

# **Data Driven Simulation for Porous Media Flows. Solution Update**

Yalchin Efendiev

Texas A& M University

Collaborators: Craig C. Douglas (Yale and Kentucky), Robert Lodder (Kentucky), Paul Dostert (Texas A& M), Richard Ewing (Texas A& M), Victor Ginting (Texas A& M), Raytcho Lazarov (Texas A& M), Guan Qin (Texas A& M), Martin Cole (Utah), Greg Jones (Utah), Chris Johnson (Utah), Mohamed Iskandarani (Miami)

# Introduction

$$\frac{\partial C}{\partial t} - L(C) = S,$$

where  $S$  is source term.

$S$ , initial data  $C(x, 0)$ , ... are unknown.

- Our goal is to predict the solution in real time using sensor information (values of the solution at fixed or moving sparse locations).
- The fast forward computations are required.
- We would like to avoid solving differential equations for large “past” time interval and only use the information from a few previous time steps.
- Because the solution update is performed based on short time interval, the equations can be linearized using priors. Nonlinear equation and inversion can be used to obtain accurate prior information.

# Porous media flow and transport.

Linear transport model dominated by convection and diffusion:

$$\frac{\partial C}{\partial t} + v \cdot \nabla C - \nabla \cdot (D \nabla C) = 0 \text{ in } \Omega.$$

By Darcy's Law,  $v = -k \nabla p$ . Pressure  $p$  satisfies

$$-\nabla \cdot (k \nabla p) = 0$$

with some prescribed boundary conditions and initial condition/data  $C(\mathbf{x}, 0) = C^0(\mathbf{x})$ .  $C(\mathbf{x}, t)$  is the contaminant concentration over the porous medium  $\Omega$  and at time level  $t$ ,  $k$  is the permeability of the porous medium.  $D$  is the diffusion coefficient. The methodology described next is applicable to any linear operator  $L$ .

# Setting up objective function

- $N_s$  is the number of sensors installed in various points in the porous medium and  $\{\mathbf{x}_j\}_{j=1}^{N_s}$  denote such points.
- $N_t$  is how many times the concentration is measured in time and  $\{t_k\}_{k=1}^{N_t}$  denote such time levels.
- $\gamma_j(t_k)$  is the measured concentration at sensor located in  $\mathbf{x}_j$  and at time  $t_k$ .
- We seek initial data (the solution at time  $t_{n-1}$ ) in a finite dimensional space spanned by  $\tilde{C}_i^0(\mathbf{x})$ ,

$$\tilde{C}^0(\mathbf{x}) = \sum_{i=1}^{N_c} \alpha_i \tilde{C}_i^0(\mathbf{x}),$$

for some  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{N_c})$ .

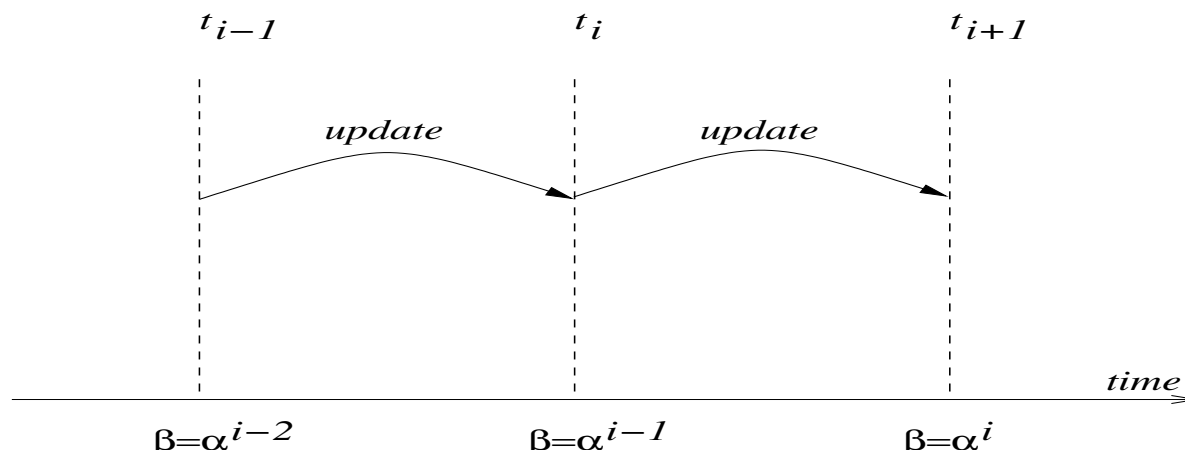
- Let  $\tilde{C}_i(\mathbf{x}, t)$  be the solution based on an initial condition  $\tilde{C}_i^0(\mathbf{x})$ . Then by superposition principle (because  $L$  is linear), the solution has the following form:

$$\tilde{C}(\mathbf{x}, t) = \sum_{i=1}^{N_c} \alpha_i \tilde{C}_i(\mathbf{x}, t).$$

# Objective function

$$F(\alpha) = \sum_{l=1}^{N_t} \sum_{j=1}^{N_s} \left( \sum_{i=1}^{N_c} \alpha_i \tilde{C}_i(\mathbf{x}_j, t_l) - \gamma_j(t_l) \right)^2 + \sum_{i=1}^{N_c} \kappa_i (\alpha_i - \beta_i)^2.$$

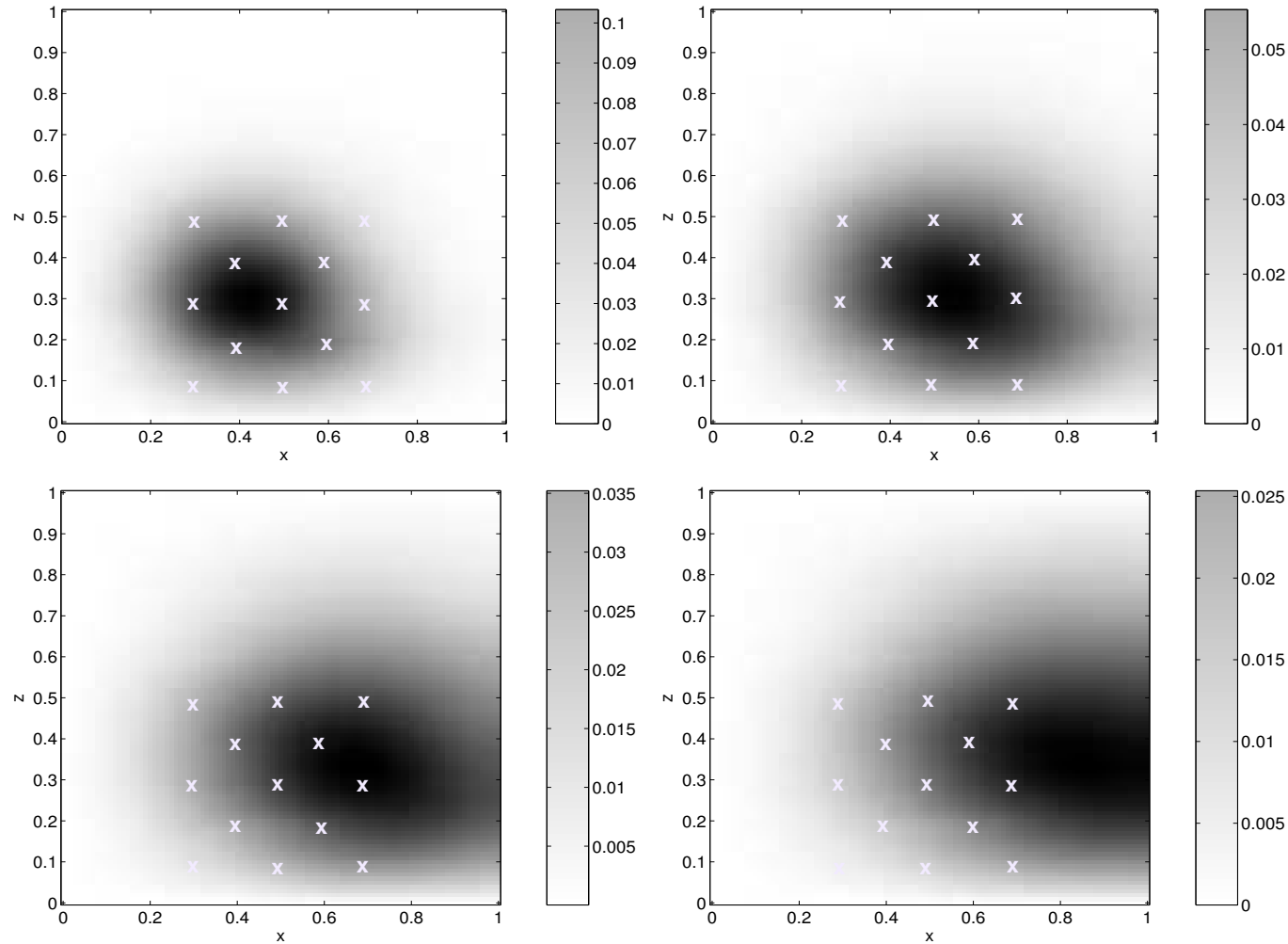
- Includes a penalty term that contains the prior information related to the initial data to regularize the problem.
- $\kappa = (\kappa_1, \kappa_2, \dots, \kappa_{N_c})$  is the penalty coefficients for  $\beta = (\beta_1, \beta_2, \dots, \beta_{N_c})$ .
- The prior information will be updated during the simulation to achieve higher accuracy. In particular,  $\beta_i = \alpha_i^{n-1}$  and  $\kappa_i = \kappa_i^n$ .



# Numerical setting

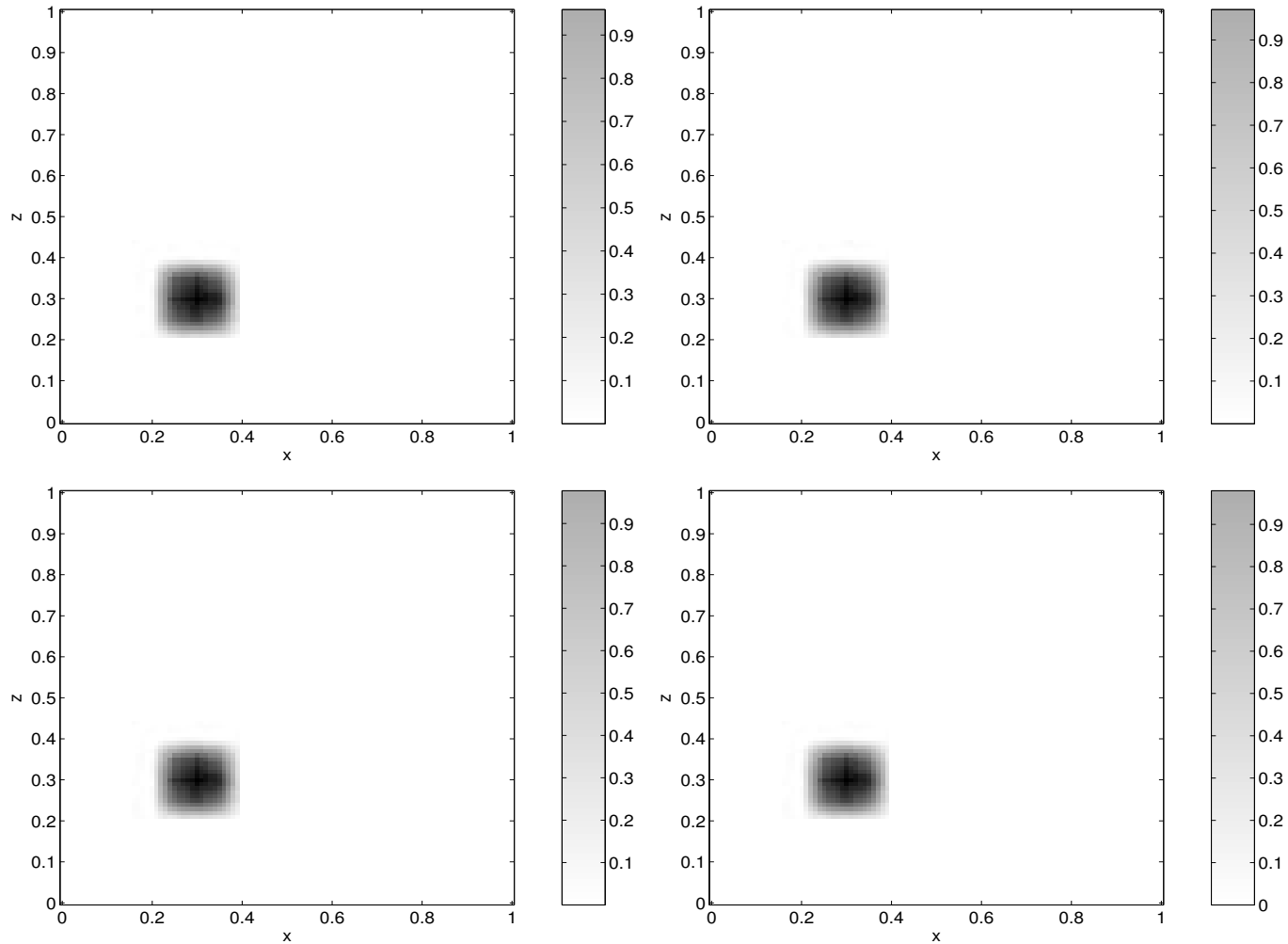
- The permeability  $k$  is generated with given correlation length  $l_x = 0.2$  and  $l_z = 0.02$ , with a spherical variogram using GSLIB algorithms.
- We set  $D = 0.1$ . Zero concentration at the inlet, bottom, and top edges, and a zero diffusion, i.e.,  $(D\nabla C) \cdot \vec{n} = 0$ , at the outlet edge, with  $\vec{n}$  being the unit normal vector pointing outward on the outlet edge.

# Numerical Results



Concentration at several time levels:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right). The (x) indicates the sensor location.

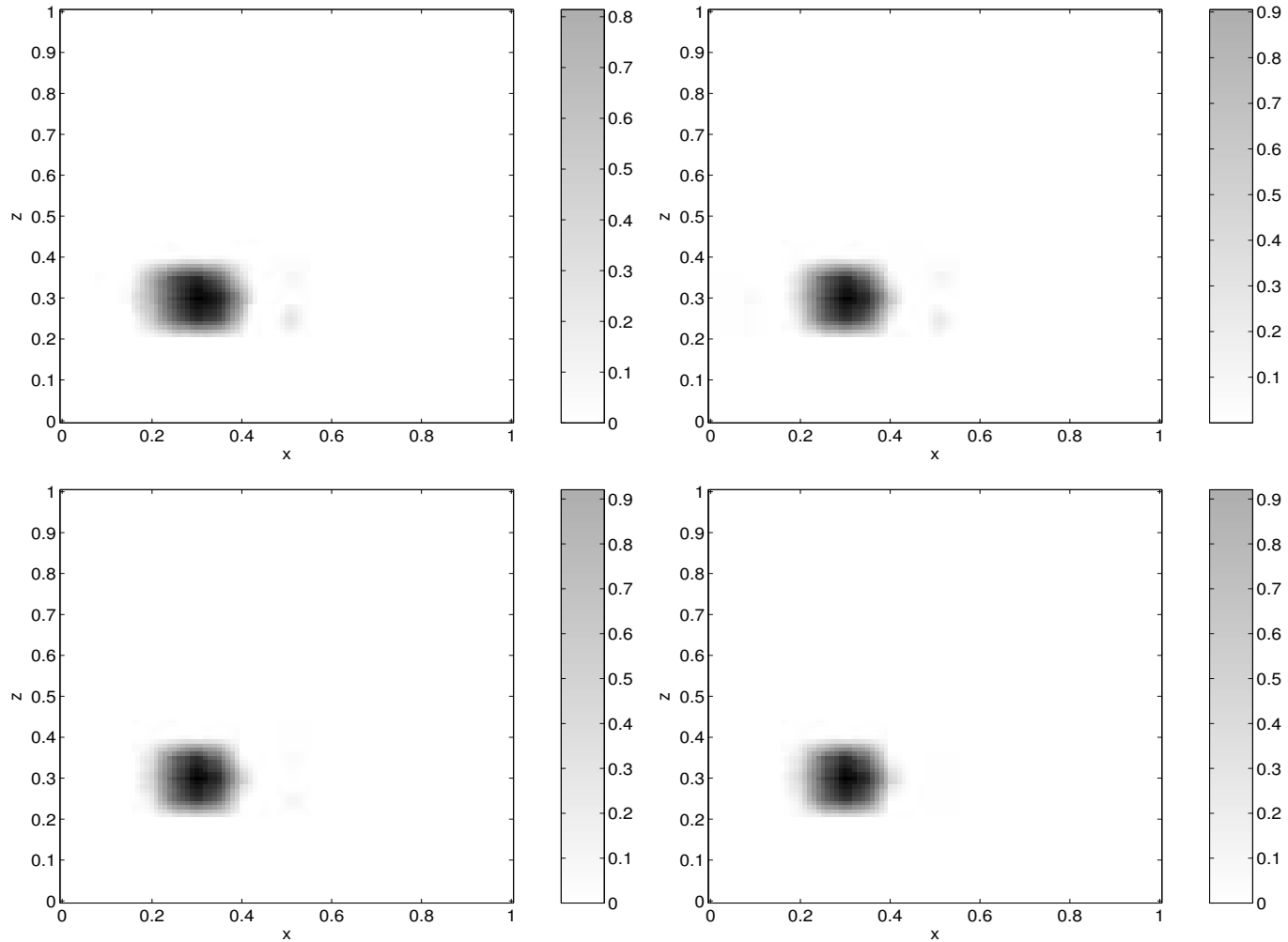
# Numerical Results



Updated initial data:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right). The prior assumes a known support.

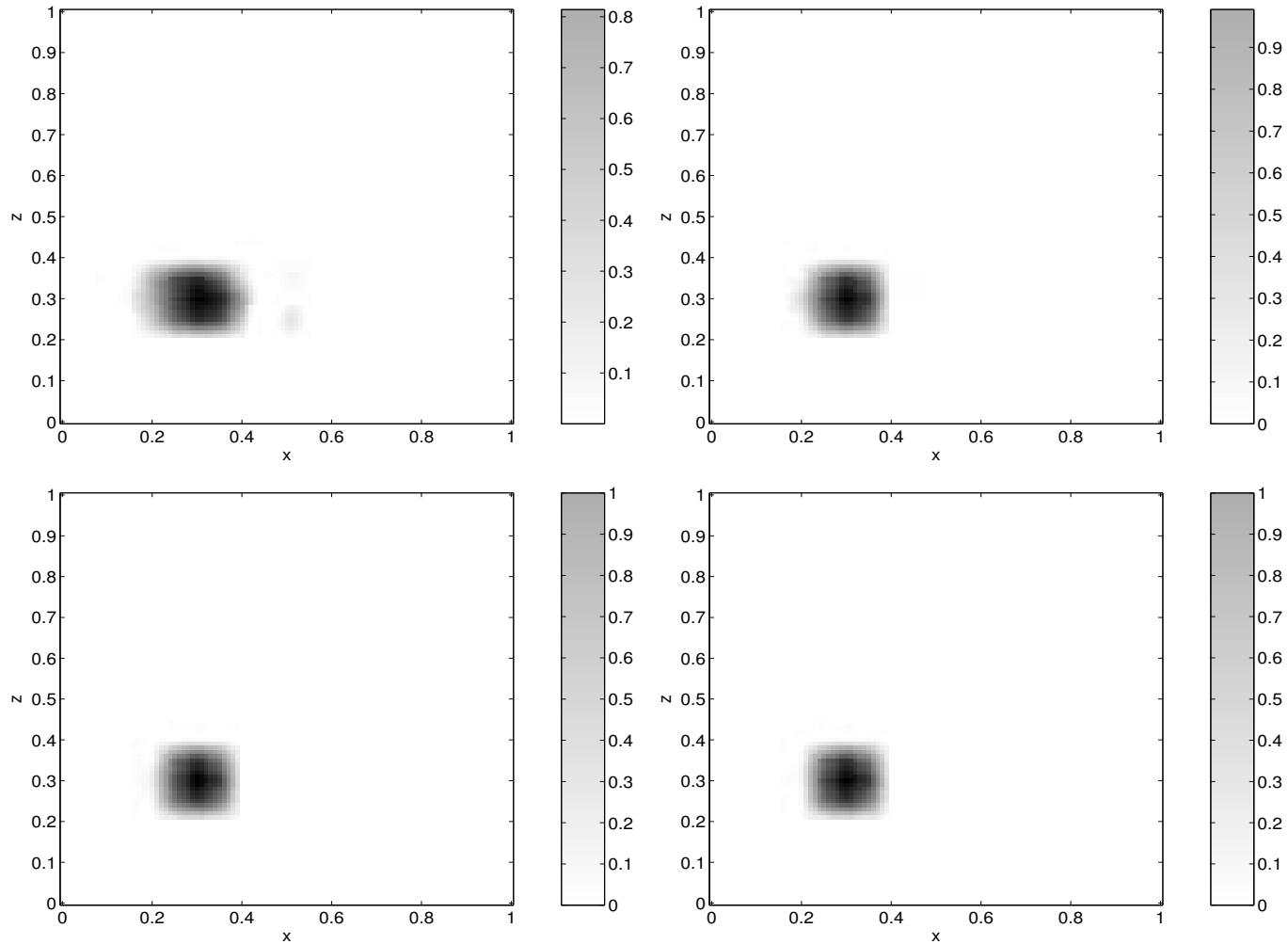


# Numerical Results



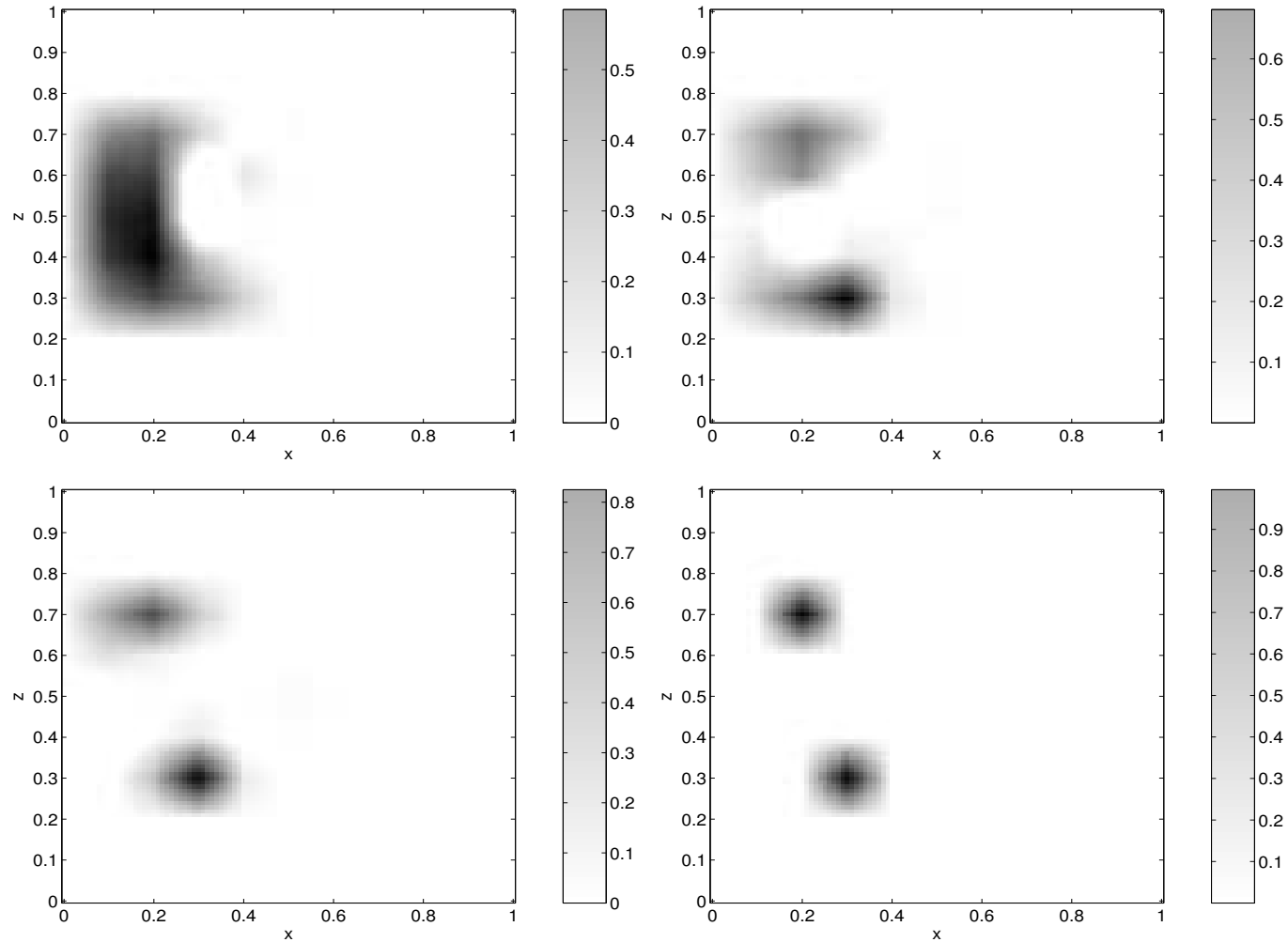
Updated initial data:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right). The prior for  $\beta$  assumes a larger support.

# Numerical Results



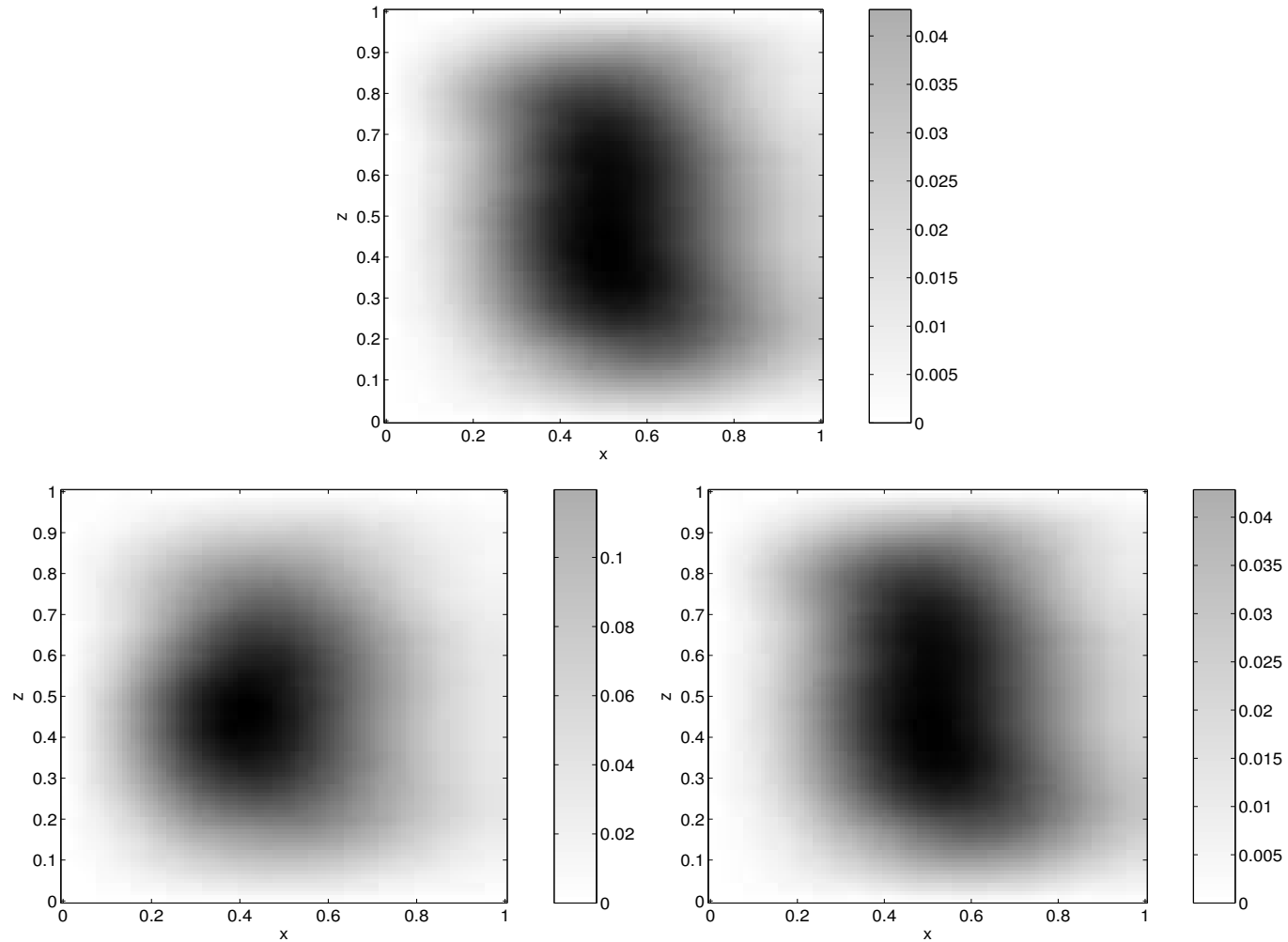
Updated initial data:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right). The prior for  $\beta$  assumes a larger support. Both  $\beta$  and  $\kappa$  are updated.

# Numerical Results



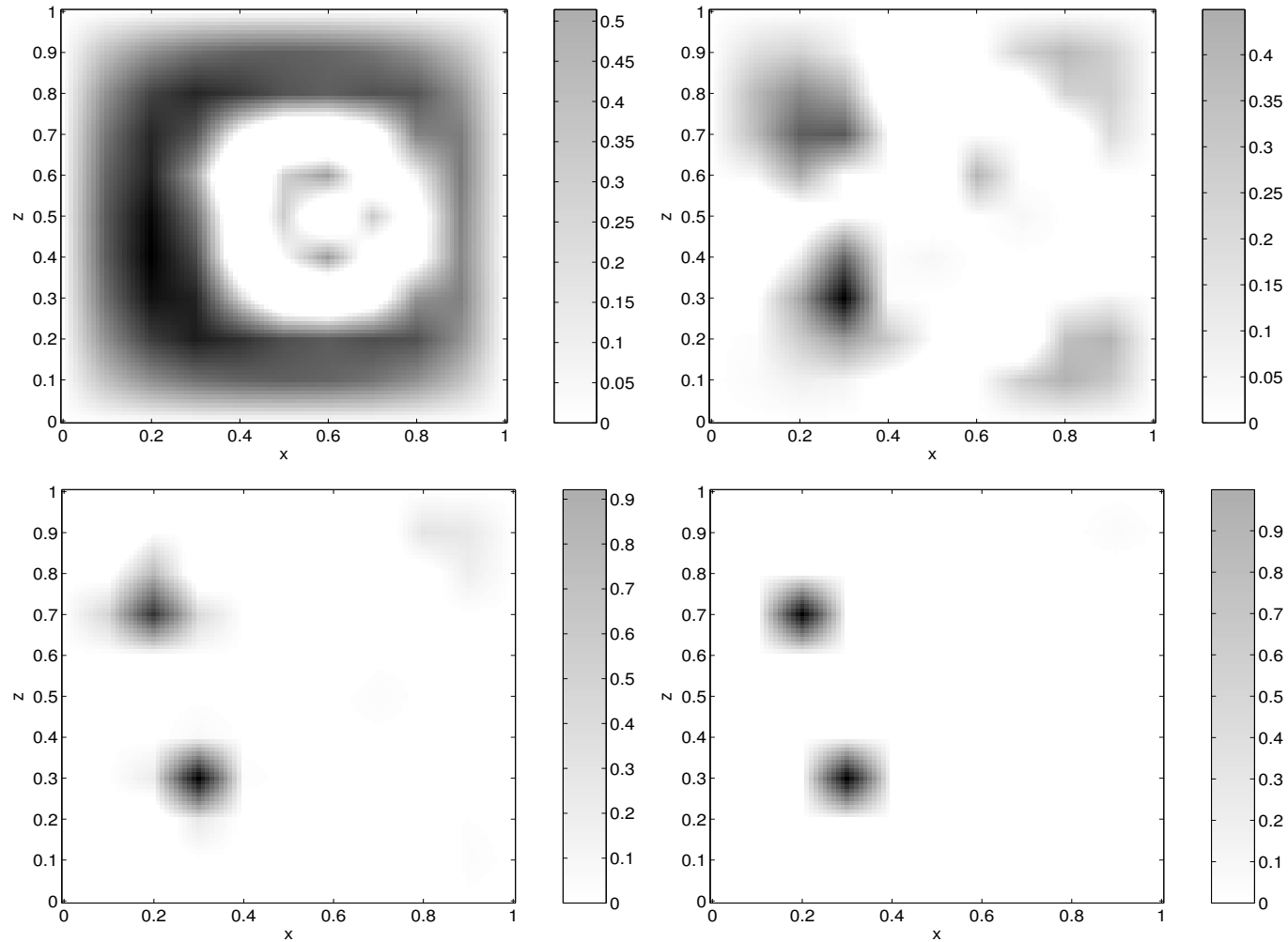
Updated initial data for two-peak initial data:  $t = 0.04$  (top left),  $t = 0.08$  (top right),  $t = 0.12$  (bottom left),  $t = 0.16$  (bottom right). The prior for  $\beta$  assumes a larger support that covers the support of both peaks. Both  $\beta$  and  $\kappa$  are updated.

# Numerical Results



Comparison of predicted concentration at  $t = 0.2$  for the case presented in previous figure (top) true concentration, (bottom left) concentration using first update, (bottom right) concentration using fourth update.

# Numerical Results



Updated initial data for two-peak initial data:  $t = 0.04$  (top left),  $t = 0.12$  (top right),  $t = 0.20$  (bottom left),  $t = 0.28$  (bottom right). The prior for  $\beta$  assumes a support equal to  $\Omega$  ( $\Omega_c = \Omega$ ). Both  $\beta$  and  $\kappa$  are updated.

# Efficient techniques

- The problem becomes more ill-posed as the size of parameter space increases.
- In general, the support of prior initial data can be taken to be large.
- Instead of approximating “initial” condition with piece-wise linear or piece-wise constant functions, we can seek its approximation with multi-scale basis functions. This will allow to reduce the size of parameter space and is applicable if the heterogeneities do not change and also for “large” times. If heterogeneities also change, a few local basis functions will be modified.
- The basis functions can be constructed based on the solution of

$$v \cdot \phi_i - D\Delta\phi_i = 0,$$

with  $\phi_i = \phi_i^0$ .

# Uncertainty propogation

- In general, the data contains the measurement errors. To account for these measurement errors, we use randomized maximum liklihood (RML) method.
- Denote prior by

$$P(\alpha) = \exp \left( - \sum_{i=1}^{N_c} \kappa_i (\alpha_i - \beta_i)^2 \right)$$

and the likelihood by

$$P(\gamma|\alpha) = \exp \left( - \sum_{j=1}^{N_s} \left( \sum_{i=1}^{N_c} \alpha_i \tilde{C}_i(\mathbf{x}_j, t) - \gamma_j(t) \right)^2 \right).$$

Then the posterior dostrubution, we would like to sample is given by

$$P(\alpha|\gamma) \propto P(\gamma|\alpha)P(\alpha) = \exp(-F(\alpha|\gamma))$$

where

$$F(\alpha) = F(\alpha|\gamma) = \sum_{j=1}^{N_s} \left( \sum_{i=1}^{N_c} \alpha_i \tilde{C}_i(\mathbf{x}_j, t) - \gamma_j(t) \right)^2 + \sum_{i=1}^{N_c} \kappa_i (\alpha_i - \beta_i)^2.$$

# Uncertainty propagation

- We need rigorous sampling from the posterior distribution (it is Gaussian).
- If we denote  $\sum_{i=1}^{N_c} \alpha_i \tilde{C}_i(\mathbf{x}_j, t) = H\alpha$ , where  $H$  is the matrix, then  $F(\alpha)$  can be written as

$$F(\alpha) = \sum_{i=1}^{N_s} (H\alpha - \gamma_i(t))^2 + \sum_{i=1}^{N_c} \kappa_i (\alpha_i - \beta_i)^2 =$$
$$(\gamma - H\alpha)^T C_1^{-1} (\gamma - H\alpha) + (\alpha - \beta)^T C_2^{-1} (\alpha - \beta)$$

- The sampling from  $\exp(-F(\alpha|\gamma))$  can be accomplished in the following way:  
(1) Choose unconditional realizations of  $\beta_{uc}$  and  $\gamma_{uc}$  from

$$\exp(-(\beta_{uc} - \beta)^T C_2^{-1} (\beta_{uc} - \beta)) \quad \text{and} \quad \exp(-(\gamma_{uc} - \gamma)^T C_1^{-1} (\gamma_{uc} - \gamma))$$

respectively.

- (2) Solve the minimization problem for objective functional

$$(\gamma_{uc} - H\alpha)^T C_1^{-1} (\gamma_{uc} - H\alpha) + (\alpha - \beta_{uc})^T C_2^{-1} (\alpha - \beta_{uc})$$



# Uncertainty propagation

- It can be shown that the above approach provides rigorous sampling of the posterior.
- The above approach is similar to Ensemble Kalman Filter.
- Kalman filter samples from

$$P(\alpha|\gamma) \propto \exp(-(\gamma - H\alpha)^T C_1^{-1}(\gamma - H\alpha)) \exp(-(\alpha - \beta)^T C_2^{-1}(\alpha - \beta)).$$

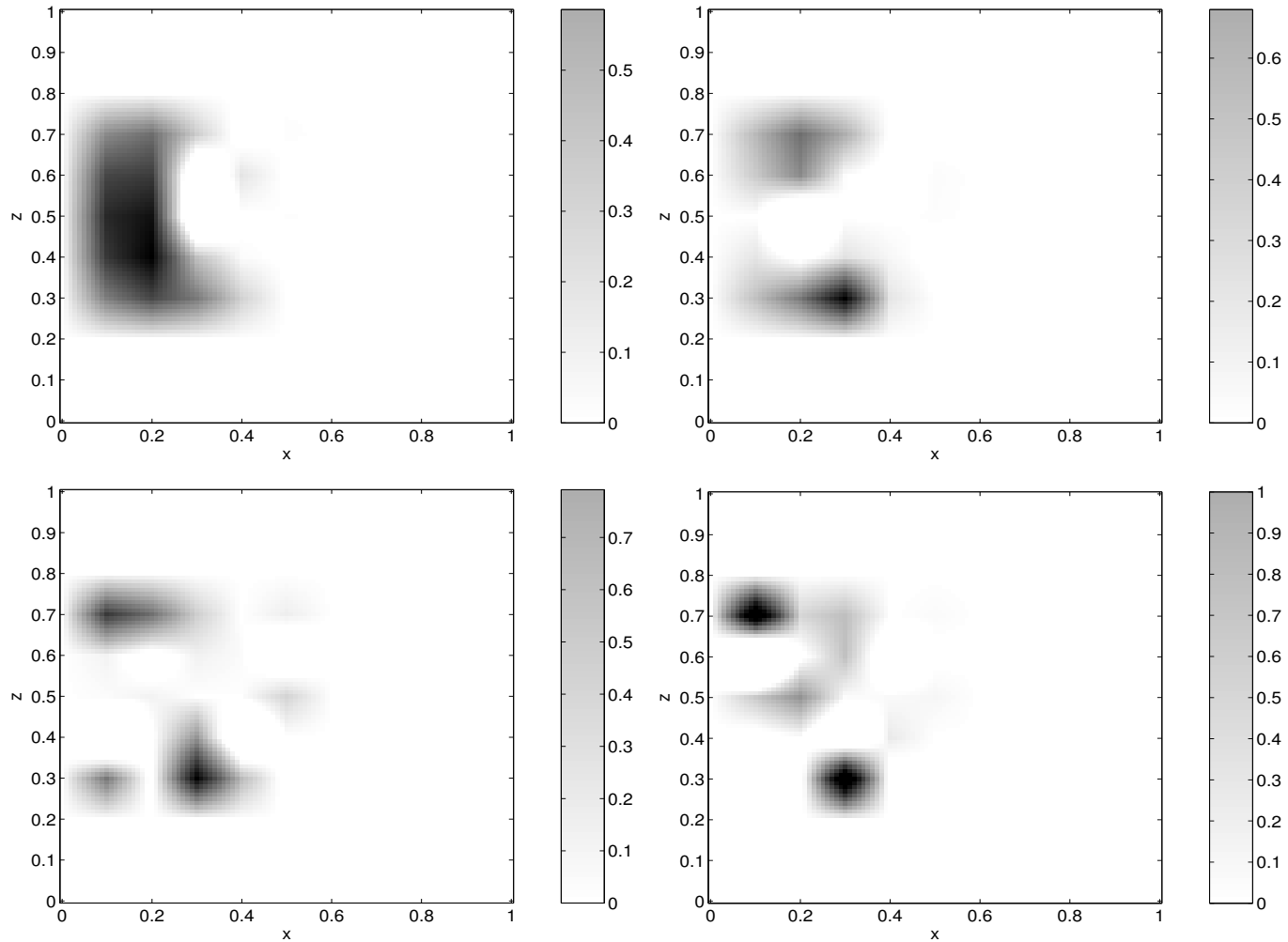
The above posterior distribution is Gaussian

$$P(\alpha|\gamma) \propto \exp(-(\alpha - \bar{\alpha})^T C_3^{-1}(\alpha - \bar{\alpha}))$$

where  $\bar{\alpha}$  solves the minimization problem  $\min F(\alpha)$ .

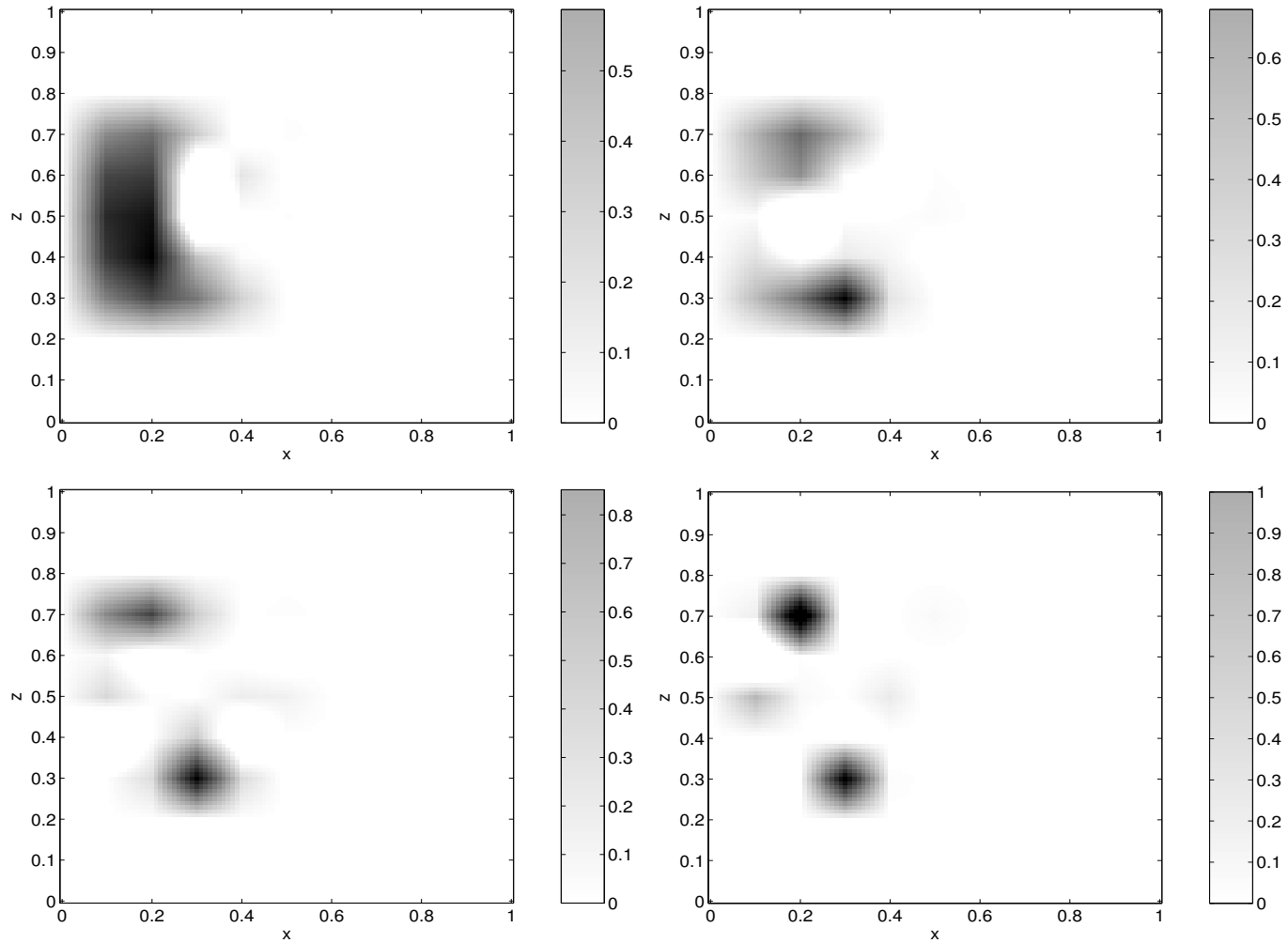
- Monte Carlo version of Kalman Filter is called Ensemble Kalman Filter (EnKF).
- RML can replace EnKF. It was found that RML can perform more accurate sampling for nonlinear posteriors compared to EnKF.
- We can use RML for sampling from nonlinear distribution when linearity assumption is violated, or we can use more rigorous and fast sampling techniques which will be described in permeability estimation.

# Numerical Results



First realization of updated initial data:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right).

# Numerical Results



Second realization of updated initial data:  $t = 0.1$  (top left),  $t = 0.2$  (top right),  $t = 0.3$  (bottom left),  $t = 0.4$  (bottom right).

# Generalization

Consider

$$\frac{\partial C}{\partial t} - L(C) = S, C(x, 0) = C^0(x).$$

We split the problem into two problems:

$$\frac{\partial C_1}{\partial t} - L(C_1) = 0, C_1(x, 0) = C^0(x),$$

$$\frac{\partial C_2}{\partial t} - L(C_2) = S, C_2(x, 0) = 0.$$

$$C = C_1 + C_2.$$

For  $C_1$ , representing initial condition by  $C^0(x) = \sum_{i=1}^{N_D} \lambda_i \varphi_i(x)$ , we have

$C_1 = \sum_{i=1}^{N_D} \lambda_i \varphi_i(x, t)$ , where

$$\frac{\partial \varphi_i(x, t)}{\partial t} - L(\varphi_i(x, t)) = 0, \varphi_i(x, 0) = \varphi_i^0(x).$$

Thus,  $C_1(x, t) = \sum_{i=1}^{N_D} \lambda_i \varphi_i(x, t)$ .

For  $C_2$ , we solve

$$\frac{\partial \psi_k(x, t)}{\partial t} - L(\psi_k(x, t)) = \delta_k(x), \quad C(x, 0) = 0,$$

where  $\delta_k(x)$  is unit pulse at location  $x = x_k$ . Then

$$C_2(x, t) = \sum_{k=1}^{N_c} \alpha_k \psi_k(x, t).$$

Thus,

$$C(x, t) = \sum_{i=1}^{N_D} \lambda_i \varphi_i(x, t) + \sum_{k=1}^{N_c} \alpha_k \psi_k(x, t),$$

where  $\lambda_i$  and  $\alpha_k$  are unknown.

Note that in general we don't know the temporal functionality of the pulse. In our computations, the latter is linearized.

# Objective function

$$F(\alpha, \lambda) = \sum_{j=1}^{N_s} \left[ \left( \sum_{k=1}^{N_c} \alpha_k \psi_k(x_j, t) + \sum_{k=1}^{N_D} \lambda_k \varphi_k(x_j, t) - \gamma_j(t) \right)^2 \right] + \sum_{k=1}^{N_c} \tilde{\kappa}_k (\alpha_k - \tilde{\beta}_k)^2 + \sum_{k=1}^{N_D} \hat{\kappa}_k (\lambda_k - \hat{\beta}_k)^2.$$

If we denote  $N = N_c + N_d$ ,  $\mu = [\alpha_1, \dots, \alpha_{N_c}, \lambda_1, \dots, \lambda_{N_D}]$ ,  $\eta(x, t) = [\psi_1, \dots, \psi_{N_c}, \varphi_1, \dots, \varphi_{N_D}]$ ,  $\beta = [\tilde{\beta}_1, \dots, \tilde{\beta}_{N_c}, \hat{\beta}_1, \dots, \hat{\beta}_{N_D}]$  and  $\kappa = [\tilde{\kappa}_1, \dots, \tilde{\kappa}_{N_c}, \hat{\kappa}_1, \dots, \hat{\kappa}_{N_D}]$  then

$$F(\mu) = \sum_{j=1}^{N_s} \left[ \left( \sum_{k=1}^N \mu_k \eta_k(x_j, t) - \gamma_j(t) \right)^2 \right] + \sum_{k=1}^N \kappa_k (\mu_k - \beta_k)^2.$$

# Permeability estimation

Denote

$$M(\gamma) = \{\gamma_j(t_k), j = 1, \dots, N_s, k = 1, \dots, N_t\}.$$

Then,

$$P(k(x) | M(\gamma)) \propto P(M(\gamma) | k(x))P(k(x)),$$

This is nonlinear posterior.

**Algorithm (Metropolis-Hasting MCMC)**

- Step 1. At  $k_n$  generate  $k$  from  $q(k|k_n)$  (instrumental proposal distribution).
- 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  $1 - p(k_n, k)$ .

This algorithm constructs Markov chain with steady state distribution  $\pi(k)$ .

# Construction of priors

Prior for permeability field is constructed based on two-point correlation function of  $Y(x)$ , where  $k(x) = \exp(Y(x))$ . Denote the covariance function of  $Y$  as

$$R(x, y) = E[Y(x)Y(y)].$$

Consider a complete basis in  $L_2$

$$\int_{\Omega} R(x, y)\phi_k(y)dy = \lambda_k\phi_k(x), \quad k = 1, 2, \dots,$$

where  $\lambda_k = E[Y_k^2] > 0$ . Furthermore, we have

$$R(x, y) = \sum_{k=1}^{\infty} \lambda_k \phi_k(x)\phi_k(y).$$

Denote  $\theta_k = Y_k/\sqrt{\lambda_k}$ , then  $\theta_k$  satisfies  $E(\theta_k) = 0$  and  $E(\theta_i\theta_j) = \delta_{ij}$ . Then

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



Typically, the eigenvalues decay fast for log-Gaussian fields,  
At sensor locations, the permeability field is known. This condition is imposed by setting

$$\sum_k \sqrt{\lambda_k} \theta_k \phi_k(x_j) = \alpha_j,$$

where  $\alpha_j$  are prescribed constants.

# Efficient permeability estimation

- Direct (full) MCMC simulations are usually prohibitively expensive, because each proposal requires a fine-scale computation.
- Direct MCMC usually requires many (thousand) of iterations for the convergence to a steady state, where each iteration involves the computation of the fine-scale solution over a large time interval.
- The acceptance rate of direct MCMC is usually small, i.e., most proposals will be rejected.
- One way to improve direct MCMC is to increase the acceptance rate by modifying the proposal. We propose an algorithm, where the proposal distribution is modified using coarse-scale model.

# Preconditioning of MCMC

## Algorithm (preconditioned MCMC)

- Step 1. At  $k_n$ , generate  $k$  from distribution  $q(k|k_n)$ .
- Step 2. Accept  $k$  as a proposal for *the fine-scale model* with probability

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

i.e. pass  $k$  or  $k_n$  as a proposal to the fine-scale model with probability  $g(k_n, k)$  or  $1 - g(k_n, k)$  respectively. Here  $\pi^*(k) = P(k^*|F)$ . Therefore, the final proposal to the fine-scale model is generated from the effective instrument distribution

$$Q(k|k_n) = g(k_n, k)q(k|k_n) + \left( 1 - \int g(k_n, k)q(k|k_n)dk \right) \delta_{k_n}(k).$$

- Step 3. Accept  $k$  as a sample with probability

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

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

# Convergence of modified Markov Chain

First it can be shown that

$$\rho(k_n, k) = \min \left( 1, \frac{\pi(k)\pi^*(k_n)}{\pi(k_n)\pi^*(k)} \right).$$

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

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

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

# Convergence of modified Markov Chain

The transition probability of the chain is defined by

$$K(k_n, k) = \rho(k_n, k)Q(k|k_n), \quad k \neq k_n, \quad \text{and} \quad K(k_n, \{k_n\}) = 1 - \int_{k \neq k_n} \rho(k_n, k)Q(k|k_n)dk$$

**Lemma.**

$$\pi(k_n)K(k_n, k) = \pi(k)K(k, k_n)$$

for any  $k, k_n \in \mathcal{E}$ .

**Lemma.** If the proposal  $q(y|x)$  satisfies positivity condition, then the chain  $\{k_n\}$  generated by the preconditioned MCMC method is strongly  $\pi$ -irreducible and aperiodic.

**Theorem.** Suppose proposal is positive, then the chain generated by the preconditioned MCMC method is ergodic: for any  $h \in L^1(\pi)$ ,

$$\lim_{T \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N h(k_n) = \int h(k)\pi(k)dk,$$

and the distribution of the chain  $\{k_n\}$  converges to  $\pi(k)$  in the total variation norm

$$\lim_{n \rightarrow \infty} \sup_A |K^n(k_0, A) - \pi(A)| = 0$$

# Instrumental distribution

- Several types of instrumental proposal distribution can be considered: (1) independent sampler; (2) random walk sampler; (3) Langevin sampler.
- In the case of independent sampler, the proposal distribution  $q(k|k_n)$  is chosen to be independent of  $k_n$  and equal to the prior distribution.
- In random walk sampler, the proposal distribution depends on the previous value of the permeability field and given by

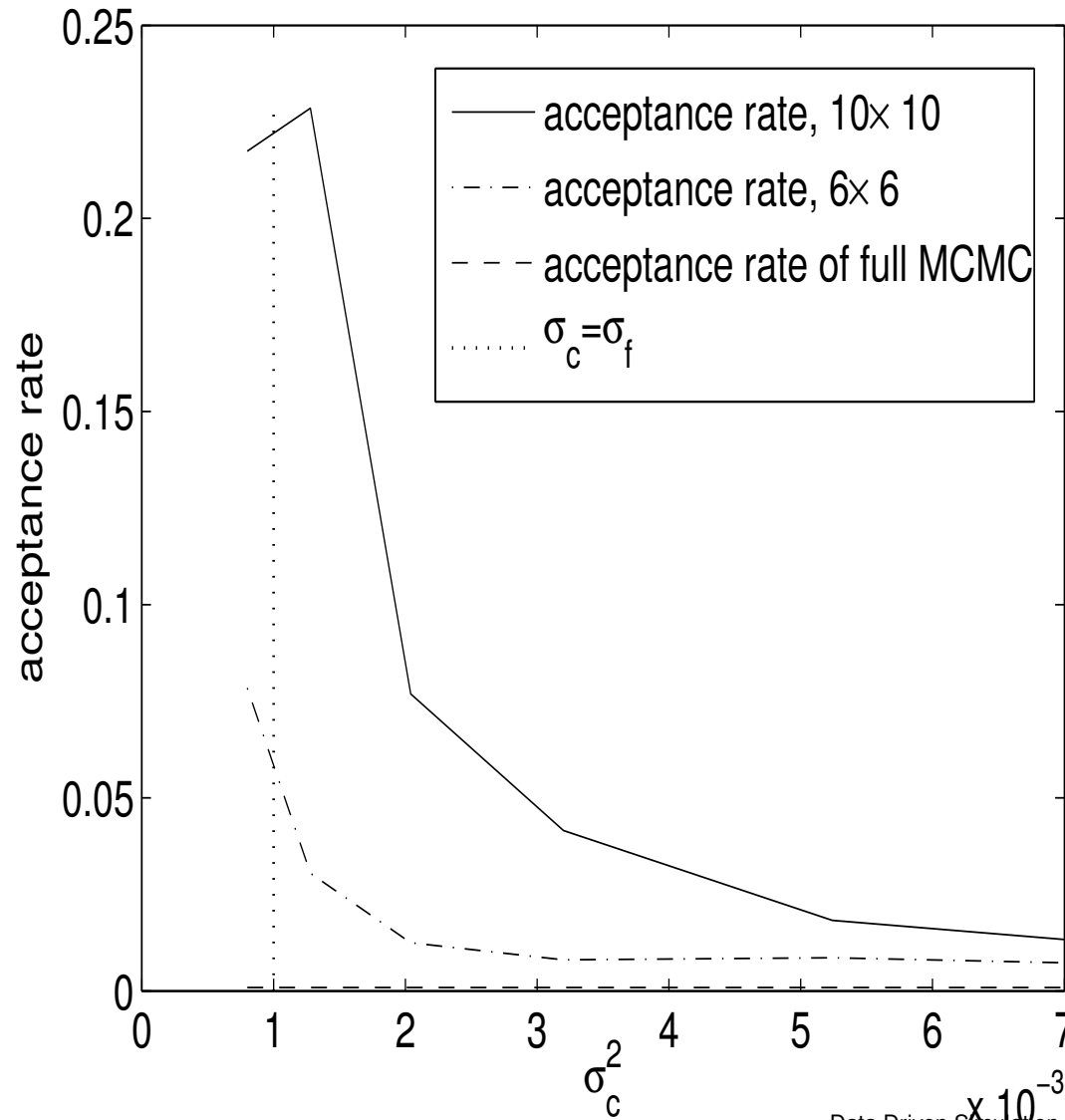
$$Y = k_n + \epsilon_n,$$

where  $\epsilon_n$  is a random perturbation with prescribed variance.

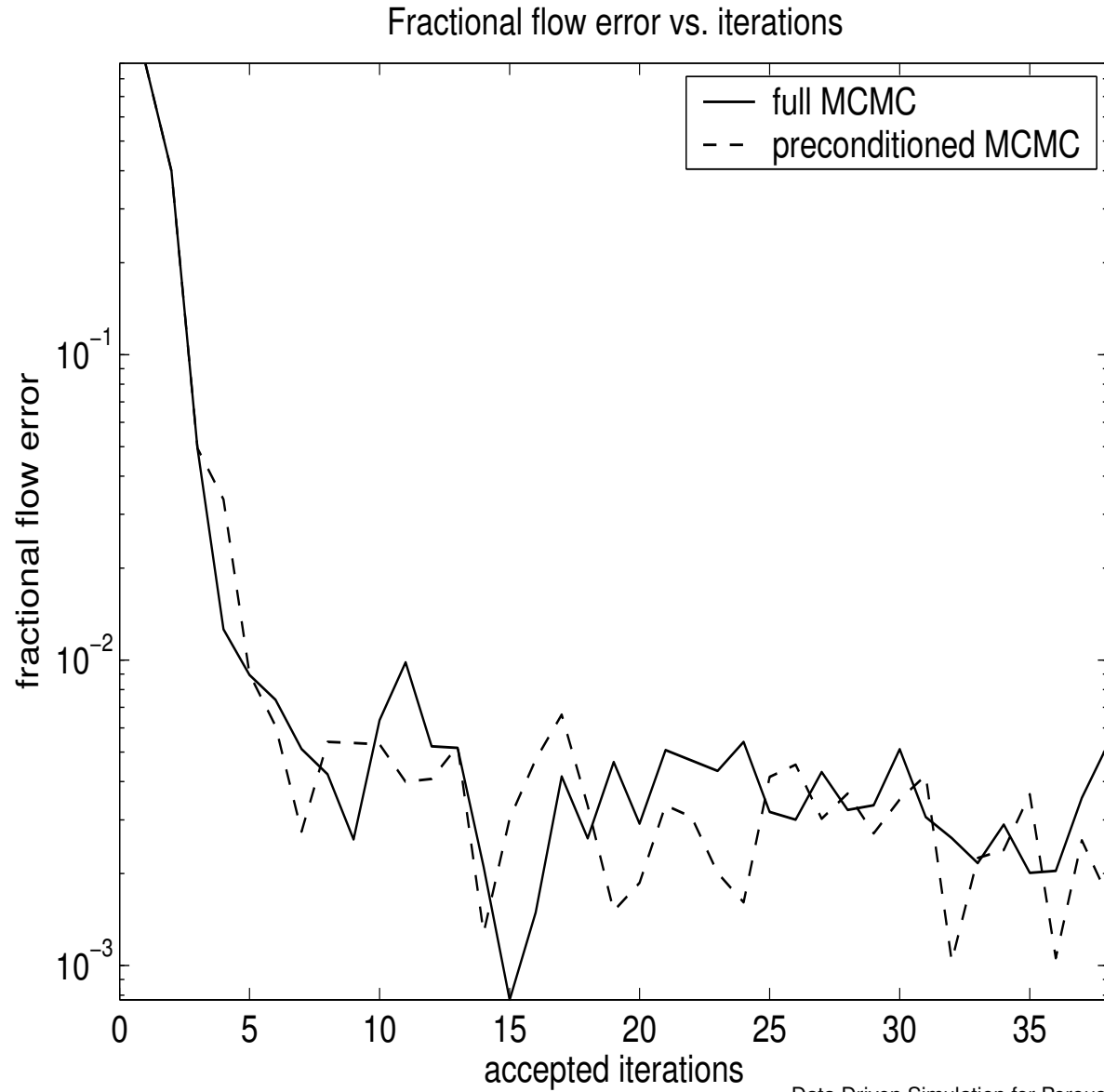
- In Langevin sampler, the proposal  $Y$  is chosen as

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

# Independent sampler

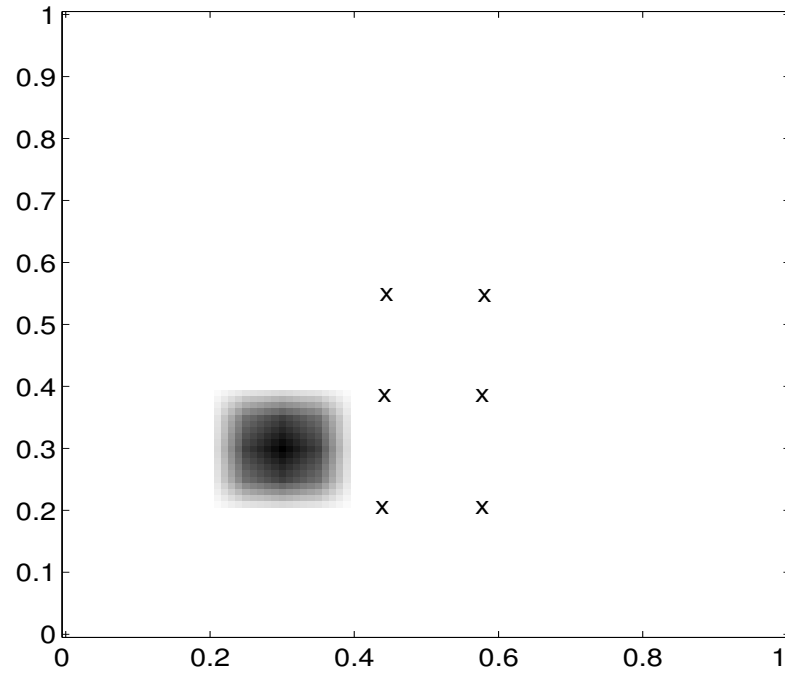


# Numerical results

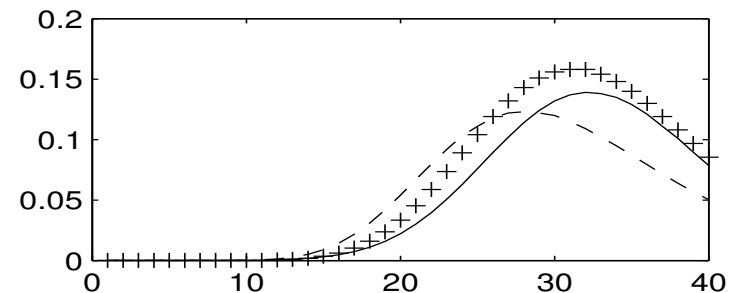
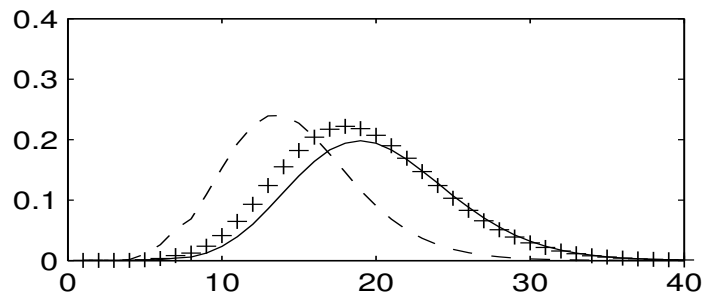
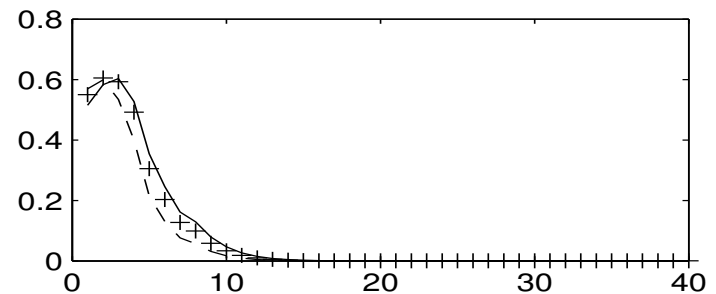
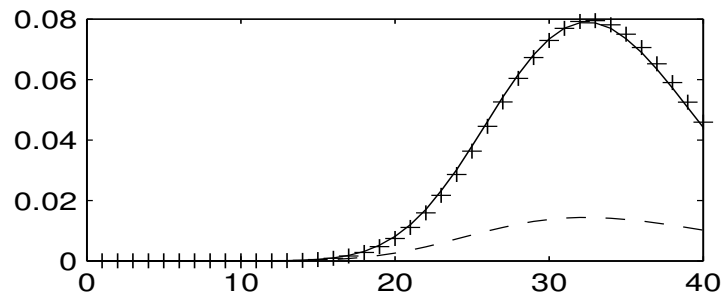
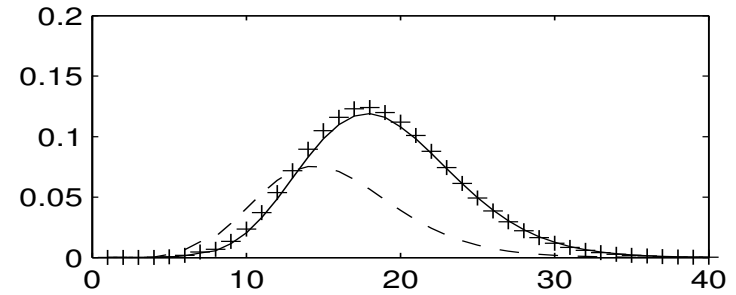
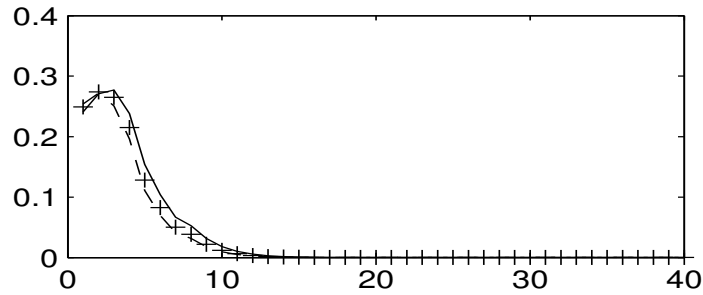




# Numerical results

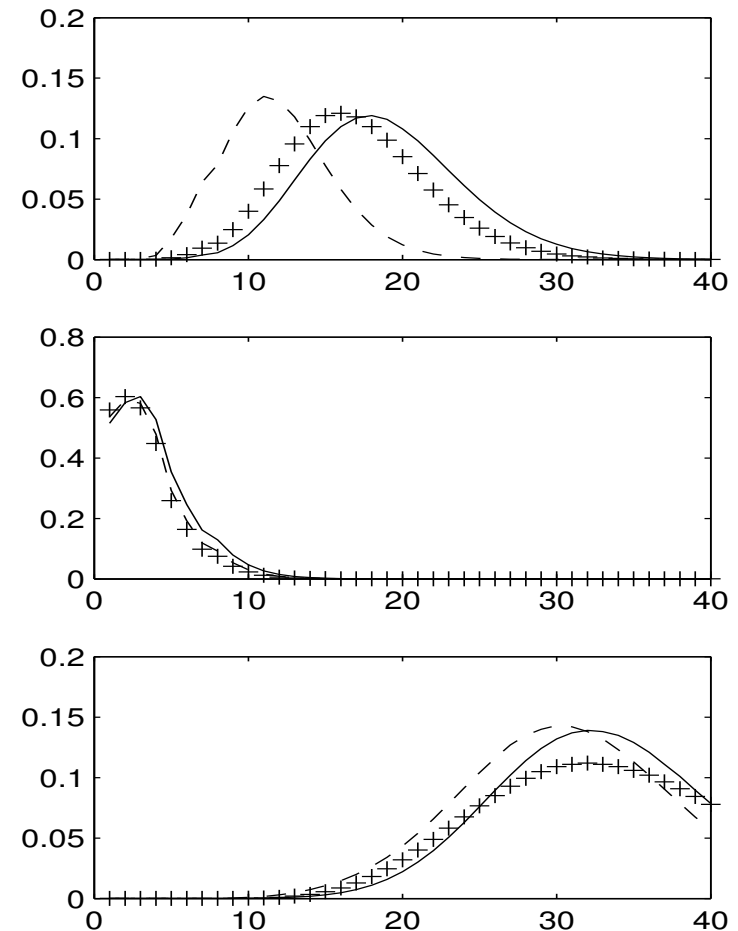
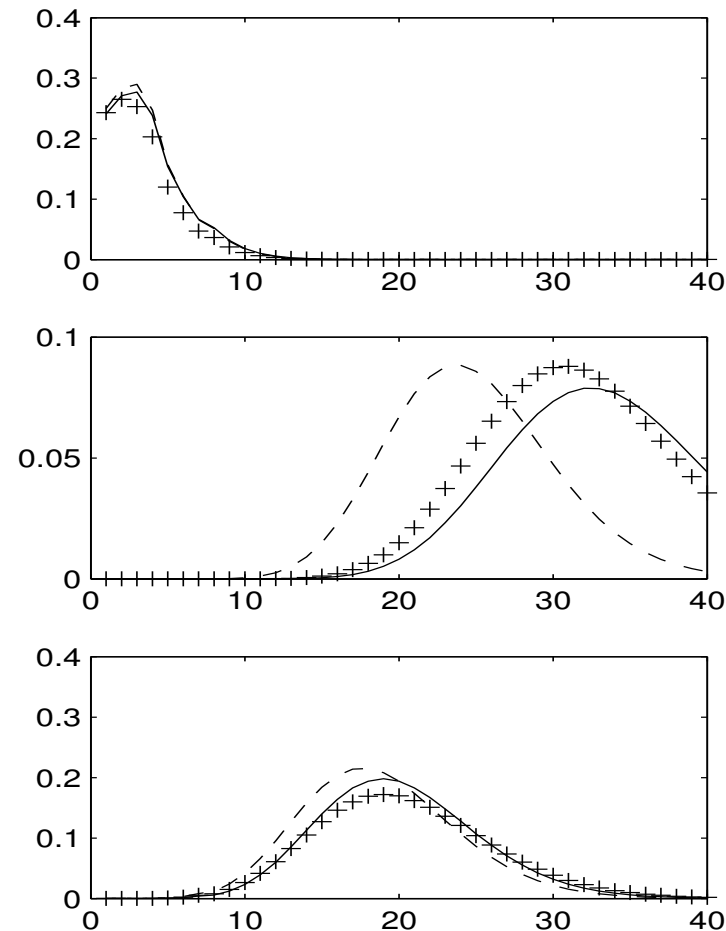


# Numerical results



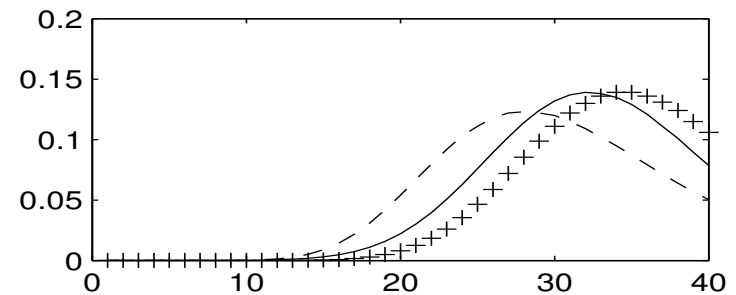
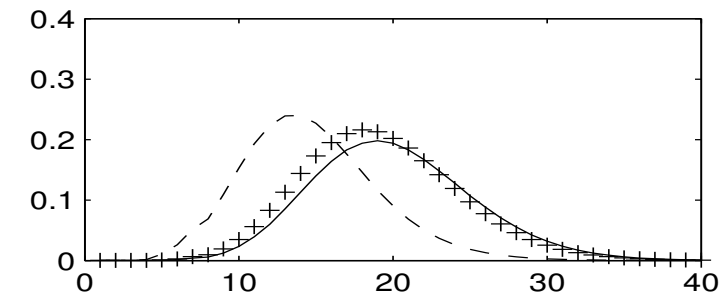
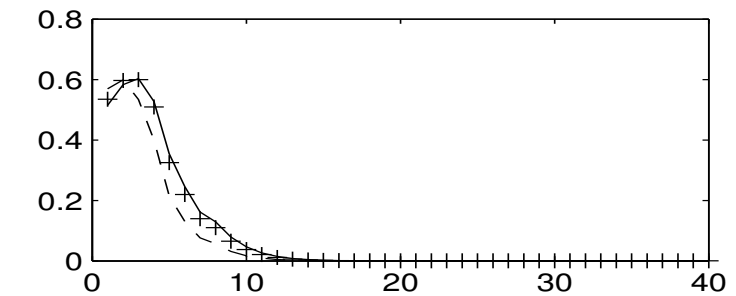
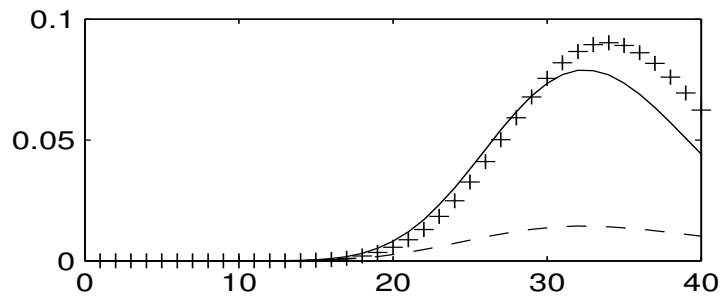
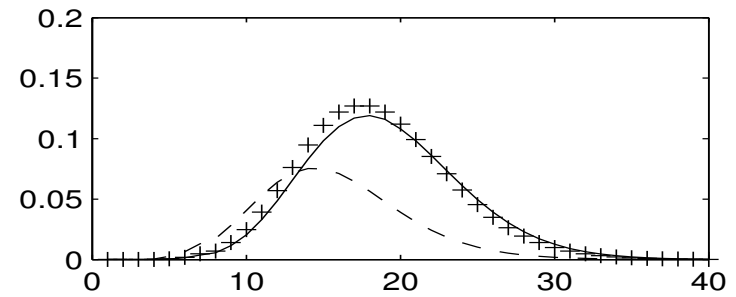
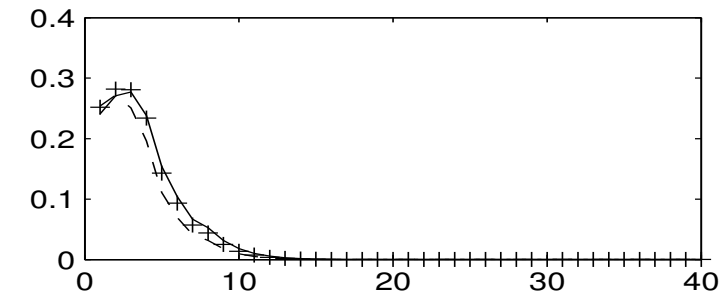
Concentration at different time instances for random walk sampler with  $\sigma = 0.01$  - the solid line designates the observed concentration, the dashed line designates the first match, and the data marked with '+' designates the concentration after the measurement information is incorporated into the simulations.

# Numerical results



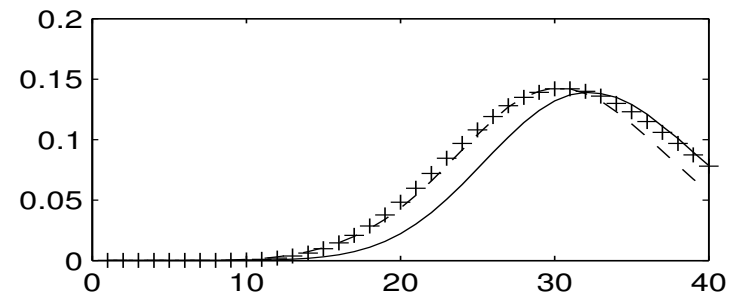
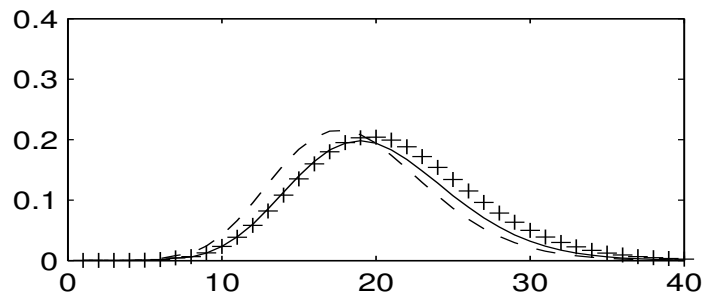
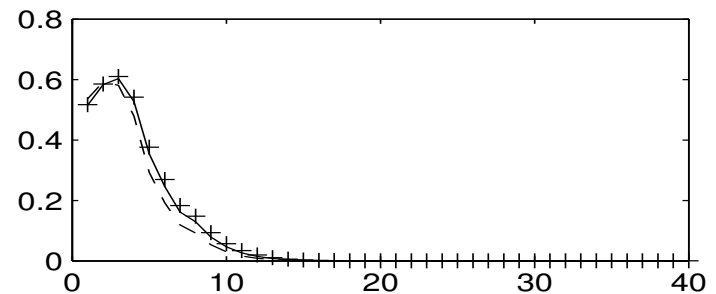
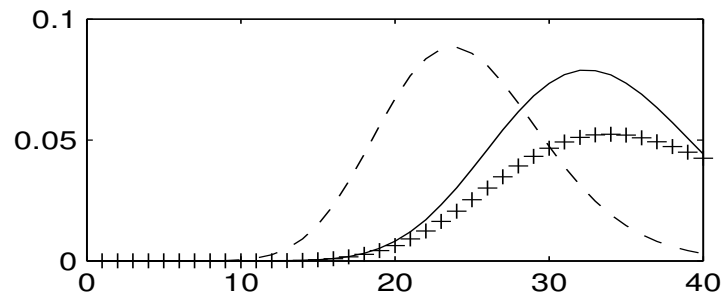
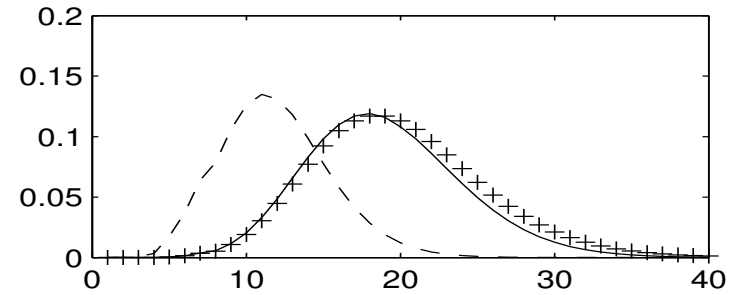
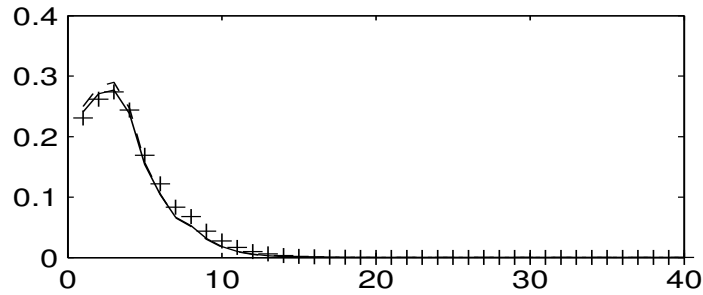
Concentration at different time instances for independent sampler with  $\sigma = 0.01$  - the solid line designates the observed concentration, the dashed line designates the first match, and the data marked with '+' designates the concentration after the measurement information is incorporated into the simulations.

# Numerical results



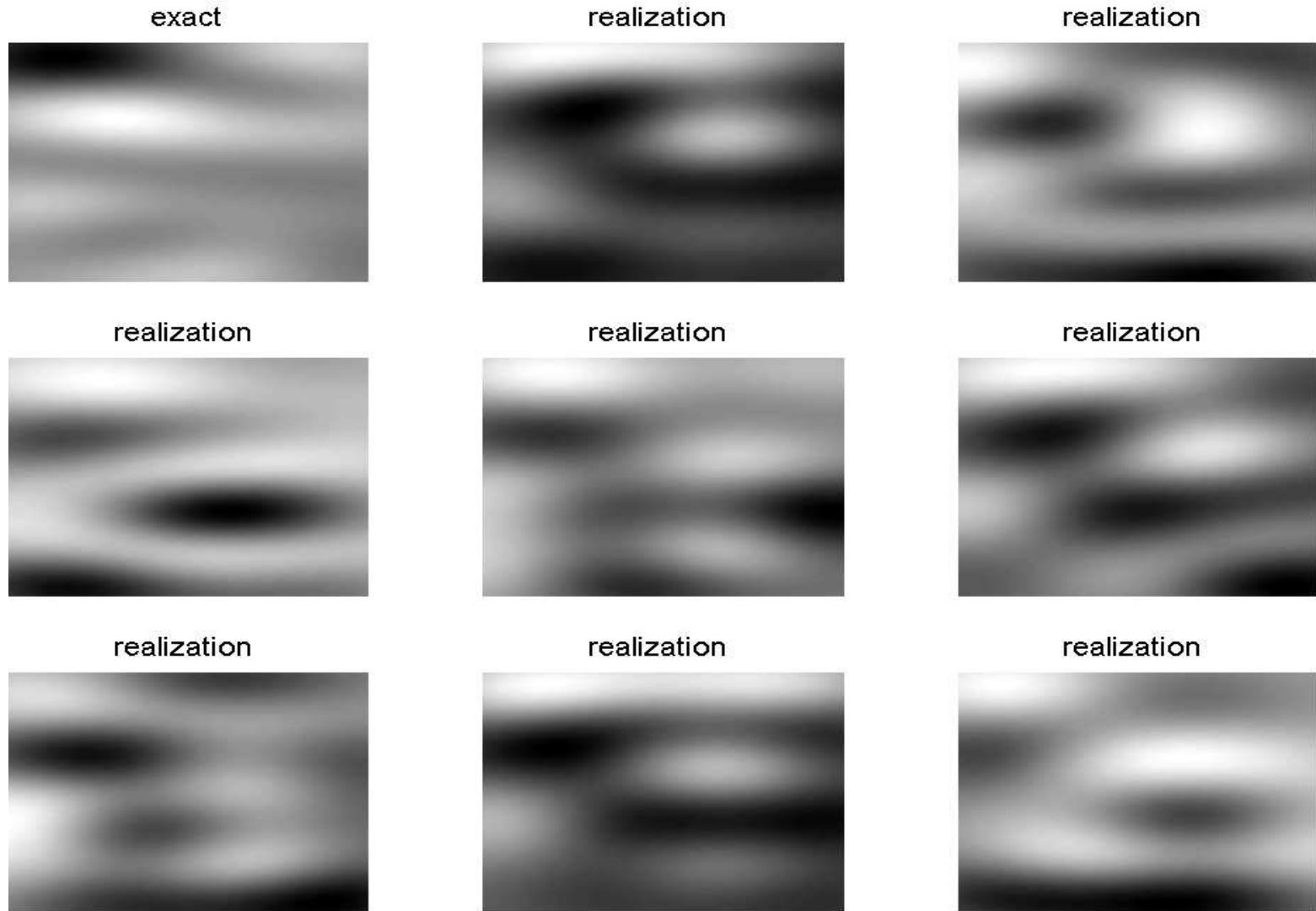
Concentration at different time instances for random walk sampler with  $\sigma = 0.005$  - the solid line designates the observed concentration, the dashed line designates the first match, and the data marked with '+' designates the concentration after the measurement information is incorporated into the simulations.

# Numerical results



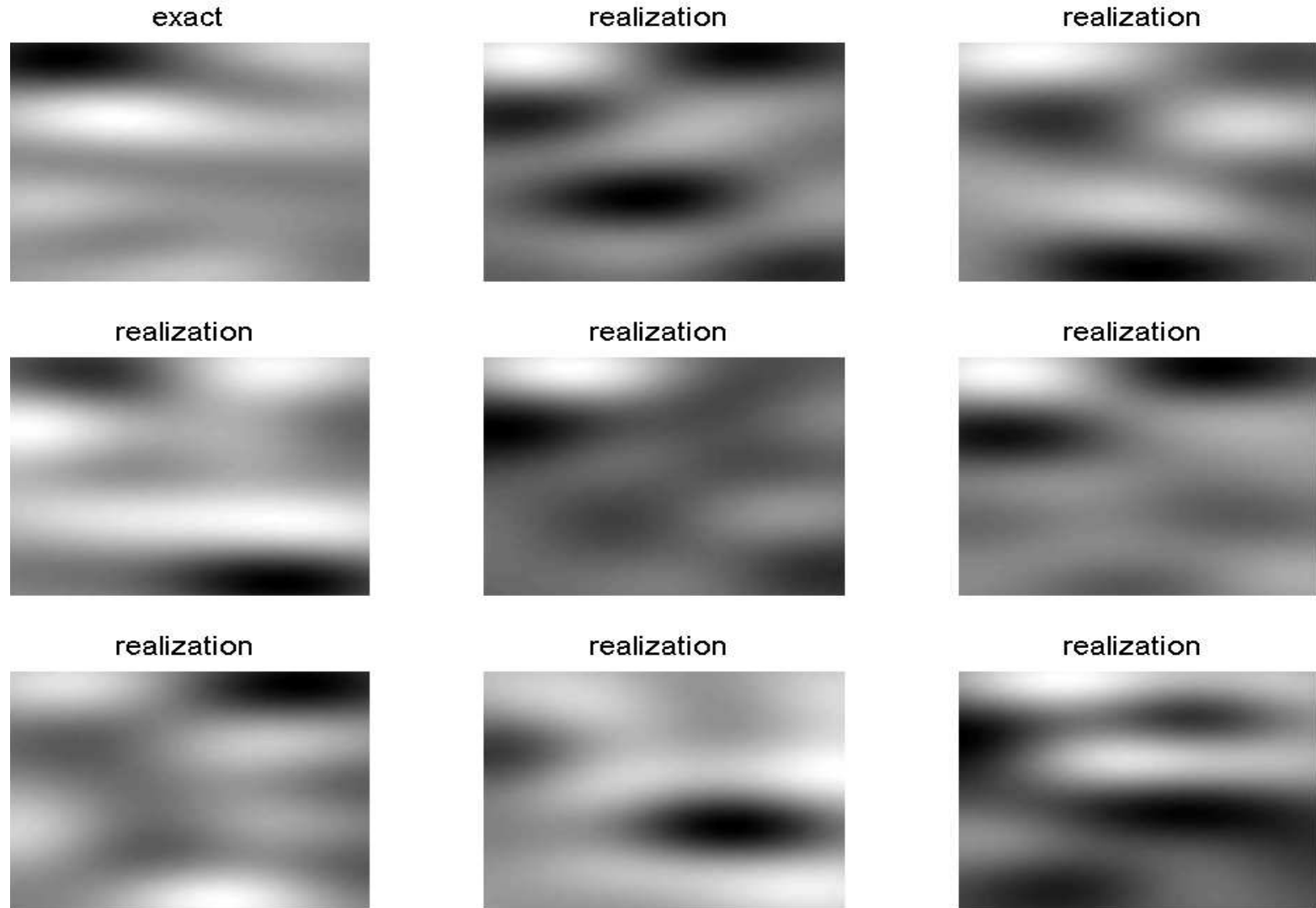
Concentration at different time instances for independent sampler with  $\sigma = 0.005$  - the solid line designates the observed concentration, the dashed line designates the first match, and the data marked with '+' designates the concentration after the measurement information is incorporated into the simulations.

# Numerical results



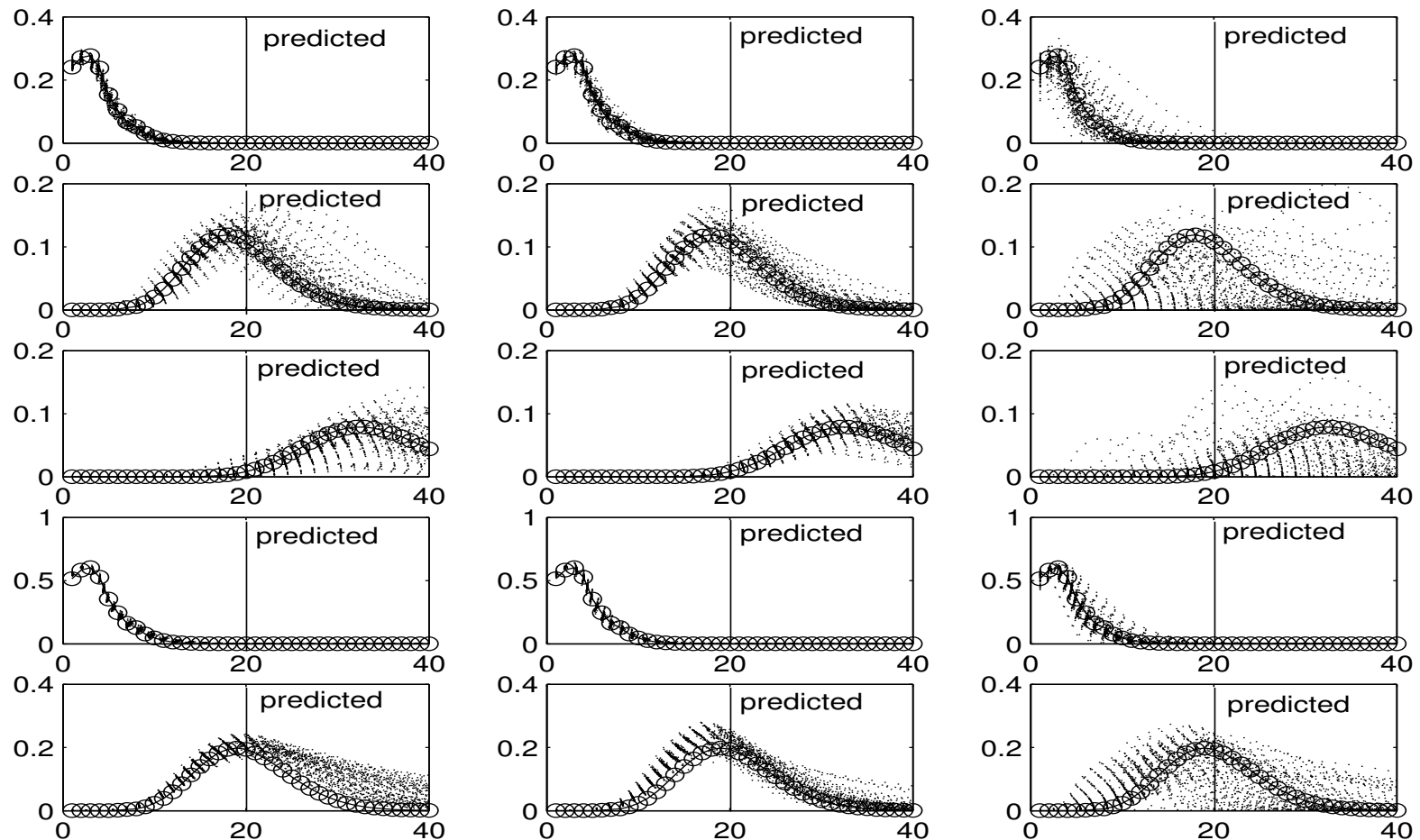
Permeability realizations after each update. Random walk sampler with  $\sigma = 0.01$  is used.

# Numerical results



Permeability realizations after each update. Independent sampler with  $\sigma = 0.01$  is used.

# Numerical results



Each row corresponds to the time history  $[0, 20]$  of the concentration at the measurement point and forward prediction. The (o) denotes the measured concentration, whereas the dotted line corresponds to the predicted concentration using the sampled permeability. The first column uses MCMC with  $\sigma = 10^{-2}$ , the second column uses MCMC with  $\sigma = 0.005$ , and the third column uses realizations from the prior distribution.



# Conclusions

- The solution prediction. “Initial” data update.
- Uncertainty propagation.
- Permeability estimation. Efficient sampling techniques.

# Future work

- Comparison with EnKF. The use of preconditioners for nonlinear EnKF.
- The use of multiscale basis functions for efficient prediction as well as preconditioning.
- Moving sensors.
- Robust priors.
- Applications to near coastal water contaminant transport.