Hierarchical Blackbox Inversion

Andrew Stuart[†]

[†]California Institute of Technology

Data Assimilation, Uncertainty Reduction and Optimization for Subsurface Flow, IPAM May 22nd 2017

EnKF: N Chada, M Iglesias, KJH Law, L Roininen, C Schillings

MCMC: M Dunlop, M Iglesias

May 21, 2017

Overview

EnKF Inversion

Ensemble Kalman Filter (EnKF) and Inversion

Problem Statement

Find u from y where $\mathcal{G} : X \mapsto Y$, η is noise and

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathrm{N}(0, \Gamma).$$

- EnKF widely applied: geosciences, NWP and oceanography.
- Origins in sequential data assimilation. ([7] Evensen (1994)).
- Application to inverse problems. ([10] Iglesias (2016)).
- Requires only black box forward model $\mathcal{G}$.
- Very limited theoretical understanding.

Heuristic

Minimize, over a subspace defined via the linear span of an ensemble,

$$\Phi(u) := \frac{1}{2}\|y - \mathcal{G}(u)\|_\Gamma^2.$$

Basic EnKF Inversion ([9] Iglesias et al (2013))

- Initial Ensemble $\{u_0^{(j)}\}_{j=1}^J \subset X$.
- Ensemble First and Second Order Moments

  Means:

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

  Covariances:

$$C_n^{ww} = \frac{1}{J}\sum_{\ell=1}^{J}\big(\mathcal{G}(u_n^{(\ell)}) - \overline{w}_n\big) \otimes \big(\mathcal{G}(u_n^{(\ell)}) - \overline{w}_n\big),$$

$$C_n^{uw} = \frac{1}{J}\sum_{\ell=1}^{J}\big(u_n^{(\ell)} - \overline{u}_n\big) \otimes \big(\mathcal{G}(u_n^{(\ell)}) - \overline{w}_n\big).$$

- Update step $n \mapsto n+1$:

$$u_{n+1}^{(j)} = u_n^{(j)} + C_n^{uw}\big(C_n^{ww} + \Gamma\big)^{-1}\big(y - \mathcal{G}(u_n^{(j)})\big).$$

Linear span of initial ensemble is preserved:
[14] Li and Reynolds (2011), [9] Iglesias et al (2013).
Implement iteratively regularized form: [10] Iglesias (2016).

Continuous Time Limit

- Linear Case $\mathcal{G}(\cdot) = A \cdot$.
- $\Gamma = h^{-1}\Gamma_0, h \to 0.$
- Least Squares Functional

$$\Phi(u) = \frac{1}{2}\|y - Au\|^2_{\Gamma_0}.$$

- Gradient Structure

$$\frac{du^{(j)}}{dt} = -C\nabla\Phi(u^{(j)}),$$

$$C = \frac{1}{J}\sum_{\ell=1}^{J}\big(u^{(\ell)} - \overline{u}\big) \otimes \big(u^{(\ell)} - \overline{u}\big). \tag{1}$$

Theorem. ([13] Schillings and S (2016))

Algorithm minimizes $\Phi(\cdot; y)$ over a finite dimensional subspace defined by the linear span of the initial ensemble.

Numerical Results

Groundwater Flow

Forward Problem

Given $\kappa \in X := L^\infty(D;\mathbb{R}^+)$ find $p \in H_0^1(D;\mathbb{R})$ such that:

$$\begin{aligned} -\nabla \cdot (\kappa \nabla p) &= f, \quad x \in D, \\ p &= 0, \quad x \in \partial D. \end{aligned}$$

Inverse Problem

Set $\kappa = \exp(u)$. Given K linear functionals of the pressure $\mathcal{G}_k(u) = \ell_k(p)$, $\ell_k \in H^{-1}(D;\mathbb{R})$, find u from noisy measurements y where:

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathrm{N}(0, \Gamma).$$

Numerical Experiments: Groundwater Flow

Unknown channel geometry ($d \in \mathbb{R}^5$); unknown fields ($v \in L^2(D; \mathbb{R}^2)$).
Initial ensemble: draw i.i.d. d uniform and v Gaussian.

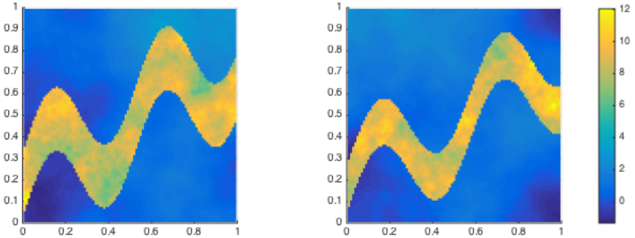


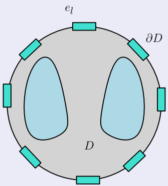
Figure: Left: true permeability. Right: reconstruction

Electrical Impedance Tomography (EIT)

Forward Problem

Given $(\kappa, I) \in L^\infty(D;\mathbb{R}^+) \times \mathbb{R}^m$ find $(\nu, V) \in H^1(D) \times \mathbb{R}^m$:

$$\begin{aligned} -\nabla \cdot (\kappa \nabla \nu) = 0 \quad &\in \quad D, \\ \nu + z_\ell \kappa \nabla \nu \cdot n = V_\ell \quad &\in \quad e_\ell, \quad \ell = 1, \dots, m, \\ \nabla \nu \cdot n = 0 \quad &\in \quad \partial D \setminus \cup_{\ell=1}^m e_\ell, \\ \int \kappa \nabla \nu \cdot n \, ds = I_\ell \quad &\in \quad e_\ell, \quad \ell = 1, \dots, m. \end{aligned}$$



Ohm's Law: $V = R(\kappa) \times I$.

Inverse Problem

Set $\kappa = \exp(u)$. Given a set of K noisy measurements of voltage V_k from currents I_k, and $\mathcal{G}_k(u) = R(\kappa) \times I_j$, find u from y where:

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathrm{N}(0, \Gamma).$$

Numerical Experiments: EIT

Unknown inclusion geometry ($d \in \mathbb{R}^3$); unknown conductivities ($\kappa \in \mathbb{R}^2$).
Initial ensemble: draw i.i.d. d uniform and κ uniform.

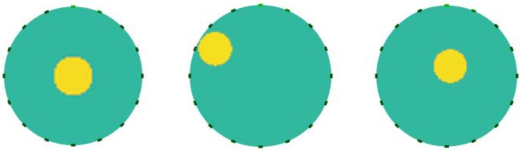


Figure: Left: Truth. Middle: $J = 3$. Right: $J = 50$.

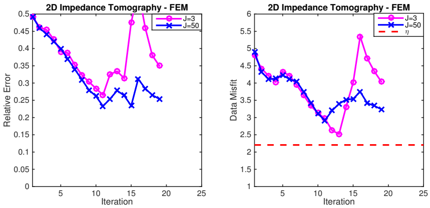


Figure: Left: relative error. Right: data misfit.

Hierarchical EnKF

Whittle-Matérn Initial Ensembles

- Create initial ensemble of functions via Gaussian random fields.
- Common choice: Matérn family

$$c_{\sigma,\nu,\tau}(x,x') := \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\tau|x-x'|\right)^{\nu} K_\nu\left(\tau|x-x'|\right).$$

  - Smoothness parameter: $\nu \in \mathbb{R}^+$.
  - Inverse length-scale parameter: $\tau \in \mathbb{R}^+$.
  - Amplitude parameter: $\sigma \in \mathbb{R}$.

- Corresponding covariance operator

$$\mathcal{C}_{\sigma,\nu,\tau} \propto \sigma^2 \tau^{2\nu} (\tau^2 I - \triangle)^{-\nu-\frac{d}{2}}.$$

  - $\nu = \alpha - \frac{d}{2}$.

- Hierarchical: invert for parameters such as σ, ν, τ as well as field itself.

Centred vs Non-centred

- Define $\theta = (\alpha, \tau)$.
- Generate samples v by solving the SPDE

$$(\tau^2 I - \triangle)^{\frac{\alpha}{2}} v = \sigma \tau^{\alpha - \frac{d}{2}} \xi,$$

where $\xi \sim \mathrm{N}(0, I)$ is white noise.

See [4] Lindgren et al (2011).

- Hierarchical: invert for parameters θ as well as field v.
- Centred approach: view (v, θ) as unknowns.
- Non-centred approach: view (ξ, θ) as unknowns.

See [11] Papaspiliopoulos et al (2007).

Groundwater Flow Revisited

Unknowns are:

- unknown channel geometry ($d \in \mathbb{R}^5$);
- unknown fields ($v \in L^2(D; \mathbb{R}^2)$);
- unknown hyperparameters ($\theta \in \mathbb{R}^4$).

Initial ensemble and truth given by:

- initial ensemble: draw i.i.d. d uniform, v Gaussian and θ uniform;
- true regularity parameters: $(\alpha_1, \alpha_2) = (2.0, 2.8)$;
- true inverse length scales $(\tau_1, \tau_2) = (30.0, 10.0)$.

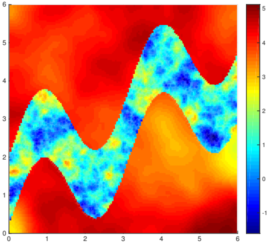


Figure: True log-permeability.

Numerical Experiments: Groundwater Flow

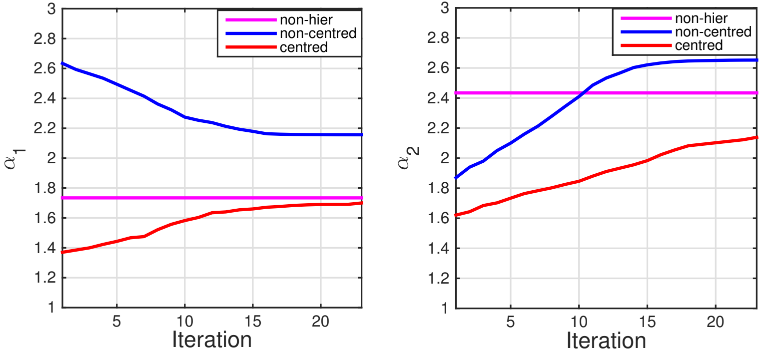


Figure: α_1 (truth 2) and α_2 (truth 2.8).

Numerical Experiments: Groundwater Flow

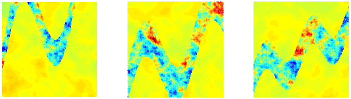


Figure: Nonhierarchical EnKF.

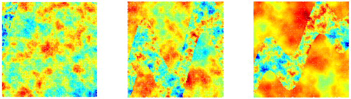


Figure: Centred hierarchical EnKF.

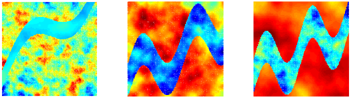


Figure: Non-centred hierarchical EnKF.

Non-Stationary Hyperparameters

- Treating the length scale $\tau^{-1} = \ell$ as a field.
- To ensure positivity write

$$\ell(x) := \exp(w(x)).$$

- Now we generate samples ξ from the SPDE

$$(1 - \ell(x; w)^2 \Delta)^{\alpha/2} u = \sigma \sqrt{\ell(x; w)^d} \xi.$$

- Centred approach: view (u, w) as unknown.
- Non-centred approach: view (ξ, w) as unknown.
- We show only non-centred results.

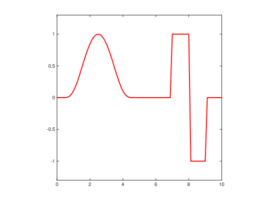
See [12] Roininen et al (2016).

Linear Inverse Problem

Forward Problem

Given $u \in L^2(D;\mathbb{R})$ find $p \in H^2(D) \cap H_0^1(D)$:

$$
\begin{aligned}
-\frac{d^2 p}{dx^2} + p &= u \qquad \in D,\\
p &= 0 \qquad \in \partial D.
\end{aligned}
$$



Inverse Problem

Given $Au := \{p(x_j)\}_{j=1}^{J}$, find u from y where:

$$y = Au + \eta, \quad \eta \sim \mathrm{N}(0, \Gamma).$$

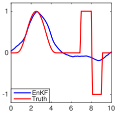


Figure: Non-hierarchical method.

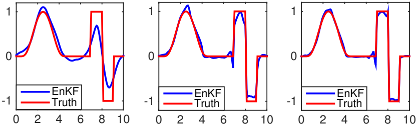
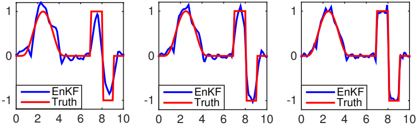


Figure: Above: Gaussian prior on ℓ. Below: Cauchy prior on ℓ.



Bayesian Inversion

The Inverse Problem

Problem Statement

Find $\mathbb{P}(u|y)$ from y where $u \sim \mathbb{P}(u) = \mathsf{N}(0, C)$,
$\mathcal{G} : X \mapsto Y$ and $y|u \sim \mathbb{P}(y|u) = \mathsf{N}\big(\mathcal{G}(u), \Gamma\big)$.

- Considerably more computationally intensive than optimization.
- Allows soluton of inverse problems with uncertainty quantification.
- Monte Carlo Markov Chain is a key methodology.
- Mesh-independent convergence of MCMC is crucial.

Theorem. ([5] S (2010), [6] Iglesias et al (2016).)

Bayes' Rule for functions:

$$\mathbb{P}(du|y) = \frac{1}{Z} \exp\left(-\Phi(u)\right) \mathbb{P}(\mathrm{d}u).$$

$\mathbb{P}(du|\cdot)$ Lipschitz in Hellinger metric.

The Basic Function Space MCMC Algorithm

See [1] Cotter et al (2013).

- Initialize Pick $u^{(0)} \in X$.
- Propose $u^{\star} = (1-\beta^2)^{\frac{1}{2}} u^{(n)} + \beta \xi^{(n)}, \quad \xi^{(n)} \sim N(0, C)$.
- Accept Set $u^{(n+1)} = u^{\star}$ with probability

$$\min\left\{1, \exp\left(\Phi(u^{(n)}) - \Phi(u^{\star})\right)\right\}.$$

- Otherwise set $u^{(n+1)} = u^{(n)}$.

Theorem. ([3] Hairer et al (2013).)

This algorithm converges with random constant K independent of mesh used to approximate forward model:

$$\mathbb{P}\left(g(u)|y\right) = \frac{1}{N}\sum_{n=1}^{N} g(u^{(n)}) + \frac{K(\xi)}{\sqrt{N}}$$

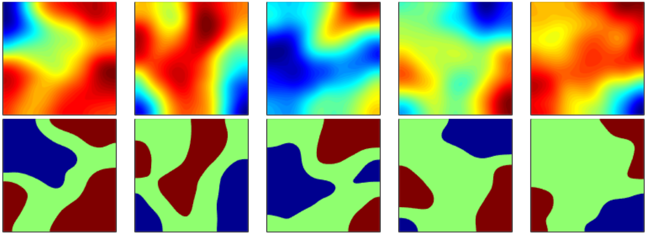
(Not true for RWM: $u^{\star} = u^{(n)} + \beta \xi^{(n)}, \quad \xi \sim N(0, C)$.)

Numerical Results

Groundwater Flow Revisited

- Level set method: allows use of Gaussian priors to represent piecewise constant functions. See [6] Iglesias et al (2016), [2] Dunlop et al (2016).
- Threshold a continuous function at a number of levels.
- Let $X = C^0(D)$, and define $T : X \mapsto Z$ by

$$T(u)(x) = \begin{cases} \kappa_1 & u(x) \in (-\infty, -1) \\ \kappa_2 & u(x) \in [-1, 1) \\ \kappa_3 & u(x) \in [1, \infty). \end{cases}$$

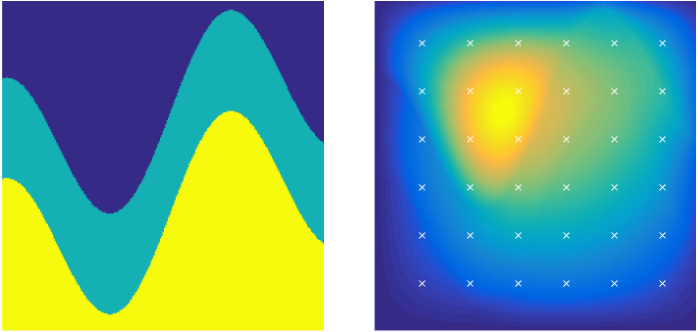


Numerical Experiments: Groundwater Flow

- The forward map $\mathcal{G} : X \mapsto \mathbb{R}^K$ is given by $\mathcal{G} = \mathcal{F} \circ T$, $\mathcal{F} : \kappa \mapsto p$.
- We choose $K = 36$, and observe data

$$y = \mathcal{G}(u^\dagger) + \eta, \quad \eta \sim N(0, 0.05^2).$$

- True permeability $T(u^\dagger)$ and observed data y are shown below.

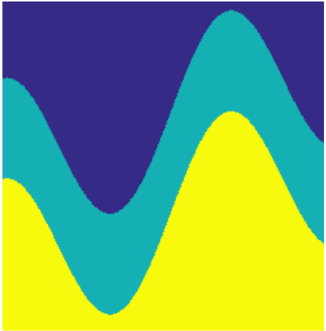


Numerical Experiments: Groundwater Flow

- The forward map $\mathcal{G} : X \mapsto \mathbb{R}^K$ is given by $\mathcal{G} = \mathcal{F} \circ T$, $\mathcal{F} : \kappa \mapsto p$.
- We choose $K = 36$, and observe data

$$y = \mathcal{G}(u^\dagger) + \eta, \quad \eta \sim N(0, 0.05^2).$$

- True permeability $T(u^\dagger)$ and estimated mean $T(\mathbb{E}(u))$ are shown below.



Hierarchical Bayesian Inversion

Hierarchical Models

- Use Whittle-Matern Gaussian random field.
- Recall $\theta = (\alpha, \tau)$ denotes parameters encoding regularity and length-scale.
- Prior

$$\mathbb{P}(u, \theta) = \mathbb{P}(u|\theta)\mathbb{P}(\theta).$$

- Posterior

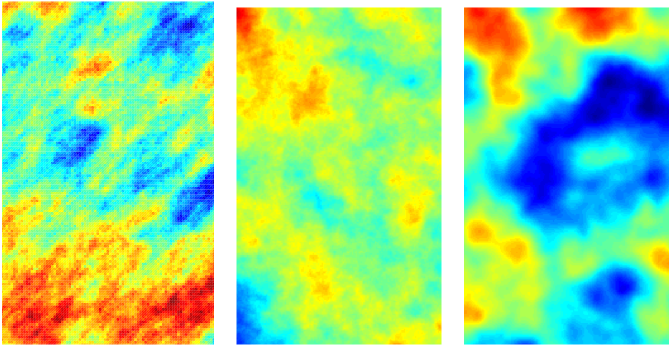
$$\mathbb{P}(u, \theta|y) = \frac{1}{Z} \exp\left(-\Phi(u)\right) \mathbb{P}(u, \theta).$$

- Use of centred (u, θ) or non-centred (ξ, θ) variables possible:

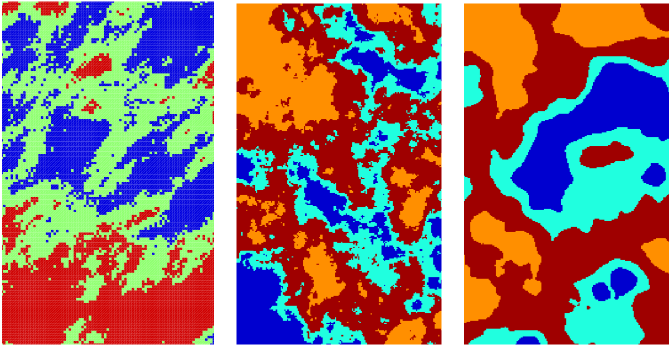
$$(\tau^2 I - \triangle)^{\frac{\alpha}{2}} u = \sigma \tau^{\alpha - \frac{d}{2}} \xi,$$

- Use Metropolis-within-Gibbs to update $\xi|\theta, y$ and $\theta|\xi, y$.

Hierarchical Level Set Function

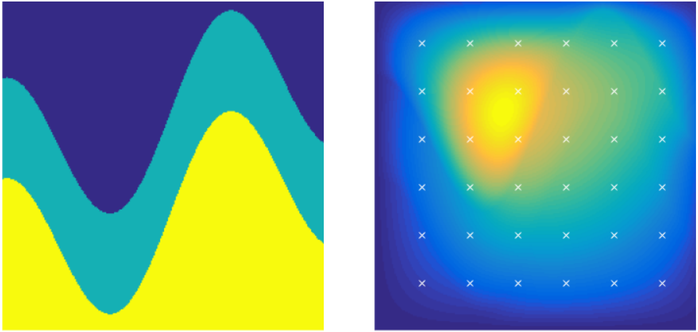


Hierarchical Level Set Thresholded



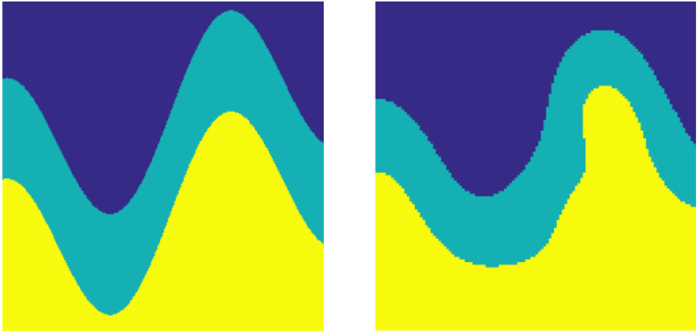
Numerical Example: Groundwater Flow Revisited

- We return to the groundwater flow example. We use the same data y, but now use a hierarchical Gaussian prior in which the length scale and regularity of samples are treated as additional unknowns in the problem.
- True permeability $T(u^\dagger)$ and observed data y are shown below.



Numerical Example: Groundwater Flow Revisited

- We return to the groundwater flow example. We use the same data y, but now use a hierarchical Gaussian prior in which the length scale and regularity of samples are treated as additional unknowns in the problem.
- True permeability $T(u^\dagger)$ and estimated mean $T(\mathbb{E}(u))$ are shown below.



Numerical Example: Groundwater Flow Revisited

- We return to the groundwater flow example. We use the same data y, but now use a hierarchical Gaussian prior in which the length scale and regularity of samples are treated as additional unknowns in the problem.
- Recall estimated mean $T(\mathbb{E}(u))$ without hierarchical estimation.

Conclusions

- Hierarchical inversion: considerable benefits at little extra cost.
- Optimization: use EnKF.
- Sampling: use function space MCMC.
- Non-centred methods key to hierarchical success.

References

[1] S. L. Cotter, G. O. Roberts, A. M. Stuart and D. White
MCMC methods for functions: modifying old algorithms to make them faster
Statistical Science, 28(3):424-446, 2013.

[2] M. M. Dunlop, M. A. Iglesias and A. M. Stuart
Hierarchical Bayesian level set inversion,
Statistics and Computing, 2016.

[3] M. Hairer, A. M. Stuart and S. J. Vollmer
Spectral gaps for Metropolis-Hastings algorithms in infinite dimensions
Annals of Applied Probability, 24:2455-2490, 2014.

[4] F. Lindgren, H. Rue and J. Lindström
An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach
JRSS-B (Statisical Methodology), 73 (4):423-498, 2011.

[5] A. M. Stuart
Inverse Problems: a Bayesian perspective
Acta Numerica, 19(May):451-559, 2010.

[6] M. Iglesias, Y. Lu and A.M. Stuart
A Bayesian level set method for geometric inverse problems.
Interfaces and Free Boundaries, **18**(2016), 181-217.

[7] G. Evensen
Data Assimilation: the Ensemble Kalman Filter,
Springer, 2009.

References

[8] N. K. Chada, M. Iglesias, L. Roininen, A. M. Stuart
Geometric and Hierarchical Ensemble Kalman Inversion,
In preparation.

[9] M. A. Iglesias, K. Law, A. M. Stuart
Ensemble Kalman method for inverse problems,
Inverse Problems, 29 (4), 045001, 2013.

[10] M. A. Iglesias
A regularising iterative ensemble Kalman method for PDE-constrained inverse problems,
Inverse Problems, 32 (2), 025002, 2016.

[11] O. Papaspiliopoulos, G. O. Roberts, and M. Sköld
A general framework for the parametrization of hierarchical models,
Statistical Science, pages 59-73, 2007.

[12] L. Roininen, M. Girolami, S. Lasanen, M. Markkanen
Hyperpriors for Matern fields with applications in Bayesian inversion,
arXiv:1612.02989.

[13] C. Schillings, A. M. Stuart
Analysis of the ensemble Kalman filter for inverse problems,
arXiv:1602.02020v1, SIAM Num. Analysis, to appear.

[14] G. Li and A.C. Reynolds
Iterative ensemble Kalman filters for data assimilation
SPE J. 14 (3) 496505 SPE-109808-PA

[15] A.A. Emerick
Towards a hierarchical parametrization to address prior uncertainty in ensemble-based data assimilation
Comput Geosci (2016) 20:3547