4DVAR by Ensemble Kalman Smoothers

Parallel 4DVAR without tangents and adjoints

J. Mandel, S. Gratton, and E. Bergou

University of Colorado Denver, INP-ENSEEIHT, and CERFACS

Supported by Fondation STAE project ADTAO
and NSF grants AGS-0835579 and DMS-1216481

SIAM UQ, March 31, 2014

Motivation

- 4DVAR sets up a very large nonlinear least squares problem.
- Expensive: iterations, each evaluating the model, tangent and adjoing operators, and solving large linear least squares.
- Extra code for tangent and adjoint operators.
- Iterations may not converge, not even locally.
- Need to parallelize.

Method

- Solve the linear least squares from 4DVAR by EnKS, naturally parallel over the ensemble members.
- Linear algebra glue is cheap, or use parallel dense libraries.
- Finite differences $\Rightarrow$ no tangent and adjoint operators needed.
- Add Tikhonov regularization to the linear least squares $\Rightarrow$ Levelberg-Marquardt method, guaranteed convergence.
- Cheap and simple implementation of Tikhonov regularization within EnKS as an additional observation.

Weak Constraint 4DVAR

- We want to determine $x_0, \ldots, x_k$ (x_i = state at time i) approximately from model and observations (data)

$$
\begin{array}{rcll}
x_0 & \approx & x_\mathrm{b} & \text{state at time 0} \;\approx\; \text{the background} \\
x_i & \approx & \mathcal{M}_i\left(x_{i-1}\right) & \text{state evolution} \;\approx\; \text{by the model} \\
\mathcal{H}_i\left(x_i\right) & \approx & y_i & \text{value of observation operator} \;\approx\; \text{the data}
\end{array}
$$

- quantify "$\approx$" by covariances

$$x_0 \approx x_\mathrm{b} \;\Leftrightarrow\; \left\|x_0 - x_\mathrm{b}\right\|^2_{\mathbf{B}^{-1}} = \left(x_0 - x_\mathrm{b}\right)^\mathrm{T} \mathbf{B}^{-1} \left(x_0 - x_\mathrm{b}\right) \approx 0 \text{ etc.}$$

- $\Rightarrow$ **nonlinear least squares problem**

$$\left\|x_0 - x_\mathrm{b}\right\|^2_{\mathbf{B}^{-1}} + \sum_{i=1}^{k} \left\|x_i - \mathcal{M}_i\left(x_{i-1}\right)\right\|^2_{\mathbf{Q}_i^{-1}} + \sum_{i=1}^{k} \left\|y_i - \mathcal{H}_i\left(x_i\right)\right\|^2_{\mathbf{R}_i^{-1}} \to \min_{x_{0:k}}$$

Incremental 4DVAR

- Incremental approach (Courtier et al., 1994): linearization

$$\mathcal{M}_i\left(x_{i-1}+\delta x_{i-1}\right) \approx \mathcal{M}_i\left(x_{i-1}\right)+\mathcal{M}_i'\left(x_{i-1}\right)\delta x_{i-1}$$
$$\mathcal{H}_i\left(x_i+\delta x_i\right) \approx \mathcal{H}_i\left(x_i\right)+\mathcal{H}_i'\left(x_i\right)\delta x_i$$

- gives the **Gauss-Newton method** (Bell, 1994), iterations

$$x_{0:k} \leftarrow x_{0:k} + \delta x_{0:k}$$

with the **linear least squares** problem for the increments $\delta x_{0:k}$

$$\left\|x_0+\delta x_0-x_\mathrm{b}\right\|_{\mathbf{B}^{-1}}^2+\sum_{i=1}^{k}\left\|x_i+\delta x_i-\mathcal{M}_i\left(x_{i-1}\right)-\mathcal{M}_i'\left(x_{i-1}\right)\delta x_{i-1}\right\|_{\mathbf{Q}_i^{-1}}^2$$
$$+\sum_{i=1}^{k}\left\|y_i-\mathcal{H}_i\left(x_i\right)-\mathcal{H}_i'\left(x_i\right)\delta x_i\right\|_{\mathbf{R}_i^{-1}}^2 \rightarrow \min_{\delta x_{0:k}}$$

Linearized 4DVAR as Kalman smoother

The linear least squares problem for the increments $\delta x_{0:k}$

$$\|\delta x_0 - z_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \sum_{i=1}^{k} \|\delta x_i - \mathbf{M}_i \delta x_{i-1} - m_i\|^2_{\mathbf{Q}_i^{-1}} + \sum_{i=1}^{k} \|d_i - \mathbf{H}_i \delta x_i\|^2_{\mathbf{R}_i^{-1}} \to \min_{\delta x_{0:k}}$$

where

$$z_\mathrm{b} = x_\mathrm{b} - x_0, \, m_i = \mathcal{M}_i(x_{i-1}) - x_i, \, d_i = y_i - \mathcal{H}_i(x_i),$$
$$\mathbf{M}_i = \mathcal{M}_i'(x_{i-1}), \, \mathbf{H}_i = \mathcal{H}_i'(x_i)$$

is **equivalent to the Kalman smoother**

$$\begin{aligned} Z_0 &= z_\mathrm{b} & +V_0, \quad & V_0 \sim N(0, \mathbf{B}) \\ Z_i &= \mathbf{M}_i Z_{i-1} + m_i & +V_i, \quad & V_i \sim N(0, \mathbf{Q}_i) \\ d_i &= \mathbf{H}_i Z_i & +W_i, \quad & W_i \sim N(0, \mathbf{R}_i) \end{aligned}$$

$$\delta x_{0:k} = E(Z_{0:k} | d_{1:k}).$$

Ensemble Kalman smoother (EnKS)

Ensembles: $U^N = [u^1, \ldots, u^N]$. $V^N \sim N(m, \mathbf{A})$ is i.i.d. from $N(m, \mathbf{A})$, $Z^N_{i|k}$ is ensemble of states at time i, conditioned on all data up to time i.

Ensemble Kalman filter (EnKF): The analysis step makes linear combinations (transformation by a $\mathbf{T}^N_i$):

$$Z^N_{i|i} = Z^N_{i|i-1}\mathbf{T}^N_i, \quad \mathbf{T}^N_i \in \mathbb{R}^{N\times N}.$$

EnKS = EnKF + transform the history exactly the same way:

$$Z^N_{0:i|i} = Z^N_{0:i|i-1}\mathbf{T}^N_i.$$

Derivative-free implementation of the EnKS

Linearization by finite differences at the previous iteration, step size $\tau > 0$ towards ensemble member Z^n

Model operator:

$$\mathbf{M}_i Z_{i-1}^n + m_i = \mathcal{M}_i' \left(x_{i-1}\right) Z_{i-1}^n + \mathcal{M}_i \left(x_{i-1}\right) - x_i$$
$$\approx \frac{\mathcal{M}_i \left(x_{i-1} + \tau Z_{i-1}^n\right) - \mathcal{M}_i \left(x_{i-1}\right)}{\tau} + \mathcal{M}_i \left(x_{i-1}\right) - x_i$$

Observation function:

$$\mathbf{H}_i Z_i^n \approx \frac{\mathcal{H}_i \left(x_{i-1} + \tau Z_{i-1}^n\right) - \mathcal{H}_i \left(x_{i-1}\right)}{\tau}$$

Requires $N + 1$ evaluations of $\mathcal{M}_i$ and $\mathcal{H}_i$, at x_{i-1} and $x_{i-1} + \tau \delta x_{i-1}^n$.
Accurate in the limit $\tau \to 0$.
For $\tau = 1$, recover the nonlinear EnKS (no progress of 4DVAR, EnKS independent of the point of linearization)

Tikhonov regularization and Levenberg-Marquardt method

- Gauss-Newton may not converge, even locally. Add a penalty (Tikhonov regularization) to control the size of the increments δx_i :

$$\|\delta x_0 - z_\mathrm{b}\|^2_{\mathbf{B}^{-1}} + \sum_{i=1}^{k} \|\delta x_i - \mathbf{M}_i \delta x_{i-1} - m_i\|^2_{\mathbf{Q}_i^{-1}} + \sum_{i=1}^{k} \|d_i - \mathbf{H}_i \delta x_i\|^2_{\mathbf{R}_i^{-1}}$$
$$+ \gamma \sum_{i=0}^{k} \|\delta x_i\|^2_{\mathbf{S}_i^{-1}} \to \min_{\delta x_{0:k}}$$

- Becomes the Levenberg-Marquardt method, which is guaranteed to converge for large enough γ.
- Implement the regularization as independent observations $\delta x_i \approx 0$ with error covariance $\mathbf{S}_i$: simply run the analysis step the second time (Johns and Mandel, 2008). Here, this implementation is statistically exact because the distributions in the Kalman smoother are gaussian.

Convergence of the ensemble Kalman smoother

- Consider reference random vector $Z_{i|k}$, the state at time i conditioned exactly on all data up to time k.
- **Algorithm** (Kalman smoother on reference random vectors) Initialize $Z_{0|0} \sim N(z_{\mathrm{b}}, \mathbf{B})$.
  For $i = 1, \ldots, k$, advance in time

$$Z_{i|i-1} = \mathbf{M}_i Z^N_{i-1|i-1} + V_i, \quad V_i \sim N(m_i, \mathbf{Q}_i)$$

  followed by the analysis step with the exact covariance

$$Z_{0:|i} = Z_{0:i|i-1} - \mathbf{P}_{0:i,i}\mathbf{H}_i^{\mathrm{T}}(\mathbf{H}_i\mathbf{P}_i\mathbf{H}_i^{\mathrm{T}} + \mathbf{R}_i)^{-1}(\mathbf{H}_i Z^n_{i|i-1} - D_i),$$
$$\text{where } \mathbf{P}_{0:i,i} = \mathrm{Cov}(Z_{0:i|i-1}, Z_{i|i-1}), \quad D_i \sim N(d_i, \mathbf{R}_i)$$

- **Theorem** $Z_{i|i}$ has the filtering distribution $p(z_i | d_{1:i})$.
- **Theorem** $Z_{0:i|i}$ has the smoothing distribution $p(z_{0:i} | d_{1:i})$.
- **Theorem** (Convergence of the ensemble Kalman smoother)

$$\mathbf{P}^N \to \mathbf{P}_i, \quad Z^j_{i|i} \to Z_{i|i} \text{ as } N \to \infty, \text{for all } i, \text{ in all } L^p,\ 1 \le p < \infty.$$

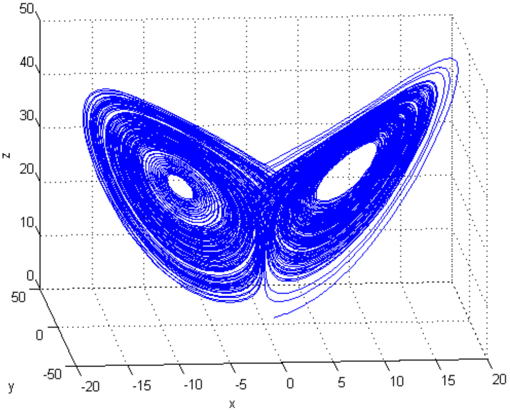
Convergence of the outer iterations

Theorem

Each Gauss-Newton-Levenberg-Marquardt iteration produced by EnKS-4DVAR converges to the corresponding iteration of 4DVAR (which solves exactly the incremental linearized least squares problem with the exact derivative), in all L^p, $1 \leq p < \infty$, as the ensemble size $N \to \infty$ and the finite difference step $\tau \to 0$.

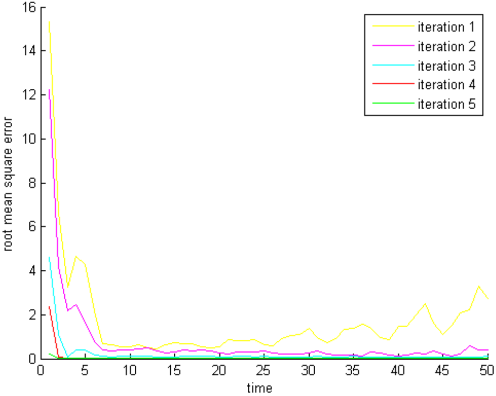
Computational results

Lorenz 63 model



$$\frac{dx}{dt} = -\sigma(x - y)$$
$$\frac{dy}{dt} = \rho x - y - xz$$
$$\frac{dz}{dt} = xy - \beta z$$

EnKS-4DVAR for Lorenz 63 model



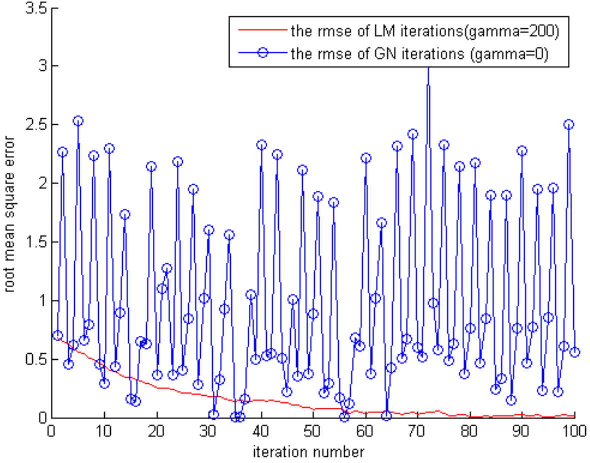
Root mean square error of EnKS-4DVAR iterations over 50 timesteps

Iteration	1	2	3	4	5	6
RMSE	20.16	15.37	3.73	2.53	0.09	0.09

An example where Gauss-Newton does not converge

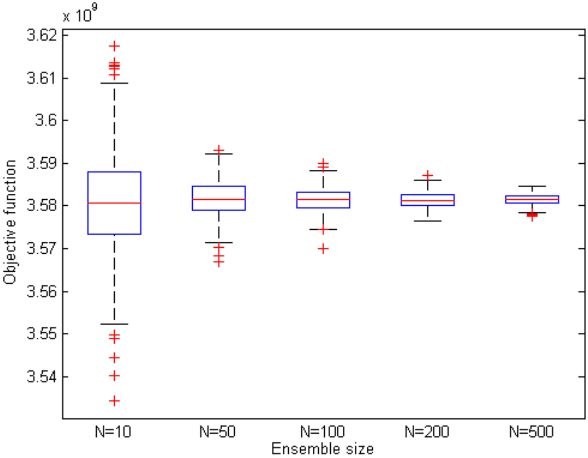
$$(x_0 - 2)^2 + (3 + x_1^2)^2 + 10^6(x_0 - x_1)^2 \to \min$$

4DVAR with $x_b = 2$, $\mathbf{B} = \mathbf{I}$, $M_1 = I$, $\mathcal{H}_1(x) = x^2$, $y_1 = 3$, $\mathbf{Q}_1 = 10^{-6}$



Convergence for increasing ensemble size

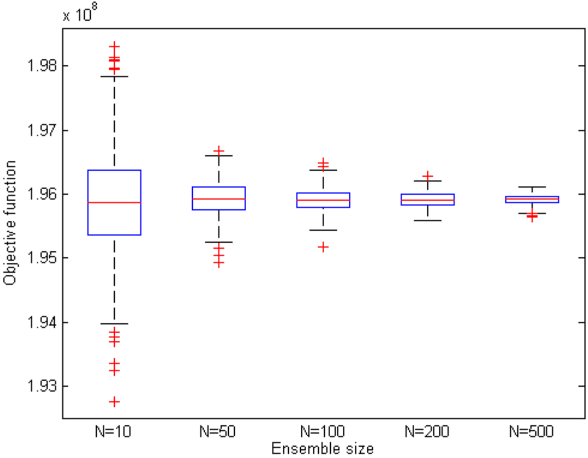
Lorenz 63



Gauss-Newton iteration 1

Convergence for increasing ensemble size

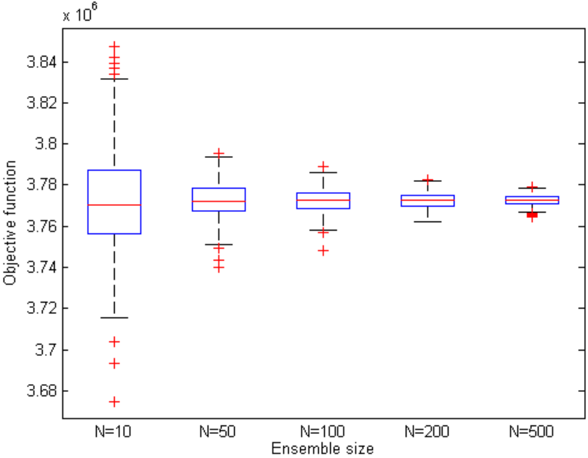
Lorenz 63



Gauss-Newton iteration 2

Convergence for increasing ensemble size

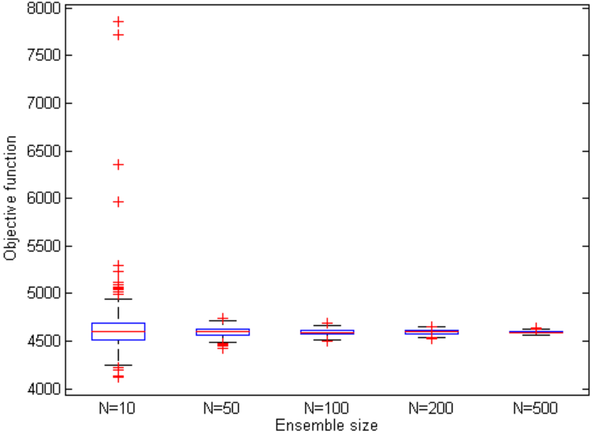
Lorenz 63



Gauss-Newton iteration 3

Convergence for increasing ensemble size

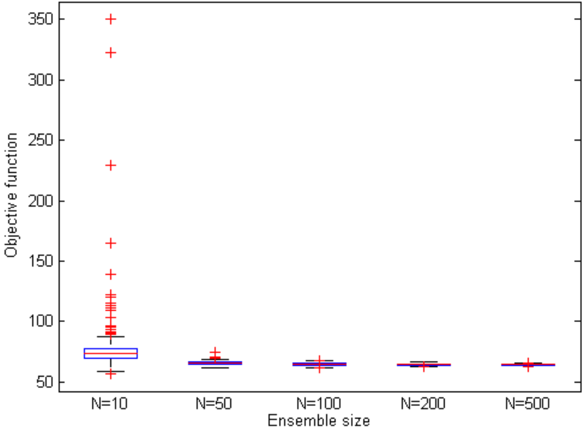
Lorenz 63



Gauss-Newton iteration 4

Convergence for increasing ensemble size

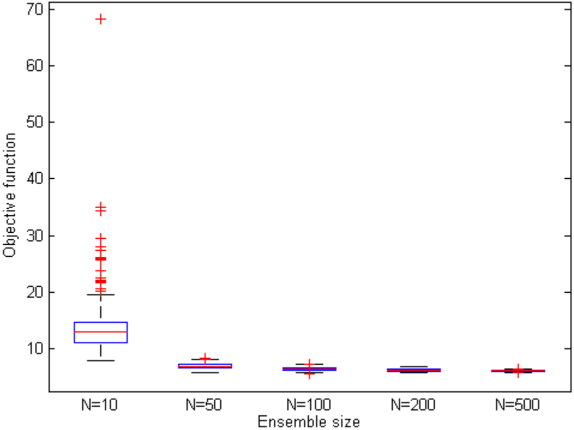
Lorenz 63



Gauss-Newton iteration 5

Convergence for increasing ensemble size

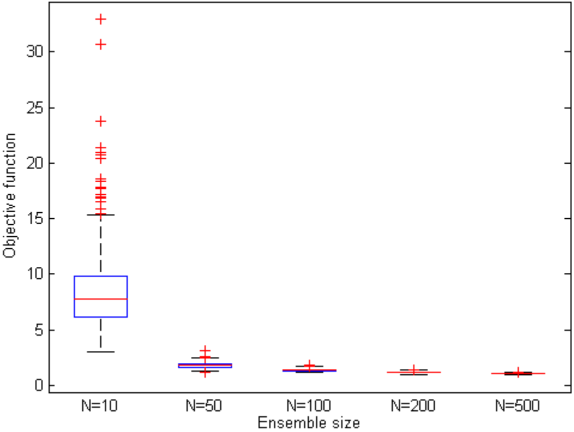
Lorenz 63



Gauss-Newton iteration 6

Convergence for increasing ensemble size

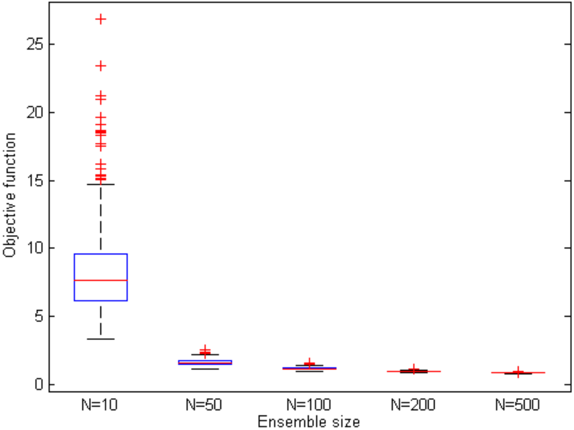
Lorenz 63



Gauss-Newton iteration 7

Convergence for increasing ensemble size

Lorenz 63



Gauss-Newton iteration 8

Conclusions from numerical experiments

- As small finite difference step τ as rounding errors allow is the best, particularly for strongly nonlinear problems.
- Scale all covariances by a small constant α, again subject to rounding errors.
- For small α and τ, the Monte Carlo simulation error $1/\sqrt{N}$ for ensemble size N remains.
- Gauss-Newton-Levenberg-Marquardt method converges up to the accuracy allowed by Monte Carlo simulation.

Related work

- The equivalence between weak constraint 4DVAR and Kalman smoothing is approximate for nonlinear problems, but still useful (Fisher et al., 2005).
- Estimated background covariance from anensemble for 4DVAR.
- Gradient methods in the span of the ensemble for one analysis cycle (i.e., 3DVAR) include Zupanski (2005), Sakov et al. (2012) (with square root EnKF as a linear solver in Newton method), and Bocquet and Sakov (2012), who added regularization and use LETKF-like approach to minimize the nonlinear cost function over linear combinations of the ensemble.
- Liu et al. (2008, 2009) use strong constraint 4DVAR, and minimize in the in the span of the ensemble
- Zhang et al. (2009) use EnKF to obtain the covariance for 4DVAR, and 4DVAR to feed the mean analysis into EnKF.

References

B. Bell. The iterated Kalman smoother as a Gauss-Newton method. *SIAM Journal on Optimization*, 4(3):626–636, 1994. doi: 10.1137/0804035.

M. Bocquet and P. Sakov. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlinear Processes in Geophysics*, 19(3): 383–399, 2012. doi: 10.5194/npg-19-383-2012.

P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4d-var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387, 1994. ISSN 1477-870X. doi: 10.1002/qj.49712051912.

M. Fisher, M. Leutbecher, and G. A. Kelly. On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131(613, Part c):3235–3246, OCT 2005. ISSN 0035-9009. doi: {10.1256/qj.04.142}.

C. J. Johns and J. Mandel. A two-stage ensemble Kalman filter for smooth data assimilation. *Environmental and Ecological Statistics*, 15:101–110, 2008. doi: 10.1007/s10651-007-0033-0.

C. Liu, Q. Xiao, and B. Wang. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Monthly Weather Review*, 136(9):3363–3373, 2008. doi: 10.1175/2008MWR2312.1.

C. Liu, Q. Xiao, and B. Wang. An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with Advanced Research WRF (ARW). *Monthly Weather Review*, 137(5):1687–1704, 2009. doi: 10.1175/2008MWR2699.1.

H. E. Rauch, F. Tung, and C. T. Striebel. Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450, 1965.

P. Sakov, D. S. Oliver, and L. Bertino. An iterative EnKF for strongly nonlinear systems. *Monthly Weather Review*, 140(6):1988–2004, 2012. doi: 10.1175/MWR-D-11-00176.1.

F. Zhang, M. Zhang, and J. Hansen. Coupling ensemble Kalman filter with four-dimensional variational data assimilation. *Advances in Atmospheric Sciences*, 26(1):1–8, 2009. doi: 10.1007/s00376-009-0001-8.

M. Zupanski. Maximum likelihood ensemble filter: Theoretical aspects. *Monthly Weather Review*, 133(6):1710–1726, 2013/01/01 2005. doi: 10.1175/MWR2946.1.