

The combined parameter and state estimation problem

Geir Evensen and Laurent Bertino

Hydro Research Centre, Bergen, Norway,
Nansen Environmental and Remote Sensing Center, Bergen, Norway

- Combined parameter and state estimation

“Find the pdf of the parameters and the model solution conditional on a set of measurements.”

- Bayesian formulation.

- Ensemble solutions.

Bayesian formulation

Model equations and measurements:

$$\frac{\partial \psi}{\partial t} = g(\psi, \alpha) + \mathbf{q},$$

$$\psi|_{t_0} = \Psi_0 + \mathbf{a},$$

$$\alpha = \alpha_0 + \alpha',$$

$$\mathcal{M}(\psi, \alpha) = \mathbf{d} + \epsilon.$$

Bayes theorem becomes:

$$f(\psi, \alpha, \psi_0 | \mathbf{d}) \propto f(\psi | \alpha, \psi_0) f(\psi_0) f(\alpha) f(\mathbf{d} | \psi, \alpha).$$

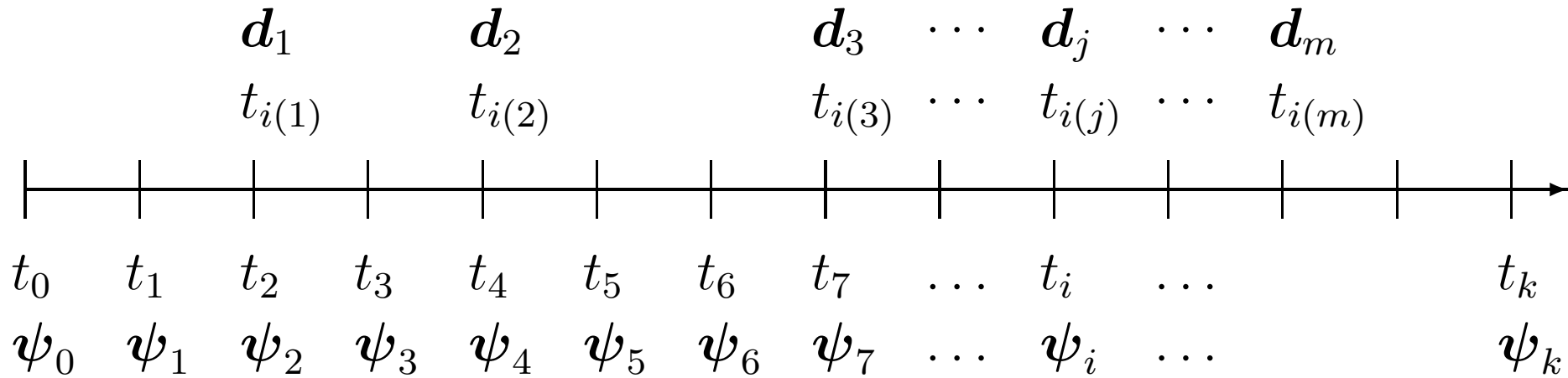
Gaussian priors

$$f(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0 | \mathbf{d}) \propto \exp \left\{ -\frac{1}{2} \mathcal{J}[\boldsymbol{\psi}, \boldsymbol{\alpha}] \right\},$$

$$\begin{aligned} \mathcal{J}[\boldsymbol{\psi}, \boldsymbol{\alpha}] &= \left(\frac{\partial \boldsymbol{\psi}}{\partial t} - \mathbf{g}(\boldsymbol{\psi}, \boldsymbol{\alpha}) \right)^{\text{T}} \bullet \mathbf{W}_{qq} \bullet \left(\frac{\partial \boldsymbol{\psi}}{\partial t} - \mathbf{g}(\boldsymbol{\psi}, \boldsymbol{\alpha}) \right) \\ &+ (\boldsymbol{\psi}_0 - \boldsymbol{\Psi}_0)^{\text{T}} \circ \mathbf{W}_{aa} \circ (\boldsymbol{\psi}_0 - \boldsymbol{\Psi}_0) \\ &+ (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^{\text{T}} \circ \mathbf{W}_{\alpha\alpha} \circ (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) \\ &+ (\mathbf{d} - \mathcal{M}[\boldsymbol{\psi}])^{\text{T}} \mathbf{W}_{\epsilon\epsilon} (\mathbf{d} - \mathcal{M}[\boldsymbol{\psi}]). \end{aligned}$$

- MLH solution, hard to solve and to compute error estimates.

Discretisation in time



Assume

- Model is first order Markov process.

$$f(\psi_1, \dots, \psi_k, \alpha, \psi_0) \propto$$

$$f(\alpha) f(\psi_0) \prod_{i=1}^k f(\psi_i | \psi_{i-1}, \alpha).$$

- Measurement errors are independent in time

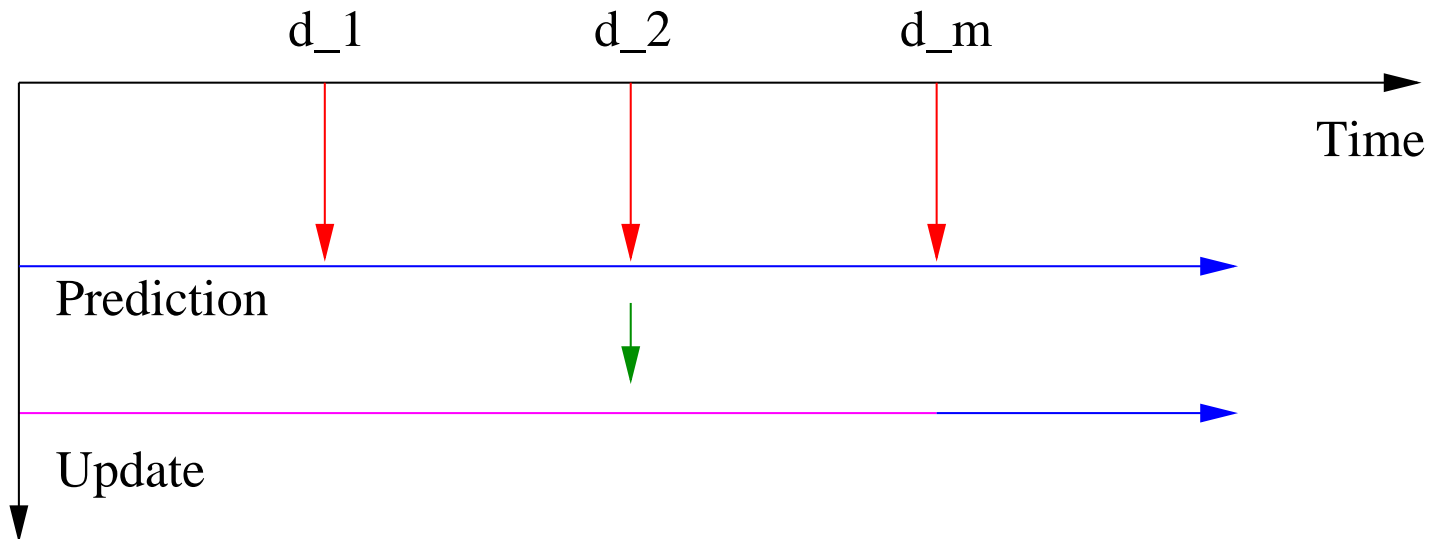
$$f(\mathbf{d} | \psi, \alpha) = \prod_{j=1}^m f(\mathbf{d}_j | \psi_{i(j)}, \alpha).$$

Bayes for discrete state

- Bayes theorem then becomes

$$f(\psi_1, \dots, \psi_k, \alpha, \psi_0 | \mathbf{d}) \propto f(\alpha) f(\psi_0)$$

$$\prod_{i=1}^k f(\psi_i | \psi_{i-1}, \alpha) \prod_{j=1}^m f(\mathbf{d}_j | \psi_{i(j)}, \alpha),$$



Rewrite as:

$$f(\psi_1, \dots, \psi_k, \alpha, \psi_0 | \mathbf{d}) \propto f(\alpha) f(\psi_0)$$

$$\prod_{i=1}^{i(1)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_1 | \psi_{i(1)}, \alpha)$$

$$\prod_{i=i(1)+1}^{i(2)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_2 | \psi_{i(2)}, \alpha) \cdots$$

$$\prod_{i=i(m-1)+1}^{i(m)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_m | \psi_{i(m)}, \alpha)$$

$$\prod_{i=i(m)+1}^k f(\psi_i | \psi_{i-1}, \alpha)$$

Sequential processing of measurements

- First update

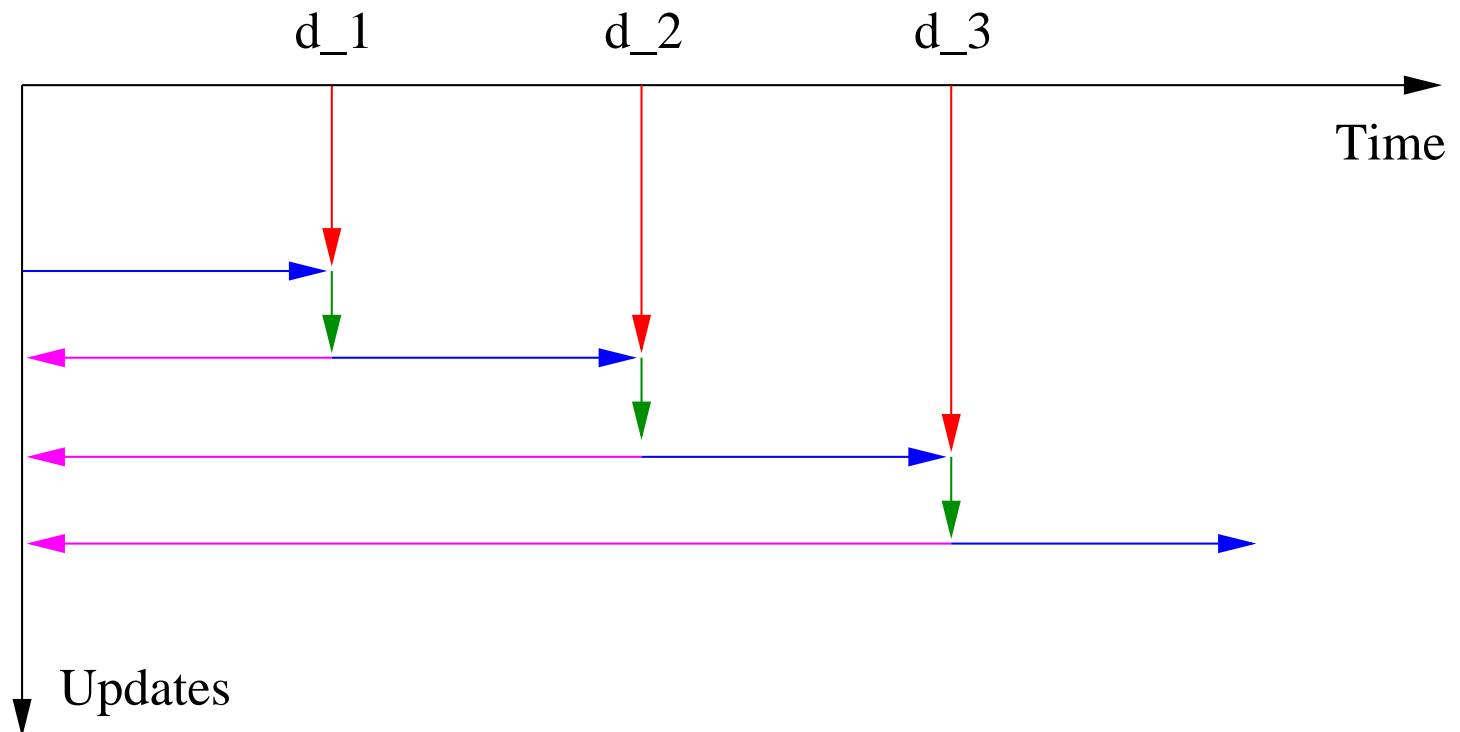
$$f(\psi_1, \dots, \psi_{i(1)}, \alpha, \psi_0 | \mathbf{d}_1) \propto$$
$$f(\alpha) f(\psi_0)$$
$$\prod_{i=1}^{i(1)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_1 | \psi_{i(1)}, \alpha)$$

- Second update

$$f(\psi_1, \dots, \psi_{i(2)}, \alpha, \psi_0 | \mathbf{d}_1, \mathbf{d}_2) \propto$$
$$f(\psi_1, \dots, \psi_{i(1)}, \alpha, \psi_0 | \mathbf{d}_1)$$
$$\prod_{i=i(1)+1}^{i(2)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_2 | \psi_{i(2)}, \alpha)$$

Summary

- Independent measurements processed sequentially in time.
- Sequence of inverse problems.
- Solution of one sub-problem is prior for next.
- Hard to solve using traditional variational methods.
- Well suited for ensemble methods.



Variance minimizing analysis

Given

- First guess, ψ^f , and prediction error covariance, $C_{\psi\psi}^f$.
- Measurements, d , and error covariance, $C_{\epsilon\epsilon}$.

The variance minimizing analysis is minimum of

$$\begin{aligned} \mathcal{J}[\psi^a] = & \left(\psi^a - \psi^f \right)^T \bullet C_{\psi\psi}^{-1} \bullet \left(\psi^a - \psi^f \right) \\ & + \left(d - \mathcal{M}\psi^a \right)^T C_{\epsilon\epsilon}^{-1} \left(d - \mathcal{M}\psi^a \right). \end{aligned}$$

Analysis equations

Update of estimate

$$\psi^a = \psi^f + C_{\psi\psi}^f \mathcal{M}^T \left(\mathcal{M} C_{\psi\psi}^f \mathcal{M}^T + C_{\epsilon\epsilon} \right)^{-1} \left(d - \mathcal{M} \psi^f \right)$$

Error covariance for update

$$C_{\psi\psi}^a = C_{\psi\psi}^f - C_{\psi\psi}^f \mathcal{M}^T \left(\mathcal{M} C_{\psi\psi}^f \mathcal{M}^T + C_{\epsilon\epsilon} \right)^{-1} \mathcal{M} C_{\psi\psi}^f$$

Ensemble methods

ES: Ensemble Smoother

EnKS: Ensemble Kalman Smoother

EnKF: Ensemble Kalman Filter

- Ensemble representation for pdfs.
- Ensemble prediction for time evolution of pdfs.
- Linear ensemble analysis scheme:
 - “Variance minimizing”.
 - Assumes Gaussian pdf for model prediction.
 - No resampling!

The error covariance matrix

Define ensemble covariances around the ensemble mean

$$\mathbf{C}_{\psi\psi} = \overline{(\psi - \bar{\psi})(\psi - \bar{\psi})^T}$$

- The ensemble mean, $\bar{\psi}$, is the best-guess.
- The error variance is defined by the ensemble spread.
- The smoothness of members defines the covariance.



Representation by a finite ensemble of model states.

Ensemble representation

- Define ensemble matrix

$$\mathbf{A}(\mathbf{x}, t_i) = \begin{pmatrix} \psi^1(\mathbf{x}, t_i) & \psi^2(\mathbf{x}, t_i) & \dots & \psi^N(\mathbf{x}, t_i) \\ \alpha^1(\mathbf{x}, t_i) & \alpha^2(\mathbf{x}, t_i) & \dots & \alpha^N(\mathbf{x}, t_i) \end{pmatrix}.$$

- Parameters augmented to model state.
- The ensemble covariance becomes

$$\mathbf{C}_{\psi\psi}^e(\mathbf{x}_1, \mathbf{x}_2, t_i) = \frac{\mathbf{A}'(\mathbf{x}_1, t_i) \mathbf{A}'^T(\mathbf{x}_2, t_i)}{N - 1}.$$

Dynamical evolution of error statistics

- Members evolve according to stochastic model dynamics

$$d\psi = \mathbf{g}(\psi, \alpha)dt + \mathbf{h}(\psi)dq$$
$$d\alpha = 0.$$

- The pdf evolves according to Kolmogorov's equation

$$\frac{\partial f}{\partial t} + \sum_i \frac{\partial (g_i f)}{\partial \psi_i} = \frac{1}{2} \sum_{i,j} \frac{\partial^2 f (\mathbf{h} \mathbf{Q} \mathbf{h}^T)_{ij}}{\partial \psi_i \partial \psi_j}.$$

- Fundamental equation for evolution of error statistics.
- May be solved using Monte Carlo methods.

Ensemble analysis equations

- Define randomized measurements $D = d + E$.
- Variance minimizing ensemble update becomes

$$A^a = A$$

$$+ C_{\psi\psi}^e \mathcal{M}^T \left(\mathcal{M} C_{\psi\psi}^e \mathcal{M}^T + C_{\epsilon\epsilon} \right)^{-1} (D - \mathcal{M}A)$$

$$= A$$

$$+ A' (\mathcal{M}A')^T \left(\mathcal{M}A' (\mathcal{M}A')^T + C_{\epsilon\epsilon} \right)^{-1} (D - \mathcal{M}A)$$

⋮

$$= AX$$

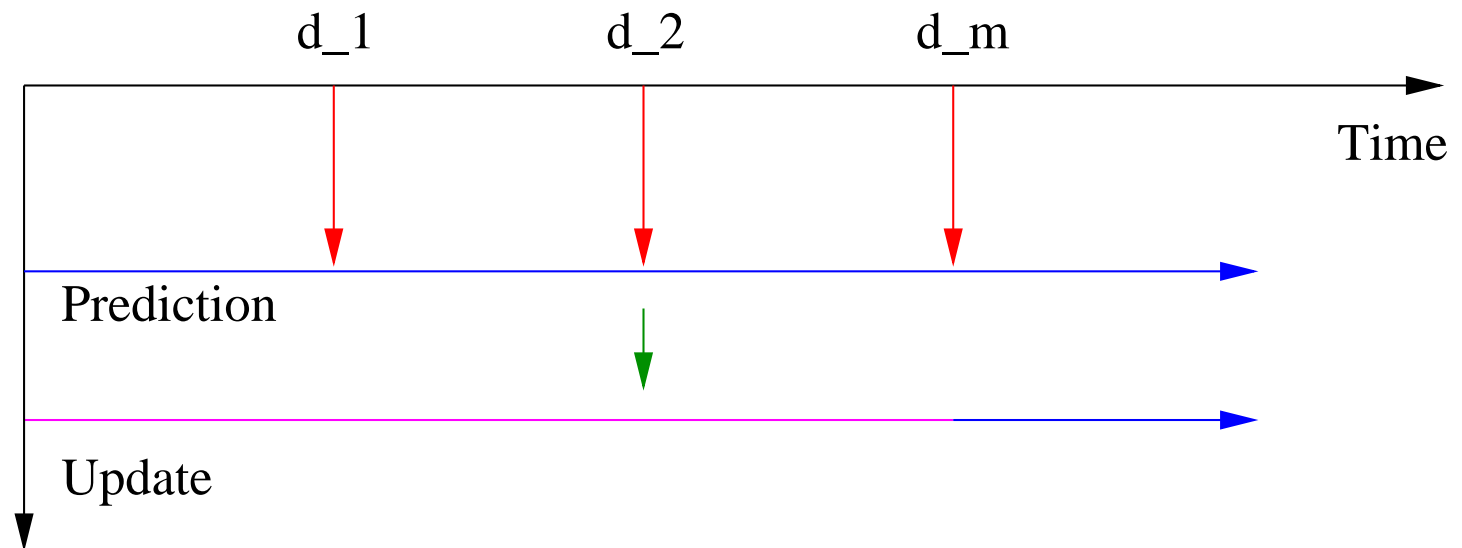
⇒ Updated ensemble with correct mean and covariance!

ES: The Ensemble Smoother

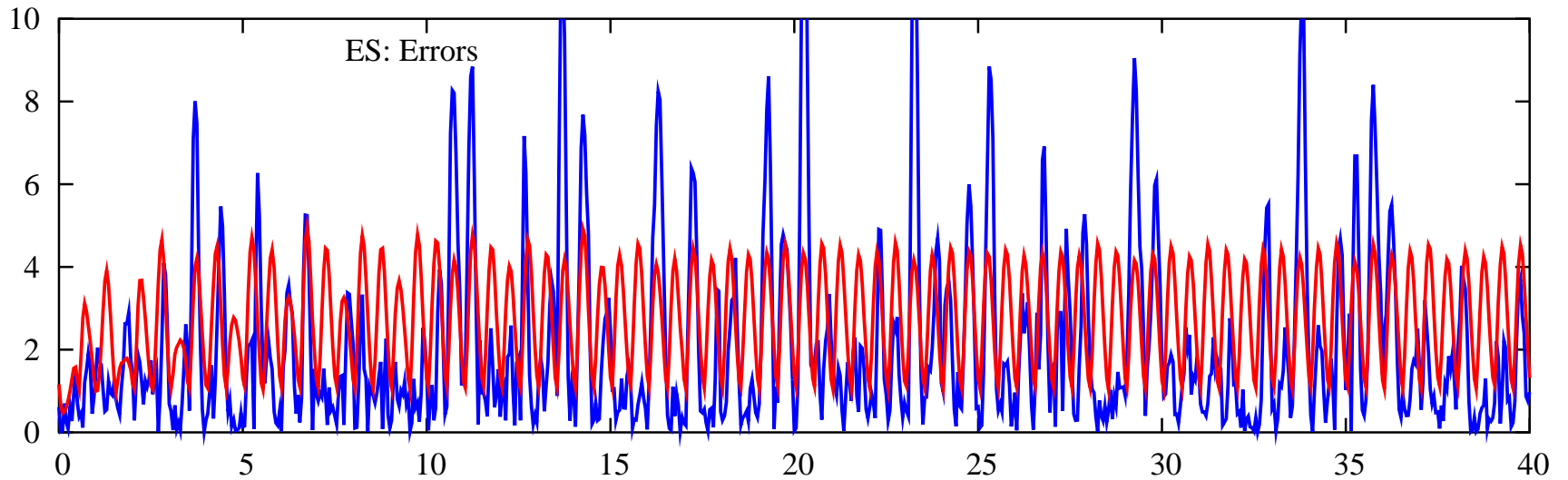
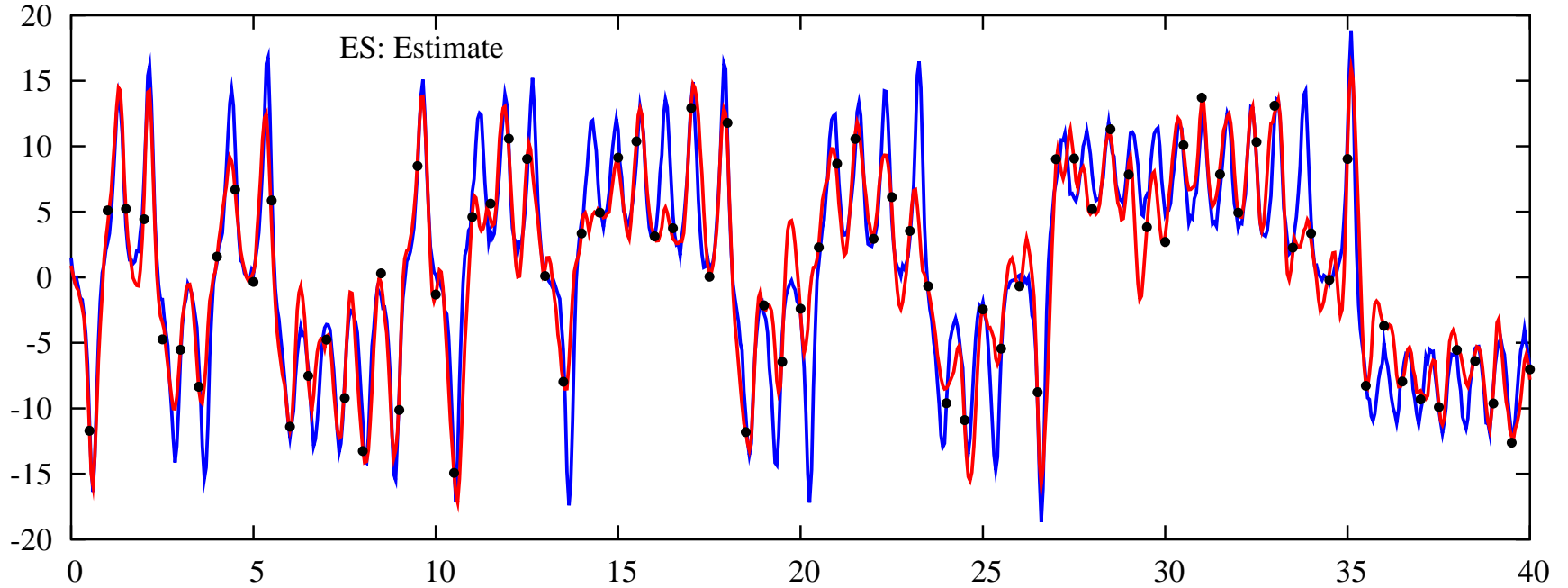
Computes the Bayesian update using linear ensemble equation

$$f(\psi_1, \dots, \psi_k, \alpha, \psi_0 | \mathbf{d}) =$$

$$f(\psi_1, \dots, \psi_k, \alpha, \psi_0) \prod_{j=1}^m f(\mathbf{d}_j | \psi_{i(j)}, \alpha),$$



ES: Example with Lorenz equations



ES: summary

Gauss–Markov interpolation in space and time.

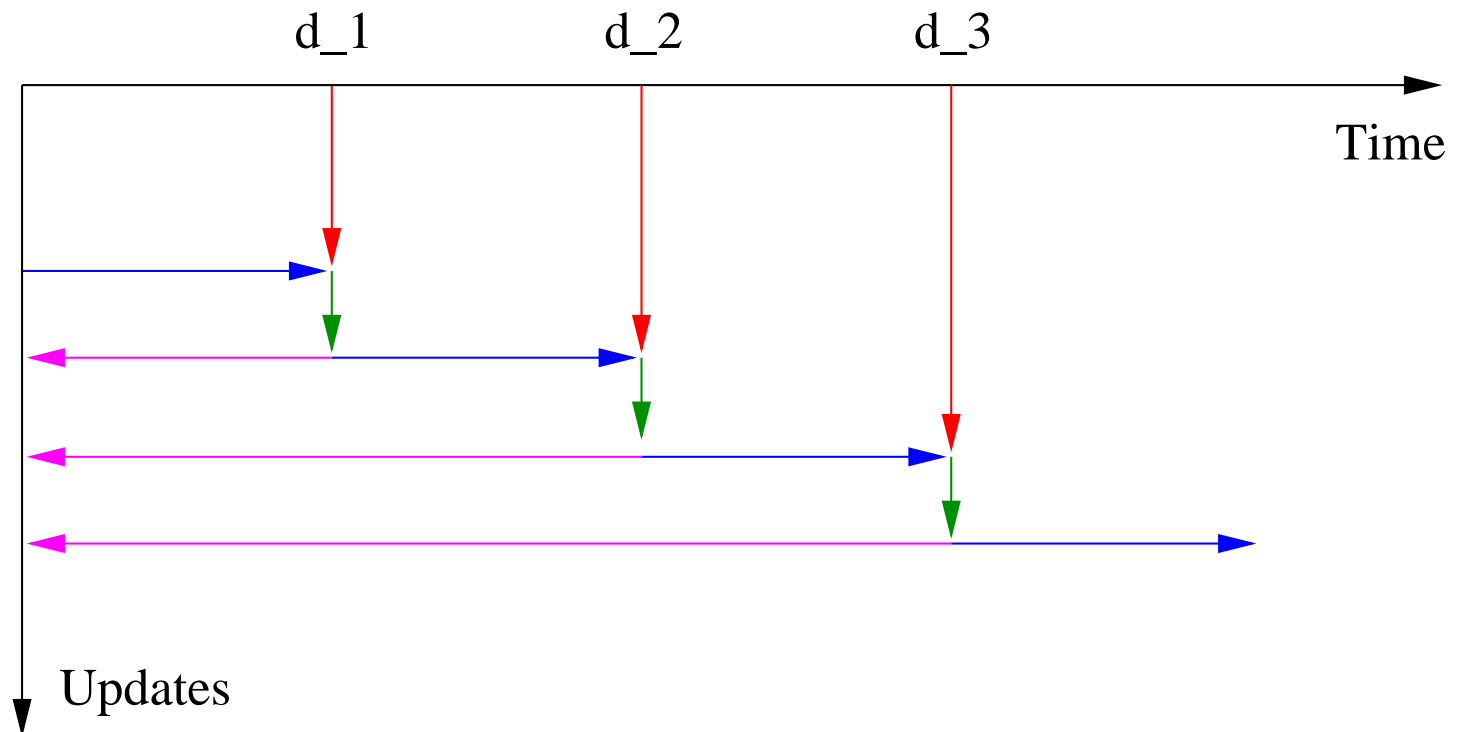
- Creates an ensemble for the model evolution.
- Assumes Gaussian pdf for model evolution.
- Computes variance minimizing ensemble analysis.
- Exact solution for linear problems.

EnKS: The ensemble Kalman smoother

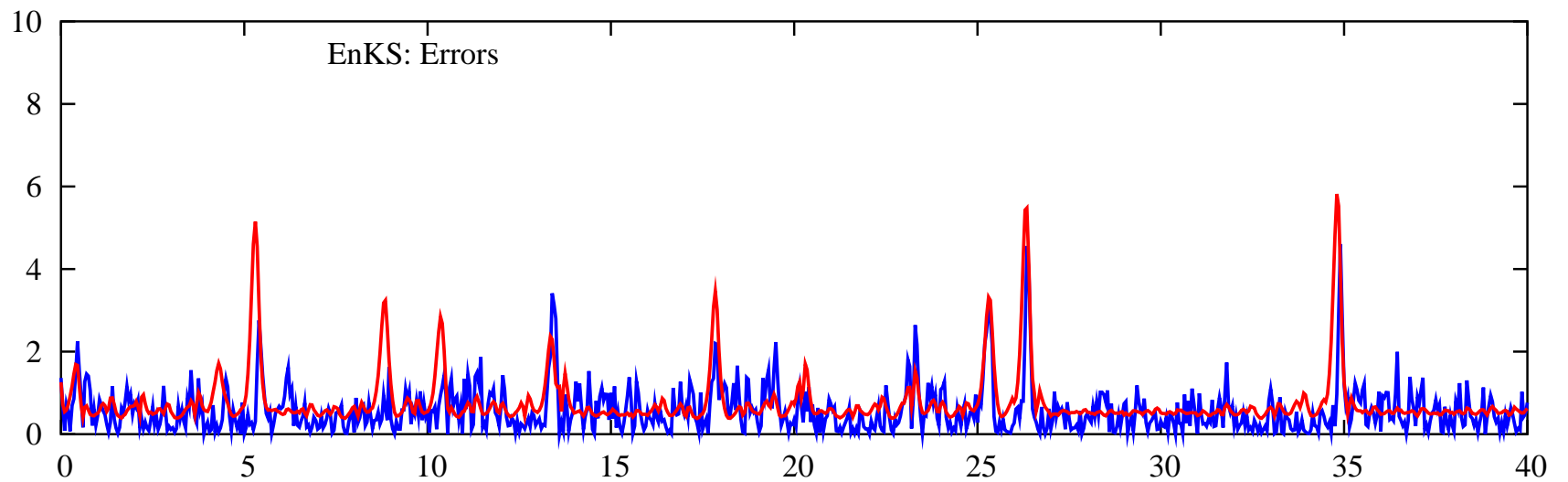
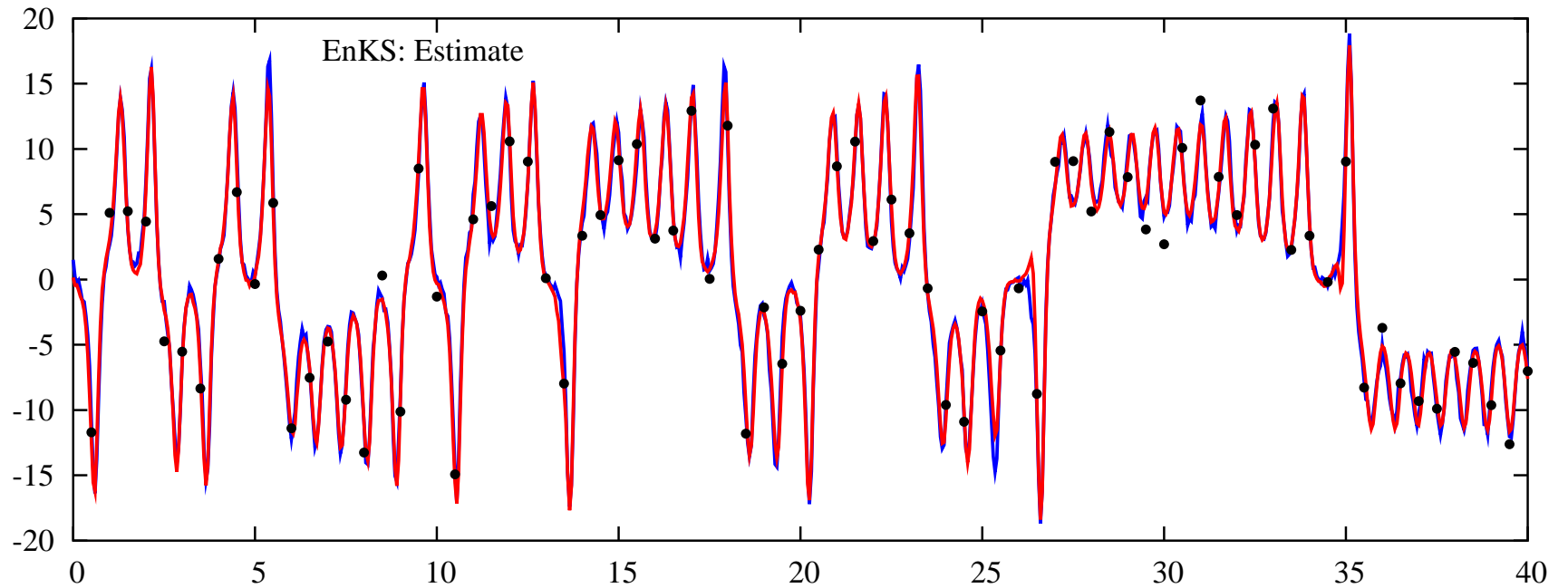
- Bayes with sequential processing of data: First update is

$$f(\psi_1, \dots, \psi_{i(1)}, \alpha, \psi_0 | \mathbf{d}_1) \propto$$

$$f(\alpha) f(\psi_0) \prod_{i=1}^{i(1)} f(\psi_i | \psi_{i-1}, \alpha) f(\mathbf{d}_1 | \psi_{i(1)}, \alpha),$$



EnKS solution

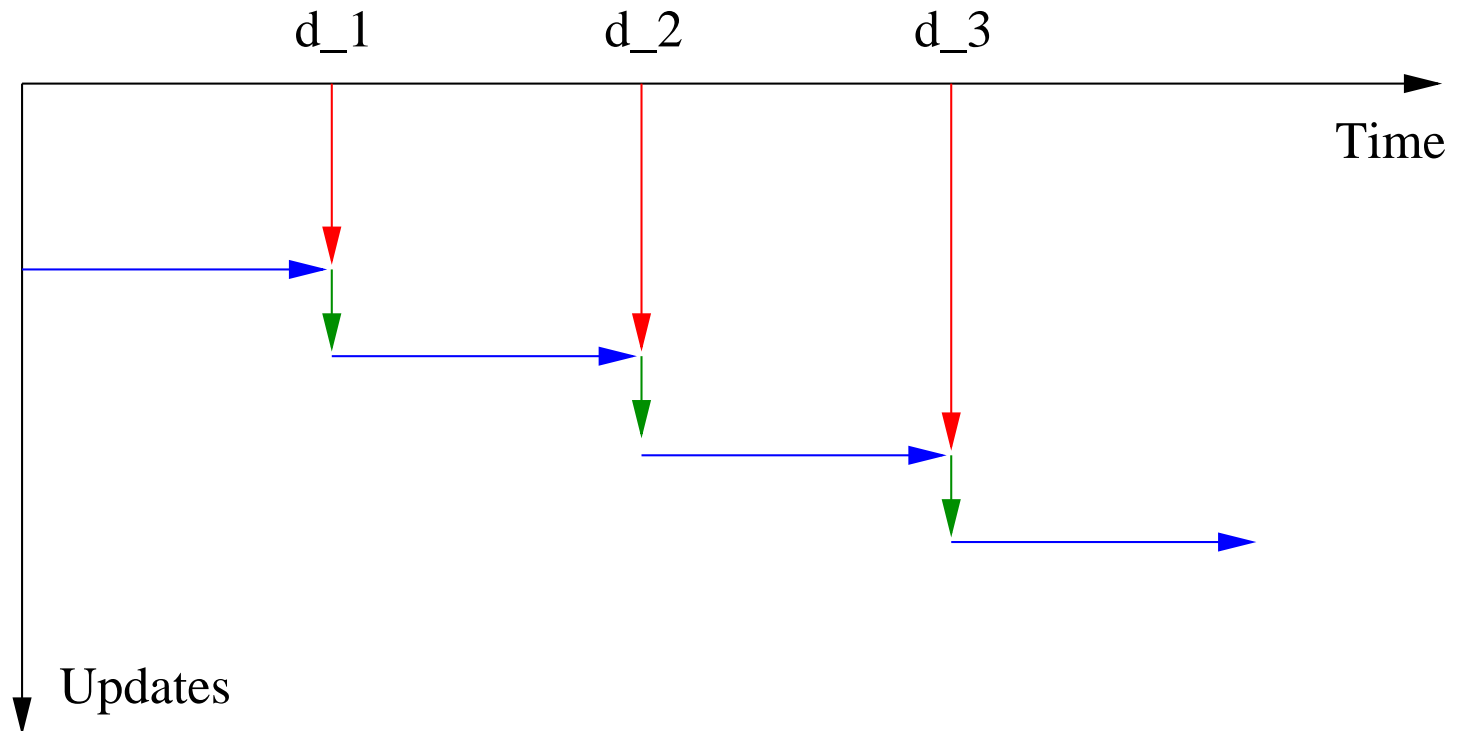


EnKS summary

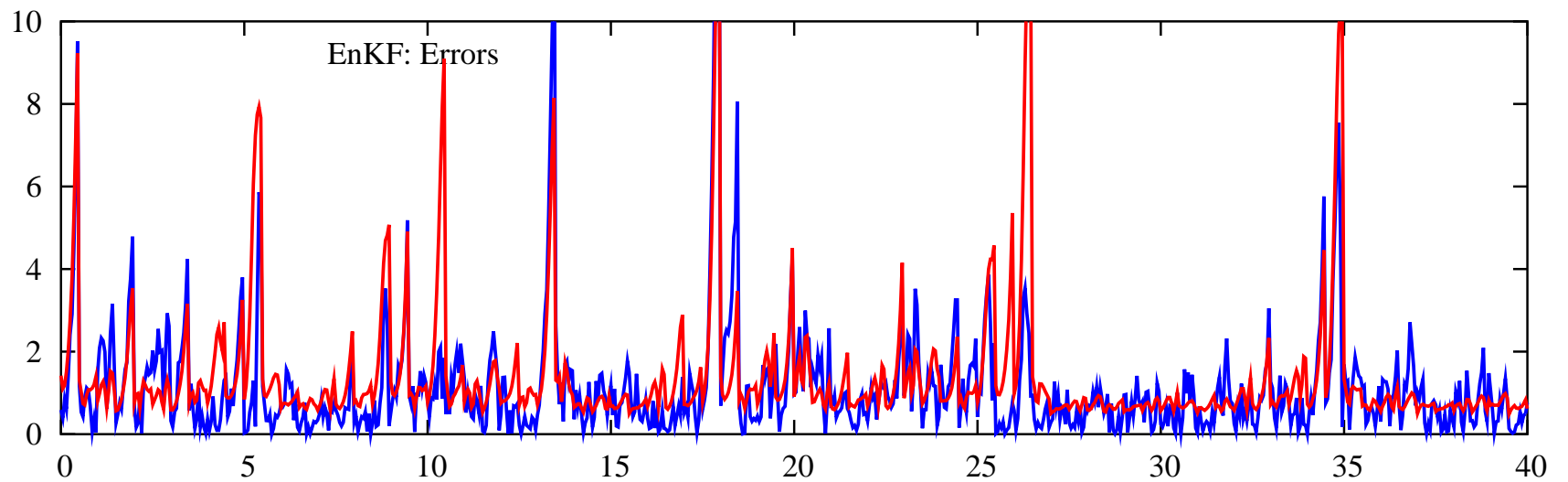
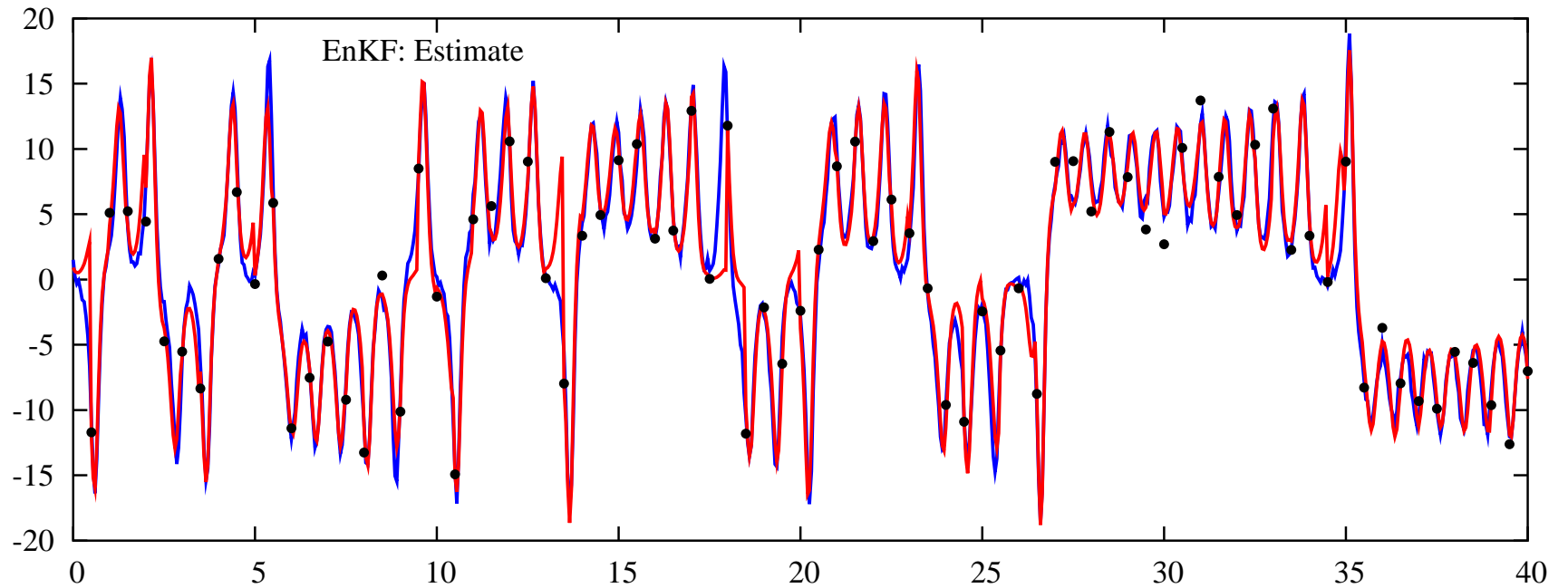
- ES and EnKS give identical results for linear models.
- EnKS is superior to the ES with nonlinear models.
 - Sequential processing of measurements introduces “Gaussianity”.
 - Ensemble is kept close to the true state.

EnKF: Ensemble Kalman Filter

- Special case of EnKS.
- Only update state at data times
- EnKF forms a prior for the EnKS.
- A prediction from EnKF and EnKS will be identical.



EnKF solution



Example

- Scalar model

$$\frac{\partial \psi}{\partial t} = 1 - \alpha + q,$$

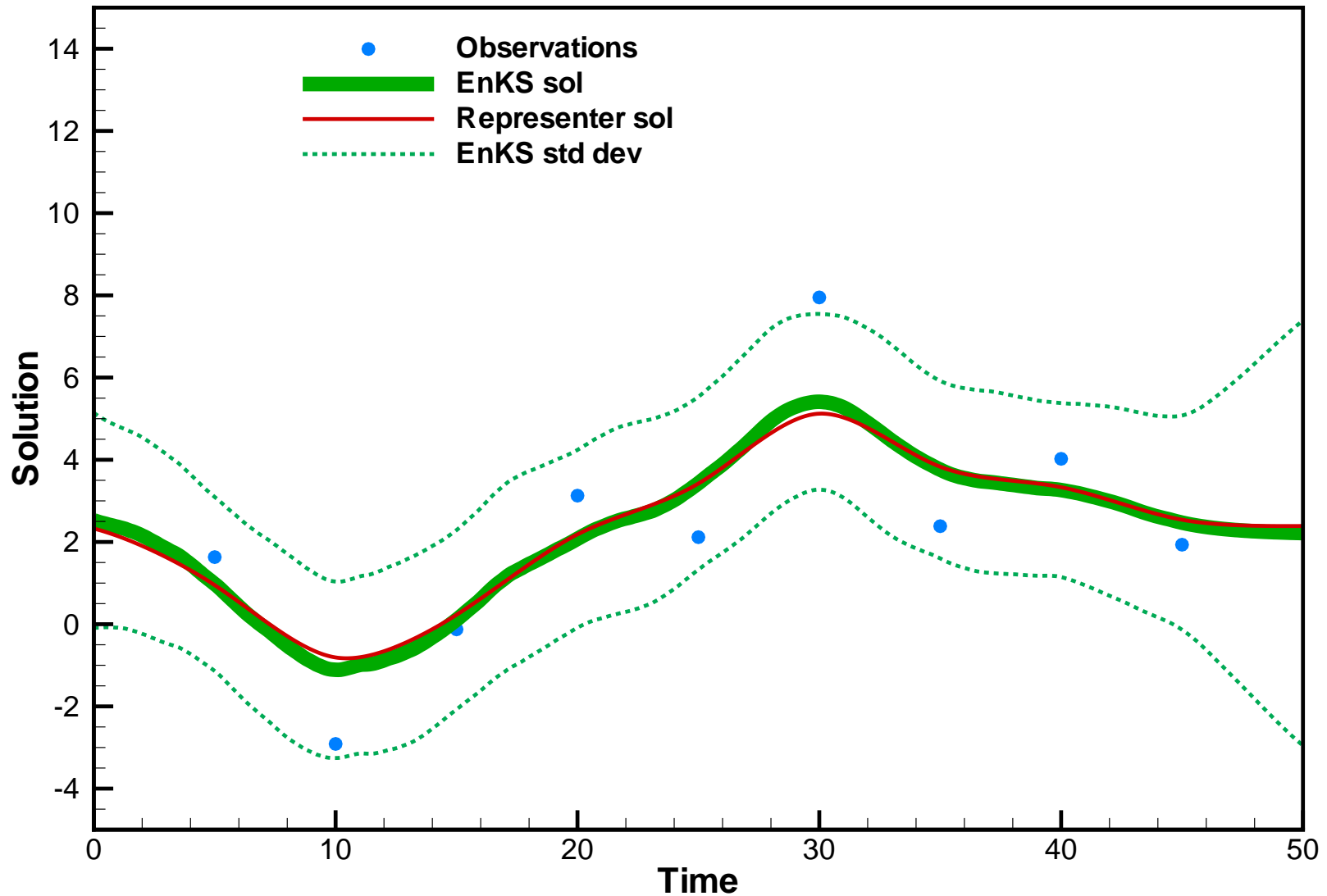
$$\psi(t = 0) = 3 + a,$$

$$\alpha = 0 + \alpha',$$

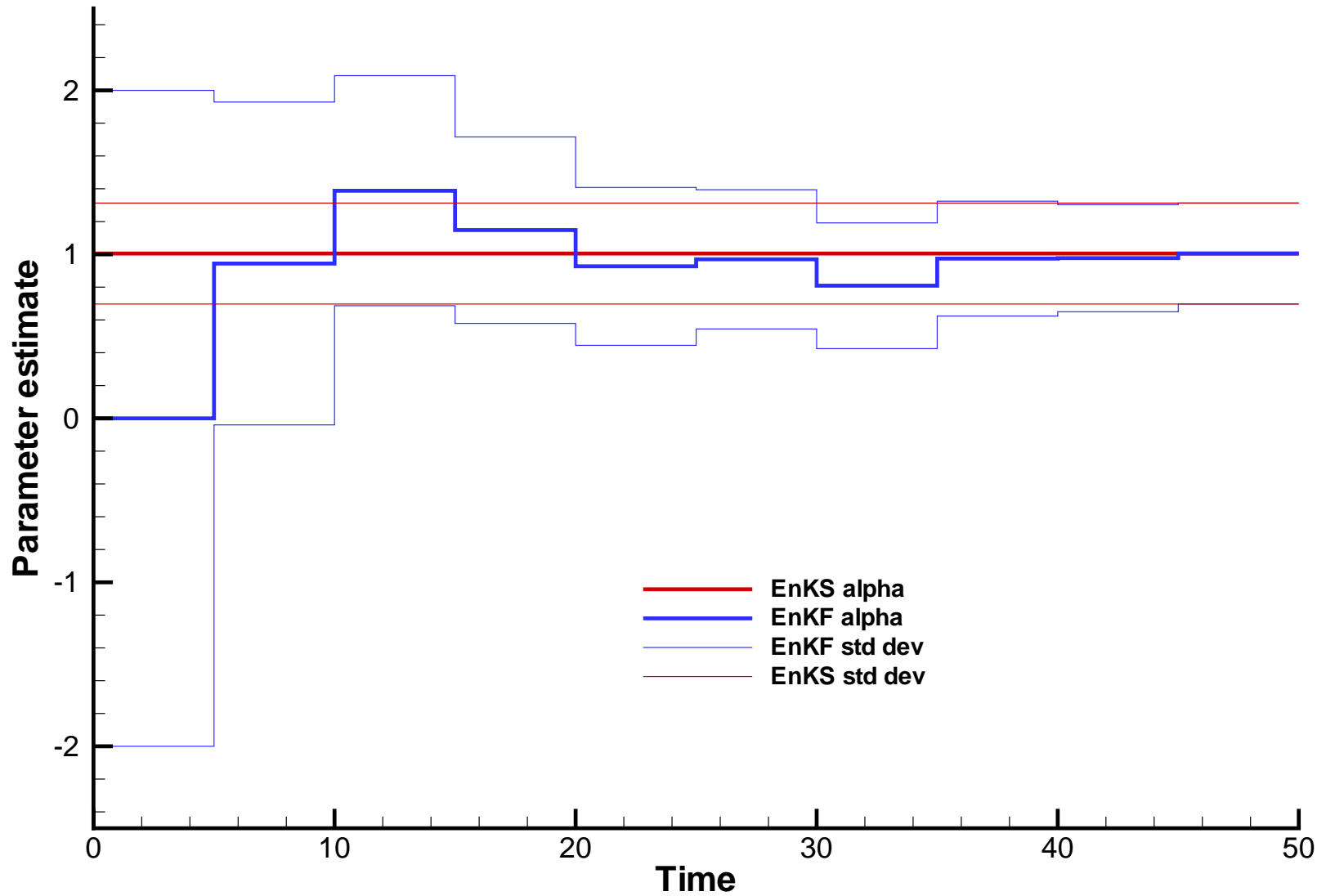
$$\mathcal{M}(\psi) = d + \epsilon.$$

- True parameter value is $\alpha = 1$.
- Model operator is linear and independent of ψ .
- Solved using EnKF, EnKS and Representer methods.
- Exponential time correlation for model errors.

State and parameter estimation



Estimate of parameter



Summary

www.nersc.no/~geir/EnKF

- Joint state and parameter estimation problem!
- Bayesian formulation:
 - Variational methods (multiple minima; MLH).
 - Hard to compute error statistics!
 - Do not allow for sequential processing of measurements!
 - Ensemble methods (Gaussianity assumption; mean).
 - Provide estimate with error statistics.
 - Allow for sequential processing of measurements.
 - Introduces Gaussianity!
 - Advantage with nonlinear dynamics.