Evaluating Data Assimilation Algorithms with MCMC: Two Case studies

Kody Law[1]

[1]SRI UQ Center
KAUST

Data Assimilation Shortcourse
KAUST, February 20th, 2014

Collaboration with Andrew Stuart and Marco Iglesias (Warwick).

Funded by EPSRC, ERC and ONR

Center for Uncertainty Quantification

Outline

1. NSE
   - INVERSE PROBLEM
   - ALGORITHMS
   - RESULTS
   - STABILITY
2. SUBSURFACE
   - INVERSE PROBLEM
   - ALGORITHMS
   - RESULTS
3. CONCLUSIONS

Outline

Navier Stokes Equations and Data

Write NSE as ODE in $H = \{u \in L^2(\mathbb{T}^2) \big| \nabla \cdot u = 0, \int_{\mathbb{T}^2} u\, dx = 0\}$:

$$\frac{dv}{dt} + \nu A v + B(v, v) = f, \qquad v(0) = u$$

$$v(t) = \Psi(u; t), \quad \Psi^{(j)}(u) := \Psi(u; j\tau)$$

Find u given noisy observations y_j:

$$y_j = H\Psi^{(j)}(u) + \eta_j,$$

$$\eta_j \sim \mathcal{N}(0, \Gamma)$$

$$Y_j = \{y_i\}_{i=1}^{j}.$$

Navier Stokes Equations and Data

Write NSE as ODE in $H = \left\{u \in L^2(\mathbb{T}^2) \middle| \nabla \cdot u = 0, \int_{\mathbb{T}^2} u\, dx = 0\right\}$:

$$\frac{dv}{dt} + \nu Av + B(v,v) = f, \qquad v(0) = u$$
$$v(t) = \Psi(u;t), \quad \Psi^{(j)}(u) := \Psi(u; j\tau)$$

Find u given noisy observations y_j:

$$y_j = H\Psi^{(j)}(u) + \eta_j,$$
$$\eta_j \sim \mathcal{N}(0,\Gamma)$$
$$Y_j = \{y_i\}_{i=1}^{j}.$$

Bayesian Formulation

- Prior $\mathbb{P}_0$ on u:

$$\mathbb{P}_0 = \mathcal{N}(m_0, C_0).$$

- Bayes formula:

$$\frac{\mathbb{P}(du|Y_J)}{\mathbb{P}_0(du)} \propto \mathbb{P}(Y_J|u).$$

- Here

$$\mathbb{P}(Y_J|u) \propto \exp\left(-\Phi(u; Y_J)\right)$$

$$\Phi(u; Y_J) = \frac{1}{2}\sum_{j=1}^{J}\left\|\Gamma^{-\frac{1}{2}}\left(y_j - \Psi^{(j)}(u)\right)\right\|^2$$

Bayesian Formulation

- Prior $\mathbb{P}_0$ on u:

$$\mathbb{P}_0 = \mathcal{N}(m_0, C_0).$$

- Bayes formula:

$$\frac{\mathbb{P}(du|Y_J)}{\mathbb{P}_0(du)} \propto \mathbb{P}(Y_J|u).$$

- Here

$$\mathbb{P}(Y_J|u) \propto \exp\big(-\Phi(u; Y_J)\big)$$

$$\Phi(u; Y_J) = \frac{1}{2}\sum_{j=1}^{J}\left\|\Gamma^{-\frac{1}{2}}\big(y_j - \Psi^{(j)}(u)\big)\right\|^2$$

Bayesian Formulation

- Prior $\mathbb{P}_0$ on u:

$$\mathbb{P}_0 = \mathcal{N}(m_0, C_0).$$

- Bayes formula:

$$\frac{\mathbb{P}(du|Y_J)}{\mathbb{P}_0(du)} \propto \mathbb{P}(Y_J|u).$$

- Here

$$\begin{aligned} \mathbb{P}(Y_J|u) \propto & \quad \exp\big(-\Phi(u; Y_J)\big) \\ \Phi(u; Y_J) = & \quad \tfrac{1}{2}\textstyle\sum_{j=1}^{J} \left\|\Gamma^{-\frac{1}{2}}\big(y_j - \Psi^{(j)}(u)\big)\right\|^2 \end{aligned}$$

Bayesian Formulation

- Prior $\mathbb{P}_0$ on u:

$$\mathbb{P}_0 = \mathcal{N}(m_0, C_0).$$

- Bayes formula:

$$\frac{\mathbb{P}(du|Y_J)}{\mathbb{P}_0(du)} \propto \mathbb{P}(Y_J|u).$$

- Here

$$\begin{aligned} \mathbb{P}(Y_J|u) \propto & \quad \exp\big(-\Phi(u; Y_J)\big) \\ \Phi(u; Y_J) = & \quad \tfrac{1}{2}\textstyle\sum_{j=1}^{J} \left\|\Gamma^{-\frac{1}{2}}\big(y_j - \Psi^{(j)}(u)\big)\right\|^2 \end{aligned}$$

Posterior Distribution

- pCN MCMC - Gold standard :

Propose $u^* = m_0 + (1-2\beta)^{\frac{1}{2}}(u^{(n-1)} - m_0) + \sqrt{2\beta}\mathcal{N}(0, C_0)$

$$u^{(n)} = \left\{ \begin{array}{cc} u^* & \text{with probability } 1 \wedge \exp\{\Phi(u^{(n-1)}) - \Phi(u^*)\} \\ u^{(n-1)} & else. \end{array} \right\}$$

- 4DVAR - Gaussian at MAP Estimator :

$$u \approx \mathcal{N}(m_0', C_0'), \qquad C_0' = \left(D^2\Phi(m_0') + \hat{C}_0^{-1}\right)^{-1}$$

$$m_0' = \mathrm{argmin}_u \left(\Phi(u) + \tfrac{1}{2}\|\hat{C}_0^{-\frac{1}{2}}(u - m_0)\|^2\right).$$

Posterior Distribution

- pCN MCMC - Gold standard :

Propose $u^* = m_0 + (1-2\beta)^{\frac{1}{2}}(u^{(n-1)} - m_0) + \sqrt{2\beta}\mathcal{N}(0, C_0)$

$$u^{(n)} = \left\{ \begin{array}{cc} u^* & \text{with probability } 1 \wedge \exp\{\Phi(u^{(n-1)}) - \Phi(u^*)\} \\ u^{(n-1)} & else. \end{array} \right\}$$

- 4DVAR - Gaussian at MAP Estimator :

$$u \approx \mathcal{N}(m_0', C_0'), \qquad C_0' = \left(D^2\Phi(m_0') + \hat{C}_0^{-1}\right)^{-1}$$

$$m_0' = \mathrm{argmin_u}\Big(\Phi(u) + \frac{1}{2}\|\hat{C}_0^{-\frac{1}{2}}(u - m_0)\|^2\Big).$$

Approximate Gaussian Filters I

- Impose the Gaussian approximation:

$$\mathbb{P}(v_{j-1}|Y_{j-1}) = \Psi^{j-1} \star \mathbb{P}(u|Y_{j-1}) \approx N(m_{j-1}, C_{j-1})$$
$$\mathbb{P}(v_j|Y_{j-1}) = \Psi^{j} \star \mathbb{P}(u|Y_{j}) \approx N(\Psi(m_{j-1}), \widehat{C}_j).$$

- Find update rule: $(m_{j-1}, C_{j-1}) \mapsto (m_j, C_j)$.

$$m_j = (I - K_j H)\Psi(m_{j-1}) + K_j y_j,$$
$$K_j = \widehat{C}_j H^\top (H \widehat{C}_j H^\top + \Gamma)^{-1},$$
$$C_j = (I - K_j H)\widehat{C}_j.$$

Approximate Gaussian Filters I

- Impose the Gaussian approximation:

$$\begin{aligned} \mathbb{P}(v_{j-1}|Y_{j-1}) = & \quad \Psi^{j-1} \star \mathbb{P}(u|Y_{j-1}) \approx N(m_{j-1}, C_{j-1}) \\ \mathbb{P}(v_j|Y_{j-1}) = & \quad \Psi^{j} \star \mathbb{P}(u|Y_j) \approx N(\Psi(m_{j-1}), \widehat{C}_j). \end{aligned}$$

- Find update rule: $(m_{j-1}, C_{j-1}) \mapsto (m_j, C_j)$.

$$\begin{aligned} m_j = & \quad (I - K_j H)\Psi(m_{j-1}) + K_j y_j, \\ K_j = & \quad \widehat{C}_j H^\top (H\widehat{C}_j H^\top + \Gamma)^{-1}, \\ C_j = & \quad (I - K_j H)\widehat{C}_j. \end{aligned}$$

Approximate Gaussian Filters II

In all cases

$$C_j^{-1} = \widehat{C}_j^{-1} + H^\top \Gamma^{-1} H$$

- **3DVAR** :

$$C_j \equiv C_0.$$

- **FDF**: For C chosen from Gaussian SPDE parameter fit:

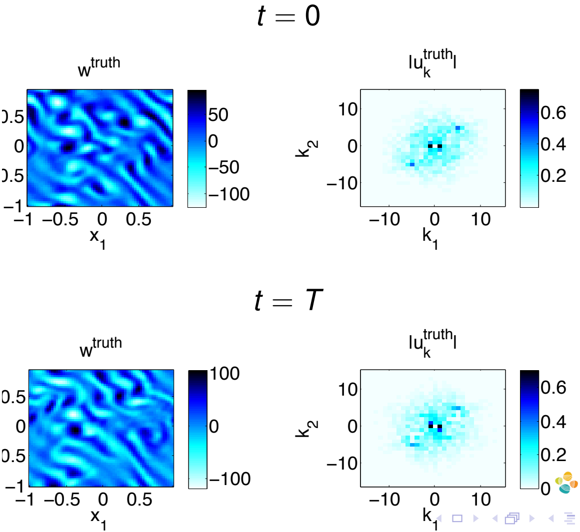
$$C_j \equiv C.$$

- **(LR)ExKF**: (Low rank approximation of):

$$\widehat{C}_j = D\Psi(m_{j-1}) C_{j-1} D\Psi(m_{j-1})^\top.$$

- **EnKF**: Particle approximations for m_j and C_j.

Truth



Relative Error in Mean c/w Posterior

Mean

method	e_{mean}
4DVAR($t = 0$)	0.000731491
4DVAR($t = T$)	0.00130112
3DVAR	0.0634553
FDF	0.165732
LRExKF	0.00614573
EnKF	0.035271

Relative Error in Variance c/w Posterior

Variance

method	$e_{variance}$
4DVAR($t = 0$)	0.0932748
4DVAR($t = T$)	0.220154
3DVAR	6.34057
FDF	28.9155
LRExKF	0.195101
EnKF	0.274428

Accuracy Theorem (3DVAR)

- Recall true signal $v_j = \Psi^{(j)}(u) = v(j\tau)$.
- Define $\sup_{j\geq 1} \|\xi_j\| = \epsilon$.

Theorem (BLLMSS11)

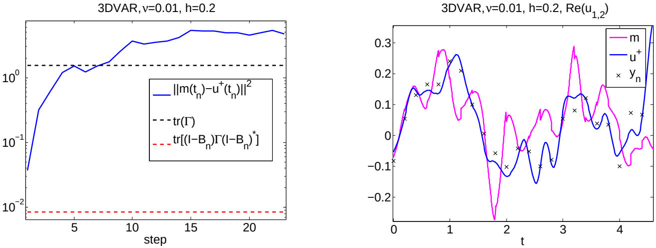
Assume that:

- $m_0 \in B_{\mathrm{V}}(0, R)$;
- *Enough low wavenumbers observed, and τ small enough.*
- $\Gamma = \eta \widehat{C}$;

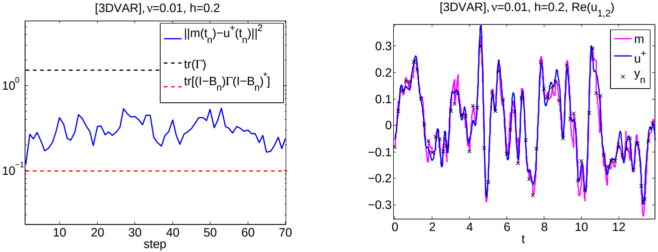
Then $\exists\ \eta_c = \eta_c(r)$, $r \in (0,1)$, and $c \in (0, \infty)$ such that, for all $\eta < \eta_c$,

$$\|\hat{m}_j - v_j\| \leq r^j \|\hat{m}_0 - v_0\| + c\epsilon.$$

Inaccurate


3DVAR, ν=0.01, h=0.2
10^0
10^−1
10^−2
5
10
15
20
step
||m(t_n)−u^+(t_n)||^2
tr(Γ)
tr[(I−B_n)Γ(I−B_n)^*]
3DVAR, ν=0.01, h=0.2, Re(u_1,2)
0.3
0.2
0.1
0
−0.1
−0.2
0
1
2
3
4
t
m
u^+
y_n

Accurate


[3DVAR], ν=0.01, h=0.2
||m(t_n)−u^+(t_n)||^2
tr(Γ)
tr[(I−B_n)Γ(I−B_n)^*]
10^0
10^−1
10
20
30
40
50
60
70
step
[3DVAR], ν=0.01, h=0.2, Re(u_1,2)
m
u^+
y_n
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
0
2
4
6
8
10
12
t

Evaluation wrt gold standard, fixed Interval ($T = J\tau$): Relative Error c/w Posterior: stabilized

From [LS11]:

method	e_{mean}	$e_{variance}$
3DVAR	0.458527	1.8214
[3DVAR]	0.27185	6.62328
LRExKF	0.632448	0.4042
[LRExKF]	0.201327	11.2449
EnKF	0.901703	0.554611
[EnKF]	0.169262	4.07238
FDF	0.189832	11.4573

Outline

Two-phase incompressible flow and Data

$$-\nabla \cdot \left[\lambda(s) e^{u} \nabla p\right] = f_1(p, s) \quad \text{in } D \times [0, T],$$
$$\psi \frac{\partial s}{\partial t} - \nabla \cdot \left[\lambda_w(s) e^{u} \nabla p\right] = f_2(p, s) \quad \text{in } D \times [0, T].$$

h pointwise measurements of total flow rate at production wells and bottom-hold pressure at injection wells.

$$y_j = h_0(p(j\tau), s(j\tau)) + \eta_j,$$
$$\eta_j \sim \mathcal{N}(0, \Gamma)$$
$$Y_j = \{y_i\}_{i=1}^{j}.$$

ertainty

Two-phase incompressible flow and Data

$$
\begin{aligned}
-\nabla\cdot\left[\lambda(s)e^{u}\nabla p\right] &= f_1(p,s) \quad \text{in } D\times[0,T],\\
\psi\frac{\partial s}{\partial t}-\nabla\cdot\left[\lambda_w(s)e^{u}\nabla p\right] &= f_2(p,s) \quad \text{in } D\times[0,T].
\end{aligned}
$$

h pointwise measurements of total flow rate at production wells and bottom-hold pressure at injection wells.

$$
\begin{aligned}
y_j &= h_0(p(j\tau), s(j\tau)) + \eta_j,\\
\eta_j &\sim \mathcal{N}(0,\Gamma)\\
Y_j &= \{y_i\}_{i=1}^{j}.
\end{aligned}
$$

Filtering

Augmented System Dynamics

Define

$$z = \begin{pmatrix} u \\ p \\ s \end{pmatrix}, \qquad \Psi(z) = \begin{pmatrix} u \\ \psi_p(u,p,s) \\ \psi_s(u,p,s) \end{pmatrix}, \qquad z_{j+1} = \Psi(z_j),$$

and then data is, for $h(\cdot) = (0, h_0(\cdot))$,

$$y_j = h(z_j) + \eta_j \qquad \text{where } \eta_j \sim N(0, \Gamma).$$

State Estimation

Aim to recover posterior distribution on z_j given $Y_j = \{y_i\}_{i=1}^{j}$. In particular $u_j | Y_j$.

Posterior Distribution

- MCMC - Gold standard :

Propose $u^* = m_0 + (1-2\beta)^{\frac{1}{2}}(u^{(n-1)} - m_0) + \sqrt{2\beta}\mathcal{N}(0, C_0)$

$$u^{(n)} = \left\{ \begin{array}{cc} u^* & \text{with probability } 1 \wedge \exp\{\Phi(u^{(n-1)}) - \Phi(u^*)\} \\ u^{(n-1)} & \mathit{else}. \end{array} \right\}$$

- 4DVAR - MAP Estimator :

$$u \approx \mathcal{N}(m_0', C_0'), \qquad C_0' = \left(D^2\Phi(m_0') + \hat{C}_0^{-1}\right)^{-1}$$

$$m_0' = \mathrm{argmin_u}\left(\Phi(u) + \tfrac{1}{2}\|\hat{C}_0^{-\frac{1}{2}}(u - m_0)\|^2\right).$$

Posterior Distribution

- MCMC - Gold standard :

Propose $u^* = m_0 + (1 - 2\beta)^{\frac{1}{2}}(u^{(n-1)} - m_0) + \sqrt{2\beta}\mathcal{N}(0, C_0)$

$$u^{(n)} = \left\{ \begin{array}{ll} u^* & \text{with probability } 1 \wedge \exp\{\Phi(u^{(n-1)}) - \Phi(u^*)\} \\ u^{(n-1)} & \textit{else}. \end{array} \right\}$$

- 4DVAR - MAP Estimator :

$$u \approx \mathcal{N}(m_0', C_0'), \qquad C_0' = \left(D^2\Phi(m_0') + \hat{C}_0^{-1}\right)^{-1}$$

$$m_0' = \mathrm{argmin}_u \left(\Phi(u) + \frac{1}{2}\|\hat{C}_0^{-\frac{1}{2}}(u - m_0)\|^2\right).$$

Approximate Gaussian Filters

- EnKF (with and without localization)
- Square-root EnKF (with and without localization)
- Randomized Maximum Likelihood (RML)

The posterior distribution is approximated by an ensemble $\{\hat{u}^n\}_{n=1}^N$ generated by solving

$$\hat{u}^{(n)} = \operatorname{argmin}_u \Big(\Phi(u; Y_J^{(n)}) + \frac{1}{2}||C^{-1/2}(u - u^{(n)})||^2 \Big)$$

where

$$\begin{aligned} u^{(n)} &\sim \mathbb{P}_0, \\ Y_J^{(n)} &= Y_J + \eta_J^{(n)}, \qquad \eta_J^{(n)} \sim \mathcal{N}(0, \oplus_{j=1}^J \Gamma) \end{aligned} \tag{1}$$

Approximate Gaussian Filters

- EnKF (with and without localization)
- Square-root EnKF (with and without localization)
- Randomized Maximum Likelihood (RML)

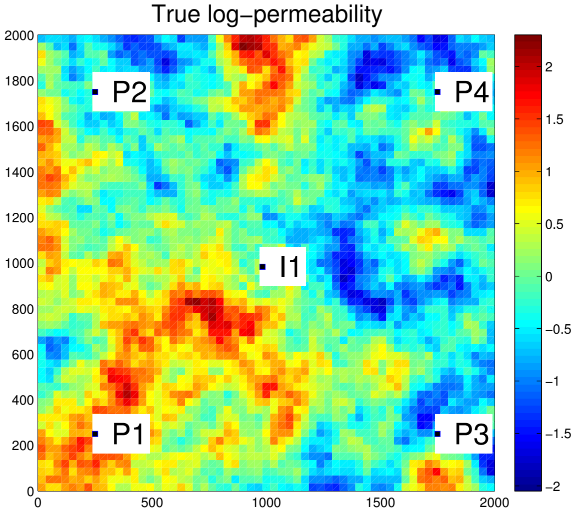
The posterior distribution is approximated by an ensemble $\{\hat{u}^n\}_{n=1}^N$ generated by solving

$$\hat{u}^{(n)} = \operatorname{argmin}_u \Big(\Phi(u; Y_J^{(n)}) + \tfrac{1}{2} ||C^{-1/2}(u - u^{(n)})||^2 \Big)$$

where

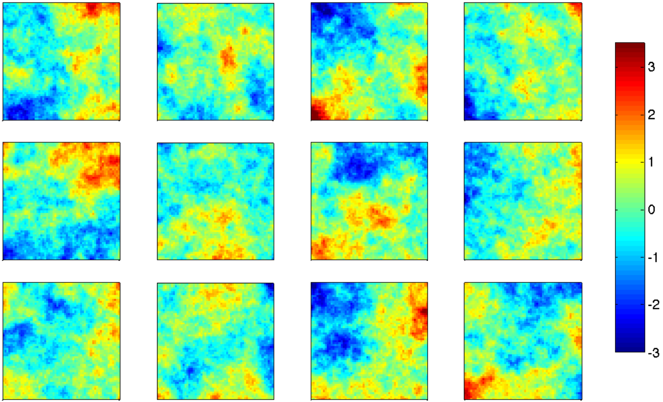
$$\begin{aligned} u^{(n)} &\sim \mathbb{P}_0, \\ Y_J^{(n)} &= Y_J + \eta_J^{(n)}, \qquad \eta_J^{(n)} \sim \mathcal{N}(0, \oplus_{j=1}^J \Gamma) \end{aligned} \quad (1)$$

The Truth


True log−permeability
P1
P2
P3
P4
I1
0
200
400
600
800
1000
1200
1400
1600
1800
2000
500
1000
1500
2000
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2

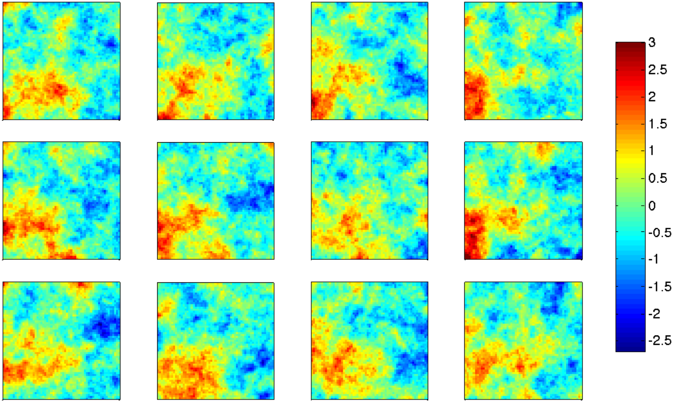
Prior Geology

Monte Carlo samples of the prior distribution $\mathbb{P}_0$ look like:



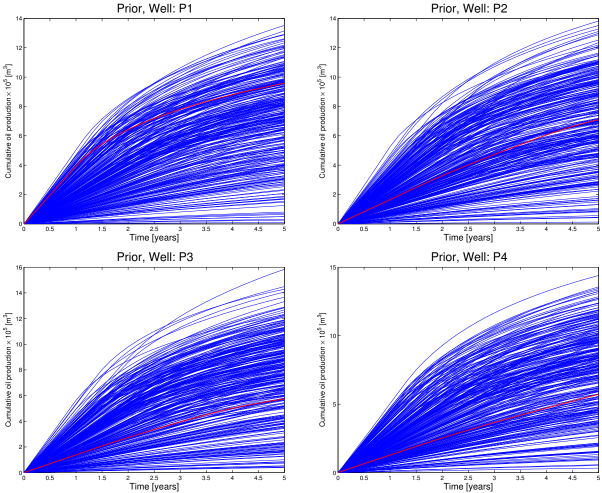
Posterior Geology

MCMC samples of the posterior $\mathbb{P}(\cdot|y)$ look like:



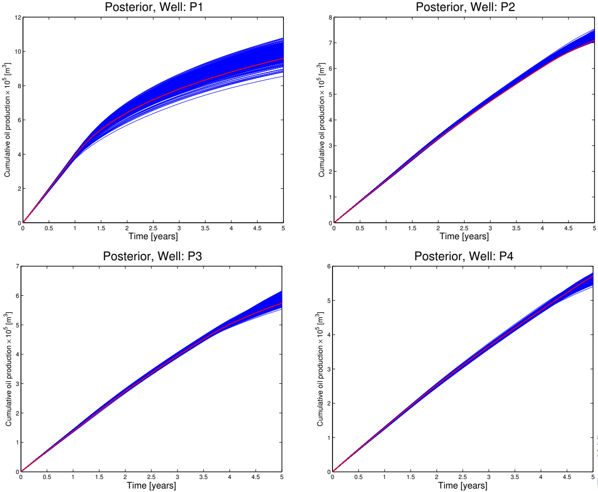
Uncertainty Under The Prior

Large uncertainty in cumulative oil production:



Uncertainty Under The Posterior

Reduction of the uncertainty via data:



Evaluation wrt gold standard, fixed interval

From [ILS13]:

Method	e_{mean}	$e_{variance}$	Cost (FM)
MCMC	0.000	0.000	5.5×10^7
MAP	0.277	0.143	6.0×10^1
RML ($N_e = 50$)	0.235	0.415	3.0×10^3
EnKF ($N_e = 50$)	0.834	0.391	5.0×10^1
EnKF (loc, $N_e = 50$)	0.563	0.252	5.0×10^1
EnKF ($N_e = 8000$)	0.337	0.216	8.0×10^3

FM: Forward Model evaluations.

Outline

Conclusions

- Approximate filters:
  - can reproduce the posterior mean accurately;
  - may not reproduce covariance accurately;
  - can exhibit instability (e.g. on longer time-intervals).
- This instability can cause loss of accuracy in even mean prediction;
- Filter stabilization, via variance inflation can be used to ameliorate instability. Intrusive: changes the covariance.
- Spatial localization techniques can improve accuracy of ensemble method, hence ameliorate instability: mean-field EnKF more stable than finite ensemble.

Conclusions

- Approximate filters:
  - can reproduce the posterior mean accurately;
  - may not reproduce covariance accurately;
  - can exhibit instability (e.g. on longer time-intervals).
- This instability can cause loss of accuracy in even mean prediction;
- Filter stabilization, via variance inflation can be used to ameliorate instability. Intrusive: changes the covariance.
- Spatial localization techniques can improve accuracy of ensemble method, hence ameliorate instability: mean-field EnKF more stable than finite ensemble.

Conclusions

- Approximate filters:
  - can reproduce the posterior mean accurately;
  - may not reproduce covariance accurately;
  - can exhibit instability (e.g. on longer time-intervals).
- This instability can cause loss of accuracy in even mean prediction;
- Filter stabilization, via variance inflation can be used to ameliorate instability. Intrusive: changes the covariance.
- Spatial localization techniques can improve accuracy of ensemble method, hence ameliorate instability: mean-field EnKF more stable than finite ensemble.

Conclusions

- Approximate filters:
  - can reproduce the posterior mean accurately;
  - may not reproduce covariance accurately;
  - can exhibit instability (e.g. on longer time-intervals).
- This instability can cause loss of accuracy in even mean prediction;
- Filter stabilization, via variance inflation can be used to ameliorate instability. Intrusive: changes the covariance.
- Spatial localization techniques can improve accuracy of ensemble method, hence ameliorate instability: mean-field EnKF more stable than finite ensemble.

Extensions

- Use this technique to assess other algorithms, e.g. particle filters.
- Develop better/more efficient MCMC algorithms.
- Challenge: improve estimates of variance or, more optimistically, general observables.
- Decompose state-space: use e.g. ExKF in the (changing) stable subspace, and a more accurate method in the unstable subspace.

Extensions

- Use this technique to assess other algorithms, e.g. particle filters.
- Develop better/more efficient MCMC algorithms.
- Challenge: improve estimates of variance or, more optimistically, general observables.
- Decompose state-space: use e.g. ExKF in the (changing) stable subspace, and a more accurate method in the unstable subspace.

Extensions

- Use this technique to assess other algorithms, e.g. particle filters.
- Develop better/more efficient MCMC algorithms.
- Challenge: improve estimates of variance or, more optimistically, general observables.
- Decompose state-space: use e.g. ExKF in the (changing) stable subspace, and a more accurate method in the unstable subspace.

Extensions

- Use this technique to assess other algorithms, e.g. particle filters.
- Develop better/more efficient MCMC algorithms.
- Challenge: improve estimates of variance or, more optimistically, general observables.
- Decompose state-space: use e.g. ExKF in the (changing) stable subspace, and a more accurate method in the unstable subspace.

References 1

- **[LS12]**: K.J.H.Law and A.M.Stuart. "Evaluating Data Assimilation Algorithms." Monthly Weather Review 140, 3757-3782 (2012)
- **[ILS13]**: M.A. Iglesias, K.J.H.Law and A.M.Stuart. "Evaluation of Gaussian approximations for data assimilation in reservoir models." Computational Geosciences, 17:5, 851-885 (2013).
- **MCMC**: S. L. Cotter, G. O. Roberts, A. M. Stuart, D. White. "MCMC methods for functions: modifying old algorithms to make them faster." Statistical Science 28 (2013) 424–446.
- **Stability**: C. E. A. Brett, K. F. Lam, K. J. H. Law, D. S. McCormick, M. R. Scott and A. M. Stuart. "Accuracy and Stability of Filters for Dissipative PDEs." Physica D 245, 1, 34-45 (2013).

References 2

- **4DVAR**: Bennett A. "Inverse modeling of the ocean and atmosphere." Cambridge. (2002).
- **3DVAR**: Kalnay E. "Atmospheric modeling, data assimilation, and predictability." Cambridge Univ Pr. (2003).
- **FDF** Majda A, Harlim J, Gershgorin B. "Mathematical strategies for filtering turbulent dynamical systems." Dynamical Systems **27**(2): 441-486. (2010)
- **ExKF** Jazwinski A. "Stochastic processes and filtering theory." Academic Pr. (1970).
- **EnKF** Evensen G. "Data assimilation: the ensemble kalman filter." Springer Verlag. (2009).