A Regularizing Ensemble Kalman Method for PDE-constrained Inverse Problems

Marco Iglesias

School of Mathematical Sciences, University of Nottingham, U.K.

AMCS Seminar,
Center for UQ, CEMSE, KAUST, KSA
September 8th, 2015
1 Introduction

2 Numerical Investigation of the Scheme

3 Applications
Outline

1 Introduction

2 Numerical Investigation of the Scheme

3 Applications
General Aim

Inferring/estimating functions which are inputs for a PDE model, given measurements/observations form the output.

PDE-constrained applications

*Porous media flow*  
*Electrical Impedance Tomography*

M. A. Iglesias

A regularizing ensemble Kalman method for PDE-constrained inverse problems.


M. A. Iglesias

Iterative regularization for ensemble data assimilation in reservoir modeling

*Computational Geosciences, (2015) 19:177-212*
Abstract Setting

Let $G : X \rightarrow \mathbb{R}^J$.

Forward Model

Given $u \in X$ compute

$$y = G(u).$$

Let $\eta \in \mathbb{R}^J$ be a realization of an observational noise.

Inverse Problem

Given $y \in \mathbb{R}^J$ find $u \in X$:

$$y = G(u) + \eta.$$
Groundwater Flow: Forward and Inverse Problem

Forward groundwater flow model:

**Forward Problem: Darcy flow**

\[
\begin{align*}
- \nabla \cdot \kappa \nabla p &= f \quad \text{in } D, \\
- \kappa \nabla p \cdot n &= B_N \quad \text{in } \Gamma_N, \\
p &= B_D \quad \text{in } \Gamma_D
\end{align*}
\]

where \( \partial D = \Gamma_N \cup \Gamma_D \).

\[
\begin{align*}
u &= \log(\kappa(x)) \in X \equiv L^\infty(D) \rightarrow \mathcal{G}(u) = \{p(x_i)\}_{i=1}^J \in \mathbb{R}^J
\end{align*}
\]

**Inverse Problem**

Given \( y \in \mathbb{R}^J \) find \( u \in X \):

\[
y = \mathcal{G}(u) + \eta.
\]
Bayesian Inversion

**Prior**

Probabilistic information about $u$ before data is collected:

$$\mu_0(u) = \mathbb{P}(u)$$

**Likelihood**

Since $y = G(u) + \eta$, if $\eta \sim N(0, \Gamma)$, then $\mathbb{P}(y|u) = N(G(u), \Gamma)$. Then (Γ-weighted) model-data misfit $\Phi$ is the negative log-likelihood:

$$\Phi(u; y) = \frac{1}{2} \left\| \Gamma^{-1/2} (y - G(u)) \right\|^2$$

**Posterior**

Probabilistic information about $u$ after data is collected:

$$\mu^y(u) = \mathbb{P}(u|y).$$

$$\frac{\mu^y(u)}{\mu_0(u)} \propto \exp \left( - \Phi(u; y) \right)$$

(University of Nottingham)
Bayesian Inversion

**Posterior**

Probabilistic information about $u$ after data is collected:

$$
\mu^y(u) = \mathbb{P}(u|y).
$$

$$
\frac{\mu^y(u)}{\mu_0(u)} \propto \exp\left( - \Phi(u; y) \right)
$$

**Challenge**

To explore the probability measure $\mu^y$.

- $\mathcal{G}$ is highly nonlinear; $\mu^y$ cannot be characterized with a few parameters.

- The problem is high dimensional ($X$ is discretized with $10^6$-$10^9$ cells).

- Standard sampling methods for Bayesian inference do not work.

- Infinite-dimensional Bayesian framework [Stuart, 2010]; MCMC method for functions (pcn-MCMC) [Cotter, et-al, 2013].

- Well-known for continuous $\mathcal{G}$.
**The Classical (deterministic) formulation of the Inverse Problem**

Given data \( y \in Y \) find

\[
u = \arg \min_{u \in X} \| \Gamma^{-1/2} (y - G(u)) \|^2 \rightarrow \min
\]

For most PDE-constrained applications \( G : X \to \mathbb{R}^M \) is compact (unless \( X \) is finite dimensional)

**Lack of continuity (lack of stability) with respect to the data**

We can construct a sequence \( u_n \in X \) such that

\[
 u_n \not\rightarrow u \quad \text{but} \quad G(u_n) \rightarrow G(u)
\]

If we want to compute the minimizer above with standard optimization we may observe semiconvergence behavior \([Kirsch, 1996]\)
Regularization Approaches (for nonlinear operators)

- Regularize-then-compute (e.g. Tikhonov, TSVD)
- Compute while regularizing (Iterative Regularization) [Kaltenbacher, 2010]
  - regularizing Levenberg-Marquardt
  - Landweber iteration
  - truncated Newton-CG
  - iterative regularized Gauss-Newton method

Introduce noise level

\[ ||\Gamma^{-1/2}(y - G(u^\dagger))|| \leq \eta \]

Regularization

Construct an approximation \( u^n \) that is stable, i.e. such that

\[ u^n \rightarrow u \quad \text{as} \quad \eta \rightarrow 0 \]

where

\[ G(u) = G(u^\dagger) \]
Merging the Bayesian and the Classical approach

Consider \( \mu_0(u) = \mathbb{P}(u) = \mathcal{N}(0, C) \) the prior on \( u \) and
\[
    y = G(u) + \xi, \quad \xi \sim \mathcal{N}(0, \Gamma)
\]

The Bayesian Inverse Problem

Characterize the posterior \( \mu^y(u) = \mathbb{P}(u|y) \):
\[
    \frac{\mu^y(u)}{\mu_0(u)} \propto \exp \left( - \frac{1}{2} ||\Gamma^{-1/2}(y - G(u))||^2 \right) \quad y^{(j)} = y + \eta^{(j)} \sim \mathcal{N}(0, \Gamma)
\]

Ensemble Approximating the Bayesian posterior \( \mu^y(u) = \mathbb{P}(u|y) \)

Randomizing least-squares, e.g.
\[
    \frac{1}{2} ||\Gamma^{-1/2}(y^{(j)} - G(u))||^2 \rightarrow \min \quad y^{(j)} = y + \eta^{(j)} \sim \mathcal{N}(0, \Gamma)
\]

or, for example, Randomized Maximum Likelihood
\[
    \frac{1}{2} ||\Gamma^{-1/2}(y^{(j)} - G(u^{(j)}))||^2 + ||C^{-1/2}(u - u^{(j)})||_X^2 \rightarrow \min \quad u^{(j)} \sim \mathcal{N}(0, C)
\]
Overview of this work

**Classical (deterministic) Inversion**

- **Least-squares**
  \[ \Phi(u; y) = \frac{1}{2} \| \Gamma^{-1/2}(y - G(u)) \|^2 \]

**Bayesian Inversion**

- **Characterize the posterior**
  \[ \frac{d\mu^y}{d\mu_0}(u) \propto \exp \left( - \Phi(u; y) \right) \]

**Iterative Regularization**

- (e.g. regularizing LM, landweber iteration)

**Ensemble Kalman-based methods**

- \[ u^{(j,a)} = u^{(j,f)} + K(y^{(j)} - G(u^{(j,f)})) \]

---

M. A. Iglesias

A regularizing ensemble Kalman method for PDE-constrained inverse problems.


M. A. Iglesias

Iterative regularization for ensemble data assimilation in reservoir modeling

*Computational Geosciences, (2015) 19:177-212*
Bayesian Inverse Problem

Given a prior $\mu_0(u)$ of on $u$ and data

$$y = G(u) + \eta \quad \eta \sim N(0, \Gamma)$$

find $\mu(u) = P(u|y)$.

$$\mu(u) \propto \mu_0(u) \exp \left\{-\frac{1}{2} ||\Gamma^{-1/2}(y - G(u))||^2 \right\}$$

Define

$$z = \begin{pmatrix} u \\ G(u) \end{pmatrix}, \quad y = Hz + \eta, \quad H = (0, I)$$

Alternative Bayesian Inverse Problem

Given a prior on $z$, $\mu_0(z)$ and data $y$, find $\mu(z) = P(z|y)$

$$\frac{\mu(z)}{\mu_0(z)} \propto \exp \left\{-\frac{1}{2} ||\Gamma^{-1/2}(y - G(u))||^2 \right\}$$
Bayesian Inverse Problem

Given a prior on $z$, $\mu_0(z)$ and data $y$, find $\mu(z) = \mathbb{P}(z|y)$

$$
\mu(u) \propto \mu_0(u) \exp \left\{ -\frac{1}{2} \| \Gamma^{-1/2} (y - G(u)) \|^2 \right\}
$$

Construct an initial ensemble

$$
z_0^{(j,f)} = \begin{pmatrix} u_0^{(j)} \\ G(u_0^{(j)}) \end{pmatrix}, \quad \{u_0^{(j)}\}^N_{j=1} \sim \mu_0
$$

Compute mean and covariance

$$
\bar{z}^f = \frac{1}{N_e} \sum_{j=1}^{N_e} z^{(j,f)} \quad C^f = \frac{1}{(N_e - 1)} \sum_{j=1}^{N_e} z^{(j,f)}(z^{(j,f)})^T - \bar{z}^f(\bar{z}^f)^T
$$

Gaussian Approximation: $\mu_0(z) = \mathcal{N}(\bar{z}^f, C^f)$,
Since $\mu_0(z) = N(\bar{z}^f, C^f)$, then

$$
\mu(u) \propto \mu_0(u) \exp \left\{ -\frac{1}{2} ||\Gamma^{-1/2}(y - G(u))||^2 \right\} = N(\bar{z}^a, C^a)
$$

where

$$
\bar{z}^{(a)} = \bar{z}^{(f)} + C^f H^T \left( H C^f H^T + \Gamma \right)^{-1} (y - H \bar{z}^{(f)})
$$

$$
C^a = (I - K) C^f
$$

Updated each ensemble according to

$$
z^{(j,a)} = z^{(j,f)} + C^f H^T \left( H C^f H^T + \Gamma \right)^{-1} (y^{(j)} - H z^{(j,f)})
$$

with

$$
y^{(j)} = y + \eta^{(j)}, \quad \eta^{(j)} \sim N(0, \Gamma)
$$
Updated each ensemble according to

\[ z^{(j,a)} = z^{(j,f)} + C^f H^T \left( HC^f H^T + \Gamma \right)^{-1} (y^{(j)} - Hz^{(j,f)}) \]

Claim 1: \( \{z^{(j,a)}\}_{j=1}^{N_e} \approx P(z|y) \).

Recall that \( z = \begin{pmatrix} u \\ G(u) \end{pmatrix} \). Then,

\[ u^{(j)} = u_0^{(j)} + C^{uw} (C^{ww} + \Gamma)^{-1} (y^{(j)} - G(u_0^{(j)})) \]

\[ C^{uw} = \frac{1}{(N_e - 1)} \sum_{j=1}^{N_e} (u_0^{(j)} - \bar{u}_0)(G(u_0^{(j)}) - G(\bar{u}_0))^T \]

\[ C^{ww} = \frac{1}{(N_e - 1)} \sum_{j=1}^{N_e} (G(u_0^{(j)}) - G(\bar{u}_0))(G(u_0^{(j)}) - G(\bar{u}_0))^T \]

Claim 2: \( \{u^{(j)}\}_{j=1}^{N_e} \approx P(u|y) \).

Observation: \( \{u^{(j)}\}_{j=1}^{N_e} = P(u|y) \) if \( G \) is linear and \( \mu_0 \) is Gaussian.
Issues with Ensemble Kalman Smoother

It kind of works....sometimes.

\[ u^{(j)} = u^{(j)}_0 + C^{uw}(C^{ww} + \Gamma)^{-1}(y^{(j)} - G(u^{(j)}_0)) \]

Underestimates the uncertainty

**Iterative smoothers**

\[ u^{(j)}_n = u^{(j)}_{n-1} + C^{uw}_{n-1}(C^{ww}_{n-1} + \Gamma)^{-1}(y^{(j)} - G(u^{(j)}_{n-1})) \]

Overestimates the uncertainty

**Ad-hoc fixes of Iterative smoothers**

\[ u^{(j)}_n = u^{(j)}_{n-1} + \rho \circ C^{uw}_{n-1}(C^{ww}_{n-1} + \alpha \Gamma)^{-1}(y^{(j)} - G(u^{(j)}_{n-1})) \]

\( \rho \) is a localization matrix and \( \alpha \) is an inflation parameter.
Understanding the iterative ensemble smoother with iterative regularization

Suppose we are interested in solving

\[
u = \arg \min_{u \in \mathcal{X}} ||\Gamma^{-1/2}(y - G(u))||^2\]

where \(\mathcal{X}\) has norm \(||C^{-1/2} \cdot||\)

Levenberg-Marquardt

\(u_{n+1}\) iteration level is given by

\[
u_{n+1} = u_n + \arg \min_{v \in \mathcal{X}} ||\Gamma^{-1/2}(y - G(u_n) - DG(u_n)v)||^2 + \alpha||C^{-1/2}v||_{\mathcal{X}}^2\]

After some computations

\[
u_{n+1} = u_n + CDG^*(u_n)(DG(u_n)C DG^*(u_n) + \alpha \Gamma)^{-1}(y - G(u_n))\]
Hanke proposed a way to select $\alpha$ and a stopping criteria ($\delta=\text{noise level}$):

$$||\Gamma^{-1/2}(y - G(u_n))|| \approx \delta$$

so that

$$u_{n+1} = u_n + CDG^*(u_n)(DG(u_n)CDG^*(u_n) + \alpha \Gamma)^{-1}(y - G(u_n))$$

generates a stable approximation to the solution of the classical inverse problem.

**Theorem [Hanke 1997]**

The LM scheme terminates after a finite number of iterations $n^*$ and

$$u_{n^*} \rightarrow u \quad \text{as} \quad \eta \rightarrow 0 \quad \text{(where} \quad G(u) = G(u^\dagger))$$

$u^\dagger$ is the truth
Consider an initial ensemble \( \{ u_0^{(j)} \}_{j=1}^{N_e} \subseteq X \)

**Linearize around the ensemble mean**

\[
\overline{u}_n \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} u_n^{(j)}
\]

\[
w_n^{(j)} \equiv g(u_n^{(j)}) \approx g(\overline{u}_n) + DG(\overline{u}_n)(u_n^{(j)} - \overline{u}_n)
\]

\[
C_{uu}^n DG^*(\overline{u}_n)v \approx C_{uw}^n v \\
DG(\overline{u}_n)C_{uu}^n DG(\overline{u}_n)^*v \approx C_{ww}^n v
\]

**Recall the update formula for the regularizing LM scheme**

\[
u_{n+1} = u_n + CDG^*(u_n)(DG(u_n)C DG^*(u_n) + \alpha \Gamma)^{-1}(y \ominus g(u_n))
\]

Replace

\[
\begin{align*}
  u_n & \quad \mapsto \quad \overline{u}_n, \\
  CDG^*(u_n) & \quad \mapsto \quad C_{uu}^n DG^*(\overline{u}_n) \approx C_{uw}^n \\
  DG(u_n)C DG^*(u_n) & \quad \mapsto \quad DG(\overline{u}_n)C_{uu}^n DG^*(\overline{u}_n) \approx C_{ww}^n
\end{align*}
\]
Regularizing ensemble Kalman method

**Update formula for the mean of the ensemble**

\[
\bar{u}_{n+1} = \bar{u}_n + C_n^{uw}(C_n^{ww} + \alpha \Gamma)^{-1}(y - \bar{w}_n)
\]

where \(\bar{w}_n \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} \mathcal{G}(u_n^{(j)})\).

We propose to update each ensemble in a consistent fashion

\[
u_{n+1}^{(j)} = u_n^{(j)} + C_n^{uw}(C_n^{ww} + \alpha \Gamma)^{-1}(y^{(j)} - \mathcal{G}(u_n^{(j)}))
\]

Selection of \(\alpha\):

\[
\rho \|\Gamma^{-1/2}(y - \bar{w}_n)\|_Y \leq \alpha \|\Gamma^{1/2}(C_n^{ww} + \alpha \Gamma)^{-1}(y - \bar{w}_n)\|_Y
\]

Stopping criteria

\[
\|\Gamma^{-1/2}(y - \bar{w}_n)\|_Y \approx \delta
\]
Let $\rho < 1$ and $\tau > 1/\rho$. Generate an initial ensemble $u_0^{(j)} \sim \mu_0$

**A regularizing Kalman method**

1. **Prediction Step:** Evaluate $w_m^{(j,f)} = G(u_m^{(j)})$ define $\overline{w}_m^f$
2. **Stopping criteria.** If
   
   $$||\Gamma^{-1/2}(y - \overline{w}_m^f)|| \leq \tau \eta$$
   
   Stop. Otherwise: define $C_m^{uw}$, $\overline{u}_m$, $C_m^{ww}$ and
3. **Analysis step:** Compute the updated ensembles

   $$u_{m+1}^{(j)} = u_m^{(j)} + C_m^{uw} (C_m^{ww} + \alpha_m \Gamma)^{-1} (y^{(j)} - w_m^{(j,f)})$$

   for $\alpha_m$ such that

   $$\alpha_m ||\Gamma^{1/2}(C_m^{ww} + \alpha_m \Gamma)^{-1}(y^\eta - \overline{w}_m^f)|| \leq \rho ||\Gamma^{-1/2}(y^\eta - \overline{w}_m^f)||$$
1 Introduction

2 Numerical Investigation of the Scheme

3 Applications
Synthetic experiment with Darcy flow model

Initial ensemble generated from a prior $\mathbb{P}(u) = N(\bar{u}, C)$. 
$\mathcal{G}(u)$ be the forward operator that arises from Darcy flow.

Consider a truth $u^\dagger \sim \mathbb{P}(u)$ from which synthetic data are generated by 
$y = \mathcal{G}(u^\dagger) + \xi$ \quad $\xi \sim N(0, \Gamma)$ (prescribed $\Gamma$ covariance of the Gaussian noise).

For the numerical investigation with respect to the approximation properties of the Bayesian posterior see

M. A. Iglesias

Iterative regularization for ensemble-based data assimilation in reservoir models. 
Synthetic experiment with Darcy flow model

Initial ensemble generated from a prior $\mathbb{P}(u) = N(\bar{u}, C)$.

$G(u)$ be the forward operator that arises from Darcy flow.

Some elements from the initial ensemble
Results from the standard ES choice $\alpha = 1$.

Reconstructing the truth with the mean of an ensemble of $N_e = 75$ (with small noise)
\[
\bar{u}_n \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} u_n^{(j)}
\]

\[
\| \Gamma^{-1/2} (y - G(\bar{u}_n)) \|_2^2 \quad \| \bar{u}_n - u^\dagger \|_{L^2(D)}
\]

Data misfit

Error w.r.t truth

(University of Nottingham)

A regularizing ensemble Kalman method

Marco Iglesias
Reconstructing the truth with the mean of an ensemble of $N_e = 75$ (with small noise)
Performance

$$\bar{u}_n \equiv \frac{1}{N} \sum_{j=1}^{N} u^{(j)}$$

$$\| \Gamma^{-1/2} (y - G(\bar{u}_n)) \|_{l^2}$$  \hspace{1cm}  $$\| \bar{u}_n - u^\dagger \|_{L^2(D)}$$

Data misfit

Error w.r.t truth

(U)nal of Nottingham)  A regularizing ensemble Kalman method  Marco Iglesias 29 / 47
Regularization parameter $\alpha$

Plot of log $\alpha$

- $\rho = 0.7$
- $\rho = 0.5$

(iteration vs. log alpha)
Regularizing properties as a function of the ensemble size

\( N_e = 50, \ \rho = 0.7 \)
Regularizing properties as a function of the ensemble size

$N_e = 75, \quad \rho = 0.7$

$N_e = 75, \quad \rho = 0.7$
Regularizing properties as a function of the ensemble size

\[ N_e = 100, \ \rho = 0.7 \]
Regularizing properties as a function of the ensemble size

$N_e = 150, \, \rho = 0.7$

![Graph showing log-data misfit and relative error](image)
Regularizing properties as a function of the ensemble size

\( N_e = 200, \ \rho = 0.7 \)

\[ \text{iteration} \]

\[ \text{log-data misfit} \]

\[ \text{relative error} \]

\( N_e = 200, \ \rho = 0.7 \)
Regularizing properties as a function of the ensemble size

\( N_e = 300, \ \rho = 0.7 \)

- Left panel: Log-data misfit vs. iteration
- Right panel: Relative error vs. iteration

(University of Nottingham)
Convergence as the noise level decreases

\[ \bar{u}_n \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} u_n^{(j)} \]

\[ \| \Gamma^{-1/2} (y - G(\bar{u}_n)) \|_2 \]

\[ \| \bar{u}_n - u^\dagger \|_{L^2(D)} \]
The proposed ES as an approximate regularizing LM scheme

Comparing ES with the regularizing LM scheme (on the same subspace)
Outline

1. Introduction
2. Numerical Investigation of the Scheme
3. Applications
### Forward map: Resin Transfer Molding

\[
- \nabla \cdot e^u \nabla p = f \quad \text{in } D(t) \\
p = p_{in} \quad \text{on } \Gamma_{in} \\
p = p_f \quad \text{on } \Gamma_s(t) \\
-e^u \nabla p \cdot n = 0 \quad \text{on } \Gamma_N
\]

\[u = \log(\kappa(x)) \in X \equiv L^\infty(D) \longrightarrow G(u) = \{p(x_i)\}_{i=1}^N \in Y \equiv \mathbb{R}^M\]

### The Inverse Problem

Find \( u \in X \) given

\[y = G(u) + \eta \quad \eta \sim N(0, \Gamma)\]
The Forward model
Electrical Impedance Tomography

Complete Electrode Model: Forward and Inverse Problem

Given $\kappa$, $\{z_m\}_{m=1}^{ne}$ and $I = \{I_m\}_{m=1}^{ne}$ compute $v$ and $V = \{V_m\}_{m=1}^{ne}$

\[ \nabla \cdot \kappa \nabla v = 0 \quad \text{in} \quad D, \]

\[ v + z_m \kappa \nabla v \cdot \nu = V_m \quad \text{on} \quad e_m, \quad m = 1, \ldots, ne, \]

\[ \nabla v \cdot \nu = 0 \quad \text{on} \quad \partial D \setminus \bigcup_{m=1}^{ne} e_m, \]

\[ \int_{e_m} \kappa \nabla v \cdot \nu \ ds = I_m \quad m = 1, \ldots, ne, \]

Inverse Problem:

Given $I^{(1)}, \ldots, I^{(N)}$ and the observations of voltages $V^{(1)}, \ldots, V^{(N)}$ find $\kappa$ and $z_m$
Geometric Parameterization with Level-Sets

Permeability $\kappa$ defined through level set function $u$:

$$\kappa(x) = \kappa_1 \chi_{\{u < 0\}}(x) + \kappa_2 \chi_{\{u \geq 0\}}(x).$$

$u \mapsto \kappa$ is discontinuous

Forward Map and initial ensemble

$$u \mapsto \kappa \mapsto G(u) = \{p(x_i)\}_{i=1}^{N} \in \mathbb{R}^{M}$$

Gaussian prior $\mu_0 = N(0, C_0)$ on the level-set function. Covariance $C_0$ reflects the regularity of the shape.

M. A. Iglesias, Y. Lu and A. M. Stuart
A level-set approach to Bayesian geometric inverse problems.
Iterative regularization provides strategies for regularizing Kalman based methods.

Regularization has strong effect in the robustness and accuracy of ensemble methods for solving both classical and Bayesian inverse problems.

The stabilization of the proposed method is suitable for solving level-set based geometric inverse problems.

Further investigations are required to establish the mathematical properties of these approximations.
M.A. Iglesias, K. Lin and A.M. Stuart  
Well-posed Bayesian geometric inverse problems arising in subsurface flow.  
*Inverse Problems*, 30 (2014) 114001

M. A. Iglesias  
A regularizing ensemble Kalman method for PDE-constrained inverse problems.  

M. A. Iglesias, Y. Lu and A. M. Stuart  
A level-set approach to Bayesian geometric inverse problems.  

M. A. Iglesias  
Iterative regularization for ensemble data assimilation in reservoir modeling  

M. Iglesias, K. Law and A.M. Stuart  
Ensemble Kalman methods for inverse problems.  

MCMC methods for functions: modifying old algorithms to make them faster.  