lyapunov vectors at t_0 associated to the unstable and neutral directions are the columns of $\mathbf{V}_{+,0} \in \mathbb{R}^{n\times n_0}$. The condition reads

$$\mathrm{rank}\left(\mathbf{X}_0^\top \mathbf{V}_{+,0}\right) = n_0. \tag{18}$$

  - **Condition 2:** The model is sufficiently observed so that the unstable and neutral directions remain under control, that is

$$\mathbf{U}_{+,k}^\top \mathbf{\Gamma}_k \mathbf{U}_{+,k} > \varepsilon \mathbf{I} \tag{19}$$

    where $\mathbf{U}_{+,k}$ is a matrix whose columns are the backward Lyapunov vectors related to non-negative exponents and $\varepsilon > 0$ is a positive number.

  - **Condition 3:** For any neutral backward Lyapunov vector $\mathbf{u}_k$, we have

$$\lim_{k\to\infty} \mathbf{u}_k^\top \mathbf{\Gamma}_k \mathbf{u}_k = \infty, \tag{20}$$

    i.e. the neutral modes should be sufficiently observed.

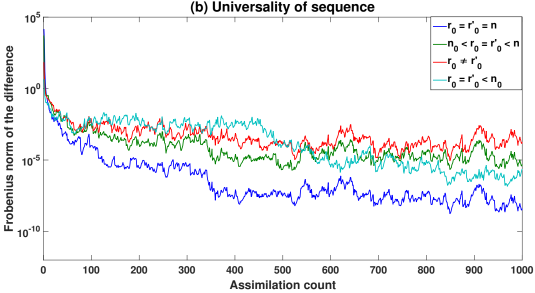
Result 4: Asymptotics of $\mathbf{P}_k$

Under these three conditions, we obtain

$$\lim_{k\to\infty}\left\{\mathbf{P}_k - \mathbf{U}_{+,k}\left[\mathbf{U}_{+,k}^\top \mathbf{\Gamma}_k \mathbf{U}_{+,k}\right]^{-1}\mathbf{U}_{+,k}^\top\right\} = \mathbf{0}. \tag{21}$$

The asymptotic sequence does not depend on $\mathbf{P}_0$, only $\mathbf{\Gamma}_k$!

- Peculiar role of the neutral modes (arithmetic convergence).
- Numerical illustration and verification



Linearized Lorenz-96 model around a Lorenz-96 trajectory.

Frobenius norm of the difference between two different $\mathbf{P}_0$ when the conditions are satisfied, *i.e.* $\|\mathbf{P}_k^{\mathrm{a}} - \mathbf{P}_k^{\prime\mathrm{a}}\|$.

From the degenerate KF to the square-root EnKF

- Normalised perturbation decomposition:

$$\mathbf{P}_k = \mathbf{X}_k \mathbf{X}_k^\top . \tag{22}$$

- Square-root formulation; right-transform update formula:

$$\mathbf{X}_k = \mathbf{M}_{k:0}\mathbf{X}_0 \left[\mathbf{I} + \mathbf{X}_0^\top \boldsymbol{\Theta}_k \mathbf{X}_0\right]^{-1/2} \boldsymbol{\Psi}_k , \tag{23}$$

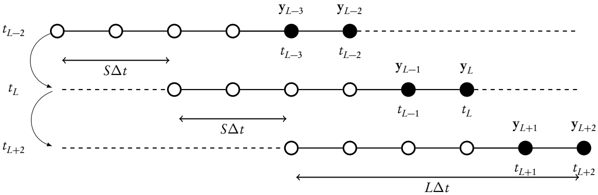
where $\boldsymbol{\Psi}_k$ is an orthogonal matrix.

- Square-root formulation; left-transform update formula:

$$\mathbf{X}_k = \left[\mathbf{I} + \mathbf{M}_{k:0}\mathbf{P}_0\mathbf{M}_{k:0}^\top \boldsymbol{\Gamma}_k\right]^{-1/2} \mathbf{M}_{k:0}\mathbf{X}_0\boldsymbol{\Psi}_k . \tag{24}$$

- With linear models, Gaussian observation and initial errors, the (square-root) degenerate KF is equivalent to the square-root EnKF and can serve as a proxy to the EnKF applied to nonlinear models.

Degenerate square root Kalman smoother



- The scheme at a glance, variational correspondence ($\mathbf{x} = \overline{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}$) :

$$\widetilde{\mathscr{J}}(\mathbf{w}) = \frac{1}{2}\sum_{l=k+L-S+1}^{k+L} \left\|\mathbf{y}_l - \mathbf{H}_l\mathbf{M}_{l:k}\left(\overline{\mathbf{x}}_k + \mathbf{X}_k\mathbf{w}\right)\right\|^2_{\mathbf{R}_l} + \frac{1}{2}\left\|\mathbf{w}\right\|^2$$

- From the Hessian of $\widetilde{\mathscr{J}}$,

$$\mathbf{I}_N + \mathbf{X}_k^\top \widehat{\mathbf{\Omega}}_k \mathbf{X}_k \qquad \text{where} \quad \widehat{\mathbf{\Omega}}_k \triangleq \sum_{l=k+L-S+1}^{k+L} \mathbf{M}_{l:k}^\top \mathbf{\Omega}_l \mathbf{M}_{l:k}\,,$$

we infer

$$\mathbf{X}_{k+S} = \mathbf{M}_{k+S:k}\mathbf{X}_k\left(\mathbf{I}_N + \mathbf{X}_k^\top \widehat{\mathbf{\Omega}}_k \mathbf{X}_k\right)^{-\frac{1}{2}}\mathbf{\Psi}_k\,.$$

Degenerate square root Kalman smoother

▶ The convergence rate of the collapse of $\mathbf{P}_k$ of the smoother is not expected to be faster than the filter's: the bounding rate is the same.

▶ However the accuracy of the smoother for re-analysis is expected to be better which should impact the asymptotic sequences. Indeed we have, for $k = pS$, $p = 0, 1, \ldots$:

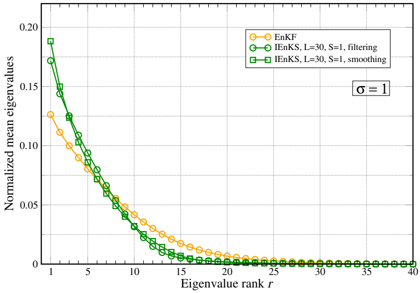
$$\lim_{k\to\infty} \left\{ \mathbf{X}_k - \mathbf{U}_{+,k} \left[\mathbf{U}_{+,k}^\top \widehat{\mathbf{\Gamma}}_k \mathbf{U}_{+,k} \right]^{-\frac{1}{2}} \mathbf{\Psi}_k \right\} = \mathbf{0} .$$

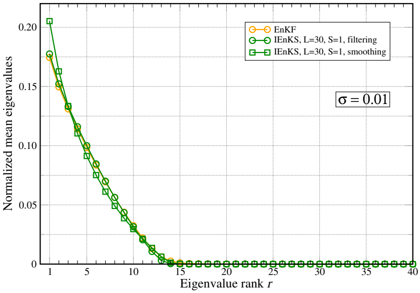
▶ The only difference is in the observability matrix $\widehat{\mathbf{\Gamma}}_k$, for $k = pS$, $p = 0, 1, \ldots$:

$$\widehat{\mathbf{\Gamma}}_k = \mathbf{\Gamma}_k + \sum_{l=k}^{k+L-S} \mathbf{M}_{k:l}^{-\top} \mathbf{\Omega}_l \mathbf{M}_{k:l}^{-1} .$$

Outline

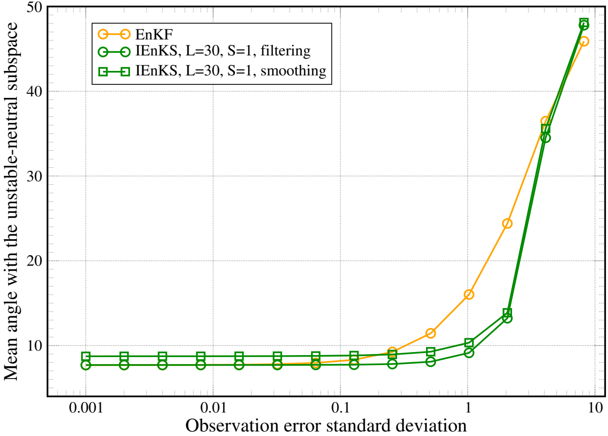
Spectrum of the analysis error covariance matrix





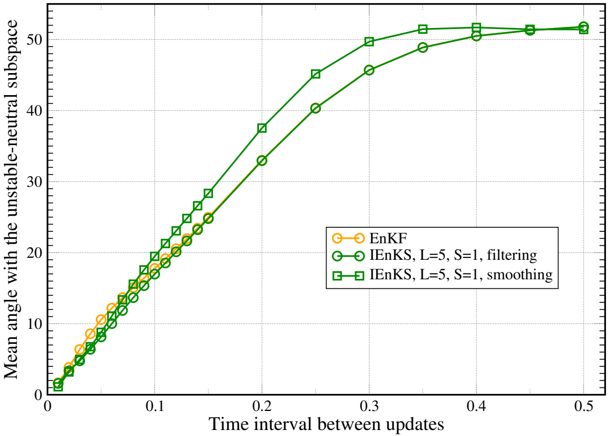
- Time-average spectra of $\mathbf{P}^{\mathrm{a}}_k$: A visible transition at $r = 15$.

Nonlinear chaotic model



▶ Average angle (in degrees) between a perturbation (from the ensemble) and the unstable-neutral subspace as a function of the observation error (EnKF and IEnKS, Lorenz-96, $\Delta t = 0.05$, $\mathbf{R} = \sigma^2 \mathbf{I}$ $N = 20$).

Nonlinear chaotic model



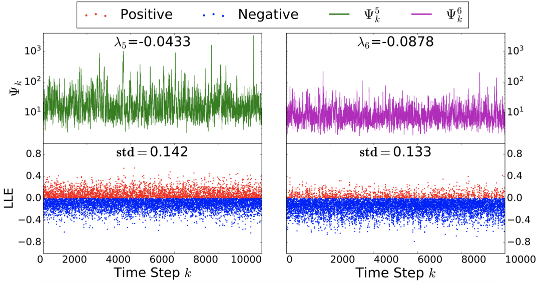
▶ Average angle (in degrees) between a perturbation (from the ensemble) and the unstable-neutral subspace as a function of the interval between updates (EnKF and IEnKS, Lorenz-96, $\mathbf{R} = 10^{-4}\mathbf{I}$, $N = 20$).

Outline

Error in stochastic models: role of the instabilities?

$$\mathbf{x}_k = M_{k:k-1}(\mathbf{x}_{k-1}) + \eta_k, \qquad \eta_k \in \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$$

- Asymptotic uncertainty in the stable BLVs no longer zero, but still bounded.
- However, the error bounds depend on [Grudzien et al. 2018a]
(i) the model error size (*i.e.* $||\mathbf{Q}||$), and (ii) the variance of the local LEs (LLEs).



$m = 10$ and $n_0 = 4$

- If the noise is large and/or the LLEs have high variance, the bounds will be impractically large.
- In noisy systems it is necessary to include weakly stable BLVs of high variance.

Error in stochastic models: an upwelling effect

▶ Will the *necessary* increase $\mathbf{N} = \mathbf{n_0} \to \mathbf{n_0} + \mathbf{n_{ws}}$ also be *sufficient*?

▶ Write the model propagator in the basis of the BLVs using the recursive QR decomposition

$$\mathbf{M}_k = \mathbf{E}_k \mathbf{U}_k \mathbf{E}_k^{\mathrm{T}}, \quad \mathbf{E}_k = (\mathbf{E}_k^{\mathrm{f}}\, \mathbf{E}_k^{\mathrm{u}}) \text{ with } \mathbf{U}_k = \begin{pmatrix} \mathbf{U}_k^{\mathrm{ff}} & \mathbf{U}_k^{\mathrm{fu}} \\ 0 & \mathbf{U}_k^{\mathrm{uu}} \end{pmatrix}$$

and partition the error into **filtered**/**unfiltered** variables $\varepsilon_k = \mathbf{E}_k^{\mathrm{f}} \varepsilon_k^{\mathrm{f}} + \mathbf{E}_k^{\mathrm{u}} \varepsilon_k^{\mathrm{u}}$

▶ The error in the filtered space ("seen" by DA) is given recursively by [Grudzien et al. 2018b]

$$\varepsilon_{k+1}^{\mathrm{f}} = (\mathbf{U}_{k+1}^{\mathrm{ff}} - \mathbf{U}_{k+1}^{\mathrm{ff}} \mathbf{K}_k \mathbf{H}_k \mathbf{E}_k^{\mathrm{f}}) \varepsilon_k^{\mathrm{f}} - \mathbf{U}_{k+1}^{\mathrm{ff}} \mathbf{K}_k \varepsilon_k^{\mathrm{obs}} + \eta_k^{\mathrm{f}} + (\mathbf{U}_{k+1}^{\mathrm{fu}} - \mathbf{U}_{k+1}^{\mathrm{ff}} \mathbf{K}_k \mathbf{H}_k \mathbf{E}_k^{\mathrm{u}}) \varepsilon_k^{\mathrm{u}}$$

▶ **The terms in black** correspond to the usual KF-like recursion and highlight the stabilizing effect of DA [Carrassi et al. 2008].

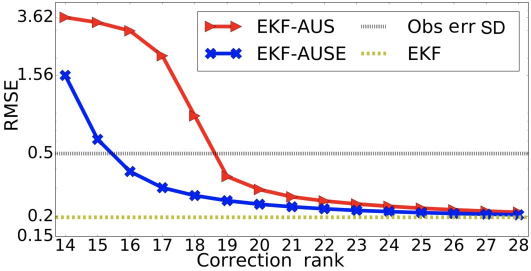
▶ **The terms in red** disappear when the filtered subspace is the entire state space ($n = m$).

Error in stochastic models: an upwelling effect

▶ When $n < m$, they represent the **dynamical upwelling** of the unfiltered error into the filtered variables [Grudzien et al. 2018b].

▶ This phenomenon **occurs whenever** $n < m$, but is **exacerbated by stochastic noise**.

▶ Leads to underestimating the error in the (En)KF $\Rightarrow$ Inflation required



- **EKF** solves the *full-rank* recursion.
- **EKF-AUS** solves the *low-rank* recursion **without upwelling** (black terms only).
- **EKF-AUSE** solves the *low-rank* recursion **with upwelling** (black+red terms).

Outline

Conclusions and more

- We have shown that, in deterministic/noiseless dynamics, the (En)KF/(En)KS and their iterative variants naturally project the uncertainty on the unstable-neutral subspace $\Rightarrow N = n_0$ members are sufficient.

- This shows that the EnKF/EnKS naturally implement the AUS program (without expliciting the Lyapunov filtration).

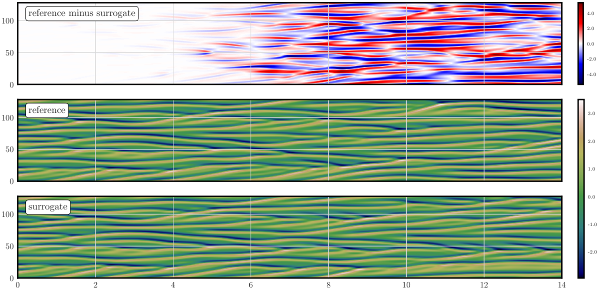
- In stochastic/noisy dynamics, weakly stable modes of high variance must be included. Furthermore we have demonstrated the existence of an upwelling of uncertainty from unfiltered-to-filtered subspace that motivates the need for multiplicative inflation.

- Much more on the topic in the minisymposia MS172 Data and Dynamics: Dynamical Systems Techniques in Data Assimilation - Part I & II, this afternoon, Ballroom 1.

Conclusions and more

▶ This was the state of the art 2 years ago about DA and DS . . .

▶ Since then, machine learning made its way to data assimilation, and a new hot topic is the convergence of DA, DS and ML. For instance, DA could be used to infer the ODEs or PDEs of dynamical systems from partial and noisy observations [Bocquet et al., 2019], or use deep learning in combinaison with DA [Brajard et al., 2019].



Time (Lyapunov unit)

Many open questions: How many required dof in the surrogate model? Ergodic properties of the surrogate models? Numerical stability (stiffness)? Can it be used as a substitute for the model in DA schemes?

References

[1] M. Bocquet, J. Brajard, A. Carrassi, and L. Bertino, *Data assimilation as a deep learning tool to infer ode representations of dynamical models*, Nonlin. Processes Geophys. Discuss., 2019 (2019), pp. 1–29.

[2] M. Bocquet and A. Carrassi, *Four-dimensional ensemble variational data assimilation and the unstable subspace*, Tellus A, 69 (2017), p. 1304504.

[3] M. Bocquet, K. S. Gurumoorthy, A. Apte, A. Carrassi, C. Grudzien, and C. K. R. T. Jones, *Degenerate Kalman filter error covariances and their convergence onto the unstable subspace*, SIAM/ASA J. Uncertainty Quantification, 5 (2017), pp. 304–333.

[4] J. Brajard, A. Carrassi, M. Bocquet, and L. Bertino, *Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: a case study with the Lorenz 96 model*, Geosci. Model Dev. Discuss., 2019 (2019), pp. 1–21.

[5] A. Fillion, M. Bocquet, and S. Gratton, *Quasi static ensemble variational data assimilation: a theoretical and numerical study with the iterative ensemble Kalman smoother*, Nonlin. Processes Geophys., 25 (2018), pp. 315–334.

[6] C. Grudzien, A. Carrassi, and M. Bocquet, *Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error*, SIAM/ASA J. Uncertainty Quantification, 6 (2018), p. 1335–1363.

[7] C. Grudzien, A. Carrassi, and M. Bocquet, *Chaotic dynamics and the role of covariance inflation for reduced rank Kalman filters with model error*, Nonlin. Processes Geophys., 25 (2018), pp. 633–648.

[8] K. S. Gurumoorthy, C. Grudzien, A. Apte, A. Carrassi, and C. K. R. T. Jones, *Rank deficiency of Kalman error covariance matrices in linear time-varying system with deterministic evolution*, SIAM Journal on Control and Optimization, 55 (2017), pp. 741–759.

[9] G.-H. C. Ng, D. McLaughlin, D. Entekhabi, and A. Ahanin, *The role of model dynamics in ensemble Kalman filter performance for chaotic systems*, Tellus A, 63 (2011), pp. 958–977.

[10] L. Palatella, A. Carrassi, and A. Trevisan, *Lyapunov vectors and assimilation in the unstable subspace: theory and applications*, J. Phys. A: Math. Theor., 46 (2013), p. 254020.

[11] L. Palatella and A. Trevisan, *Interaction of Lyapunov vectors in the formulation of the nonlinear extension of the Kalman filter*, Phys. Rev. E, 91 (2015), p. 042905.

[12] A. Trevisan, M. D'Isidoro, and O. Talagrand, *Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension*, Q. J. R. Meteorol. Soc., 136 (2010), pp. 487–496.