Extending the square root method to account for model noise in the ensemble Kalman filter

Patrick Nima Raanes[*,1,2],
Alberto Carrassi[1], and Laurent Bertino[1]

[1]Nansen Environmental and Remote Sensing Center
[2]Mathematical Institute, University of Oxford

Os, June 9, 2015

[*]email: patrick.n.raanes@gmail.com

Paper

MWR special issue on DA, 2015 ?

Abstract: A square root approach is considered for the problem of accounting for model noise in the forecast step of the ensemble Kalman filter (EnKF) and related algorithms. Primarily intended to replace additive, pseudo-random noise simulation, the core method is based on the analysis step of ensemble square root filters, and consists in the deterministic computation of a transform matrix. The theoretical advantages regarding dynamical consistency are surveyed, applying equally well to the square root method in the analysis step. A fundamental problem due to the limited size of the ensemble subspace is discussed, and novel solutions that complement the core method are suggested and studied. Benchmarks from twin experiments with simple, low-order dynamics indicate improved performance over standard approaches such as additive, simulated noise and multiplicative inflation.

Model noise – Problem statement

Assume

$$\boldsymbol{x}^{t+1} = f(\boldsymbol{x}^t) + \boldsymbol{q}^t\,, \qquad \text{where} \qquad \boldsymbol{q}^t \sim \mathcal{N}(0, \mathbf{Q})\,, \tag{1}$$

with f and $\mathbf{Q} = \mathbb{C}\text{ov}(\boldsymbol{q})$ perfectly known.

Model noise – Problem statement

Assume

$$\boldsymbol{x}^{t+1} = f(\boldsymbol{x}^t) + \boldsymbol{q}^t \,, \qquad \text{where} \qquad \boldsymbol{q}^t \sim \mathcal{N}(0, \mathbf{Q}) \,, \tag{1}$$

with f and $\mathbf{Q} = \mathbb{C}\text{ov}(\boldsymbol{q})$ perfectly known.

Then we want the forecast ensemble to satisfy

$$\bar{\mathbf{P}}^f = \bar{\mathbf{P}} + \mathbf{Q} \,, \tag{2}$$

where

$$\bar{\mathbf{P}} = \frac{1}{N-1} \sum_n (\boldsymbol{x}_n - \bar{\boldsymbol{x}})(\boldsymbol{x}_n - \bar{\boldsymbol{x}})^{\mathrm{T}} \,. \tag{3}$$

Outline

SQRT-CORE

Initial comparisons

Residual noise treatment

Further experiments

Outline

Sqrt-Core

Initial comparisons

Residual noise treatment

Further experiments

Lessons learnt the past 15 years

Any square root update, $\mathbf{A} \mapsto \mathbf{AT}$, will

- deterministically match covariance relations
- preserve the ensemble subspace
- satisfy linear, homogeneous, equality constraints $\star$

Lessons learnt the past 15 years

Any square root update, $\mathbf{A} \mapsto \mathbf{AT}$, will

- deterministically match covariance relations
- preserve the ensemble subspace
- satisfy linear, homogeneous, equality constraints $\star$

Furthermore, the "symmetric choice", $\mathbf{A} \mapsto \mathbf{AT}_s$, will

- preserve the mean
- satisfy linear, inhomogeneous constraints $\star$
- satisfy the first-order approximation to non-linear constraints $\star$
- minimise ensemble displacement $\star$
- yield equally likely realisations $\star$

$\star$: (plausibly) improves "dynamical consistency" of realisations.

Sqrt-Core

$\bar{\mathbf{P}}^f = \bar{\mathbf{P}} + \mathbf{Q}$ can be rewritten using $\bar{\mathbf{P}} = \frac{1}{N-1}\mathbf{A}\mathbf{A}^{\mathrm{T}}$, yielding:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N-1)\mathbf{Q}\,. \tag{4}$$

Sqrt-Core

$\bar{\mathsf{P}}^f = \bar{\mathsf{P}} + \mathsf{Q}$ can be rewritten using $\bar{\mathsf{P}} = \frac{1}{N-1}\mathsf{A}\mathsf{A}^{\mathrm{T}}$, yielding:

$$\mathsf{A}^f {\mathsf{A}^f}^{\mathrm{T}} = \mathsf{A}\mathsf{A}^{\mathrm{T}} + (N-1)\mathsf{Q}\,. \tag{4}$$

(Brutally) factorising out A using the M-P pseudoinverse, A^+:

$$\mathsf{A}^f {\mathsf{A}^f}^{\mathrm{T}} = \mathsf{A}\left(\mathsf{I}_N + (N-1)\mathsf{A}^+\mathsf{Q}(\mathsf{A}^{\mathrm{T}})^+\right)\mathsf{A}^{\mathrm{T}}\,, \tag{5}$$

Sqrt-Core

$\bar{\mathbf{P}}^f = \bar{\mathbf{P}} + \mathbf{Q}$ can be rewritten using $\bar{\mathbf{P}} = \frac{1}{N-1}\mathbf{A}\mathbf{A}^{\mathrm{T}}$, yielding:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N-1)\mathbf{Q}\,. \tag{4}$$

(Brutally) factorising out $\mathbf{A}$ using the M-P pseudoinverse, $\mathbf{A}^+$:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\left(\mathbf{I}_N + (N-1)\mathbf{A}^+\mathbf{Q}(\mathbf{A}^{\mathrm{T}})^+\right)\mathbf{A}^{\mathrm{T}}\,, \tag{5}$$

we get Sqrt-Core:

$$\mathbf{A}^f = \mathbf{A}\mathbf{T}_s^f \tag{6}$$

where $\mathbf{T}_s^f$ is the sym. square root of the middle factor in eqn. (5).

Sqrt-Core

$\bar{\mathbf{P}}^f = \bar{\mathbf{P}} + \mathbf{Q}$ can be rewritten using $\bar{\mathbf{P}} = \frac{1}{N-1}\mathbf{A}\mathbf{A}^{\mathrm{T}}$, yielding:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N-1)\mathbf{Q}\,. \tag{4}$$

(Brutally) factorising out $\mathbf{A}$ using the M-P pseudoinverse, $\mathbf{A}^+$:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\left(\mathbf{I}_N + (N-1)\mathbf{A}^+\mathbf{Q}(\mathbf{A}^{\mathrm{T}})^+\right)\mathbf{A}^{\mathrm{T}}\,, \tag{5}$$

we get Sqrt-Core:

$$\mathbf{A}^f = \mathbf{A}\mathbf{T}_s^f \tag{6}$$

where $\mathbf{T}_s^f$ is the sym. square root of the middle factor in eqn. (5).

We also see that the problem of eqn. (4) is ill-posed...

SQRT-CORE

In fact, SQRT-CORE only satisfies

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N-1)\hat{\mathbf{Q}} \tag{7}$$

where $\hat{\mathbf{Q}} = \mathbf{\Pi_A Q \Pi_A}$, and $\mathbf{\Pi_A} = \mathbf{A}\mathbf{A}^{+}$ is the orthogonal projector onto the column space of $\mathbf{A}$.

Outline

Sqrt-Core

Initial comparisons

Residual noise treatment

Further experiments

Overview of alternatives

Method	$\mathbf{A}^f =$	**where**
ADD-Q	$\mathbf{A} + \mathbf{D}$	$\mathbf{D} = \mathbf{Q}^{1/2}\Xi, \quad \xi_n \sim \mathcal{N}(0, \mathbf{I}_m)$
MULT-1	$\lambda\mathbf{A}$	$\lambda^2 = \frac{\mathrm{trace}(\bar{\mathbf{P}}+\mathbf{Q})}{\mathrm{trace}(\bar{\mathbf{P}})}$
MULT-m	$\mathbf{\Lambda}\mathbf{A}$	$\mathbf{\Lambda}^2 = \mathrm{diag}(\bar{\mathbf{P}})^{-1}\,\mathrm{diag}(\bar{\mathbf{P}} + \mathbf{Q})$
SQRT-CORE	$\mathbf{A}\mathbf{T}$	$\mathbf{T} = \left(\mathbf{I}_N + (N-1)\mathbf{A}^{+}\mathbf{Q}\mathbf{A}^{+\mathrm{T}}\right)_s^{1/2}$

Overview of alternatives

Method	$\mathbf{A}^f =$	where	thus satisfying
Add-Q	$\mathbf{A} + \mathbf{D}$	$\mathbf{D} = \mathbf{Q}^{1/2}\Xi, \quad \xi_n \sim \mathcal{N}(0, \mathbf{I}_m)$	$\mathbb{E}_{\mathbf{D}}$(eqn. (4))
Mult-1	$\lambda\mathbf{A}$	$\lambda^2 = \frac{\mathrm{trace}(\bar{\mathbf{P}}+\mathbf{Q})}{\mathrm{trace}(\bar{\mathbf{P}})}$	trace(eqn. (4))
Mult-m	$\mathbf{\Lambda}\mathbf{A}$	$\mathbf{\Lambda}^2 = \mathrm{diag}(\bar{\mathbf{P}})^{-1}\,\mathrm{diag}(\bar{\mathbf{P}} + \mathbf{Q})$	diag(eqn. (4))
Sqrt-Core	$\mathbf{A}\mathbf{T}$	$\mathbf{T} = \left(\mathbf{I}_N + (N-1)\mathbf{A}^{+}\mathbf{Q}\mathbf{A}^{+\mathrm{T}}\right)_s^{1/2}$	$\mathbf{\Pi}_{\mathbf{A}}$(eqn. (4))$\mathbf{\Pi}_{\mathbf{A}}$

Overview of alternatives

Method	$\mathbf{A}^f =$	where	thus satisfying
Add-Q	$\mathbf{A}+\mathbf{D}$	$\mathbf{D}=\mathbf{Q}^{1/2}\Xi, \quad \xi_n \sim \mathcal{N}(0,\mathbf{I}_m)$	$\mathbb{E}_{\mathbf{D}}$(eqn. (4))
Mult-1	$\lambda\mathbf{A}$	$\lambda^2 = \frac{\mathrm{trace}(\bar{\mathbf{P}}+\mathbf{Q})}{\mathrm{trace}(\bar{\mathbf{P}})}$	trace(eqn. (4))
Mult-m	$\mathbf{\Lambda}\mathbf{A}$	$\mathbf{\Lambda}^2 = \mathrm{diag}(\bar{\mathbf{P}})^{-1}\,\mathrm{diag}(\bar{\mathbf{P}}+\mathbf{Q})$	diag(eqn. (4))
Sqrt-Core	$\mathbf{A}\mathbf{T}$	$\mathbf{T}=\left(\mathbf{I}_N+(N-1)\mathbf{A}^{+}\mathbf{Q}\mathbf{A}^{+\mathrm{T}}\right)_s^{1/2}$	$\mathbf{\Pi}_{\mathbf{A}}$(eqn. (4))$\mathbf{\Pi}_{\mathbf{A}}$

Also:

- Complete resampling
- 2nd-order exact sampling (Pham, 2001)
- A similar (but distinct) square root method (Nakano, 2013)
- Relaxation (Zhang et al., 2004)
- Forcings fields or boundary conditions (Shutts, 2005)
- SEIK, with forgetting factor (Pham, 2001)
- RRSQRT, with orthogonal ensemble (Heemink et al., 2001)

Snapshot comparison

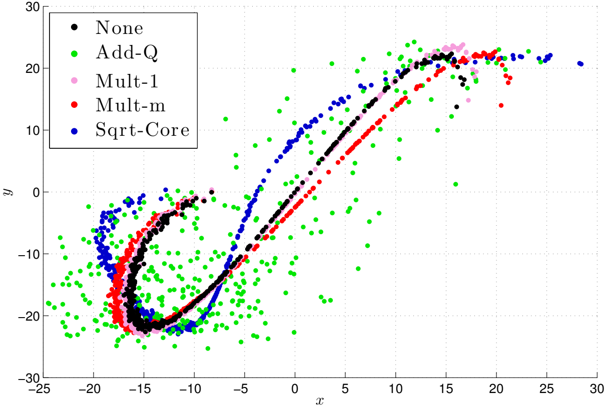


Figure: Snapshot of ensemble forecasts with the Lorenz-63 system after model noise incorporation.

Experimental setup

- Twin experiment: tracking a simulated "truth", $\boldsymbol{x}^t$
- RMSE $= \sqrt{\frac{1}{m}\|\bar{\boldsymbol{x}}^t - \boldsymbol{x}^t\|_2^2}$
- Analysis update:
  - ETKF (using the symmetric square root)
  - No localisation
  - Inflation (for analysis update errors): tuned for $\textsc{Add-Q}$
- Baselines

Lorenz-63 – system

Integrated with RK4:

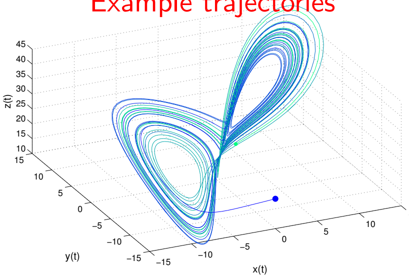
$r = 28$, $\sigma = 10$, and $b = 8/3$.

$$\dot{x} = \sigma(y - x)\,,$$
$$\dot{y} = rx - y - xz\,,$$
$$\dot{z} = xy - bz\,,$$

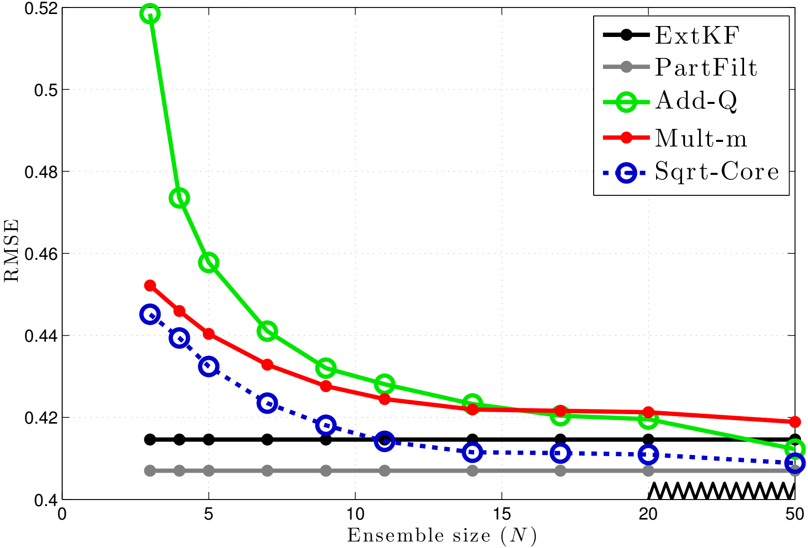
Direct observations of the entire state, with $\mathbf{R} = 2\mathbf{I}_3$.

$$\mathbf{Q} = \begin{bmatrix} 10 & -2 & 3 \\ -2 & 5 & 3 \\ 3 & 3 & 5 \end{bmatrix} /10\,.$$

Example trajectories

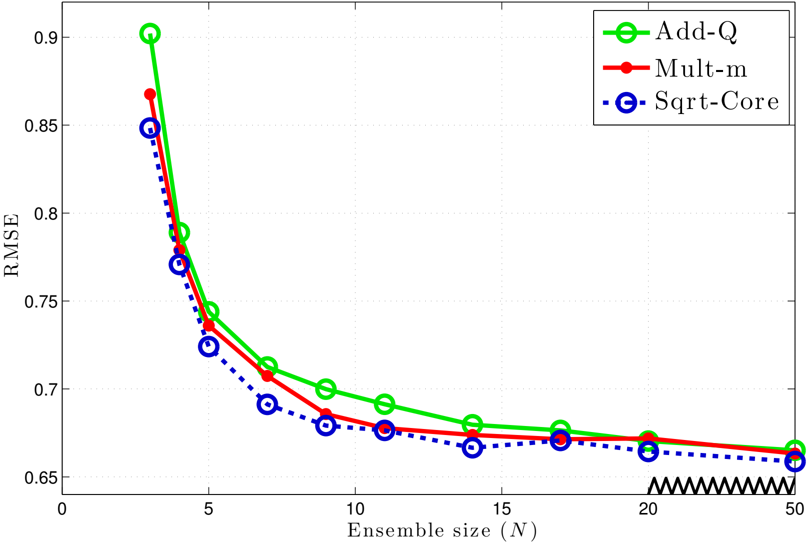


Lorenz-63 – vs N, with $\Delta t_{\text{obs}} = 0.05$



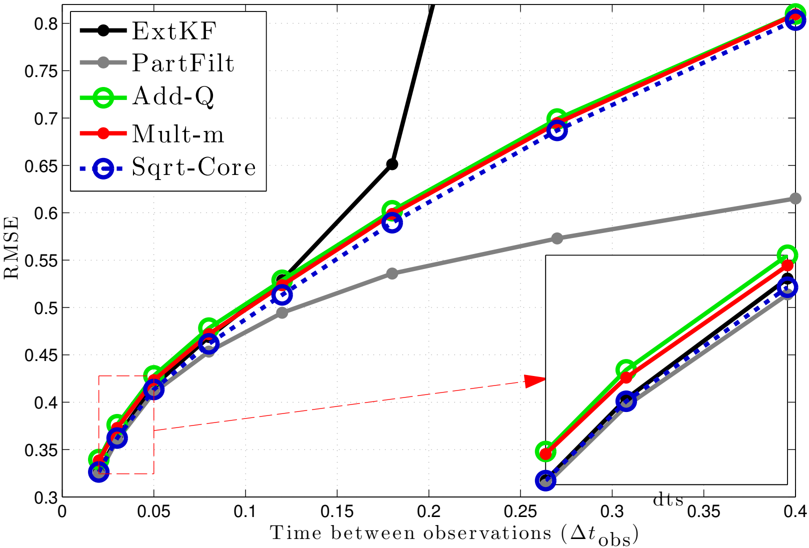
where the particle filter uses $N = 10^4$.

Lorenz-63 – vs N, with $\Delta t_{\text{obs}} = 0.25$

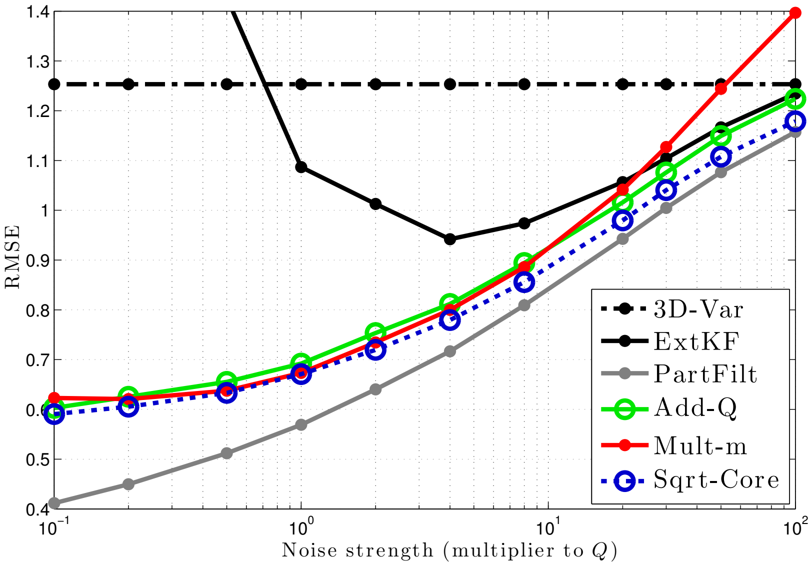


Particle filter RMSE: 0.57. Extended Kalman filter RMSE: 1.4.

Lorenz-63 – vs Δt_{obs}, with $N = 12$



Lorenz-63 – vs **Q** multiplier, with $N = 12$, $\Delta t_{\text{obs}} = 21$



Outline

SQRT-CORE

Initial comparisons

Residual noise treatment

Further experiments

Improving SQRT-CORE: Residual noise treatment

After SQRT-CORE there is still $[\mathbf{Q} - \hat{\mathbf{Q}}]$ unaccounted for.

$\implies$ Residual noise problem:

$$\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N{-}1)[\mathbf{Q} - \hat{\mathbf{Q}}]\,. \tag{8}$$

Note: notation recycled from original problem.

A first approach – Sqrt-Add-Z

1. Define $\mathbf{Z} = (\mathbf{I}_m - \mathbf{\Pi}_{\mathbf{A}})\mathbf{Q}^{1/2}$.
2. Add $\tilde{\boldsymbol{q}}_n = \mathbf{Z}\tilde{\boldsymbol{\xi}}_n$ to realisation n, with $\tilde{\boldsymbol{\xi}}_n \sim \mathcal{N}(0, \mathbf{I}_m)$.

A first approach – Sqrt-Add-Z

1. Define $\mathbf{Z} = (\mathbf{I}_m - \mathbf{\Pi}_{\mathbf{A}})\mathbf{Q}^{1/2}$.
2. Add $\tilde{\boldsymbol{q}}_n = \mathbf{Z}\tilde{\boldsymbol{\xi}}_n$ to realisation n, with $\tilde{\boldsymbol{\xi}}_n \sim \mathcal{N}(0, \mathbf{I}_m)$.

But due to cross-terms, $\mathbf{Z}$ is not a square root of $[\mathbf{Q} - \hat{\mathbf{Q}}]$, and therefore Sqrt-Add-Z is biased:

$$\mathbb{E}_{\{\tilde{\xi}\}}\left(\mathbf{A}^f {\mathbf{A}^f}^{\mathrm{T}}\right) = \mathbf{A}\mathbf{A}^{\mathrm{T}} + (N-1)[\mathbf{Q} - \hat{\mathbf{Q}}] - \left(\hat{\mathbf{Q}}^{1/2}\mathbf{Z}^{\mathrm{T}} + \mathbf{Z}\hat{\mathbf{Q}}^{\mathrm{T}/2}\right) .$$

Compare to eqn. (8).

The underlying problem: replacing one draw with two

As an analogy to the "core+residual" problem, define

$$\boldsymbol{q} = \hat{\mathsf{Q}}^{1/2}\boldsymbol{\xi} + \mathsf{Z}\boldsymbol{\xi}\,, \tag{9}$$

$$\boldsymbol{q}^{\perp\!\!\!\perp} = \hat{\mathsf{Q}}^{1/2}\hat{\boldsymbol{\xi}} + \mathsf{Z}\tilde{\boldsymbol{\xi}}\,, \tag{10}$$

where $\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}, \tilde{\boldsymbol{\xi}} \sim \mathcal{N}(0, \mathsf{I}_m)$ are all independent.

Note that

$$\mathbb{C}\mathrm{ov}(\boldsymbol{q}) = \mathsf{Q} = \underbrace{\hat{\mathsf{Q}} + \mathsf{Z}\mathsf{Z}^{\mathrm{T}}}_{\mathbb{C}\mathrm{ov}(\boldsymbol{q}^{\perp\!\!\!\perp})} + \left(\hat{\mathsf{Q}}^{1/2}\mathsf{Z}^{\mathrm{T}} + \mathsf{Z}\hat{\mathsf{Q}}^{\mathrm{T}/2}\right)\,. \tag{11}$$

Reintroducing dependence – Sqrt-Dep

Let $\mathbf{\Pi}$ be any orthogonal projection matrix, and define

$$\boldsymbol{\xi}^{\perp\!\!\!\perp} = \mathbf{\Pi}\hat{\boldsymbol{\xi}} + (\mathbf{I}_m - \mathbf{\Pi})\tilde{\boldsymbol{\xi}}\,, \tag{12}$$

where, as before, $\hat{\boldsymbol{\xi}}, \tilde{\boldsymbol{\xi}} \sim \mathcal{N}(0, \mathbf{I}_m)$ are independent.

But, $\boldsymbol{\xi}^{\perp\!\!\!\perp} \sim \mathcal{N}(0, \mathbf{I}_m)$ too (no cross terms)!

Choose $\mathbf{\Pi}$ so that $\mathbf{Z\Pi} = 0$. Rather than eqn. (9), redefine $\boldsymbol{q}$:

$$\boldsymbol{q} = \mathbf{Q}^{1/2}\boldsymbol{\xi}^{\perp\!\!\!\perp}\,. \tag{13}$$

Then,

$$\mathbb{C}\text{ov}(\boldsymbol{q}) = \mathbf{Q}\,. \tag{14}$$

The solution: reintroducing dependence – SQRT-DEP

But also:

$$\boldsymbol{q} = (\hat{\mathbf{Q}}^{1/2} + \mathbf{Z}) \left(\mathbf{\Pi} \hat{\boldsymbol{\xi}} + (\mathbf{I}_m - \mathbf{\Pi}) \tilde{\boldsymbol{\xi}} \right) \tag{15}$$

$$= \hat{\mathbf{Q}}^{1/2} \hat{\boldsymbol{\xi}} + \mathbf{Z} \left(\mathbf{\Pi} \hat{\boldsymbol{\xi}} + (\mathbf{I}_m - \mathbf{\Pi}) \tilde{\boldsymbol{\xi}} \right) . \tag{16}$$

Hence, while maintaining $\mathbb{C}\mathrm{ov}(\boldsymbol{q}) = \mathbf{Q}$, the influence of $\tilde{\boldsymbol{\xi}}$ has been confined to $\mathrm{span}(\mathbf{Z}) = \mathrm{span}(\mathbf{A})^{\perp}$.

The solution: reintroducing dependence – Sqrt-Dep

But also:

$$\boldsymbol{q} = (\hat{\mathbf{Q}}^{1/2} + \mathbf{Z}) \left(\mathbf{\Pi} \hat{\boldsymbol{\xi}} + (\mathbf{I}_m - \mathbf{\Pi}) \tilde{\boldsymbol{\xi}} \right) \tag{15}$$

$$= \hat{\mathbf{Q}}^{1/2} \hat{\boldsymbol{\xi}} + \mathbf{Z} \left(\mathbf{\Pi} \hat{\boldsymbol{\xi}} + (\mathbf{I}_m - \mathbf{\Pi}) \tilde{\boldsymbol{\xi}} \right) . \tag{16}$$

Hence, while maintaining $\mathbb{C}\mathrm{ov}(\boldsymbol{q}) = \mathbf{Q}$, the influence of $\tilde{\boldsymbol{\xi}}$ has been confined to $\mathrm{span}(\mathbf{Z}) = \mathrm{span}(\mathbf{A})^{\perp}$.

Algorithm: for each realisation:

1. Compute $\hat{\boldsymbol{\xi}}_n$ corresponding to Sqrt-Core
2. Draw $\tilde{\boldsymbol{\xi}}_n$
3. Total (core+residual) update: eqn. (16)

Outline

SQRT-CORE

Initial comparisons

Residual noise treatment

Further experiments

Overview of alternatives

Method	$\mathbf{A}^f =$	where
ADD-Q	$\mathbf{A} + \mathbf{D}$	$\mathbf{D} = \mathbf{Q}^{1/2}\Xi$, each column of Ξ drawn from $\mathcal{N}(0, \mathbf{I}_m)$
MULT-1	$\lambda\mathbf{A}$	$\lambda^2 = \frac{\text{trace}(\bar{\mathbf{P}}+\mathbf{Q})}{\text{trace}(\bar{\mathbf{P}})}$
MULT-m	$\mathbf{\Lambda}\mathbf{A}$	$\mathbf{\Lambda}^2 = \text{diag}(\bar{\mathbf{P}})^{-1}\,\text{diag}(\bar{\mathbf{P}} + \mathbf{Q})$
SQRT-CORE	$\mathbf{A}\mathbf{T}$	$\mathbf{T} = \left(\mathbf{I}_N + (N-1)\mathbf{A}^{+}\mathbf{Q}\mathbf{A}^{+\mathrm{T}}\right)_s^{1/2}$

Also:

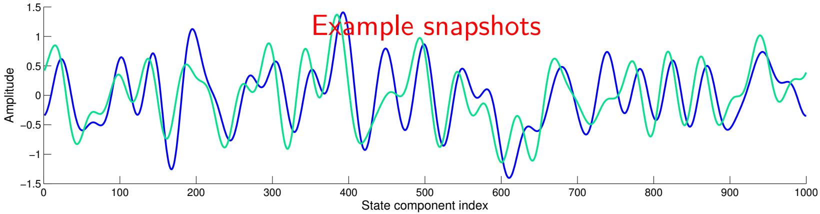
- SQRT-ADD-Z
- SQRT-DEP

Linear advection – system

For $t = 0, 1, \ldots$ and $i = 1, \ldots, m$, and with periodic BCs,

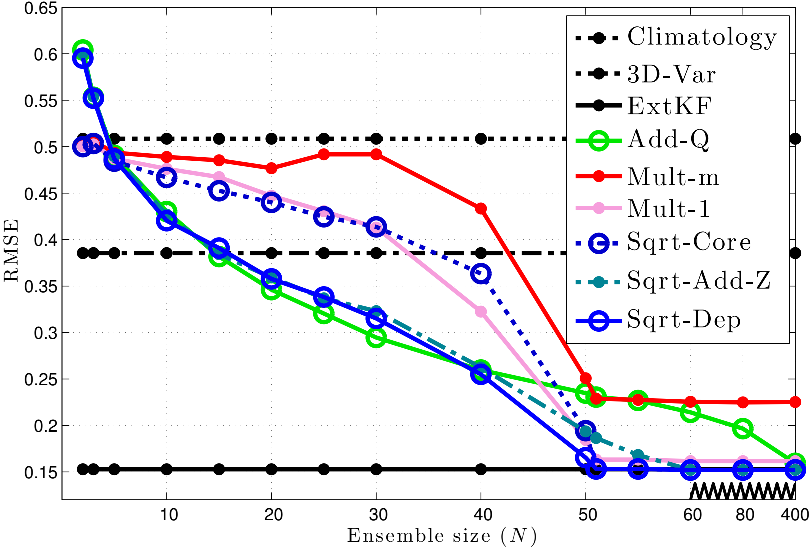
$$x_i^{t+1} = 0.98 x_{i-1}^{t} \, . \tag{17}$$

Direct observation of the truth at $p = 40$ equidistant locations with $\mathbf{R} = 0.01\mathbf{I}_p$, every seventh time step: $\Delta t_{\text{obs}} = 7\Delta t = 4.9$.



Q is such that although $m = 1000$,
the system subspace only has 50 dimensions.

Linear advection – results



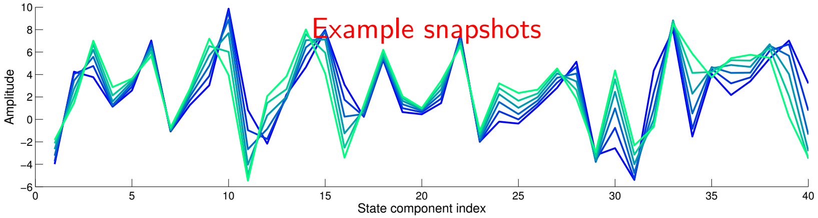
Lorenz-96 – system

Integrated with RK4,

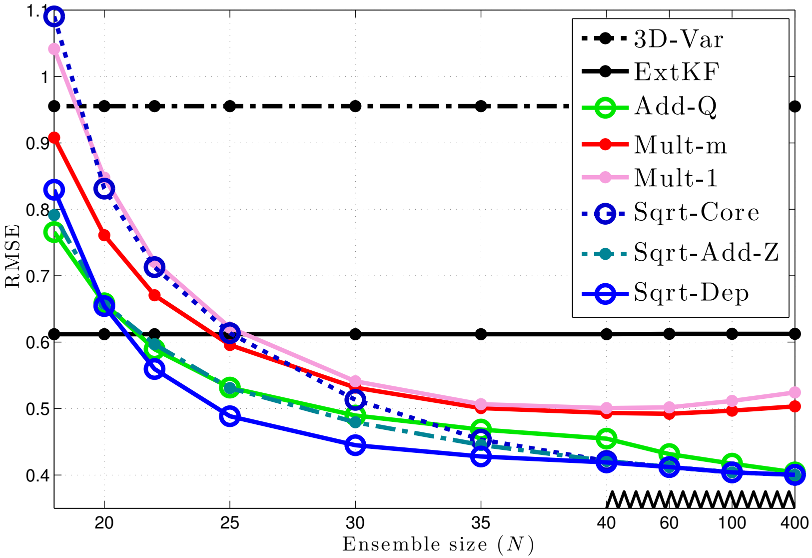
$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})\, x_{i-1} - x_i + F\,, \tag{18}$$

with periodic BCs, $i = 1, \ldots, m$, $m = 40$, and

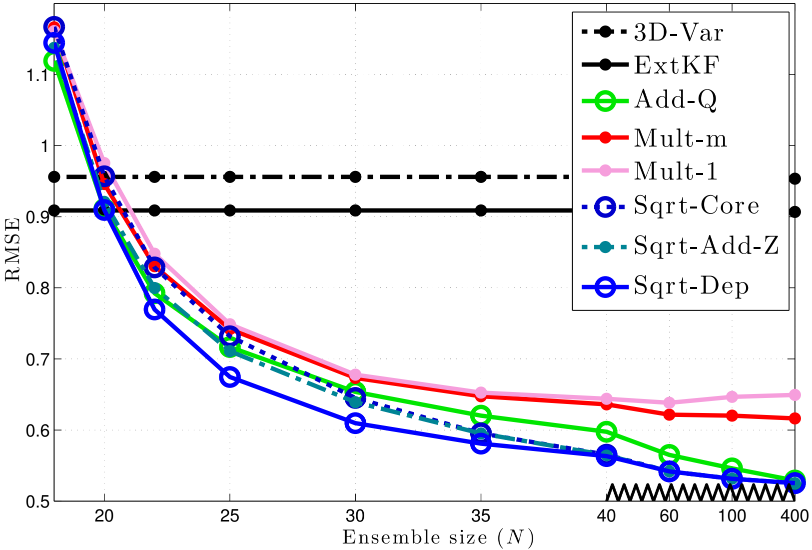
$$Q_{i,j} = \exp\left(-1/30\|i-j\|_2^2\right) + 0.1\delta_{i,j}\,. \tag{19}$$



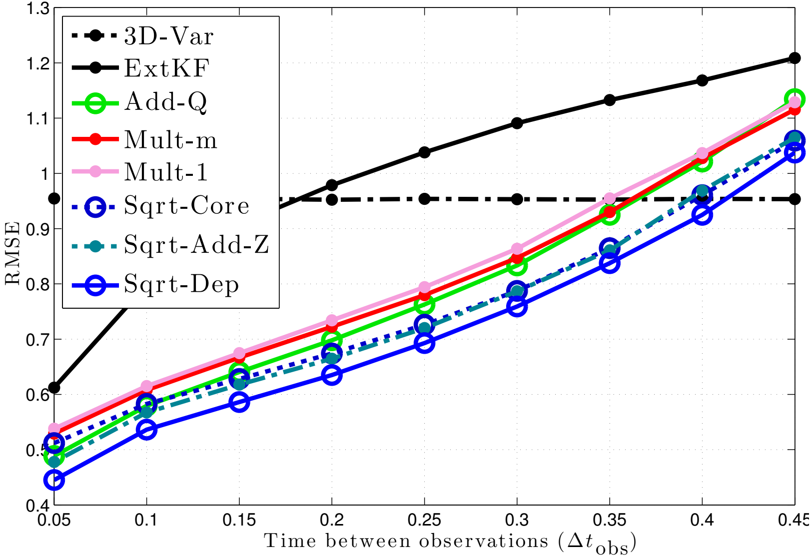
Lorenz-96 – vs N, with $\Delta t_{\text{obs}} = 0.05$



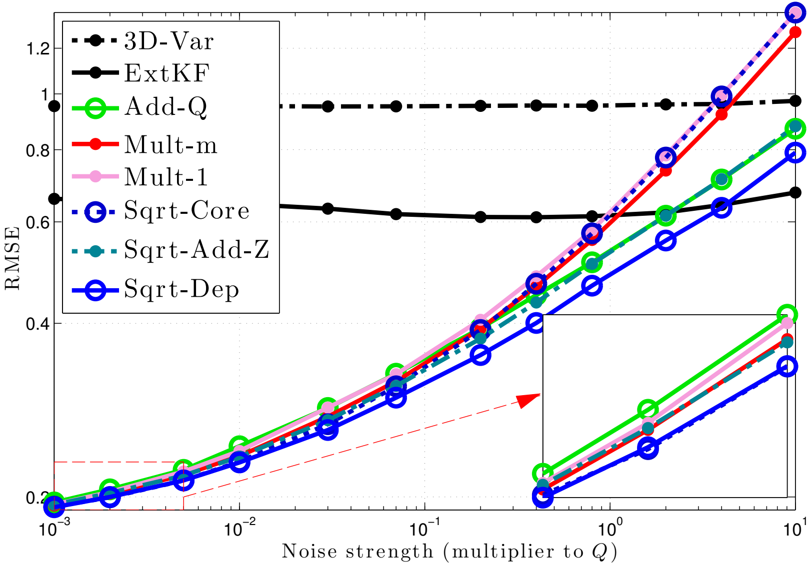
Lorenz-96 – vs N, with $\Delta t_{\text{obs}} = 0.15$



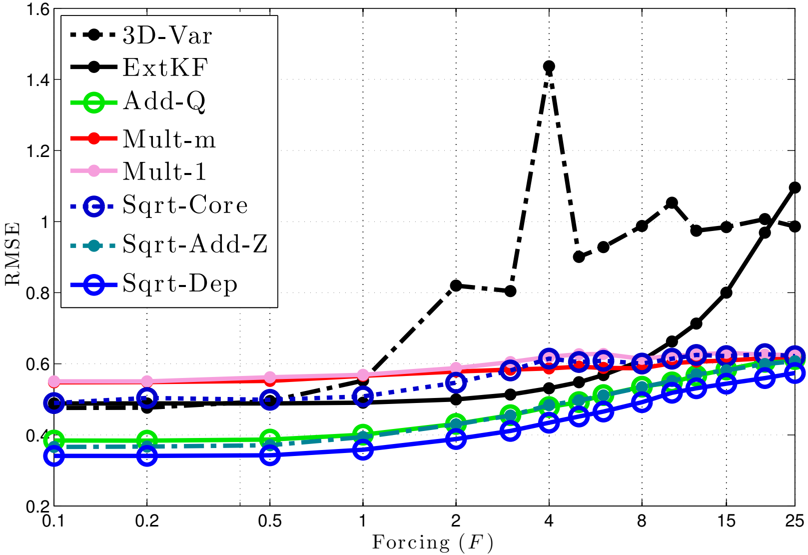
Lorenz-96 – vs Δt_{obs}, with $N = 30$



Lorenz-96 – vs **Q** multiplier, with $N = 25$, $\Delta t_{\text{obs}} = 0.05$



Lorenz-96 – vs F, with $N = 25$, $\Delta t_{\text{obs}} = 0.05$



Summary

- Extending the square root method to the forecast step
  - Main aim: eliminate sampling errors
  - Secondary benefit: dynamical consistency

Summary

- Extending the square root method to the forecast step
  - Main aim: eliminate sampling errors
  - Secondary benefit: dynamical consistency
- Sqrt-Core is deficient when $[\mathbf{Q} - \hat{\mathbf{Q}}]$ is significant
  - Sqrt-Add-Z is simple and efficient, but biased
  - Sqrt-Dep is costly, but more satisfactory
  - Both methods perform robustly better than Mult-m and Add-Q

Summary

- Extending the square root method to the forecast step
  - Main aim: eliminate sampling errors
  - Secondary benefit: dynamical consistency
- Sqrt-Core is deficient when $[\mathbf{Q} - \hat{\mathbf{Q}}]$ is significant
  - Sqrt-Add-Z is simple and efficient, but biased
  - Sqrt-Dep is costly, but more satisfactory
  - Both methods perform robustly better than Mult-m and Add-Q
- Future directions
  - Experiments on larger models and more realistic model error
  - Improvements to Sqrt-Add-Z and Sqrt-Dep
  - Investigate perspectives from Nakano (2013) and M. Bocquet

References

Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance reduced ensemble Kalman filtering. *Monthly Weather Review*, **129**, 1718–1728.

Nakano, S., 2013: A prediction algorithm with a limited number of particles for state estimation of high-dimensional systems. *Information Fusion (FUSION), 2013 16th International Conference on*, IEEE, 1356–1363.

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review*, **129**, 1194–1207.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3079–3102.

Whitaker, J. S. and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, **140**, 3078–3089.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Monthly Weather Review*, **132**, 1238–1253.