Optimal linearization trajectories

Roel Stappers and Jan Barkmeijer

stappers@knmi.nl

The 9^{th} ADJOINT WORKSHOP

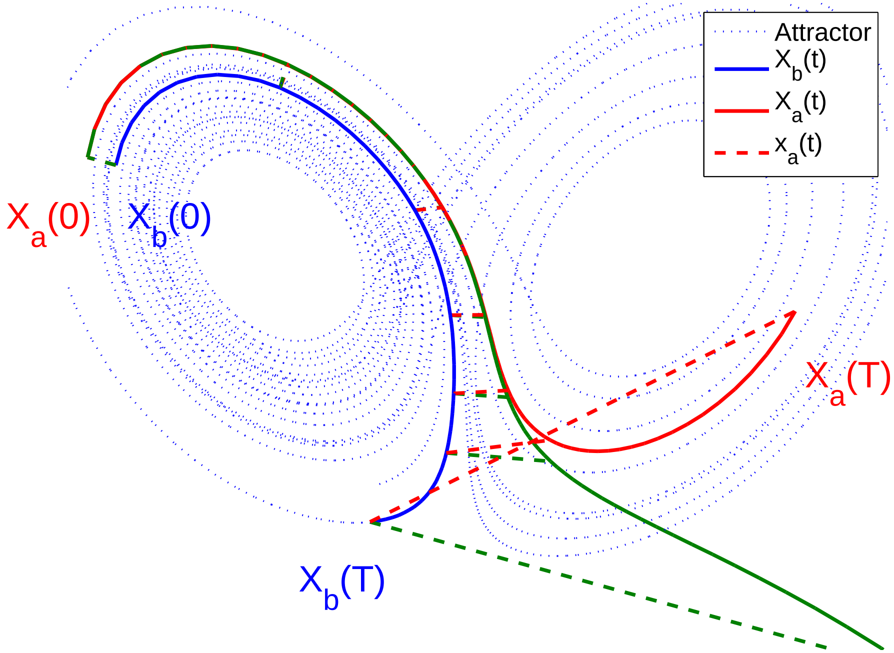
Cefalu, Sicily, Italy

10-14 October 2011

Overview

- Notation/Introduction
- Nonlinear differential equations (advection)
- Optimal linearization trajectories
- Iterative relinearization
- Forecast sensitivity
- Concluding remarks

Notation



- Nonlinear trajectory: $\mathbf{X}_b(t) = \mathcal{M}(\mathbf{X}_b(0), t, 0)$
- Nonlinear trajectory: $\mathbf{X}_a(t) = \mathcal{M}(\mathbf{X}_a(0), t, 0)$
- Increment trajectory: $\mathbf{x}_a(t) = \mathbf{X}_a(t) - \mathbf{X}_b(t)$
- Tangent linear trajectory: $\hat{\mathbf{x}}_a(t) = \mathbf{M}_{\mathbf{X}_b}(t, 0)\mathbf{x}_a(0)$

Use of linear models in NWP

- Variational data assimilation (4D-VAR)
- Ensemble prediction (Singular vectors)
- Diagnostics (Adjoint based observation impact)

Use of linear models in NWP

- Variational data assimilation (4D-VAR)
- Ensemble prediction (Singular vectors)
- Diagnostics (Adjoint based observation impact)

Main limitation

- Linear models are only useful in the “short” range

Use of linear models in NWP

- Variational data assimilation (4D-VAR)
- Ensemble prediction (Singular vectors)
- Diagnostics (Adjoint based observation impact)

Main limitation

- Linear models are only useful in the “short” range

(Research question)

- Can we exactly simulate the nonlinear growth of perturbations with the linear model?

Nonlinear differential equations

Consider the differential equation

$$\dot{\mathbf{X}} = q(\mathbf{X}, \mathbf{X}) + b(\mathbf{X}) + c,$$

where q is linear in both arguments, b is linear and c is a (time dependent) forcing.
In NWP $q(\mathbf{X}, \mathbf{X})$ is typically a result of the advection part of the total derivative, e.g.

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} \qquad \frac{D\phi}{Dt} = \frac{\partial \phi}{\partial t} + (\mathbf{u} \cdot \nabla)\phi$$

Nonlinear differential equations

Consider the differential equation

$$\dot{\mathbf{X}} = q(\mathbf{X}, \mathbf{X}) + b(\mathbf{X}) + c,$$

where q is linear in both arguments, b is linear and c is a (time dependent) forcing.

- Initial condition 1 (background): $\mathbf{X}_b(0)$
- Initial condition 2 (analysis): $\mathbf{X}_a(0)$
- Corresponding trajectories $\mathbf{X}_b(t)$ and $\mathbf{X}_a(t)$
- Define $\mathbf{X}_a(t) = \mathbf{X}_b(t) + \mathbf{x}_a(t)$

Derivation of the Opt. lin. trajectory

The analysis trajectory satisfies

$$\dot{\mathbf{X}}_a = q(\mathbf{X}_a, \mathbf{X}_a) + b(\mathbf{X}_a) + c$$

Derivation of the Opt. lin. trajectory

The analysis trajectory satisfies

$$\dot{\mathbf{X}}_a = q(\mathbf{X}_a, \mathbf{X}_a) + b(\mathbf{X}_a) + c$$

Substitution of $\mathbf{X}_a(t) = \mathbf{X}_b(t) + \mathbf{x}_a(t)$ gives

$$\dot{\mathbf{X}}_b + \dot{\mathbf{x}}_a = q(\mathbf{X}_b + \mathbf{x}_a, \mathbf{X}_b + \mathbf{x}_a) + b(\mathbf{X}_b + \mathbf{x}_a) + c$$

Derivation of the Opt. lin. trajectory

The analysis trajectory satisfies

$$\dot{\mathbf{X}}_a = q(\mathbf{X}_a, \mathbf{X}_a) + b(\mathbf{X}_a) + c$$

Substitution of $\mathbf{X}_a(t) = \mathbf{X}_b(t) + \mathbf{x}_a(t)$ gives

$$\dot{\mathbf{X}}_b + \dot{\mathbf{x}}_a = q(\mathbf{X}_b + \mathbf{x}_a, \mathbf{X}_b + \mathbf{x}_a) + b(\mathbf{X}_b + \mathbf{x}_a) + c$$

Using bilinearity of q and linearity of b and

$$\dot{\mathbf{X}}_b = q(\mathbf{X}_b, \mathbf{X}_b) + b(\mathbf{X}_b) + c$$

to eliminate $\dot{\mathbf{X}}_b$
gives the exact time evolution for perturbations

$$\boxed{\dot{\mathbf{x}}_a = q(\mathbf{X}_b, \mathbf{x}_a) + q(\mathbf{x}_a, \mathbf{X}_b) + b(\mathbf{x}_a) + q(\mathbf{x}_a, \mathbf{x}_a)}$$

Derivation of the Opt. lin. trajectory

From the previous slide

$$\dot{\mathbf{x}}_a = \underbrace{q(\mathbf{X}_b, \mathbf{x}_a) + q(\mathbf{x}_a, \mathbf{X}_b) + b(\mathbf{x}_a)}_{\mathbf{Df}(\mathbf{X}_b)\mathbf{x}_a} + q(\mathbf{x}_a, \mathbf{x}_a) \tag{1}$$

Retaining only terms linear in $\mathbf{x}_a$ gives the familiar tangent linear (TL) model

$$\dot{\hat{\mathbf{x}}}_a = \mathbf{Df}(\mathbf{X}_b)\hat{\mathbf{x}}_a \tag{2}$$

Here $\mathbf{Df}(\mathbf{X}_b)$ is the Jacobian evaluated along the trajectory $\mathbf{X}_b(t)$.

The key point of this presentation

From the previous slide

$$\dot{\mathbf{x}}_a = \underbrace{q(\mathbf{X}_b, \mathbf{x}_a) + q(\mathbf{x}_a, \mathbf{X}_b) + b(\mathbf{x}_a)}_{\mathbf{Df}(\mathbf{X}_b)\mathbf{x}_a} + q(\mathbf{x}_a, \mathbf{x}_a) \qquad (3)$$

The key observation is that this can be written as

$$\boxed{\dot{\mathbf{x}}_a = \mathbf{Df}(\mathbf{X}_b + \mathbf{x}_a/2)\mathbf{x}_a}$$

i.e. we obtain the exact time evolution of perturbations in the TL-model if we integrate around the trajectory $\mathbf{X}_b + \mathbf{x}_a/2$ instead of $\mathbf{X}_b$.

The key point of this presentation

From the previous slide

$$\dot{\mathbf{x}}_a = \underbrace{q(\mathbf{X}_b, \mathbf{x}_a) + q(\mathbf{x}_a, \mathbf{X}_b) + b(\mathbf{x}_a)}_{\mathbf{Df}(\mathbf{X}_b)\mathbf{x}_a} + q(\mathbf{x}_a, \mathbf{x}_a) \tag{4}$$

The key observation is that this can be written as

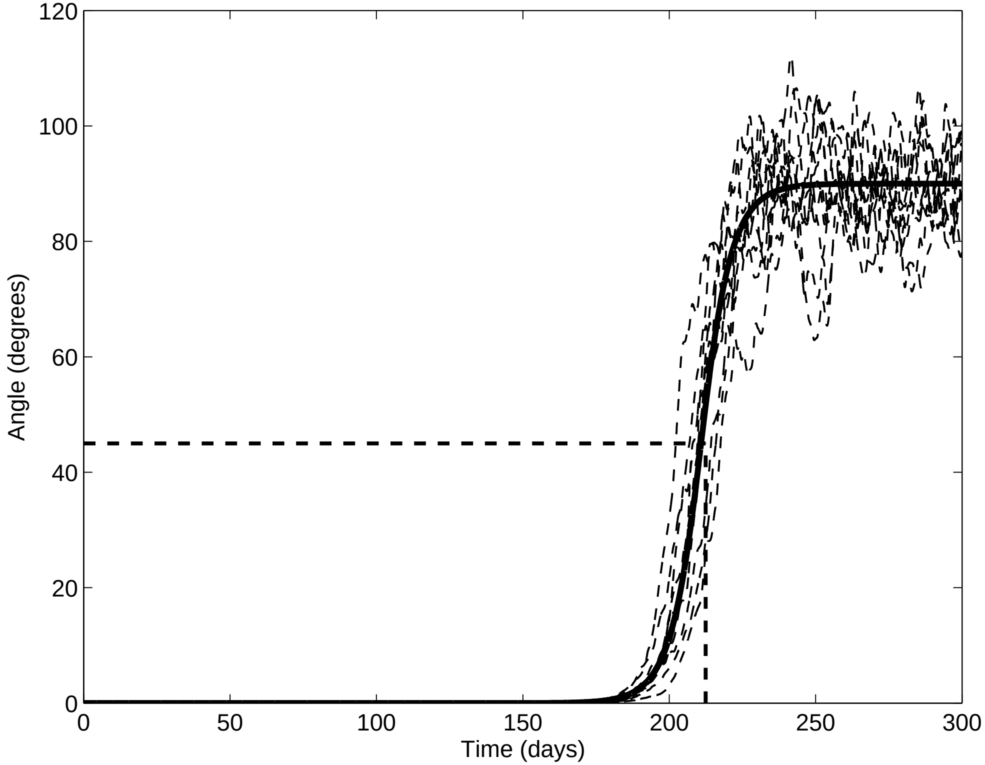
$$\boxed{\dot{\mathbf{x}}_a = \mathbf{Df}(\mathbf{X}_b + \mathbf{x}_a/2)\mathbf{x}_a}$$

i.e. we obtain the exact time evolution of perturbations in the TL-model if we integrate around the trajectory $\mathbf{X}_b + \mathbf{x}_a/2$ instead of $\mathbf{X}_b$. In integral form

$$\boxed{\mathbf{x}_a(t) = \mathbf{M}_{\mathbf{X}_b+\mathbf{x}_a/2}(t, 0)\mathbf{x}_a(0)}$$

Opt. lin. trajectories in the QG model

Quasi geostrophic model (Marshall and Molteni, 1993)



$$
\begin{aligned}
\mathbf{x}_a(t) &= \mathbf{X}_a(t) - \mathbf{X}_b(t) \\
\hat{\mathbf{x}}_a(t) &= \mathbf{M}_{\mathbf{X}_b + \mathbf{x}_a/2}(t,0)\mathbf{x}_a(0)
\end{aligned}
$$

The relinearization method

Exact time evolution of perturbations

$$\boxed{\mathbf{x}_a = \mathbf{M}_{\mathbf{X}_b + \mathbf{x}_a/2} \mathbf{x}_a(0)}$$

For a given estimate of the increment trajectory $\hat{\mathbf{x}}_a^{k-1}$ the TL-model can be used in the following iterative procedure

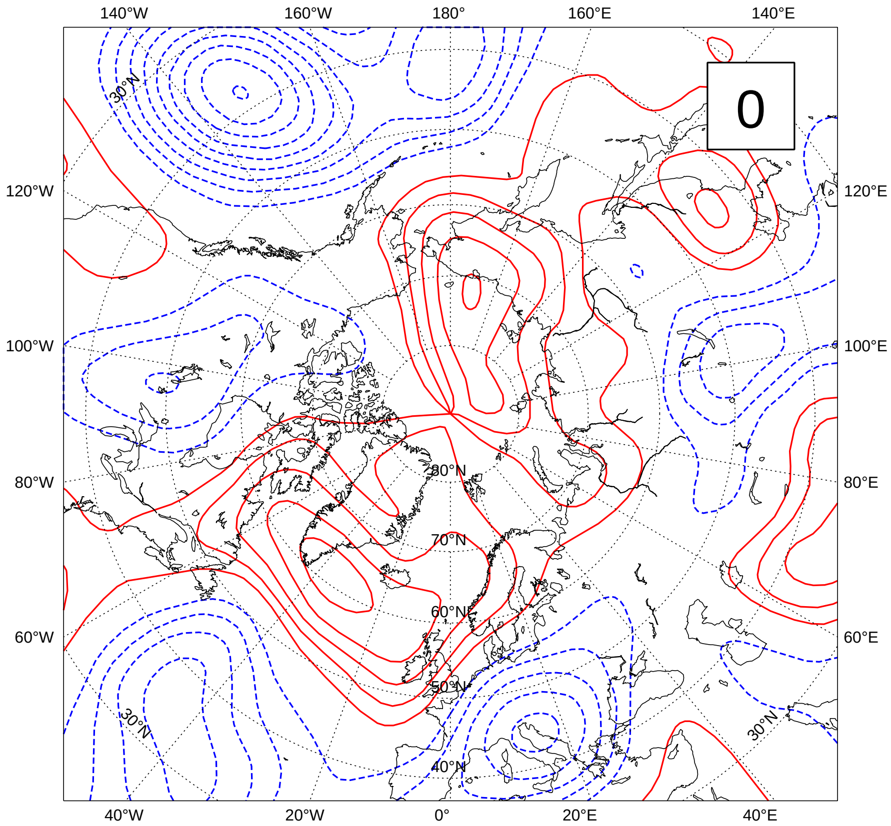
$$\boxed{\hat{\mathbf{x}}_a^k = \mathbf{M}_{\mathbf{X}_b + \hat{\mathbf{x}}_a^{k-1}/2} \mathbf{x}_a(0)}$$

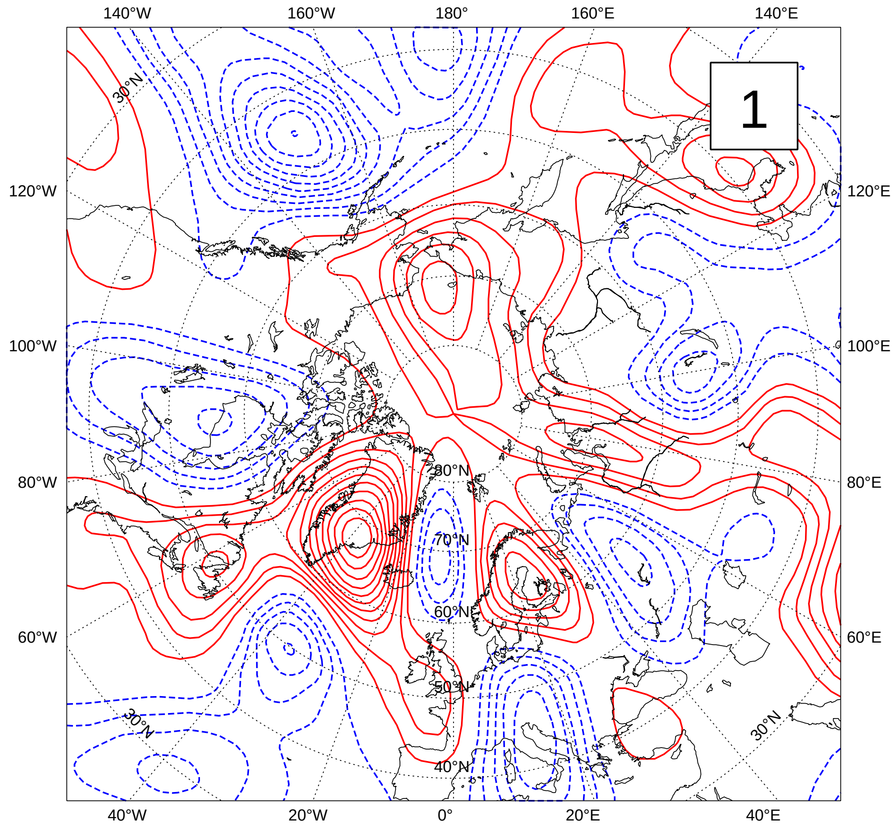
i.e. we consider the complete trajectory $\hat{\mathbf{x}}_a^{k-1}(t)$ as the independent variable $\hat{\mathbf{x}}_a^k = \mathbf{T}_{\mathbf{X}_b}(\hat{\mathbf{x}}_a^{k-1})$.

Prediction QG model T21L3 (2 days)

Nonlinear

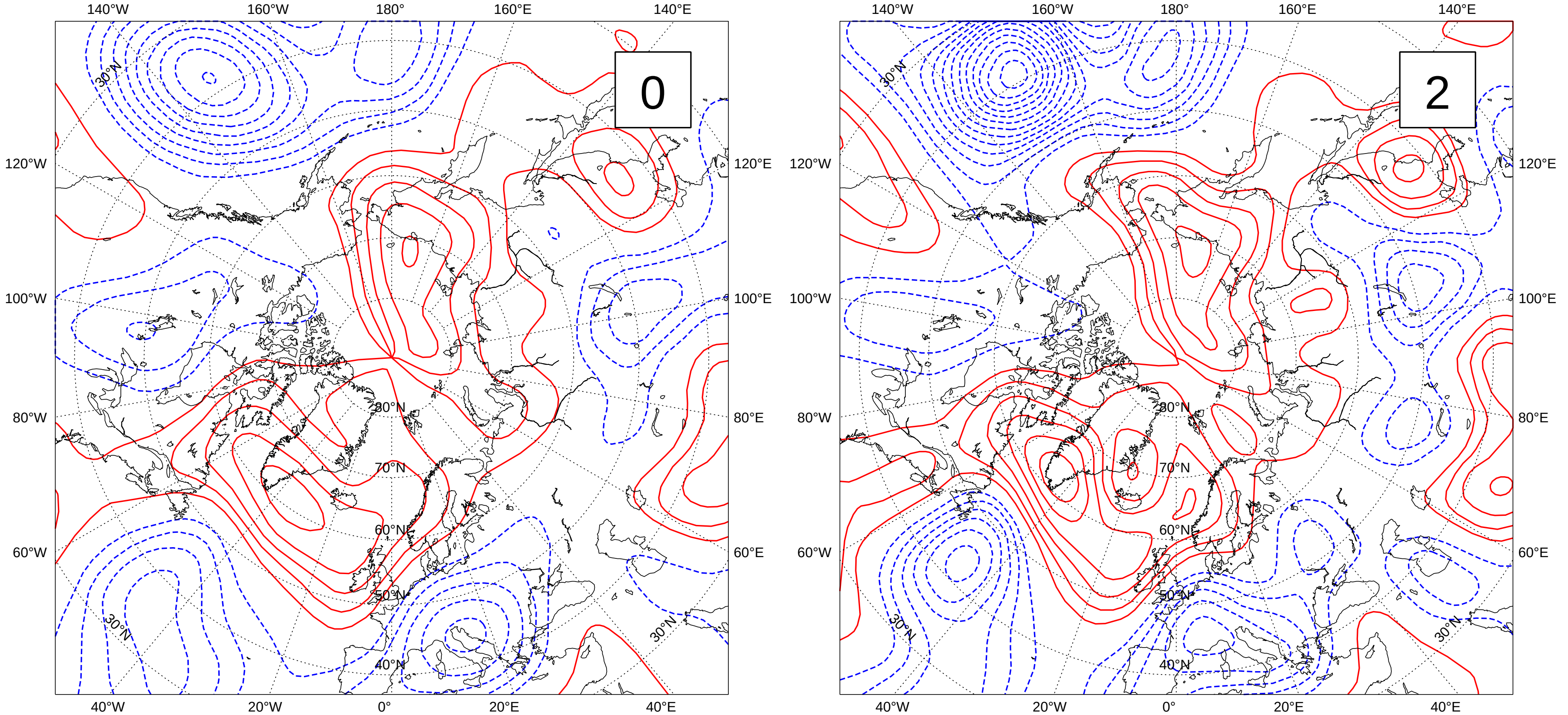
Tangent linear iteration 1 ($\mathbf{x}_a^0(t) = 0$)





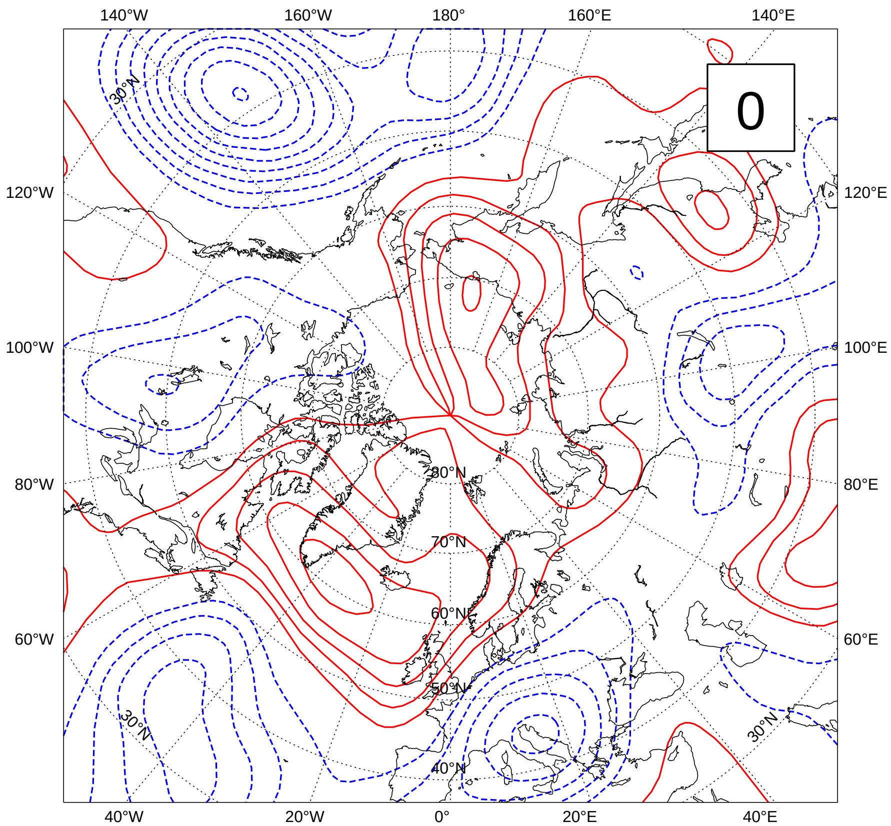
Prediction QG model T21L3 (2 days)

Nonlinear | Tangent linear iteration 2

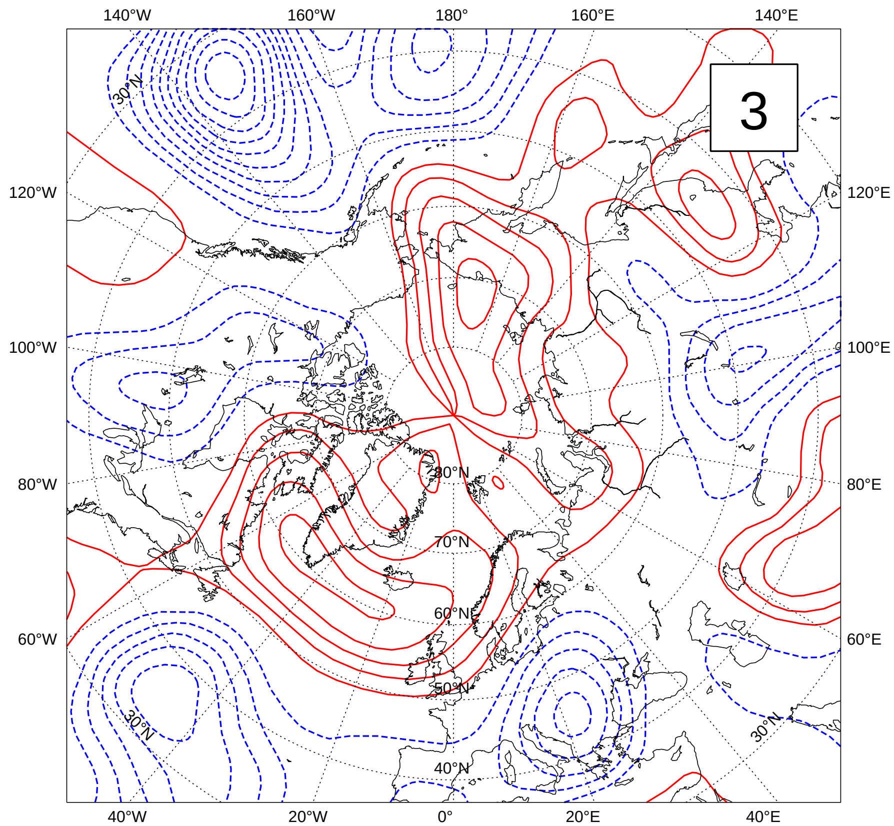


Prediction QG model T21L3 (2 days)

Nonlinear

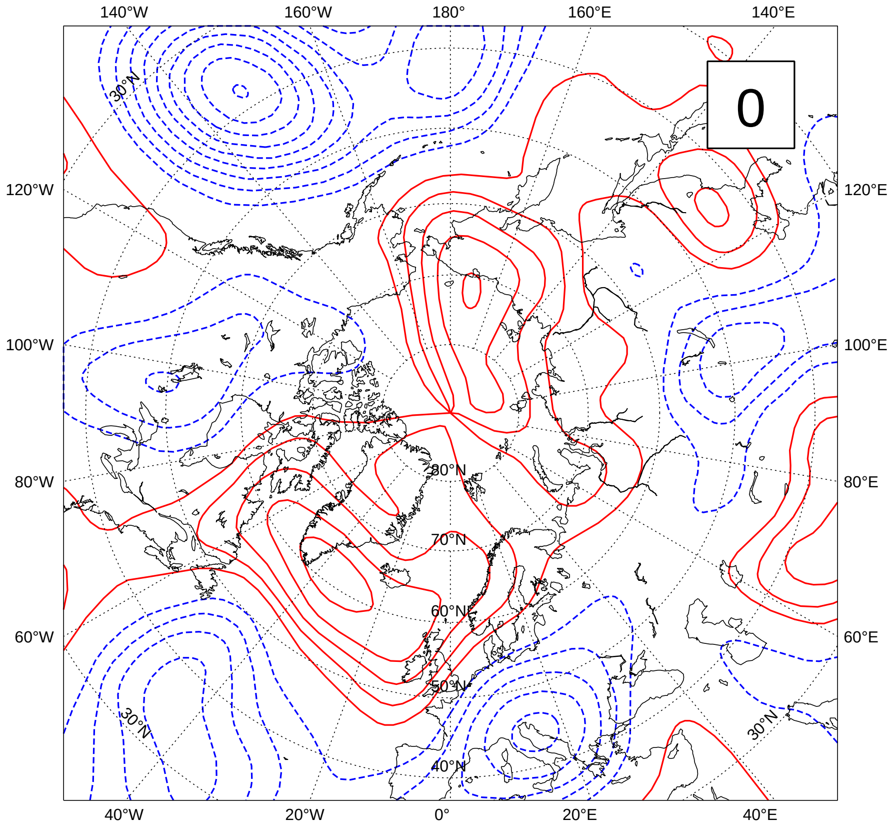


Tangent linear iteration 3

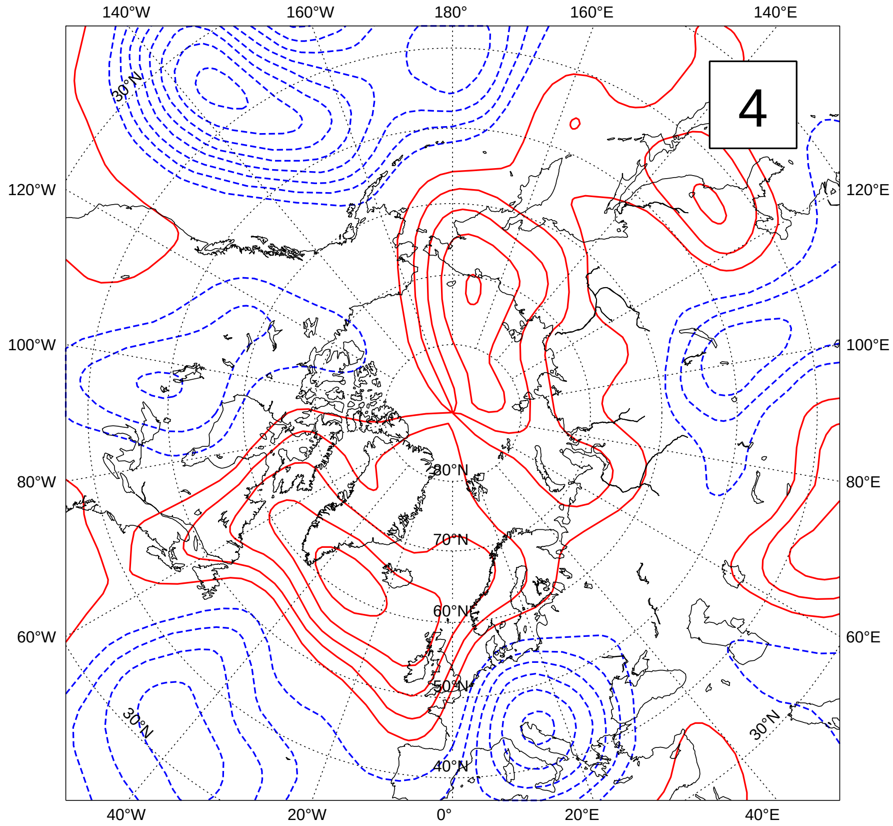


Prediction QG model T21L3 (2 days)

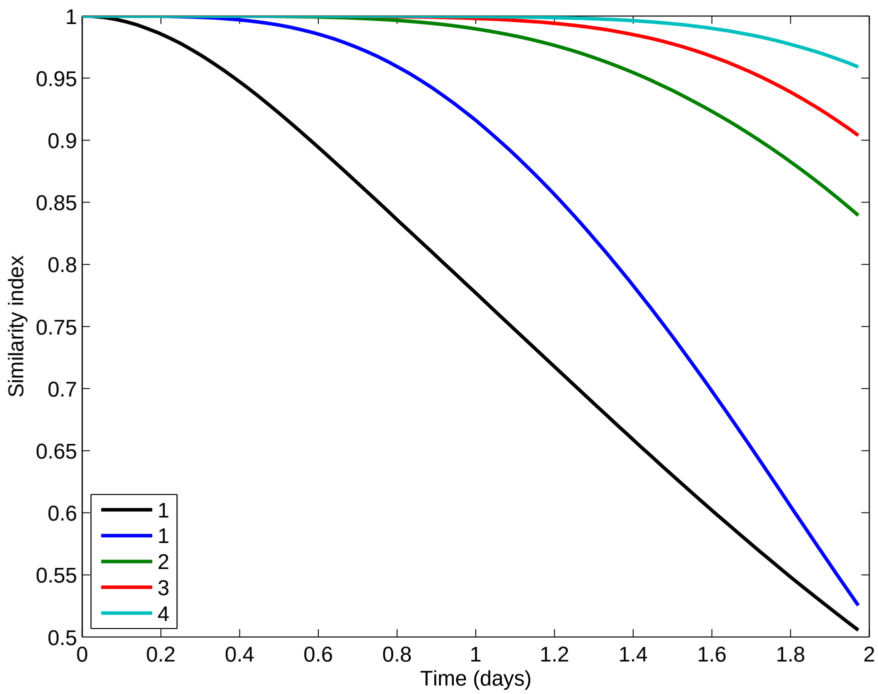
Nonlinear

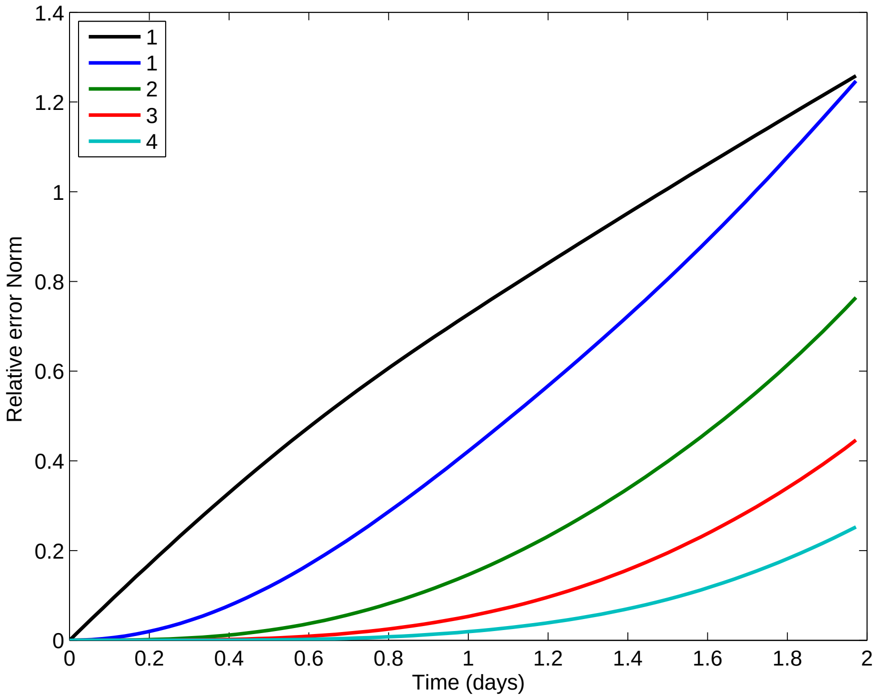


Tangent linear iteration 4



Prediction QG model T21L3 (2 days)





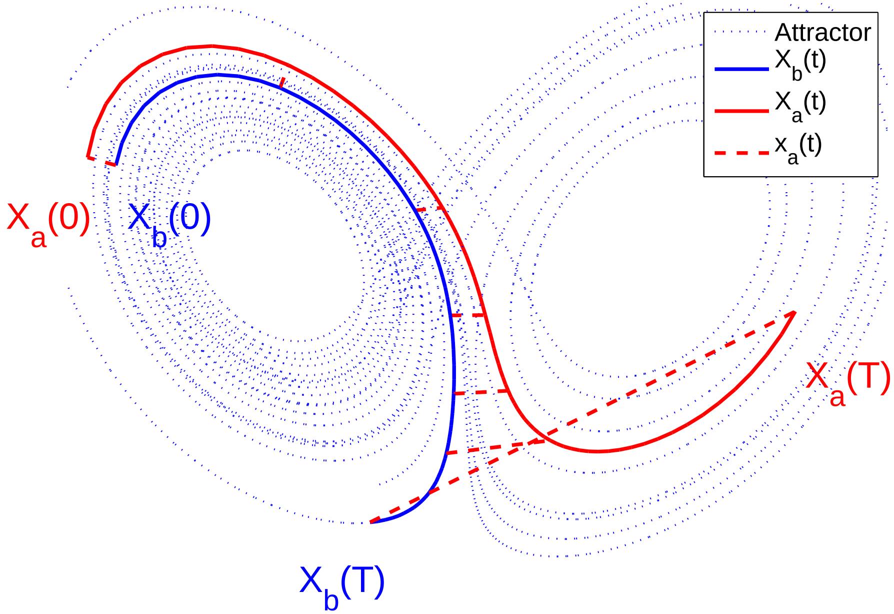
Similarity index

Relative error norm

Applications

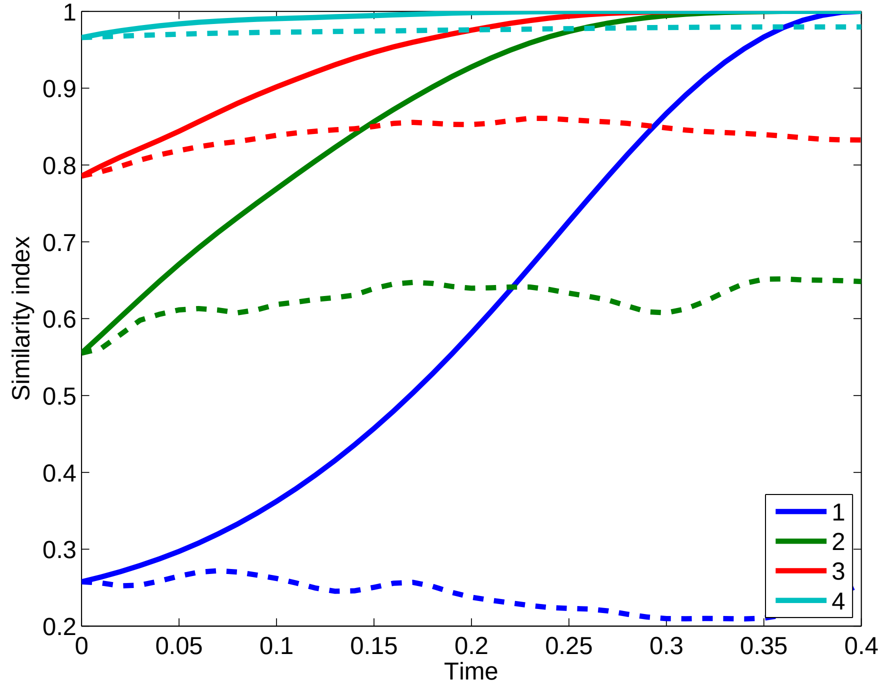
- Forecast sensitivity
- Variational data assimilation
- Adjoint based observation impact

Forecast sensitivity



- Nonlinear trajectories: $\mathbf{X}_b(t)$,$\mathbf{X}_a(t)$
- Measurement at time T: $\mathbf{X}_a(T)$
- Innovation at time T: $\mathbf{x}_a(T) = \mathbf{X}_a(T) - \mathbf{X}_b(T)$
- Estimate at time $t = 0$: $\hat{\mathbf{x}}_a(0) = \mathbf{M}^{-1}_{\mathbf{X}_b}(T, 0)\mathbf{x}_a(T)$

Forecast sensitivity (L96 model, 2days)



- Iteratively solve $\hat{\mathbf{x}}_a^k(t) = \mathrm{M}^{-1}_{\mathbf{X}_b + \hat{\mathbf{x}}_a^{k-1}/2} \mathbf{x}_a(T)$
- NL model can not be used to update $\mathbf{X}_b(t)$ (4D-VAR).

Concluding remarks

- The nonlinear time evolution of perturbations can be described by linear models
- In QG model opt. lin. traj. can be used for 200 days
- The incremental 4D-VAR algorithm can be revised s.t.
  - Computational cheaper
  - Allows merging of inner and outer loops
  - Easier to analyse by linear algebra methods
  - More suitable for high resolutions/long windows
- Higher order nonlinearities can be taken into account by linearizing around an ensemble of trajectories
- Exact adjoint based observation impact (including effect of multiple outer loops)

Concluding remarks

- The nonlinear time evolution of perturbations can be described by linear models
- In QG model opt. lin. traj. can be used for 200 days
- The incremental 4D-VAR algorithm can be revised s.t.
  - Computational cheaper
  - Allows merging of inner and outer loops
  - Easier to analyse by linear algebra methods
  - More suitable for high resolutions/long windows
- Higher order nonlinearities can be taken into account by linearizing around an ensemble of trajectories
- Exact adjoint based observation impact (including effect of multiple outer loops)

Thank you for your attention

An alternative 4D-VAR formulation

$$\mathbf{K}_{\mathbf{X}} = \left(\mathbf{H}_{\mathbf{X}}^T \mathbf{R}^{-1} \mathbf{H}_{\mathbf{X}} + \mathbf{B}^{-1}\right)^{-1} \mathbf{H}_{\mathbf{X}}^T \mathbf{R}^{-1}$$

4D-VAR

$$\hat{\mathbf{x}}_g^k = \mathbf{K}_{\mathbf{X}_g^{k-1}} \mathbf{y}^{k-1} \quad \text{with} \quad \mathbf{y}^{k-1} = \mathbf{Y} - \mathcal{H}(\mathbf{X}_g^{k-1})$$

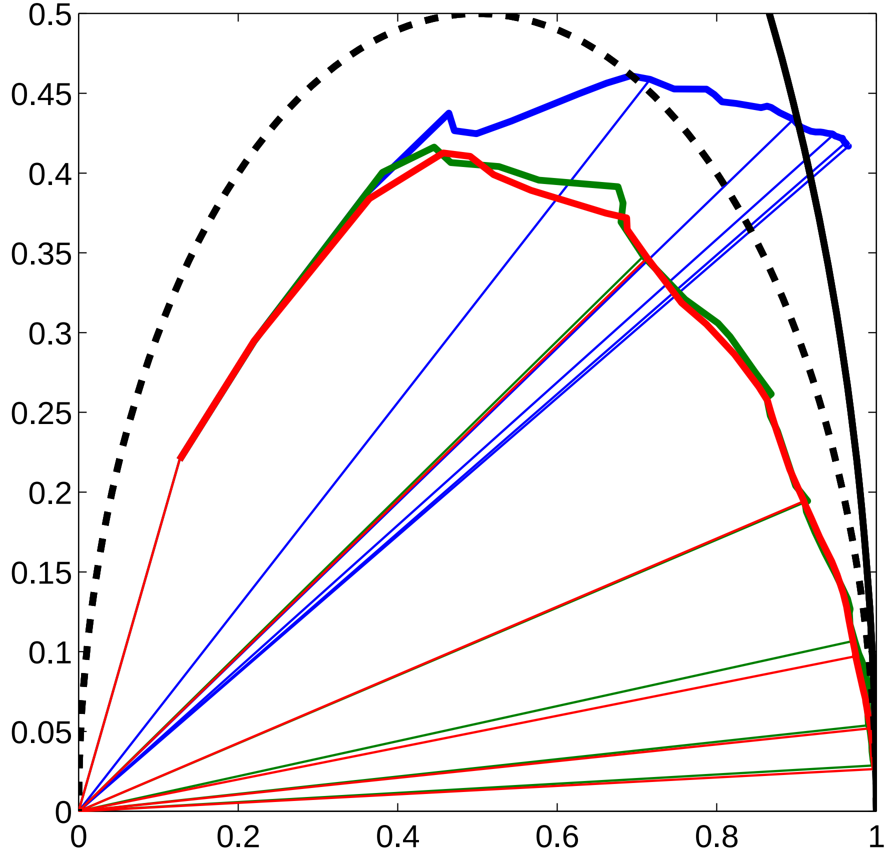
Alternative 4D-VAR

$$\hat{\mathbf{x}}_a^k = \mathbf{K}_{\mathbf{X}_b + \hat{\mathbf{x}}_a^{k-1}/2} \mathbf{y} \quad \text{with} \quad \mathbf{y} = \mathbf{Y} - \mathcal{H}(\mathbf{X}_b)$$

Both methods account for nonlinearity by modifying the linearization trajectory but

- Possibility to update trajectory in inner loops
- Proposed method does not modify the innovation vector
- More suitable for high resolution/long window DA?

Merging inner/outer loops QG model



without inner loop updates (blue) with inner loop updates (green) and solving the nonlin. problem directly (red) M1QN3

Beyond advection

For bilinear models

$$\boxed{\dot{\hat{\mathbf{x}}}_a = \mathbf{Df}(\mathbf{X}_b + \tfrac{1}{2}\mathbf{x}_a)\hat{\mathbf{x}}_a}$$

Higher order nonlinearities

$$\dot{\mathbf{X}} = s_0 + s_1(\mathbf{X}) + s_2(\mathbf{X}, \mathbf{X}) + s_3(\mathbf{X}, \mathbf{X}, \mathbf{X}) + \dots$$

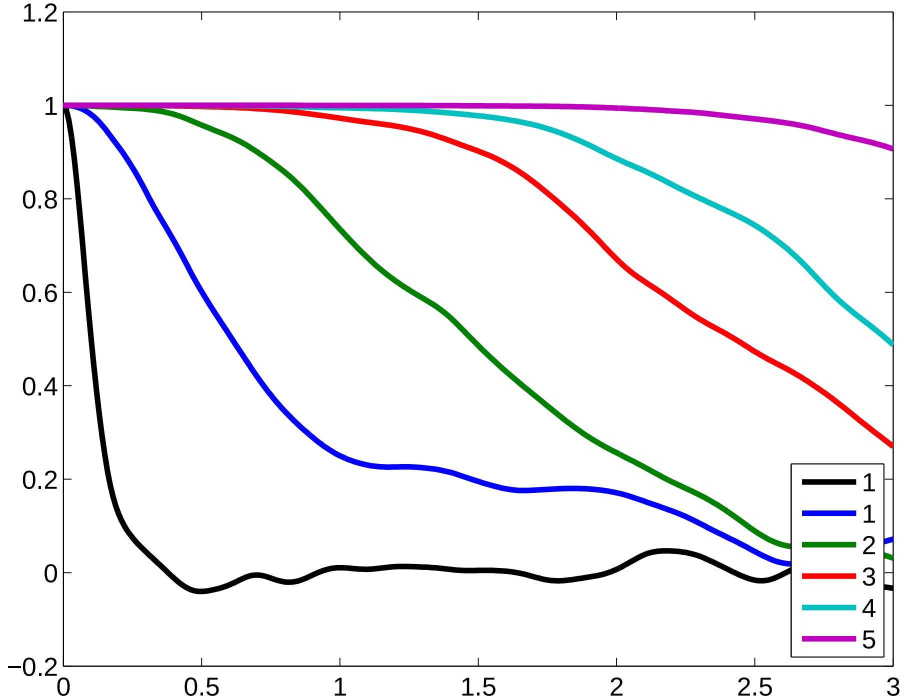
can be taken into account by

$$\dot{\hat{\mathbf{x}}}_a = \sum_{j=1}^{J} \alpha_j \mathbf{Df}(\mathbf{X}_b + \beta_j \mathbf{x}_a)\hat{\mathbf{x}}_a$$

Using Gaussian quadrature we can take into account all terms up to order s_{2J}. The TL-model is integrated around an ensemble of trajectories simultaneously.

Ensemble of trajectories in L96

Similarity index Lorenz 96 with higher order nonlinearities



$$
\begin{aligned}
\mathbf{x}_a(t) &= \mathbf{X}_a(t) - \mathbf{X}_b(t) \\
\dot{\hat{\mathbf{x}}}_a &= \sum_{j=1}^{J} \alpha_j \mathbf{Df}(\mathbf{X}_b + \beta_j \mathbf{x}_a)\hat{\mathbf{x}}_a
\end{aligned}
$$

Adjoint observation impact