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

J. Erhel [1]
joint work with
A. Beaudoin [2], J.-R. de Dreuzy [3], G. Pichot [1],
J. Charrier [4], A. Debussche [5], M. Oumouni [1,6], Z. Mghazli [6]

[1]Inria, Rennes, France, [2]Institut Pprime, Poitiers, France, [3]Geosciences, Rennes, France, [4]University of Marseille, France, [5]ENS Rennes, France, [6]University of Kenitra, Morocco

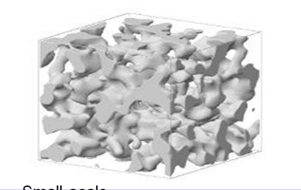
Spatial Statistics and Uncertainty Quantification on Supercomputers

Bath, May 2014

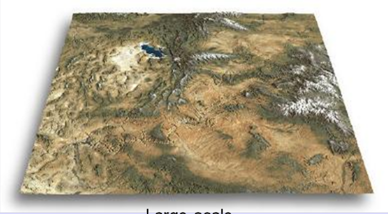
1. Physical model and examples of results

2. Numerical methods and HPC

Heterogeneous porous media



Small scale



Large scale

Meso scale

- Heterogeneity at all scales
- Lack of data

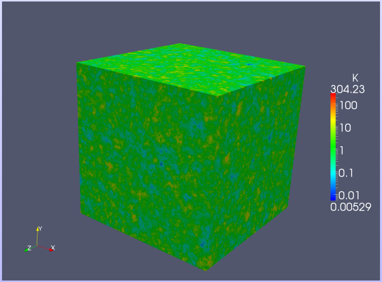
Models and methods must deal with uncertainty: **random data**

A. Beaudoin et al., WRR 2007, WRR 2008, WRR 2010, WRR 2013

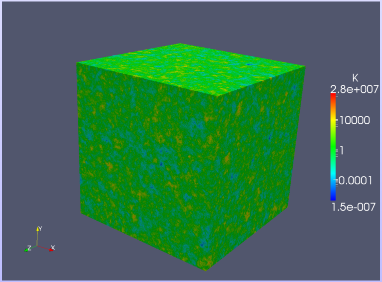
Random conductivity field

$$\begin{cases} a(\omega, x) = \exp(Y(\omega, x)) \text{ in } \Omega \times O, \\ cov[Y](x,y) = \sigma^2 \exp(-(\frac{\|x-y\|}{l})^{\delta}) \text{ in } O, \end{cases} \quad (1)$$

with O open domain in $\mathbb{R}^2$ or $\mathbb{R}^3$,
Ω probability space,
Y a gaussian field with mean μ and covariance $cov[Y]$



$\sigma = 1, l = 5$, 5 orders or magnitude



$\sigma = 3, l = 5$, 14 orders of magnitude

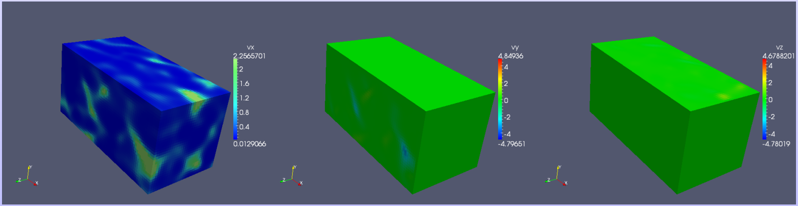
Flow equations

Single-phase steady-state flow in a saturated porous medium

Darcy's law and mass conservation law

$$\begin{cases} \theta(x)v(\omega,x) = -a(\omega,x)\nabla p(\omega,x) \text{ in } \Omega\times O, \\ \nabla.(\theta(x)v(\omega,x)) = 0 \text{ in } \Omega\times O, \\ \text{Boundary conditions}, \end{cases} \quad (2)$$

with θv the Darcy velocity, p the hydraulic head,
θ the porosity, a the hydraulic conductivity.



Domain $L_x = 64, L_y = 32, L_z = 32,$ Gaussian covariance $l = 5, \sigma = 2$

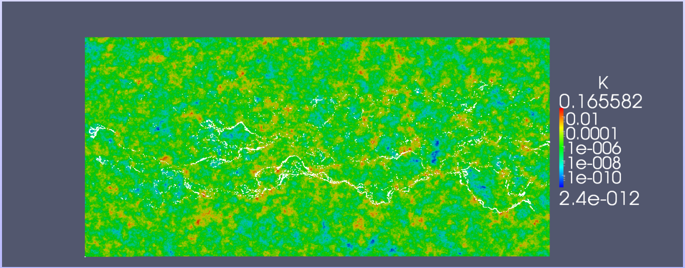
Transport by advection-diffusion-dispersion

$$\begin{cases} \frac{\partial(\theta(x)c(\omega,x,t))}{\partial t} + \nabla.(\theta(x)v(\omega,x)c(\omega,x,t)) - \nabla.(\theta(x)D_v(\omega,x)\nabla c(\omega,x,t)) = 0, \\ \text{boundary conditions,} \\ \text{initial condition,} \end{cases} \quad (3)$$

with c the solute concentration, D_v the dispersion tensor given by

$$D_v(\omega,x) = (\alpha_T(x)\,\|v(\omega,x)\| + D_m(x))I + (\alpha_L(x) - \alpha_T(x))\frac{v(\omega,x)v(\omega,x)^T}{\|v(\omega,x)\|}, \quad (4)$$

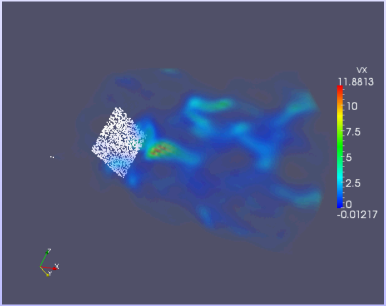
with $D_m(x)$ the isotropic molecular diffusion coefficient,
$\alpha_T(x)$ and $\alpha_L(x)$ the transversal and longitudinal dispersion coefficients.



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($D_m = 0.005, \sigma = 1, l = 5$)

Spreading and macro-dispersion

$$\begin{cases} G_k(\omega, t) = \int_O c(\omega, x, t) x_k dx, \\ S_k(\omega, t) = \int_O c(\omega, x, t)(x_k - G_k(\omega, t))^2 dx, \\ D_k(\omega, t) = \frac{1}{2} \frac{dS_k(\omega, t)}{dt}. \end{cases}$$

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

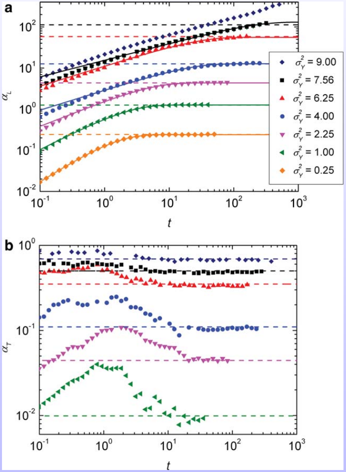
$$\begin{cases} S_k(t) = \mathbb{E}_\omega[S_k(\omega, t)], \\ D_k(t) = \mathbb{E}_\omega[D_k(\omega, t)]. \end{cases}$$

Asymptotic values for $t \to \infty$

3D numerical results

- 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: $p_{in} = L_x, \ p_{out} = 0$
- Inflow homogeneous Neumann and outflow homogeneous Dirichelt transport boundary condition
- Pure advection
- 128 up to 512 processors

A. Beaudoin et al., WRR, 2013



Mean macro dispersion for various heterogeneities

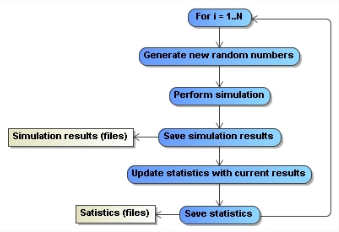
J. Erhel joint work with A. Beaudoin , J.-R. de Dreuzy , G. Pichot , J. Charrier , A. Debussche , M. Oumouni , Z. Mghazli

Numerical methods and HPC

Monte-Carlo method

- Finite number of samples $S_k(t) \simeq \frac{1}{N}\sum_{i=1}^{N} S_{i,k}(t)$,
- 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).

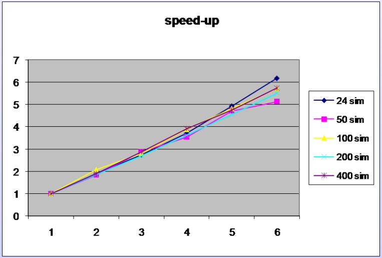
Erhel et al., ParCFD, 2010



Monte-Carlo loop

- Random generation of streams with RNGSTREAM library,
- embarrassingly parallel loop,
- failure recovery with checkpoints.

2D parallel Monte Carlo



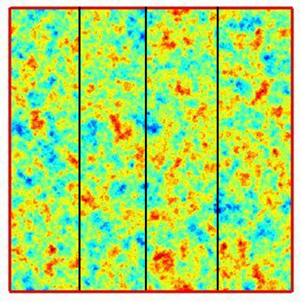
Speedup on a cluster (sequential simulations)

- 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 Δx,
- 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



Domain decomposition: one domain per processor

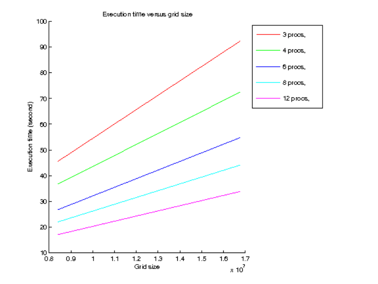
- Ghost cells for communication,
- overlapping of SEND and RECV non blocking operations,
- number of subdomains fixed by memory requirements.

$a = \exp(\mu + \sigma Z)$ 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 $O(N_m \log(N_m))$.

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



Speed-up of random generation

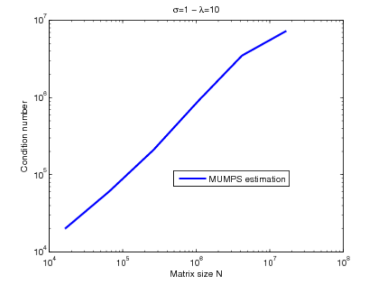
- extended domain for FFT,
- redistribution of data.

Mixed hybrid finite element method

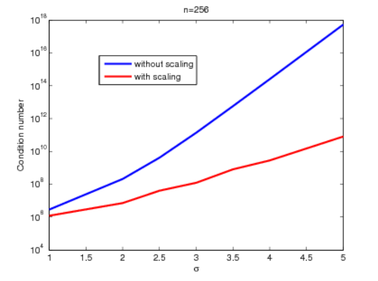
- local and global mass conservation,
- can deal with anisotropic conductivity,
- can deal with unstructured mesh,
- equivalent to a finite volume method in a regular case,
- sparse symmetric positive definite matrix,
- mesh step Δx and approximate velocity $\tilde{v}$,
- complexity in $O(N_m)$.

Matrix condition number

- 2D numerical experiments



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



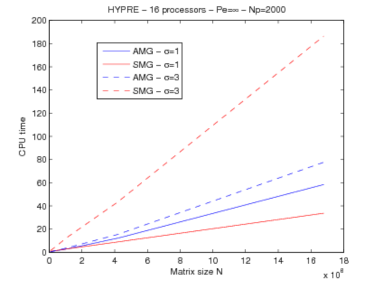
Condition number versus σ: $O(\exp(\sigma))$

Flow method: system solving

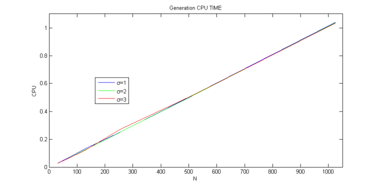
Sparse linear solver

- 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

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

$$\begin{cases} \frac{\partial c(x,t))}{\partial t} + v(x)\nabla c(x,t) - D_m \Delta c(x,t)) = 0, \\ c(x,0) = c_0(x) \end{cases} \tag{5}$$

Stochastic Differential Equation

$$\begin{cases} dX_t = v(X_t)dt + \sqrt{2D_m}dW_t, \\ X_0 \text{ given} \end{cases} \tag{6}$$

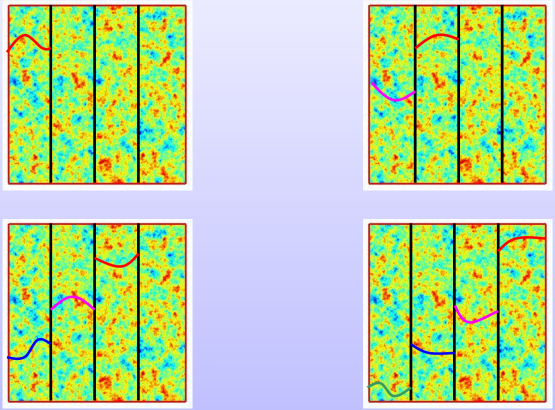
If X_0 admits $c_0(x)$ as density, then $X(t)$ admits $c(x)$ as density.

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

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

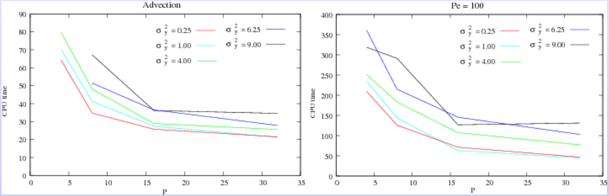
$$\begin{cases} \widetilde{X}_{i,j}(t+\Delta t) = \widetilde{X}_{i,j}(t) + \widetilde{v}_i(\widetilde{X}_{i,j}(t))\Delta t + \sqrt{2D_m \Delta t}\widetilde{W}_{i,j}(t), \\ \widetilde{X}_{i,j}(0) \text{ given.} \end{cases} \tag{7}$$

Injection of particles by packets



Local and global communications

Parallel particle tracker



Physical parameters

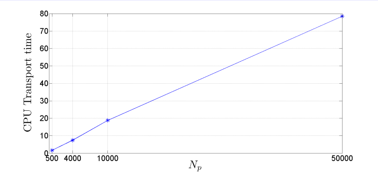
- 3D domain of size $1024 \times 256 \times 256$
- Log-normal conductivity field Gaussian covariance and $\mu = 0, \sigma = 1, l = 10$
- Periodic transversal boundary conditions for flow and transport
- Dirichlet longitudinal flow boundary conditions: $p_{in} = 1024$, $p_{out} = 0$
- Inflow homogeneous Neumann and outflow homogeneous Dirichelt 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 0.8×256 and longitudinal size Δx
- Local time step
- Use of 128 cores of a cluster (in Poitiers)

A. Beaudoin et al., ESAIM Proceedings, 2013

CPU time for one simulation



CPU time of random walker versus the number of particles

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

Approximation of macro-dispersion for one simulation

J. Erhel joint work with A. Beaudoin , J.-R. de Dreuzy , G. Pichot , J. Charrier , A. Debussche , M. Oumouni , Z. Mghazli

Spreading and macro dispersion

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

Particle Monte-Carlo applied to simulation i

$$\begin{cases} \widetilde{G}_{i,k}(t) = \frac{1}{M} \sum_{j=1}^{M} \widetilde{X}_{i,j,k}(t), \\ \\ \widetilde{S}_{i,k}(t) = \frac{1}{M} \sum_{j=1}^{M} (\widetilde{X}_{i,j,k}(t) - \widetilde{G}_{i,k}(t))^2, \\ \\ \widetilde{V}_{i,k}(t) = \frac{1}{M} \sum_{j=1}^{M} \widetilde{v}_{i,k}(\widetilde{X}_{i,j}(t)), \\ \\ \widetilde{D}_{i,k}(t) = \mathrm{trace}(D_m) + \frac{1}{M} \sum_{j=1}^{M} \widetilde{X}_{i,j,k}(t) \widetilde{v}_{i,k}(\widetilde{X}_{i,j}(t)) - \widetilde{G}_{i,k}(t) \widetilde{V}_{i,k}(t). \end{cases}$$

Approximation of mean macro dispersion

N Monte-Carlo simulations

$$\begin{cases} \widetilde{S}_k(t) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{S}_{i,k}(t), \\ \widetilde{D}_k(t) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{D}_{i,k}(t) \end{cases}$$

Theoretical results

with some assumptions

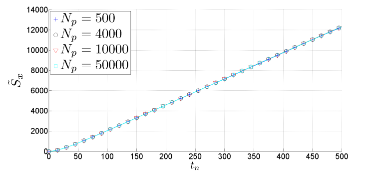
$$\left\| D(t) - \widetilde{D}(t) \right\|_{L^2_{\omega,\xi}} \leq C(\Delta t^{\frac{1+\alpha}{2}} + \Delta x \left|\log(\Delta x)\right| + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}})$$

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

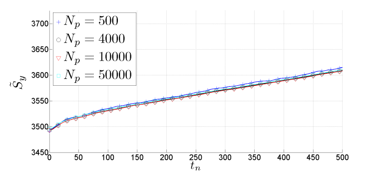
M. Oumouni, Ph-D, 2013

Convergence of spreading with the number of particles

$N = 500$ simulations



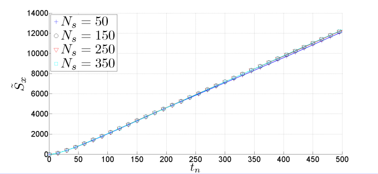
Longitudinal spreading



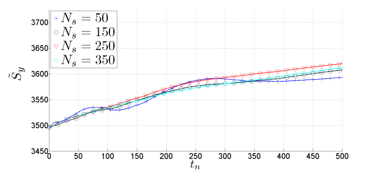
Transversal spreading

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)