Ensemble redrawing in strongly nonlinear systems

Pavel Sakov [1] Marc Bocquet [2 3]

[1]Bureau of Meteorology, Melbourne, Australia

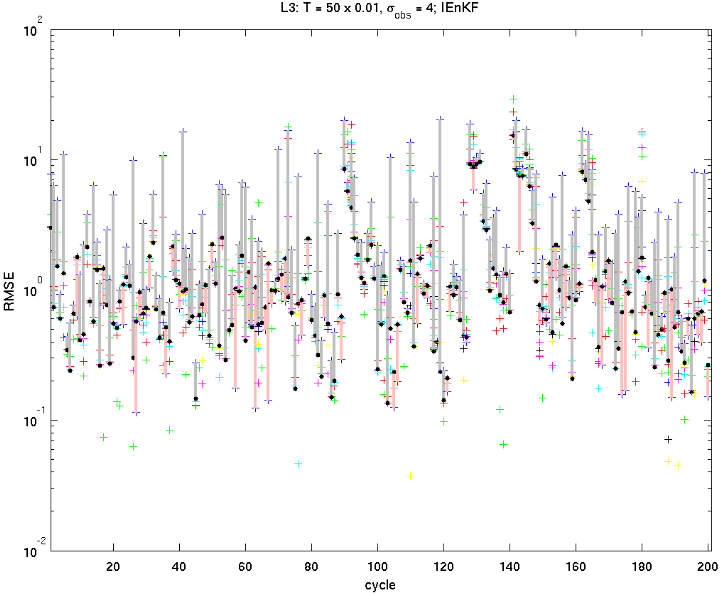
[2]Université Paris-Est, CEREA, Champs-sur-Marne, France

[3]INRIA, Paris Rocquencourt research centre, France

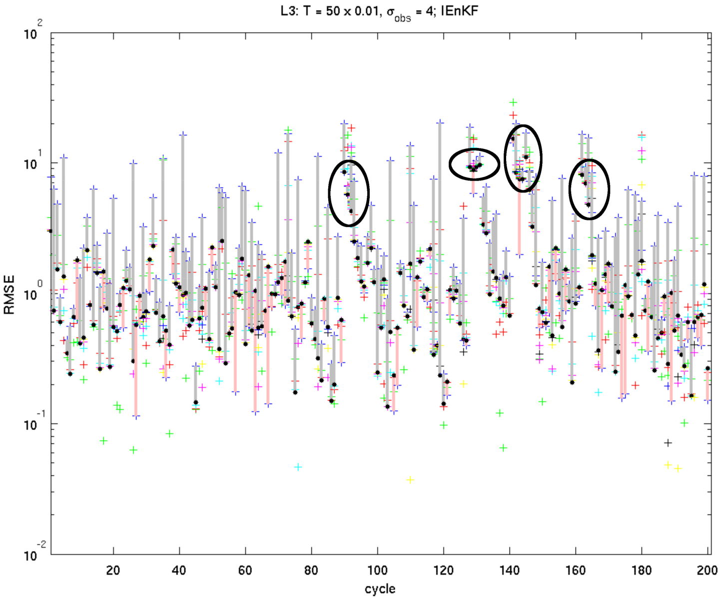
EnKF workshop, Bergen 22-24 May 2013

Outline

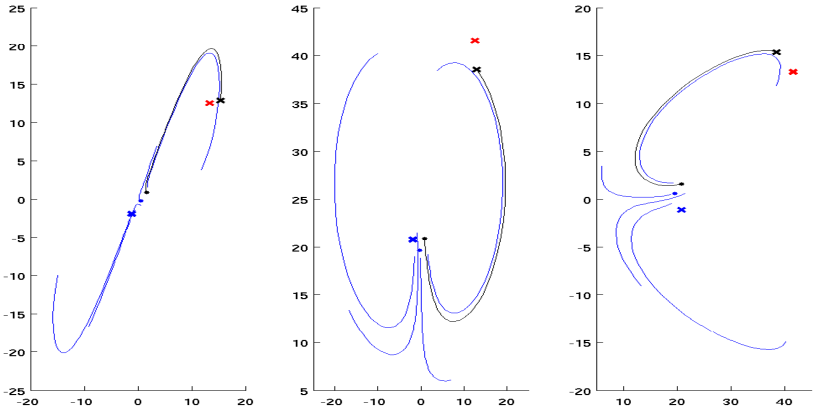
Motivation: an example with 3-variable Lorenz model



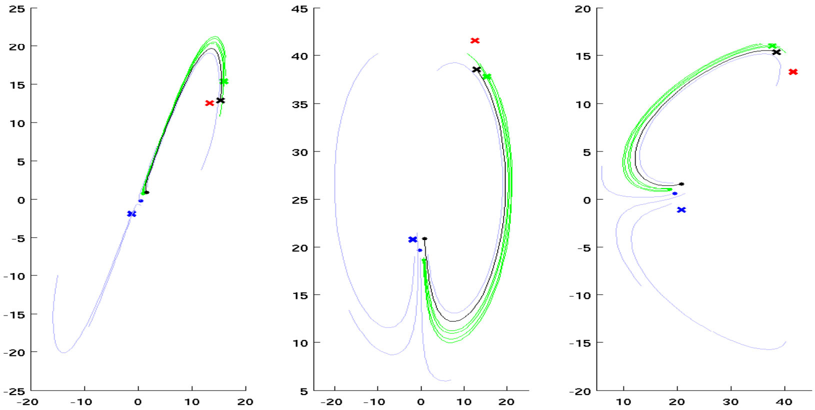
Motivation: an example with 3-variable Lorenz model



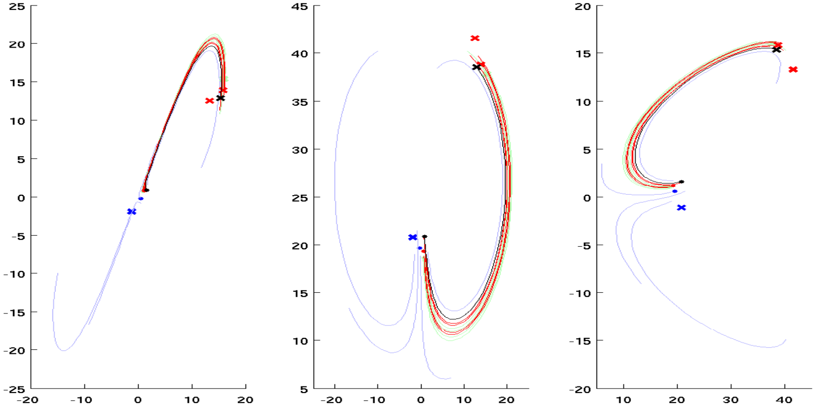
Example 1: fast convergence



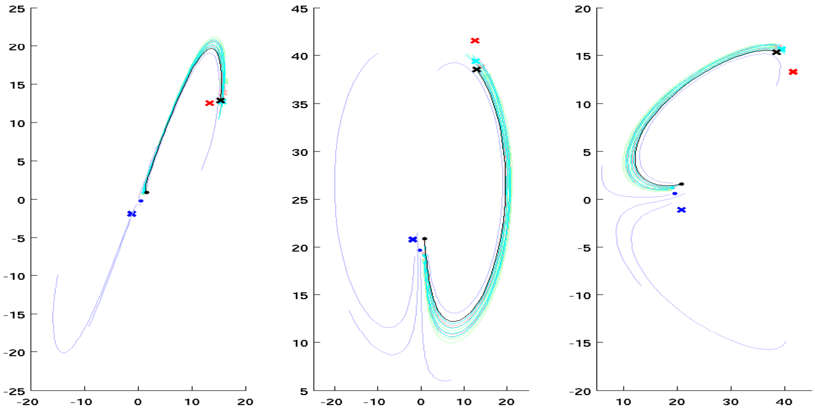
Example 1: fast convergence



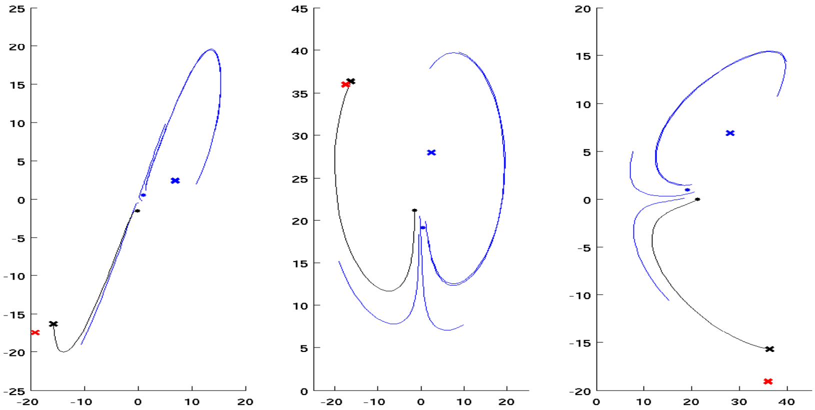
Example 1: fast convergence



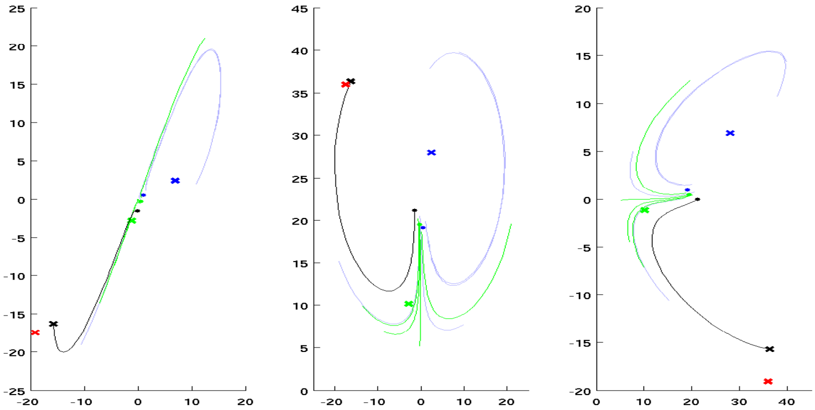
Example 1: fast convergence



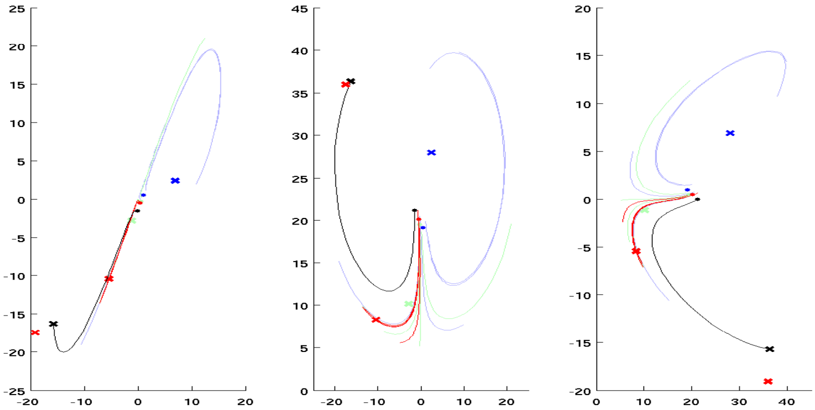
Example 2: Slower convergence



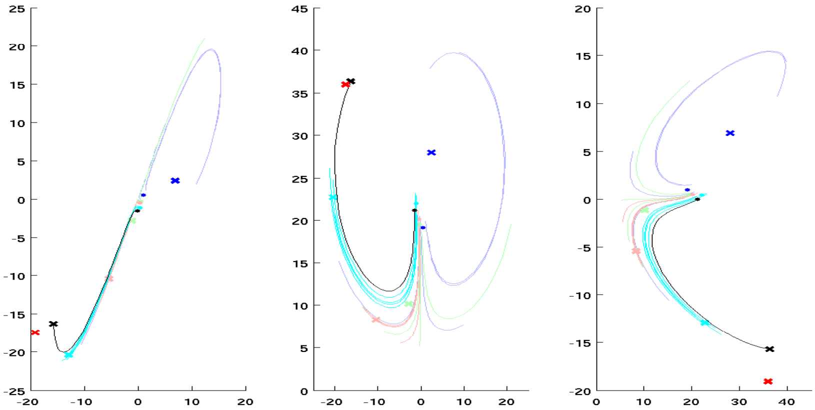
Example 2: Slower convergence



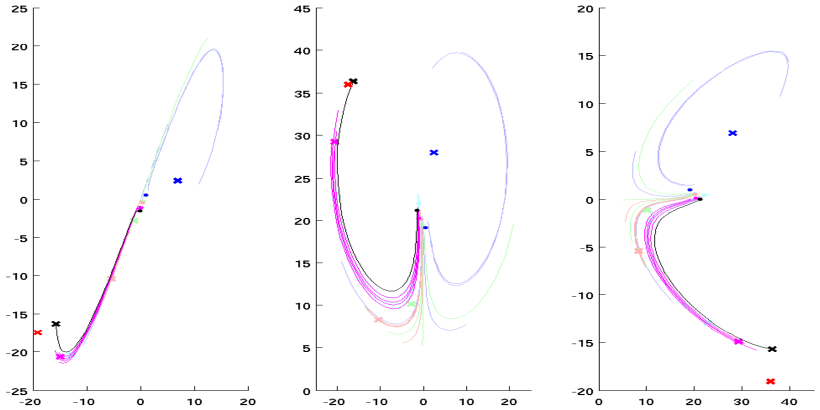
Example 2: Slower convergence



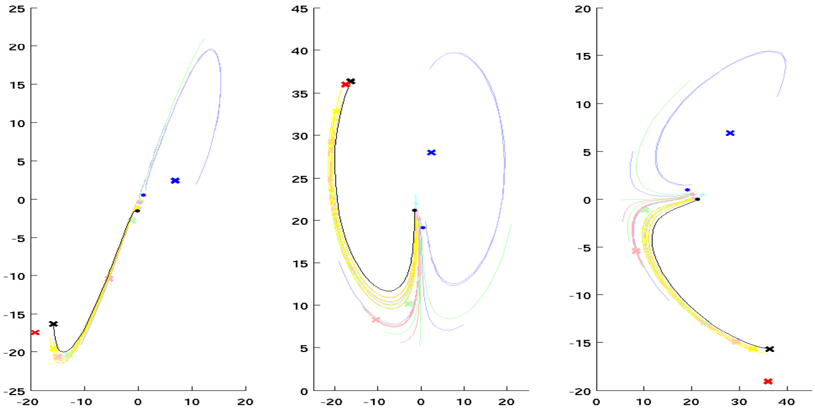
Example 2: Slower convergence



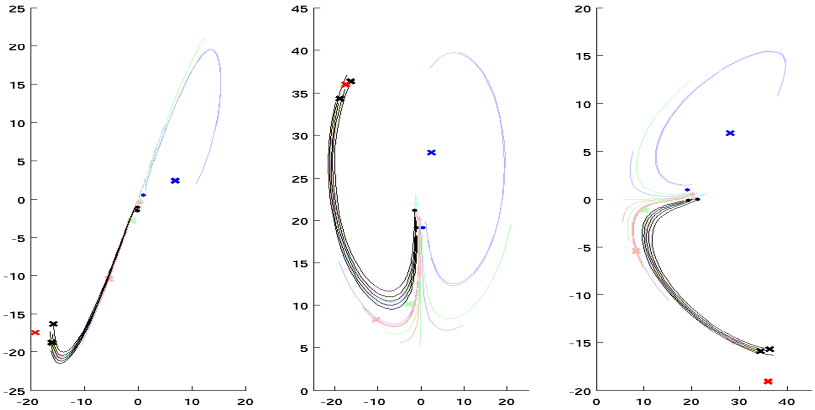
Example 2: Slower convergence



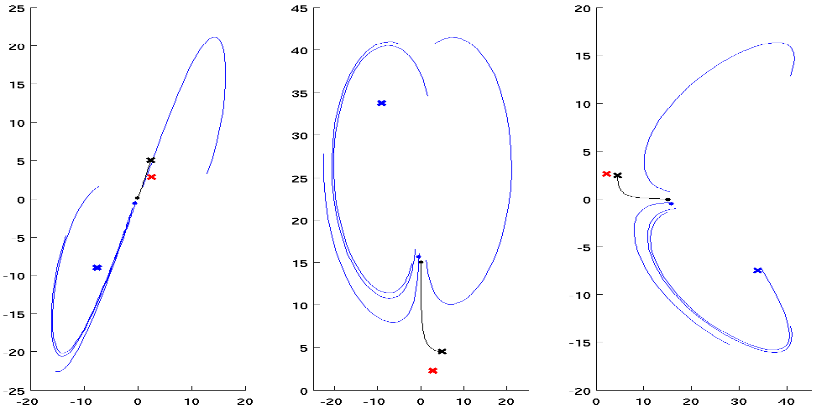
Example 2: Slower convergence



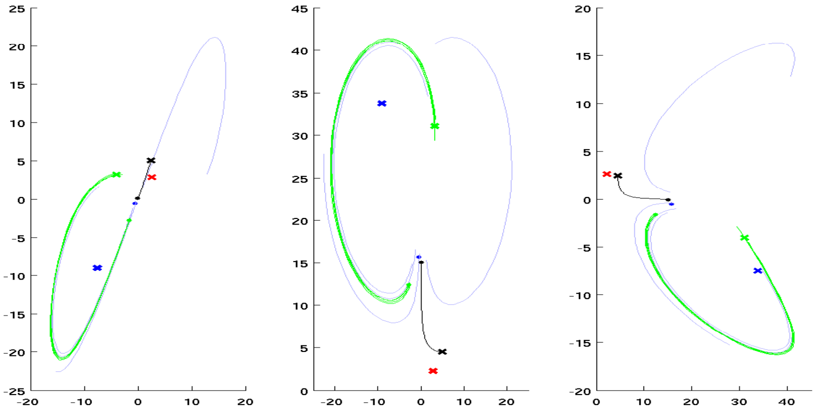
Example 2: Slower convergence



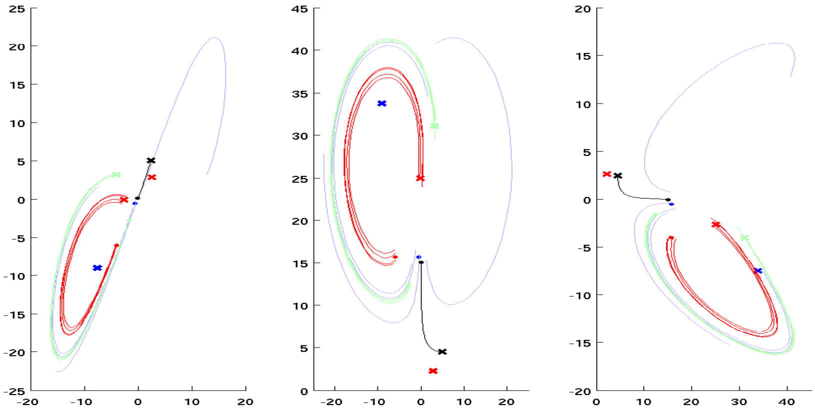
Example 3: Divergence



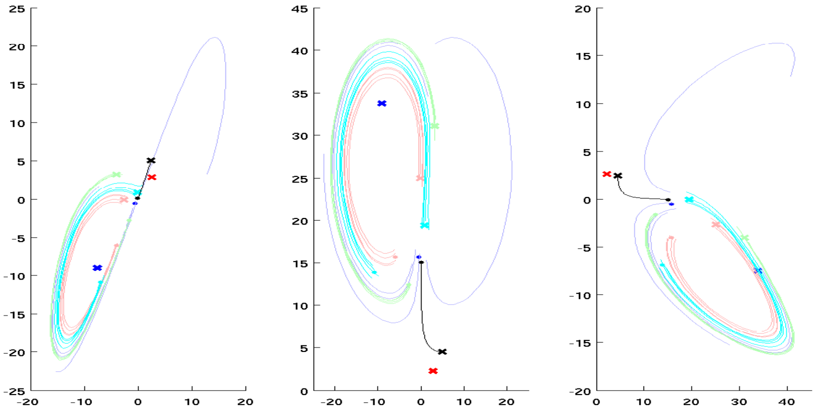
Example 3: Divergence



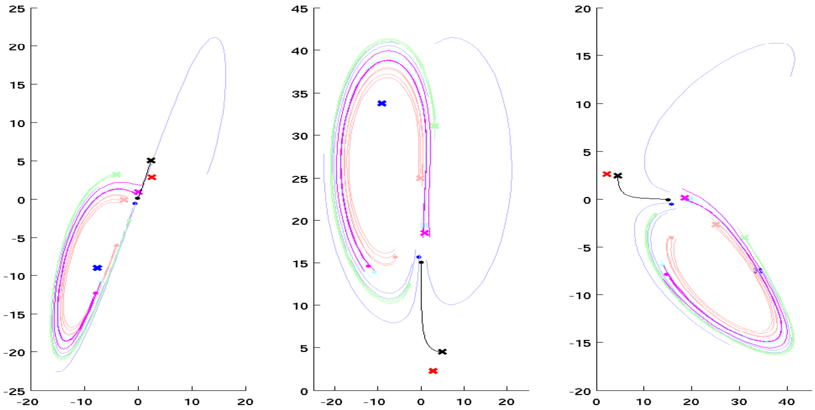
Example 3: Divergence



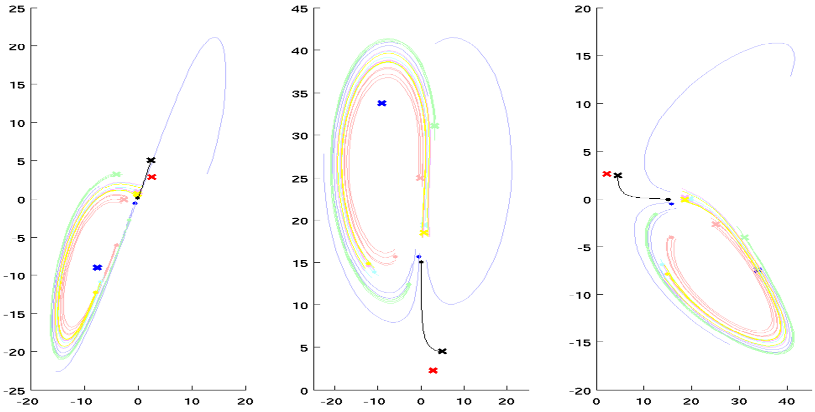
Example 3: Divergence



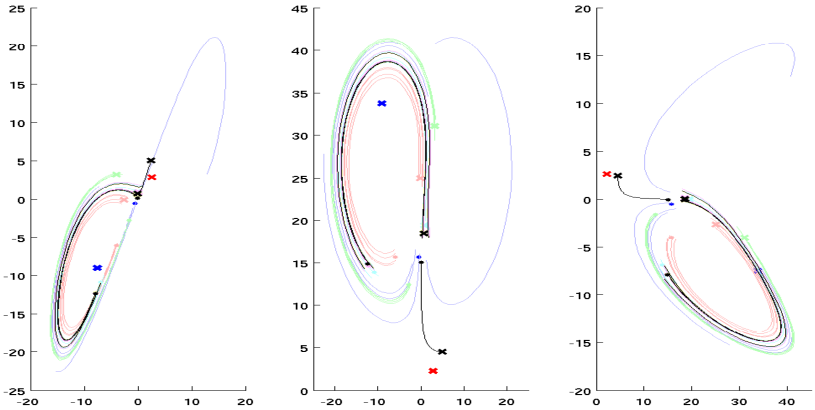
Example 3: Divergence



Example 3: Divergence

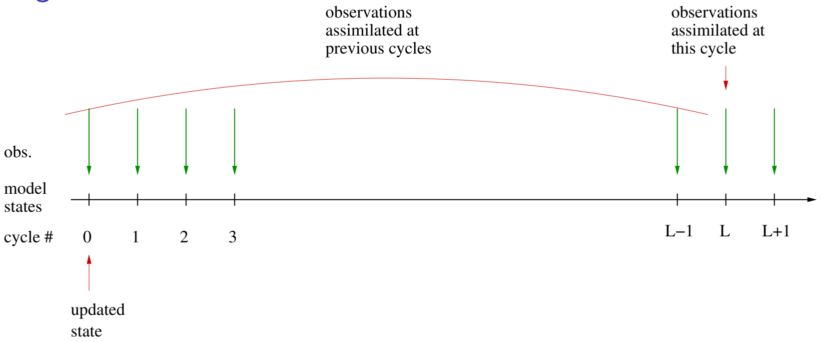


Example 3: Divergence



IEnKS

Lag-L smoother



IEnKS (Gauss-Newton, transform)

$$\mathbf{x}_0^{(0)} = \mathbf{E}_0\,\mathbf{1}/m, \quad \mathbf{A}_0^{(0)} = \mathbf{E}_0 - \mathbf{x}_0^{(0)}, \quad \mathbf{w} = \mathbf{0}, \quad \mathbf{T} = \mathbf{I}$$

repeat

$$\mathbf{x}_0 = \mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\mathbf{w}$$

$$\mathbf{E}_0 = \mathbf{x}_0\mathbf{1}^{\mathrm{T}} + \mathbf{A}_0^{(0)}\mathbf{T}$$

$$\mathbf{E}_L = \mathcal{M}_{0\to L}(\mathbf{E}_0)$$

$$\mathbf{Hx} = \mathcal{H}(\mathbf{E}_L)\,\mathbf{1}/m$$

$$\mathbf{HA} = [\mathcal{H}(\mathbf{E}_L) - \mathbf{Hx}]\,\mathbf{T}^{-1}$$

$$\nabla J = (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{Hx})/(m-1) + [(\mathbf{A}_0^{(0)})^{\mathrm{T}}\mathbf{A}_0^{(0)}]^{\dagger}(\mathbf{A}_0^{(0)})^{\mathrm{T}}(\mathbf{x}_0^{(0)} - \mathbf{x}_0)$$

$$\mathbf{M} = \mathbf{I} + (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HA}/(m-1)$$

$$\Delta\mathbf{w} = \mathbf{M}^{-1}\nabla J$$

$$\mathbf{w} = \mathbf{w} + \Delta\mathbf{w}$$

$$\mathbf{T} = \mathbf{M}^{-1/2}$$

until $\|\Delta\mathbf{w}\| < \varepsilon$

$$\mathbf{E}_1 = \mathcal{M}_{0\to 1}(\mathbf{E}_0)$$

inflate $\mathbf{E}_1$

IEnKS (Gauss-Newton, regression)

$$\mathbf{x}_0^{(0)} = \mathbf{E}_0\,\mathbf{1}/m, \quad \mathbf{A}_0^{(0)} = \mathbf{E}_0 - \mathbf{x}_0^{(0)}, \quad \mathbf{w} = \mathbf{0}, \quad \mathbf{T} = \mathbf{I}$$

repeat

$$\mathbf{x}_0 = \mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\mathbf{w}$$

$$\mathbf{E}_0 = \mathbf{x}_0\mathbf{1}^{\mathrm{T}} + \mathbf{A}_0^{(0)}\mathbf{T}$$

$$\mathbf{E}_L = \mathcal{M}_{0\to L}(\mathbf{E}_0)$$

$$\mathbf{Hx} = \mathcal{H}(\mathbf{E}_L)\,\mathbf{1}/m$$

$$\mathbf{HA} = [\mathcal{H}(\mathbf{E}_L) - \mathbf{Hx}], \quad \mathbf{HM} = \mathbf{HA}(\mathbf{A}_0^{(0)}\mathbf{T})^{\dagger}, \quad \mathbf{HA} = \mathbf{HM}\,\mathbf{A}_0^{(0)}$$

$$\nabla J = (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - \mathbf{Hx} + \mathbf{HM}(\mathbf{x}_0 - \mathbf{x}_0^{(0)})]/(m-1)$$

$$\mathbf{M} = \mathbf{I} + (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HA}/(m-1)$$

$$\Delta\mathbf{w} = \mathbf{M}^{-1}\nabla J$$

$$\mathbf{w} = \mathbf{w} + \Delta\mathbf{w}$$

$$\mathbf{T} = \mathbf{M}^{-1/2}$$

until $\|\Delta\mathbf{w}\| < \varepsilon$

$$\mathbf{E}_1 = \mathcal{M}_{0\to 1}(\mathbf{E}_0)$$

inflate $\mathbf{E}_1$

IEnKS (Gauss-Newton, transform, revisited)

$$\mathbf{x}_0^{(0)} = \mathbf{E}_0\, \mathbf{1}/m, \quad \mathbf{A}_0^{(0)} = \mathbf{E}_0 - \mathbf{x}_0^{(0)}, \quad \mathbf{w} = \mathbf{0}, \quad \mathbf{T} = \mathbf{I}$$

repeat

$$\mathbf{x}_0 = \mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\mathbf{w}$$

$$\mathbf{E}_0 = \mathbf{x}_0\mathbf{1}^{\mathrm{T}} + \mathbf{A}_0^{(0)}\mathbf{T}$$

$$\mathbf{E}_L = \mathcal{M}_{0\to L}(\mathbf{E}_0)$$

$$\mathbf{Hx} = \mathcal{H}(\mathbf{E}_L)\,\mathbf{1}/m$$

$$\mathbf{HA} = \left[\mathcal{H}(\mathbf{E}_L) - \mathbf{Hx}\right]\mathbf{T}^{-1}$$

$$\nabla J = (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{Hx})/(m-1) - \mathbf{w}$$ (Bocquet and Sakov, 2012)

$$\mathbf{M} = \mathbf{I} + (\mathbf{HA})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HA}/(m-1)$$

$$\Delta\mathbf{w} = \mathbf{M}^{-1}\nabla J$$

$$\mathbf{w} = \mathbf{w} + \Delta\mathbf{w}$$

$$\mathbf{T} = \mathbf{M}^{-1/2}$$

until $\|\Delta\mathbf{w}\| < \varepsilon$

$$\mathbf{E}_1 = \mathcal{M}_{0\to 1}(\mathbf{E}_0)$$

inflate $\mathbf{E}_1$

IEnKS: two formulations

State space formulation:

$$\mathbf{x}_0^a = \arg\min_{\{\mathbf{x}_0\}} \Big\{ (\mathbf{x}_0 - \mathbf{x}_0^{(0)})^{\mathrm{T}} (\mathbf{P}_0^{(0)})^{-1} (\mathbf{x}_0 - \mathbf{x}_0^{(0)}) + \left[\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)\right]^{\mathrm{T}} (\mathbf{R}_L)^{-1} \left[\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)\right] \Big\},$$

$$\mathbf{x}_L = \mathcal{M}_{0\to L}(\mathbf{x}_0)$$

Ensemble space formulation:

$$\mathbf{w} = \arg\min_{\{\mathbf{w}\}} \Big\{ \mathbf{w}^{\mathrm{T}}\mathbf{w} + \left[\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)\right]^{\mathrm{T}} (\mathbf{R}_L)^{-1} \left[\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)\right] \Big\},$$

$$\mathbf{x}_L = \mathcal{M}_{0\to L}(\mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\mathbf{w})$$

(Equivalence - see Hunt et al. 2007)

EnKF solution space and IEnKF solution

- Let $\mathbf{A}$ be analysed ensemble anomalies in the EnKF
- Then $\tilde{\mathbf{A}} = \mathbf{A}\,\mathbf{U}$, where $\mathbf{U}: \ \mathbf{U}^{\mathrm{T}} = \mathbf{I}, \ \mathbf{U}\,\mathbf{1} = \mathbf{1}$ is also a KF solution

$$\mathbf{w} = \arg\min_{\{\mathbf{w}\}} \left\{ \mathbf{w}^{\mathrm{T}}\mathbf{w} + [\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)]^{\mathrm{T}} (\mathbf{R}_L)^{-1} [\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)] \right\},$$
$$\mathbf{x}_L = \mathcal{M}_{0\to L}(\mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\mathbf{w})$$

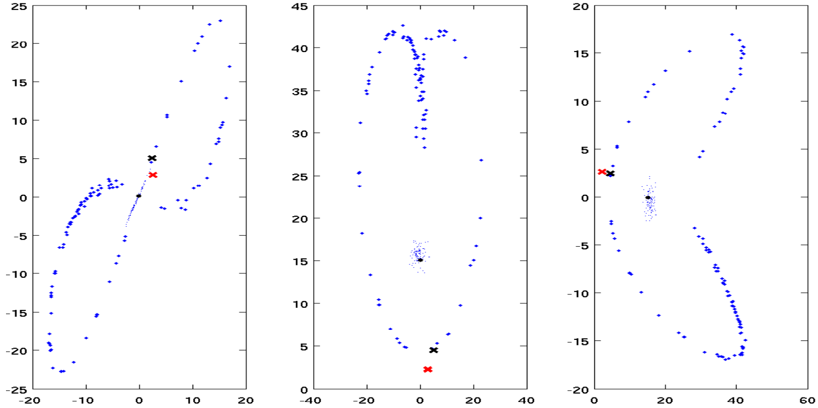
$\mathbf{A} \leftarrow \mathbf{AU}$:

$$\tilde{\mathbf{w}} = \arg\min_{\{\tilde{\mathbf{w}}\}} \left\{ \tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{w}} + [\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)]^{\mathrm{T}} (\mathbf{R}_L)^{-1} [\mathbf{y}_L - \mathcal{H}_L(\mathbf{x}_L)] \right\},$$
$$\mathbf{x}_L = \mathcal{M}_{0\to L}(\mathbf{x}_0^{(0)} + \mathbf{A}_0^{(0)}\tilde{\mathbf{w}}),$$

where $\quad \tilde{\mathbf{w}} \equiv \mathbf{U}\mathbf{w}$

Hence the ensemble redrawing can result in:

- A slightly different solution due to estimating the sensitivities from an ensemble of finite spread
- Convergence to another minimum due to the different initial state of the system



25
20
15
10
5
0
-5
-10
-15
-20
-25
-20
-10
0
10
20
45
40
35
30
25
20
15
10
5
0
-40
-20
0
20
40
20
15
10
5
0
-5
-10
-15
-20
0
20
40
60

Algorithm

IEnKF cycle with redrawing:

```
repeat
    <iterate>
    if <diverged>
        <re-initialise the cycle>
        <redraw the forecast ensemble>
        continue
    end if
until <success>
```

Algorithm

Rolling back:

repeat

<IEnKF cycle with redrawing>

if <failed>

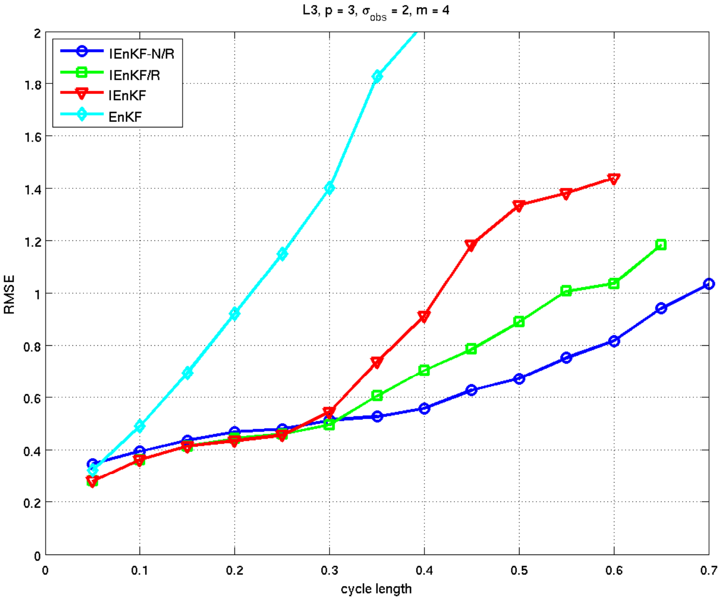
<go back >

<redraw the analysed ensemble>

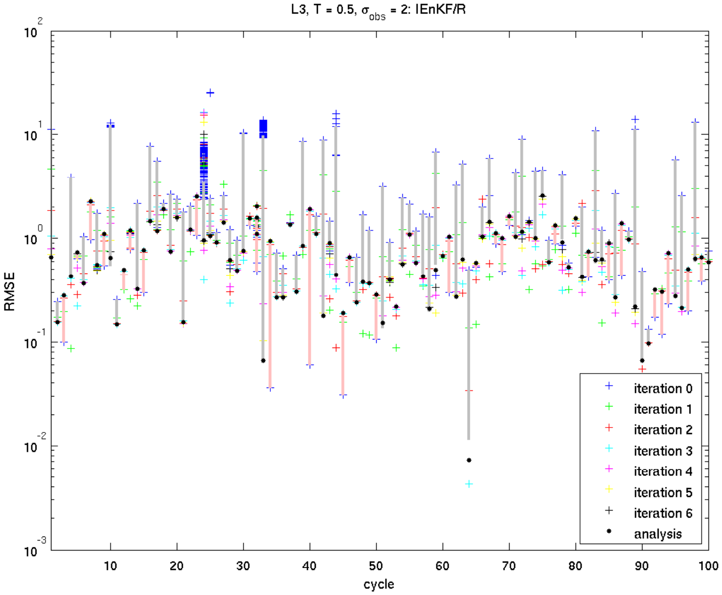
end if

until <success >

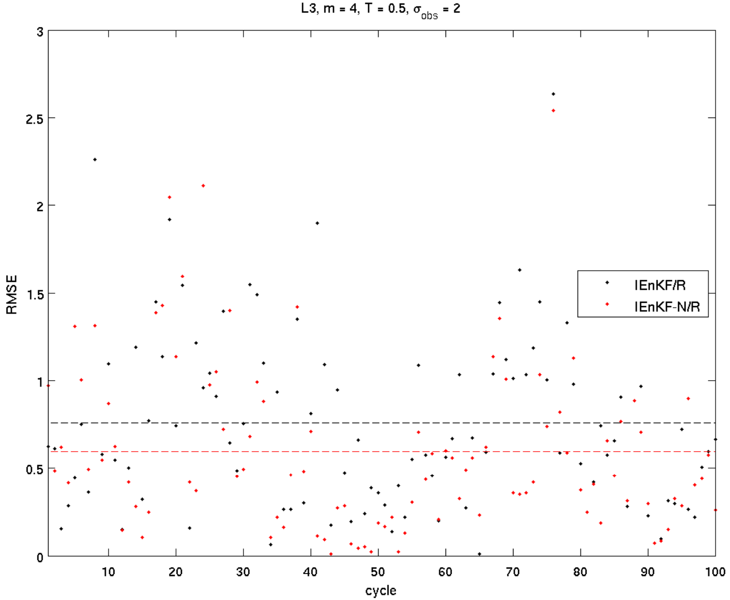
Performance with L3



IEnKF: performance log



IEnKF vs. IEnKF-N: baseline peformance



IEnKS-N

...

repeat

...

$$\nabla J = (\mathbf{HA})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Hx})/(m-1) - \mathbf{w}$$

becomes

$$\nabla J = (\mathbf{HA})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Hx})/(m-1) - m\,\mathbf{w}/(\varepsilon_N + \mathbf{w}^{\mathrm{T}}\mathbf{w})/(m-1)$$

...

until $\|\Delta \mathbf{w}\| < \varepsilon$

$\mathbf{E}_1 = \mathcal{M}_{0\to 1}(\mathbf{E}_0)$
inflate $\mathbf{E}_1$

becomes

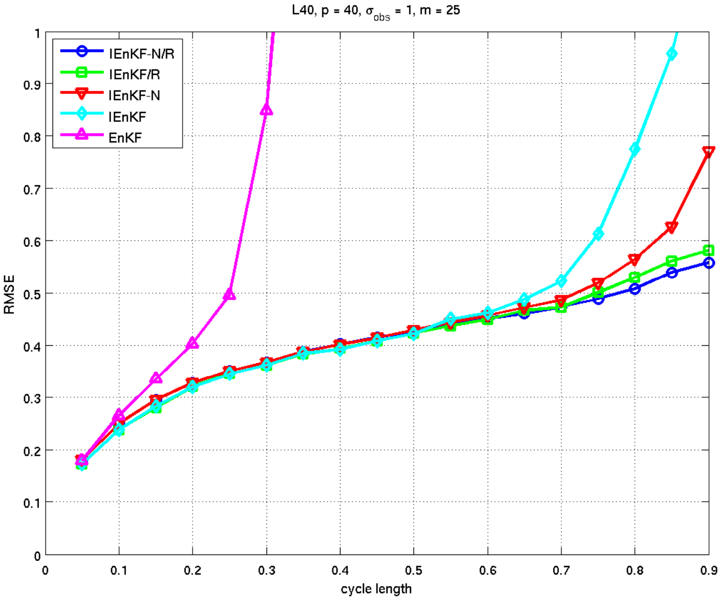
$$c = \varepsilon_N + \mathbf{w}^{\mathrm{T}}\mathbf{w}$$

$$\mathbf{M} = m\,(c\,\mathbf{I} - 2\mathbf{w}\mathbf{w}^{\mathrm{T}})/c^2/(m-1) + (\mathbf{HA})^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{HA}/(m-1)$$

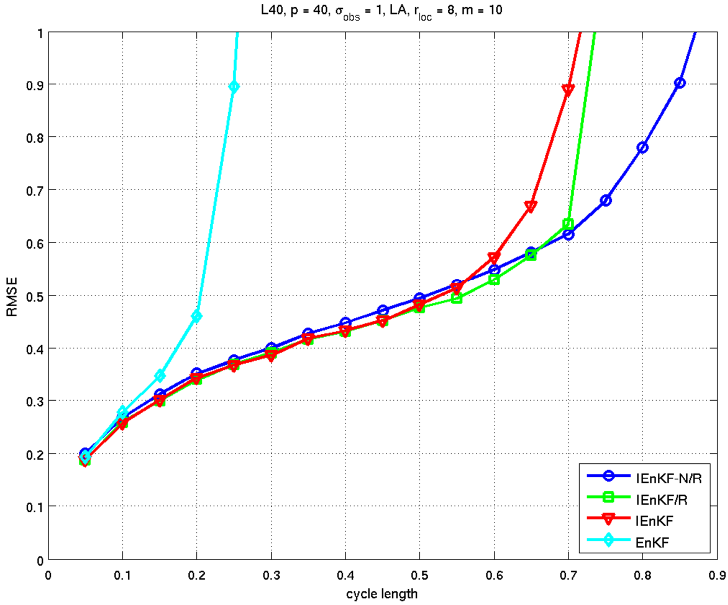
$$\mathbf{E}_0 = \mathbf{x}_0 \mathbf{1}^{\mathrm{T}} + \mathbf{A}_0^{(0)} \mathbf{M}^{-1/2}$$

$$\mathbf{E}_1 = \mathcal{M}_{0\to 1}(\mathbf{E}_0)$$

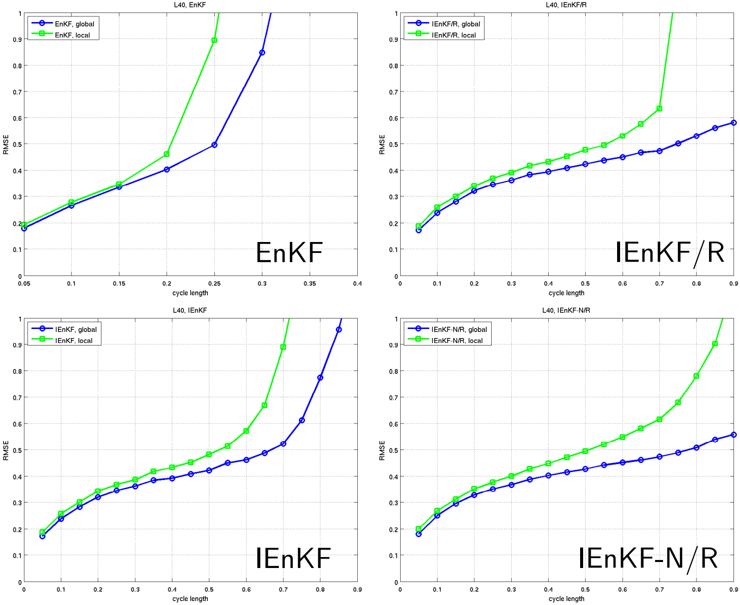
Performance with L40 (global)



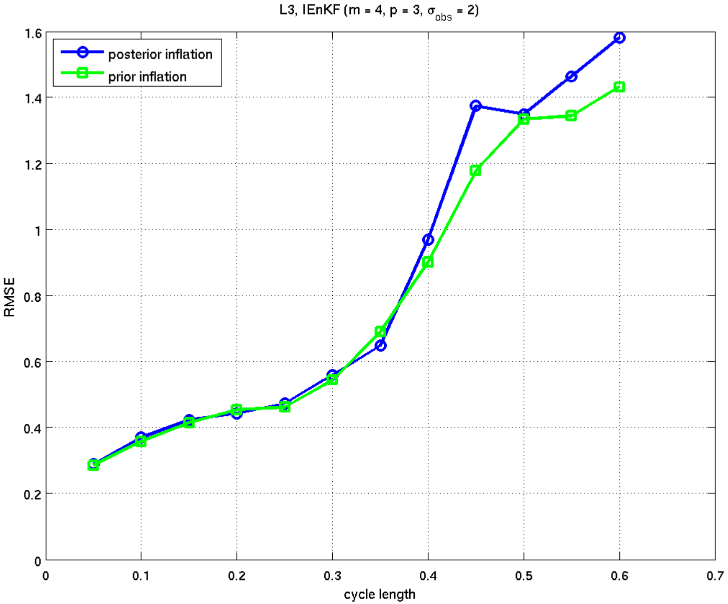
Performance with L40 (local)



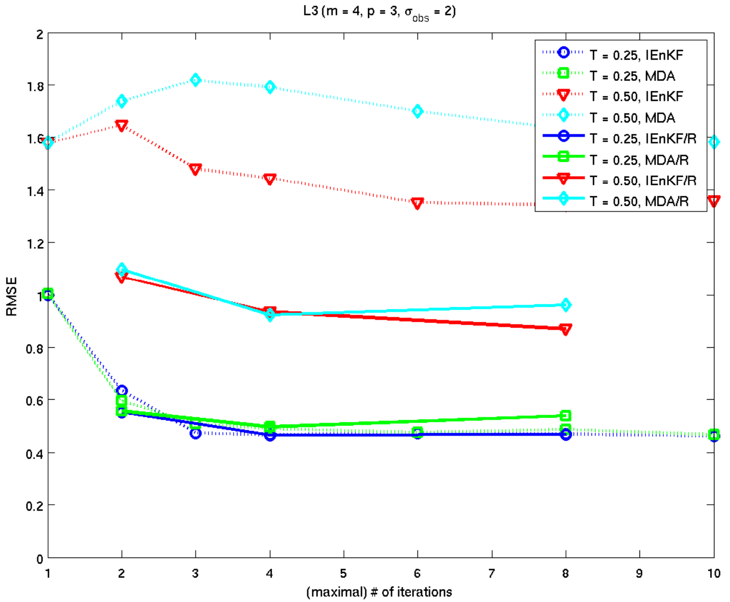
How bad is going local?



Inflation: prior or posterior?



IEnKF vs. MDA


L3 (m = 4, p = 3, σ_obs = 2)
T = 0.25, IEnKF
T = 0.25, MDA
T = 0.50, IEnKF
T = 0.50, MDA
T = 0.25, IEnKF/R
T = 0.25, MDA/R
T = 0.50, IEnKF/R
T = 0.50, MDA/R
RMSE
(maximal) # of iterations
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
1
2
3
4
5
6
7
8
9
10

Scaled rotations

Givens rotation:

$$\hat{\mathbf{U}}(i,j,\theta) = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \cos(\theta) & \dots & -\sin(\theta) & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \sin(\theta) & \dots & \cos(\theta) & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

(col. i) (col. j) (row i) (row j)

$$\hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{T}} = 1$$

General rotation:

$$\mathbf{U} = \prod_{i,j=1;\ i\geq j}^{m} \hat{\mathbf{U}}(i,j,\theta_{ij})$$

Scaled rotation:

$$\mathbf{U}(\varepsilon) = \prod_{i,j=1;\ i\geq j}^{m} \mathbf{U}(i,j,\varepsilon\theta_{ij}), \qquad \varepsilon \in [0,1]$$

Conclusions

- Ensemble redrawing is a useful option for iterative schemes
- Can be particularly useful in strongly nonlinear situations
- It can also be useful to handle occasional instabilities
- A system rollback is also often required to achieve positive results
- The ensemble redrawing can be localised by using scaled ensemble rotations

Other results:

- IEnKF-N seems to have a better baseline performance in strongly nonlinear situations
- Prior inflation seems to work better than the posterior inflation in stronlgy nonlinear situations
- MDA can often perform almost equally with the IEnKF

Thank you!

References

Bocquet, M., 2011: Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Proc. Geoph.*, **18**, 735–750.

Bocquet, M. and P. Sakov, 2012: Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlinear Proc. Geoph.*, **19**, 383–399.

Emerick, A. A. and A. C. Reynolds, 2013: History matching time-lapse seismic data using the ensemble kalman filter with multiple data assimilations. *Computat. Geosci.*, **16**, 639–659.

Gu, Y. and D. S. Oliver, 2007: An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE Journal*, **12**, 438–446.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Sakov, P., D. S. Oliver, and L. Bertino, 2012: An iterative EnKF for strongly nonlinear systems. *Mon. Wea. Rev.*, **140**, 1988–2004.

Yang, S.-C., E. Kalnay, and B. Hunt, 2012: Handling nonlinearity and non-Gaussianity in the Ensemble Kalman Filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, **140**, 2628–2646.