Iterative Regularization for Data Assimilation in Petroleum Reservoirs

Marco Iglesias

Mathematics Institute
University of Warwick

June 17th, 2013
University of Warwick

Outline

Overview

Main Objective:

Develop efficient and robust implementations of iterative regularization methods for data assimilation in subsurface flow applications.

M. Iglesias
Iterative regularization for data assimilation in reservoir models. *In preparation.*

Key aspects of the proposed work:

- Apply existing iterative regularization (IR) techniques [Kaltenbacher, 2010] for the solution of inverse problems in subsurface flow.
- Use ideas from IR to develop stable RML and Kalman-based ensemble methods that capture aspects of the posterior.
- Synthetic experiments show substantial benefits in the robust estimation of unknown geologic properties.

Background of the proposed work

- Iterative regularization methods for history matching (joint with Clint Dawson, UT Austin).

  M. A. Iglesias and C. Dawson
  The regularizing Levenberg-Marquardt scheme for history matching of petroleum reservoirs, *submitted to Computational Geosciences, http://arxiv.org/abs/1302.3501. 2013*

- Ensemble Kalman method for the solution of generic inverse problems (joint with Kody Law (now KAUST) and Andrew Stuart (Warwick)).

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

- Infinite-dimensional Bayesian framework for evaluating approximations of the posterior (joint with Kody Law (now KAUST) and Andrew Stuart (Warwick)).

  M. Iglesias, K. Law and A.M. Stuart,
  Evaluating Gaussian approximations for data assimilation in reservoir models applications. *to appear in Computational Geosciences, 2013.*

Outline

Standard approach for history matching

Compute $u_{MAP} = \arg\min_{u \in X} J(u)$

$$J \equiv \frac{1}{2}||\Gamma^{-1/2}(y^{\eta} - G(u))||_Y^2 + \frac{1}{2}||C^{-1/2}(u - \overline{u})||_X^2$$

where

- u is the unknown (geologic properties).
- $\overline{u}$ priori mean of the unknown.
- C prior covariance of the unknown.
- $G(u)$ is the model predictions of production data (flow rates, BHP).
- y^{η} production data corrupted by noise.
- Γ covariance of the observational noise.

In the Bayesian framework: u_{MAP} maximizes the posterior $\mu(u) \propto \exp(-J(u))$ (i.e. MAP estimator)

Standard approach for history matching

Deterministic framework: Tikhonov

$$J(u, y^\eta) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2 + \underbrace{\frac{1}{2}||C^{-1/2}(u - \overline{u})||_X^2}_{\textit{regularization}}$$

In theory we have stability

It can be shown that if $y_k \to y^\eta$ then

$$u_k \equiv \arg\min_{u\in X} J(u, y^k) \to u \equiv \arg\min_{u\in X} J(u, y^\eta)$$

...but in practice

The stability in the computation of $\arg\min_{u\in X} J(u, y)$ depends on the selection of Γ and C. Typical choices may lead to lack of robustness and even numerical instabilities.

Proposed approach: Iterative regularization (IR)

General idea

Compute stable and robust approximations

$$u_{IR} = \arg\min_{u\in X} \Phi(u, y^\eta) \equiv \arg\min_{u\in X} \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2$$

Enforcing geologic constraints

Prior mean is used to initialize the IR approach (i.e. $u_0 = \overline{u}$).
Prior error covariance is used in the regularization applied in the inner loop of the algorithm (via penalization of the increment in the norm $||C^{-1/2}\cdot||_X$).

Iterative regularization methods have been widely used in the solution of inverse problems [Hanke 1997, Kaltenbacher, 2010].
We extend this ideas to RML and ensemble Kalman methods where the aim is to capture aspects of the posterior.

Numerical Instabilities of standard approaches for minimizing $J(u, y^\eta) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2 + \frac{1}{2}||C^{-1/2}(u - \overline{u})||_X^2$

Example: Gauss-Newton (GN) $u_{m+1} = u_m + \kappa_m \Delta u$

$$\left[DG^*(u_m)\Gamma^{-1}DG(u_m) + C^{-1}\right]\Delta u$$

$$= DG^*(u^n)\Gamma^{-1}[y^\eta - G(u_m)) - C^{-1}(u_m - \overline{u})]$$

It is well known that GN for history matching results in “poor estimates”.

$DG^*(u_m)\Gamma^{-1}DG(u_m) + C^{-1}$ this is ill-conditioned for the typical choices of Γ and C.

Partial fix: Levenberg-Marquardt

$$\left[DG^*(u_m)\Gamma^{-1}DG(u_m) + C^{-1} + \lambda_m C^{-1}\right]\Delta u$$

$$= DG^*(u^n)\Gamma^{-1}[y^\eta - G(u_m)) - C^{-1}(u_m - \overline{u})]$$

for some $\lambda_m > 0$.

Standard Levenberg-Marquardt

The LM parameter

$\lambda_0 = \max\{\sqrt{J(u_0)/N_d}, J(u_0)/N_d\}$, where N_d is the number of observations.

$$\lambda_{m+1} = \begin{cases} \lambda_m/10 & \text{if } J(u_{m+1}) < J(u_m) \\ 10\lambda_m & \text{if } J(u_{m+1}) \geq J(u_m) \end{cases}$$

The LM stopping criteria

$$\frac{|J(u_{m+1}) - J(u_m)|}{J(u_{m+1})} \leq \epsilon_0, \qquad \frac{||u_{m+1} - u_m||_X}{||u_{m+1}||_X} \leq \epsilon_1$$

Failure of standard LM (without further regularization)

Example of how some choices of Γ and C may lead to instabilities on the minimizer of

$$J(u, y^\eta) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2 + \frac{1}{2}||C^{-1/2}(u - \overline{u})||_X^2$$

Let us fix Γ (consistent with our synthetic data) and parameterize C as

$$C = \kappa^{-1} C_0$$

with C_0 fixed (spherical covariance function).

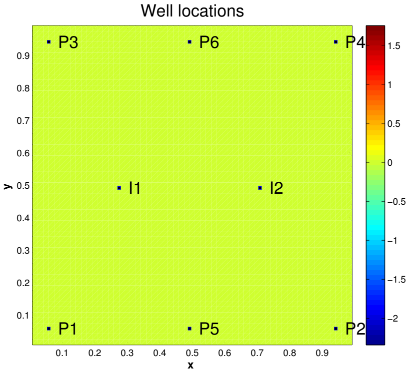
Remark

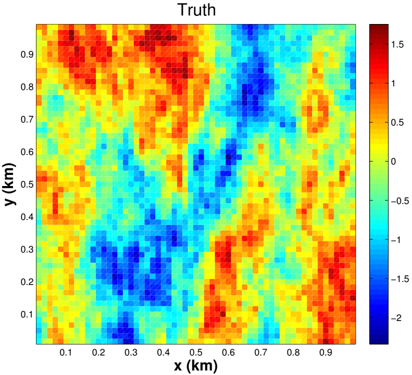
The parameter κ enables us to control the size of the prior term relative to the data misfit

$$J(u, y^\eta) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2 + \frac{\kappa}{2}||C_0^{-1/2}(u - \overline{u})||_X^2$$

Synthetic experiment

Model: Oil-water incompressible reservoir. Injection wells subject to water injection. Production wells subject to total flow rates.
Measure variables: BHP at the injectors and oil/water rates at the producers.

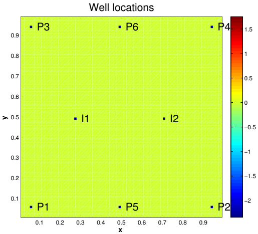


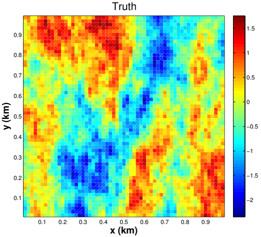


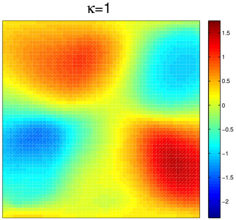
True is generated from $N(\overline{u}, C_0)$.

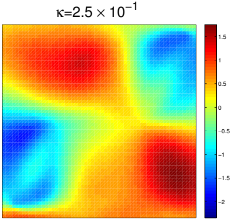
Lack of stability with small $\kappa' s$

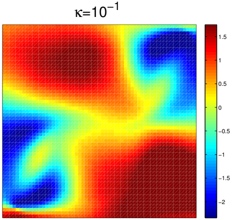
$$J(u, y^{\eta}) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^{\eta} - G(u))||_Y^2 + \frac{\kappa}{2}||C_0^{-1/2}(u - \overline{u})||_X^2$$



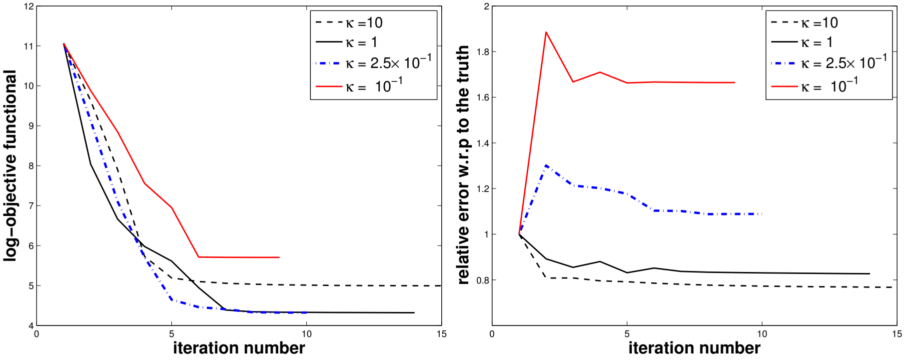








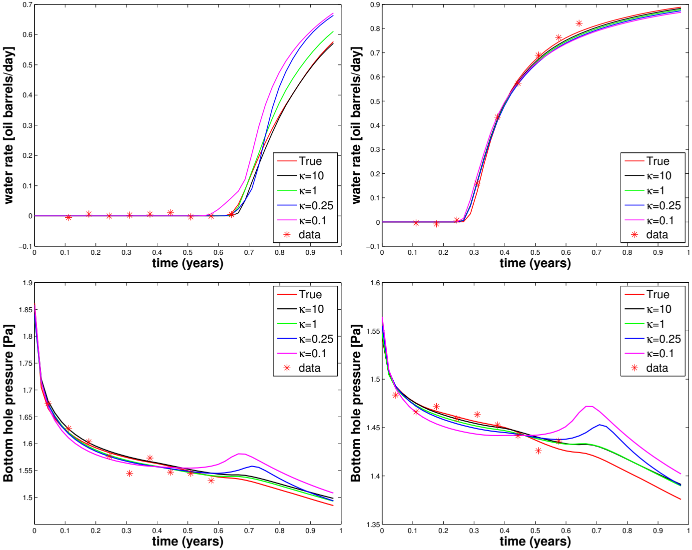
Lack of stability with small $\kappa' s$



These criteria fail to provide “reasonable” estimates.

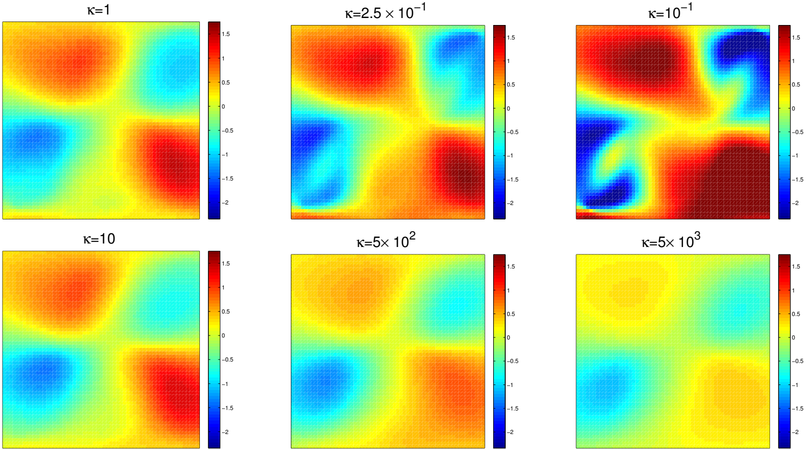
$$\frac{|J(u_{m+1}) - J(u_m)|}{J(u_{m+1})} \leq \epsilon_0, \qquad \frac{||u_{m+1} - u_m||_X}{||u_{m+1}||_X} \leq \epsilon_1$$

History matching with small $\kappa' s$



Standard LM without further regularization

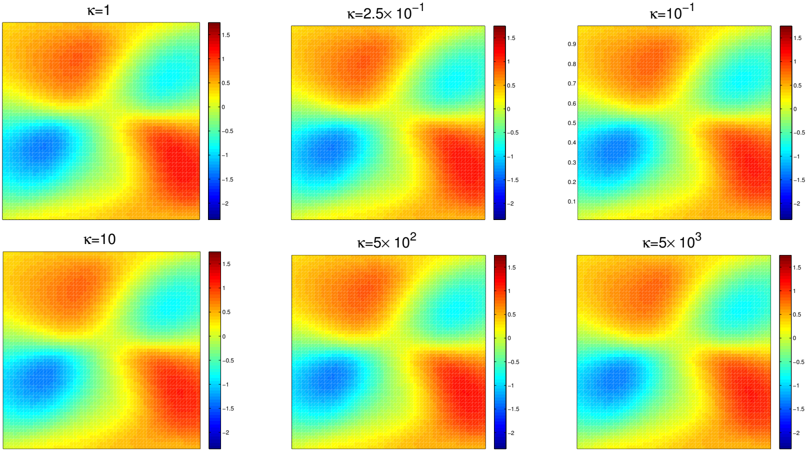
$(C = \kappa^{-1} C_0)$. Small κ= lack of stability. Large κ =lack of fidelity.



Standard LM results could be improved if further regularization methods (e.g. TSVD) is applied.

Iterative regularization (Regularizing LM)

$u_{IR} = \arg\min_{u \in X} \frac{1}{2} ||\Gamma^{-1/2}(y^{\eta} - G(u))||^2_Y$. Robust under the selection of κ. C is used in the inner loop for regularization.



Outline

Iterative regularization

We want to compute stable approximations to the minimizer of

$$\Phi(u, y^\eta) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u))||_Y^2$$

We are interested in nonlinear IP so we do it iteratively

Compute $u_{m+1}^\eta = \Delta u_m^\eta + u_m^\eta$ such that Δu_m^η minimizes

$$J^m(w) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u_m^\eta) - DG(u_m^\eta)w)||_Y^2$$

The linearized IP is also ill-posed.

Regularizing LM of Hanke [Hanke, 1997, 2010]

Compute $\Delta u_m^\eta(\alpha_m) = \mathrm{argmin}_{w\in X} J_{LM}^m(w, \alpha_m)$, where

$$J_{LM}^m(w, \alpha_m) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u_m^\eta) - DG(u_m^\eta)w)||_Y^2 + \frac{1}{2}\alpha_m||C^{-1/2}w||_X^2$$

Key aspects of the Regularizing LM

Regularizing LM of Hanke

Compute $\Delta u_m^\eta(\alpha_m) = \mathrm{argmin}_{w\in X} J_{LM}^m(w,\alpha_m)$, where

$$J_{LM}^m(w,\alpha_m) \equiv \frac{1}{2}||\Gamma^{-1/2}(y^\eta - G(u_m^\eta) - DG(u_m^\eta)w)||_Y^2 + \frac{1}{2}\alpha_m||C^{-1/2}w||_X^2$$

Regularizing LM parameter (discrepancy principle)

$$||\Gamma^{-1/2}(y^\eta - G(u_m^\eta) - DG(u_m^\eta)\Delta u_m^\eta(\alpha_m))||_Y^2 \geq \rho^2||\Gamma^{-1/2}(y^\eta - G(u_m^\eta))||_Y^2$$

for some $\rho \in (0,1)$.

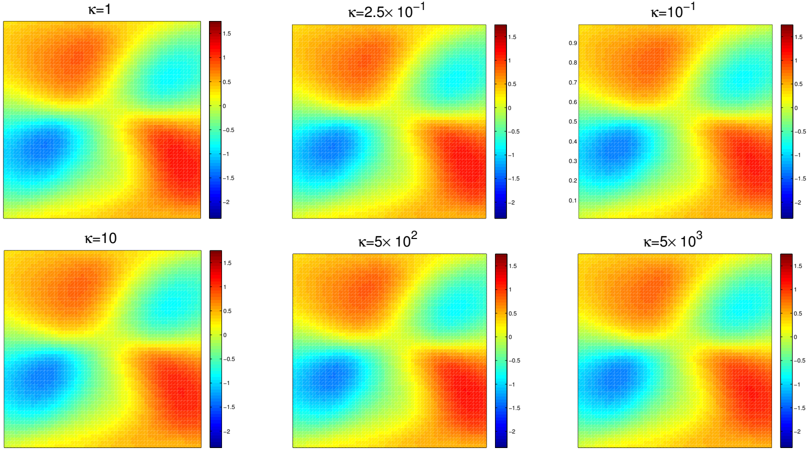
Regularizing LM stopping criteria (discrepancy principle)

$$||\Gamma^{-1/2}(y^\eta - G(u_{k+1}^\eta))||_Y \leq \tau\delta \leq ||\Gamma^{-1/2}(y^\eta - G(u_k^\eta))||_Y$$

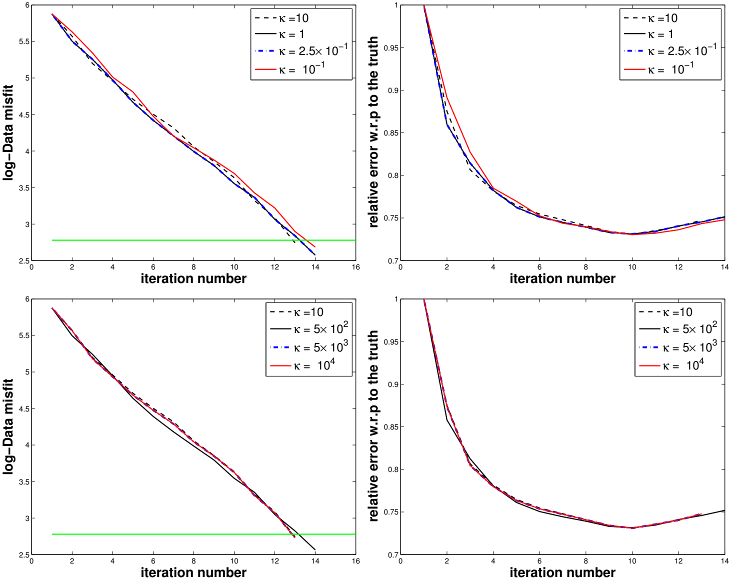
for $\tau > 1/\rho$ where $\delta = ||\Gamma^{-1/2}\eta||_Y$ is the noise level

Numerical Experiments. Iterative regularization (Regularizing LM)

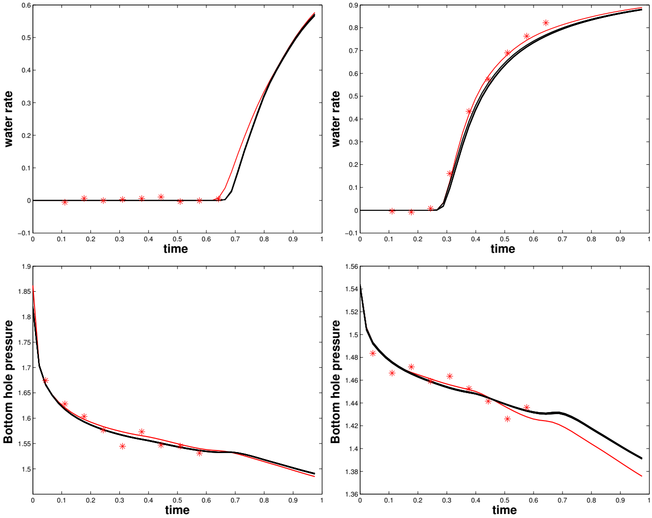
$u_{IR} = \arg\min_{u \in X} \frac{1}{2} ||\Gamma^{-1/2}(y^{\eta} - G(u))||_Y^2$. Robust under the selection of κ.



Small and big $\kappa's$



History matching with all $\kappa' s$



Outline

Applications of IR for approximating the Bayesian posterior

Recall the Randomized Maximum Likelihood (RML) method

$$\hat{u}^{(j)} = \arg\min_{u \in X} \left\{ \frac{1}{2} ||\Gamma^{-1/2}(y^{(j)} - G(u))||_Y^2 + \frac{1}{2} ||C^{-1/2}(u - u^{(j)}|| \right\}$$

Randomizing IR

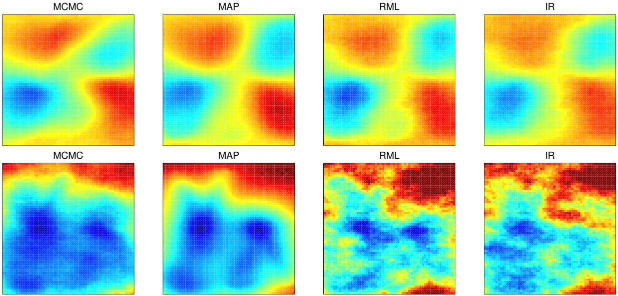
$$\hat{u}^{(j)} = \arg\min \left\{ \frac{1}{2} ||\Gamma^{-1/2}(y^{(j)} - G(u))||_Y^2 \right\}$$

with initial guess $u_0^{(j)} = u^{(j)}$ and (enforcing geologic constraint C as before).

We assess the proposed approach in terms of capturing the Bayesian posterior. We use an MCMC method for functions that scales well with respect to the parameter space and enable us to develop Benchmarks of moderate size [Iglesias/Law/Stuart, 2013].

Approximating the posterior. Case 1.

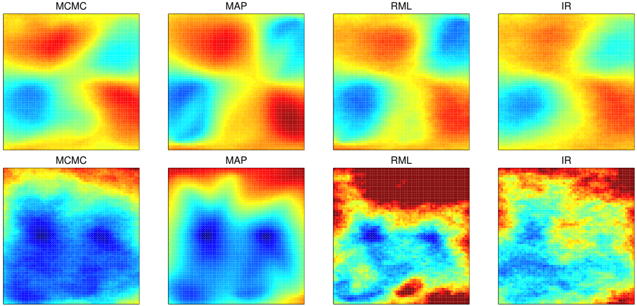
Here RML is computed with the standard LM of the previous section.



Method	$\frac{\|\|u^{pos}-u\|\|}{\|\|u^{pos}\|\|}$	$\frac{\|\|\sigma^{pos}-\sigma\|\|}{\|\|\sigma^{pos}\|\|}$	cost
MCMC	0.000	0.000	5.0×10^7 FM
MAP	0.396	0.192	n
RML ($N_e = 100$)	0.401	0.419	100n
IR ($N_e = 100$)	0.318	0.384	100n

FM=forward model evaluations; n=cost of computing the MAP

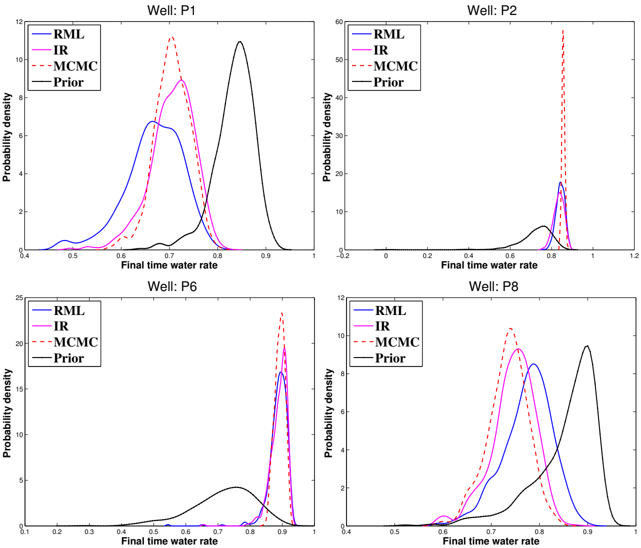
Approximating the posterior. Case 2 (more measurements in time).



Method	$\frac{\|\|u^{pos}-u\|\|}{\|\|u^{pos}\|\|}$	$\frac{\|\|\sigma^{pos}-\sigma\|\|}{\|\|\sigma^{pos}\|\|}$	cost
MCMC	0.000	0.000	5.0×10^7 FM
MAP	0.553	0.2338	n
RML ($N_e = 100$)	0.453	0.855	100n
IR ($N_e = 100$)	0.326	0.519	100n

FM=forward model evaluations; n=cost of computing the MAP

Approximating the posterior. Case 2 (more measurements in time).



Outline

Regularizing ensemble Kalman-based method

Inverse problem

Fin u given $y^\eta = G(u) + \eta$

Artificial Dynamics via state augmentation

Define

$$z = \begin{pmatrix} u \\ w \end{pmatrix}, \qquad \Xi(z) = \begin{pmatrix} u \\ G(u) \end{pmatrix}, \qquad z_{n+1} = \Xi(z_n).$$

and then data is, for $H = (0, I)$,

$$y_n = Hz_n + \eta_n \qquad \text{where } \eta_n \sim N(0, \Gamma).$$

State Estimation

Try to estimate z_n given $y_n = y^\eta + \xi_n, \quad \xi_n \sim N(0, \Gamma)$ (perturbed observations).

A Regularizing Kalman method

Prediction Step

$$z_{n+1}^{(j,f)} = \Xi(z_n^{(j,a)}), \quad j \in \{1, \cdots, J\}.$$

Compute ensemble mean and covariance $z_{n+1}^f = \frac{1}{J}\sum_{j=1}^J z_{n+1}^{(j,f)}$
$C_{n+1} = \frac{1}{J}\sum_{j=1}^J (z_{n+1}^{(j,f)} - z_{n+1}^f)(z_{n+1}^{(j,f)} - z_{n+1}^f)^T$

Analysis Step

$$z_{n+1}^{(j,a)} = \mathrm{argmin}_z \left(\frac{1}{2}||\Gamma^{-1/2}(y_{n+1} - Hz)||^2 + \frac{\alpha}{2}||C_{n+1}^{-\frac{1}{2}}(z - z_{n+1}^{(j,f)})||^2 \right)$$

$$z_{n+1}^{(j,a)} = (HC_{n+1}H^T + \alpha\Gamma)^{-1}(y^j - Hz_n^{(j,f)}), \quad j \in \{1, \cdots, J\}.$$

Perturbed Observations Data

$$y_n^j = y_n + \eta_n^j, \qquad \eta_n^j \sim N(0, \Gamma)$$

A Regularizing Kalman method

Analysis Step

$$z_{n+1}^{(j,a)} = \mathrm{argmin}_z \Big(\frac{1}{2} ||\Gamma^{-1/2}(y_{n+1} - Hz)||^2 + \frac{\alpha}{2} ||C_{n+1}^{-\frac{1}{2}}(z - z_{n+1}^{(j,f)})||^2 \Big)$$

$$z_{n+1}^{(j,a)} = (HC_{n+1}H^T + \alpha\Gamma)^{-1}(y^j - Hz_n^{(j,f)}), \quad j \in \{1, \cdots, J\}.$$

Regularizing LM parameter

$$||\Gamma^{-1/2}(y^\eta - Hz_{n+1}^a)||_Y \geq \rho ||\Gamma^{-1/2}(y^\eta - Hz_{n+1}^f)||_Y$$

for some $\rho \in (0, 1)$.

Regularizing LM stopping criteria

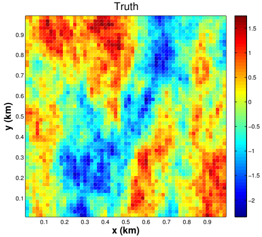
Inspired by iterative regularization methods, we require

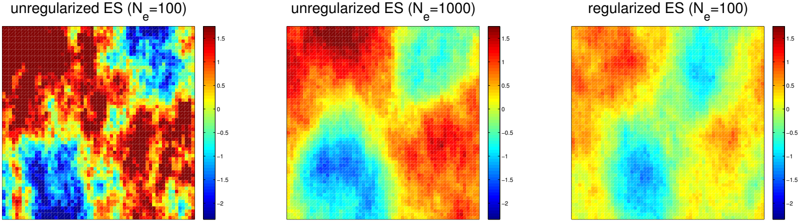
$$||\Gamma^{-1/2}(y^\eta - G(u_{n+1})||_Y \leq \tau ||\Gamma^{-1/2}\eta||_Y \leq ||\Gamma^{-1/2}(y^\eta - G(u_n))||_Y$$

for $\tau > 1/\rho$ where $u_n = H^T z_n^a$.

Performance of the regularizing EnKF

ϵ=relative error with respect to the truth.



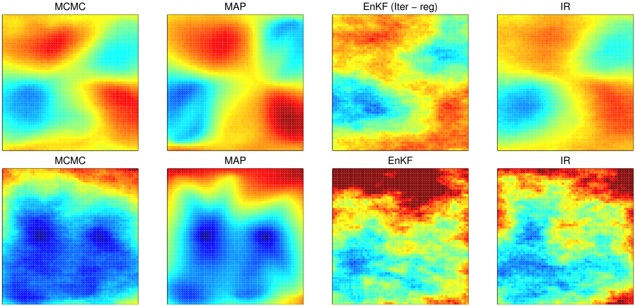


$\epsilon = 174\%$ $cost = 100FM$

$\epsilon = 91\%$ $cost = 1000FM$

$\epsilon = 78\%$ $cost = 1000FM$

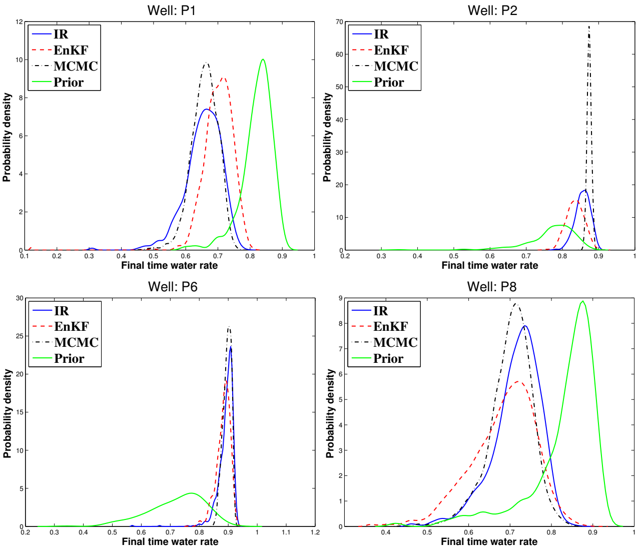
Approximating the posterior



Method	$\frac{\|\|u^{pos}-u\|\|}{\|\|u^{pos}\|\|}$	$\frac{\|\|\sigma^{pos}-\sigma\|\|}{\|\|\sigma^{pos}\|\|}$	cost
MCMC	0.000	0.000	5.0×10^7 FM
MAP	0.553	0.233	n
EnKF ($N_e = 100$)	0.540	0.510	1000FM
IR ($N_e = 100$)	0.326	0.487	100n

M=forward model evaluations; n=cost of computing the MAP

Approximating the posterior with the regularizing EnKF



Summary

- Iterative regularization provides robust algorithms to compute approximate solutions to inverse problems in reservoir modeling.
- Ideas of IR methods can be extended to ensemble-based techniques that can be use to approximate the Bayesian posterior.
- In particular, a regularizing ES based on IR ideas has been proposed and synthetic experiments indicate this method is a robust derivative-free iterative solver for the solution of nonlinear ill-posed inverse problems.
- Further investigations are required to establish the mathematical properties of these approximations.

References I

M Hanke
A Regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems, *Inverse Problems* 13, 1997.

M Hanke
The regularizing Levenberg-Marquardt scheme is of optimal order,*J. Integral Equations Appl*, 22 (2), 2010.

M. A. Iglesias and C. Dawson
The regularizing Levenberg-Marquardt scheme for history matching of petroleum reservoirs, *submitted to Computational Geosciences, http://arxiv.org/abs/1302.3501. 2013*

M. A. Iglesias and C. Dawson
An iterative representer-based scheme for data inversion in reservoir modeling,
Inverse Problems, 25(2009), 035006 (34pp)

M. A. Iglesias and D. McLaughlin
Data inversion in coupled subsurface flow and geomechanics models,
Inverse Problems, 28 (2012) 115009.

M. A. Iglesias and D. McLaughlin
Level-set techniques for facies identification in reservoir modeling,
Inverse Problems, 27(2011) 035008.

References II

M. A. Iglesias and C. Dawson

The regularizing Levenberg-Marquardt scheme for history matching of petroleum reservoirs, *submitted to Computational Geosciences, http://arxiv.org/abs/1302.3501. 2013*

M. Iglesias, K. Law and A.M. Stuart,

Ensemble Kalman methods for inverse problems. *Inverse Problems.* 29 (2013) 045001

M. Iglesias, K. Law and A.M. Stuart,

Evaluating Gaussian approximations for data assimilation in reservoir models applications. *to appear in Computational Geosciences, 2013.*