Analysis of the Ensemble Kalman Filter for Inverse Problems

Mike Christie, Claudia Schillings, Andrew Stuart

11th International EnKF Workshop

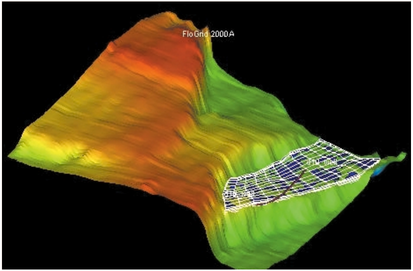
research supported by the Engineering and Physical Sciences Research Council

Outline

Reservoir Modeling

Teal South

- Reservoir in the Gulf of Mexico
- Monthly production rates of oil, water and gas avalaible



Source: Christie et al.

Model

- Five geological layers with uniform properties
- 9 unknown parameters (porosity, horizontal permeability multipliers for each layer, vertical to horizontal permeability ratio, rock compressibility, aquifer strength)
- Matching to the field oil production rate
- Eclipse used to simulate the flow in porous media

Reservoir Modeling

Results (100 forward simulations)

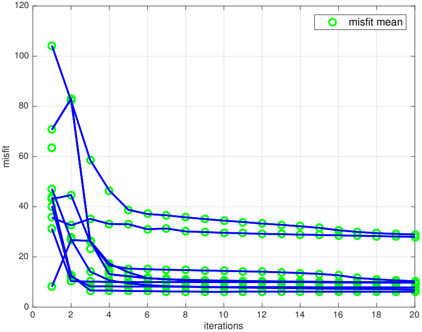


Figure: Misfit of the mean to the (noisy) observational data, J=5, 20 iterations.

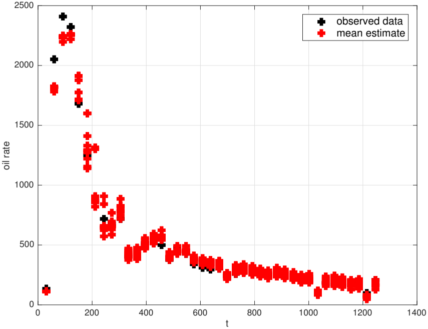


Figure: Prediction of the mean compared to the (noisy) observational data, J=5, 20 iterations.

Inverse Problem

Find the unknown data $u \in X$ from noisy observations

$$y = \mathcal{G}(u) + \eta \qquad \text{with } \eta \sim \mathcal{N}(0, \Gamma)$$

- $u \in X$ parameter vector / parameter function
- $\mathcal{G} : X \to Y$ forward response operator; X, Y separable Hilbert spaces
- y result / observations
- Evaluation of $\mathcal{G}$ expensive

Inverse Problem

Find the unknown data $u \in X$ from noisy observations

$$y = \mathcal{G}(u) + \eta \qquad \text{with } \eta \sim \mathcal{N}(0, \Gamma)$$

- $u \in X$ parameter vector / parameter function
- $\mathcal{G} : X \to Y$ forward response operator; X, Y separable Hilbert spaces
- y result / observations
- Evaluation of $\mathcal{G}$ expensive

Deterministic optimisation problem

$$\min_u \frac{1}{2} \|y - \mathcal{G}(u)\|^2 + R(u)$$

- R regularisation term

Bayesian inverse problem

- u, η, y random variables / fields
- Prior μ_0, posterior μ^y

Bayesian Inverse Problem

Find the unknown data $u \in X$ from noisy observations

$$y = \mathcal{G}(u) + \eta$$

Bayes' Theorem (A. M. Stuart 2010)

Assuming $\mathcal{G} \in C(X, Y)$ and $\mu_0(X) = 1$, then the posterior measure μ^y on $u|y$ is absolutely continuous w.r. to the prior on u and

$$\mu^y(du) = \frac{1}{Z} \exp(-\Phi(u))\mu_0(du)$$

with $\Phi : X \mapsto \mathbb{R}$, $\Phi(u) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$ and $Z = \int \exp(-\Phi(u))\mu_0(du)$.

Bayesian Inverse Problem

Find the unknown data $u \in X$ from noisy observations

$$y = \mathcal{G}(u) + \eta$$

Ensemble Kalman Filter

- Fully Bayesian inversion is often too expensive.
- EnKF is widely used.
- Currently, very little analysis of the EnKF is available.

Aim: Build analysis of properties of EnKF for fixed ensemble size.

Bayesian Inverse Problem

Find the unknown data $u \in X$ from noisy observations

$$y = \mathcal{G}(u) + \eta$$

Ensemble Kalman Filter

Optimisation viewpoint

Study of the properties of the EnKF as a regularisation technique for minimisation of the least-squares misfit functional

Continuous time limit

Analysis of the properties of the differential equations

EnKF for Inverse Problems

Sequence of Interpolating Measures

For $N \in \mathbb{N}, h := 1/N$, we define a sequence of measures $\mu_n \ll \mu_0$, $n = 1, \ldots, N$, which evolve the prior μ_0 into the posterior distribution $\mu_N = \mu^y$, by

$$\mu_{n+1}(du) = \frac{Z_n}{Z_{n+1}} \exp(-h\Phi(u))\mu_n(du) \Leftrightarrow \mu_{n+1} = L_n \mu_n$$

with nonlinear operator L_n corresponding to application of Bayes' theorem and normalisation constant $Z_n = \int \exp(-nh\Phi(u))\mu_0(du)$ with $\Phi(u) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$.

Ensemble of Interacting Particles

Initial ensemble $\{u_0^{(j)}\}_{j=1}^J$ constructed by prior knowledge, $u^{(j)} \sim \mu_0$ iid for $J < \infty$.

Linearisation of $\mathrm{L_n}$ and approximation of μ_n by a J-particle Dirac measure leads to the EnKF method.

EnKF for Inverse Problems

Sequence of Interpolating Measures

For $N \in \mathbb{N}, h := 1/N$, we define a sequence of measures $\mu_n \ll \mu_0$, $n = 1, \ldots, N$, which evolve the prior μ_0 into the posterior distribution $\mu_N = \mu^y$, by

$$\mu_{n+1}(du) = \frac{Z_n}{Z_{n+1}} \exp(-h\Phi(u))\mu_n(du) \Leftrightarrow \mu_{n+1} = L_n \mu_n$$

with nonlinear operator L_n corresponding to application of Bayes' theorem and normalisation constant $Z_n = \int \exp(-nh\Phi(u))\mu_0(du)$ with $\Phi(u) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$.

Ensemble of Interacting Particles

Initial ensemble $\{u_0^{(j)}\}_{j=1}^J$ constructed by prior knowledge, $u^{(j)} \sim \mu_0$ iid for $J < \infty$.

Linearisation of $\mathrm{L_n}$ and approximation of μ_n by a J-particle Dirac measure leads to the EnKF method.

EnKF for Inverse Problems

Sequence of Interpolating Measures

For $N \in \mathbb{N}, h := 1/N$, we define a sequence of measures $\mu_n \ll \mu_0$, $n = 1, \ldots, N$, which evolve the prior μ_0 into the posterior distribution $\mu_N = \mu^y$, by

$$\mu_{n+1}(du) = \frac{Z_n}{Z_{n+1}} \exp(-h\Phi(u))\mu_n(du) \Leftrightarrow \mu_{n+1} = L_n \mu_n$$

with nonlinear operator L_n corresponding to application of Bayes' theorem and normalisation constant $Z_n = \int \exp(-nh\Phi(u))\mu_0(du)$ with $\Phi(u) = \frac{1}{2}|y - \mathcal{G}(u)|_\Gamma^2$.

Ensemble of Interacting Particles

Initial ensemble $\{u_0^{(j)}\}_{j=1}^J$ constructed by prior knowledge, $u^{(j)} \sim \mu_0$ iid for $J < \infty$.

Linearisation of $\mathbf{L_n}$ and approximation of $\mu_\mathbf{n}$ by a J-particle Dirac measure leads to the EnKF method.

EnKF for Inverse Problems

Update of the EnKF for Inverse Problems

$$u_{n+1}^{(j)} = u_n^{(j)} + C_{n+1}^{up}(C_{n+1}^{pp} + \frac{1}{h}\Gamma)^{-1}(y_{n+1}^{(j)} - \mathcal{G}(u_n^{(j)}))$$

with empirical covariances

$$C_{n+1}^{up} = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)} \otimes \mathcal{G}(u_n^{(j)}) - \overline{u}_n \otimes \overline{\mathcal{G}}(u_n)$$

$$C_{n+1}^{pp} = \frac{1}{J}\sum_{j=1}^{J} \mathcal{G}(u_n^{(j)}) \otimes \mathcal{G}(u_n^{(j)}) - \overline{\mathcal{G}}(u_n) \otimes \overline{\mathcal{G}}(u_n),$$

mean $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}, \quad \overline{\mathcal{G}}(u_n) = \frac{1}{J}\sum_{j=1}^{J} \mathcal{G}(u_n^{(j)})$

and observations $y_{n+1}^{(j)} = y + \eta_{n+1}^{(j)}$, $\eta_{n+1}^{(j)} \sim N(0, \frac{1}{h}\Gamma)$.

Properties of the EnKF for Inverse Problems e.g. [Iglesias, Law, Stuart 2013]

- The ensemble parameter estimate lies in the linear span of the initial ensemble.
- This linear span property implies that the accuracy of the EnKF estimate is bounded from below by the best approximation in $\mathrm{span}\{u_0^{(1)}, \ldots, u_0^{(J)}\}$.
- In the linear case, the EnKF estimate converges in the limit $J \to \infty$ to the solution of the regularised least-squares problem.

EnKF for Inverse Problems

Update of the EnKF for Inverse Problems

$$u_{n+1}^{(j)} = u_n^{(j)} + C_{n+1}^{up}(C_{n+1}^{pp} + \frac{1}{h}\Gamma)^{-1}(y_{n+1}^{(j)} - \mathcal{G}(u_n^{(j)}))$$

with empirical covariances

$$C_{n+1}^{up} = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)} \otimes \mathcal{G}(u_n^{(j)}) - \overline{u}_n \otimes \overline{\mathcal{G}}(u_n)$$

$$C_{n+1}^{pp} = \frac{1}{J}\sum_{j=1}^{J} \mathcal{G}(u_n^{(j)}) \otimes \mathcal{G}(u_n^{(j)}) - \overline{\mathcal{G}}(u_n) \otimes \overline{\mathcal{G}}(u_n),$$

mean $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}, \quad \overline{\mathcal{G}}(u_n) = \frac{1}{J}\sum_{j=1}^{J} \mathcal{G}(u_n^{(j)})$

and observations $y_{n+1}^{(j)} = y + \eta_{n+1}^{(j)}$, $\eta_{n+1}^{(j)} \sim N(0, \frac{1}{h}\Gamma)$.

Properties of the EnKF for Inverse Problems e.g. [Iglesias, Law, Stuart 2013]

- The ensemble parameter estimate lies in the linear span of the initial ensemble.
- This linear span property implies that the accuracy of the EnKF estimate is bounded from below by the best approximation in $\mathrm{span}\{u_0^{(1)}, \dots, u_0^{(J)}\}$.
- In the linear case, the EnKF estimate converges in the limit $J \to \infty$ to the solution of the regularised least-squares problem.

Continuous Time Limit

Update of the Iterates

$$u_{n+1}^{(j)} = u_n^{(j)} + h\ C_{n+1}^{up}\big(h\ C_{n+1}^{pp} + \Gamma\big)^{-1}(y^\dagger - \mathcal{G}(u_n^{(j)})) \\ + h^{\frac{1}{2}}\ C_{n+1}^{up}\big(h\ C_{n+1}^{pp} + \Gamma\big)^{-1}\Gamma^{\frac{1}{2}}\zeta_{n+1}^{j}$$

with $\zeta_{n+1} \sim \mathcal{N}(0, id)$.

Limiting SDE

Interpreting the iterate as $u_n^{(j)} \approx u^{(j)}(nh)$ gives

$$\mathrm{d}u^{(j)} = C^{up}\Gamma^{-1}(y^\dagger - \mathcal{G}(u^{(j)}))\,\mathrm{d}t + C^{up}\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

where $W^{(1)}, \ldots, W^{(J)}$ are pairwise independent cylindrical Wiener processes and $y^\dagger$ denotes the noisy observational data $\mathcal{G}(u^\dagger) + \eta^\dagger$ with $\eta^\dagger \sim \mathcal{N}(0, \Gamma)$.

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Limiting SDE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Noise-free Case

Limiting ODE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger + \eta - u^{(j)})\,\mathrm{d}t + C(u)A^*\Gamma^{-\frac{1}{2}}\,\mathrm{d}W^{(j)}\,,$$

Continuous Time Limit (Linear Case)

Assumption: Linear response operator $\mathcal{G}(u) = Au$ with $A \in \mathcal{L}(X, Y)$

$$u_{n+1}^{(j)} = u_n^{(j)} + hC(u_n)A^*\Gamma^{-1}(y_{n+1}^{(j)} - Au_{n+1}^{(j)})$$

with $\quad C(u_n) = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$ and $\quad \overline{u}_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}$.

Noise-free Case

Limiting ODE

$$\mathrm{d}u^{(j)} = C(u)A^*\Gamma^{-1}A(u^\dagger - u^{(j)})\,\mathrm{d}t\,,$$

or equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}u^{(j)} = -C(u)D_u\Phi(u^{(j)}; y)$$

with potential $\Phi(u; y) = \frac{1}{2}\|\Gamma^{-\frac{1}{2}}(y - Au)\|^2$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \dots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^J$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0, \infty); \mathcal{X}_0)$ for $j = 1, \dots, J$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \dots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^{J}$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0,\infty);\mathcal{X}_0)$ for $j = 1, \dots, J$.

Sketch of Proof

Quantities

$$e^{(j)} = u^{(j)} - \overline{u}\,, \qquad r^{(j)} = u^{(j)} - u^\dagger\,,$$

$$E_{lj} = \langle Ae^{(l)}, Ae^{(j)}\rangle_\Gamma\,, \qquad R_{lj} = \langle Ar^{(l)}, Ar^{(j)}\rangle_\Gamma\,, \qquad F_{lj} = \langle Ar^{(l)}, Ae^{(j)}\rangle_\Gamma\,.$$

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \dots, J$ and define $\mathcal{X}_0$ to be the linear span of the $\{u^{(j)}(0)\}_{j=1}^J$.

Then, the limiting ODE has a unique solution $u^{(j)}(\cdot) \in C([0,\infty); \mathcal{X}_0)$ for $j = 1, \dots, J$.

Sketch of Proof

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{(j)} = -\frac{1}{J} \sum_{k=1}^J E_{jk} d^{(k)}, \quad \frac{\mathrm{d}}{\mathrm{d}t} r^{(j)} = -\frac{1}{J} \sum_{k=1}^J F_{jk} r^{(k)}, \qquad j = 1, \dots, J$$

$$\frac{\mathrm{d}}{\mathrm{d}t} E = -\frac{2}{J} E^2, \qquad \frac{\mathrm{d}}{\mathrm{d}t} R = -\frac{2}{J} F F^\top, \qquad \frac{\mathrm{d}}{\mathrm{d}t} F = -\frac{2}{J} F E$$

Global existence of $\mathbf{E}$, $\mathbf{R}$ and $\mathbf{F}$ $\Rightarrow$ global existence of $\mathbf{r}$ and $\mathbf{e}$

Long-time Behaviour (Linear Case)

(b) Ensemble Collapse

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \dots, J$.

The solution of

$$\frac{\mathrm{d}}{\mathrm{d}t} E = -\frac{2}{J} E^2$$

with initial cond. $E(0) = X \Lambda_0 X^*$, $\Lambda_0 = \text{diag}\{\lambda_0^{(1)}, \dots, \lambda_0^{(J)}\}$, $X \in \mathbb{R}^{J \times J}$ orthogonal, is given by $E(t) = X \Lambda(t) X^*$.

$\Lambda(t)$ satisfies the following decoupled ODE

$$\frac{\mathrm{d}\lambda^{(j)}}{\mathrm{d}t} = -\frac{2}{J} (\lambda^{(j)})^2$$

with solution $\lambda^{(j)}(t) = \left(\frac{2}{J} t + \frac{1}{\lambda_0^{(j)}}\right)^{-1}$, if $\lambda_0^{(j)} \neq 0$, otherwise $\lambda^{(j)}(t) = 0$.

Long-time Behaviour (Linear Case)

(b) Ensemble Collapse

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A. Let $u^{(j)}(0) \in \mathcal{X}$ for $j = 1, \ldots, J$.

The solution of

$$\frac{\mathrm{d}}{\mathrm{d}t} E = -\frac{2}{J} E^2$$

with initial cond. $E(0) = X\Lambda_0 X^*$, $\Lambda_0 = \mathrm{diag}\{\lambda_0^{(1)}, \ldots, \lambda_0^{(J)}\}$, $X \in \mathbb{R}^{J\times J}$ orthogonal, is given by $E(t) = X\Lambda(t)X^*$.

$\Lambda(t)$ satisfies the following decoupled ODE

$$\frac{\mathrm{d}\lambda^{(j)}}{\mathrm{d}t} = -\frac{2}{J}(\lambda^{(j)})^2$$

with solution $\lambda^{(j)}(t) = \left(\frac{2}{J}t + \frac{1}{\lambda_0^{(j)}}\right)^{-1}$, if $\lambda_0^{(j)} \neq 0$, otherwise $\lambda^{(j)}(t) = 0$.

The rate of convergence of E and F is algebraic with a constant growing with larger ensemble size J.

Long-time Behaviour (Linear Case)

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A.Let $Y^\parallel$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^J$ and let $Y^\perp$ denote the orthogonal complement of $Y^\parallel$ in $\mathcal{Y}$ with respect to the inner product $\langle \cdot, \cdot \rangle_\Gamma$ and assume that the initial ensemble members are chosen so that $Y^\parallel$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

Then $Ar^{(j)}(t)$ may be decomposed uniquely as

$$Ar_\parallel^{(j)}(t) + Ar_\perp^{(j)}(t) \quad \text{with } Ar_\parallel^{(j)} \in Y^\parallel \text{ and } Ar_\perp^{(j)} \in Y^\perp.$$

Furthermore $Ar_\parallel^{(j)}(t) \to 0$ as $t \to \infty$ and $Ar_\perp^{(j)}(t) = Ar_\perp^{(j)}(0) = Ar_\perp^{(1)}$.

Long-time Behaviour (Linear Case)

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A.Let $Y^\parallel$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^J$ and let $Y^\perp$ denote the orthogonal complement of $Y^\parallel$ in $\mathcal{Y}$ with respect to the inner product $\langle \cdot, \cdot \rangle_\Gamma$ and assume that the initial ensemble members are chosen so that $Y^\parallel$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

Then $Ar^{(j)}(t)$ may be decomposed uniquely as

$$Ar_\parallel^{(j)}(t) + Ar_\perp^{(j)}(t) \quad \text{with } Ar_\parallel^{(j)} \in Y^\parallel \text{ and } Ar_\perp^{(j)} \in Y^\perp.$$

Furthermore $Ar_\parallel^{(j)}(t) \to 0$ as $t \to \infty$ and $Ar_\perp^{(j)}(t) = Ar_\perp^{(j)}(0) = Ar_\perp^{(1)}$.

Adaptive choice of the initial ensemble to ensure convergence of the residuals.

Long-time Behaviour (Linear Case)

Idea of Proof

Subspace property

$$Ae^{(j)}(t) = \sum_{k=1}^{J} \ell_{jk}(t) Ae^{(k)}(0)$$

where the matrix $L = \{\ell_{jk}\}$ is invertible.

Decomposition of the residual

$$Ar^{(j)}(t) = \sum_{k=1}^{J} \alpha_k Ae^{(k)}(t) + Ar_{\perp}^{(1)}$$

Convergence of the residuals

Boundedness of the coefficient vector

$$|\alpha(t)|^2 \leq \frac{\lambda_0^{(J)}}{\lambda_0^{\min}} |\alpha(0)|^2$$

gives convergence of the residuals.

Long-time Behaviour for Noisy Data (Linear Case)

Find the parameters u from (noisy) observations $y^\dagger$

$$y^\dagger = Au^\dagger + \eta^\dagger$$

Global Existence of Solutions ✓

Ensemble Collapse ✓

Convergence of Residuals $\rightarrow$ **convergence of the misfit**

Variants on EnKF

Variance Inflation

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -\big(\alpha C_0 + C(u)\big) D_u \Phi(u^{(j)}; y), \quad j = 1, \ldots, J\,,$$

where C_0 is a self-adjoint, strictly positive operator.

Localisation

Randomised Search

Variants on EnKF

Variance Inflation

Localisation

$$\rho : D \times D \to \mathbb{R} , \qquad \rho(x,y) = \exp(-|x-y|^r) ,$$

where $D \subset \mathbb{R}^d$ denotes the physical domain and $|\cdot|$ is a suitable norm in D, $r \in \mathbb{N}$.

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C^{\mathrm{loc}}(u) D_u \Phi(u^{(j)}; y), \quad j = 1, \dots, J ,$$

where $C^{\mathrm{loc}}(u)\phi(x) = \int_D \phi(y) k(x,y) \rho(x,y) \,\mathrm{d}y$ with k being the kernel of $C(u)$, $\phi \in \mathcal{X}$.

Randomised Search

Variants on EnKF

Variance Inflation

Localisation

Randomised Search

$$\mu_{n+1} = L_n P_n \mu_n \ .$$

where P_n is any Markov kernel which preserves μ_n.

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = \frac{1}{J}\sum_{k=1}^{J}\big\langle \mathcal{G}(u^{(k)}) - \overline{\mathcal{G}}, y - \mathcal{G}(u^{(j)})\big\rangle_{\Gamma}\big(u^{(k)} - \overline{u}\big)$$

$$-u^{(j)} - tC_0 D_u \Phi(u^{(j)}; y) + \sqrt{2C_0}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t} \ .$$

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0,\pi)\,,\ p = 0 \quad \text{in } \partial D\ ,$$

where

$A = \mathcal{O} \circ L^{-1}$ with $L = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + id$ and $D(L) = H^2(D) \cap H_0^1(D)$
$\mathcal{O} : X \mapsto \mathbb{R}^K$, equispaced observation points in D with spacing $\tau_N^{\mathcal{O}} = 2^{-N_K}$ at
$x_k = \frac{k}{2^{N_K}},\ k = 1, \ldots, 2^{N_K} - 1$, $o_k(\cdot) = \delta(\cdot - x_k)$ with $K = 2^{N_K} - 1$.

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0,\pi)\,,\ p = 0 \quad \text{in } \partial D\,.$$

The goal of computation is to recover the unknown data $u^\dagger$ from observations

$$y \;=\; \mathcal{O}L^{-1}u^\dagger + \eta = Au^\dagger + \eta\,.$$

Numerical Experiments (Linear Case)

1-dimensional elliptic equation

$$-\frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + p = u \quad \text{in } D := (0,\pi)\,,\ p = 0 \quad \text{in } \partial D\,.$$

The goal of computation is to recover the unknown data $u^\dagger$ from observations

$$y \;=\; \mathcal{O}L^{-1}u^\dagger + \eta = Au^\dagger + \eta\,.$$

Computational Setting

- Noisy case, $\Gamma = I$.
- $u \sim \mathcal{N}(0, C)$ with $C = \beta(A - id)^{-1}$ and with $\beta = 10$.
- Finite element method using continuous, piecewise linear ansatz functions on a uniform mesh with meshwidth $h = 2^{-8}$ (the spatial discretisation leads to a discretisation of u, i.e. $u \in \mathbb{R}^{2^8-1}$).
- The space $\mathcal{A} = \mathrm{span}\{u_0^{(j)}\}_{j=1}^{J}$ is chosen based on the KL expansion of $C = \beta(A - id)^{-1}$ (in red and green) and in an adaptive way minimising $Ar_\perp^{(j)}(t)$ (in blue) .

Numerical Experiments (Linear Case)

Underdetermined case, $\mathbf{K = 2^4 - 1, J = 5}$

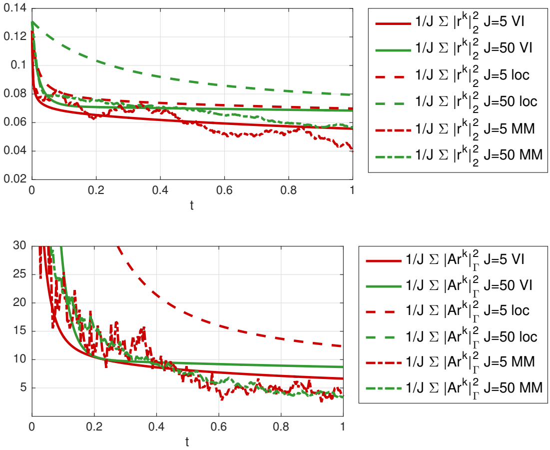


Figure: Quantities $|r|_2^2$, $|Ar|_\Gamma^2$ w.r. to time t, $J = 5$ (red) and $J = 50$ (green) for the discussed variants, $\beta = 10$, $\beta = 10$, $K = 2^4 - 1$, initial ensemble chosen based on KL expansion of $C = \beta(A - id)^{-1}$.

Numerical Experiments (Linear Case)

Underdetermined case, $\mathbf{K = 2^4 - 1}$, $\mathbf{J = 5}$

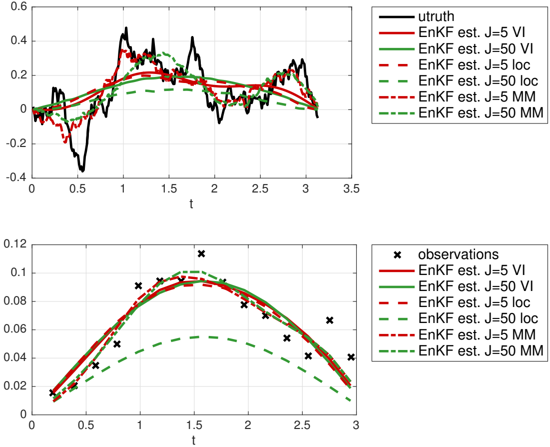


Figure: Comparison of the EnKF estimate with the truth and the observations, $J = 5$ (red) and $J = 50$ (green) for the discussed variants, $\beta = 10$, $K = 2^4 - 1$, initial ensemble chosen based on KL expansion of $C = \beta(A - id)^{-1}$.

Conclusions and Outlook

- Deriving the continuous time limit allows to determine the asymptotic behaviour of important quantities of the algorithm.
- The continuous approach offers the possibility to improve the performance of the method by choosing appropriate numerical discretisation schemes based on the properties of the solution.
- Generalisation of the results to noisy observational data, i.e. $Au^\dagger + \eta^\dagger$.
- Improving the performance of the algorithm by controlling the approximation quality of the subspace spanned by the ensemble.
- Analysis of EnKF variants
  - Variance inflation
  - Localisation
  - Iterative regularisation
  - Markov mixing

References

M. Bocquet, and P. Sakov 2014 An iterative ensemble Kalman smoother. *Q.J.R. Meteorol. Soc.* **140** 682.

O.G. Ernst, B. Sprungk, and H. Starkloff 2015 Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems. *http://arxiv.org/abs/1504.03529*.

G Evensen 2009 Data Assimilation: The Ensemble Kalman Filter *Springer*.

M A Iglesias, K J H Law and A M Stuart 2013 Ensemble Kalman methods for inverse problems *Inverse Problems* **29** 045001.

F Le Gland, V Monbet and V-D Tran 2011 Large sample asymptotics for the ensemble Kalman filter *Dan Crisan, Boris Rozovskii. The Oxford Handbook of Nonlinear Filtering, Oxford University Press*, 598-631.

D T B Kelly, K J H Law and A M Stuart 2014 Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time *Nonlinearity* **27** 2579-2603.

C Schillings and A M Stuart 2015 Analysis of the Ensemble Kalman Filter for Inverse Problems *(submitted)*.

A. S. Stordal, and A. H. Elsheikh 2015 Iterative ensemble smoothers in the annealed importance sampling framework *Advances in Water Resources* **86**.