# Analysis of the Ensemble Kalman Filter for Inverse Problems

Matthieu Bulté

17.01.2019

## Inverse problems

Given a continuous map $\mathcal{G} : \mathcal{X} \to \mathcal{Y}$ between two separable Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, we would like to identify the *parameter* $u \in \mathcal{X}$ satisfying

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

where $y \in \mathcal{Y}$ is the *observation* and $\eta \in \mathcal{Y}$ is the *observational noise*. The operator $\Gamma$ here is a symmetrical positive-definite operator describing the correlation structure of the observational noise. For simplicity, we limit ourselves to the case where $\mathcal{Y} = \mathbb{R}^K$ for $K \in \mathbb{N}$.

A key element in understanding and solving inverse problems is the *least squares* functional $\Phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, measuring the model-data misfit, given by

$$\Phi(u; y) = \frac{1}{2} \left\| \Gamma^{-1/2}(y - \mathcal{G}(u)) \right\|_{\mathcal{Y}}^2, \tag{LS}$$

where $\Gamma$ is a positive semi-definite operator describing the covariance structure of the noise $\eta$.

## Regularization

In general, minimizing the model-data misfit (LS) is *ill-posed*: that the solution might not exist or not be unique. To avoid this, we employ *regularization* techniques to improve the solvability of the system. We consider the following two approaches

- **Optimization approach.** Replace the minimization problem (LS) by

  $$u^* = \arg\min_{u \in \mathcal{X}} \Phi(u; y) + \|u - u_0\|_{\mathcal{C}}^2,$$

  for some $u_0 \in \mathcal{X}$ and covariance operator $\mathcal{C}$.

- **Bayesian approach.** Consider $u \sim \mu_0$ and $\eta \sim \mathrm{N}(0, \Gamma)$ in (IP) to obtain

  $$u^* = u | y \sim \mu,$$

  where $\mu(\mathrm{d}u) \propto \exp(-\Phi(u; y))\mu_0(\mathrm{d}u)$.

A comparison and motivation for the approaches is done by T.J. Sullivan [5]. In this talk, we focus on the Bayesian approach to regularization of inverse problems.

## Gaussian Process

We wish to recover a function $u^\dagger \in H_0^1([0,1])$ from noisy evaluation $y_1, \dots, y_{K-1}$ of $u^\dagger$ at nodes $x_1, \dots, x_{K-1}$ where $K \in \mathbb{N}$ and $x_k = k/K$. This translates to the following inverse problem

$$y = \mathcal{O}(u^\dagger) + \eta,$$

where $y \in \mathbb{R}^K$ and $\mathcal{O}(u) = (u(x_1), \dots, u(x_{K-1}))$. We regularize the problem using the Bayesian approach: we let $\eta \sim \mathrm{N}(0, \gamma 1)$ with $\gamma = 0.01$ and $u \sim \mathrm{N}(0, (-\Delta)^{-1})$.

This example will be used along the presentation to illustrate different concepts and results.

## EnKF for inverse problems

The EnKF algorithm approximates partially, noisily observed dynamical systems by evolving an ensemble of $J \in \mathbb{N}$ particles $\{u_n^{(j)}\}_{j=1}^J$ for each time step $n \in \mathbb{N}$. We construct an artificial dynamic by discretizing the interpolation between $\mu_0$ and $\mu$: given a step size $h \in \mathbb{R}$, we define

$$\mu_n \propto \exp(-nh\Phi(u;y))\mu_0(\mathrm{d}u) \text{ for } n = 0, \dots, h^{-1}.$$

Applying the EnKF to this dynamic gives the following update rule

$$u_{n+1}^{(j)} = u_n^{(j)} + C^{\mathrm{up}}(u_n)[C^{\mathrm{pp}}(u_n) + h^{-1}\Gamma]^{-1}(y_{n+1}^{(j)} - G(u_n^{(j)})), \qquad \text{(ENKF-IP)}$$

with

$$\begin{aligned} y_{n+1}^{(j)} &= y + \xi_{n+1}^{(j)}, & C^{\mathrm{up}} &= \widehat{\mathrm{cov}}(u, \mathcal{G}(u)), \\ \{\xi_n^{(j)}\}_{n,j} &\sim \mathrm{N}(0, \Sigma) \text{ i.i.d }, & C^{\mathrm{pp}} &= \widehat{\mathrm{cov}}(\mathcal{G}(u), \mathcal{G}(u)). \end{aligned}$$

The initial ensemble $\{u_0^{(j)}\}_{j=1}^J$ can be constructed from $J$ i.i.d. draws of the distribution $\mu_0$. However, we will see later that alternative constructions of the initial ensemble might be preferable. An analysis of the large ensemble limit can be found in [2, 3, 1].

## Continuous-time limit

Instead of considering how well the ensemble constructed by the EnKF approximates $u|y$ in the large ensemble limit, we study the algorithm for $h \to 0$. By rewriting the update rule (ENKF-IP) and taking $h \to 0$, one recognizes the Euler-Maruyama discretization of the following coupled system of SDEs

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

where $\{W^{(j)}\}_{j=1}^J$ are independent Brownian motions. Assuming that $\mathcal{G} = A$ is linear and that the observations are noise-free, that is $\Sigma = 0$, the coupled system of SDEs become the following system of ODEs

$$\begin{aligned}\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} &= \hat{\mathrm{cov}}(u, Au)\Gamma^{-1}(y - Au^{(j)}) \\ &= \frac{1}{J}\sum_{k=1}^{J}\langle A(u^{(k)} - \overline{u}), y - Au^{(j)}\rangle_\Gamma (u^{(k)} - \overline{u}) \\ &= A^T\Gamma^{-1}(y - Au^{(j)})\hat{\mathrm{cov}}(u, u) \\ &= -\hat{\mathrm{cov}}(u, u)\mathrm{D}_u\Phi(u^{(j)}; y).\end{aligned} \tag{ODE}$$

Hence, each particle follows a preconditioned gradient flow for $\Phi(\cdot; y)$. This draws a link between a link between the optimization and Bayesian approaches to inverse problems.

# Analysis

While the construction of the continuous model only justifies considering $t$ up to 1, we now study properties of the ensemble for $t \to \infty$. To simplify the analysis, we limit ourselves to the linear noise-free case, where $y = Au^\dagger$. We start by introducing the following useful quantities:

$$e^{(j)} = u^{(j)} - \overline{u}, \qquad r^{(j)} = u^{(j)} - u^\dagger, \qquad E_{ij} = \langle Ae^{(i)}, Ae^{(j)}\rangle_\Gamma.$$

We can then cite the following results from [4] about the behaviour of the system (ODE).

**Theorem 1.** *Let $\mathcal{X}_0$ be the linear span of $\{u^{(j)}(0)\}_{j=1}^J$, then the system of equations (ODE) possesses a unique ensemble of solutions $\{u^{(j)}\}_{j=1}^J$ with $u^{(j)} \in C([0, \infty); \mathcal{X}_0)$ for every $j = 1, \ldots, J$.*

**Theorem 2.** *The matrix $E$ converges to 0 for $t \to \infty$ with $\|E\|_2 = \mathcal{O}(Jt^{-1/2})$.*

**Theorem 3.** *Let $\mathcal{Y}_0$ denote the linear span of $\{Au^{(j)}(0)\}_{j=1}^J$, and $\mathcal{Y}_\perp$ denote the orthogonal completement of $\mathcal{Y}_0$ w.r.t. $\langle\cdot,\cdot\rangle_\Gamma$. Assume further that the initial ensemble has been chosen such that $dim(\mathcal{Y}_\perp) \le \min(dim(\mathcal{Y}), J - 1)$. Then we can decompose $Ar^{(j)} = Ar_0^{(j)} + Ar_\perp^{(j)}$ with $Ar_0^{(j)} \in \mathcal{Y}_0$ and $Ar_\perp^{(j)} \in \mathcal{Y}_\perp$ and obtain $Ar_0^{(j)} \to 0$ as $t \to \infty$.*

# Numerical results

We now wish to observe the convergence properties of the continuous-time limite of the EnKF algorithm. Since the Gaussian process system is not finite dimensional, a discretization is required. We base this discretization on a spectral Karhunen–Loève expansion of the Gaussian process which is then cut to a finite number of terms, $N = 100$.

In the algorothm, the initial ensemble is constructed based on the first $J$ eigenpairs $(\lambda_j, \psi_j)$ of the KL expansion. Thus $u^{(j)} = \sqrt{\lambda_j}\psi_j z_j$ where $\{z_j\}_{j=1}^J$ are i.i.d. realizations of $\mathrm{N}(0, 1)$ and

$$\lambda_j = (\pi j)^{-2} \qquad \psi_j(x) = \sqrt{2}\sin(\pi n x).$$

Results of the experiments are to be found in figures  and .

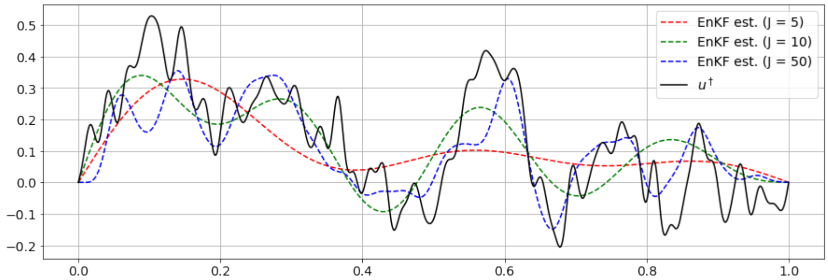


Figure 1: Comparison of the truth to several EnKF estimates for $J = 5, 10, 50$

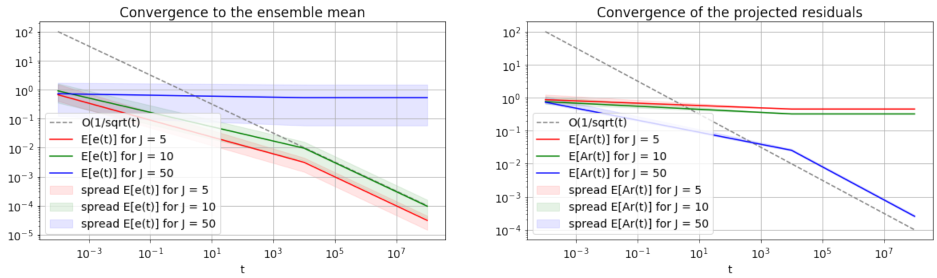


Figure 2: Convergence of EnKF the estimates for $J = 5, 10, 50$

# References

[1] O. G. Ernst, B. Sprungk, and H.-J. Starkloff. Analysis of the ensemble and polynomial chaos kalman filters in bayesian inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):823–851, 2015.

[2] M. Goldstein and D. Wooff. *Bayes linear statistics: Theory and methods*, volume 716. John Wiley & Sons, 2007.

[3] K. J. Law, H. Tembine, and R. Tempone. Deterministic mean-field ensemble kalman filtering. *SIAM Journal on Scientific Computing*, 38(3):A1251–A1279, 2016.

[4] C. Schillings and A. M. Stuart. Analysis of the ensemble kalman filter for inverse problems. *SIAM Journal on Numerical Analysis*, 55(3):1264–1290, 2017.

[5] T. J. Sullivan. *Introduction to uncertainty quantification*, volume 63. Springer, 2015.