The parameter estimation problem revisited

Geir Evensen

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

based on
"The combined parameter and state estimation problem"
submitted to Computational Geosciences

2005 SIAM Conference on
Mathematical and Computational Issues in the Geosciences

Outline

- Parameter and state estimation problem.
- Bayesian formulation.
- Variational inverse problem.
- Ensemble methods.
- Examples.

Model equations

$$
\begin{aligned}
\frac{\partial \boldsymbol{\psi}}{\partial t} &= \boldsymbol{g}(\boldsymbol{\psi}, \boldsymbol{\alpha}), \\
\boldsymbol{\psi}|_{t_0} &= \boldsymbol{\Psi}_0, \\
\boldsymbol{\psi}|_{\partial \mathcal{D}} &= \boldsymbol{\psi}_b, \\
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0.
\end{aligned}
$$

$\Longrightarrow$ Well-posed problem with unique deterministic solution.

- $\boldsymbol{\psi}(\boldsymbol{x}, t)$ is model state vector.
- $\boldsymbol{\alpha}(\boldsymbol{x})$ is vector of model parameters.

Model equations and measurements

$$
\begin{aligned}
\frac{\partial \boldsymbol{\psi}}{\partial t} &= \boldsymbol{g}(\boldsymbol{\psi}, \boldsymbol{\alpha}), \\
\boldsymbol{\psi}|_{t_0} &= \boldsymbol{\Psi}_0, \\
\boldsymbol{\psi}|_{\partial \mathcal{D}} &= \boldsymbol{\psi}_b, \\
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0. \\
\boldsymbol{d} &= \mathcal{M}(\boldsymbol{\psi}, \boldsymbol{\alpha}).
\end{aligned}
$$

$\Longrightarrow$ Over-determined problem with no solution.

Example of direct measurement of $\psi(t)$:

$$\psi(t_i) = \mathcal{M}_i(\psi) = \int \psi(t)\delta(t - t_i)dt.$$

Allow for errors

Assume stochastic errors $\boldsymbol{q}(\boldsymbol{x},t)$, $\boldsymbol{a}(\boldsymbol{x})$, $\boldsymbol{b}(\boldsymbol{s},t)$, $\boldsymbol{\alpha}(\boldsymbol{x})$, and $\boldsymbol{\epsilon}$:

$$\begin{aligned}
\frac{\partial \boldsymbol{\psi}}{\partial t} &= \boldsymbol{g}(\boldsymbol{\psi},\boldsymbol{\alpha}) + \boldsymbol{q}, \\
\boldsymbol{\psi}|_{t_0} &= \boldsymbol{\Psi}_0 + \boldsymbol{a}, \\
\boldsymbol{\psi}|_{\partial\mathcal{D}} &= \boldsymbol{\psi}_b + \boldsymbol{b}, \\
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}', \\
\boldsymbol{d} &= \mathcal{M}(\boldsymbol{\psi}^{\mathrm{t}},\boldsymbol{\alpha}^{\mathrm{t}}) + \boldsymbol{\epsilon}.
\end{aligned}$$

$\Longrightarrow$ Infinitively many solutions.

- Must specify statistics for error terms!
- Least squares problem.
- Find estimate for $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$ which “minimizes” errors.

State estimation

"Find an estimate of the state given a dynamical model and measurements."

- Standard data assimilation problem.
- Parameters fixed.
- Minimize errors in model and measurements.
- Solved by e.g. adjoint, representer or Kalman filter methods.

Parameter estimation

"Find an estimate of a set of parameters, resulting in a model solution which is close to the measurements."

- Standard parameter estimation problem.
- Minimization of a cost function.
- Solved using e.g. gradient methods.
- Solution is a single model realization.
- Model assumed perfect given the parameters.

A more general problem

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

- Combined parameter and state estimation problem.
- Bayesian formulation.
- Optimal solution is the conditional pdf.
  - Different estimators for parameter and state.
  - Provides uncertainty estimates.

Part I

Bayesian formulation

Bayesian formulation

- Bayes theorem

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

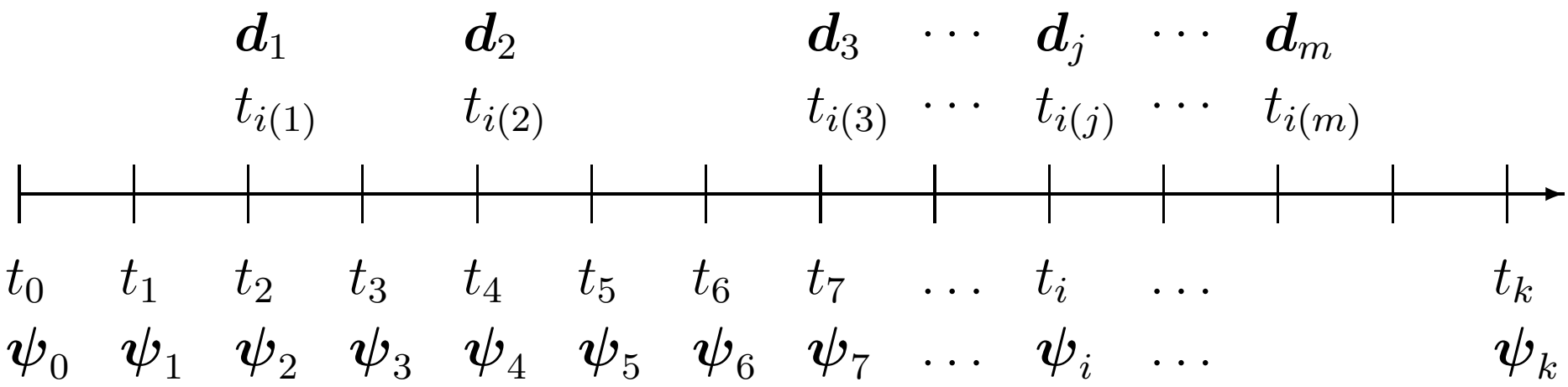
- The pdf for the model prediction is

$$f(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) = f(\boldsymbol{\psi}|\boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) f(\boldsymbol{\alpha}).$$

- Bayes theorem then becomes

$$\boxed{f(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b|\boldsymbol{d}) \propto f(\boldsymbol{\psi}|\boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) f(\boldsymbol{\alpha}) f(\boldsymbol{d}|\boldsymbol{\psi}, \boldsymbol{\alpha}).}$$

Discretisation in time



- Time nodes t_i for $i = 0, \ldots, k$.
- Define $\boldsymbol{\psi}_i = \boldsymbol{\psi}(t_i)$.
- Then $\boldsymbol{\psi}$ is represented by $(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k)$
- Measurements $\boldsymbol{d}_j$ available at times, $t_{i(j)}$, $j = 1, \ldots, m$.

Assume

- Model is first order Markov process.

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) \propto$$

$$f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_b) f(\boldsymbol{\psi}_0) \prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}, \boldsymbol{\psi}_b).$$

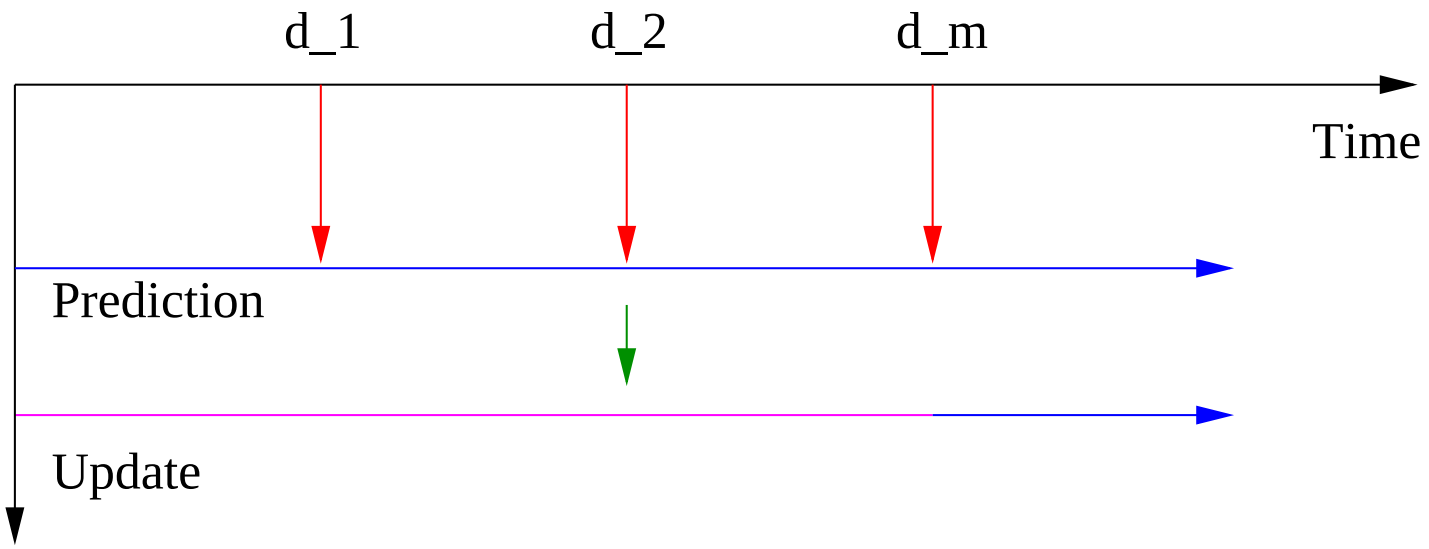
- Measurement errors are independent in time

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

Bayes for discrete state

- Bayes theorem then becomes

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}),$$



Rewrite as:

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b)$$

$$\prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha})$$

$$\prod_{i=i(1)+1}^{i(2)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_2 | \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha}) \cdots$$

$$\prod_{i=i(m-1)+1}^{i(m)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_m | \boldsymbol{\psi}_{i(m)}, \boldsymbol{\alpha})$$

$$\prod_{i=i(m)+1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha})$$

Sequential processing of measurements

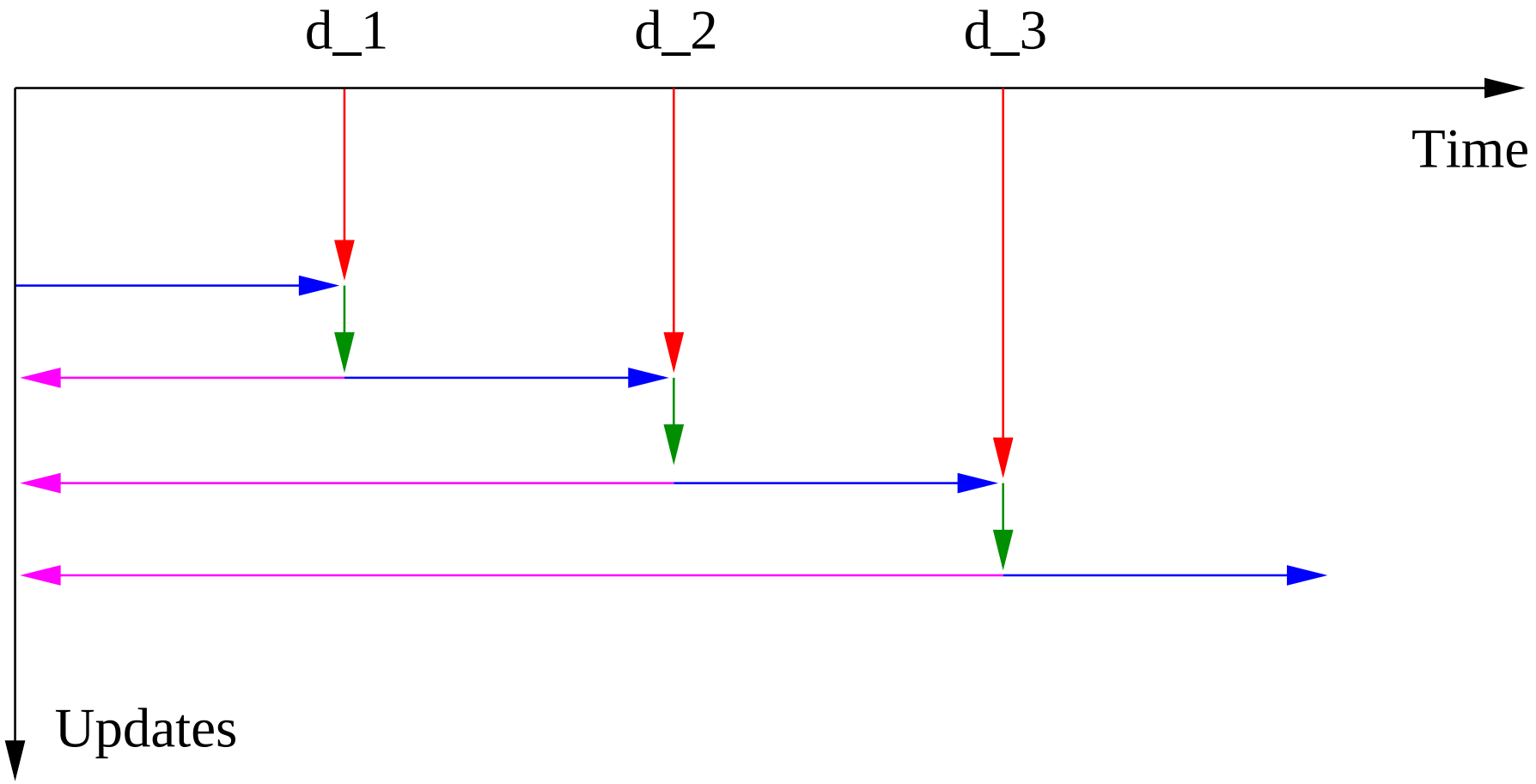
- First update

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha})$$

- Second update

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1, \boldsymbol{d}_2) \propto f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \prod_{i=i(1)+1}^{i(2)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_2 | \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha})$$

Sequential processing of measurements



Summary

- Independent measurements can be processed sequentially in time.
- Sequence of inverse problems.
- Solution of one sub-problem is prior for next.
- No approximations.

Part II

The generalized inverse

Generalized inverse formulation

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b)$$

$$\prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}),$$

Gaussian priors.

$\Downarrow$

Least squares variational inverse problem.

Gaussian priors

Prior for parameters:

- Prior estimate: $\boldsymbol{\alpha}_0(\boldsymbol{x})$.
- Error covariance: $\boldsymbol{C}_{\alpha\alpha}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ with inverse $\boldsymbol{W}_{\alpha\alpha}(\boldsymbol{x}_1, \boldsymbol{x}_2)$

$$\int_{\mathcal{D}} \boldsymbol{C}_{\alpha\alpha}(\boldsymbol{x}_1, \boldsymbol{x}_3)\boldsymbol{W}_{\alpha\alpha}(\boldsymbol{x}_3, \boldsymbol{x}_2)d\boldsymbol{x}_3 = \boldsymbol{\delta}(\boldsymbol{x}_1 - \boldsymbol{x}_2).$$

- Prior pdf for $\boldsymbol{\alpha}$:

$$f(\boldsymbol{\alpha}) \propto \exp\left\{-\frac{1}{2}\iint_{\mathcal{D}} (\boldsymbol{\alpha}(\boldsymbol{x}_1) - \boldsymbol{\alpha}_0(\boldsymbol{x}_1))^{\mathrm{T}} \boldsymbol{W}_{\alpha\alpha}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\cdots)d\boldsymbol{x}_1 d\boldsymbol{x}_2\right\}.$$

Conditional joint density

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

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

- Minimum of $\mathcal{J}$ is the MLH solution for $\boldsymbol{\psi}(\boldsymbol{x}, t)$ and $\boldsymbol{\alpha}(\boldsymbol{x})$.

Example with scalar model

Least squares penalty function

$$\begin{aligned}\mathcal{J}[\psi,\alpha] = &\left(\frac{\partial\psi}{\partial t} - g(\psi,\alpha)\right) \bullet W_{qq} \bullet \left(\frac{\partial\psi}{\partial t} - g(\psi,\alpha)\right)\\ &+ (\psi(t_0) - \Psi_0) W_{aa} (\psi(t_0) - \Psi_0)\\ &+ (\alpha - \alpha_0) W_{\alpha\alpha} (\alpha - \alpha_0)\\ &+ (\boldsymbol{d} - \mathcal{M}[\psi])^{\mathrm{T}} \boldsymbol{W}_{\epsilon\epsilon} (\boldsymbol{d} - \mathcal{M}[\psi])\end{aligned}$$

with

$$\bullet \equiv \int_{t_0}^{t_k} dt.$$

Euler-Lagrange equations

$$\frac{\partial \psi}{\partial t} = g(\psi, \alpha) + C_{qq} \bullet \lambda$$

$$\psi(t_0) = \Psi_0 + C_{aa}\lambda(t_0)$$

$$\frac{\partial \lambda}{\partial t} = -\left.\frac{\partial g}{\partial \psi}\right|_{\substack{\psi \\ \alpha}} \lambda - \mathcal{M}^{\mathrm{T}}[\delta]\boldsymbol{W}_{\epsilon\epsilon}(\boldsymbol{d} - \mathcal{M}[\psi])$$

$$\lambda(t_k) = 0$$

$$\alpha = \alpha_0 + C_{\alpha\alpha} \int_{t_0}^{t_k} \left.\frac{\partial g}{\partial \alpha}\right|_{\substack{\psi \\ \alpha}} \lambda \, dt$$

Euler-Lagrange equations

State estimation for each iterate α_{l+1}

$$\frac{\partial \psi}{\partial t} = g(\psi, \alpha) + C_{qq} \bullet \lambda$$

$$\psi(t_0) = \Psi_0 + C_{aa}\lambda(t_0)$$

$$\frac{\partial \lambda}{\partial t} = -\left.\frac{\partial g}{\partial \psi}\right|_{\substack{\psi \\ \alpha}} \lambda - \mathcal{M}^{\mathrm{T}}[\delta]\boldsymbol{W}_{\epsilon\epsilon}(\boldsymbol{d} - \mathcal{M}[\psi])$$

$$\lambda(t_k) = 0$$

Parameter estimation using iteration in α

$$\alpha_{l+1} = \alpha_l - \gamma \left(\alpha_l - \alpha_0 - C_{\alpha\alpha} \int_{t_0}^{t_k} \left.\frac{\partial g}{\partial \alpha}\right|_{\substack{\psi_l \\ \alpha_l}} \lambda_l \, dt \right)$$

Parameter estimation

State estimation with $C_{qq} = 0$ and $C_{aa} = 0$

$$\frac{\partial \psi}{\partial t} = g(\psi, \alpha)$$
$$\psi(t_0) = \Psi_0$$
$$\frac{\partial \lambda}{\partial t} = -\left.\frac{\partial g}{\partial \psi}\right|_{\substack{\psi \\ \alpha}} \lambda - \mathcal{M}^{\mathrm{T}}[\delta] \boldsymbol{W}_{\epsilon\epsilon}(\boldsymbol{d} - \mathcal{M}[\psi])$$
$$\lambda(t_k) = 0$$

$$\alpha_{l+1} = \alpha_l - \gamma \left(\alpha_l - \alpha_0 - C_{\alpha\alpha} \int_{t_0}^{t_k} \left.\frac{\partial g}{\partial \alpha}\right|_{\substack{\psi_l \\ \alpha_l}} \lambda_l \, dt \right)$$

- $\Rightarrow$ Adjoint method.

Reformulation

Define Lagrangian

$$\mathcal{L}[\psi, \alpha, \lambda, \mu] = \int_{t_0}^{t_k} \left(\frac{\partial \psi}{\partial t} - g(\psi, \alpha) \right) \lambda \, dt$$
$$+ (\psi(t_0) - \Psi_0)\mu$$
$$+ (\alpha - \alpha_0) W_{\alpha\alpha} (\alpha - \alpha_0)$$
$$+ (\boldsymbol{d} - \mathcal{M}[\psi])^{\mathrm{T}} \boldsymbol{W}_{\epsilon\epsilon} (\boldsymbol{d} - \mathcal{M}[\psi])$$

Variation with respect to:

- λ returns model.
- μ returns initial condition.
- α returns E-L equations for parameter estimation problem.

Summary

- The generalized inverse:
  - Combined parameter and state estimation.
  - MLH solution consistent with Gaussian priors.
- State estimation:
  - Solves only for model state (representer method).
- Parameter estimation:
  - Solves only for parameters (adjoint method).
- Issues:
  - Highly nonlinear problem!
  - Results in a single model realization!
  - Hard to compute uncertainty estimates!
  - Does not allow for sequential processing of measurements!

Part III

Ensemble methods

ES: Ensemble Smoother

EnKS: Ensemble Kalman Smoother

EnKF: Ensemble Kalman Filter

- Ensemble representation for pdfs.
- Ensemble predition for time evolution of pdfs.
- Assume Gaussian pdf for model prediction and use “variance minimizing” analysis schemes.

The error covariance matrix

Define ensemble covariances around the ensemble mean

$$\boldsymbol{C}_{\psi\psi} = \overline{\left(\boldsymbol{\psi} - \overline{\boldsymbol{\psi}}\right)\left(\boldsymbol{\psi} - \overline{\boldsymbol{\psi}}\right)^{\mathrm{T}}}$$

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

$\Downarrow$

Representation by a finite ensemble of model states.

Ensemble representation

- Define ensemble matrix

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

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

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

Dynamical evolution of error statistics

- Members evolve according to stochastic model dynamics

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

- The pdf evolves according to Kolmogorov's equation

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

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

Variance minimizing analysis

Given

- First guess, $\boldsymbol{\psi}^{\mathrm{f}}$, and prediction error covariance, $\boldsymbol{C}^{\mathrm{f}}_{\psi\psi}$.
- Measurements, $\boldsymbol{d}$, and error covariance, $\boldsymbol{C}_{\epsilon\epsilon}$.

The variance minimizing analysis is minimum of

$$\mathcal{J}[\boldsymbol{\psi}^{\mathrm{a}}] = \left(\boldsymbol{\psi}^{\mathrm{a}} - \boldsymbol{\psi}^{\mathrm{f}}\right)^{\mathrm{T}} \bullet \boldsymbol{C}^{-1}_{\psi\psi} \bullet \left(\boldsymbol{\psi}^{\mathrm{a}} - \boldsymbol{\psi}^{\mathrm{f}}\right) + \left(\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}^{\mathrm{a}}]\right)^{\mathrm{T}} \boldsymbol{C}^{-1}_{\epsilon\epsilon} \left(\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}^{\mathrm{a}}]\right).$$

Analysis equations

Update of estimate

$$\boldsymbol{\psi}^{\mathrm{a}} = \boldsymbol{\psi}^{\mathrm{f}} + \mathcal{M}^{\mathrm{T}} \left[\boldsymbol{C}^{\mathrm{f}}_{\psi\psi} \right] \left(\mathcal{M}^{\mathrm{T}} \left[\mathcal{M} \left[\boldsymbol{C}^{\mathrm{f}}_{\psi\psi} \right] \right] + \boldsymbol{C}_{\epsilon\epsilon} \right)^{-1} \left(\boldsymbol{d} - \mathcal{M} \left[\boldsymbol{\psi}^{\mathrm{f}} \right] \right)$$

Error covariance for update

$$\boldsymbol{C}^{\mathrm{a}}_{\psi\psi} = \boldsymbol{C}^{\mathrm{f}}_{\psi\psi} - \mathcal{M}^{\mathrm{T}} \left[\boldsymbol{C}^{\mathrm{f}}_{\psi\psi} \right] \left(\mathcal{M}^{\mathrm{T}} \left[\mathcal{M} \left[\boldsymbol{C}^{\mathrm{f}}_{\psi\psi} \right] \right] + \boldsymbol{C}_{\epsilon\epsilon} \right)^{-1} \mathcal{M} \left[\boldsymbol{C}^{\mathrm{f}}_{\psi\psi} \right]$$

Ensemble analysis equations

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

$$
\begin{aligned}
\boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + \mathcal{M}^{\mathrm{T}}\left[\boldsymbol{C}^{\mathrm{e}}_{\psi\psi}\right] \\
&\quad \left(\mathcal{M}^{\mathrm{T}}\left[\mathcal{M}\left[\boldsymbol{C}^{\mathrm{e}}_{\psi\psi}\right]\right] + \boldsymbol{C}_{\epsilon\epsilon}\right)^{-1}\left(\boldsymbol{D} - \mathcal{M}\left[\boldsymbol{A}\right]\right) \\
&= \boldsymbol{A} + \boldsymbol{A}'\mathcal{M}^{\mathrm{T}}\left[\boldsymbol{A}'\right] \\
&\quad \left(\mathcal{M}\left[\boldsymbol{A}'\right]\mathcal{M}^{\mathrm{T}}\left[\boldsymbol{A}'\right] + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{D} - \mathcal{M}\left[\boldsymbol{A}\right]\right) \\
&\vdots \\
&= \boldsymbol{A}\boldsymbol{X}
\end{aligned}
$$

$\Rightarrow$ Updated ensemble with correct mean and covariance!

ES: The Ensemble Smoother

- Joint pdf for model prediction

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) = f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}).$$

- Ensemble representation for the prior pdfs.
- Ensemble integration gives the joint pdf.

Problem: how to compute the Bayesian update

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) =$$

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}),$$

ES: Ensemble representation

- Define ensemble matrix for joint state from t_0 to t_k as

$$\widetilde{\boldsymbol{A}}_k = \begin{pmatrix} \boldsymbol{A}(\boldsymbol{x}, t_0) \\ \vdots \\ \boldsymbol{A}(\boldsymbol{x}, t_k) \end{pmatrix}$$

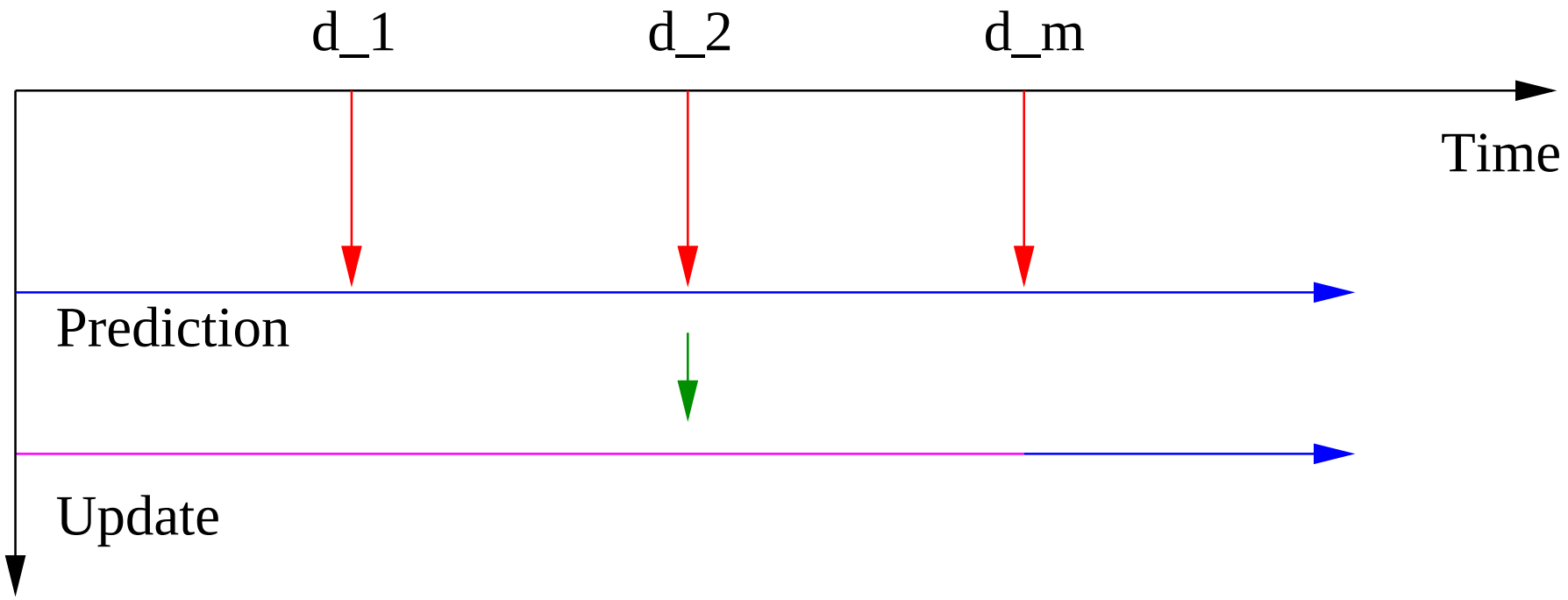
which defines the space–time ensemble covariance

$$\widetilde{\boldsymbol{C}}^{\mathrm{e}}_{\psi\psi}(\boldsymbol{x}_1, t_1, \boldsymbol{x}_2, t_2) = \frac{\widetilde{\boldsymbol{A}}'_k(\boldsymbol{x}_1, t_1)\widetilde{\boldsymbol{A}}'^{\mathrm{T}}_k(\boldsymbol{x}_2, t_2)}{N-1}$$

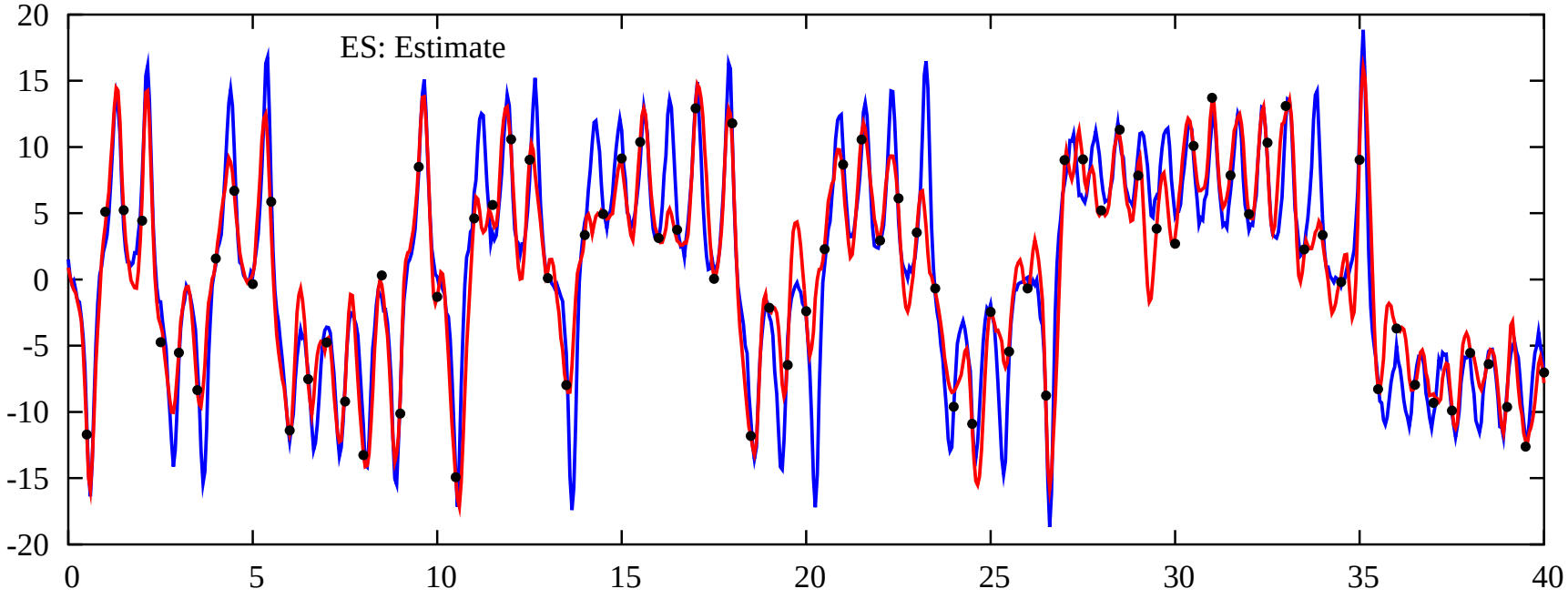
ES: Analysis scheme

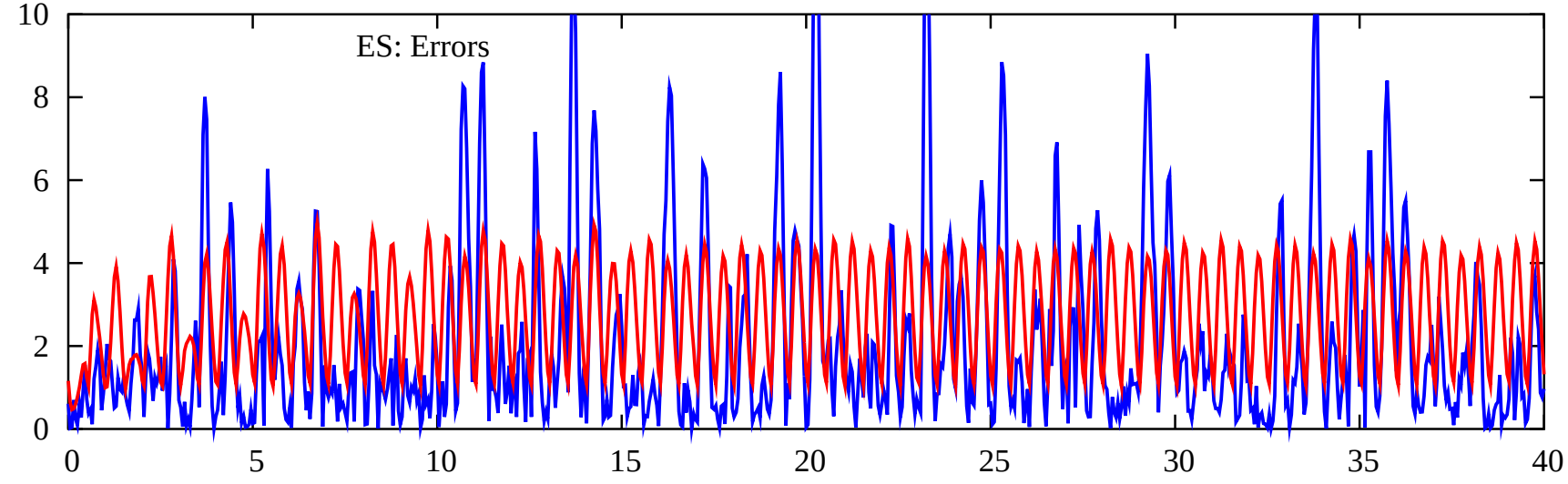
$$
\begin{aligned}
\widetilde{\boldsymbol{A}}_k^{\mathrm{a}} &= \widetilde{\boldsymbol{A}}_k + \mathcal{M}^{\mathrm{T}} \left[\widetilde{\boldsymbol{C}}_{\psi\psi}^{\mathrm{e}} \right] \\
&\quad \left(\mathcal{M}^{\mathrm{T}} \left[\mathcal{M} \left[\widetilde{\boldsymbol{C}}_{\psi\psi}^{\mathrm{e}} \right] \right] + \boldsymbol{C}_{\epsilon\epsilon} \right)^{-1} \left(\boldsymbol{D} - \mathcal{M} \left[\widetilde{\boldsymbol{A}}_k \right] \right) \\
&= \widetilde{\boldsymbol{A}}_k + \widetilde{\boldsymbol{A}}_k' \mathcal{M}^{\mathrm{T}} \left[\widetilde{\boldsymbol{A}}_k' \right] \\
&\quad \left(\mathcal{M} \left[\widetilde{\boldsymbol{A}}_k' \right] \mathcal{M}^{\mathrm{T}} \left[\widetilde{\boldsymbol{A}}_k' \right] + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}} \right)^{-1} \left(\boldsymbol{D} - \mathcal{M} \left[\widetilde{\boldsymbol{A}}_k \right] \right) \\
&\vdots \\
&= \widetilde{\boldsymbol{A}}_k \boldsymbol{X}
\end{aligned}
$$

ES solution procedure



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

- Starts from Bayes with sequential processing of data

E.g. first update:

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \propto$$

$$f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}),$$

Use ensemble representations

$$\widetilde{\boldsymbol{A}}_{i(1)} = \begin{pmatrix} \boldsymbol{A}(\boldsymbol{x}, t_0) \\ \vdots \\ \boldsymbol{A}(\boldsymbol{x}, t_{i(1)}) \end{pmatrix} \quad \text{and} \quad \boldsymbol{D}_1.$$

EnKS: Analysis equation

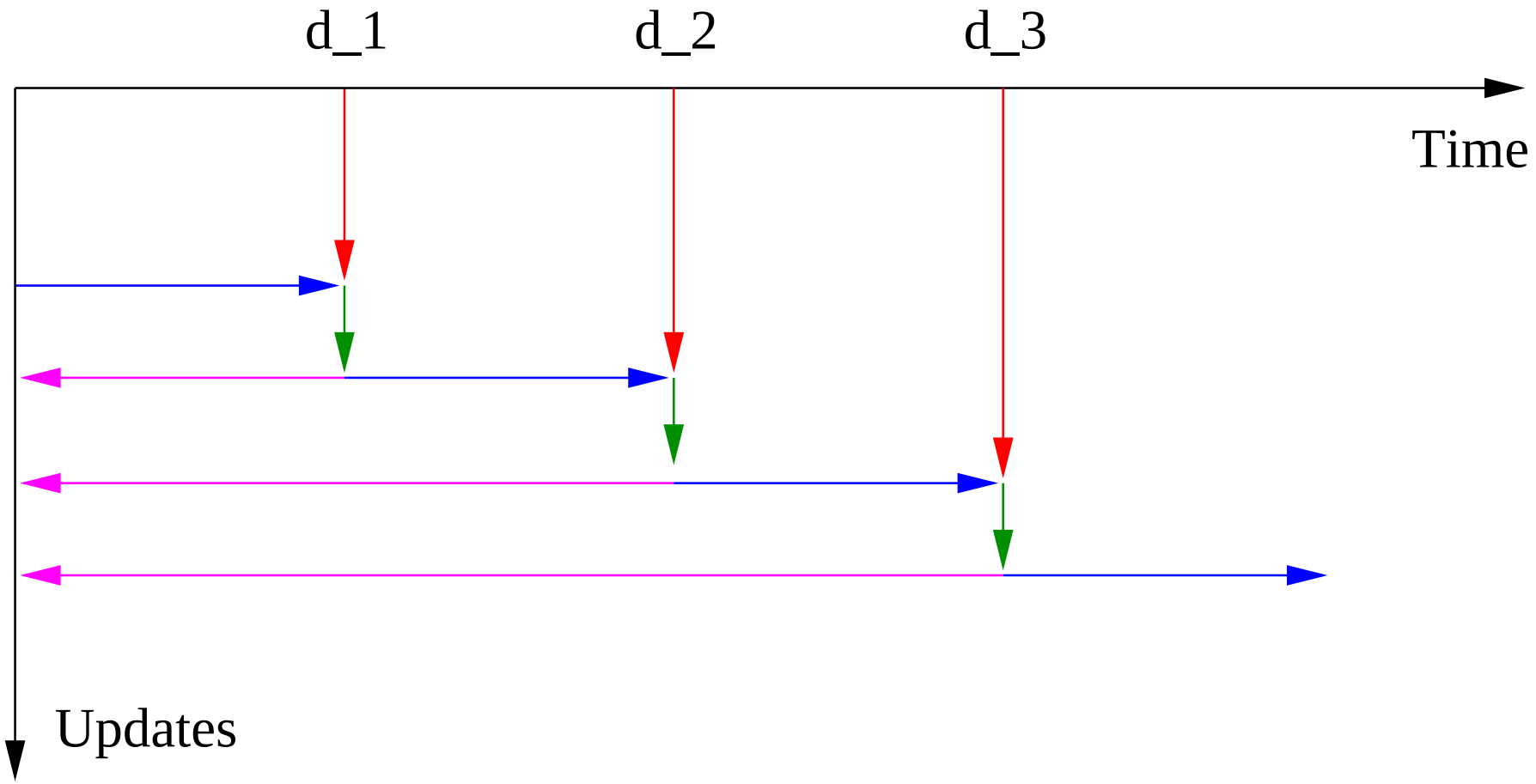
- The EnKS analysis equation for the first update is

$$\widetilde{\boldsymbol{A}}^{\mathrm{a}}_{i(1)} = \widetilde{\boldsymbol{A}}_{i(1)} + \widetilde{\boldsymbol{A}}'_{i(1)} \mathcal{M}_1^{\mathrm{T}} \left[\boldsymbol{A}'_{i(1)} \right]$$

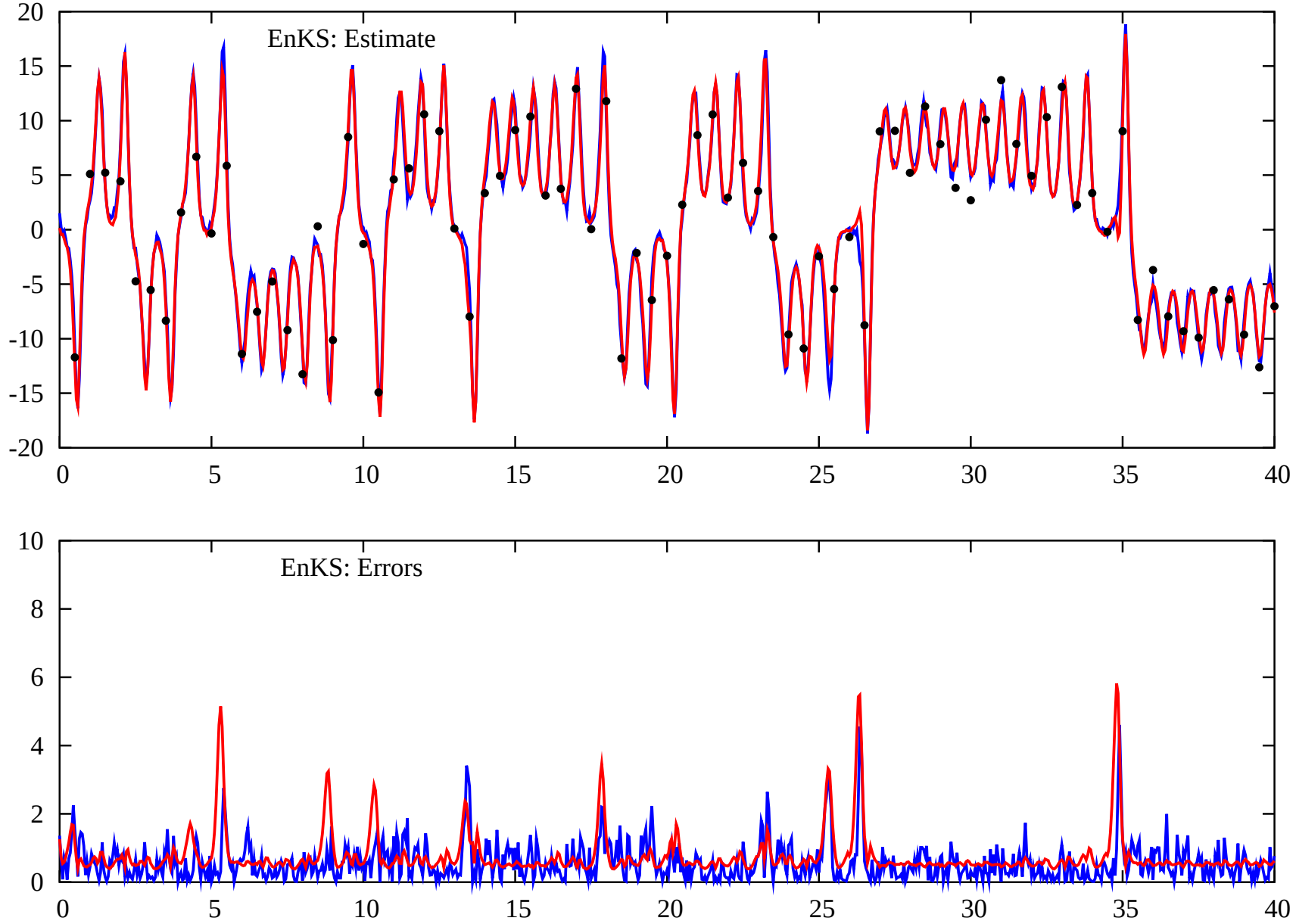
$$\left(\mathcal{M}_1 \left[\boldsymbol{A}'_{i(1)} \right] \mathcal{M}_1^{\mathrm{T}} \left[\boldsymbol{A}'_{i(1)} \right] + \boldsymbol{E}_1 \boldsymbol{E}_1^{\mathrm{T}} \right)^{-1} \left(\boldsymbol{D}_1 - \mathcal{M}_1 \left[\boldsymbol{A}_{i(1)} \right] \right)$$

$$= \widetilde{\boldsymbol{A}}_{i(1)} \boldsymbol{X}_1$$

EnKS solution procedure



EnKS solution



EnKS summary

- ES and EnKS gives 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.
- Accuracy of analysis is determined by:
  - the accuracy of $\boldsymbol{X}_j$,
  - the properties of the ensemble error space.

EnKF: Ensemble Kalman Filter

- Special case of EnKS.
- Only update state at data times

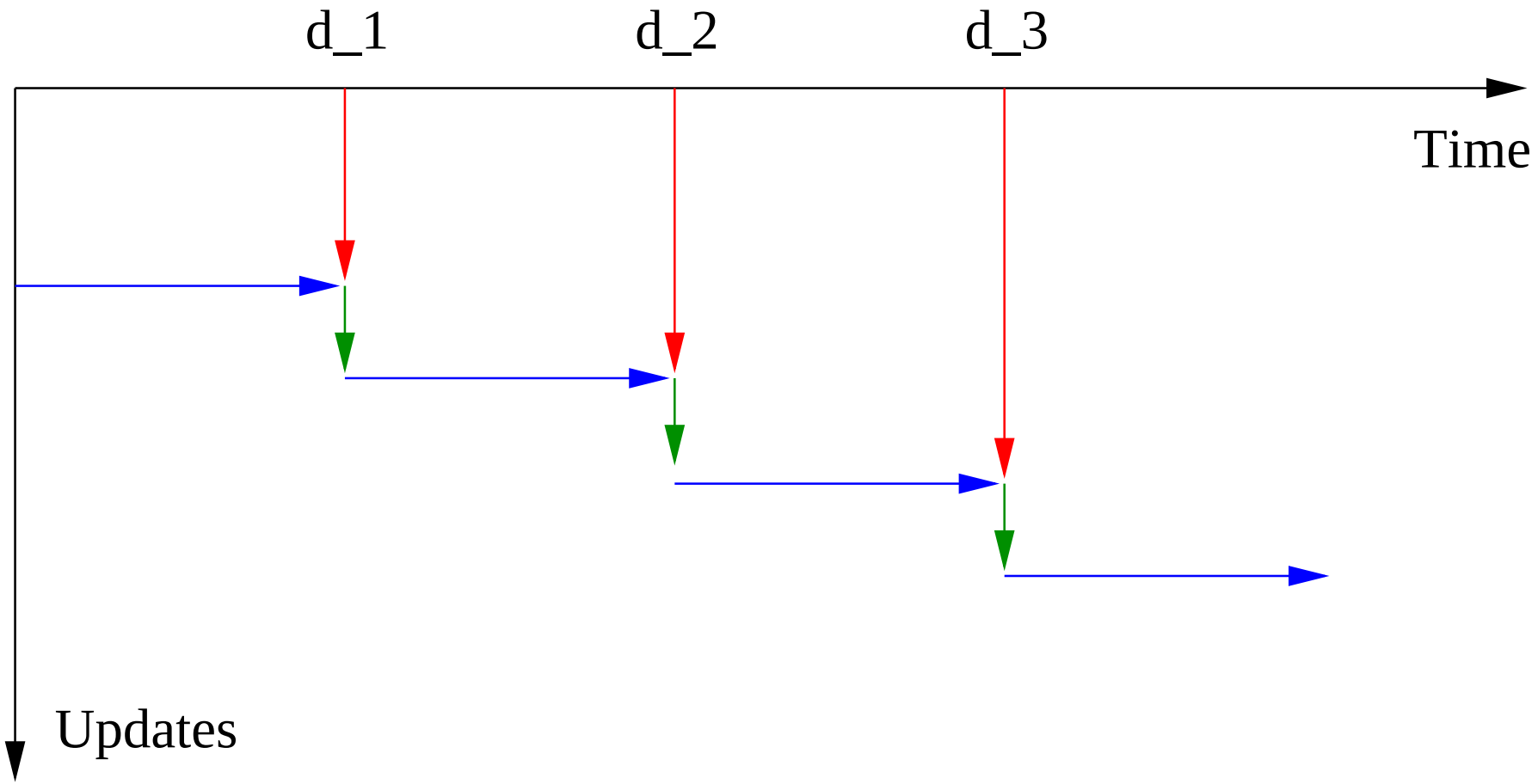
$$\boldsymbol{A}^{\mathrm{a}}_{i(1)} = \boldsymbol{A}_{i(1)} + \boldsymbol{A}'_{i(1)} \mathcal{M}_1^{\mathrm{T}} \left[\boldsymbol{A}'_{i(1)} \right]$$

$$\left(\mathcal{M}_1 \left[\boldsymbol{A}'_{i(1)} \right] \mathcal{M}_1^{\mathrm{T}} \left[\boldsymbol{A}'_{i(1)} \right] + \boldsymbol{E}_1 \boldsymbol{E}_1^{\mathrm{T}} \right)^{-1} \left(\boldsymbol{D}_1 - \mathcal{M}_1 \left[\boldsymbol{A}_{i(1)} \right] \right)$$

$$= \boldsymbol{A}_{i(1)} \boldsymbol{X}_1$$

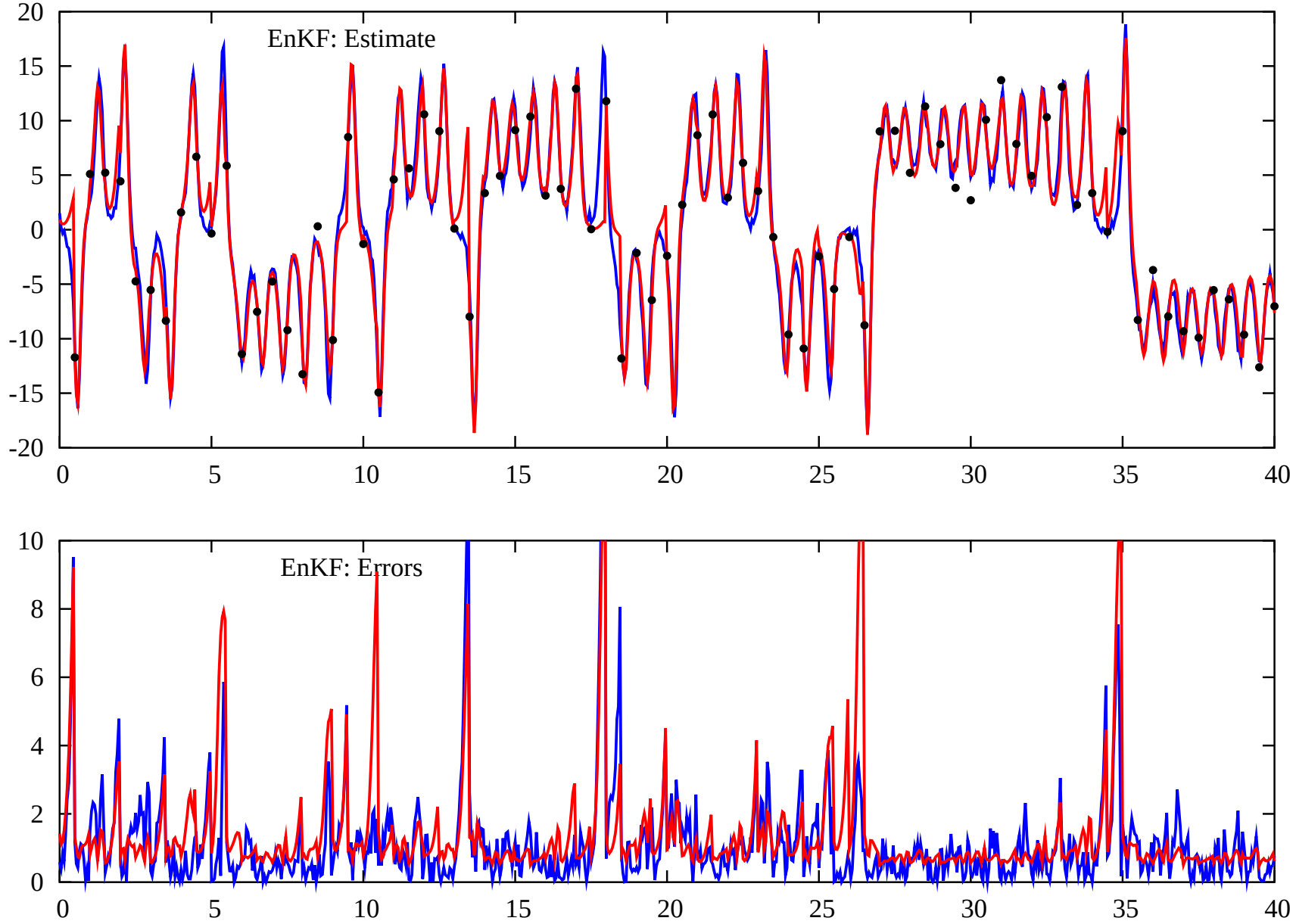
- EnKF forms a prior for the EnKS.

EnKF solution procedure



$$\boldsymbol{A}_{\mathrm{EnKS}}(\boldsymbol{x}, t_l) = \boldsymbol{A}_{\mathrm{EnKF}}(\boldsymbol{x}, t_l) \prod_{j'=j}^{m} \boldsymbol{X}_{j'}.$$

EnKF solution



EnKF summary

- A prediction from EnKF and EnKS will be identical.
- EnKF suitable for forecasting purposes.
- The EnKS solution can be computed using the EnKF solution as a prior.

Part IV

A simple example

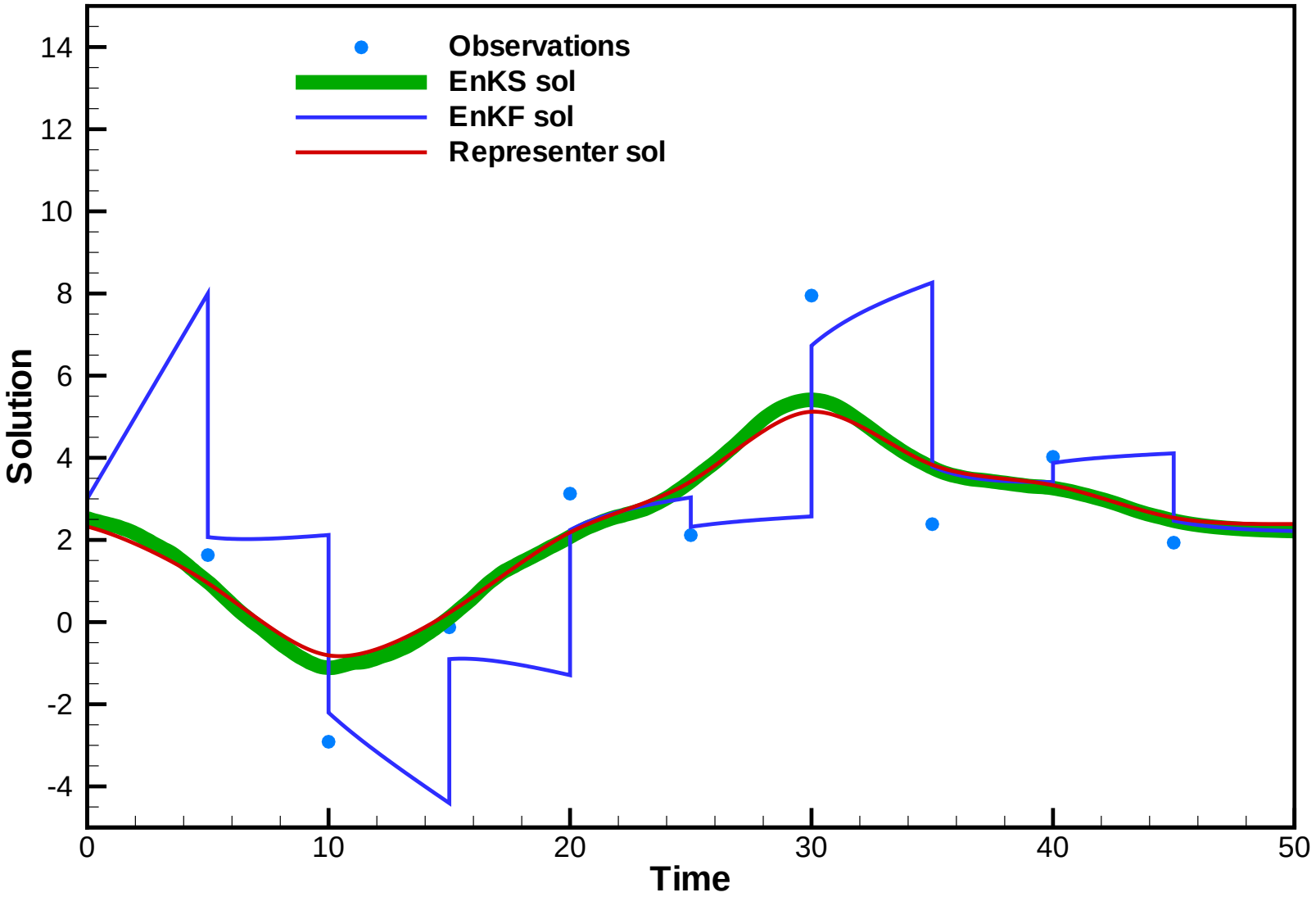
Example

- Scalar model

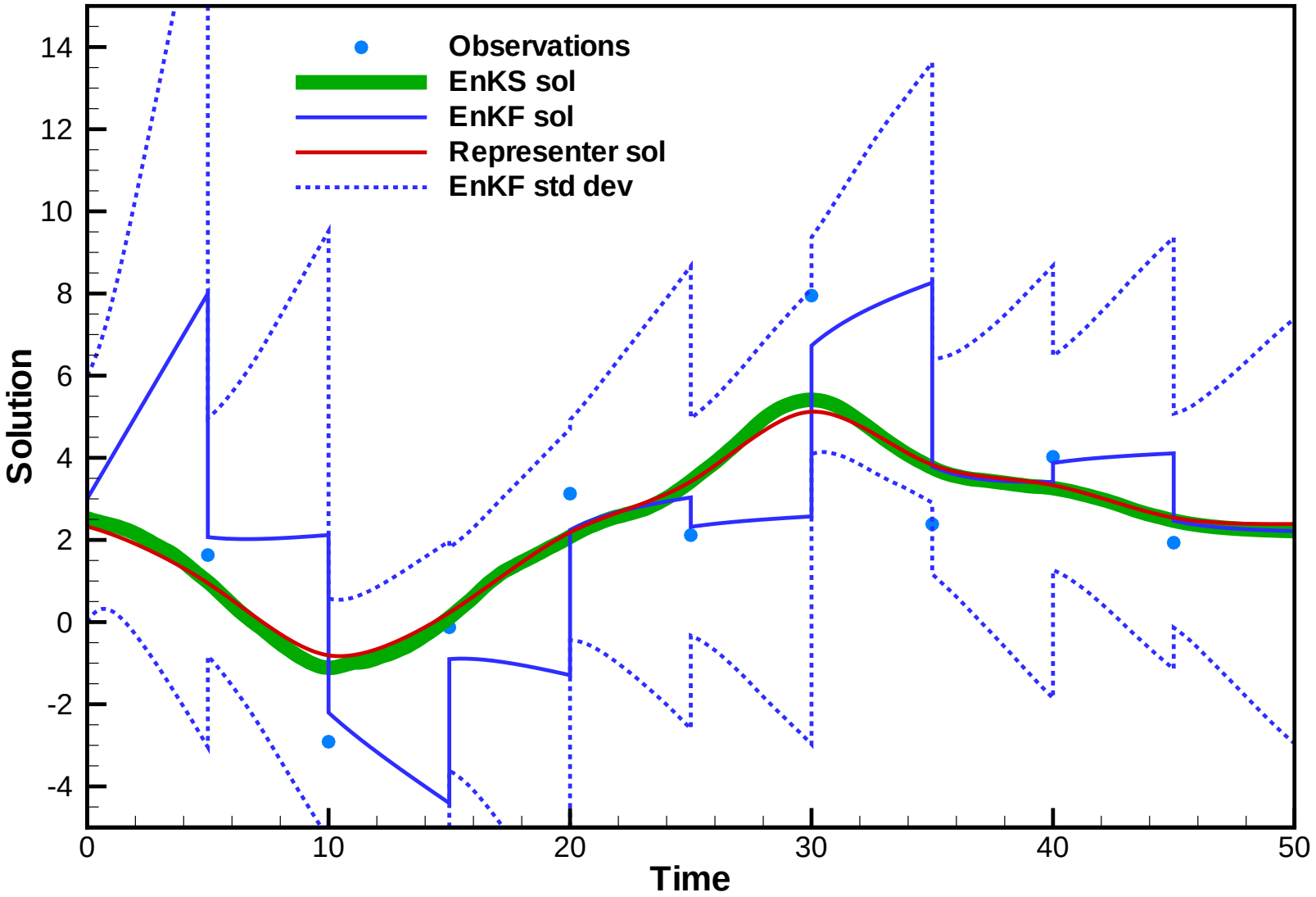
$$\frac{\partial \psi}{\partial t} = 1 - \alpha + q,$$
$$\psi(t=0) = 3 + a,$$
$$\alpha = 0 + \alpha',$$
$$\boldsymbol{d} = \mathcal{M}(\psi^t) + \boldsymbol{\epsilon}.$$

- 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.

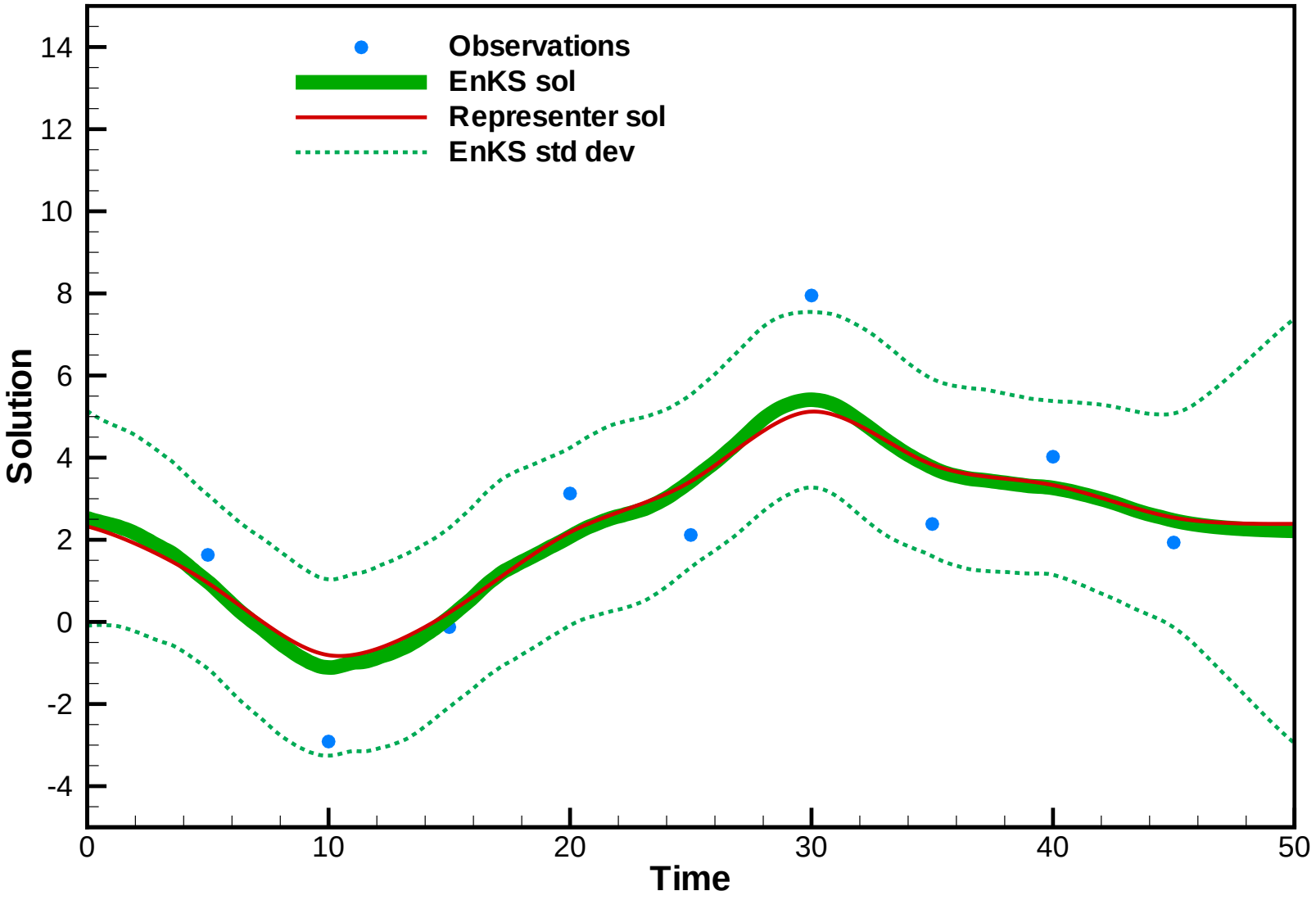
State and parameter estimation



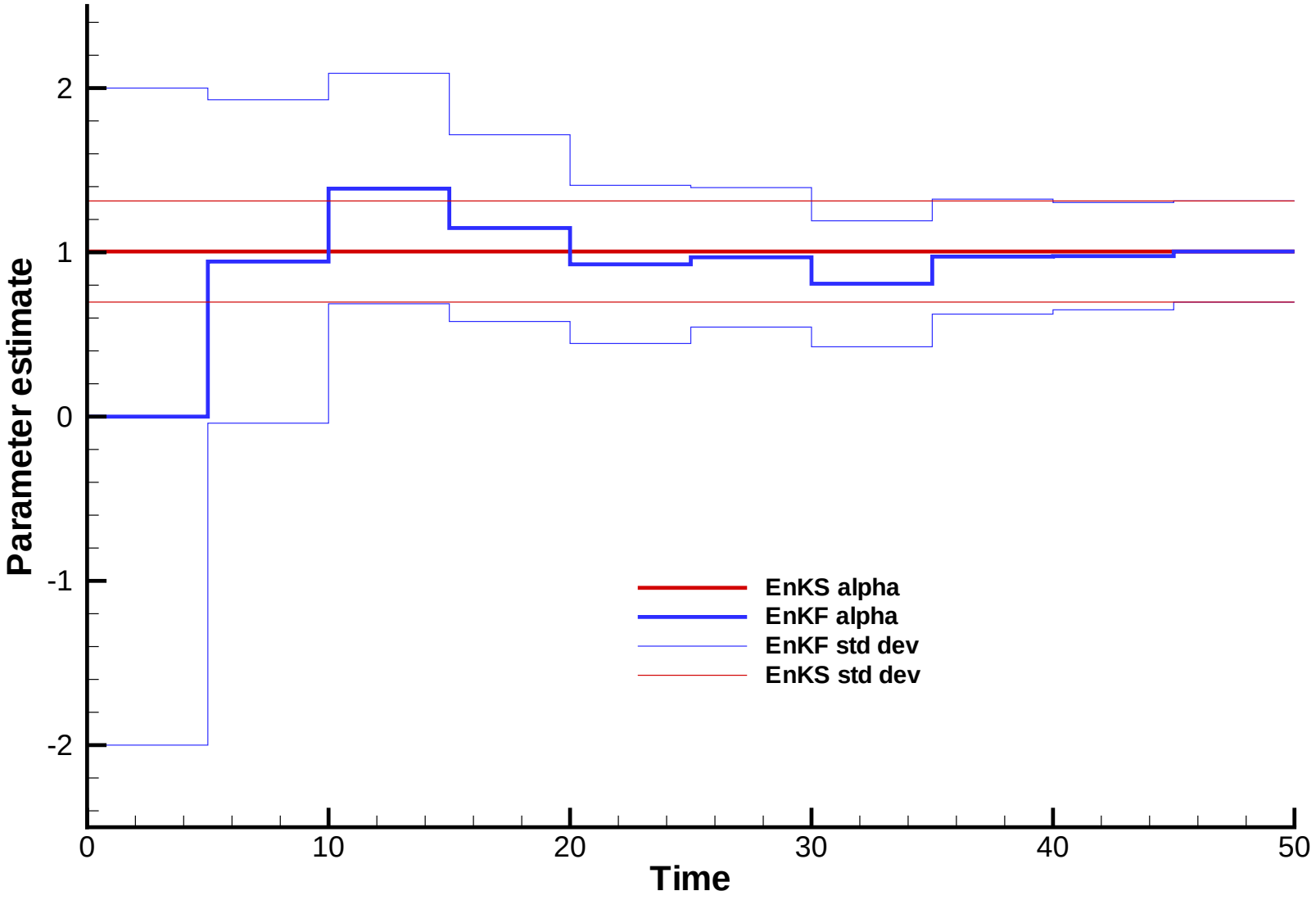
State and parameter estimation



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 measuremets.
      - Introduces Gaussianity!
      - Advantage with nonlinear dynamics.
- There is more to parameter estimation than just estimating the parameters!