The use of coarse-scale models in uncertainty quantification

Yalchin Efendiev

Department of Mathematics

Texas A&M University

Collaborators: A. Datta-Gupta (TAMU), P. Dostert (TAMU), T. Hou (Caltech), V. Ginting (TAMU), B. Mallick (TAMU), W. Luo (Caltech)

Introduction

- Dynamic data integration in subsurface applications consists of integrating large-scale data (e.g., production data, tracer data) in order to reduce uncertainity and achieve realistic sampling of subsurface properties.

- An integrated response is usually measured with some precision. Trying to obtain hydraulic conductivity samples based on this integrated response is an ill-posed problem.

- The problem reduces to sampling from a complicated distribition involving the solutions of coupled nonlinear partial differential equations.

- Metropolis-Hasting Markov chain Monte Carlo (MCMC) methods can be used as an umbrella sampling method. MCMC used in a straightforward way is very CPU demanding.

- We propose and analyze approaches for efficient sampling which employ spatial multi-scale models.

Prototypical model

We consider two-phase flow in a reservoir under the assumption that the displacement is dominated by viscous effects.
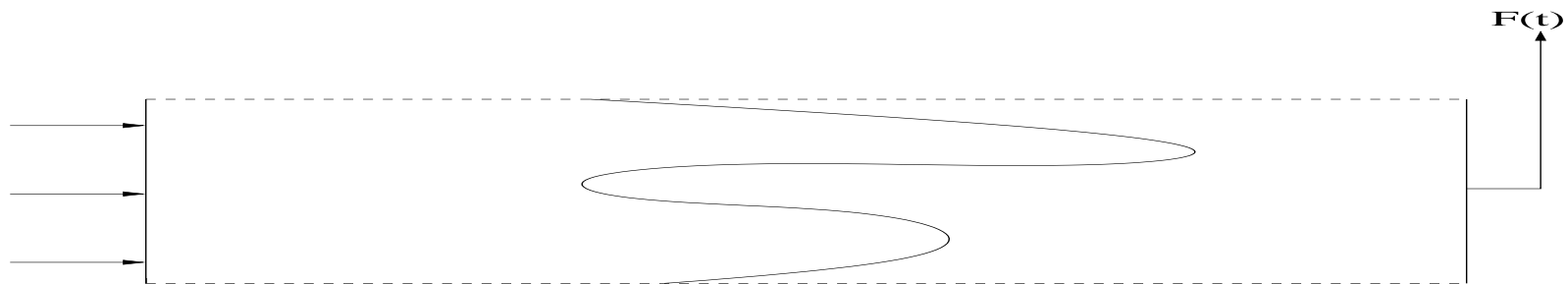
$$v_j = -\frac{k_{rj}(S)}{\mu_j}k \cdot \nabla p, \ \ j = w, o$$

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

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

Measure coarse-scale data:

$$F(t) = \frac{\int_{out} vf(S)dl}{\int_{out} vdl}$$

Coarse-scale spatial models (recap)

- We employ multiscale finite element methods as a single-phase upscaling technique. Multiscale methods, as traditional upscaling techniques, pre-compute effective parameters (basis functions) that are repeatedly used for different boundary condition, sources and mobilities.

- The pressure equation is upscaled using multiscale finite volume method and coarse-scale velocity field is calculated and used for solving the saturation equation. Basis functions are constructed only at time zero.

- This provides a very inexpensive approximation for the solution.

Coarse-grid   Fine-grid
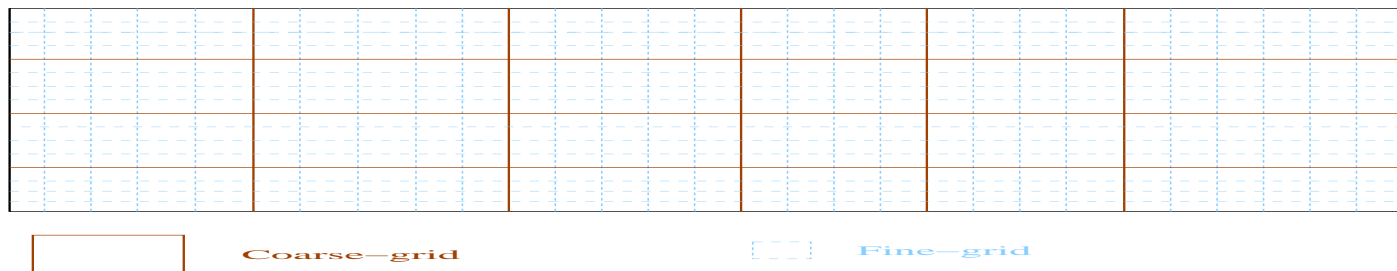
Multiscale methods (recap)

- Basis functions are constructed by solving the leading order homogeneous equation in an element K (coarse grid or RVE)

$$div(k(x)\nabla \phi^i) = 0 \quad \text{in} \quad K$$

- Boundary conditions are very important for accuracy of subgrid capturing error. Choices: (1) local boundary conditions (the information only within the target coarse block is taken into account); (2) oversampling (the information in slightly larger than the target coarse block domain is taken into account); (3) limited global information

- Pressure equation is solved on the coarse grid using ϕ_i and the saturation equation is also solved on the coarse grid. The resulting approach is similar to solving the system of equations with k replaced by k^*, $div(\lambda(S^*)k^*\nabla p^*) = q$, $S_t^* + v^* \cdot \nabla f(S^*) = q_w$, $v^* = -\lambda(S^*)k^*\nabla p^*$.

- Finally, the stochastic MsFEM is developed where a few realizations of $k(x, \omega)$ are used to construct multiscale basis functions for the whole ensemble.

Problem setting

- Given the fractional flow information (coarse-scale data) $F(t)$ 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 $P(F|k) = \exp(-\frac{\|F_k(t)-F^{obs}(t)\|^2}{\sigma_f^2})$.

- Typical prior can be $P(k) = \exp(-\frac{\|k-k_{obs}\|^2}{\sigma_k^2})$, where k_{obs} is a coarse-scale permeability. Thus, the posterior distribution is

$$P(k|F) \propto \exp(-\frac{\|F_k(t) - F^{obs}(t)\|^2}{\sigma_f^2}) \exp(-\frac{\|k - k_{obs}\|^2}{\sigma_k^2}).$$

Examples of prior

$Y(x, \omega) = log(k(x, \omega))$ is described by two-point correlation functions
$R(x, y) = E(Y(x, \omega)Y(y, \omega))$. In general, $R(x, y)$ can be defined on the fine-grid. Here
are examples:

- Normal

$$R(x, y) = \sigma^2 \exp\Big(-\frac{|x_1 - y_1|^2}{2L_1^2} - \frac{|x_2 - y_2|^2}{2L_2^2}\Big),$$

where L_1 and L_2 are called correlation lengths and σ is the variance.
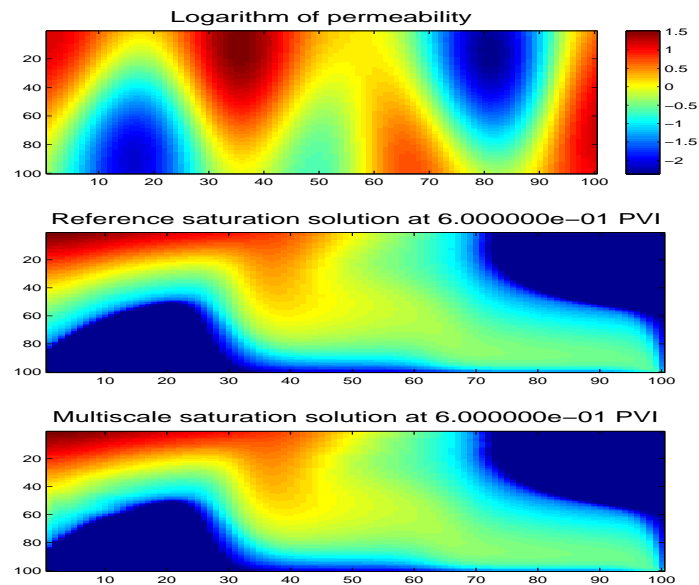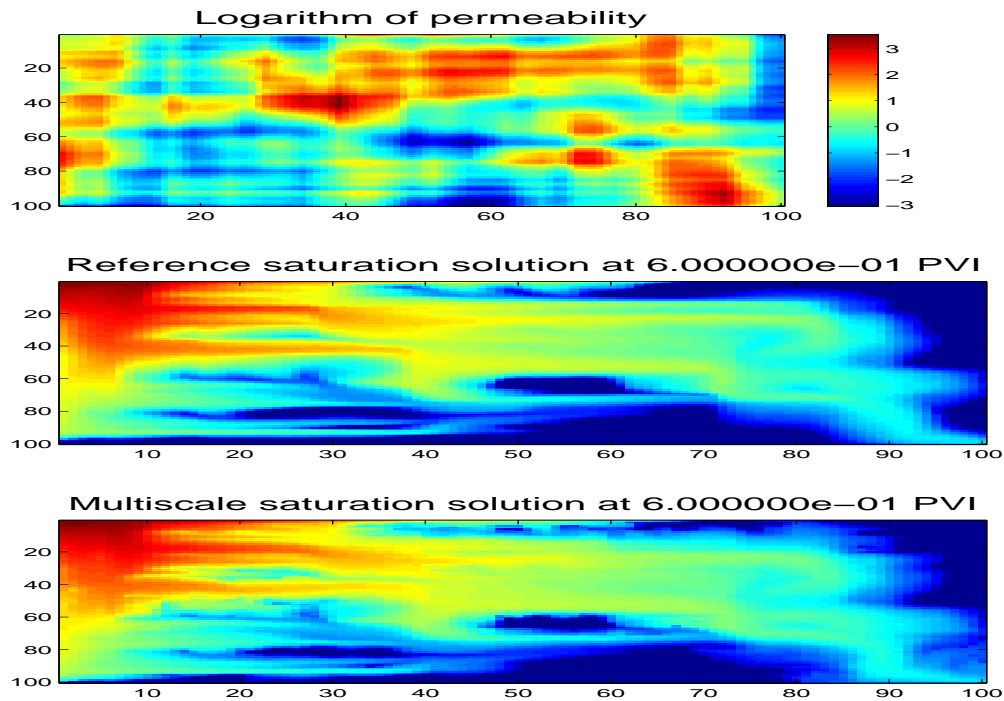

$L_1 = 0.1, L_2 = 0.5.$

# Examples

Exponential variogram (less smooth).

$$R(x, y) = \sigma^2 \exp\left(-\frac{|x_1 - y_1|}{L_1} - \frac{|x_2 - y_2|}{L_2}\right).$$

Representation of the conductivity fields

- Karhunen-Loeve expansion. $Y(x) = \sum_{i=1}^{M} \theta_i \sqrt{\lambda_i} \Phi_i(x)$, where λ_i and $\Phi_i(x)$ are eigenvalues and eigenvectors of covariance matrix.

- For normal distribution, λ_i decay very fast and only a few terms survive in K-L expansion, i.e., M is small.

- For rougher fields, λ_i decay slower and M is large.

Eigenvalues and eigenvectors

first

third

eight

twenty
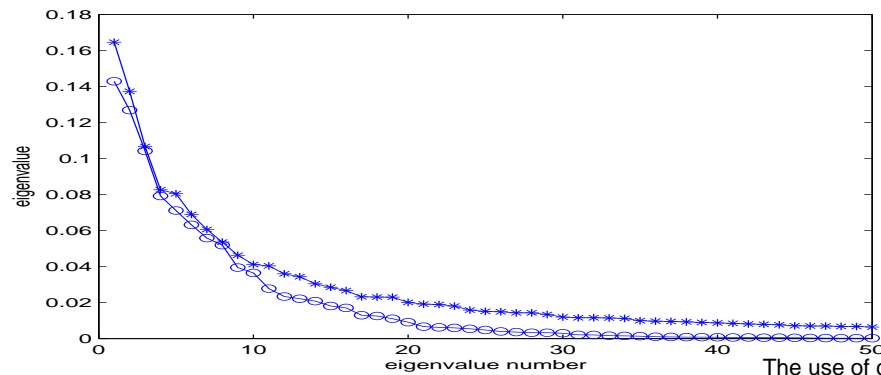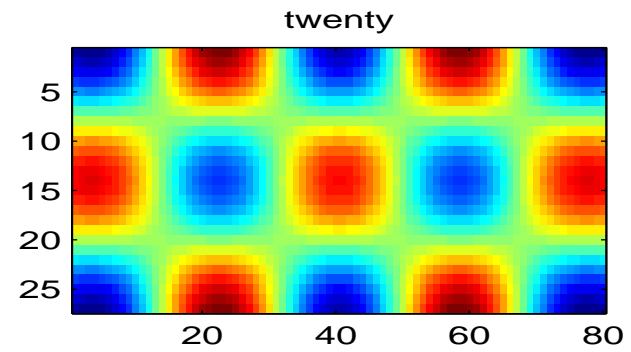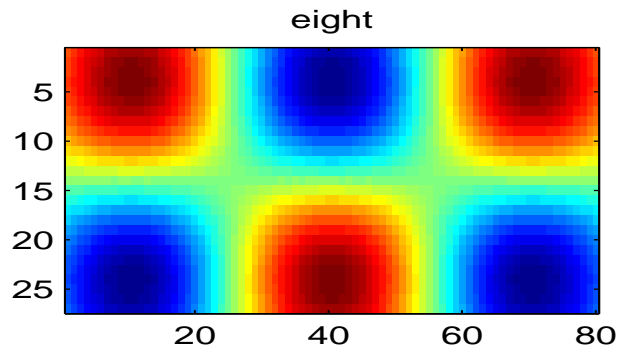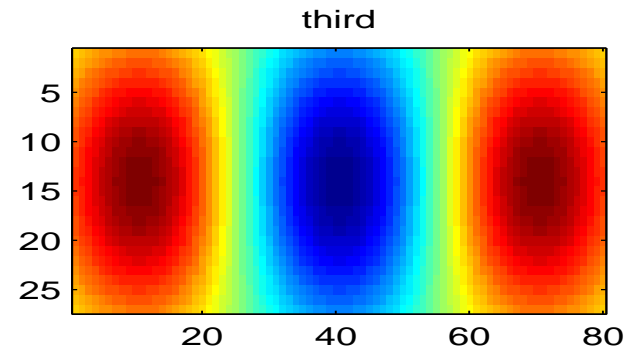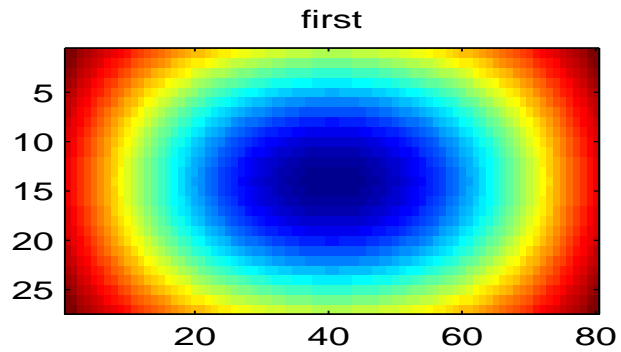
Difficulties

- $\pi(k) = P(k|F)$ can be multi-modal and high dimensional.

- $\pi(k) = P(k|F)$ is not given analytically and involves the solution of nonlinear pde system.
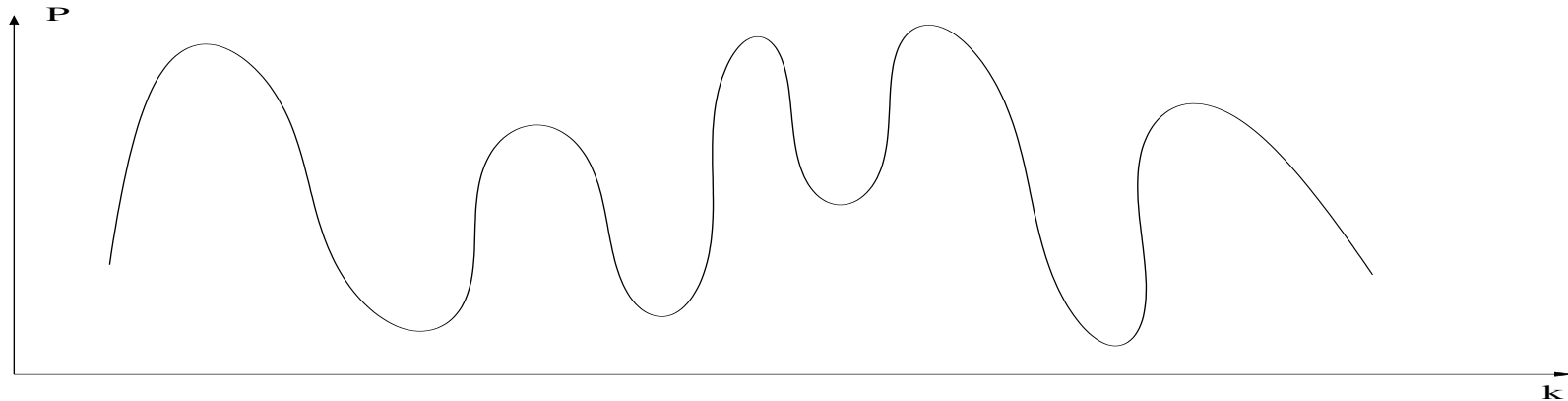
Metropolis-Hastings MCMC

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 propose an algorithm, where the proposal distribution is modified using coarse-scale spatial models.

Remark

Coarse-scale posterior is smoother and do not have all local maxima of fine-scale posterior
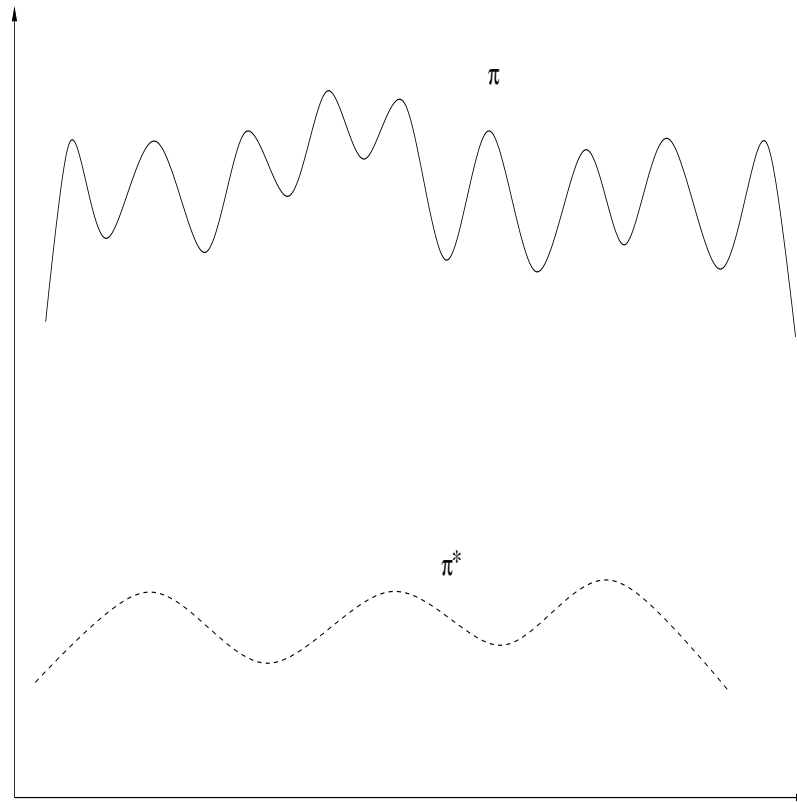
Remark
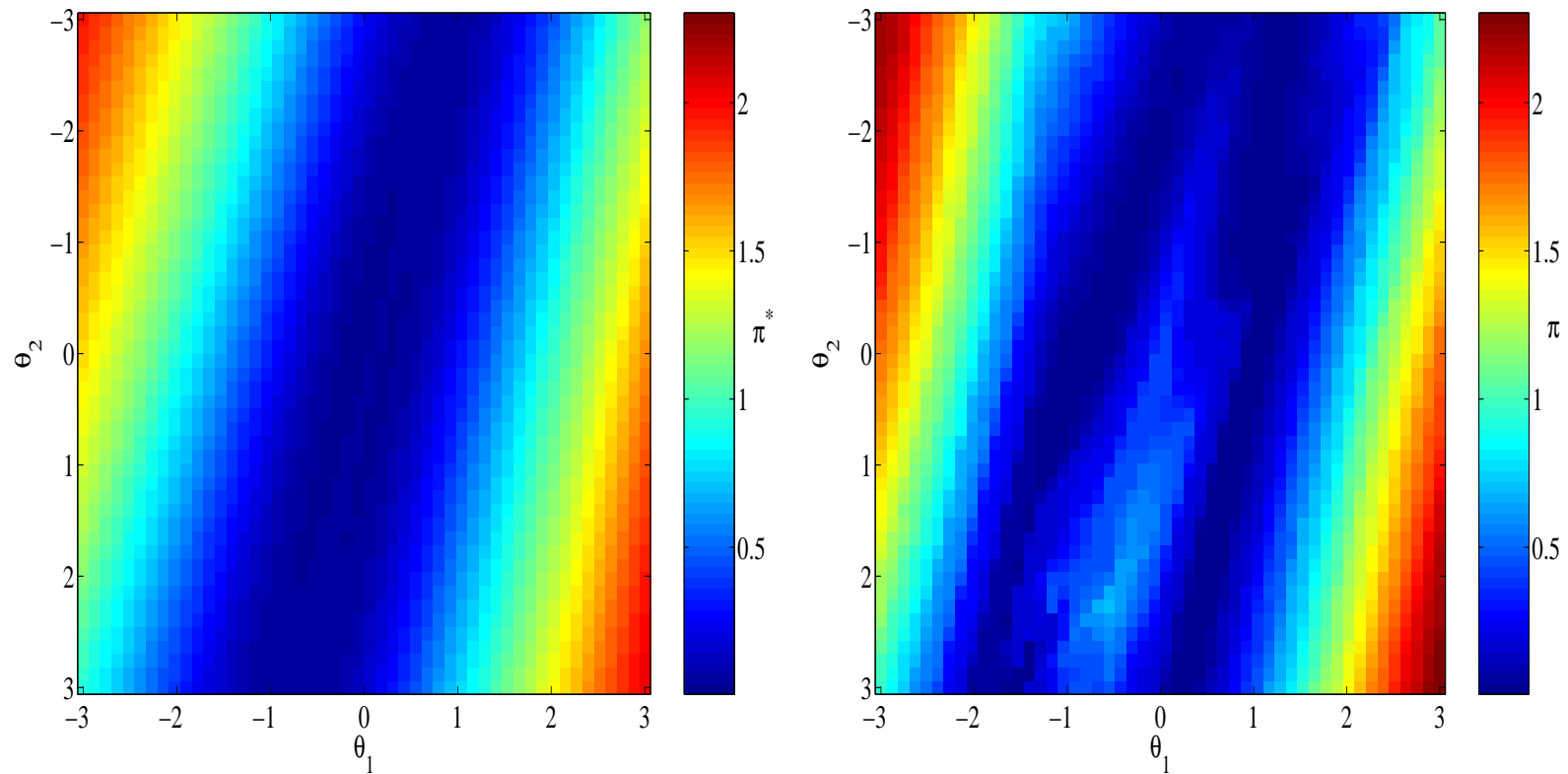
Left: Coarse-scale response surface π^* restricted to a 2-D hyperplane. Right: Fine-scale response surface π restricted to the same 2-D hyperplane.

Instrumental distribution

- Some simple instrumental distributions are independent sampler and random walk sampler.

- In the case of independent sampler, the proposal distribution $q(k|k_n)$ is chosen to be independent of k_n.

- In random walk sampler, the proposal distribution depends on the previous value of the permeability field and given by

$$k = k_n + \epsilon_n,$$

where ϵ_n is a random perturbation with prescribed variance.

Langevin instrumental distribution

An important type of proposal distribution can be derived from the Langevin diffusion. The Langevin diffusion is defined by the stochastic differential equation

$$dk(\tau) = \frac{1}{2}\nabla \log \pi(k(\tau))d\tau + dW_\tau,$$

where W_τ is the standard Brownian motion vector with independent components. The solutions of this stochastic differential equation are from $\pi(k)$.
A discretization of the equation,

$$k_{n+1} = k_n + \frac{\Delta\tau}{2}\nabla \log \pi(k_n) + \sqrt{\Delta\tau}\epsilon_n,$$

where ϵ_n are independent standard normal distributions.
The proposal is chosen to be

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

Langevin instrumental distribution

The transition distribution of the proposal is

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

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

The reasons for using Langevin:

- Pde's describing the physical model allow us to compute the gradients.

- The use of gradients is common in "stochastic" subsurface applications, e.g., Randomized Maximum Likelihood (RML). This approach samples the measurement data and the prior information independently and then minimize the posterior functional with these samples.

- The use of Langevin proposals usually yields higher mixing rates compared to e.g., random walk sampler.

Preconditioned coarse-gradient Langevin algorithm

The main idea: (1) use coarse-scale simulations to compute the gradient and make a proposal; (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

$$k = \begin{cases} Y \text{ with probability } g(k_n, Y), \\ k_n \text{ with probability } 1 - g(k_n, Y), \end{cases}$$

  where

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

- 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.

Preconditioned coarse-gradient Langevin algorithm

- The transition distribution of the coarse-grid proposal is

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

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

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 $\mathcal{E} \subseteq \Omega$.

Numerical setting

- We consider log-normal permeability fields $k(x) = \exp(Y(x))$, where $Y(x)$ is prescribed with a covariance matrix (e.g., normal or exponential).

- The permeability field is parameterized via Karhunen-Loève Expansion

$$Y(x, \omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \theta_k(\omega) \phi_k(x),$$

where $E(\theta_k) = 0$, $E(\theta_i \theta_j) = \delta_{ij}$, λ_k and $\phi_k(x)$ are eigenvalues and eigenvectors of covariance matrix.

- First step parameter reduction is performed by neglecting "small" eigenvalues.

- The permeability field can be conditioned at well locations.

Numerical result
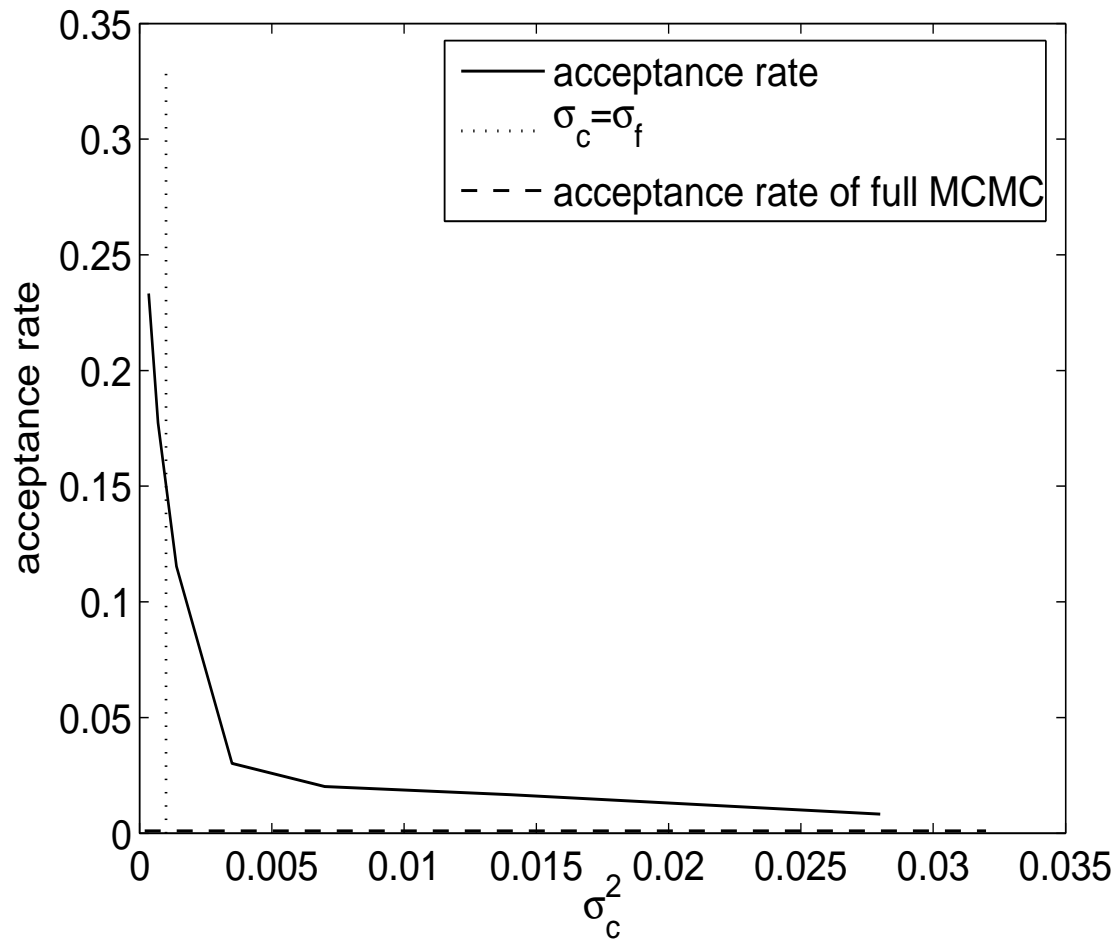
Acceptance rate vs. different coarse-scale precisions for the preconditioned MCMC. Single-phase flow and $\sigma_f^2 = 0.001$.

Numerical result
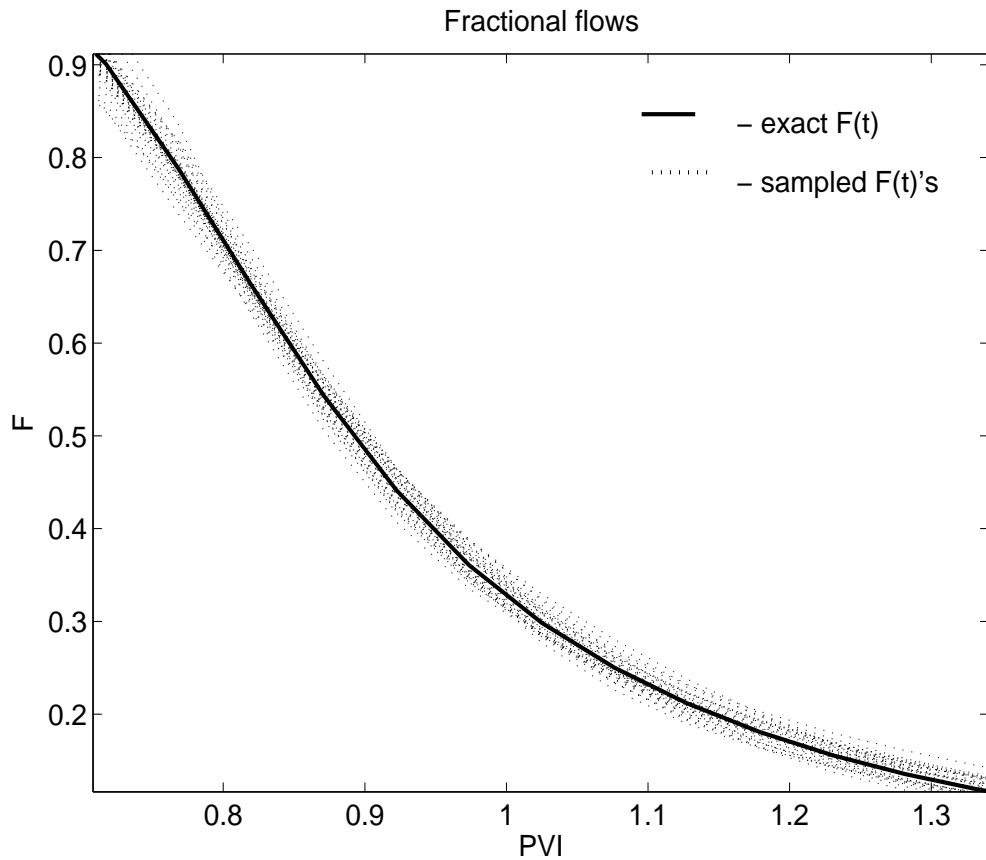
Fractional flows

Fractional flow comparisons. Left: Cross-plot between the reference fractional flow and sampled fractional flows. Right: Solid line designates the fine-scale reference fractional flow, and dotted lines designate fractional flows corresponding to sampled permeability fields.

Numerical result

Fractional flow error vs. iterations
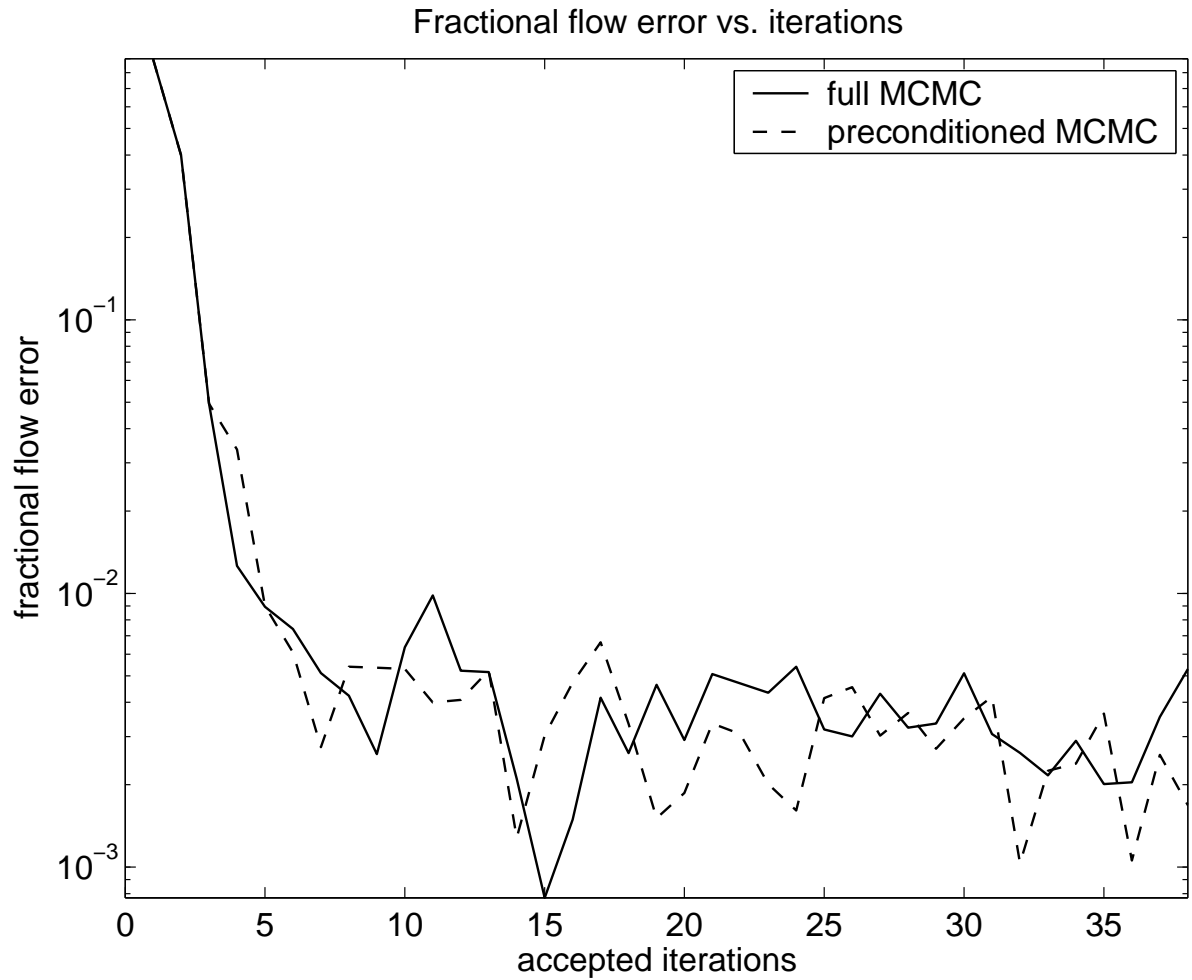
Fractional flow errors vs. accepted iterations.

Numerical result
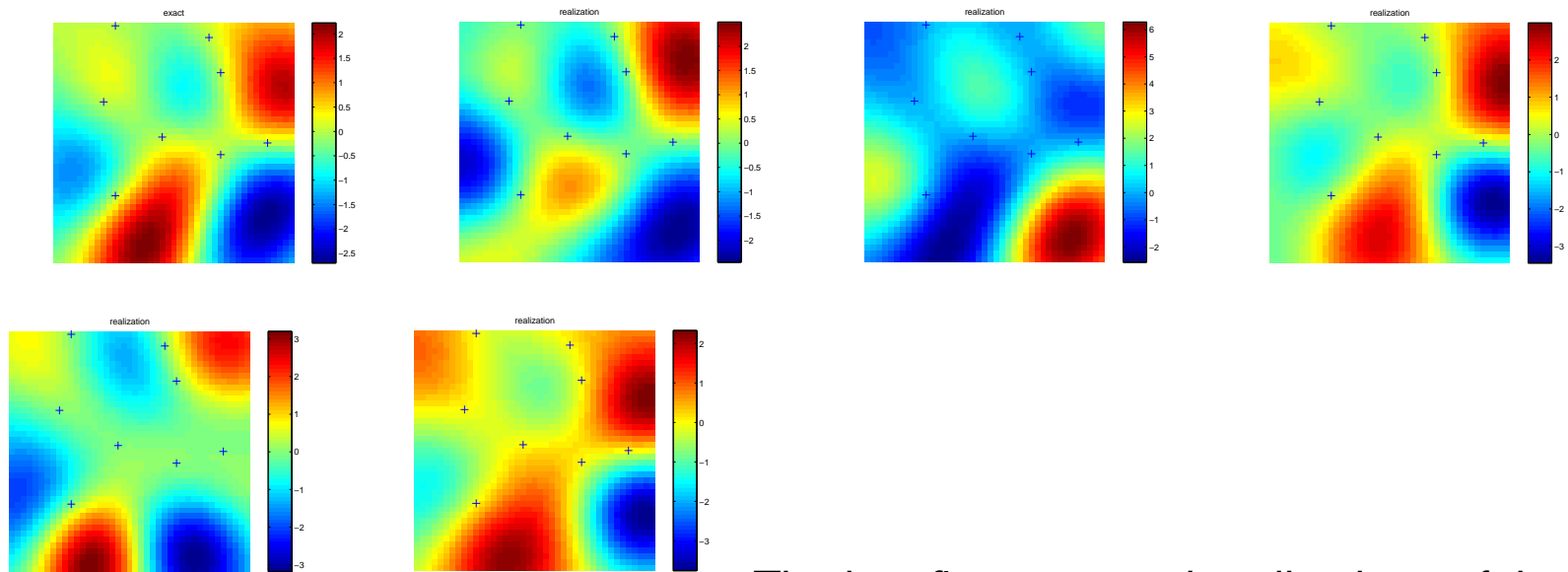
The last five accepted realizations of the log permeability field. The "+" sign marks the locations of the hard data.
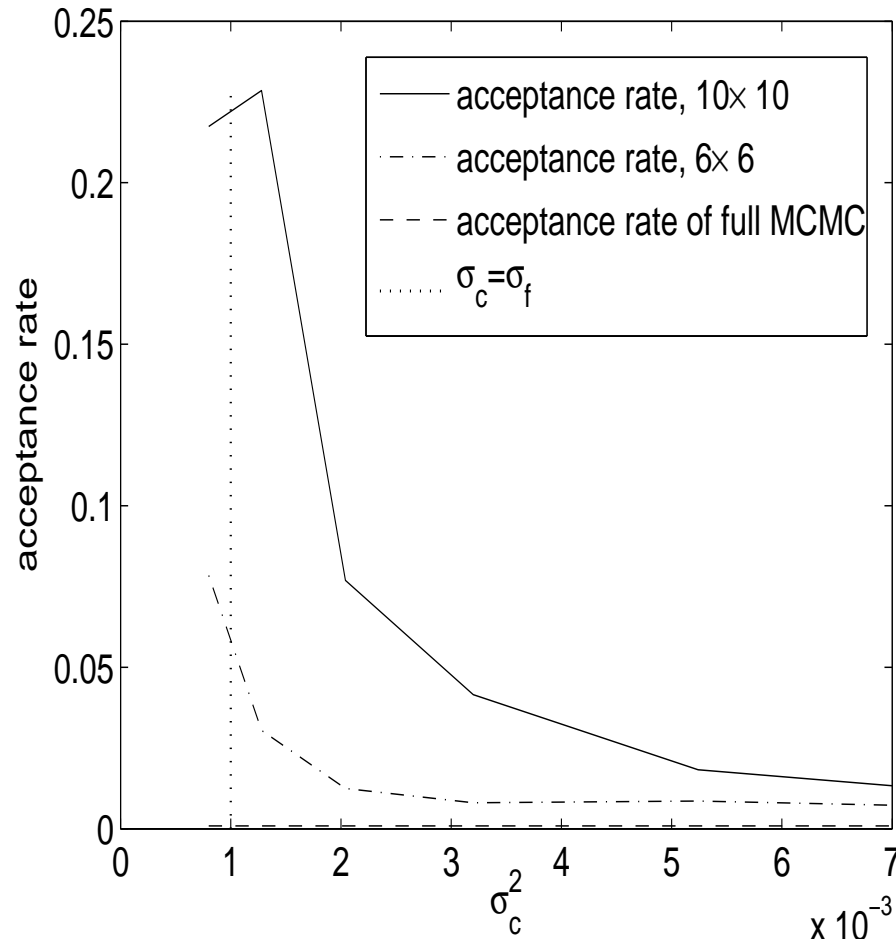
Numerical result

Acceptance rate vs. different coarse-scale precisions of MCMC using 6×6 and 10×10 coarse-scale models. Anisotropic single-phase flow and $\sigma_f^2 = 0.001$.
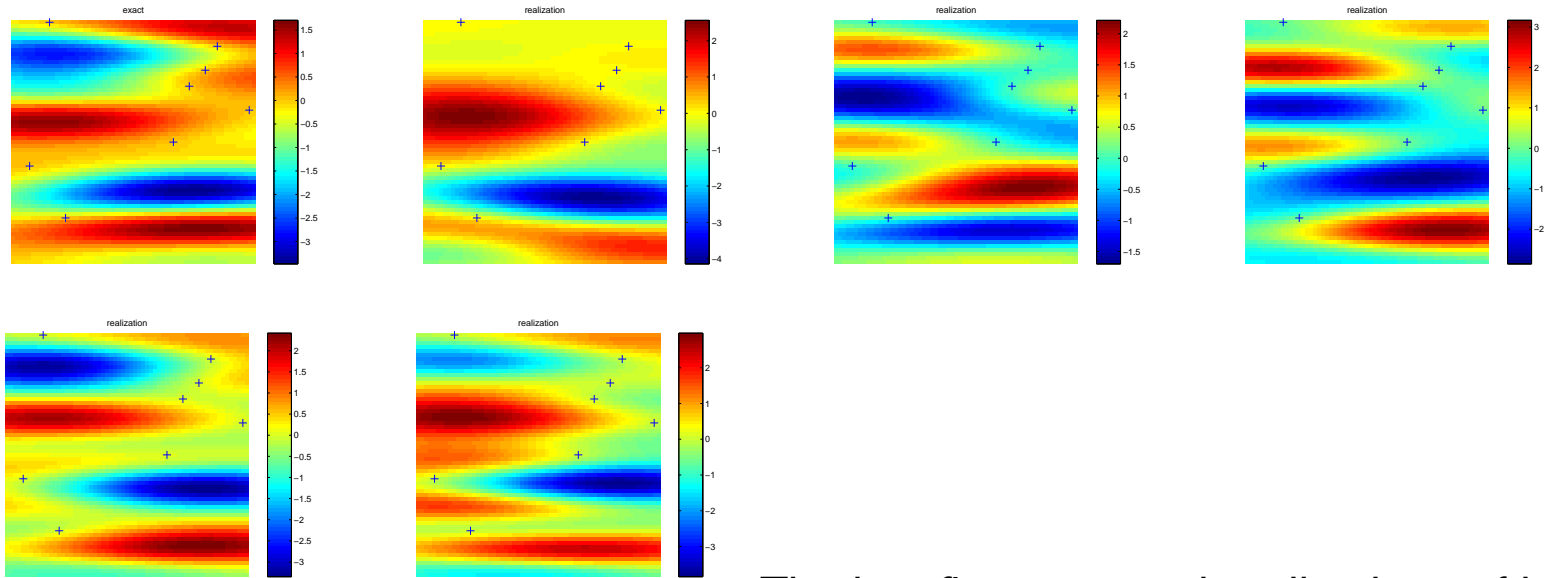
Numerical result

The last five accepted realizations of log permeability field for anisotropic case. The "+" sign marks the locations of the hard data.
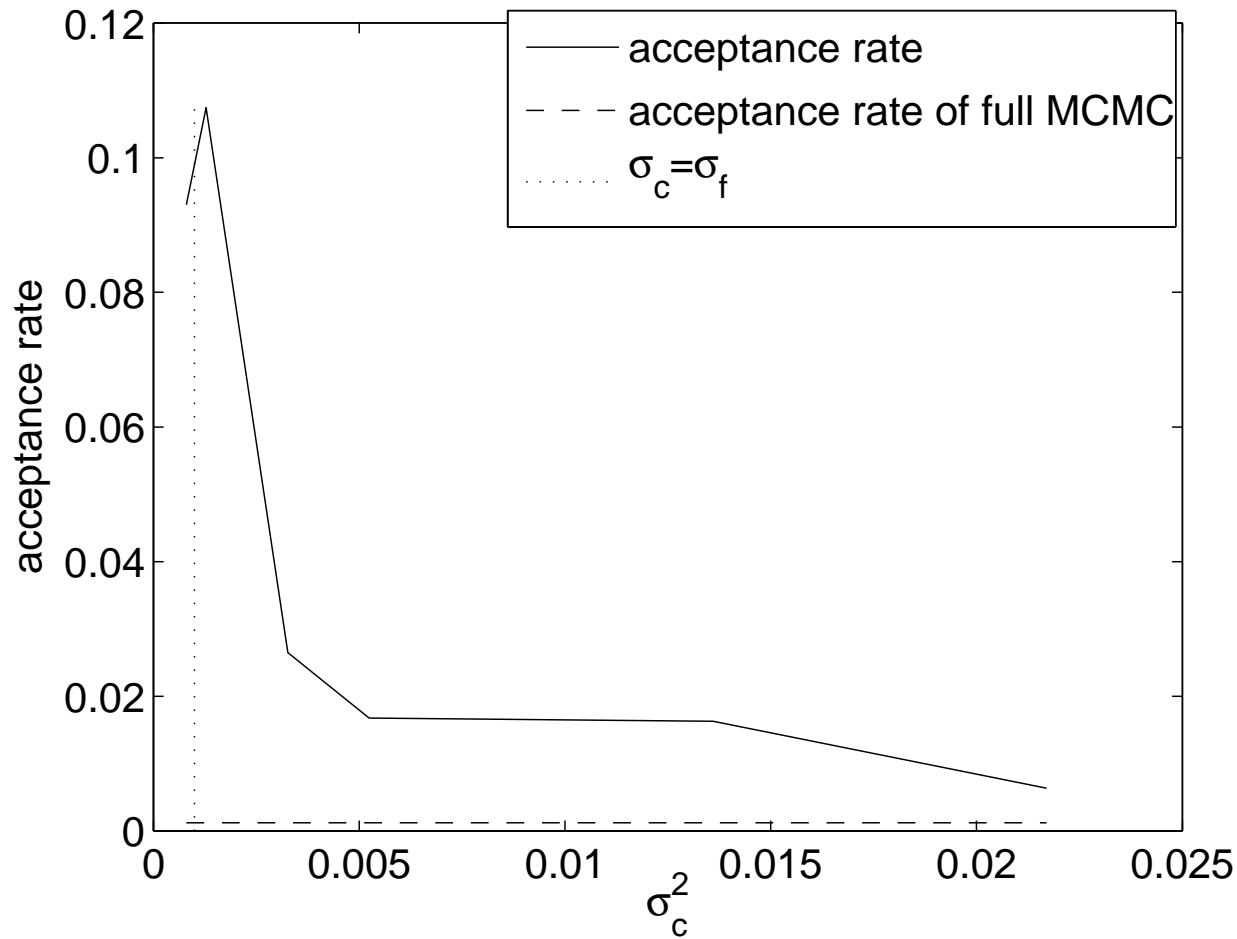
Numerical result

Acceptance rate vs. coarse-scale precision of the MCMC method. Two-phase flow and $\sigma_f^2 = 0.001$.

Numerical result

Fractional flow error vs. iterations
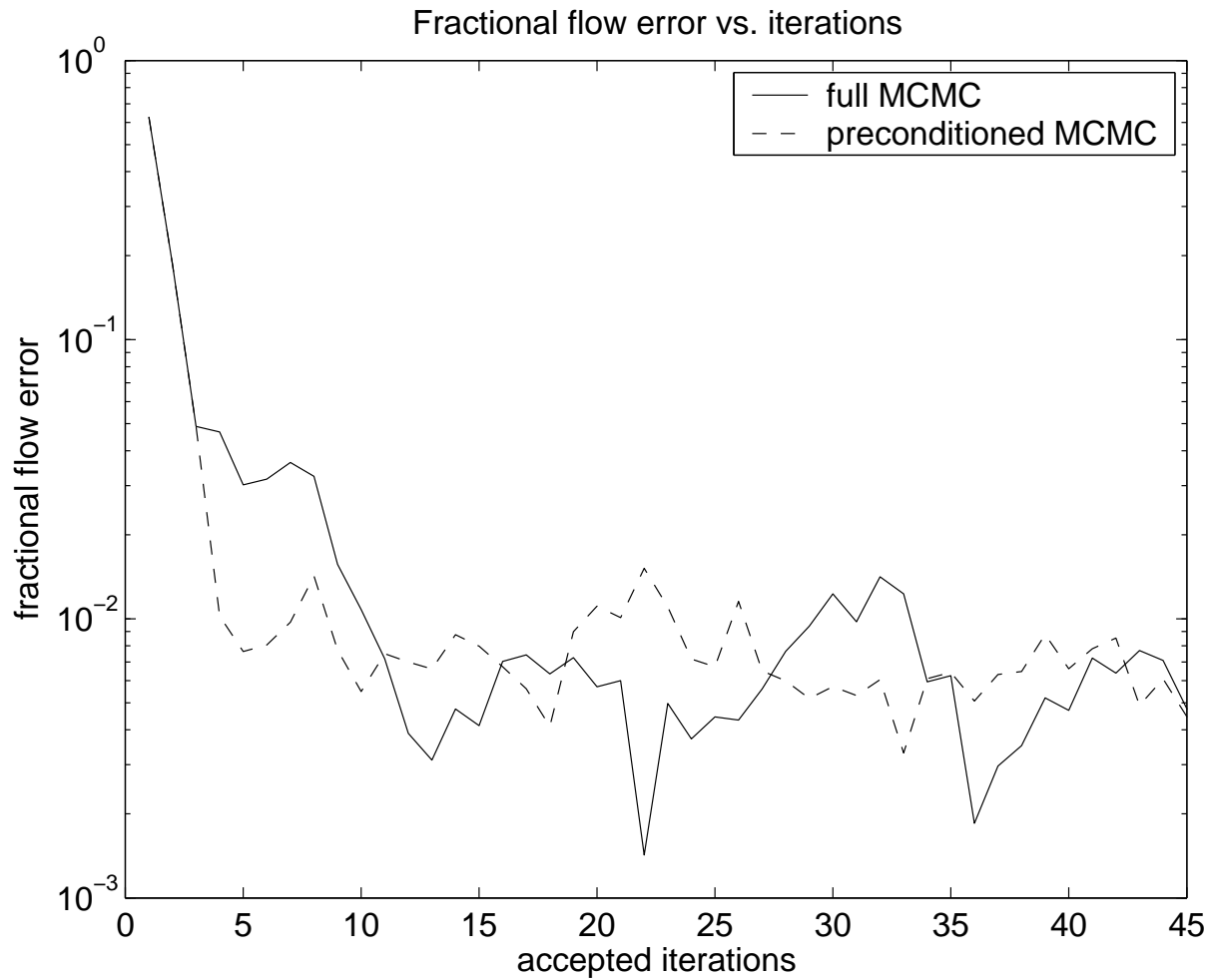
Fractional flow errors vs. accepted iterations in two phase-flow.

Numerical result
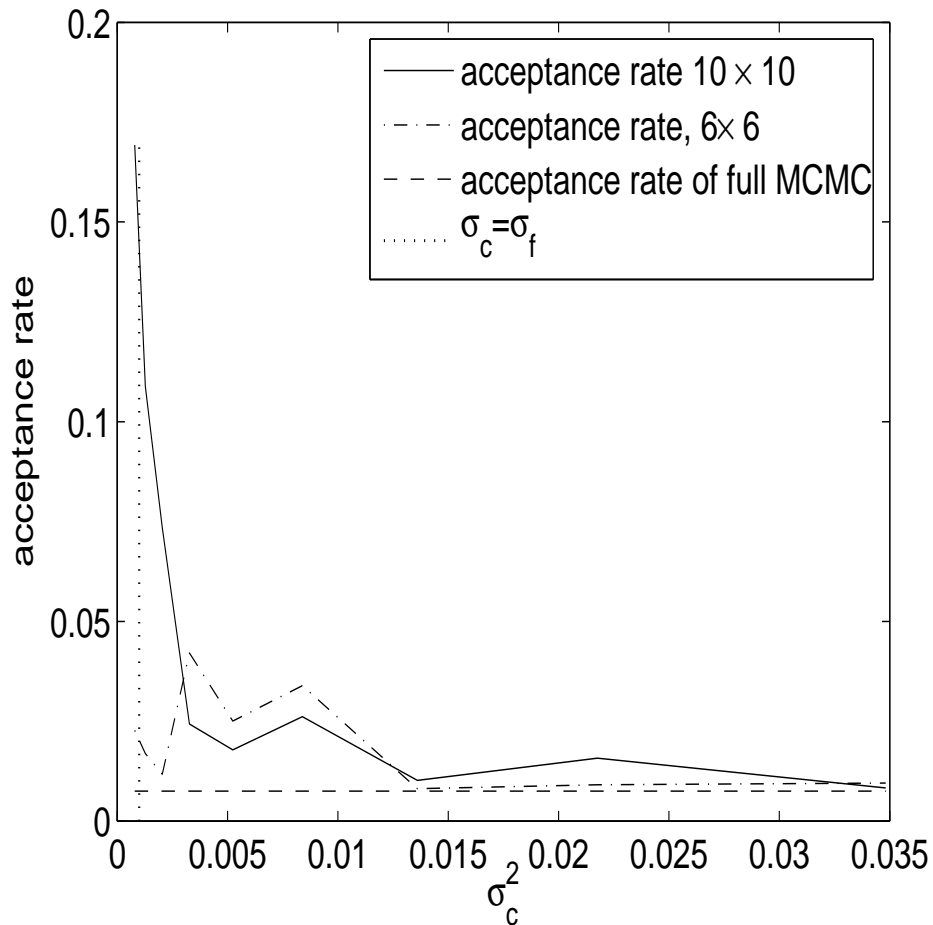
Acceptance rate vs. coarse-scale precision of MCMC using 6×6 and 10×10 coarse-scale models. Single-phase flow and $\sigma_f^2 = 0.001$. Random walk sampler is used as the proposal distribution.
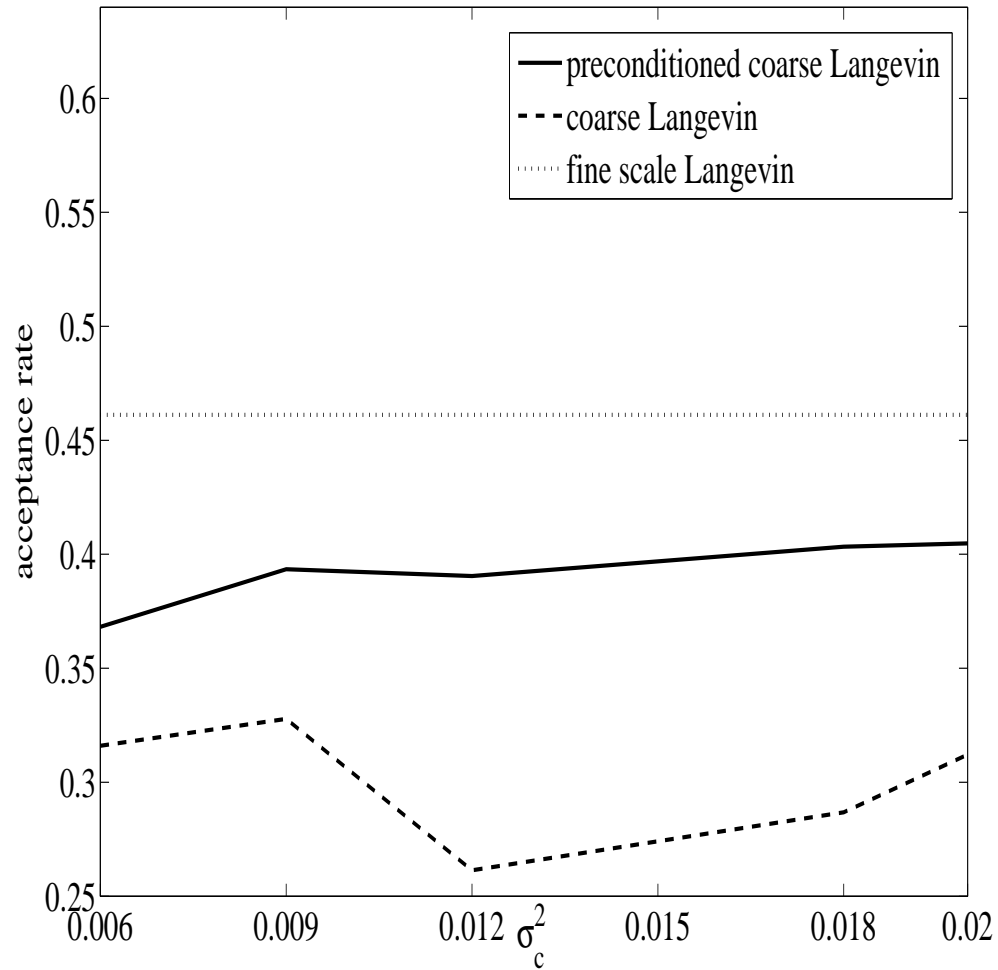
Numerical result

Fractional flow error vs. iterations
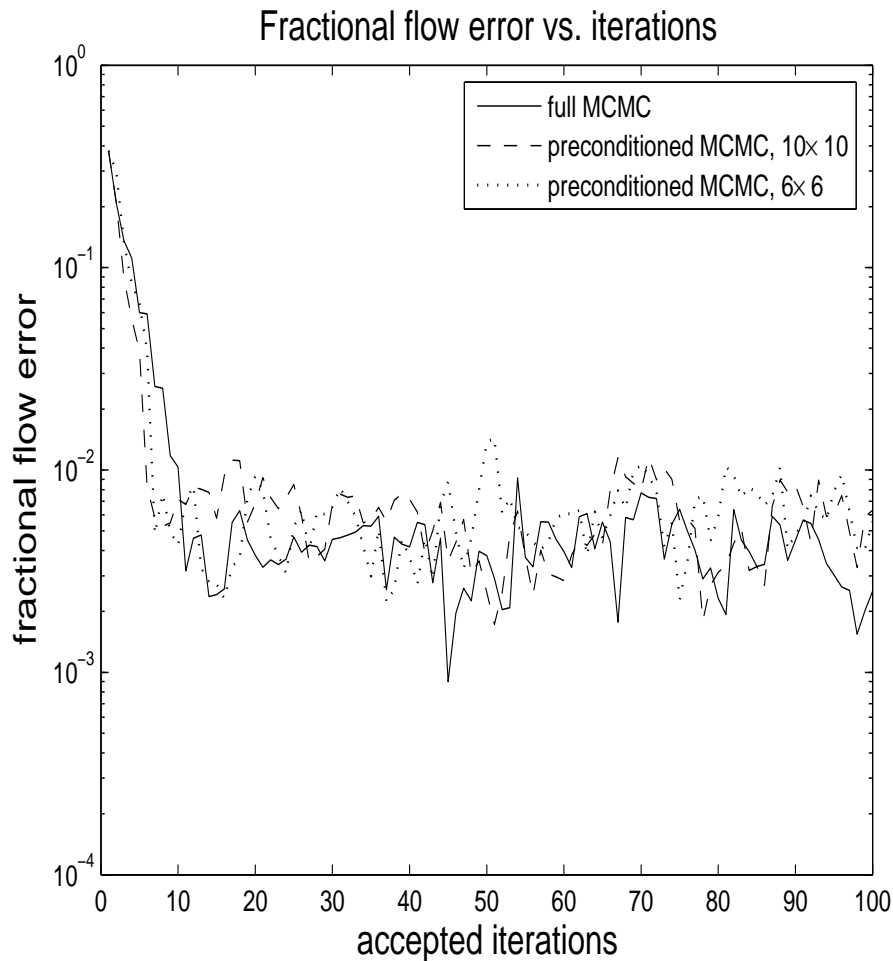
Fractional flow errors vs. accepted iterations for 6×6 and 10×10 coarse-scale models. Single-phase flow and $\sigma_f^2 = 0.001$. Random walk sampler is used as the proposal distribution.

Numerical result
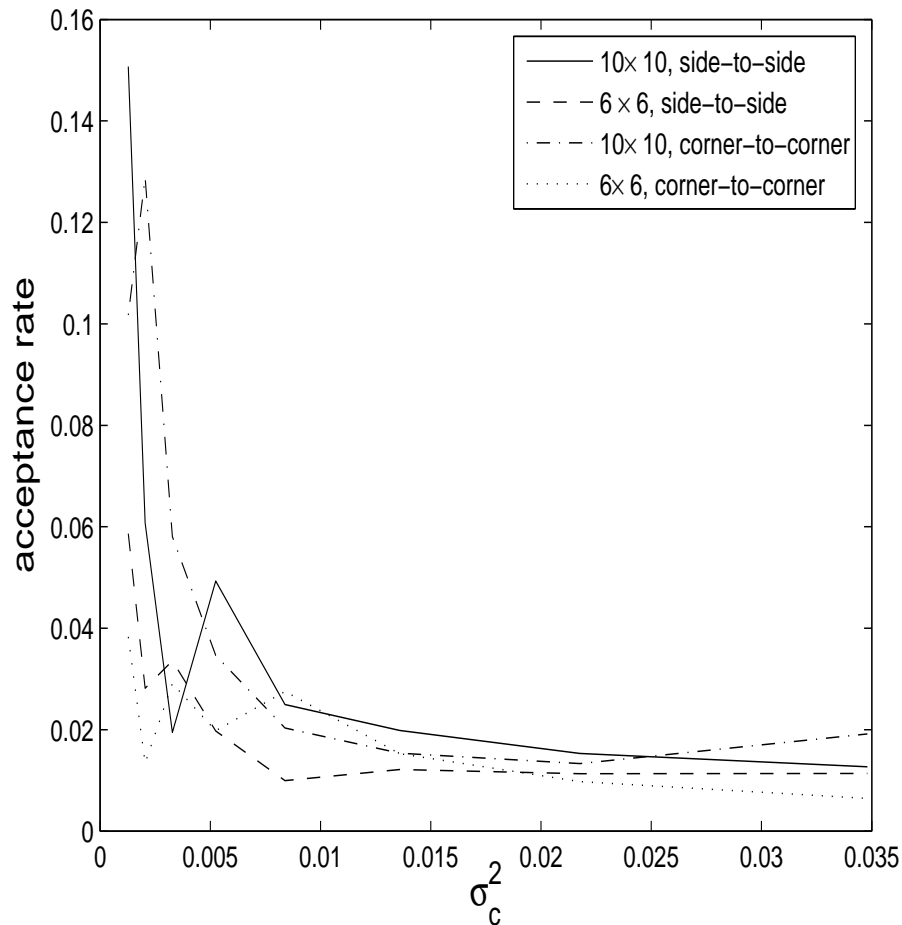
Acceptance rates of the preconditioned MCMC method using 6×6 and 10×10 coarse-scale models for side-to-side and corner-to-corner boundary conditions. Single-phase flow and $\sigma_f^2 = 0.001$. Random walk sampler is used as the proposal distribution.

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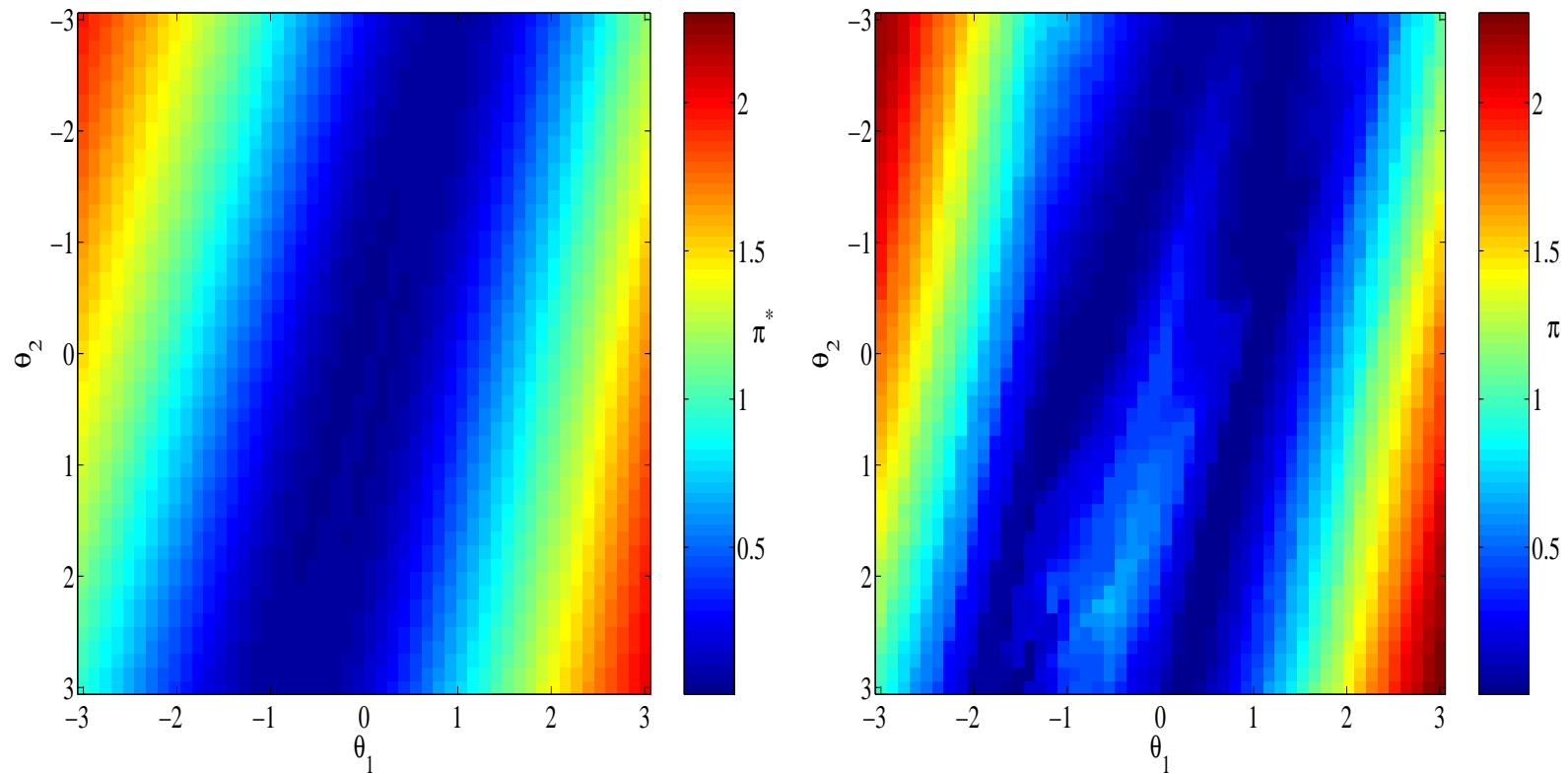
Left: Coarse-scale response surface π^* restricted to a 2-D hyperplane. Right: Fine-scale response surface π restricted to the same 2-D hyperplane.
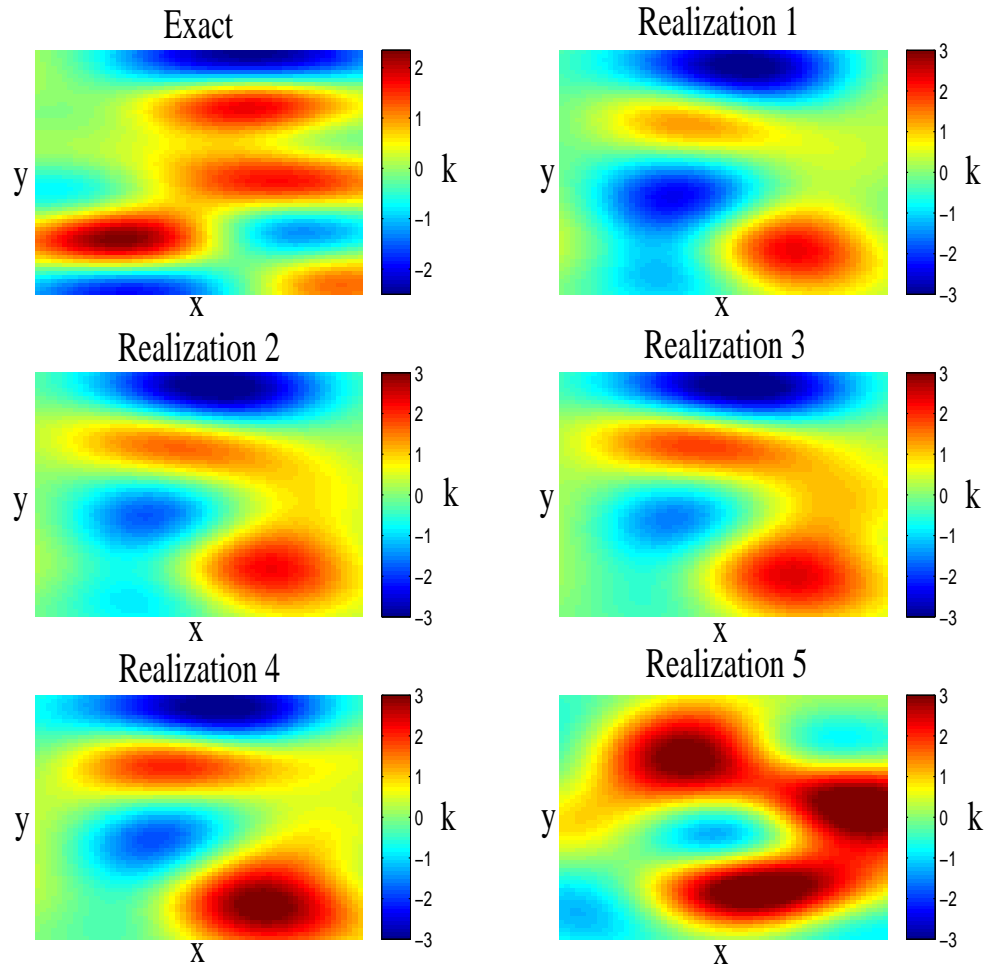
Numerical results

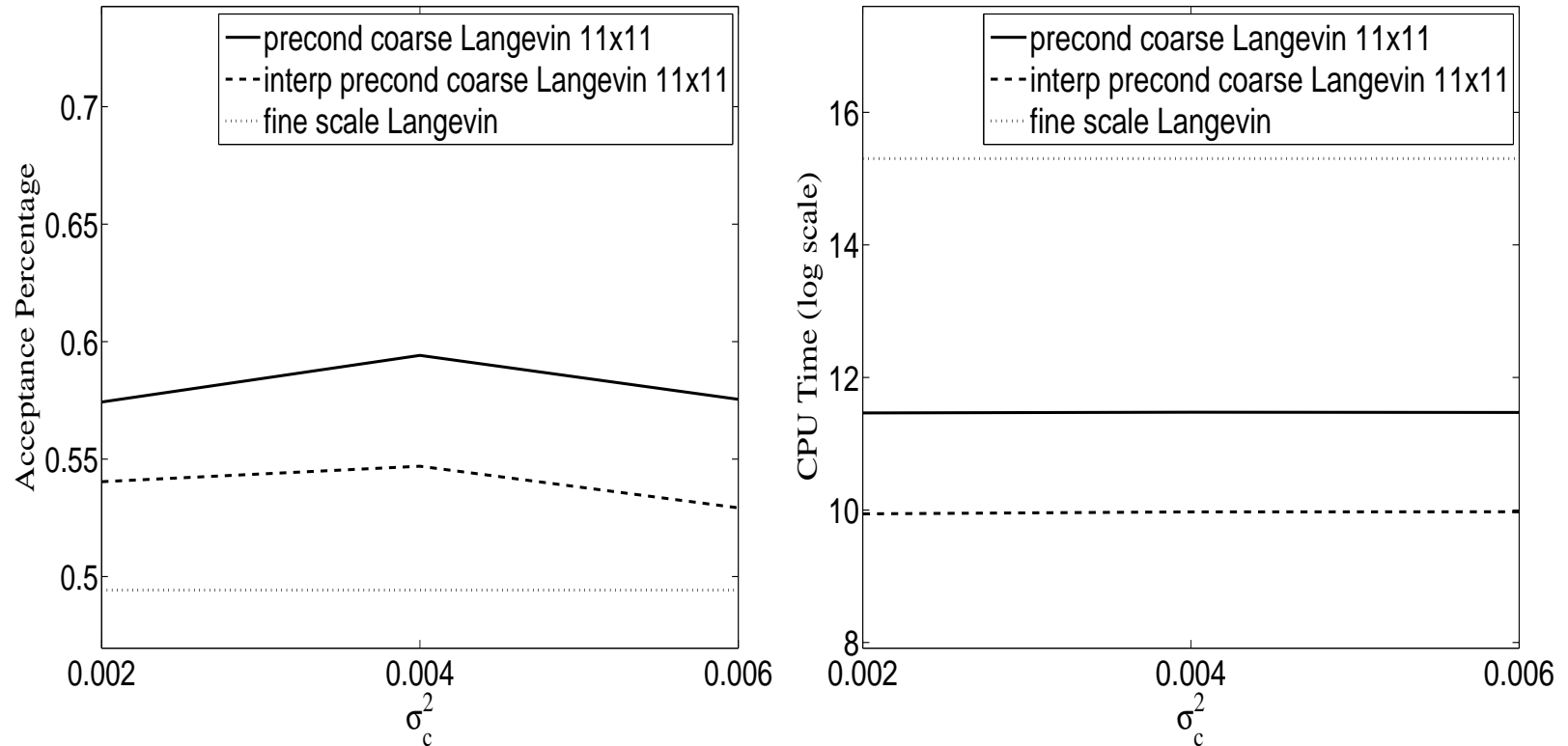Acceptance rate comparison, $\delta = 0.1$.
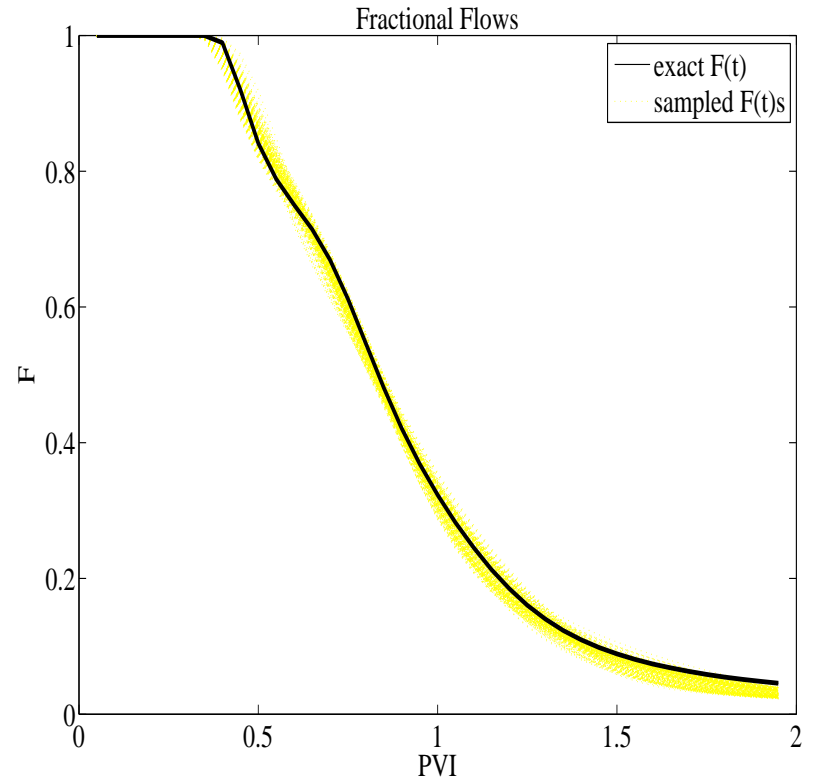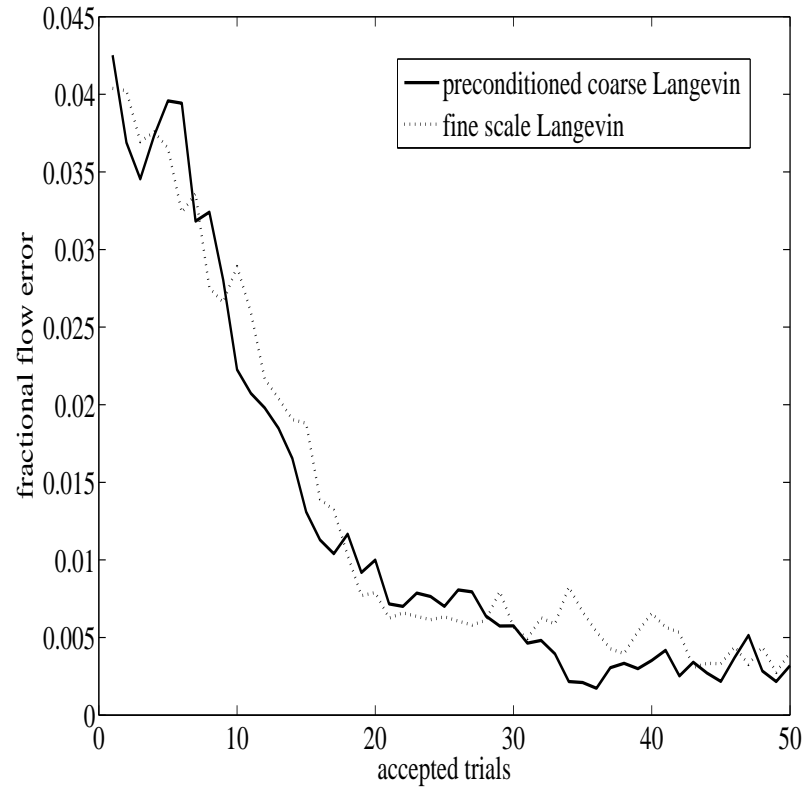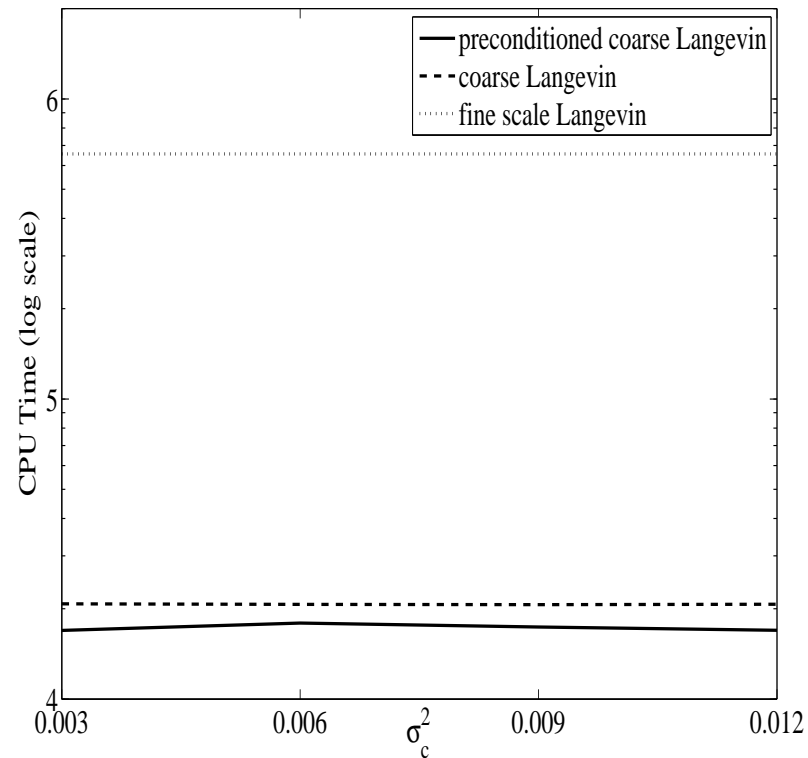
Numerical results

Numerical results

Exact

Realization 1

Realization 2

Realization 3

Realization 4
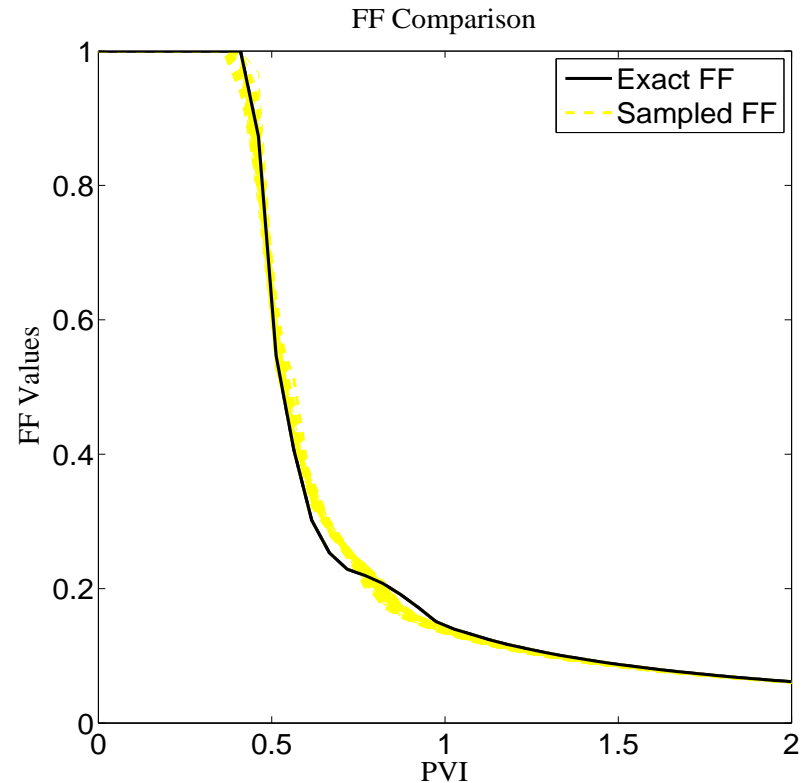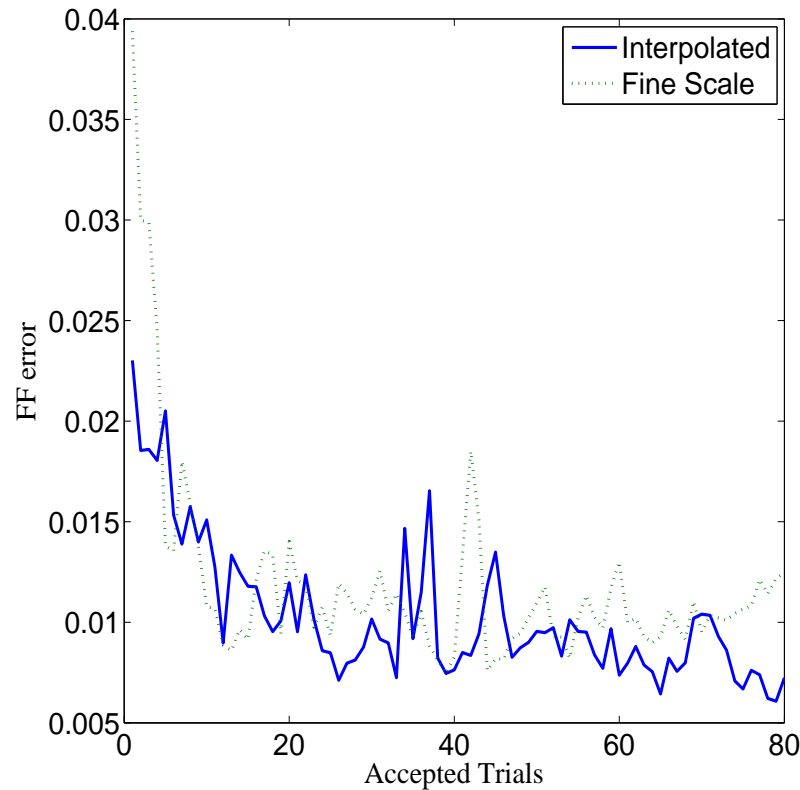
Realization 5

Numerical Results

Left: Acceptance rate comparison. Right: Natural log of CPU time (seconds) comparison. In each plot $\delta = 0.05$ and $\sigma_f^2 = 0.002$.
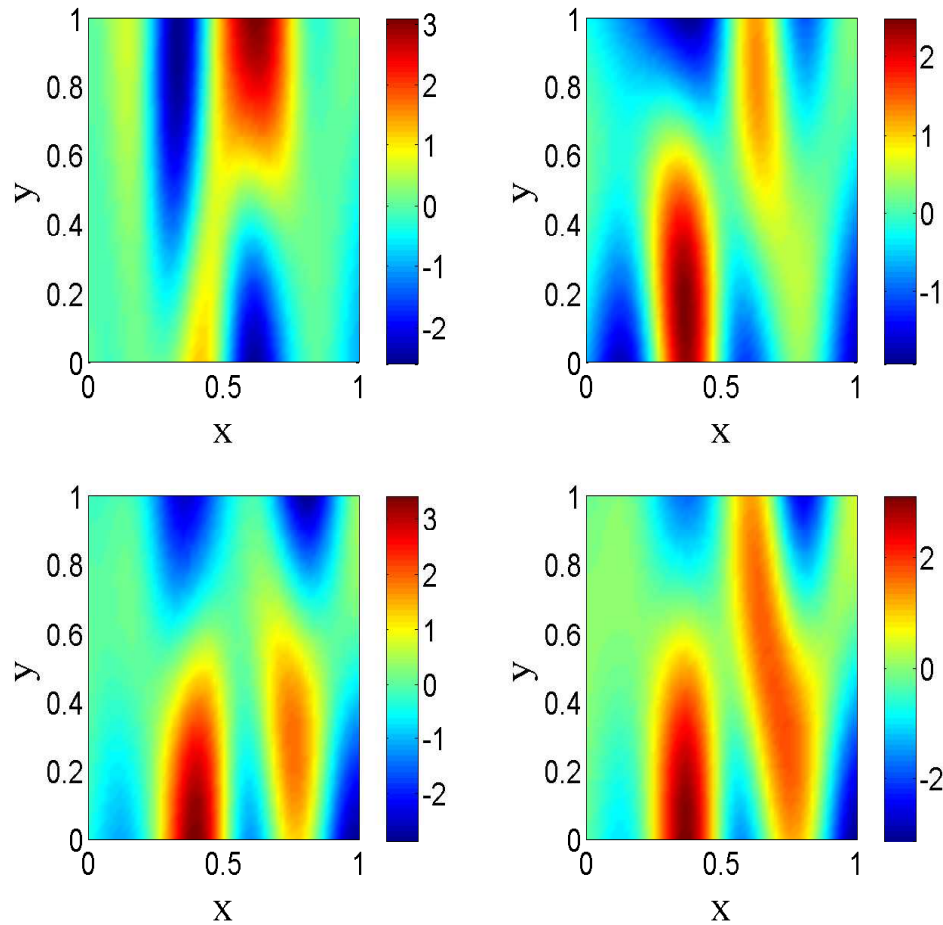
Numerical Results

CPU times (seconds) for Langevin algorithms. $\sigma_f^2 = 0.003$, $\delta = 0.05$, 7×7 coarse-grid.
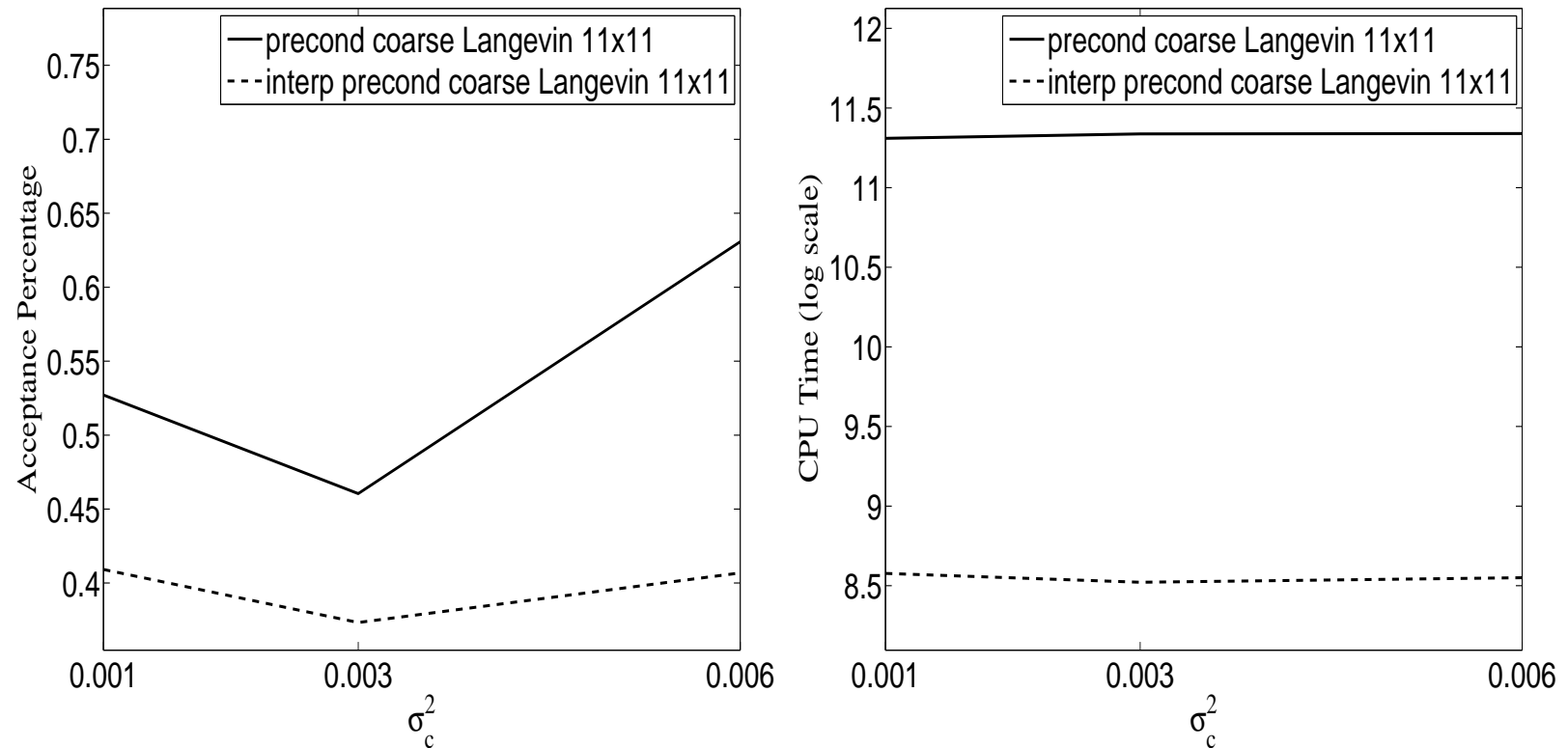
Numerical Results

Left: The fractional flow errors for the fine Langevin algorithm compared with interpolated Langevin algorithm. Right: The fractional flows of sampled realizations and the reference fractional flow. In these numerical tests, $\delta = 0.05$ and $\sigma_f^2 = 0.002$.
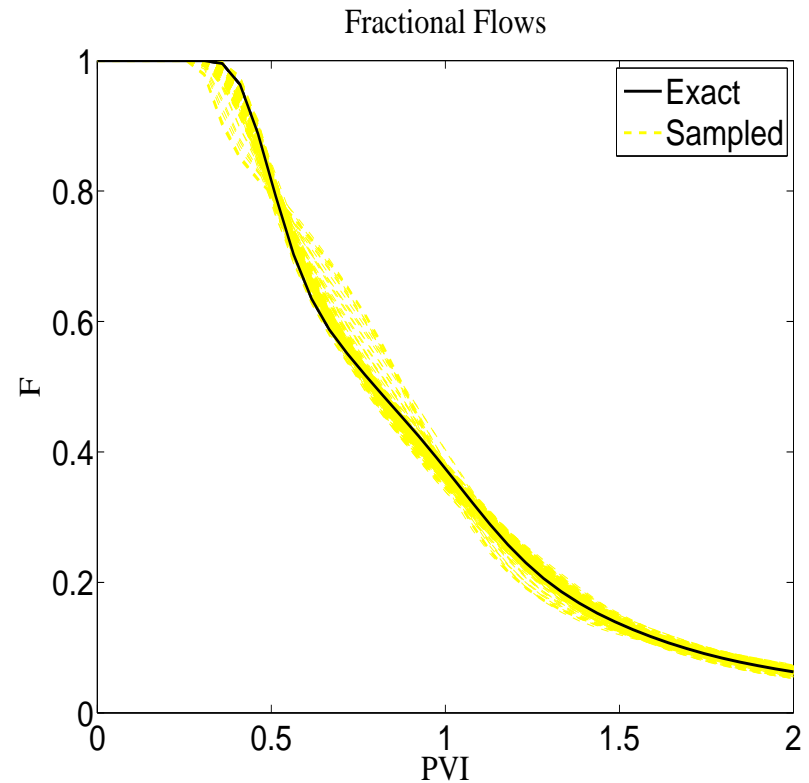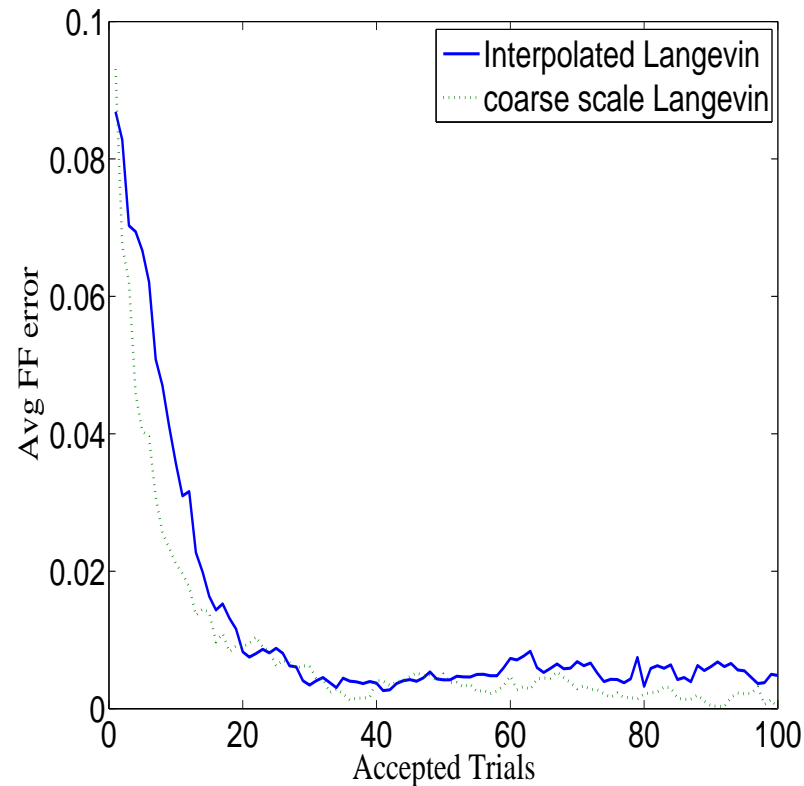
Numerical Results

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

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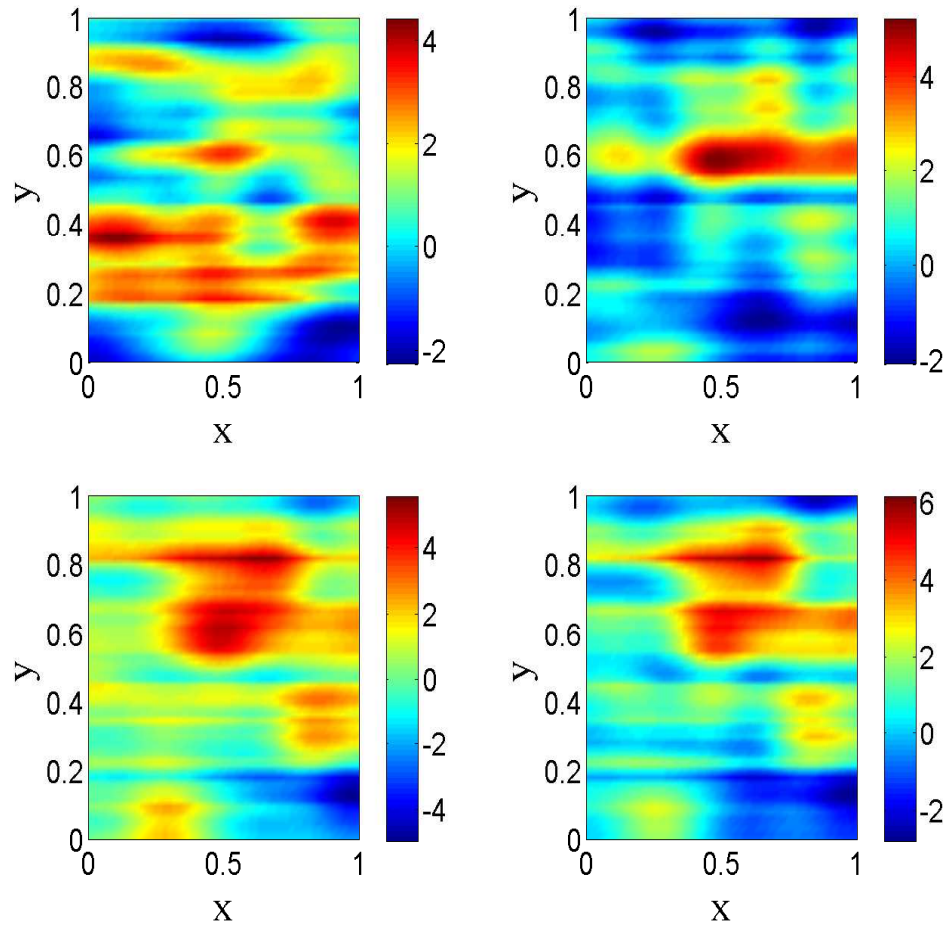
Upper left plot is the reference permeability. The other three plots are examples of accepted permeability realizations.

Conclusions

- Direct sampling using MH MCMC approaches is expensive

- Inexpensive coarse-scale models can be used to precondition Langevin MH simulations.

- Coarse-scale simulations are based on multiscale finite element type methods.

- Multiscale basis functions can be constructed to represent an ensemble of permeability fields.

- Numerical discretization of Langevin equation can help to improve the mixing of MC.

- Numerical results demonstrate CPU time can be reduced by two orders of magnitude.

- We have applied the proposed techniques for hydraulic conductivity estimation based on soil moisture data at different scales.