Convergence Analysis of Ensemble Kalman Inversion

Claudia Schillings

MS 76 Nonlinear Filtering and Data Assimilation in Complex Dynamical Systems

SIAM UQ 2018

UNIVERSITÄT MANNHEIM

Joint work with Dirk Blömker (U Augsburg), Andrew Stuart (Caltech), Philipp Wacker (FAU), Simon Weissmann (U Mannheim)

Outline

1. Inverse Problem
2. Ensemble Kalman Inversion
3. Continuous Time Limit
4. Long-time Behaviour (Linear Case)
5. Summary

Bayesian Inverse Problem

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

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

- $u \in X$ parameter function
- $\mathcal{G} : X \to Y$ forward response operator
- y observations, here $y \in Y = \mathbb{R}^K$
- evaluation of $\mathcal{G}$ expensive
- prior $u \sim \mu_0$
- noise model $\eta \sim \mathcal{N}(0, \Gamma)$

Bayesian Inverse Problem

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

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

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

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

Algorithms

- **MCMC**
  - Dimension robust versions, multilevel strategies, improvements by local approximations
- **Approximations of the forward problem / posterior**
  - Structure exploiting approximations, best Gaussian approximations, transport maps
- Ensemble Kalman filter
- Randomized maximum likelihood , Approximate Bayesian computation, ...

Ensemble Kalman Inversion

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.

Ensemble Kalman Inversion

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

Ensemble Kalman Inversion

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

Assumption $\mathcal{G} = A \in \mathcal{L}(X, \mathbb{R}^K)$.

EnKF for Inverse Problems

Bridging Sequence

Introduction of an artificial discrete time dynamical system which maps the prior μ_0 into the posterior μ. The effective variance is amplified by $N = 1/h$ at each step, compensating for the redundant, repeated use of the data.

Ensemble of Interacting Particles

1. Use the ensemble $(u_n^{(j)})_{j=1}^J$ to define the empirical mean $\overline{u}_n = \frac{1}{J}\sum_{j=1}^J u_n^{(j)}$ and covariance $C(u_n) = \frac{1}{J-1}\sum_{j=1}^J (u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$.
2. Kalman update formulas

$$\overline{u}_{n+1} = \overline{u}_n + K_n(y - A\overline{u}_n) \qquad C(u_{n+1}) = C(u_n) - K_n A C(u_n)$$

with $K_n = C(u_n)A^*(AC(u_n)A^* + \frac{1}{h}\Gamma)^{-1}$.

3. Define $(u_{n+1}^{(j)})_{j=1}^J$ by a linear transformation D with $u_{n+1}^{(j)} = \sum_{i=1}^J u_n^{(i)} d_{ij}$ such that

$$\frac{1}{J}\sum_{j=1}^J u_{n+1}^{(j)} = \overline{u}_{n+1} \quad \text{and} \quad \frac{1}{J-1}\sum_{j=1}^J (u_{n+1}^{(j)} - \overline{u}_{n+1}) \otimes (u_{n+1}^{(j)} - \overline{u}_{n+1}) = C(u_{n+1})\,.$$

EnKF for Inverse Problems

Bridging Sequence

Introduction of an artificial discrete time dynamical system which maps the prior μ_0 into the posterior μ. The effective variance is amplified by $N = 1/h$ at each step, compensating for the redundant, repeated use of the data.

Ensemble of Interacting Particles

1. Use the ensemble $(u_n^{(j)})_{j=1}^J$ to define the empirical mean $\overline{u}_n = \frac{1}{J}\sum_{j=1}^J u_n^{(j)}$ and covariance $C(u_n) = \frac{1}{J-1}\sum_{j=1}^J (u_n^{(j)} - \overline{u}_n) \otimes (u_n^{(j)} - \overline{u}_n)$.
2. Kalman update formulas

$$\overline{u}_{n+1} = \overline{u}_n + K_n(y - A\overline{u}_n) \qquad C(u_{n+1}) = C(u_n) - K_n A C(u_n)$$

with $K_n = C(u_n)A^*(AC(u_n)A^* + \frac{1}{h}\Gamma)^{-1}$.

3. Define $(u_{n+1}^{(j)})_{j=1}^J$ by a linear transformation D with $u_{n+1}^{(j)} = \sum_{i=1}^J u_n^{(i)} d_{ij}$ such that

$$\frac{1}{J}\sum_{j=1}^J u_{n+1}^{(j)} = \overline{u}_{n+1} \quad \text{and} \quad \frac{1}{J-1}\sum_{j=1}^J (u_{n+1}^{(j)} - \overline{u}_{n+1}) \otimes (u_{n+1}^{(j)} - \overline{u}_{n+1}) = C(u_{n+1})\,.$$

EnKF for Inverse Problems

EnKF with perturbed observations

$$u_{n+1}^{(j)} = u_n^{(j)} + C(u_n)A^*(AC(u_n)A^* + \frac{1}{h}\Gamma)^{-1}(y_{n+1}^{(j)} - Au_n^{(j)})$$

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

Ensemble square root filter (ESFR)

$$u_{n+1}^{(j)} = \sum_{i=1}^{J} u_n^{(i)} d_{ij}$$

with $d_{ij} = w_i - \frac{1}{J} + s_{ij}$, where $C(u_n) = \frac{1}{J-1}P(u_n)P(u_n)^*$,
$S = (s_{ij})_{i,j} = (I + \frac{1}{J-1}(AP(u_n))^* h\Gamma^{-1}AP(u_n))^{-\frac{1}{2}}$ and
$w = \frac{1}{J}1 - \frac{1}{J-1}S^2(P(u_n))^*A^*h\Gamma^{-1}(A\overline{u}_n - y)$.

EnKF for Inverse Problems

EnKF with perturbed observations

$$u_{n+1}^{(j)} = u_n^{(j)} + C(u_n)A^*(AC(u_n)A^* + \frac{1}{h}\Gamma)^{-1}(y_{n+1}^{(j)} - Au_n^{(j)})$$

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

Ensemble square root filter (ESFR)

$$u_{n+1}^{(j)} = \sum_{i=1}^{J} u_n^{(i)} d_{ij}$$

with $d_{ij} = w_i - \frac{1}{J} + s_{ij}$, where $C(u_n) = \frac{1}{J-1}P(u_n)P(u_n)^*$,
$S = (s_{ij})_{i,j} = (I + \frac{1}{J-1}(AP(u_n))^* h\Gamma^{-1}AP(u_n))^{-\frac{1}{2}}$ and
$w = \frac{1}{J}1 - \frac{1}{J-1}S^2(P(u_n))^*A^*h\Gamma^{-1}(A\overline{u}_n - y)$.

Continuous Time Limit

EnKF with perturbed observations

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

where $W^{(1)}, \ldots, W^{(J)}$ are pairwise independent cylindrical Wiener processes and y denotes the noisy observational data.

Ensemble square root filter (ESFR)

Continuous Time Limit

EnKF with perturbed observations

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

where $W^{(1)}, \ldots, W^{(J)}$ are pairwise independent cylindrical Wiener processes and y denotes the noisy observational data.

Ensemble square root filter (ESFR)

Continuous Time Limit

EnKF with perturbed observations

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

or equivalently,

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

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

Ensemble square root filter (ESFR)

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

Continuous Time Limit

EnKF with perturbed observations

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

or equivalently,

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

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

Ensemble square root filter (ESFR)

$$\mathrm{d}\overline{u} = \frac{1}{J}YY^*A^*\Gamma^{-1}A(u^\dagger - \overline{u})\,\mathrm{d}t\,,$$

with $\mathrm{d}Y = -\frac{1}{2m-2}YY^*A^*\Gamma^{-1}AY$, e.g. [Bergemann, Reich 2009].

Continuous Time Limit

EnKF with perturbed observations

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

or equivalently,

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

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

Ensemble square root filter (ESFR)

$$\mathrm{d}u^{(j)} = -\frac{1}{2}C(u)(D_u\Phi(u^{(j)};y) + D_u\Phi(\overline{u};y))\,,$$

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

Continuous Time Limit

EnKF with perturbed observations

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

Ensemble square root filter (ESFR)

$$\frac{\mathrm{d}}{\mathrm{d}t}u^{(j)} = \frac{1}{J-1}\sum_{k=1}^{J}\langle Au^{(k)} - A\overline{u}, y - \frac{1}{2}Au^{(j)} - \frac{1}{2}A\overline{u}\rangle_{\Gamma}\big(u^{(k)} - \overline{u}\big)$$

- The ensemble parameter estimate lies in the linear span of the initial ensemble [8].
- In the linear case, the EnKF estimate converges in the limit $J \to \infty$ to the solution of the regularised least-squares problem [9, 14]. In the nonlinear setting, convergence to the mean-field Kalman filter is proven in [13].
- Ernst et al. [6] showed that the EnKF is not consistent with the Bayesian perspective in the nonlinear setting, but can be interpreted as a point estimator of the unknown parameters.
- In [4], multilevel strategies to enhance the perfomance of the EnKF are analysed.
- Kelly et al. [11, 12, 20, 19] presented an analysis of the long-time behavior and ergodicity of the ensemble Kalman filter with arbitrary ensemble size establishing time uniform bounds to control the filter divergence and ensuring in addition the existence of an invariant measure.
- Long term stability and accuracy is established for ensemble Kalman-Bucy filters applied to continuous-time filtering problems [5, 21].
- Higher order updates by polynomial chaos expansion can be found in [15].

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

Long-time Behaviour (Linear Case)

EnKF with perturbed observations / ESRF

(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 $j = 1, \ldots, J$ 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 $j = 1, \ldots, J$.

Long-time Behaviour (Linear Case)

EnKF with perturbed observations / ESRF

(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)

EnKF with perturbed observations / ESRF

(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 $j = 1, \ldots, J$ 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 $j = 1, \ldots, J$.

Sketch of Proof

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

$$\frac{\mathrm{d}}{\mathrm{d}t} E = -\frac{2}{J-1} E^2, \qquad \frac{\mathrm{d}}{\mathrm{d}t} R = -\frac{2}{J-1} F F^\top, \qquad \frac{\mathrm{d}}{\mathrm{d}t} F = -\frac{2}{J-1} F E$$

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

Long-time Behaviour (Linear Case)

EnKF with perturbed observations, ESRF

(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 $j = 1, \ldots, J$.

Then, the matrix valued quantity $E(t)$ converges to 0 for $t \to \infty$ and, indeed $\|E(t)\| = \mathcal{O}(Jt^{-1})$.

Long-time Behaviour (Linear Case)

EnKF with perturbed observations, ESRF

(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 $j = 1, \dots, J$.

Then, the matrix valued quantity $E(t)$ converges to 0 for $t \to \infty$ and, indeed $\|E(t)\| = \mathcal{O}(Jt^{-1})$.

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

Long-time Behaviour (Linear Case)

EnKF with perturbed observations

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and the forward operator A is one-to-one. 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)

EnKF with perturbed observations

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and the forward operator A is one-to-one. Let $Y^{\|}$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^{J}$ and let $Y^{\perp}$ denote the orthogonal complement of $Y^{\|}$ 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 $Y^{\|}$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

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

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

Furthermore $Ar_{\|}^{(j)}(t) \to 0$ 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)

EnKF with perturbed observations

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 (Linear Case)

ESRF

(c) Convergence of Residuals

Assume that y is the image of a truth $u^\dagger \in \mathcal{X}$ under A and the forward operator A is one-to-one. Let $Y^{\|}$ denote the linear span of the $\{Ae^{(j)}(0)\}_{j=1}^J$ and let $Y^{\perp}$ denote the orthogonal complement of $Y^{\|}$ 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 $Y^{\|}$ has the maximal dimension $\min\{J-1, \dim(\mathcal{Y})\}$.

Then $A\overline{r}(t)$ may be decomposed uniquely as

$$A\overline{r}_{\|}(t) + A\overline{r}_{\perp}(t) \quad \text{with } A\overline{r}_{\|} \in Y^{\|} \text{ and } A\overline{r}_{\perp} \in Y^{\perp}.$$

Furthermore $A\overline{r}_{\|}(t) \to 0$ as $t \to \infty$ and $A\overline{r}_{\perp}(t) = A\overline{r}_{\perp}(0)$.

Long-time Behaviour (Linear Case)

ESRF

(c) Convergence of Residuals

Then $A\overline{r}(t)$ may be decomposed uniquely as

$$A\overline{r}_{\|}(t) + A\overline{r}_{\perp}(t) \quad \text{with } A\overline{r}_{\|} \in Y^{\|} \text{ and } A\overline{r}_{\perp} \in Y^{\perp}.$$

Furthermore $A\overline{r}_{\|}(t) \to 0$ as $t \to \infty$ and $A\overline{r}_{\perp}(t) = A\overline{r}_{\perp}(0)$.

- Idea of the proof stays the same.
- Result implies the convergence of the residuals $Ar^{(i)}$.

Long-time Behaviour (Linear Case)

(a) Global Existence of Solutions

(b) Ensemble Collapse

(c) Convergence of Residuals

- **No Gaussian prior assumption (in the case of the EnKF with perturbed observations).**
- **Convergence result opens up the perspective to use the EnKF as a linear solver in case of a boundedly invertible forward operator.**
- **In the finite dimensional setting, the results can be used to characterise the parameter space informed by the data.**

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**

EnKF with perturbed observations

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

where $W^{(1)}, \ldots, W^{(J)}$ are pairwise independent cylindrical Wiener processes and y denotes the noisy observational data.

Simplified model:

$$\mathrm{d}u = -u^3 + u^2\,\mathrm{d}W$$

- EnKF seems to provide a stable numerical discretization, i.e.

$$u_{n+1} = u_n - h \cdot \frac{u_n^3}{1 + h \cdot u_n^2} + \frac{u_n^2}{1 + h \cdot u_n^2} \cdot \Delta W_{n+1}.$$

strongly approximates the simplified model problem.

EnKF with perturbed observations

$$\mathrm{d}u = -u^3 + u^2 \,\mathrm{d}W$$

Lipschitz-regularized version

$$\mathrm{d}v = -\frac{v^3}{1+\epsilon \cdot v^2}\,\mathrm{d}t + \frac{v^2}{1+\epsilon \cdot v^2}\,\mathrm{d}W.$$

and its Euler-Maruyama discretization

$$v_{n+1} = v_n + h \cdot \frac{-v_n^3}{1+\epsilon \cdot v_n^2} + \Delta W_n \cdot \frac{v_n^2}{1+\epsilon \cdot v_n^2}$$

and interpolation of v_n as

$$\overline{v}(t) = v_n + \int_{t_n}^{t} f(v_n)\,\mathrm{d}s + \int_{t_n}^{t} \sigma(v_n)\,\mathrm{d}W.$$

EnKF with perturbed observations

For any $0 < \alpha < 2$ and $0 < \eta < 1$,

$$\lim_{h=\epsilon\to 0} h^{-\frac{1}{2}\cdot\frac{3-\alpha}{3+25/3\cdot\alpha}} \cdot \Big(\sup_{t\in[0,T]} \mathbb{E}|\overline{v}(t) - u(t)|^{\alpha} \Big)^{\frac{1}{\alpha}} = 0$$

$$\lim_{h=\epsilon\to 0} h^{-\frac{1}{2}\cdot\frac{1-\eta}{1+3\eta}} \cdot \Big(\mathbb{E} \sup_{t\in[0,T]} |\overline{v}(t) - u(t)|^{\eta} \Big)^{\frac{1}{\eta}} = 0$$

- Generalization of the strong convergence results to the linear setting.
- Well-posedness and accuracy results for the SDE model (linear forward problem).
- Accuracy results of the EnkF estimate w.r. to the conditional mean.

D Blömker, C Schillings and P Wacker 2017 A strongly convergent numerical scheme from ensemble Kalman inversion arXiv:1703.06767.

EnKF with perturbed observations

For any $0 < \alpha < 2$ and $0 < \eta < 1$,

$$\lim_{h=\epsilon\to 0} h^{-\frac{1}{2}\cdot\frac{3-\alpha}{3+25/3\cdot\alpha}} \cdot \left(\sup_{t\in[0,T]} \mathbb{E}|\overline{v}(t) - u(t)|^{\alpha} \right)^{\frac{1}{\alpha}} = 0$$

$$\lim_{h=\epsilon\to 0} h^{-\frac{1}{2}\cdot\frac{1-\eta}{1+3\eta}} \cdot \left(\mathbb{E} \sup_{t\in[0,T]} |\overline{v}(t) - u(t)|^{\eta} \right)^{\frac{1}{\eta}} = 0$$

- Generalization of the strong convergence results to the linear setting.
- Well-posedness and accuracy results for the SDE model (linear forward problem).
- Accuracy results of the EnkF estimate w.r. to the conditional mean.

D Blömker, C Schillings and P Wacker 2017 A strongly convergent numerical scheme from ensemble Kalman inversion arXiv:1703.06767.

Conclusions and Outlook

- Deriving the continuous time limit allows to determine the asymptotic behaviour of important quantities of the algorithm.
- Analysis of the perturbed observation EnKF and Kalman square root filter for inverse problems .
- Generalisation of the results to noisy observational data, i.e. $Au^\dagger + \eta^\dagger$.
- Strong convergence results of the EnKF discretization.
- Analysis of EnKF variants
  - Variance inflation
  - Localization
  - Iterative regularization
  - Markov mixing

References I

K Bergemann and S Reich 2010 *A localization technique for ensemble Kalman filters Q.J.R. Meteorol. Soc.* **136** pp. 701–707.

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

D Blömker, C Schillings and P Wacker 2017 *A strongly convergent numerical scheme from ensemble Kalman inversion arXiv:1703.06767*.

A Chernov, H Hoel, K Law, F Nobile and R Tempone 2016 *Multilevel ensemble Kalman filtering for spatially extended models* arXiv:1710.07282.

P Del Moral and J Tugaut 2016 *On the stability and the uniform propagation of chaos properties of ensemble Kalman-Bucy filters*, arXiv:1605.09329.

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.

References II

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.

D. Kelly, K. Law, and A. Stuart, *Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time*, Nonlinearity, 27 (2014), p. 2579.

D. Kelly, A. J. Majda, and X. T. Tong, *Concrete ensemble Kalman filters with rigorous catastrophic filter divergence*, Proceedings of the National Academy of Sciences, 112 (2015), pp. 10589–10594.

K. Law, H. Tembine, and R. Tempone, *Deterministic mean-field ensemble Kalman filter*, SIAM J. Sci. Comput, 38 (2016), pp. A1251–A1279.

J. Mandel, L. Cobb, and J. D. Beezley, *On the convergence of the ensemble Kalman filter*, Applications of Mathematics, 56 (2011), pp. 533–541.

B. Rosic, A. Litvinenko, O. Pajonk, H.G. Matthies, *Direct Bayesian Update of Polynomial Chaos Representations*, Preprint TU Braunschweig, 2017.

References III

- C Schillings and A M Stuart 2017 *Analysis of the Ensemble Kalman Filter for Inverse Problems SIAM J Numerical Analysis* 55(3), 1264-1290.
- C Schillings and A M Stuart 2017 *Convergence analysis of ensemble Kalman inversion: the linear, noisy case Applicable Analysis*.
- A S Stordal, and A H Elsheikh 2015 *Iterative ensemble smoothers in the annealed importance sampling framework Advances in Water Resources* **86**.
- X. T. Tong, A. J. Majda, and D. Kelly 2015 *Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation*, arXiv:1507.08319.
- X. T. Tong, A. J. Majda, and D. Kelly 2016 *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29, p. 657.
- J. de Wiljes, S. Reich, and W. Stannat 2016 *Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise*, arXiv:1612.06065.

Variants on EnKF

Variance Inflation

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

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.

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = \frac{1}{J}\sum_{k=1}^{J}\big\langle \mathcal{G}(u^{(k)}) - \overline{\mathcal{G}}, y - \mathcal{G}(u^{(j)})\big\rangle_{\Gamma}\big(u^{(k)} - \overline{u}\big)$$
$$-u^{(j)} - C_0 D_u \Phi(u^{(j)}; y) + \sqrt{2C_0}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t} \ .$$

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

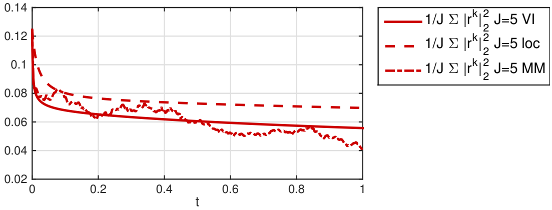
$$y \;=\; \mathcal{O}L^{-1}u^\dagger + \eta = Au^\dagger + \eta\,.$$

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 $C = \beta(A - id)^{-1}$.

Numerical Experiments (Linear Case)

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



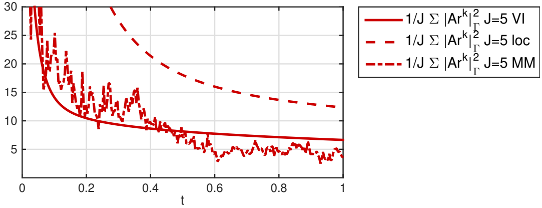


Figure: Quantities $|r|_2^2$, $|Ar|_\Gamma^2$ w.r.t. to time t, $J = 5$ (red) for the discussed variants, $\beta = 10$, $\beta = 10$, $K = 2^4 - 1$, initial ensemble chosen based on KL expansion of $C = \beta(A - id)^{-1}$.

Numerical Experiments (Linear Case)

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

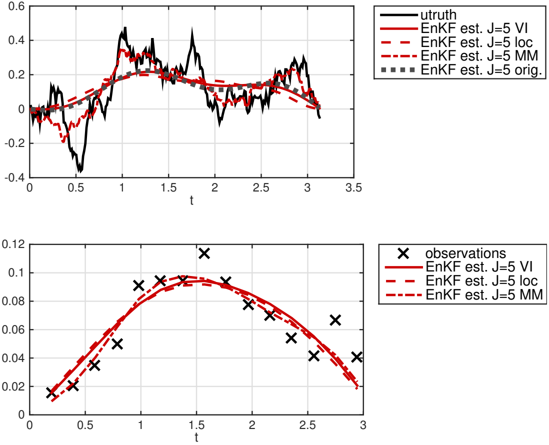


Figure: Comparison of the EnKF estimate with the truth and the observations, $J = 5$ (red) for the discussed variants, $\beta = 10$, $K = 2^4 - 1$, initial ensemble chosen based on KL expansion of $C = \beta(A - id)^{-1}$.

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

Long-time Behaviour for Noisy Data (Linear Case)

Convergence of Residuals

Assume that $y^\dagger$ is the perturbed image of a truth $u^\dagger \in X$, i.e. $y^\dagger = Au^\dagger + \eta^\dagger$ and the forward operator A is one-to-one. Let $Y^\parallel$ denote the linear span of the $\{Ad^{(j)}(0)\}_{j=1}^J$ and let $Y^\perp$ denote the orthogonal complement of $Y^\parallel$ in Y and assume that the initial ensemble members are chosen so that $Y^\parallel$ has the maximal dimension $\min\{J-1, \dim(Y)\}$.

Then $\vartheta^{(j)}(t) := Au^{(j)} - y^\dagger$ may be decomposed uniquely as

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

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

Furthermore, if $\langle \eta^\dagger, Ae^{(k)} \rangle \leq \langle Ar^{(k)}, Ae^{(k)} \rangle$, the residual is monotonically decreasing. The rate of convergence of the component of the mapped residual, which belongs to $Y^\parallel$, can be arbitrarily slow.

Numerical Experiments (Linear Case, Noisy Observations)

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

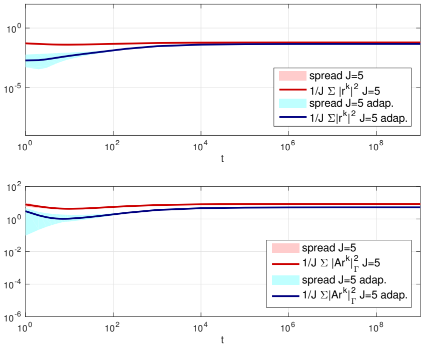


Figure: Quantities $|r|_2^2$, $|Ar|_\Gamma^2$ w.r. to t, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$.

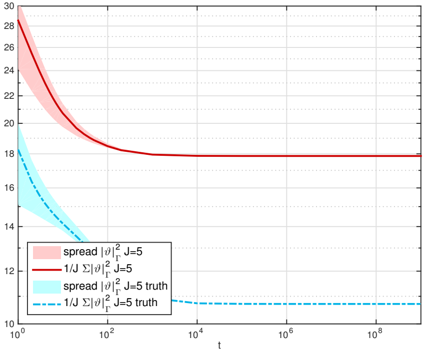


Figure: Misfit $|\vartheta|_2^2$ w.r. to time t, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$.

Numerical Experiments (Linear Case, Noisy Observations)

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

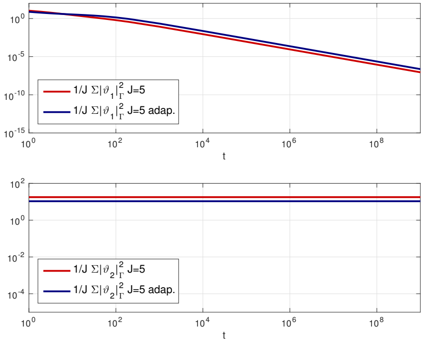


Figure: Quantities $|\vartheta_{\|}^{(j)}|_\Gamma^2$, $|\vartheta_{\perp}^{(j)}|_\Gamma^2$ w.r. to time t, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$.

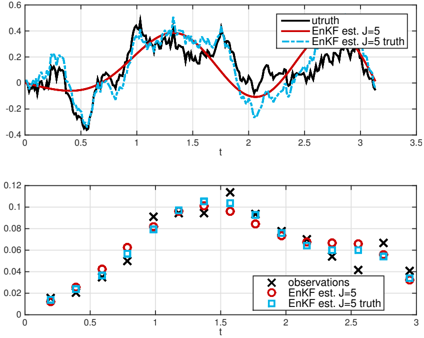


Figure: Comparison of the EnKF estimate with the truth and the observations, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$.

Numerical Experiments (Linear Case, Noisy Observations)

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

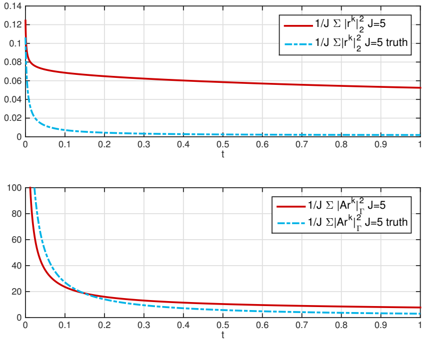


Figure: Quantities $|r|_2^2$, $|Ar|_\Gamma^2$ w.r. to time t, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$, Bayesian stopping rule.

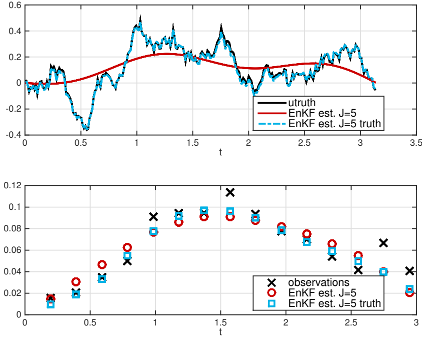


Figure: Comparison of the EnKF estimate with the truth and the observations, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $\beta = 10$, $K = 2^4 - 1$, Bayesian stopping rule.

Numerical Experiments (Linear Case, Noisy Observations)

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

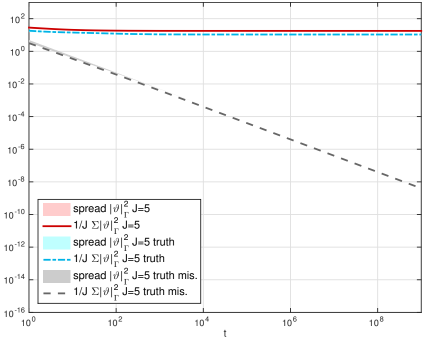


Figure: Misfit $|\vartheta|_2^2$ w.r. to time t, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $J = 5$ adaptively chosen w.r. to misfit (grey), $\beta = 10$, $K = 2^4 - 1$.

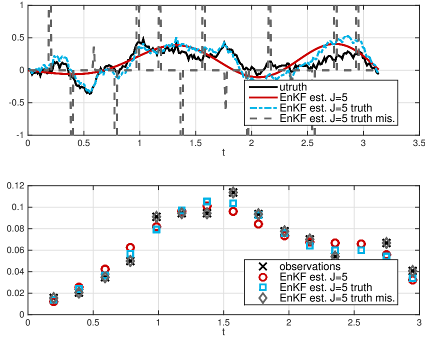


Figure: Comparison of the EnKF estimate, $J = 5$ based on KL expansion of $C = \beta(A - id)^{-1}$ (red), $J = 5$ adaptively chosen (blue), $J = 5$ adaptively chosen w.r. to misfit (grey), $\beta = 10$, $K = 2^4 - 1$.