Uncertainty Quantification and High Performance Computing for flow and transport in porous media

J. Erhel
joint work with

Outline
Physical model and examples of results
Numerical methods and HPC

Spatial Statistics and Uncertainty Quantification on Supercomputers

Bath, May 2014
Uncertainty Quantification and High Performance Computing for flow and transport in porous media

J. Erhel
joint work with

Outline

1. Physical model and examples of results
2. Numerical methods and HPC
Heterogeneous porous media

- Heterogeneity at all scales
- Lack of data

Models and methods must deal with uncertainty: **random data**

* A. Beaudoin et al., WRR 2007, WRR 2008, WRR 2010, WRR 2013
Stochastic model

Random conductivity field

\[
\begin{align*}
a(\omega, x) &= \exp(Y(\omega, x)) \quad \text{in } \Omega \times O, \\
\text{cov}[Y](x, y) &= \sigma^2 \exp\left(-\left(\frac{\|x-y\|}{l}\right)^\delta\right) \quad \text{in } O,
\end{align*}
\]

with \( O \) open domain in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \),
\( \Omega \) probability space,
\( Y \) a gaussian field with mean \( \mu \) and covariance \( \text{cov}[Y] \)

\( \sigma = 1, l = 5, 5 \) orders or magnitude

\( \sigma = 3, l = 5, 14 \) orders of magnitude
Single-phase steady-state flow in a saturated porous medium

Darcy’s law and mass conservation law

\[
\begin{align*}
\theta(x)v(\omega, x) &= -a(\omega, x)\nabla p(\omega, x) \quad \text{in } \Omega \times O, \\
\nabla \cdot (\theta(x)v(\omega, x)) &= 0 \quad \text{in } \Omega \times O,
\end{align*}
\]

Boundary conditions,

with \( \theta v \) the Darcy velocity, \( p \) the hydraulic head, \( \theta \) the porosity, \( a \) the hydraulic conductivity.

Domain \( L_x = 64, L_y = 32, L_z = 32 \), Gaussian covariance \( l = 5, \sigma = 2 \)
Transport by advection-diffusion-dispersion

\[
\begin{cases}
\frac{\partial (\theta(x)c(\omega, x, t))}{\partial t} + \nabla.(\theta(x)v(\omega, x)c(\omega, x, t)) - \nabla.\left(\theta(x)D_v(\omega, x)\nabla c(\omega, x, t)\right) = 0, \\
\text{boundary conditions,} \\
\text{initial condition,}
\end{cases}
\]

with \( c \) the solute concentration, \( D_v \) the dispersion tensor given by

\[
D_v(\omega, x) = (\alpha_T(x)\|v(\omega, x)\| + D_m(x))I + (\alpha_L(x) - \alpha_T(x)) \frac{v(\omega, x)v(\omega, x)^T}{\|v(\omega, x)\|}.
\]

with \( D_m(x) \) the isotropic molecular diffusion coefficient, \( \alpha_T(x) \) and \( \alpha_L(x) \) the transversal and longitudinal dispersion coefficients.

Evolution of the plume in a log-normal exponentially correlated field in the pure advection case (\( \sigma = 3, \ l = 10, \ \mu = -13.8155 \))
Simulation of flow and transport with random input data

- PARADIS software with Graphical User Interface
- computation of statistical quantities of interest
- large scale 2D and 3D computations, using High Performance Computing
- distributed computing with MPI communication library

Evolution of the plume by advection-diffusion\( (D_m = 0.005, \sigma = 1, l = 5) \)
Quantities of interest

**Spreading and macro-dispersion**

\[
\begin{align*}
G_k(\omega, t) &= \int_{\Omega} c(\omega, x, t)x_k \, dx, \\
S_k(\omega, t) &= \int_{\Omega} c(\omega, x, t)(x_k - G_k(\omega, t))^2 \, dx, \\
D_k(\omega, t) &= \frac{1}{2} \frac{dS_k(\omega, t)}{dt}.
\end{align*}
\]

with $G_k$ the center of mass in direction $k$, $k = 1, 2, 3$, $S_k$ the spreading, $D_k$ the macro-dispersion.

**Mean spreading and mean macro-dispersion**

\[
\begin{align*}
S_k(t) &= \mathbb{E}_\omega[S_k(\omega, t)], \\
D_k(t) &= \mathbb{E}_\omega[D_k(\omega, t)].
\end{align*}
\]

Asymptotic values for $t \rightarrow \infty$
Gaussian covariance and
\( \mu = 0, 0.5 \leq \sigma \leq 3, l = 10 \)

512 \( \leq L_x \leq 2048, L_y = L_z = 256 \)

Periodic transversal boundary conditions for flow and transport

Dirichlet longitudinal flow boundary conditions:
\( p_{\text{in}} = L_x, p_{\text{out}} = 0 \)

Inflow homogeneous Neumann and outflow homogeneous Dirichelt transport boundary condition

Pure advection

128 up to 512 processors

Mean macro dispersion for various heterogeneities

A. Beaudoin et al., WRR, 2013
Numerical methods and HPC
Monte-Carlo method

- Finite number of samples $S_k(t) \approx \frac{1}{N} \sum_{i=1}^{N} S_{i,k}(t)$,
- discretization in the physical space for flow and transport computations,
- non intrusive and parallel simulations,
- a priori large number $N$ of simulations, but fast convergence in an ergodic case (observed in numerical experiments).

**Erhel et al., ParCFD, 2010**

- Random generation of streams with RNGSTREAM library,
- embarrassingly parallel loop,
- failure recovery with checkpoints.
Large domains imply large memory requirements,
two-level parallelism: external Monte Carlo loop and internal simulations,
allocation strategy guided by computing resources.
**One parallel simulation**

- Domain meshed with a regular grid with $N_m$ cubic cells of size $\Delta x$,
- parallel generation of the random conductivity field,
- parallel computation of the flow matrix,
- parallel sparse solving,
- parallel random walker for the transport model.

*Erhel et al., ParCFD, 2008*

- Ghost cells for communication,
- overlapping of SEND and RECV non blocking operations,
- number of subdomains fixed by memory requirements.
Generation of random hydraulic conductivity

\[ a = \exp(\mu + \sigma Z) \] with \( Z \) normal field with a given covariance function.

**Spectral simulation**

- compute the discrete spectral density, using a DFT of the covariance,
- compute the discrete normal field \( Z \), using an inverse FFT,
- parallel generation with domain decomposition,
- use of FFTPACK library,
- complexity in \( O(Nm \log(Nm)) \).

*G. Lecourt, master thesis, 2012; M. Oumouni et al., Inria preprint, 2014*

**Speed-up of random generation**

- extended domain for FFT,
- redistribution of data.
Mixed hybrid finite element method

- local and global mass conservation,
- can deal with anisotropic conductivity,
- can deal with unstructured mesh,
- equivalent to a finite volume method in a regular case,
- sparse symmetric positive definite matrix,
- mesh step $\Delta x$ and approximate velocity $\tilde{v}$,
- complexity in $O(N_m)$.
Matrix condition number

- 2D numerical experiments

Condition number versus matrix size: \( O(N) \)

Condition number versus \( \sigma \): \( O(\exp(\sigma)) \)
Flow method: system solving

Sparse linear solver

- parallel direct solver for large systems,
- parallel algebraic multigrid for very large systems,
- convergence of AMG independent of heterogeneity,
- almost linear complexity in $O(N_m)$ of AMG,
- use of HYPRE library: SMG and Boomer-AMG.

Erhel et al., ParCFD, 2008; A. Beaudoin et al., ESAIM Proc., 2013

CPU time of SMG and AMG for 2D domains

CPU time of AMG for small 3D domains
**Transport method**

**Random walker method**
- Lagrangian method,
- pure advection case: characteristic method,
- pure diffusion case: Monte-Carlo method,
- no numerical diffusion,
- efficient parallel algorithm.

_Erhel et al., Europar, 2007_

**Fokker-Planck equation**
with $D_v = D_m I$ and $\theta(x) = 1$, infinite domain

\[
\left\{ \begin{array}{l}
\frac{\partial c(x,t)}{\partial t} + v(x) \nabla c(x,t) - D_m \Delta c(x,t) = 0, \\
c(x,0) = c_0(x)
\end{array} \right. \tag{5}
\]

**Stochastic Differential Equation**

\[
\left\{ \begin{array}{l}
dX_t = v(X_t) dt + \sqrt{2D_m} dW_t, \\
X_0 \text{ given}
\end{array} \right. \tag{6}
\]

If $X_0$ admits $c_0(x)$ as density, then $X(t)$ admits $c(x)$ as density.
Particle tracking

Simulation $i, i = 1, \ldots, N$
Discrete Flow equation with random data $a_i$ and approximate velocity $\tilde{v}_i$

Euler scheme with approximate velocity $\tilde{v}_i$
$M$ particles $j = 1, \ldots, M$ and time step $\Delta t$

$$
\begin{align*}
\tilde{X}_{i,j}(t + \Delta t) &= \tilde{X}_{i,j}(t) + \tilde{v}_i(\tilde{X}_{i,j}(t))\Delta t + \sqrt{2D_m\Delta t}\tilde{W}_{i,j}(t), \\
\tilde{X}_{i,j}(0) &\text{ given.}
\end{align*}
$$

(7)
Injection of particles by packets

Local and global communications
Parallel particle tracker

Outline
- Physical model and examples of results
- Numerical methods and HPC

2D parallel computations
Large 3D numerical experiments

**Physical parameters**
- 3D domain of size $1024 \times 256 \times 256$
- Log-normal conductivity field Gaussian covariance and $\mu = 0, \sigma = 1, l = 10$
- Periodic transversal boundary conditions for flow and transport
- Dirichlet longitudinal flow boundary conditions: $p_{in} = 1024$, $p_{out} = 0$
- Inflow homogeneous Neumann and outflow homogeneous Dirichlet transport boundary condition
- Molecular diffusion $D_m = 10^{-3}$

**Numerical parameters**
- Computations with PARADIS software using HPC
- Regular mesh with square cells of size $\Delta x = 1$
- Injection window of transversal size $0.8 \times 256$ and longitudinal size $\Delta x$
- Local time step
- Use of 128 cores of a cluster (in Poitiers)

*A. Beaudoin et al., ESAIM Proceedings, 2013*
CPU time for one simulation

- Random field generation CPU time: < 1 second,
- flow CPU time: about 130 seconds,
- number of V-cycles: in the range [25, 30].

CPU time of random walker versus the number of particles
Approximation of macro-dispersion for one simulation

**Spreading and macro dispersion**

\[
\begin{align*}
G_{i,k}(t) &= \mathbb{E}_\xi \left[ X_{t,i,k} \right], \\
S_{i,k}(t) &= \mathbb{E}_\xi \left[ (X_{t,i,k} - G_{i,k}(t))^2 \right], \\
V_{i,k}(t) &= \mathbb{E}_\xi \left[ v_k(X_{t,i}) \right], \\
D_{i,k}(t) &= \text{trace}(D_m) + \mathbb{E}_\xi \left[ X_{t,i,k} v_k(X_{t,i}) \right] - G_{i,k}(t)V_{i,k}(t).
\end{align*}
\]

**Particle Monte-Carlo applied to simulation \( i \)**

\[
\begin{align*}
\tilde{G}_{i,k}(t) &= \frac{1}{M} \sum_{j=1}^{M} \tilde{X}_{i,j,k}(t), \\
\tilde{S}_{i,k}(t) &= \frac{1}{M} \sum_{j=1}^{M} (\tilde{X}_{i,j,k}(t) - \tilde{G}_{i,k}(t))^2, \\
\tilde{V}_{i,k}(t) &= \frac{1}{M} \sum_{j=1}^{M} \tilde{v}_i,k(\tilde{X}_{i,j}(t)), \\
\tilde{D}_{i,k}(t) &= \text{trace}(D_m) + \frac{1}{M} \sum_{j=1}^{M} \tilde{X}_{i,j,k}(t)\tilde{v}_i,k(\tilde{X}_{i,j}(t)) - \tilde{G}_{i,k}(t)\tilde{V}_{i,k}(t).
\end{align*}
\]
Approximation of mean macro dispersion

$N$ Monte-Carlo simulations

\[
\begin{aligned}
\tilde{S}_k(t) &= \frac{1}{N} \sum_{i=1}^{N} \tilde{S}_{i,k}(t), \\
\tilde{D}_k(t) &= \frac{1}{N} \sum_{i=1}^{N} \tilde{D}_{i,k}(t)
\end{aligned}
\]

Theoretical results

with some assumptions

\[
\left\| D(t) - \tilde{D}(t) \right\|_{L^2_{\omega, \xi}} \leq C(\Delta t^{\frac{1+\alpha}{2}} + \Delta x |\log(\Delta x)| + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}})
\]

J. Charrier, Inria preprint 2011 and Marseille Univ. preprint, 2013

M. Oumouni, Ph-D, 2013
Convergence of spreading with the number of particles

\( N = 500 \) simulations

**Outline**
- Physical model and examples of results
- Numerical methods and HPC

---

**Uncertainty Quantification and High Performance Computing for flow and transport in porous media**

J. Erhel
joint work with

Longitudinal spreading

\[ N_p = 500 \]
\[ N_p = 4000 \]
\[ N_p = 10000 \]
\[ N_p = 50000 \]

Transversal spreading

\[ N_p = 500 \]
\[ N_p = 4000 \]
\[ N_p = 10000 \]
\[ N_p = 50000 \]
Monte-Carlo convergence of spreading

50000 particles

Longitudinal spreading

Transversal spreading
Concluding remarks

Current results
- Macro-dispersion in 3D domains with advection-diffusion
- Theoretical and experimental convergence analysis
- Fast Monte-Carlo convergence

Current and future work
- Random walker with discontinuous dispersion tensor (joint work with A. Lejay)
- Theoretical analysis of asymptotic spreading
- Domain decomposition for very large 3D domains (joint work with D. Tromeur-Dervout)
- Improved parallel algorithms (reduce communications and memory usage)