DIRECT AND ITERATIVE SPARSE LINEAR SOLVERS APPLIED TO GROUNDWATER FLOW SIMULATIONS

Jocelyne Erhel, Anthony Beaudoin, Jean-Raynald de Dreuzy

DIRECT AND ITERATIVE SPARSE LINEAR SOLVERS APPLIED TO GROUNDWATER FLOW SIMULATIONS

J. ERHEL *, A. BEAUDOIN , AND J.-R. DE DREUZY †

Abstract. Subsurface hydraulic properties are mainly governed by the heterogeneity of geological formations. Moreover, it is not possible to measure the permeability in the whole site. In order to take into account the effects of heterogeneity on groundwater flows and the uncertainty in the input data, we use a probabilistic model in our numerical simulations. In the flow equations, governed by the Darcy law and mass conservation, the permeability is thus a stochastic given function. The Darcy velocity and the hydraulic head become then unknown stochastic functions. When dealing with fractured rocks, the domain itself and the geometry are random variables. With non intrusive methods such as Monte Carlo sampling or stochastic collocation, statistics of outputs are obtained through many simulations of deterministic equations. High performance computing is required to solve the flow equations. We consider two different cases. In porous media, the domain has a simple geometry. We use then a regular mesh and a finite volume method, leading to a structured sparse matrix. In fractured media, as the domain is a complex network of fractures, we use an unstructured mesh and a mixed hybrid finite volume method, leading to a general sparse matrix. In both cases, we compare direct methods, Krylov iterative methods and multigrid iterative methods, used to solve the linear systems obtained by discretizing the flow equations. We study the impact of the random distribution function on the condition number of the matrix and on the accuracy of the result. We also make a complexity analysis and a scalability analysis on parallel computers.

Key words. flow equation, heterogeneity, sparse linear solvers

AMS subject classification.

1. Introduction.
groundwater flow and solute transport
heterogeneous porous media and fractured media
stochastic numerical models
large sparse linear systems

2. Numerical model.
flow equations
probability distribution of permeability variance σ
discretization with 2D structured grids and finite volume method ; number of grid points $N = n^2$
Monte-Carlo method and non intrusive UQ methods
sparse linear system
scientific platform with interfaces to libraries and to Matlab
fully parallel for large scale computations

3. Accuracy.
theoretical estimation of condition number ; impact of variance and size : cond in $O(e^{\sigma})$ and in $O(n)$
numerical estimation with MUMPS or condest : cond in $O(e^{\sigma^2})$, larger than expected
scaling and componentwise estimations with Matlab : cond in $O(e^{\sigma})$ as expected
estimation for $\sigma = 1$ with MUMPS and scaling with Matlab : cond in $O(n)$ as expected
conclusion : good accuracy for $\sigma \leq 3$ and for large n, up to 8000

*Supported by ACI GRID'5000
†Supported by

4. Complexity. CPU and memory requirements; fill-in for direct methods and number of iterations for iterative methods

impact of variance and size

theoretical estimation of CPU time for direct methods : $O(n^3)$ and no sensitivity to variance

numerical measure with UMFPACK : as expected

direct methods not sensitive to variance but very expensive in CPU and memory for large size

numerical measure with PCG / IC(0) :

not too sensitive to variance but number of iterations increases with size; plateau before superlinear convergence

PCG suitable only with multilevel preconditioner

multilevel methods : subdomain decomposition or multigrid

theory : number of V-cycles independent of n, complexity of multigrid in $O(n^2)$

theory: geometric multigrid sensitive to σ and algebraic multigrid robust to σ

numerical measure with HYPRE : as expected

SMG better for small σ and AMG better for large σ

5. Large scale computations. Clusters with distributed memory ; SPMD and MPI ; fully parallel code with data distributed from scratch

experiments with P=16 and n up to 4096

measures on clusters from Grid'5000

same conclusions as in sequential case: SMG better for small σ and AMG better for large σ

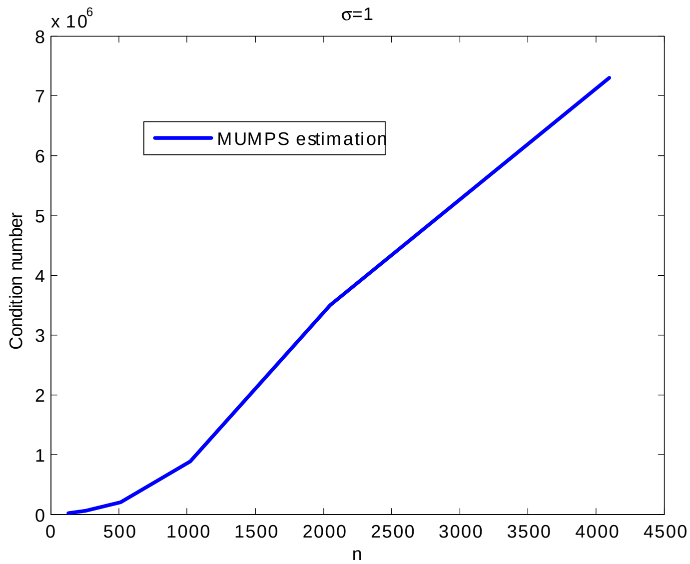
Able to solve systems with 16 millions in 1 minute

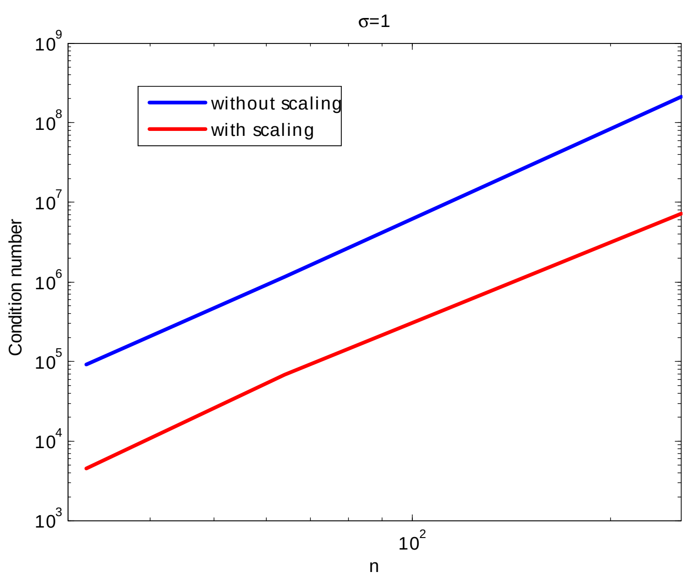
Monte-Carlo or UQ methods feasible ; multiparametric studies feasible ; done

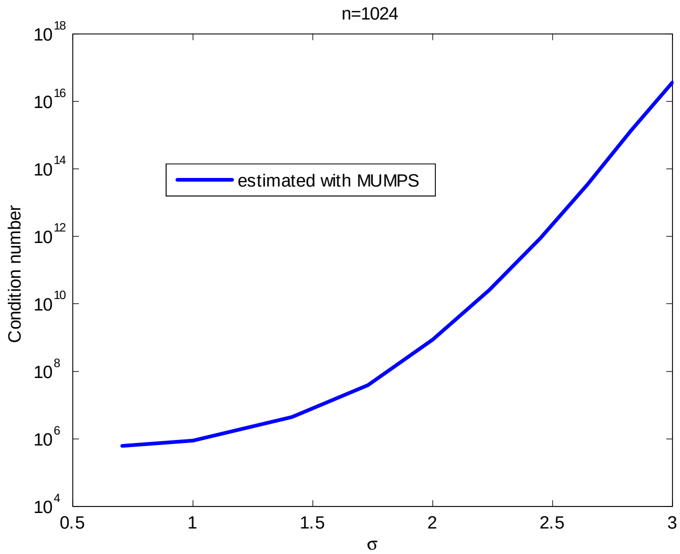
6. Future. 3D computations with subdomain decomposition

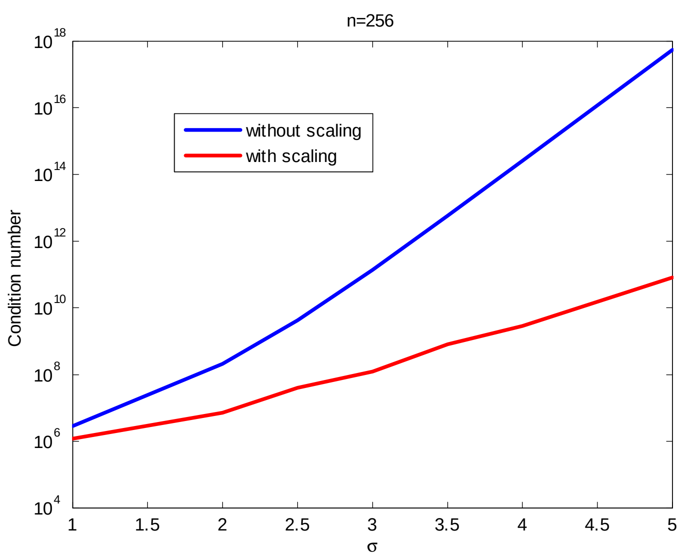
Grid computing

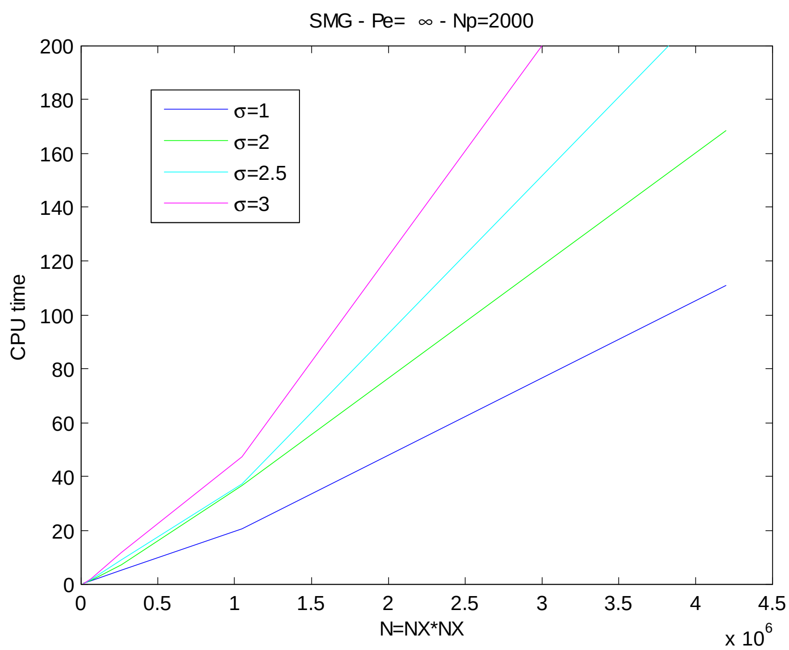
x


σ=1
x 10^6
MUMPS estimation
Condition number
n
0 1 2 3 4 5 6 7 8
0 500 1000 1500 2000 2500 3000 3500 4000 4500

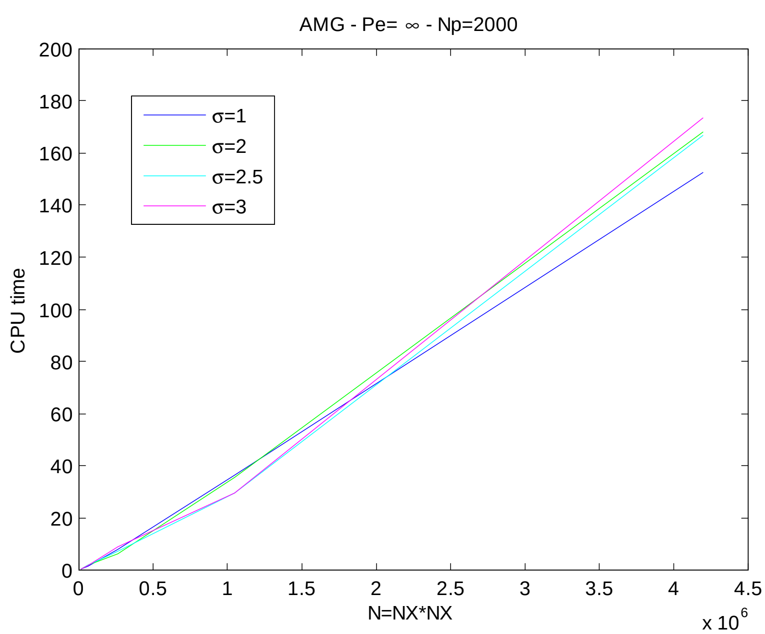

σ=1
without scaling
with scaling
Condition number
n
10^3 10^4 10^5 10^6 10^7 10^8 10^9
10^2


n=1024
Condition number
10^{18}
10^{16}
10^{14}
10^{12}
10^{10}
10^{8}
10^{6}
10^{4}
estimated with MUMPS
0.5
1
1.5
2
2.5
3
σ

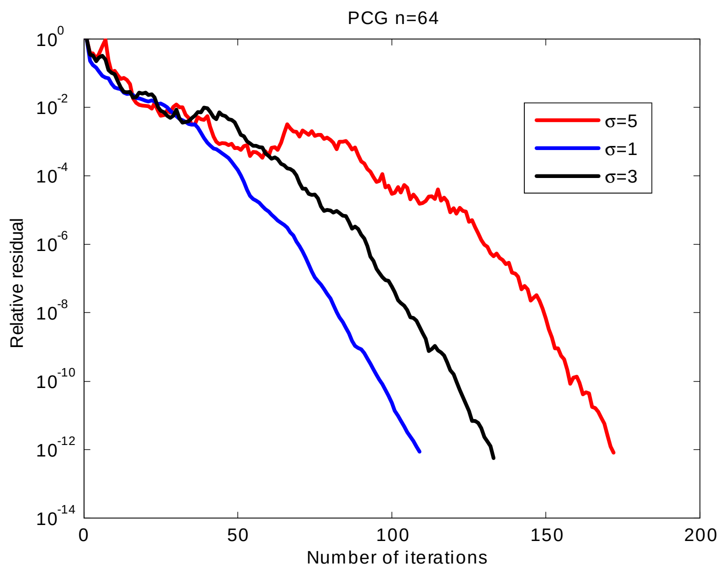

n=256
Condition number
10^{18}
10^{16}
10^{14}
10^{12}
10^{10}
10^{8}
10^{6}
10^{4}
without scaling
with scaling
1
1.5
2
2.5
3
3.5
4
4.5
5
σ

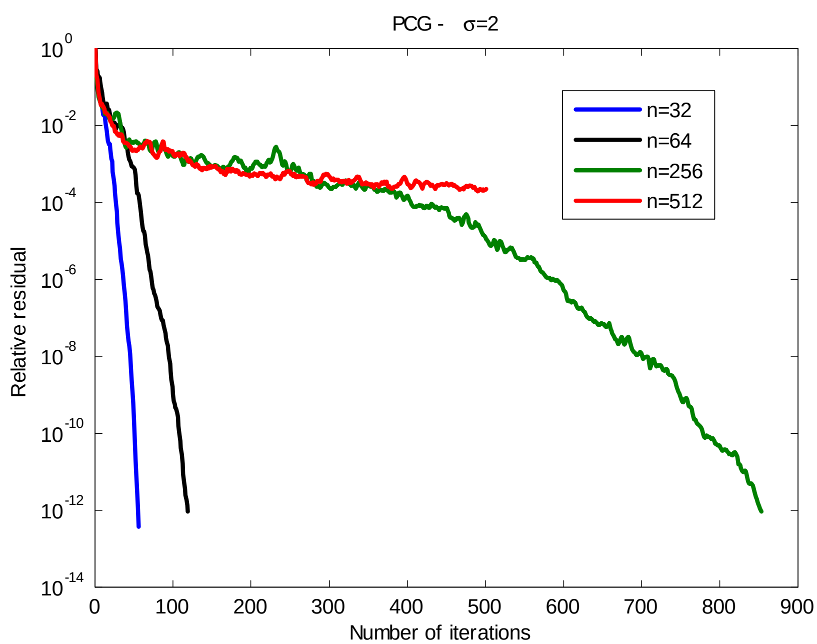


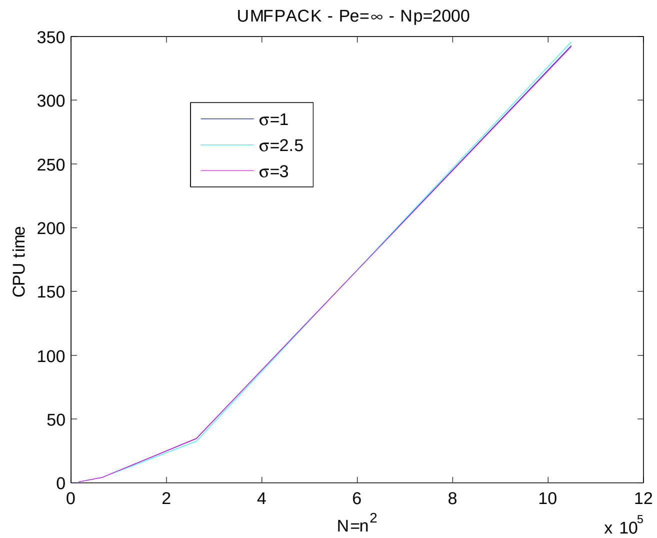
SMG - Pe= ∞ - Np=2000
CPU time
N=NX*NX
x 10^6
σ=1
σ=2
σ=2.5
σ=3

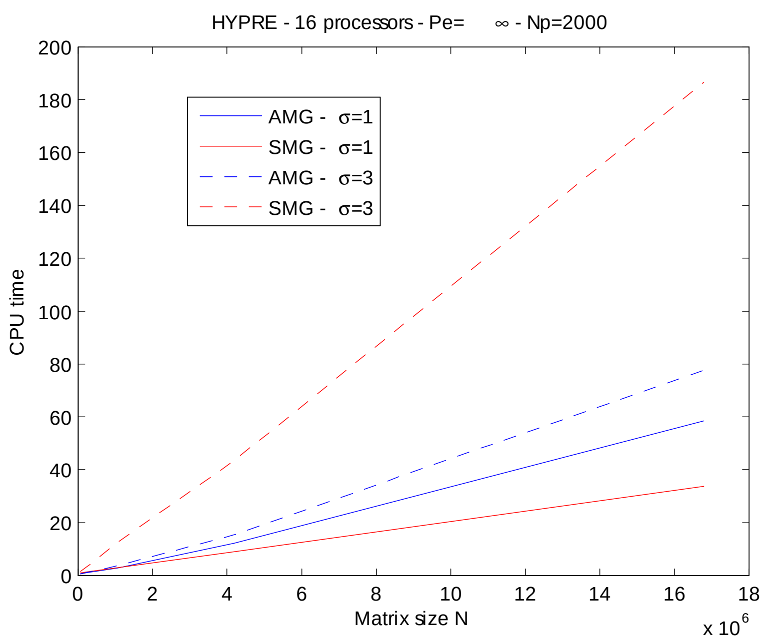


AMG - Pe= ∞ - Np=2000
CPU time
N=NX*NX
x 10^6
σ=1
σ=2
σ=2.5
σ=3


PCG n=64
σ=5
σ=1
σ=3
Relative residual
Number of iterations


PCG - σ=2
n=32
n=64
n=256
n=512
Relative residual
Number of iterations


UMFPACK - Pe=∞ - Np=2000
CPU time
N=n2
x 105
σ=1
σ=2.5
σ=3
0
50
100
150
200
250
300
350
0
2
4
6
8
10
12


HYPRE - 16 processors - Pe= ∞ - Np=2000
CPU time
Matrix size N
x 106
AMG - σ=1
SMG - σ=1
AMG - σ=3
SMG - σ=3
0
20
40
60
80
100
120
140
160
180
200
0
2
4
6
8
10
12
14
16
18