Finite-size ensemble Kalman filters (EnKF-N)

Marc Bocquet[1,2], Pavel Sakov[3,4]

(1) Université Paris-Est, CEREA, joint lab École des Ponts ParisTech and EdF R&D, France
(2) INRIA, Paris-Rocquencourt Research center, France
(3) BOM, Melbourne, Australia
(4) NERSC, Bergen, Norway

(bocquet@cerea.enpc.fr)

Outline

1. The primal EnKF-N
2. The dual EnKF-N
3. Inflation-free iterative ensemble Kalman filters (IEnKF-N)
4. Conclusions

Principle of the EnKF-N

- Empirical moments of the ensemble:

$$\overline{\mathbf{x}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}_k\,, \qquad \mathbf{P} = \frac{1}{N-1}\sum_{k=1}^{N}(\mathbf{x}_k-\overline{\mathbf{x}})(\mathbf{x}_k-\overline{\mathbf{x}})^{\mathrm{T}}\,, \tag{1}$$

- The prior of EnKF and the prior of EnKF-N:

$$p(\mathbf{x}|\overline{\mathbf{x}},\mathbf{P}) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}-\overline{\mathbf{x}})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{x}-\overline{\mathbf{x}})\right\}$$

$$p(\mathbf{x}|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N) \propto \left|(\mathbf{x}-\overline{\mathbf{x}})(\mathbf{x}-\overline{\mathbf{x}})^{\mathrm{T}} + \varepsilon_N(N-1)\mathbf{P}\right|^{-\frac{N}{2}}\,, \tag{2}$$

with $\varepsilon_N = 1$ (mean-trusting variant), or $\varepsilon_N = 1+\frac{1}{N}$ (original variant).

- Ensemble space decomposition (ETKF version of the filters): $\mathbf{x} = \overline{\mathbf{x}} + \mathbf{A}\mathbf{w}$.

- The variational principle of the analysis:

$$\mathscr{J}(\mathbf{w}) = \frac{1}{2}(\mathbf{y}-H(\overline{\mathbf{x}}+\mathbf{A}\mathbf{w}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-H(\overline{\mathbf{x}}+\mathbf{A}\mathbf{w})) + \frac{N-1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w}$$

$$\mathscr{J}(\mathbf{w}) = \frac{1}{2}(\mathbf{y}-H(\overline{\mathbf{x}}+\mathbf{A}\mathbf{w}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-H(\overline{\mathbf{x}}+\mathbf{A}\mathbf{w})) + \frac{N}{2}\ln\left(\varepsilon_N + \mathbf{w}^{\mathrm{T}}\mathbf{w}\right)\,. \tag{3}$$

EnKF-N: algorithm

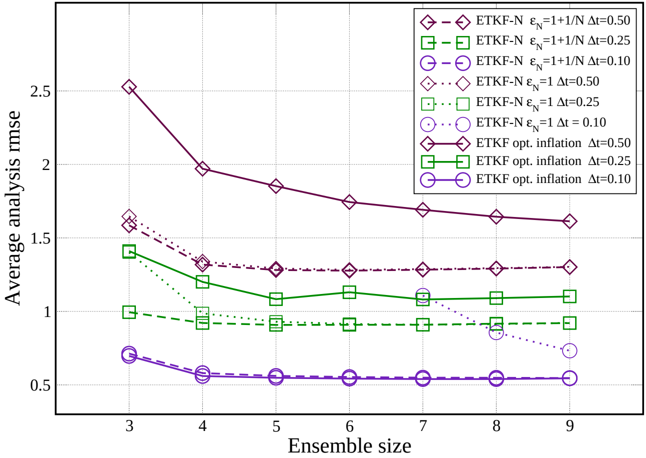
1. Requires: The forecast ensemble $\{\mathbf{x}_k\}_{k=1,\ldots,N}$, the observations $\mathbf{y}$, and error covariance matrix $\mathbf{R}$
2. Compute the mean $\overline{\mathbf{x}}$ and the anomalies $\mathbf{A}$ from $\{\mathbf{x}_k\}_{k=1,\ldots,N}$.
3. Compute $\mathbf{Y} = \mathbf{H}\mathbf{A}$, $\delta = \mathbf{y} - \mathbf{H}\overline{\mathbf{x}}$
4. Find the minimum:

$$\mathbf{w}_a = \min_{\mathbf{w}} \left\{ (\delta - \mathbf{Y}\mathbf{w})^{\mathrm{T}} \mathbf{R}^{-1} (\delta - \mathbf{Y}\mathbf{w}) + N \ln \left(\varepsilon_N + \mathbf{w}^{\mathrm{T}}\mathbf{w} \right) \right\}$$

5. Compute $\mathbf{\Omega}_a = \left(\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y} + N \frac{(\varepsilon_N + \mathbf{w}_a^{\mathrm{T}}\mathbf{w}_a)\mathbf{I}_N - 2\mathbf{w}_a\mathbf{w}_a^{\mathrm{T}}}{(\varepsilon_N + \mathbf{w}_a^{\mathrm{T}}\mathbf{w}_a)^2} \right)^{-1}$
6. Compute $\mathbf{x}^a = \overline{\mathbf{x}} + \mathbf{A}\mathbf{w}_a$.
7. Compute $\mathbf{W}^a = ((N-1)\mathbf{\Omega}_a)^{1/2}\mathbf{U}$
8. Compute $\mathbf{x}_k^a = \mathbf{x}^a + \mathbf{A}\mathbf{W}_k^a$

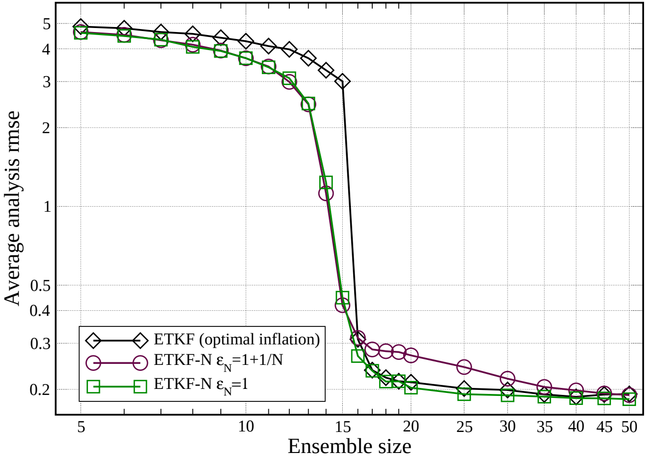
Application to the Lorenz '63 model

- Lorenz '63 toy-model: analysis rmse versus ensemble size for $\Delta t = 0.10, 0.25, 0.50$.



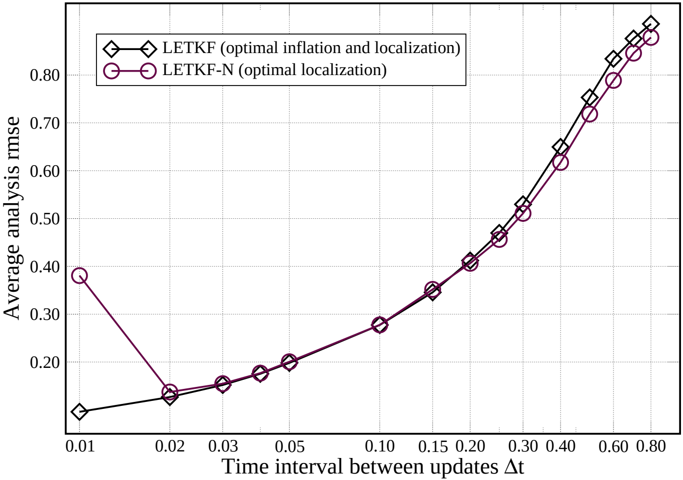
Application to the Lorenz '95 model

- EnKF-N: analysis rmse versus ensemble size, for $\Delta t = 0.05$.



Application to the Lorenz '95 model

- Local version: LETKF-N, with $N = 10$ (in the rank-deficient regime).

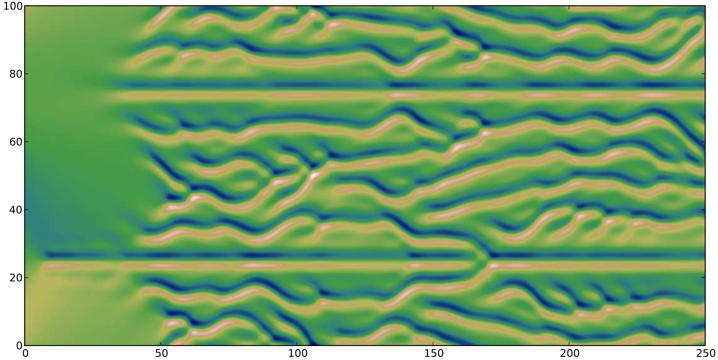


Application to Kuramato-Sivashinski model

- Complex 1D (toy-)model of turbulence [Kuramato et al., 1975; Sivashinski, 1977]

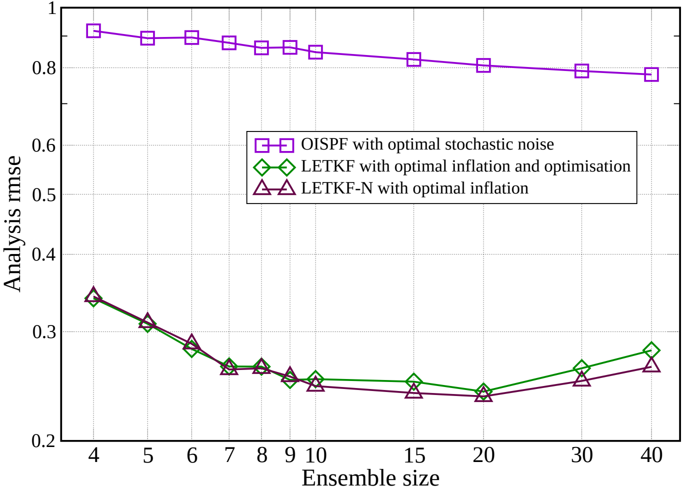
$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} + \frac{\partial^2 u}{\partial^2 x} + \nu \frac{\partial^4 u}{\partial^4 x} = 0 . \qquad (4)$$

$u(x)$ defined on an interval $[0, L]$, with cyclic boundary conditions.



Application to Kuramato-Sivashinski model ($\Delta t = 3$)

- Setup: $L = 32\pi$, $\nu = 1$, $M = 128$, $\Delta t = 3$, $T = 15000$, $\mathbf{R} = \mathbf{I}$.
- Comparison of LETKF, LETKF-N $\varepsilon_N = 1 + \frac{1}{N}$, and OISPF



Outline

Lagrangian duality

- The primal EnKF-N cost function:

$$\mathscr{J}(\mathbf{w}) = \frac{1}{2}(\mathbf{y} - H(\overline{\mathbf{x}} + \mathbf{X}\mathbf{w}))^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y} - H(\overline{\mathbf{x}} + \mathbf{X}\mathbf{w})) + \frac{N}{2} \ln\left(\varepsilon_N + \mathbf{w}^{\mathrm{T}}\mathbf{w}\right). \quad (5)$$

- Idea: We would like to split the radial degree of freedom of $\mathbf{w}$, that is $\sqrt{\mathbf{w}^{\mathrm{T}}\mathbf{w}}$, from its angular degrees of freedom, that is $\mathbf{w}/\sqrt{\mathbf{w}^{\mathrm{T}}\mathbf{w}}$.

- Lagrangian:

$$\mathscr{L}(\mathbf{w}, \rho, \zeta) = \frac{1}{2} g(\mathbf{w}) + \frac{1}{2}\zeta\left(\mathbf{w}^{\mathrm{T}}\mathbf{w} - \rho\right) + \frac{1}{2} f(\rho). \quad (6)$$

where $\delta = \mathbf{y} - \mathbf{H}\overline{\mathbf{x}}$, $g(\mathbf{w}) = (\delta - \mathbf{Y}\mathbf{w})^{\mathrm{T}} \mathbf{R}^{-1} (\delta - \mathbf{Y}\mathbf{w})$, and $f(\rho) = N \ln(\varepsilon_N + \rho)$.

- Dual cost function defined for $\zeta > 0$ by

$$\begin{aligned} \mathscr{D}(\zeta) &= \inf_{\mathbf{w}} \sup_{\rho \geq 0} \mathscr{L}(\mathbf{w}, \rho, \zeta) \\ &= \frac{1}{2}\delta^{\mathrm{T}}\left(\mathbf{R} + \mathbf{Y}\zeta^{-1}\mathbf{Y}^{\mathrm{T}}\right)^{-1}\delta + \frac{\varepsilon_N \zeta}{2} + \frac{N}{2}\ln\frac{N}{\zeta} - \frac{N}{2}. \quad (7) \end{aligned}$$

Non-convex strong duality

- Dual and primal problems:

$$\Delta = \inf_{\zeta > 0} \mathscr{D}(\zeta) \quad \text{and} \quad \Pi = \inf_{\mathbf{w}} \mathscr{J}(\mathbf{w}) . \tag{8}$$

- Strong duality result (non quadratic, non-convex case!!!):

$$\Delta = \Pi . \tag{9}$$

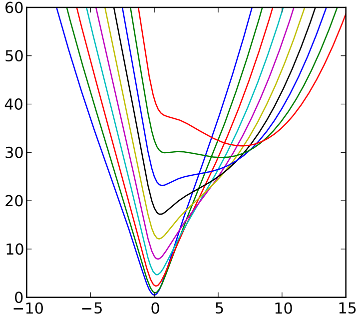
- Saddle point equations:

$$\zeta^\star = \frac{\mathrm{d}f}{\mathrm{d}\rho}(\rho^\star) = \frac{N}{\varepsilon_N + \rho^\star} , \tag{10}$$

$$\zeta^\star \mathbf{w}^\star = \nabla_{\mathbf{w}} g(\mathbf{w}^\star) = -\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}(\delta - \mathbf{Y}\mathbf{w}^\star) . \tag{11}$$

Illustration of the strong duality

Primal and dual cost functions in the one observation case and a series of innovations.



The dual EnKF-N scheme

1. Requires: The forecast ensemble $\{\mathbf{x}_k\}_{k=1,\dots,N}$, the observations $\mathbf{y}$, and error covariance matrix $\mathbf{R}$
2. Compute the mean $\overline{\mathbf{x}}$ and the anomalies $\mathbf{A}$ from $\{\mathbf{x}_k\}_{k=1,\dots,N}$.
3. Compute $\mathbf{Y} = \mathbf{H}\mathbf{A}$, $\delta = \mathbf{y} - \mathbf{H}\overline{\mathbf{x}}$
4. Find the minimum:

$$\zeta^a = \min_{\zeta \in]0,N/\varepsilon_N]} \left\{ \delta^{\mathrm{T}} \left(\mathbf{R} + \mathbf{Y}\zeta^{-1}\mathbf{Y}^{\mathrm{T}} \right)^{-1} \delta + \varepsilon_N \zeta + N \ln \frac{N}{\zeta} - N \right\} \tag{12}$$

5. Compute $\mathbf{\Omega}_a = \left(\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y} + \zeta^a \right)^{-1}$
6. Compute $\mathbf{w}^a = \mathbf{\Omega}^a \mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}\delta$.
7. Compute $\mathbf{x}_a = \overline{\mathbf{x}} + \mathbf{A}\mathbf{w}^a$.
8. Compute $\mathbf{W}^a = ((N-1)\mathbf{\Omega}_a)^{1/2}\mathbf{U}$
9. Compute $\mathbf{x}^a_k = \mathbf{x}^a + \mathbf{A}\mathbf{W}^a_k$

Assets of the dual scheme

- Efficiently finds the global minimum of the EnKF-N cost function.

- Negligible additional cost as compared to the traditional EnKF.

- The algorithm parallels the traditional EnKF.

$\zeta^\star$ can be seen as the effective size of the ensemble. It is related to an optimal inflation of the prior. Standard inflation:

$$\mathbf{x}_k \longrightarrow \overline{\mathbf{x}} + \lambda(\mathbf{x}_k - \overline{\mathbf{x}}), \tag{13}$$

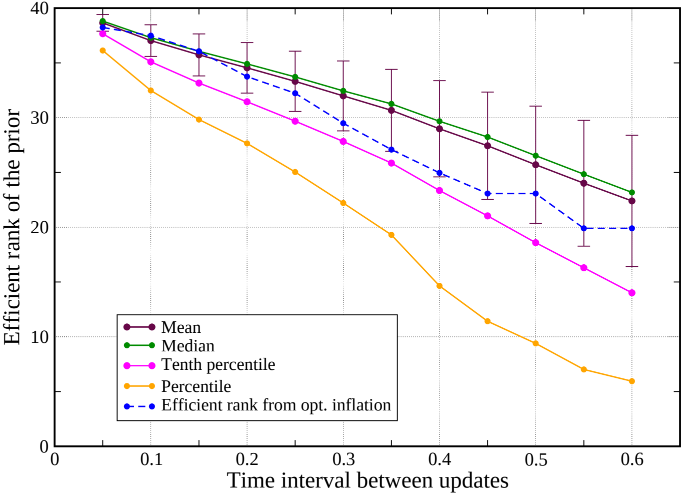
EnKF-N forecasts the following optimal prior inflation factor

$$\lambda^\star = \sqrt{\frac{N-1}{\zeta^\star}}. \tag{14}$$

- Better stability of the dual scheme as compared to the primal scheme in demanding conditions.

- Geometrical explanation of why inflation is so good at counter-acting undersampling (of variances): rotational invariance in ensemble space.

Illustration: statistics of ζ and optimally tuned inflation



Lorenz '95 model, $N = 40$.

Outline

Iterative ensemble Kalman filters

- With strongly nonlinear models, EnKF (as well as EnKF-N) cannot perform well, because the prior is too far off.

- Iterative Kalman filters
  - Original idea: Iterative extended Kalman filter [Wishner et al., 1969; Jazwinski, 1970]
  - More recently in the EnKF context [Gu & Oliver, 2007; Kalnay & Yang, 2010; Sakov, Oliver & Bertino, 2011, Bocquet & Sakov, 2012]
  - Essentially a one-lag smoother. Does the job of a one-lag 4D-Var, with dynamical error covariance matrix and without the use of the TLM and adjoint! Very efficient in very nonlinear conditions if one can afford the multiple ensemble propagations (Lorenz '63 and Lorenz '95).

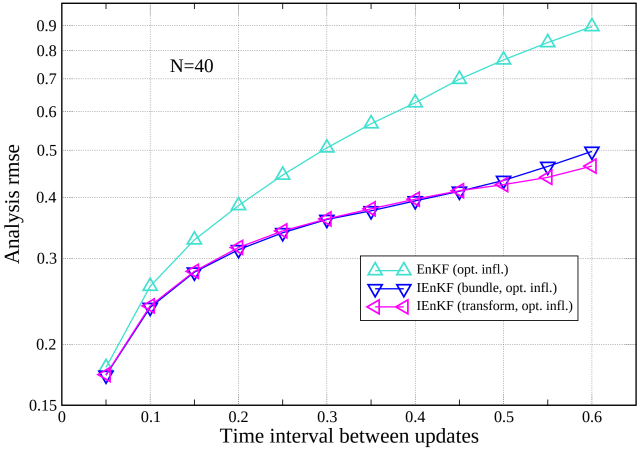
- IEnKF cost function in ensemble space:

$$\begin{aligned}\widetilde{\mathscr{J}}(\mathbf{w}) &= \frac{1}{2}\left(\mathbf{y}_2 - H(\mathscr{M}_{1\to2}(\overline{\mathbf{x}}_1 + \mathbf{A}_1\mathbf{w}))\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}_2 - H(\mathscr{M}_{1\to2}(\overline{\mathbf{x}}_1 + \mathbf{A}_1\mathbf{w}))\right) \\ &\quad + \frac{1}{2}(N-1)\mathbf{w}^{\mathrm{T}}\mathbf{w}, \qquad (15)\end{aligned}$$

and similarly for IEnKF-N.

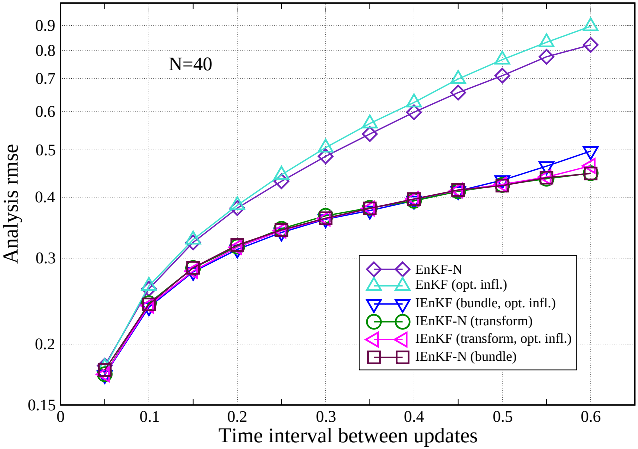
Finite-size iterative ensemble Kalman filters

- Setup: Lorenz '95, $M = 40$, $N = 40$, $\Delta t = 0.05 - 0.60$, $\mathbf{R} = \mathbf{I}$.
- Comparison of EnKF (optimal inflation), IEnKF (bundle and transform, optimal inflation), implementation different from [Sakov et al., 2011]



Finite-size iterative ensemble Kalman filters

- Setup: Lorenz '95, $M = 40$, $N = 40$, $\Delta t = 0.05 - 0.60$, $\mathbf{R} = \mathbf{I}$.
- Comparison of EnKF-N, EnKF (optimal inflation), IEnKF-N (bundle and transform), IEnKF (bundle and transform, optimal inflation)



Outline

Conclusions

- A new prior for the ensemble forecast meant to be used in an EnKF analysis has been built. It takes into account sampling errors.
- It yields a new class of filters EnKF-N, including an ensemble transform version ETKF-N, that does not seem to require inflation supposed to account for sampling errors.
- Local variants (both LA and CL) available.
- The overall performance without inflation compares well with optimally tuned ensemble filters.
- Dual variant EnKF-N is an EnKF with built-in *optimal* inflation (accounting for sampling errors).
- Combination with iterative EnKF schemes is possible.
- Almost linear regime / very small ensemble size more problematic.
- Tests planned in more complex model (e.g. shallow water).

References

1. Ensemble Kalman filtering without the intrinsic need for inflation, M. Bocquet, Nonlin. Processes Geophys., **18**, 735-750, 2011.
2. An iterative EnKF for strongly nonlinear systems, P. Sakov, D. Oliver, and L. Bertino, Mon. Wea. Rev., in press, 2012.
3. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems, M. Bocquet and P. Sakov, Nonlin. Processes Geophys., in press, 2012.