Uncertainty quantification in inverse problems with multiscale models. Applications to porous media

Yalchin Efendiev

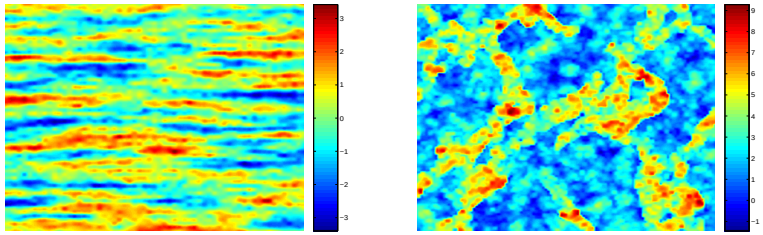
Department of Mathematics & ISC

Texas A&M University

Collaborators: A. Datta-Gupta (TAMU), P. Dostert (Arizona), T. Hou (Caltech), B. Mallick (TAMU), A. Mondal (TAMU)

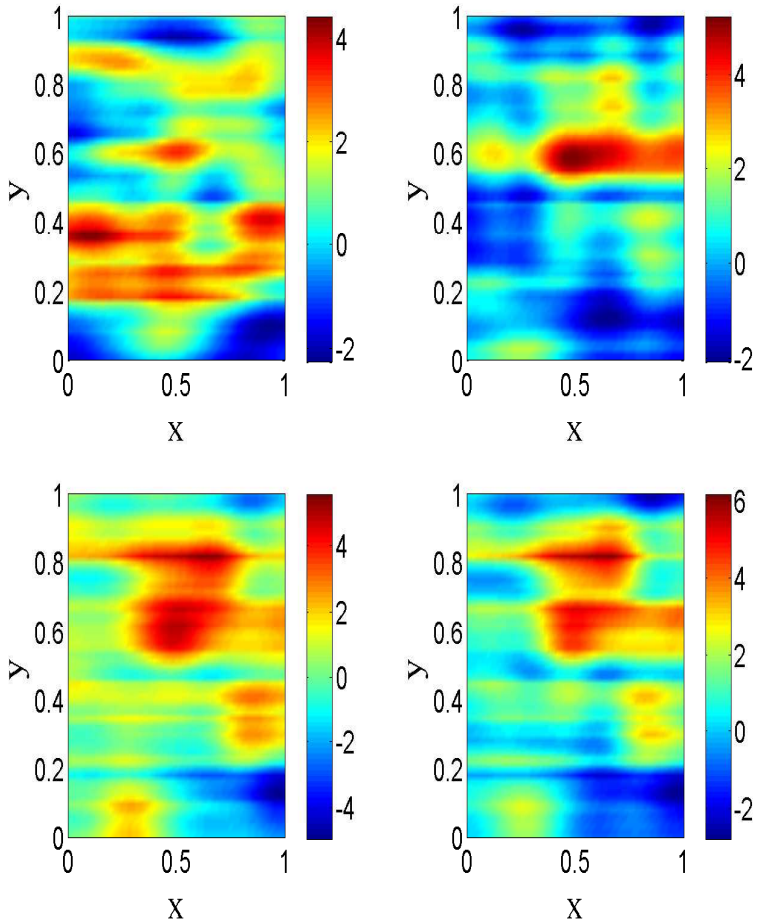
Introduction

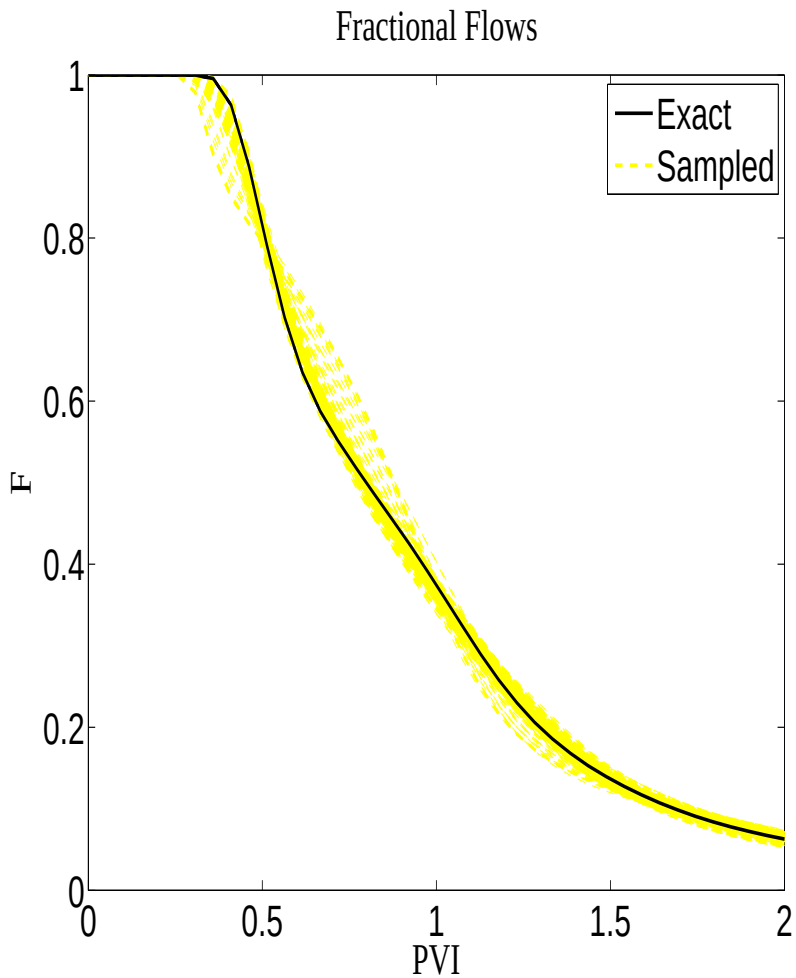
- We consider a problem of characterizing subsurface properties given "coarse-scale" measurements.
- Subsurface properties are highly heterogeneous with uncertainties at "small" scales that can affect the measurement results.



- The relation between the input (subsurface properties in this talk) and the output (e.g., oil production rate) is highly nonlinear.
- Inverse problem in a Bayesian framework (uncertainty quantification in inverse problems, ...): $P(k|D) \propto P(D|k)P(k)$, where $k(x)$ represents the subsurface property (permeability, e.g.).

Illustration





Prototypical model

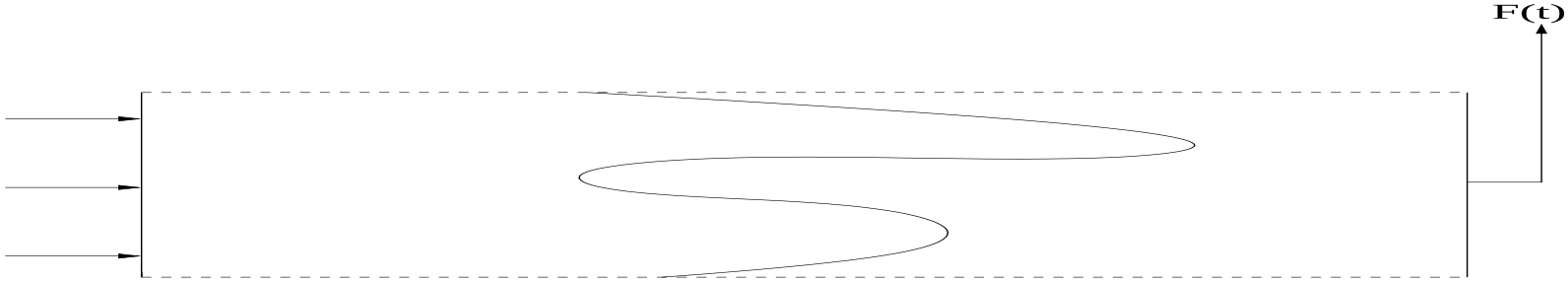
We consider two-phase flow in a heterogeneous porous formation under the assumption that the displacement is dominated by viscous effects.

$$\nabla \cdot (\lambda(S) k \nabla p) = h,$$

$$\frac{\partial S}{\partial t} + v \cdot \nabla f(S) = 0, \; v = -\lambda(S) k \nabla p.$$

Measured data:

$$F(t) = \frac{\int_{out} v f(S) dl}{\int_{out} v dl}$$

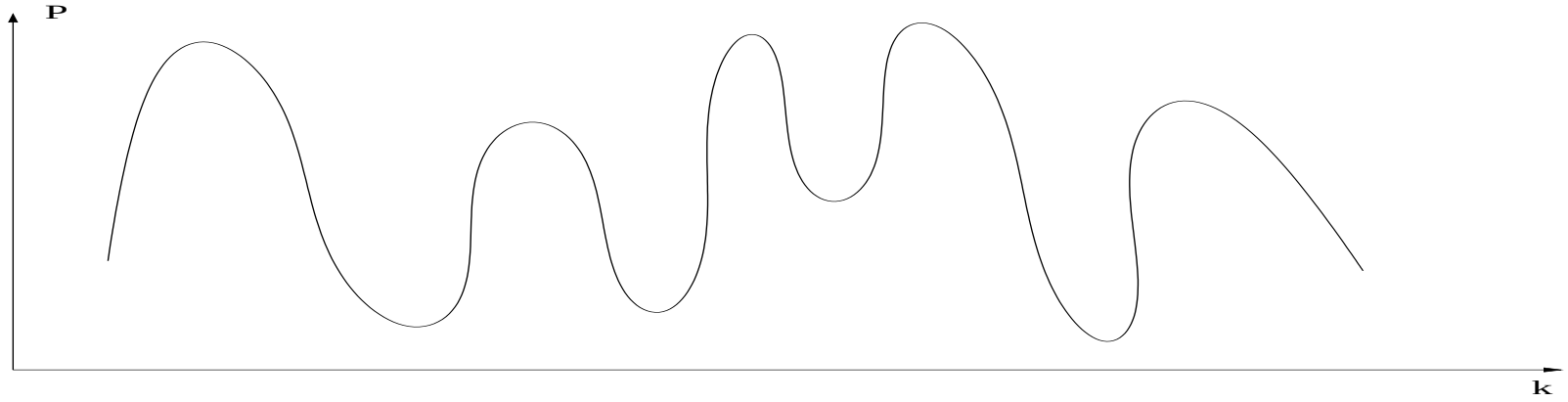


Problem setting

- Given the fractional flow information (integrated response) $F(t)$ and some precision, we would like to sample k from $P(k|F)$.
- From Bayes theorem
$$P(k|F) \propto P(F|k)P(k).$$
- Here $P(k)$ is the prior information, $P(F|k)$ is the likelihood and assumed given by $P(F|k) = \exp(-\frac{\|F_k(t)-F^{obs}(t)\|^2}{\sigma_f^2})$.
- Priors can be (1) $P(k) = \exp(-\frac{\|\log(k)-\log(k_{obs})\|^2}{\sigma_k^2})$ (2) covariance matrix is given (3) with unknown parameters in the covariance matrix (4) ...

Difficulties

- $\pi(k) = P(k|F)$ can be multi-modal and high dimensional.
- $\pi(k) = P(k|F)$ is not given analytically and involves the solution of nonlinear pde system.



Sampling

Algorithm (Metropolis-Hastings MCMC)

- Step 1. At k_n generate k from $q(k|k_n)$.
- Step 2. Accept k as a sample with probability

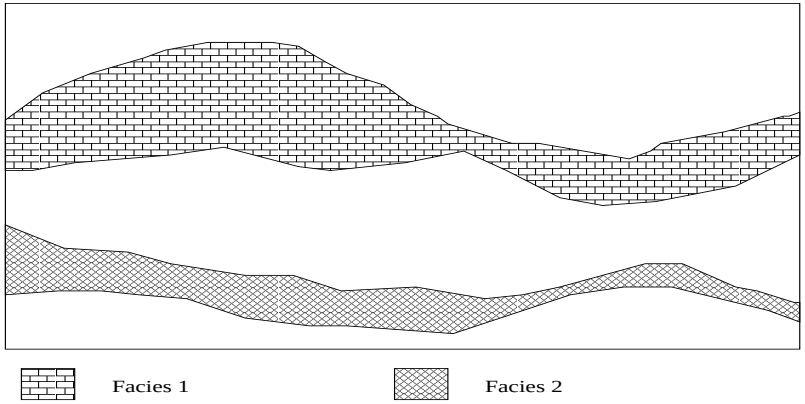
$$p(k_n, k) = \min\left(1, \frac{q(k_n|k)\pi(k)}{q(k|k_n)\pi(k_n)}\right),$$

  i.e. $k_{n+1} = k$ with probability $p(k_n, k)$, and $k_{n+1} = k_n$ with probability $1 - p(k_n, k)$.

Here $\pi(k)$ is the distribution we would like to sample.

- Direct (full) MCMC simulations are usually prohibitively expensive, because each proposal requires a fine-scale computation.
- We use algorithms where the proposal distribution is modified using coarse-scale spatial models.

Priors



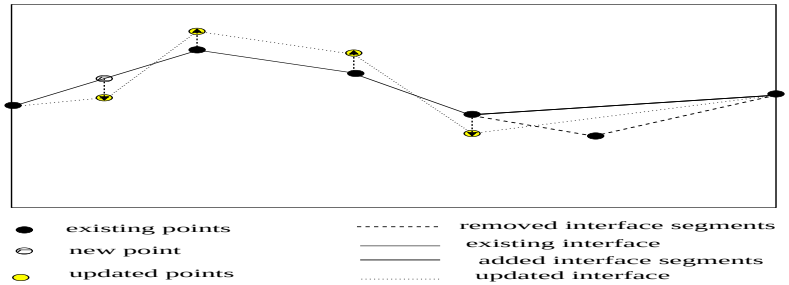
- Interface modeling with level sets ($\tau = const$)

$$\frac{\partial \tau}{\partial s} + w \cdot \nabla \tau = 0,$$

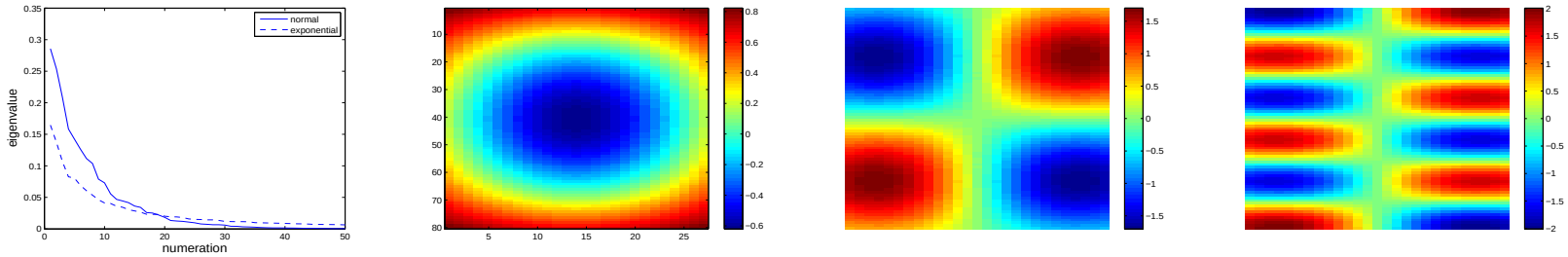
  where w is a random vector field. In simulations, we assume that the direction of the velocity field is fixed.

- Within each region, two-point correlation based log-permeability fields are used, i.e., $R(x, y) = E(Y(x, \omega)Y(y, \omega))$ is given, where $Y(x, \omega) = \log(k(x, \omega))$.

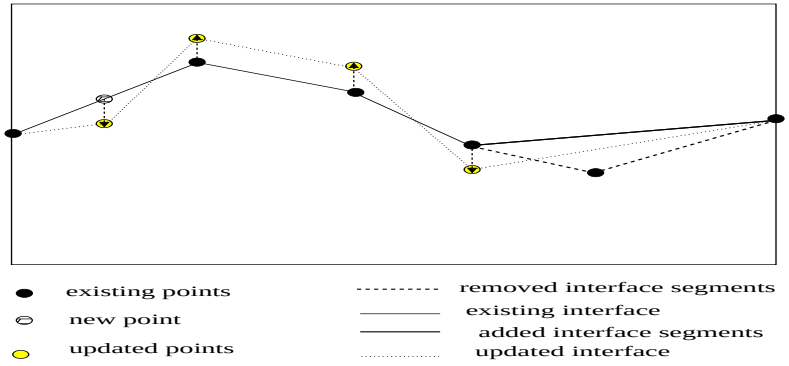
Priors



- Discretization of the interface is done with variable number of points that have unknown locations.
- Discretization of permeability field in each region is done with Karhunen-Loéve expansion, $Y(x,\omega) = \sum \theta_i(\omega)\sqrt{\lambda_i}\phi_i(x)$.
- Permeability is conditioned at some locations (well locations).



Proposal distributions



An illustration of the Birth, Death and Jump process in Reversible Jump MCMC for an interface.

- Permeability within a region ($k(x) \rightarrow Y(x) = \log(k(x)) \rightarrow \theta_i(\omega)$) is proposed either using random walk sampler or Langevin.
- Forward models based on pde's allow computing the gradients of the posterior.
- Langevin proposals are derived from the solution of $dk(\tau) = \frac{1}{2}\nabla \log \pi(k(\tau))d\tau + dW_\tau$. It is given by

$$Y = k_n + \frac{\Delta\tau}{2}\nabla \log \pi(k_n) + \sqrt{\Delta\tau}\epsilon_n.$$

Langevin Algorithms

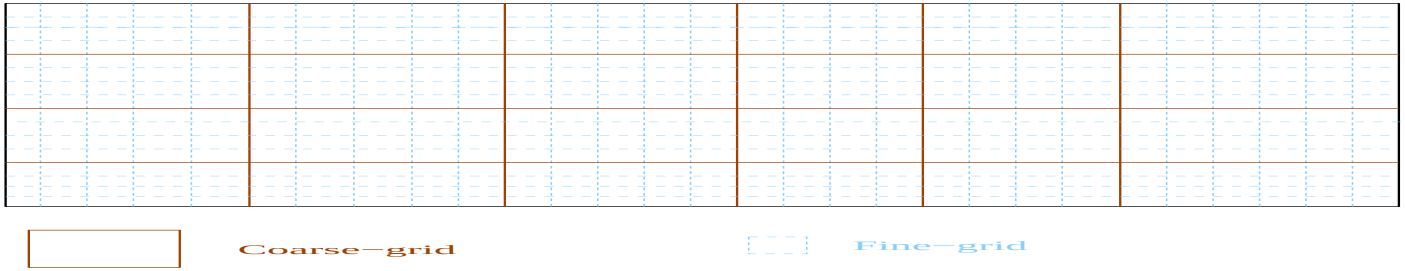
The transition distribution of the proposal is

$$q(Y|k_n) \propto \exp\left(-\frac{\|Y - k_n - \frac{\Delta\tau}{2}\nabla \log \pi(k_n)\|^2}{2\Delta\tau}\right),$$

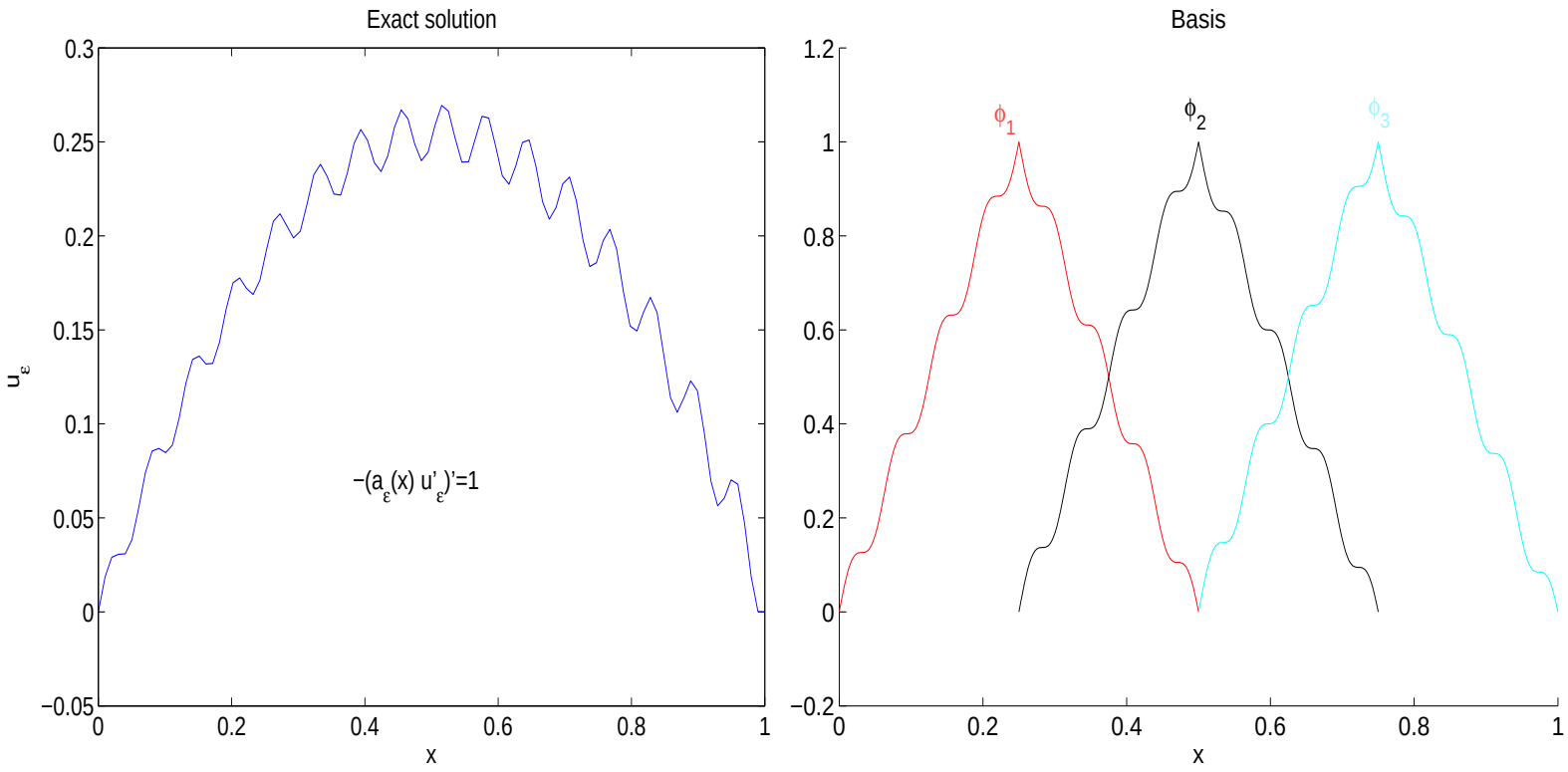
$$q(k_n|Y) \propto \exp\left(-\frac{\|k_n - Y - \frac{\Delta\tau}{2}\nabla \log \pi(Y)\|^2}{2\Delta\tau}\right).$$

Multiscale spatio-temporal models

- In multiscale simulations, we attempt to represent the solution on a coarse grid.
- Multiscale basis functions (or similarly upscaled quantities) are pre-computed on a coarse grid. The objective is to solve the problems on a coarse grid.
- Example: the solution $p(x)$ defined on 100×100 fine grid is written as $p(x) \approx \sum_{i=1}^{50} c_i \phi_i(x)$. Multiscale basis functions are computed such that it captures the local variations of the solution.
- Ensemble level multiscale methods: $p(x, \omega) \approx \sum_{i=1}^{50} c_i \phi_i(x, \omega)$.
- Some details can be found, e.g., Efendiev and Hou, Multiscale finite element methods. Theory and Applications. Springer, 2009

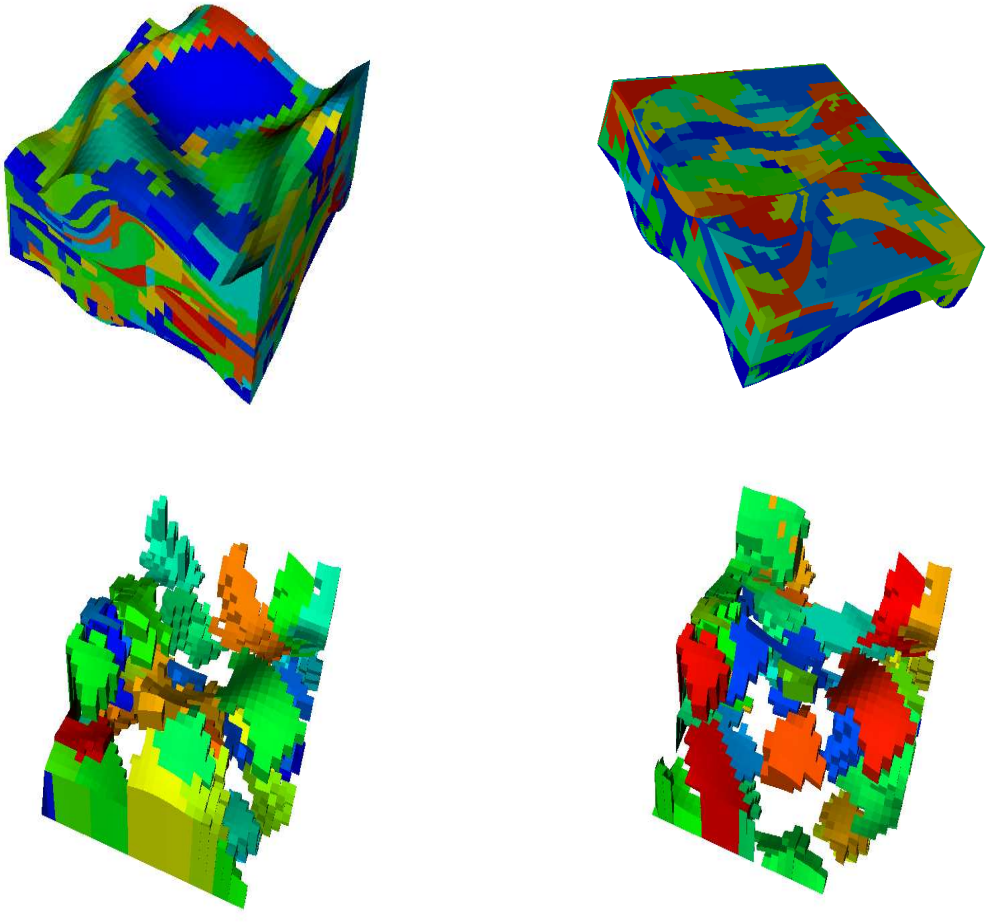


A simple example



$$(a_\epsilon(x)u')' = -1,\ u(0) = u(1) = 0,\ a_\epsilon(x) = 1/(2 + 1.99\cos(x/\epsilon)),\ \epsilon = 0.01.$$

Gridding

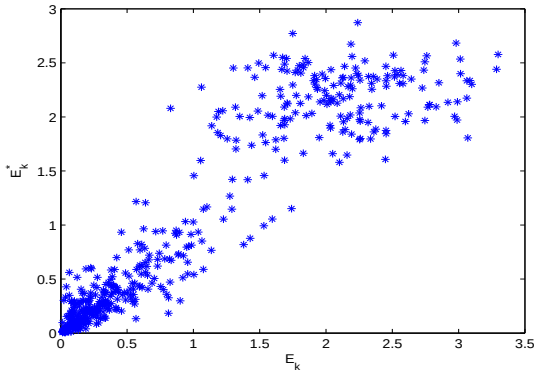


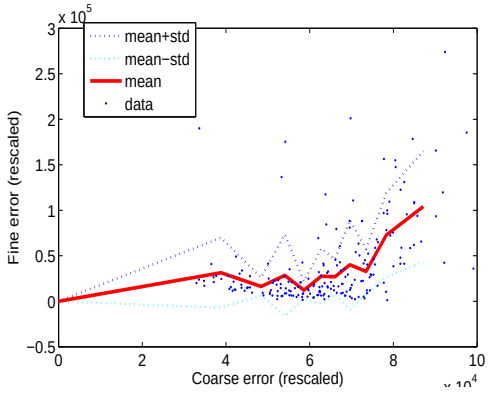
The use of multiscale models

- $\pi(k) = P(k|F) \propto P(F|k)P(k)$, where $P(F|k) = \exp(-\frac{\|F_k(t) - F^{obs}(t)\|^2}{\sigma_f^2})$.
- Based on off-line computations (fine vs. coarse), one can determine a statistical relation

$$\|F_k(t) - F_{obs}(t)\| \approx G(\|F_k^*(t) - F^{obs}(t)\|) + Noise.$$

  (error modeling, e.g., J. Glimm, M. Christie, ...)
- A simplest relation is a linear relation $G(x) = ax$.
- Introduce $\pi^*(k) \propto \exp(-\frac{G(\|F_k^*(t) - F^{obs}(t)\|)^2}{\sigma_f^2})P(k)$.





Preconditioned coarse-gradient Langevin algorithm

(1) Make a proposal (birth/death for an interface point; move the interface; populate spatial features based on "coarse gradients"); (2) run the coarse-scale simulation code and check the "appropriateness" of the sample; (3) run the "fine-scale" simulation.

- Step 1. At k_n, generate a trial proposal Y from the coarse Langevin distribution $q^*(Y|k_n)$.
- Step 2. Take the proposal k as

$$k = \begin{cases} \tilde{k} & \text{with probability } \alpha_p(k_n, \tilde{k}), \\ k_n & \text{with probability } 1 - \alpha_p(k_n, \tilde{k}), \end{cases}$$

where

$$\alpha_p(k_n, \tilde{k}) = \min\left(1, \frac{q(k_n|\tilde{k})\pi^*(\tilde{k})}{q(\tilde{k}|k_n)\pi^*(k_n)}\right).$$

Preconditioned coarse-gradient Langevin algorithm

- If we are at Birth Step

$$\alpha_p(k_n, \tilde{k}) = \min\Big\{1, \underbrace{\frac{P^*(F_{obs}|\tilde{k})}{P^*(F_{obs}|k_n)}}_{\text{likelihood ratio}} \times \underbrace{\frac{P(\tau)P(\theta)P(x^{loc}|m_n+1)P(m_n+1)}{P(\tau_n)P(\theta_n)P(x_n^{loc}|m_n)P(m_n)}}_{\text{prior ratio}}$$

$$\times \underbrace{\frac{q_\theta(\theta_n|\theta)p^{del}_{m_n+1}}{q_\theta(\theta|\theta_n)p^{add}_{m_n}q_{m_n m_n+1}(u|x_n^{loc})}}_{\text{proposal ratio}} \times \underbrace{\left|\frac{\partial g_{m_n m}(x_n^{loc}, u)}{\partial x_n^{loc}\partial u}\right|}_{\text{Jacobian}}\Big\}.$$

.....

- Step 3. Accept k as a sample with probability

$$\rho(k_n, k) = \min\left(1, \frac{Q(k_n|k)\pi(k)}{Q(k|k_n)\pi(k_n)}\right),$$

where Q is the effective proposal distribution.

Convergence of modified Markov Chain

Denote

$$\mathcal{E} = \{k;\ \pi(k) > [0]\},$$

$$\mathcal{E}^* = \{k;\ \pi^*(k) > [0]\},$$

$$\mathcal{D} = \{k;\ q(k|k_n) > [0] \text{ for some } k_n \in \mathcal{E}\},$$

To sample from $\pi(k)$ correctly, it is necessary that $\mathcal{E} \subseteq \mathcal{E}^*$. Otherwise, there will exist a subset $A \subset (\mathcal{E} \setminus \mathcal{E}^*)$ such that

$$\pi(A) = \int_A \pi(x)dx > 0 \qquad \text{and} \qquad \pi^*(A) = \int_A \pi^*(x)dx = 0.$$

As a result, the chain $\{k_n\}$ will never visit (sample from) A since the element of A will never be accepted for fine-scale run in Step 2. For the same reason, we should require that $\mathcal{E} \subseteq \Omega$.

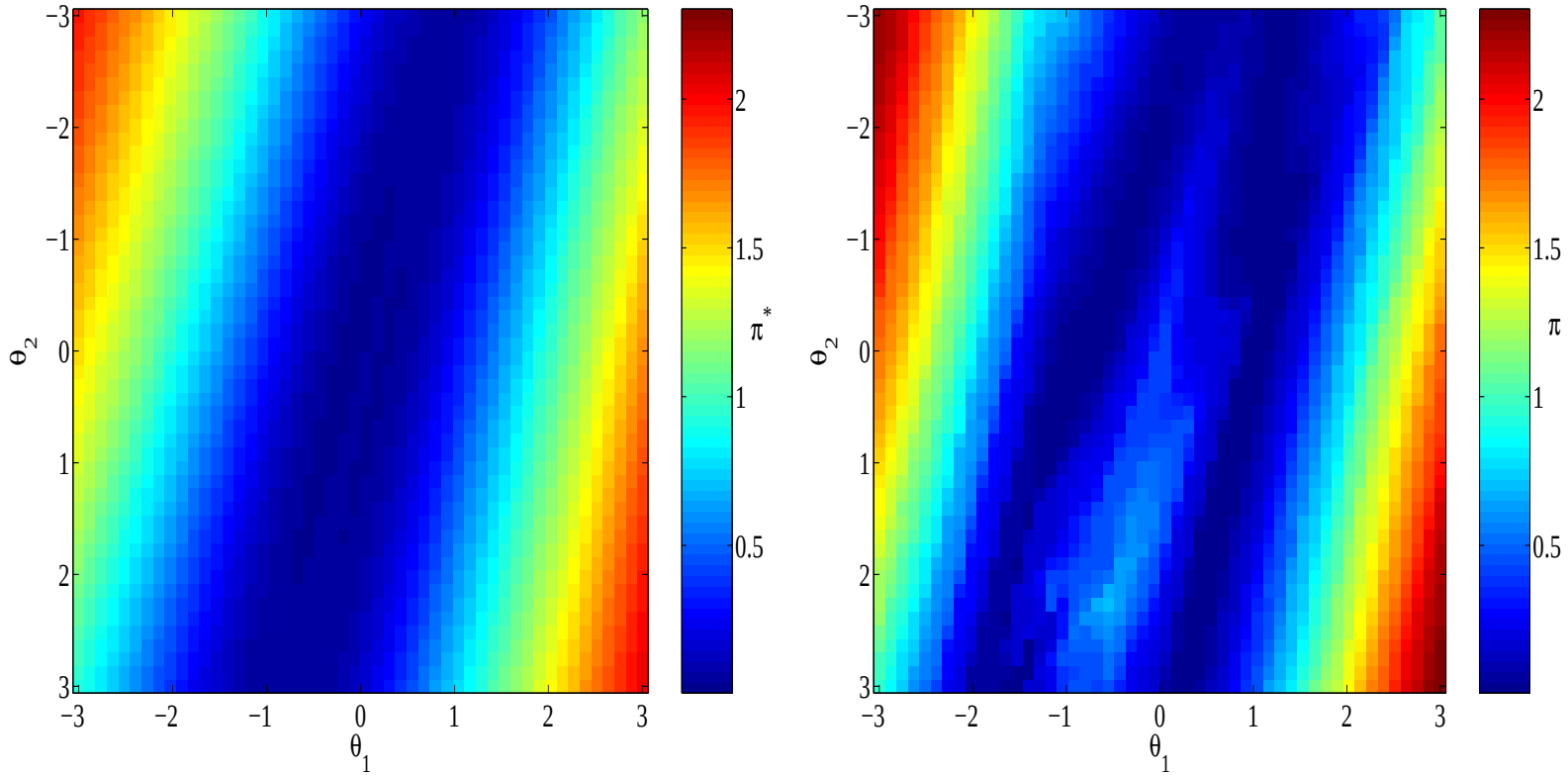
A remark

- If $\pi^*(k)$ is a smooth surface, then instead of using $\pi^*(k)$, it can be interpolated employing sparse collocation techniques

$$\tilde{\pi}^* = \sum_i \pi^*(k_i) L_i(k),$$

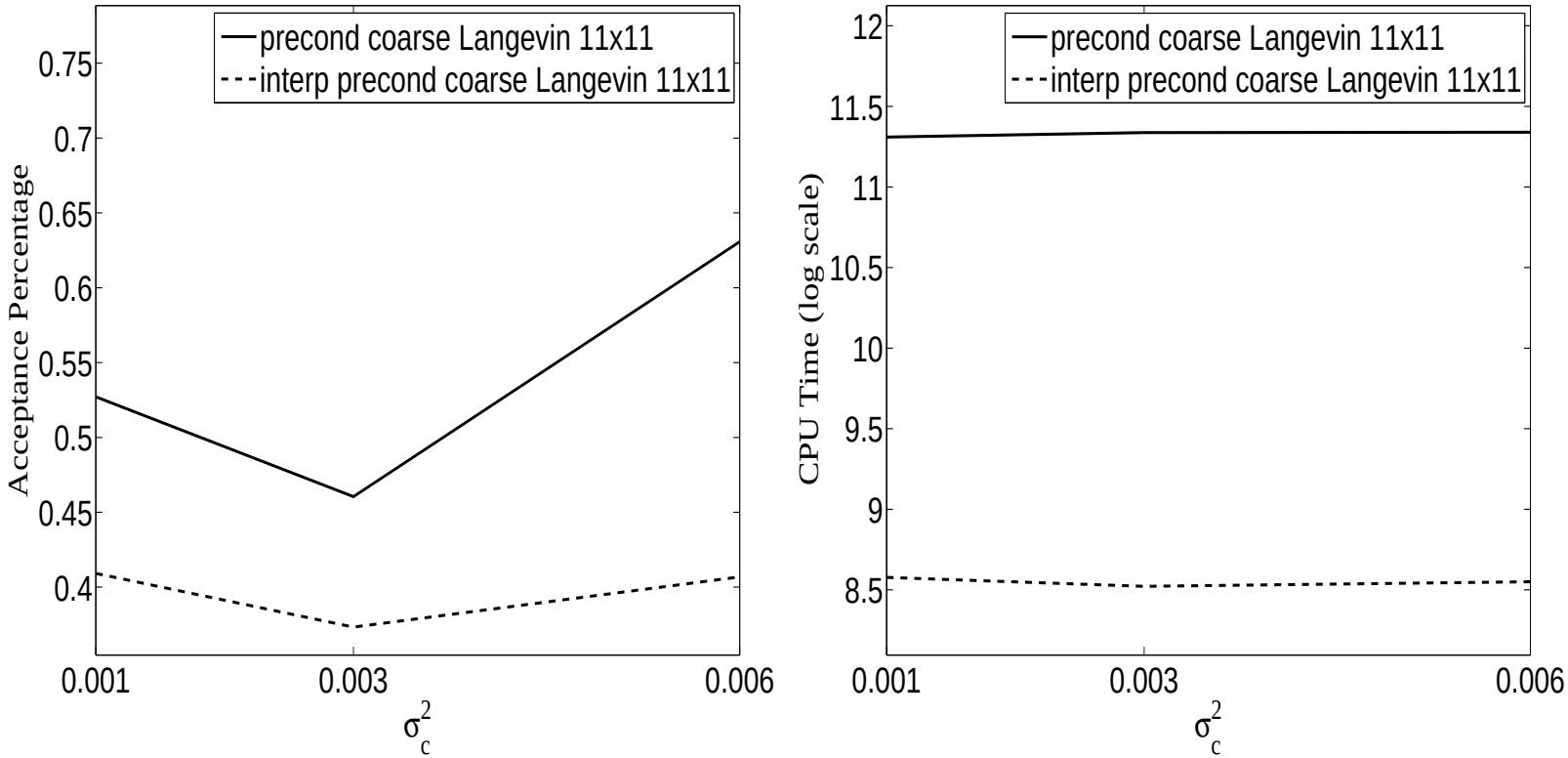
  where k_i are sparse collocation points and $L_i(k)$ are polynomials (P. Dostert et al., 2007).
- Some analysis of MH in the homogenization setting can be carried out.

Numerical Results



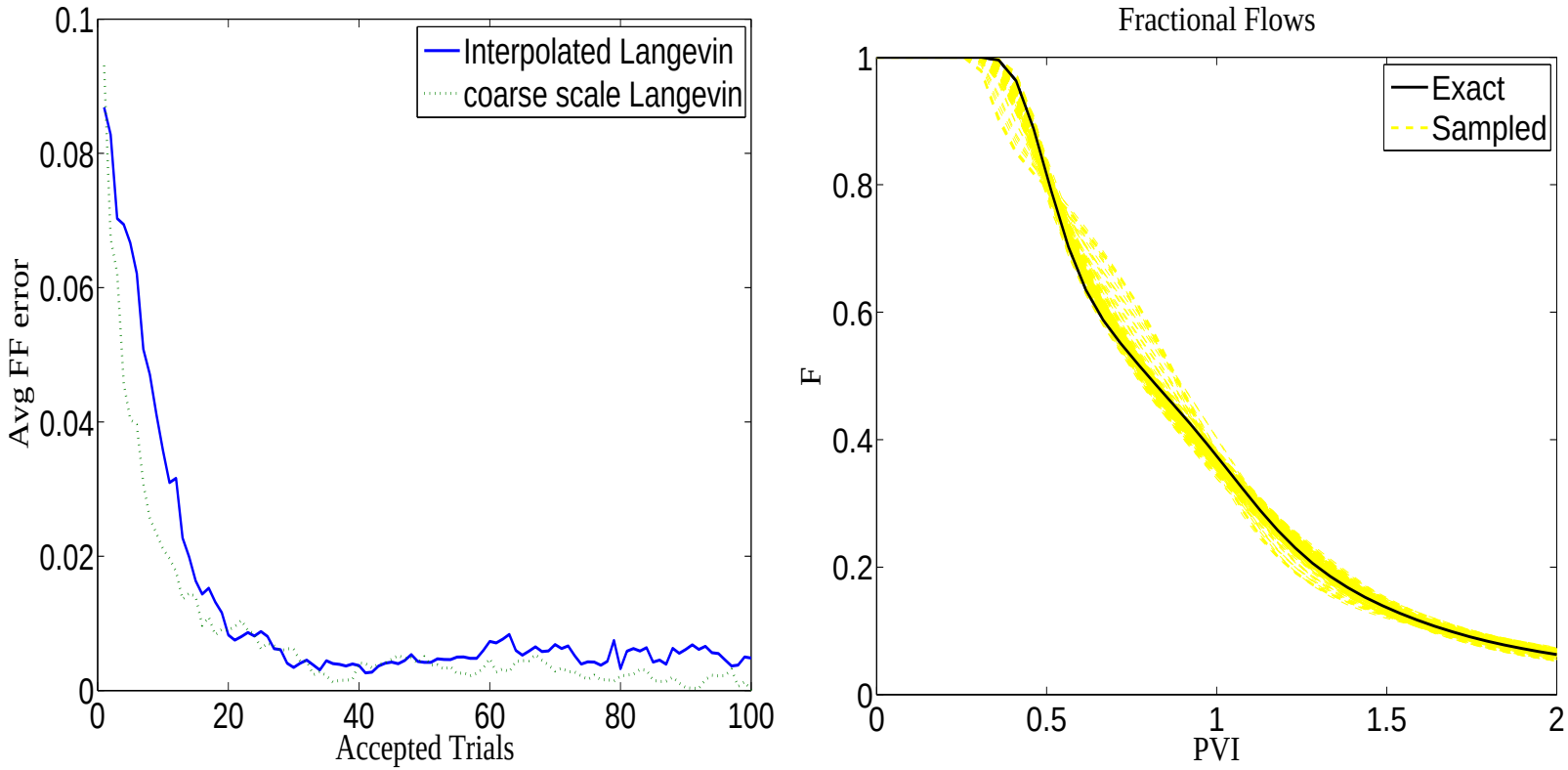
Left: Coarse-scale response surface π^* restricted to a 2-D hyperplane. Right: Fine-scale response surface π restricted to the same 2-D hyperplane.

Numerical Results



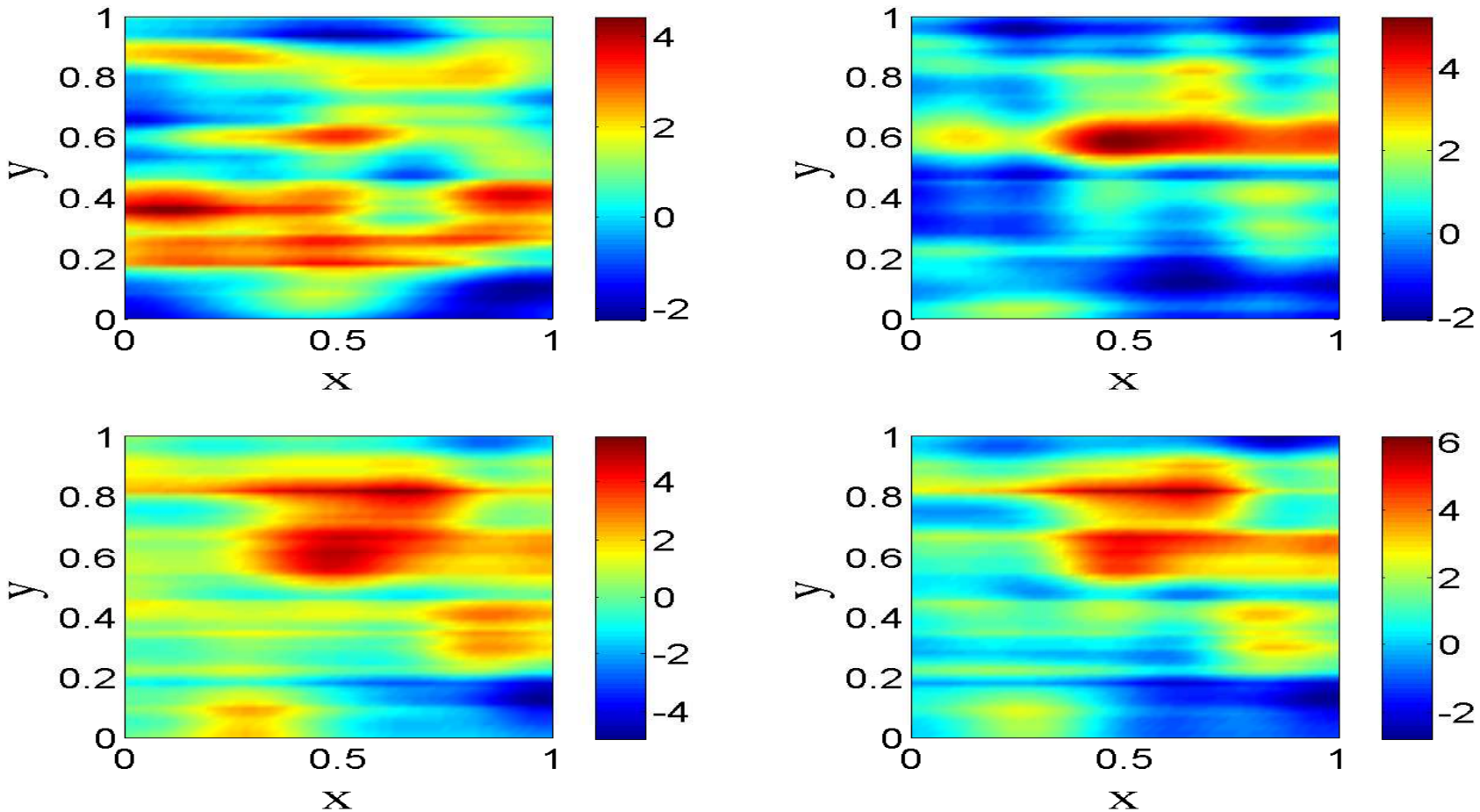
Left: Acceptance rate comparison. Right: Natural log of CPU time (seconds) comparison. In each plot we use $\delta = 0.05$ and $\sigma_f^2 = 0.001$.

Numerical Results



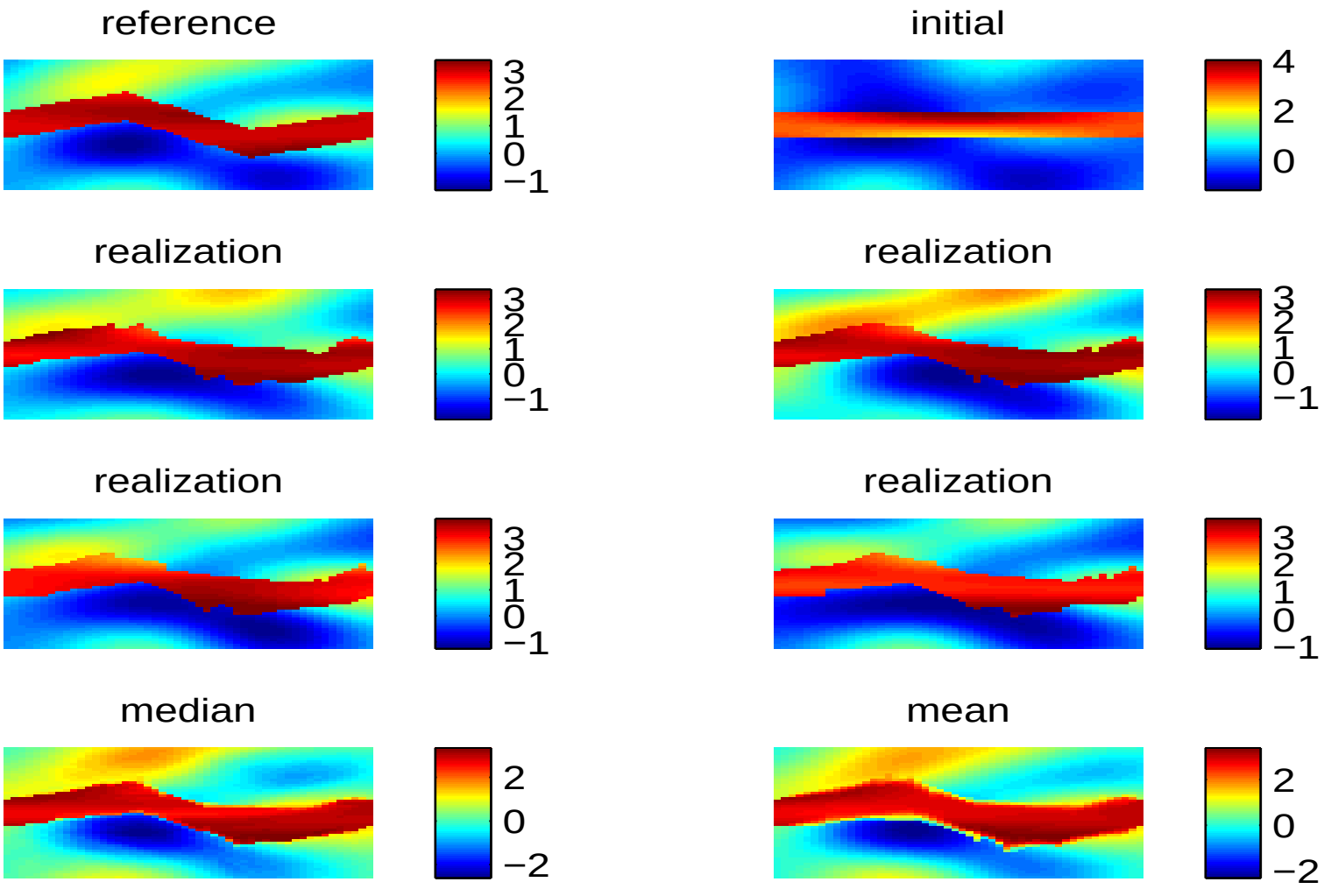
Left: The fractional flow errors for coarse Langevin compared with interpolated Langevin. Right: The fractional flows of sampled realizations and the reference fractional flow. In these numerical tests, $\delta = 0.05$, $\sigma_f^2 = 0.001$.

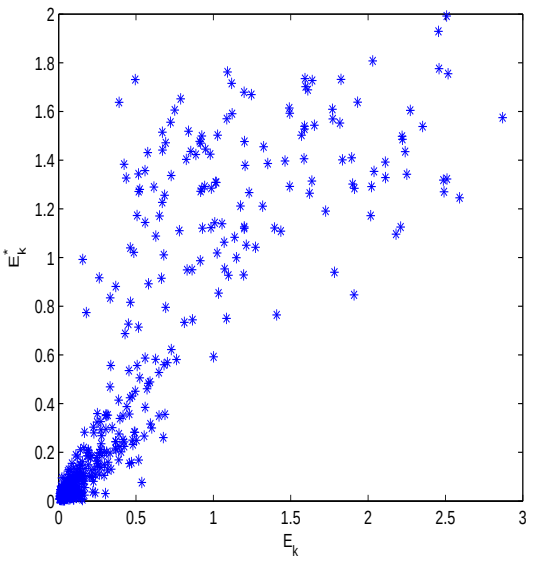
Numerical results

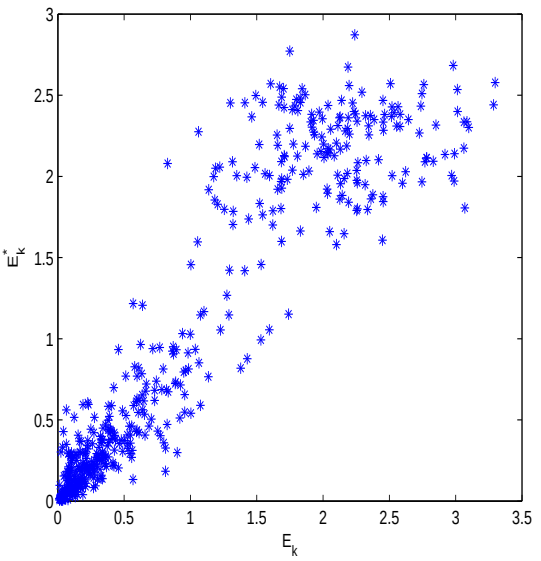


Upper left plot is the reference conductivity. The other three plots are examples of accepted conductivity realizations.

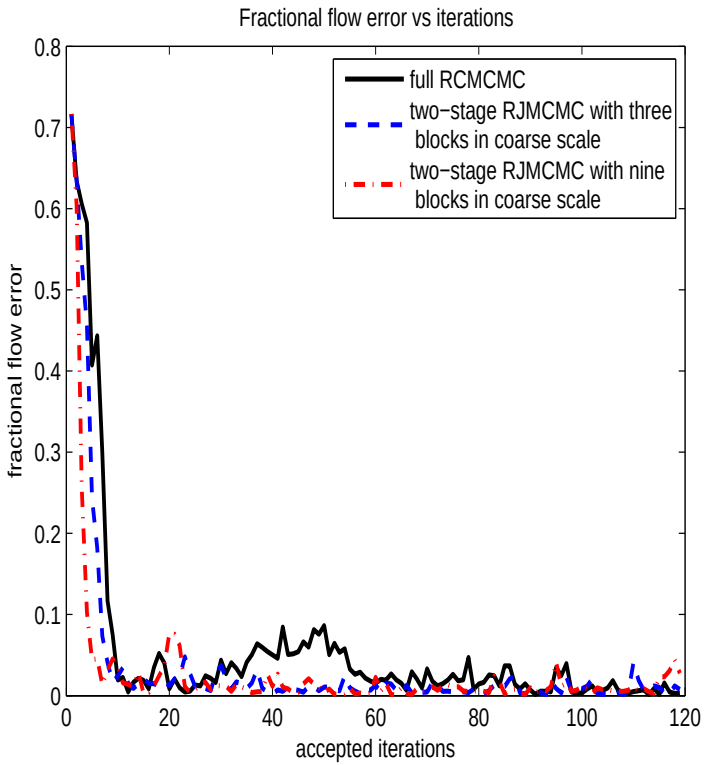
Numerical results



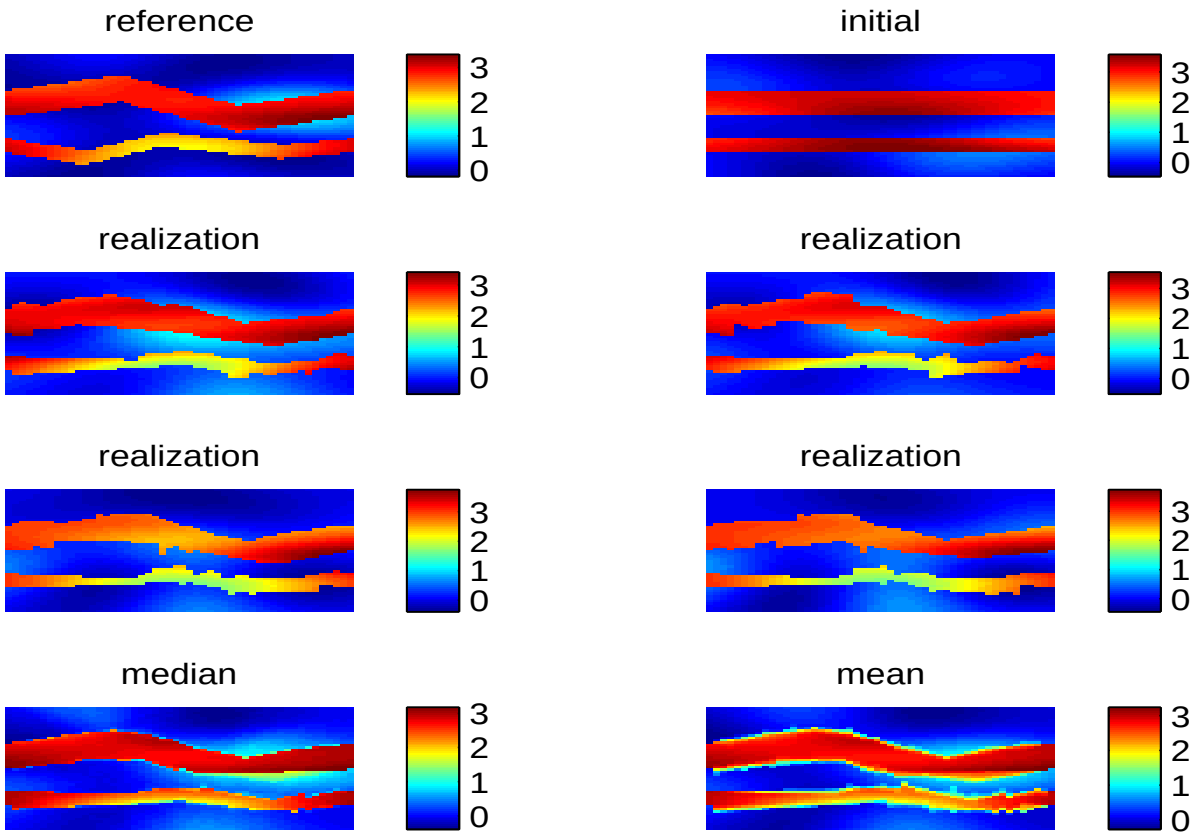

E_k^*
E_k


E_k^*
E_k

Numerical results



Numerical results



Conclusions

- Direct sampling using MH MCMC approaches is expensive
- Coarse gradient information and inexpensive coarse-scale models can be used to speed-up the simulations
- Numerical results demonstrate CPU time can be reduced by two orders of magnitude.
- Models at different scales with weights can be also used.
- Other representations for modeling interfaces
- ...