



## Asymptotic dispersion in 2D heterogeneous porous media determined by parallel numerical simulations

Jean-Raynald De Dreuzy, Anthony Beaudoin, Jocelyne Erhel

### ► To cite this version:

Jean-Raynald De Dreuzy, Anthony Beaudoin, Jocelyne Erhel. Asymptotic dispersion in 2D heterogeneous porous media determined by parallel numerical simulations. *Water Resources Research*, American Geophysical Union, 2007, 43, pp.W10439. <10.1029/2006WR005394>. <insu-00193397>

**HAL Id: insu-00193397**

**<https://hal-insu.archives-ouvertes.fr/insu-00193397>**

Submitted on 3 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Asymptotic dispersion in 2D heterogeneous porous media determined by parallel numerical simulations

Jean-Raynald de Dreuzy<sup>1</sup>, Anthony Beaudoin<sup>2,3</sup> and Jocelyne Erhel<sup>2</sup>

1- Géosciences Rennes (UMR CNRS 6118), Université de Rennes 1, 35042 Rennes cedex, France.

2- IRISA / INRIA of Rennes, Campus de Beaulieu, 35042 Rennes cedex, France.

3- Now at LMPG (Laboratoire de Mécanique, Physique et Géosciences), Université du Havre, 25 rue Philippe Lebon, BP 540, 76058 Le Havre cedex, France.

## Abstract

We determine the asymptotic dispersion coefficients in 2D exponentially-correlated lognormally-distributed permeability fields by using parallel computing. Fluid flow is computed by solving the flow equation discretized on a regular grid and transport triggered by advection and diffusion is simulated by a particle tracker. To obtain a well-defined asymptotic regime under ergodic conditions (initial plume size much larger than the correlation length of the permeability field), the characteristic dimension of the simulated computational domains was of the order of  $10^3$  correlation lengths with a resolution of ten cells by correlation length. We determine numerically the asymptotic effective longitudinal and transverse dispersion coefficients over 100 simulations for a broad range of heterogeneities  $\sigma^2 \in [0,9]$ , where  $\sigma^2$  is the lognormal permeability variance. For purely advective transport, the asymptotic longitudinal dispersion coefficient depends linearly on  $\sigma^2$  for  $\sigma^2 < 1$  and quadratically on  $\sigma^2$  for  $\sigma^2 > 1$  and the asymptotic transverse dispersion coefficient is zero. Addition of homogeneous

22 isotropic diffusion induces an increase of transverse dispersion and a decrease of longitudinal  
 23 dispersion.

## 24 **I. Introduction**

25 The determination of the large-scale dispersion coefficients has been widely debated in the  
 26 last twenty years [*Dagan*, 1989; *Gelhar*, 1993]. The classical case is the lognormal  
 27 permeability field with an exponential correlation function such as:

$$28 \quad C(r) = \sigma^2 \exp\left(-\frac{|r|}{\lambda}\right) \quad (1)$$

29 where  $\sigma^2$  is the log-normal permeability variance,  $|r|$  is the distance between two points and  
 30  $\lambda$  is the correlation length. Solute transport processes are advection and homogeneous  
 31 isotropic diffusion. We look at the asymptotic dispersion coefficient for large heterogeneity  
 32 corresponding to  $\sigma^2 \in [1,9]$ . Numerical simulations did not previously lead to definitive  
 33 solutions because of the large times and equivalent domain dimensions required for the  
 34 convergence to the asymptotic regime.

35 Two types of numerical simulations have been performed according to the derivation method  
 36 of the velocity field. The velocity field is classically computed either directly from  
 37 discretizing and solving the flow equation or from the first-order approximation of the flow  
 38 equation. The computational domain is of dimensions  $L_x$  and  $L_y$  in the two spatial dimensions  
 39  $x$  and  $y$ .  $L_x$  and  $L_y$  are counted in terms of correlation length. The correlation length  $\lambda$  is  
 40 counted in terms of grid cells. If we note  $l_m$  the dimension of the grid cell, the ratios  $L_x/\lambda$ ,  $L_y/\lambda$   
 41 and  $\lambda/l_m$  should be as large as possible. Discretizing the flow equation yields a linear system  
 42 of order proportional to the number of grid cells whatever the finite difference or finite  
 43 element scheme [*Bellin, et al.*, 1992; *Cvetkovic, et al.*, 1996; *Hassan, et al.*, 2002; *Salandin*

44 *and Fiorotto, 1998; Trefry, et al., 2003*] (Details of numerical simulations are given in  
45 Table 1). It explains why the direct solving of the flow equation has been limited to some  $10^5$   
46 cells number. It corresponds to some tens of exploitable correlation lengths that turn out to be  
47 not enough for determining directly the asymptotic dispersion coefficient. Convergence to the  
48 asymptotic regime is slow requiring very large simulations [*Bellin, et al., 1992*]. This study  
49 also shows a pronounced realization effect also obtained in [*Trefry, et al., 2003*]. The  
50 realization effect consists first in large dispersion coefficient variations and secondly in  
51 deviations from the mean behavior. It has two implications. First, the second-order moment of  
52 the solute plume requires a large number of Monte-Carlo realizations and particles to achieve  
53 convergence. Secondly, it emphasizes the problem of the relevance of the mean behavior to  
54 natural cases which are inherently single realizations requiring conditioning on measurements  
55 and the use of an inverse problem methodology.

56 The other simulation method consists in deriving the velocity field from the first order  
57 approximation of the flow equation and performing subsequently a particle tracking [*Bellin, et*  
58 *al., 1992; Dentz, et al., 2002; Rubin, 1990; Schwarze, et al., 2001*] (Table 1). This  
59 methodology does not require a grid and shortcuts the linear system solving step. Very long  
60 particle paths can be simulated and the asymptotic coefficients can be determined. In practice  
61 the average particle path length reached by this method is around hundred times larger than  
62 that obtained by the previous direct simulation method with a resolution five times finer  
63 (Table 1). However this methodology is limited to the validity domain of the first-order  
64 approximation ( $\sigma^2 < 1$ ). For larger heterogeneities, deviations of the velocity field from the  
65 normal behavior are non negligible and increase with  $\sigma^2$  [*Salandin and Fiorotto, 1998*]. The  
66 longitudinal velocity distribution becomes asymmetrical and is between the normal and  
67 lognormal distributions. The transverse velocity distribution becomes flatter with larger tails  
68 than that of the normal distribution. The first-order approximation of the velocity field does

69 not capture these effects. Moreover, the use of first-order approximated velocity field may  
 70 lead to erroneous numerical results. In fact, for even not too large heterogeneity ( $\sigma^2=1$ ), the  
 71 first-order approximation produces closed streamlines in which particles can enter either by  
 72 advection or by diffusion increasing artificially dispersion [Dentz, et al., 2003].

73 Neither the direct solution nor the first-order approximation of the flow simulation have led to  
 74 direct numerical estimates of the asymptotic transverse and longitudinal dispersion  
 75 coefficients for broad heterogeneous permeability fields ( $\sigma^2>1$ ). The only numerical estimate  
 76 is provided by Salandin and Fiorotto [1998] and concerns the dependency of the asymptotic  
 77 longitudinal dispersion coefficient  $D_{LA}$  on  $\sigma^2$ . They assume that the Lagrangian integral scale  
 78  $\lambda_{ux}$  is independent of  $\sigma^2$  and estimate numerically the Lagrangian velocity variance  $u_{xx}(0)$ . As  
 79  $D_{LA} \approx \lambda_{ux} \cdot u_{xx}(0)$  in the asymptotic regime, they found  $D_{LA} \propto \sigma^\beta$  with  $\beta=2.06, 2.19, 2.29$  and  
 80  $2.35$  respectively for  $\sigma^2$  in the intervals  $[0.05,1]$ ,  $[1,2]$ ,  $[2,3]$  and  $[3,4]$ .

81 Analytical estimates of the dispersion coefficient come from first-order and second-order  
 82 approximations of the flow and transport equations. First-order approximations yield a linear  
 83 dependence of the asymptotic longitudinal dispersion coefficient  $D_{LA}$  on  $\sigma^2$  and a zero  
 84 asymptotic transverse dispersion coefficient  $D_{TA}$  for purely advected solutes [Gelhar, 1993]:

$$85 \quad D_{LA} = u \cdot \lambda \cdot \sigma^2 \quad \text{and} \quad D_{TA} = 0 \quad (2)$$

86 where  $u$  is the mean velocity. Adding diffusion slightly reduces the asymptotic dispersion  
 87 coefficient  $D_{LA}$  for isotropic diffusion and  $Pe$  larger than 10 [Fiori, 1996]. Second-order  
 88 approximation of the transport equation has been taken into account and confirms the zero  
 89 asymptotic transverse dispersion coefficient [Hsu, et al., 1996]. For the longitudinal  
 90 dispersion coefficient and values of  $\sigma^2$  larger than 1 ( $\sigma^2=1.6$  in [Bellin, et al., 1992]), first-  
 91 order approximations of the flow and transport equations remain very close to numerical

92 results [Dagan, et al., 2003]. Adding a second-order term does not improve the approximation  
 93 of the longitudinal dispersion coefficient but on the contrary deteriorates it. It has thus been  
 94 deduced that the independent linearizations of flow and transport induce opposite deviations  
 95 from linear theoretical results that partly cancel out each other. These conclusions concerning  
 96 both the zero asymptotic transverse dispersion coefficient and the performance of the first-  
 97 order approximation were confirmed on slightly different heterogeneous media consisting in  
 98 spherical inclusions in a homogeneous medium [Dagan, et al., 2003; Jankovic, et al., 2003].  
 99 Other theoretical frameworks have been used to estimate the 2D asymptotic transverse  
 100 dispersion coefficient  $D_{TA}$ . Using volume averaging,  $D_{TA}$  is null like with the first-order  
 101 approximation [Attinger, et al., 2004], whereas  $D_{TA}$  is not null by using the conjecture of  
 102 Corrsin [Dentz, et al., 2002].

103 In this article, we compute the effective asymptotic longitudinal and transverse dispersion  
 104 coefficients for large heterogeneities ( $\sigma^2 \in [0.25, 9]$ ) both for pure advection and  
 105 homogeneous isotropic diffusion cases. To reach the asymptotic regime, we use very large  
 106 computational domains (100 times larger than the largest previously studied) under ergodic  
 107 conditions (large plume sizes compared to the correlation length of the permeability field).  
 108 We compare our results to the previous numerical results and analytical predictions.

## 109 **II. Numerical methods**

### 110 *II.1. Assumptions and notations*

111 We study 2D heterogeneous permeability field following a lognormal exponentially  
 112 correlated distribution as stated in the introduction (equation (1)). We perform the study on a  
 113 large range of  $\sigma^2$  values ( $\sigma^2 \in [0.25, 9]$ ) first because broad-range heterogeneities have been  
 114 observed in the field (for example  $\sigma^2 \sim 5$  for the Columbus site [Rehfeldt, et al., 1992]) and

115 secondly to test theoretical predictions. Solutes are transported by advection and diffusion.  
116 Diffusion is homogeneous with a diffusion coefficient noted  $d$ , the Peclet number  $Pe$   
117 expressing the ratio between advection and diffusion is equal to  $Pe = (\lambda \cdot u) / d$ , where  $u$  is  
118 the mean velocity.

119 As seen in introduction, discretizing and solving the flow equation for obtaining the velocity  
120 field computation is necessary for large heterogeneities ( $\sigma^2 > 1$ ). From previous simulation  
121 results obtained by using the first-order approximated velocity field (Table 1), the necessary  
122 domain dimensions to asymptotic regime is around a thousand of correlation lengths with a  
123 resolution of around 10 cells by correlation length [Ababou, *et al.*, 1989] leading to a number  
124 of cells of the order of  $10^8$ . Such large domains require parallel computing.

## 125 ***II.2. Permeability field generation***

126 The software must be fully parallelized as the computational domain itself cannot be stored on  
127 a unique processor. The computational domain is distributed from the beginning to the end of  
128 the simulation, according to a domain decomposition in vertical slices (figure 1). Each  
129 processor owns a well-defined part of the array corresponding to a sub-domain and keeps in  
130 local memory one layer of cells surrounding its sub-domain. These cells called “ghost cells”  
131 are necessary for the determination of the inter-cell permeability on sub-domain boundary  
132 cells. The additional cost of memory use is negligible and the communication cost between  
133 neighbouring processors is reduced.

134 The generation of the correlated lognormal field is performed via a Fourier transform  
135 [Gutjahr, 1989]. We use the software FFTW [Frigo and Johnson, 2005]. This library has a  
136 variety of composable solvers representing different FFT algorithms and implementation  
137 strategies, whose combination into a particular plan for a given size can be determined at  
138 runtime according to the characteristics of the machine/compiler in use. The construction of

139 the permeability field ends up with filling up the ghost cells, requiring the management of  
140 some communication between the processors. Permeability, velocity components and head  
141 values are all stored on the same types of array. The permeability field obtained from the  
142 Fourier transform methodology gives the right correlation length. The obtained variance is  
143 generally slightly smaller than the targeted variance [Yao, 2004]. More precisely, the variance  
144 is lowered by half the value of the mean. To avoid this bias we first generate a Gaussian  
145 correlated random field with zero mean and unitary variance. As we use a zero mean, the  
146 output variance is equal to the input targeted one. To obtain the right field, we first multiply  
147 the generated field by the standard deviation and add the logarithm of the geometric mean.  
148 We secondly take the exponential of the result. We calculated the obtained variance and  
149 found a value close at 0,02% to the input one for  $8192^2$  grids.

### 150 ***II.3. Flow computation***

151 We discretize the classical flow equation  $\nabla(K\nabla h)=0$  with  $K$  and  $h$  the permeability and  
152 hydraulic head and apply permeameter-like boundary conditions consisting in fixed head on  
153 two opposite borders and no flow on the perpendicular borders (figure 1). The flow equation  
154 is discretized according to a finite-difference scheme with harmonic inter-cell permeabilities.  
155 For regular square grids, this scheme is equivalent to mixed hybrid finite elements [Chavent  
156 and Roberts, 1991]. This equivalence ensures to these finite differences the high precision of  
157 the mixed hybrid finite elements useful for large permeability contrasts [Mosé, et al., 1994].  
158 The discrete flow equations end up to a linear system  $Ax = b$ , where  $A$  is a symmetric  
159 positive definite sparse structured matrix. The order of  $A$  is equal to the number of cells. The  
160 choice of the linear solver is essential to achieve the CPU and memory requirements for such  
161 large computational domains.

162 Several methods and solvers exist for these linear systems. They can be divided into three  
163 classes: direct, iterative and semi-iterative [Meurant, 1999; Saad, 1996]. Direct methods are  
164 highly efficient but require a large memory space. Iterative methods of Krylov type require  
165 less memory but need a scalable preconditioner to remain competitive. Iterative methods of  
166 multigrid type are often efficient and scalable, well-suited to regular grids, used by  
167 themselves or as pre-conditioners, but are sensitive to condition numbers [Wesseling, 2004].  
168 The condition number is related to the heterogeneities considered and increases very rapidly  
169 with the variance. Semi-iterative methods such as subdomain methods are hybrid  
170 direct/iterative methods which can be good tradeoffs [Toseli and Widlund, 2005]. For iterative  
171 and semi-iterative methods, the convergence and the accuracy of the results depend on the  
172 condition number which can blow up at large scale for a high variance ( $\sigma^2 > 4$ ). Because the  
173 memory space is more critical than the CPU time, we chose an iterative multigrid method. We  
174 used a numerical library HYPRE and more precisely Boomer-AMG (Algebraic MultiGrid)  
175 whose advantages are to be free, heavily used, portable and parallel [Falgout, et al., 2005].  
176 With this method, the CPU time is indeed not sensitive to the permeability variance. For a  
177 grid of  $1.3 \cdot 10^8$  nodes with  $\sigma^2 = 6.25$ , the flow computation requires around half an hour on a  
178 cluster of a 32 bi-processor AMD Opteron 2.2 GHz with 2 Go RAM each interfaced by  
179 Gigabit Ethernet.

#### 180 ***II.4. Transport simulation***

181 Transport is simulated by a particle tracker algorithm [Delay, et al., 2005]. Particle tracking is  
182 well suited for pure advection and advection-dominated transport processes because it does  
183 not introduce spurious numerical diffusion. Advection is simulated by a first order explicit  
184 scheme. We tried higher-order schemes which led to very small differences. Under this  
185 assumption of homogeneous isotropic diffusion, this method correctly models diffusion and

186 does not require any correction of the velocity term necessary for taking into account  
 187 diffusion discontinuities [Delay, et al., 2005]. Between  $t$  and  $t+dt$ , a particle moves from  
 188 positions  $M(t)$  to  $M(t+dt)$  by advection and diffusion:

$$189 \quad M(t + dt) = M(t) + v[M(t)] \cdot dt + \sqrt{2 \cdot d \cdot dt} \cdot Z \cdot r$$

190 where  $v[M(t)]$  is the velocity at the position  $M$ ,  $d$  is the diffusion coefficient,  $Z$  is a random  
 191 number drawn from a Gaussian distribution of mean 0 and variance 1 and  $r$  is a unitary vector  
 192 with uniformly distributed orientation. The time step evolves along the particle path according  
 193 to the velocity magnitude of the crossed cells. More precisely, the time step is either  
 194 proportional to the local advection time equal to the cell size  $l_m$  divided by the maximum of  
 195 the velocities computed on the cell borders noted  $v_{x+}, v_{x-}, v_{y+}, v_{y-}$  in the  $x$  and  $y$  directions or  
 196 to the diffusion time necessary to cross the cell:

$$197 \quad dt = \frac{1}{N_\alpha} \cdot \min \left[ \frac{l_m}{\max(v_{x+}, v_{x-}, v_{y+}, v_{y-})}, \frac{l_m^2}{2d} \right].$$

198  $N_\alpha$  is a positive integer representing the order of the time step number performed by the  
 199 particle in the cell. In the simulations  $N_\alpha$  is set to 10, meaning that the particle makes of the  
 200 order of 10 steps to cross the cell. The velocity  $v[M(t)]$  is obtained from a bilinear  
 201 interpolation as it is the sole interpolation method that ensures mass conservation [Pollock,  
 202 1988]. It is important to find the exit position of the particle from the cell in order that  
 203 particles always move in the cell with the velocity characteristics of the current cell and not of  
 204 the previous one [Pokrajac and Lazic, 2002]. The exit point and time from the cell are found  
 205 by linear interpolations. Diffusion is simulated by adding a random displacement of length  
 206 proportional to the square root of time and of the diffusion coefficient [Tompson and Gelhar,  
 207 1990].

208 To avoid border effects, particles are introduced at a distance  $0.05 L_x$  from the left border  
 209 (input border) of the computational domain (figure 1) corresponding for  $L_x=8192 l_m$  and  
 210  $L_y=16384 l_m$  and  $\lambda=10 l_m$  (10 cells by correlation length) to respectively 40 and 80 correlation  
 211 lengths downstream from the fixed head boundary. Particles are stopped when arriving at the  
 212 same distance upstream from the right border (output border). For all simulations, the  
 213 injection window is a thin line perpendicular to the mean flow direction of length equal to  
 214  $3277 l_m$  (i.e.  $0.4 L_y$  or around 328 correlation lengths for  $\lambda=10 l_m$ ). Particles are injected with  
 215 a uniform distribution within the injection window. The extension of the injection window is  
 216 large enough to ensure a broad sampling of the velocity field but narrow enough to prevent  
 217 particles from sampling the zones close to the no-flow boundary conditions [*Salandin and*  
 218 *Fiorotto, 1998*]. The number of particles approaching the no-flow border of the domain by  
 219 less than 15% of the domain dimension (120 correlation lengths) is recorded and found to be  
 220 null. This “exclusion zone” close to the no-flow boundaries is shown on figure 1. The  
 221 particle-tracking algorithm has been adapted for parallel simulations with the domain stored  
 222 on the different processors [*Beaudoin, et al., 2007*]. The time necessary for the simulation  
 223 transport was at most equal to the time required for the computation of flow.

### 224 **III. Dispersion computation, convergence and validation**

225 Simulations give the first two moments of the particle plume distribution  $\langle x(t) \rangle_i$  and  
 226  $\langle x^2(t) \rangle_i$ , here expressed in the longitudinal direction  $x$ :

$$227 \quad \langle x^k(t) \rangle_i = \frac{1}{N_p} \sum_{j=1}^{N_p} x_j(t)^k \quad (3)$$

228 with  $i$  the simulation number,  $x_j(t)$  the abscissa of the particle  $j$ ,  $k$  the moment order (1 or 2),  
 229 and  $N_p$  the number of particles. We compute a normalized dispersion coefficient by using the  
 230 classical formula

$$231 \quad D_L^i(t) = \frac{1}{u\lambda} \frac{1}{2} \frac{d(\langle x^2(t) \rangle_i - \langle x(t) \rangle_i^2)}{dt} \quad (4)$$

232 and discretize it on the successive time steps. The normalization factor  $u\lambda$  is logical in terms  
 233 of dimension to obtain a non-dimensional result. It is further justified for  $\sigma^2 < 1$  by the first-  
 234 order longitudinal dispersion coefficient linear in  $u\lambda$  (equation 2). In the following, the term  
 235 dispersion coefficient will refer to this normalized dispersion coefficient. We normalized the  
 236 time  $t$  as well by the characteristic time  $\lambda/u$  needed for the flux to cross a correlation length  
 237 and denote it  $t_N = ut / \lambda$ .

### 238 ***III.1. Asymptotic dispersion coefficient***

239 We determine the asymptotic dispersion coefficient  $D_{LA}^i$  from the time derivative signal (4)  
 240 according to the two following methods. Both methods rely on the late time behavior of the  
 241 dispersion coefficient  $D_L^i(t_N)$ . The first method consists in averaging  $D_L^i(t_N)$  over the time  
 242 range  $[0.5 t_{fb}, t_{fb}]$  over which  $D_L^i(t_N)$  is observed to have reached its asymptotic limit, where  
 243  $t_{fb}$  is the first breakthrough time (time for which the first particle arrives at a distance of  
 244  $0.05L_x$  from the output border). The asymptotic dispersion coefficient is the average noted  
 245  $D_{LA}^i(av)$ . The second method is a simple fit of  $D_L^i(t_N)$  over the whole time range by the  
 246 exponential function

$$247 \quad D_L^i(t_N) \approx D_{LA}^i(fit) \cdot (1 - \exp(-t_N / t_{N0}^i)) \quad (5)$$

248 where  $D_{LA}^i(\text{fit})$  is the asymptotic dispersion coefficient.  $t_{N0}^i$  is a characteristic convergence  
 249 time to the asymptotic regime. For the transverse dispersion coefficient, we derived the  
 250 realization-based coefficient  $D_T^i(t_N)$  by using the same methodology applying equations (3)  
 251 and (4) where we replace  $x$  by  $y$ . Because of the absence of any systematic time evolution, we  
 252 determine the asymptotic dispersion coefficient by averaging over the second part of the time  
 253 range  $[0.5 t_{fb}, t_{fb}]$  like for the longitudinal dispersion coefficient. Whatever the method, the  
 254 key point is to simulate transport in a sufficiently large domain to observe the stabilization of  
 255 dispersion on a time range long enough. The relevance of the asymptotic dispersion  
 256 coefficient depends on the domain dimensions counted in terms of correlation length  $L_x/\lambda$  and  
 257  $L_y/\lambda$ .

258 The mean and standard deviations of the dispersion coefficients as a function of time  
 259  $D_L(t_N) = \langle D_L^i(t_N) \rangle_{i=1..N_S}$  and  $\sigma[D_L(t_N)] = \left[ \langle D_L^i(t_N)^2 \rangle_{i=1..N_S} - \langle D_L^i(t_N) \rangle_{i=1..N_S}^2 \right]^{1/2}$  and the  
 260 mean of the asymptotic dispersion coefficients  $D_{LA} = \langle D_{LA}^i \rangle_{i=1..N_S}$  are thereafter determined  
 261 over  $N_S$  different realizations. The parameters controlling the determination of the asymptotic  
 262 dispersion coefficients are the domain dimensions  $L_x/\lambda$  and  $L_y/\lambda$ , the number of particles  $N_p$   
 263 and the number of simulations  $N_S$ . First simulations have shown that domains should be of  
 264 dimensions  $(L_x/\lambda, L_y/\lambda)$  equal to (820,820) and (1640,820) for respectively  $\sigma^2 \leq 4$  and  
 265  $\sigma^2 \geq 6.25$  to have a long enough signal. We use these values to study the convergence with  
 266  $N_p$  and  $N_S$  and verify after that these dimensions are indeed large enough. We study  
 267 successively the convergence as functions of the number of particles  $N_p$  and of the number of  
 268 simulations  $N_S$ . Two averaging methods are possible leading respectively to the effective and  
 269 ensemble dispersion coefficients. The effective dispersion is obtained by first computing the  
 270 derivative of the standard deviation of the plume concentration within a simulation and

271 secondly by averaging the computed standard deviations over the  $N_S$  simulations. The  
 272 ensemble dispersion is obtained by first computing the two first moments of the plume  
 273 concentrations over the  $N_S$  simulations and by secondly computing the derivative of the  
 274 standard deviation from the previous moments. The ensemble dispersion is larger than the  
 275 effective dispersion as it measures the plume dispersion with respect to the plume position  
 276 averaged over all simulations whereas the effective dispersion measures the plume dispersion  
 277 in each simulation with respect to the simulation mean plume position [Dentz, et al., 2000].

### 278 **III.2. Convergence with the number of particles $N_p$**

279 Figure 2 displays the dispersion coefficients  $D_L^i(t_N)$  and  $D_T^i(t_N)$  for number of particles  $N_p$   
 280 ranging from 100 to 10000. We choose an example in the most heterogeneous case ( $\sigma^2=9$ )  
 281 without diffusion (pure advection). For  $N_p=100$  (crosses), the dispersion coefficients are much  
 282 more variable than for  $N_p=1000$  (stars). Increasing the number of particles over 1000 does not  
 283 change the global tendencies of the dispersion coefficients. Finally between 5000 and 10000,  
 284 differences are very small. At a given time, the dispersion coefficient can be well approached  
 285 with  $N_p=10000$  particles. We computed also the asymptotic dispersion coefficients  $D_{LA}^i(av)$   
 286 and  $D_{TA}^i(av)$  according to the number of particles  $N_p$  in the most heterogeneous cases  
 287 ( $\sigma^2=6.25$  and 9) for Peclet numbers  $Pe$  ranging from 100 to  $\infty$ .  $Pe=\infty$  corresponds to the pure  
 288 advection case (without diffusion) (figure 3). For  $N_p \geq 2000$ ,  $D_{LA}^i(av)$  and  $D_{TA}^i(av)$  do not  
 289 vary much with the number of particles. More precisely, they vary respectively by less than  
 290 5% and 10% whatever the case. There is no systematic tendency either with the number of  
 291 particles or with the Peclet number. We kept for the pure-advection case 10000 particles and  
 292 for the advection-diffusion case 2000 particles. As convergence is not faster without diffusion  
 293 as shown by figure 3, this choice does not advantage the pure advection case more than the

294 5% and 10% precisions previously found. The global number of particles can be converted  
 295 into number of particles by correlation length at injection time. For all cases the injection  
 296 window was set to  $3277 l_m$ . For the pure-advection case  $N_p=10000$ , the number of particles by  
 297 cell is 3 on average and the number of particles by correlation length is 30 on average because  
 298 there are 10 cells by correlation length ( $\lambda=10 l_m$ ). For the advection-diffusion case  $N_p=2000$ ,  
 299 the number of particles by cell is 0.6 on average and with  $\lambda=10 l_m$ , the number of particles by  
 300 correlation length is 6 on average.

### 301 ***III.3. Convergence with the number of simulations $N_S$***

302 We study the convergence of the average and standard deviation of the dispersion coefficients  
 303 with the number of simulations  $N_S$  for the most heterogeneous cases  $\sigma^2=6.25$  and 9. The mean  
 304 longitudinal asymptotic dispersion coefficient is very close to the mean of the dispersion  
 305 coefficient taken at a given time  $t_N=600$  (figure 4a, solid and open squares compared to thick  
 306 solid and dashed grey lines). The asymptotic dispersion coefficient converges very rapidly for  
 307  $N_S \geq 20$ . The largest difference between values for 20 and 100 simulations is of the order of  
 308 2.5%. The standard deviation of the longitudinal dispersion coefficient at the same given time  
 309  $t_N=600$   $\sigma(D_L(t_N = 600))$  displays relative larger variations with the number of simulations  $N_S$   
 310 (figure 4b, lines and symbols). The maximal variation between  $N_S=20$  and  $N_S=100$  is 20%. As  
 311 variations are not monotonous, the value for the largest number of simulations cannot be more  
 312 precise than 20%. We note that both the variability of the average asymptotic dispersion  
 313 coefficients and the standard deviation of the dispersion coefficient decrease with more  
 314 diffusion (smaller Peclet numbers). Convergence of the longitudinal dispersion coefficient  
 315 with the number of simulations is thus faster with more diffusion.

316 Tendencies for the transverse dispersion coefficient are quite different. For  $N_s \geq 20$ , the  
317 asymptotic dispersion coefficient is very close to zero for the pure-advection case (figure 5a)  
318 whereas the standard deviation of the transverse dispersion coefficient at a given time is much  
319 larger around 0.6 (figure 5b). More diffusion corresponding to smaller Peclet numbers  
320 induces larger transverse asymptotic dispersion and standard deviation. The asymptotic  
321 transverse dispersion coefficient (figure 5a) converges quickly and its variations for  $N_s \geq 20$   
322 are less than 10% of its mean value. The standard deviation (figure 5b) still varies non  
323 monotonously and the amplitude of its variations can reach 25% of its mean value.

324 In the two previous sections, we have fixed the mesh size  $l_m$  and analyzed the convergence of  
325 the random walker and the Monte-Carlo simulations. For a given simulation, we verified  
326 numerically that the random walker converges when we increase the number of particles.  
327 More precisely, the dispersion coefficients  $D_L(t)$  and  $D_T(t)$  converge. We can assume a  
328 convergence in an appropriate norm; in view of the numerical results, we can also assume a  
329 uniform convergence, independent of the simulations. For a given number of particles, we  
330 verified numerically that the Monte-Carlo simulations converge when we increase the number  
331 of simulations. More precisely, we observe the convergence of the approximate first moments  
332 of the dispersion, computed with a given number of particles. Therefore, we can assume that,  
333 for a given mesh size  $l_m$ , our numerical Monte-Carlo simulations give an accurate estimation  
334 of the first moments of the two dispersion functions. However, in our simulations, the second  
335 moments do not converge correctly. There may be different reasons for this lack convergence.  
336 First the number of Monte-Carlo simulations  $N_S$  may not be large enough. Secondly  
337 dispersion coefficient may be affected by the finite volume method used for flow computation  
338 and the use of a bilinear interpolation for the velocity in the particle tracker. Thirdly, it may  
339 come from the generation of the permeability field from a truncated Fourier expansion and the  
340 assumption of a constant permeability in each grid cell. The same lack of convergence of the

341 dispersion fluctuations has already been observed and related to the finite number of Fourier  
 342 modes (figure 5 of [Eberhard, 2004]).

343 In the following, we perform 100 simulations to ensure convergence of the first moments  $D_{LA}$   
 344 and  $D_{TA}$  for each parameter set. We keep the same parameters for all simulations. As the  
 345 variations of both the longitudinal and transverse dispersion coefficients are stronger for  $\sigma^2=9$   
 346 than for  $\sigma^2=6.25$ , we checked that convergence is at least as good for lower heterogeneities  
 347 corresponding to  $\sigma^2 \leq 6.25$ .

### 348 ***III.4. Convergence of dispersion coefficients with time***

349 The domain dimensions were chosen in order to have a stabilization of the mean dispersion  
 350 coefficient as a function time  $D_L(t_N)$  over at least half the time duration of the simulation, the  
 351 maximal simulation time  $t_{fb}$  being the first breakthrough time (time for which the first  
 352 particle arrives at a distance of  $0.05L_x$  from the output border). The asymptotic regime is  
 353 maintained over around 500 time units or equivalently on a spatial range of 500 correlation  
 354 lengths. The simulations performed on domain of longitudinal dimension  $L_x=819,2 \lambda$  with  
 355  $\lambda=10 l_m$  were large enough for  $\sigma^2 \leq 4$ . For  $\sigma^2=1$ , the asymptotic regime is reached after some  
 356 tens of correlation lengths (figure 6a). However for  $\sigma^2 \geq 6.25$ , domains had to be twice  
 357 longer ( $L_x=1638,4 \lambda$ ) to obtain the same stabilization time range (figure 6b). Such long  
 358 stabilization times have also been observed in systems made of highly heterogeneous  
 359 inclusions [Jankovic, et al., 2006]. Large domain dimensions are required not only for large  
 360 values of  $\sigma^2$  but also for smaller values of  $\sigma^2$  (values around 1), although it is not obvious on  
 361 figure 6a. In fact we performed the same simulations for domains of dimensions  $L_x=102,4 \lambda$   
 362 by  $L_y=51,2 \lambda$  and found that the asymptotic regime is far from being reached although the  
 363 number of exploitable correlation lengths ( $\sim 80$ ) is large enough. There may be two reasons.

364 First the asymptotic regime is difficult to identify over some tens of correlation lengths.  
 365 Secondly, the injection window is smaller (20 correlation lengths at  $L_y=512 l_m$  compared to  
 366 327 at  $L_y=8192 l_m$ ) inducing from the beginning a lower sampling of the velocity field and a  
 367 larger convergence time to the asymptotic regime.

368 The mean  $D_L(t_N)$  and the confidence interval at 95% derived from the mean and standard  
 369 deviation are represented on figure 6 (solid and dashed lines).  $D_L^i(t_N)$  displays a large  
 370 variability but no definite trend whatever  $\sigma^2$  as shown by figure 6 (square and circle symbols).  
 371 The average  $D_L(t_N)$  of the dispersion coefficient over 100 simulations represented by the  
 372 black line smoothens the variations and indeed reaches a constant value at small times ( $t_N>30$ )  
 373 for  $\sigma^2=1$  (figure 6a) and at larger times ( $t_N>400$ ) for  $\sigma^2=9$  (figure 6b). The asymptotic regime  
 374 is well approached at least during the second half of the simulation time, i.e. in the interval  
 375  $[0.5 t_{fb}, t_{fb}]$ . The realization-based  $D_T^i(t)$  displays a strong variability around 0 but no trend,  
 376 not even at small times (square and circle points on figure 7). The average over simulations  
 377 (solid line of figure 7) does neither show any trend whereas the standard deviation (dashed  
 378 lines of figure 7) is large compared to the average values.

379 We note that several studies have used the apparent dispersion coefficient  
 380  $D_{app}(t) = 0.5 \langle x^2(t) \rangle / t$  instead of the derivative (4) to remove the oscillations of the time  
 381 derivative [Schwarze, et al., 2001; Trefry, et al., 2003]. Even though  $D_{app}$  tends to the  
 382 effective dispersion coefficient (4) for large times, the differences between these two  
 383 quantities are important and remain for very large times especially in the high variance case  
 384 as shown by figure 6 on the simulation averages (dashed-dotted lines compared to solid lines).  
 385 We thus decide to determine the asymptotic dispersion coefficient  $D_{LA}^i$  from the time  
 386 derivative signal (4).

387 **III.5. Validation**

388 We validate the numerical procedures by comparing them to theoretical and other numerical  
 389 existing results. First, for  $\sigma^2 < 1$ , we compare numerical results to first-order theoretical results  
 390 of the correlation functions of the longitudinal and transverse velocity fields (equations 9 and  
 391 10 of [Rubin, 1990]) and of the asymptotic dispersion coefficients. The velocity correlation  
 392 function is highly close at less than 5% to the first-order prediction for  $\sigma^2 < 1$  and is close at  
 393 less than 1.5% to the results of *Salandin and Fiorotto* [1998] for  $1 < \sigma^2 < 4$  (figure 8).

394 For the asymptotic longitudinal dispersion coefficient, the normalization by  $\lambda u$  enables a  
 395 direct comparison of  $D_{LA}$  numerical results with  $\sigma^2$  for  $\sigma^2 < 1$  with equation (2). The agreement  
 396 is very good as, for  $\sigma^2 = 0.25$ , equation (2) and numerical simulations give respectively 0.25  
 397 and 0.26. We compare also our results of longitudinal asymptotic dispersion coefficient in the  
 398 interval  $1 < \sigma^2 < 4$  to the variation of  $D_{LA}$  with  $\sigma$  obtained by *Salandin and Fiorotto* [1998].  
 399 *Salandin and Fiorotto* [1998] found  $D_L \propto \sigma^\beta$  with  $\beta = 2.06, 2.19, 2.29$  and  $2.35$  respectively  
 400 for  $\sigma^2$  in the intervals  $[0.05, 1]$ ,  $[1, 2]$ ,  $[2, 3]$  and  $[3, 4]$ . We find  $\beta = 2.07, 2.37$  and  $2.62$  for  $\sigma^2$  in  
 401 the intervals  $[0.05, 1]$ ,  $[1, 2.25]$ , and  $[2.25, 4]$ . We find a close agreement in the first interval but  
 402 a faster increase of the asymptotic dispersion coefficient for  $\sigma^2 > 1$ . This could be linked to the  
 403 number of correlation lengths limited to 20 in *Salandin and Fiorotto* [1998]. For advective-  
 404 diffusive transport, we validate the algorithm against the classical analytical solution obtained  
 405 in the homogeneous medium case.

## 406 IV. Results of asymptotic dispersion coefficients

### 407 IV.1. Pure advection ( $Pe=\infty$ )

408 We determine the effect of the resolution scale defined by the number of cells by correlation  
 409 length  $\lambda/l_m$ . We perform the simulations at  $\lambda/l_m=5, 10$  and  $20$  where previous simulations  
 410 used a maximum of 8 cells per correlation length (Table 1). We verify that the asymptotic  
 411 longitudinal dispersion coefficient depends linearly on  $\lambda$  for all values of  $\sigma^2$  (figure 9)  
 412 justifying furthermore the normalization of the dispersion coefficients by  $\lambda u$ . It also shows the  
 413 weak dependency of the asymptotic dispersion coefficient on the resolution of the  
 414 discretization even for high heterogeneities. 10 cells by correlation length give very similar  
 415 results as 5 or 20 cells by correlation length.

416 The asymptotic regime has been reached and maintained over at least 500 correlation lengths  
 417 whatever the value of  $\sigma^2$  (figure 10) and the asymptotic values of the dispersion coefficients  
 418 have been computed according to the procedure described in the previous section (figure 11).  
 419 Both methodologies of exponential fitting and averaging lead to similar results within an  
 420 interval of 0% to 3%. The first-order estimate of the dispersion coefficient (2) remains close  
 421 to the numerical value even for  $\sigma^2=1$  and  $2.25$  where it is lower by respectively 10% and  
 422 25%. This good performance of first-order results for values of  $\sigma^2$  significantly larger than 1  
 423 has been previously observed and explained [Bellin, et al., 1992; Dagan, et al., 2003]. The  
 424 independent linearizations of flow and transport induce opposite deviations from linear  
 425 theoretical results and may partly cancel out each other. This conclusion was confirmed on  
 426 slightly different heterogeneous media consisting of spherical inclusions in a homogeneous  
 427 medium [Dagan, et al., 2003]. For larger heterogeneity, the departure from the first-order  
 428 results increases with  $\sigma^2$ . Numerical results are respectively 50%, 90% and 150% larger than

429 the linear estimates for  $\sigma^2=4, 6.25$  and  $9$ .  $D_{LA}(av)$  is well represented by the approximate  
 430 function  $0.7 \sigma^2 + 0.2 \sigma^4$  for large heterogeneities ( $\sigma^2 > 1$ ) (dashed curve on figure 11).

431 For transverse dispersivity, numerical results show some variability around 0 without any  
 432 systematic trend neither for the realization-based result nor for the average (figures 7 and 12).  
 433 The similar transverse dispersion evolution with time for  $\sigma^2=6.25$  and  $9$  comes from the fact  
 434 that realizations are performed with the same set of seeds for the random generator. The  
 435 correlation patterns are thus identical while the magnitude of the heterogeneity changes. The  
 436 asymptotic dispersion coefficients computed by averaging over the second half of the time  
 437 chronicle  $D_{TA}(av)$  are close to zero without being systematically positive or negative  
 438 (figure 13) and the magnitude of the standard deviation is much larger than the average.  
 439 These results lead us to conclude that the asymptotic transverse dispersion coefficient is zero  
 440 on average whatever  $\sigma^2$ . This confirms theoretical conclusions obtained by volume averaging  
 441 [Attinger, et al., 2004].

442 Figures 6 and 7 show a large variability around the average both for the longitudinal and  
 443 transverse dispersion coefficients whatever the heterogeneity represented by the value of  $\sigma^2$ .  
 444 The standard deviation of the transverse dispersion coefficient converges (figure 14b) within  
 445 the computation time, whereas the convergence is not obvious for the standard deviation of  
 446 the longitudinal dispersion coefficient (figure 14a). The apparent increase of the longitudinal  
 447 dispersion coefficient remains limited to at most 30% in the time interval  $[t_{fb}/2, t_{fb}]$ , which is  
 448 close to the imprecision of 20% obtained in section III.3 because of the use of a limited  
 449 number of simulations ( $N_S$ ). Convergence would require both more realizations and longer  
 450 systems. As the increase remains limited and as  $\sigma(D_L(t_N))$  is not the main objective of the  
 451 study, we did not go further on its characterization.

452 Finally, we derive from the exponential fit of the longitudinal dispersion coefficient the  
 453 characteristic convergence time to the asymptotic regime  $t_{N0}$  (figure 15).  $t_{N0}$  does not have an  
 454 absolute meaning as it depends on the width of the injection window. We rather use  $t_{N0}$  to  
 455 compare convergence time between different values of  $\sigma^2$  in the same conditions.  $t_{N0}$   
 456 increases exponentially with the permeability variance contrarily to the first-order theory  
 457 prediction according to which  $t_{N0}$  does not depend on the medium heterogeneity  $\sigma^2$ .

#### 458 ***IV.2. Advection and diffusion ( $Pe < \infty$ )***

459 We computed the dispersion coefficient  $D_L(t_N)$  for the two Peclet number  $Pe=100$  and  $1000$ .  
 460  $D_L(t)$  reaches its asymptotic regime whatever the value of  $\sigma^2$  (figure 16). The time to reach  
 461 the asymptotic dispersion  $t_{N0}$  is smaller than in the pure-advection regime (figure 15) even if  
 462  $t_{N0}$  values are highly dispersed. Diffusion modifies only slightly the asymptotic longitudinal  
 463 dispersion coefficient  $D_{LA}$  for  $\sigma^2 \leq 1$  and let it decrease for  $\sigma^2 > 1$  (Table 2). For small values  
 464 of  $\sigma^2$  ( $\sigma^2 \leq 1$ ), the influence of diffusion is negligible as previously found [Fiori, 1996]. For  
 465  $\sigma^2=1$ , the additional dispersion induced by diffusion is not significant because the asymptotic  
 466 dispersion coefficient for  $Pe=\infty$  (larger than 1) is ten times larger than  $1/Pe$ . For  $\sigma^2 > 1$ , the  
 467 asymptotic dispersion coefficient decreases surprisingly with more diffusion. More diffusion  
 468 induces less dispersion. The decrease can be significant. For  $\sigma^2=6.25$ ,  $D_{LA}$  is 25% lower at  
 469  $Pe=100$  than its value at  $Pe=\infty$  (pure advection). For large heterogeneities, diffusion reduces  
 470 the global dispersion. This behavior was expected by Gelhar [1993] (pages 221-222) and de  
 471 Arcangelis et al. [1986] and may be explained by the following argument also invoked for  
 472 percolation systems [Koplik, et al., 1988]. Large dispersion is induced by the widely-scattered  
 473 velocity distribution. Diffusion introduces a cut-off to this distribution thus narrowing it and  
 474 letting in turn the dispersion coefficient decrease. In other words, diffusion extracts particles  
 475 from the very slow velocity zones and restricts the dispersion of particle in the medium. The

476 transverse asymptotic dispersion coefficient  $D_{TA}$  keeps a more classical behavior by  
 477 increasing with more diffusion (Table 2). However the increase of  $D_{TA}$  can be much larger  
 478 than the sole diffusion contribution  $1/Pe$ . For large heterogeneities  $\sigma^2=6.25$  and for  $Pe=100$ ,  
 479  $D_{TA}$  is 20 times larger than  $1/Pe$ . The effect of diffusion and advection cannot be simply  
 480 superposed but interact to produce a larger transverse dispersion.

## 481 V. Conclusion

482 We determine the asymptotic dispersion coefficients for 2D exponentially correlated  
 483 lognormal permeability fields on a broad range of lognormal permeability variance  $\sigma^2$   
 484 ( $\sigma^2 \in [0.25,9]$ ). We use parallel computing for simulating fluid flow and particle transport on  
 485 large domains of typical dimension from 800 to 1600 correlation lengths with a resolution of  
 486 10 cells by correlation length, where  $l_m$  is the cell characteristic dimension. Such large  
 487 domains turned out to be necessary to observe the asymptotic regime on a sufficiently long  
 488 time range for determining unambiguously the asymptotic dispersion coefficients. The  
 489 asymptotic longitudinal and transverse dispersion coefficient  $D_{LA}$  and  $D_{TA}$  have been  
 490 estimated on a realization basis by averaging over a traveled distance of at least 400  
 491 correlation lengths. We have tested an alternative derivation methodology for the asymptotic  
 492 longitudinal dispersion coefficient  $D_{LA}$  by fitting the dispersion coefficient by an exponential  
 493 function. Estimates of  $D_{LA}$  by both methodologies lead to very similar values. The  
 494 characteristic time given by the exponential fit gives an estimate of the convergence speed to  
 495 the asymptotic regime. Simulations show that it increases exponentially with the  
 496 heterogeneity  $\sigma^2$  and decreases with diffusion.

497 For pure advection ( $Pe=\infty$ ), the asymptotic longitudinal dispersion  $D_{LA}$  is larger than the first-  
 498 order estimate for high heterogeneity. More precisely, for  $\sigma^2$  equal to 4, 6.25 and 9,  $D_{LA}$  is  
 499 larger by respectively 50%, 90% and 150% than the linear estimates. For  $\sigma^2>1$ ,  $D_{LA}$  is well

500 fitted by the function  $0.7 \sigma^2 + 0.2 \sigma^4$  showing a quadratic evolution in  $\sigma^2$  for large  
501 heterogeneities. This departure from the first-order theory is probably related to the extreme  
502 flow channeling observed for high heterogeneity [Le Borgne, *et al.*, submitted; Moreno and  
503 Tsang, 1994; Salandin and Fiorotto, 1998]. Whatever the heterogeneity level, the asymptotic  
504 transverse dispersion coefficient is always zero as predicted by first-order theory for low  
505 heterogeneity and by volume averaging [Attinger, *et al.*, 2004].

506 The addition of diffusion to advection leads to two very different behaviors for longitudinal  
507 and transverse dispersions. For large heterogeneities ( $\sigma^2 > 1$ ), diffusion induces a significant  
508 longitudinal dispersion decrease and a transverse dispersion increase larger than expected. At  
509 most, for a Peclet number of 100 (advection on average hundred times larger than diffusion)  
510 and a permeability variance  $\sigma^2 = 9$ , the longitudinal dispersion decreases by a factor of 2 and  
511 the transverse dispersion is 7.5 times larger than the local diffusion.

512

513 **Acknowledgements:** This work was supported by Grid'5000 grants for executing simulations  
514 on the grid at Irista in Rennes. Comments by referees and an editor have helped  
515 improve the revised version of the paper.

## 516 **References**

517 Ababou, R., D. McLaughlin, L. W. Gelhar, and A. F. B. Tompson (1989), Numerical  
518 simulation of three-dimensional saturated flow in randomly heterogeneous porous  
519 media, *Transport in Porous Media*.

520 Attinger, S., M. Dentz, and W. Kinzelbach (2004), Exact transverse macro dispersion  
521 coefficients for transport in heterogeneous porous media, *Stochastic Environmental*  
522 *Research and Risk Assessment*.

- 523 Beaudoin, A., J. R. de Dreuzy, and J. Erhel (2007), An efficient parallel tracker for advection-  
524 diffusion simulations in heterogeneous porous media, paper presented at Europar,  
525 Rennes, France, 28-31 August 2007.
- 526 Bellin, A., P. Salandin, and A. Rinaldo (1992), Simulation of dispersion in heterogeneous  
527 porous formations: statistics, first-order theories, convergence of computations, *Water*  
528 *Resources Research*, 28, 2211-2227.
- 529 Chavent, G., and J. E. Roberts (1991), A unified physical presentation of mixed, mixed-  
530 hybrid finite elements and standard finite difference approximations for the  
531 determination of velocities in waterflow problems, *Advances in Water Resources*, 14,  
532 329-348.
- 533 Cvetkovic, V., H. Cheng, and X.-H. Wen (1996), Analysis of nonlinear effects on tracer  
534 migration in heterogeneous aquifers using Lagrangian travel time statistics, *Water*  
535 *Resources Research*, 32, 1671-1680.
- 536 Dagan, G. (1989), *Flow and Transport in Porous Formations*, 465 pp., Springer Verlag.
- 537 Dagan, G., A. Fiori, and I. Jankovic (2003), Flow and transport in highly heterogeneous  
538 formations: 1. Conceptual framework and validity of first-order approximations, *Water*  
539 *Resources Research*, 9.
- 540 de Arcangelis, L., J. Koplik, S. Redner, and D. Wilkinson (1986), Hydrodynamic Dispersion  
541 in Network Models of Porous Media, *Physical Review Letters*, 57, 996-999.
- 542 Delay, F., P. Ackerer, and C. Danquigny (2005), Solution of solute transport in porous or  
543 fractured formations by random walk particle tracking: a review, *Vadose Zone Journal*,  
544 4, 360-379.

- 545 Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2000), Temporal behavior of a  
546 solute cloud in a heterogeneous porous medium, 1, Point-like injection, *Water*  
547 *Resources Research*, 36.
- 548 Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2002), Temporal behavior of a  
549 solute cloud in a heterogeneous porous medium: 3. Numerical simulations, *Water*  
550 *Resources Research*, 7.
- 551 Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2003), Numerical studies of the  
552 transport behavior of a passive solute in a two-dimensional incompressible random flow  
553 field, *Physical Review E*, 67.
- 554 Eberhard, J. (2004), Approximations for transport parameters and self-averaging properties  
555 for point-like injections in heterogeneous media, *Journal of Physics a-Mathematical*  
556 *and General*, 37, 2549-2571.
- 557 Falgout, R. D., J. E. Jones, and U. M. Yang (2005), Pursuing scalability for HYPRE's  
558 conceptual interfaces, *ACM Transactions on mathematical software*, 31.
- 559 Fiori, A. (1996), Finite Peclet extensions of Dagan's solutions to transport in anisotropic  
560 heterogeneous formations, *Water Resources Research*, 32, 193-198.
- 561 Frigo, M., and S. G. Johnson (2005), The Design and Implementation of FFTW3,  
562 *Proceedings of the IEEE*, 93, 216-231.
- 563 Gelhar, L. W. (1993), *Stochastic Subsurface Hydrology*, Engelwood Cliffs, New Jersey.
- 564 Gutjahr, A. L. (1989), Fast Fourier transforms for random field generation (*ed.*), New Mexico  
565 Tech project report 4-R58-2690R.

- 566 Hassan, A. E., R. Andricevic, and V. Cvetkovic (2002), Evaluation of analytical solute  
567 discharge moments using numerical modeling in absolute and relative dispersion  
568 frameworks, *Water Resources Research*, 38.
- 569 Hsu, K. C., D. Zhang, and S. P. Neuman (1996), Higher-order effects on flow and transport in  
570 randomly heterogeneous porous media, *Water Resources Research*, 32, 571-582.
- 571 Jankovic, I., A. Fiori, and G. Dagan (2003), Flow and transport in highly heterogeneous  
572 formations: 3. Numerical simulations and comparison with theoretical results, *Water*  
573 *Resources Research*, 9.
- 574 Jankovic, I., A. Fiori, and G. Dagan (2006), Modeling flow and transport in highly  
575 heterogeneous three-dimensional aquifers: Ergodicity, Gaussianity, and anomalous  
576 behavior - 1. Conceptual issues and numerical simulations, *Water Resources Research*,  
577 42.
- 578 Koplik, J., S. Redner, and D. Wilkinson (1988), Transport and Dispersion in Random  
579 Networks With Percolation Disorder, *Physical Review a*, 37.
- 580 Le Borgne, T., J.-R. d. Dreuzy, P. Davy, and O. Bour (submitted), Characterization of the  
581 velocity field organization in heterogeneous media by conditional correlations, *Water*  
582 *Resources Research*.
- 583 Meurant, G. (1999), *Computer solution of large linear systems*, North Holland, Amsterdam.
- 584 Moreno, L., and C.-F. Tsang (1994), Flow channeling in strongly heterogeneous porous  
585 media: A numerical study, *Water Resources Research*, 30, 1421-1430.
- 586 Mosé, R., P. Siegel, and P. Ackerer (1994), Application of the mixed hybrid finite element  
587 approximation in a groundwater model: Luxury or necessity?, *Water Resources*  
588 *Research*, 30, 3001-3012.

- 589 Pokrajac, D., and R. Lazic (2002), An efficient algorithm for high accuracy particle tracking  
590 in finite elements, *Advances in Water Resources*, 25, 353-369.
- 591 Pollock, D. W. (1988), Semianalytical computation of path lines for finite-difference models,  
592 *Ground Water*, 26, 743-750.
- 593 Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar (1992), Field study of dispersion in a  
594 heterogeneous aquifer, 3, geostatistical analysis of hydraulic conductivity, *Water*  
595 *Resources Research*, 28.
- 596 Rubin, Y. (1990), Stochastic modeling of macrodispersion in heterogeneous porous media,  
597 *Water Resources Research*, 26, 133-141.
- 598 Saad, Y. (1996), *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company.
- 599 Salandin, P., and V. Fiorotto (1998), Solute transport in highly heterogeneous aquifers, *Water*  
600 *Resources Research*, 34, 949-961.
- 601 Schwarze, H., U. Jaekel, and H. Vereecken (2001), Estimation of Macrodispersion by  
602 Different Approximation Methods for Flow and Transport in Randomly Heterogeneous  
603 Media, *Transport in Porous Media*, 43, 265-287.
- 604 Tompson, A. F. B., and L. W. Gelhar (1990), Numerical simulation of solute transport in  
605 three-dimensional, randomly heterogeneous porous media, *Water Resources Research*,  
606 26, 2541-2562.
- 607 Toseli, A., and O. Widlund (2005), *Domain decomposition methods-algorithms and theory*,  
608 springer series in computational mathematics.
- 609 Trefry, M. G., F. P. Ruan, and D. McLaughlin (2003), Numerical simulations of  
610 preasymptotic transport in heterogeneous porous media: Departures from the Gaussian  
611 limit, *Water Resources Research*, 39.

- 612 Wesseling, P. (2004), *An Introduction to Multigrid Methods*, Edwards.
- 613 Yao, T. (2004), Reproduction of the Mean, Variance, and Variogram Model in Spectral  
614 Simulation, *Mathematical Geology*, 36, 487-506.
- 615

616 **Figure captions**

617

618 *Figure 1: Permeability field stored on four processors, boundary conditions, injection and*  
 619 *exclusion zones. The characteristics of the computational domain are  $L_x=2048.l_m=204,8.\lambda$ ,*  
 620  *$L_y=1024.l_m=102,4.\lambda$ ,  $\lambda=10.l_m$  and  $\sigma^2=2.25$  where  $\lambda$  is the correlation length and  $l_m$  is the*  
 621 *grid cell size. Permeability is increasing from blue to red. Computational domains used for*  
 622 *asymptotic dispersion determination where 4 to 8 times longer and larger than this one.*

623 *Figure 2: a) Longitudinal and b) transverse dispersion coefficients as functions of time for*  
 624 *increasing particle numbers with  $\sigma^2=9$ ,  $\lambda=10 l_m$  and  $L_x=819,2.\lambda$  and  $L_y=819,2.\lambda$  (pure*  
 625 *advection case).*

626 *Figure 3: Asymptotic longitudinal (a) and transverse (b) dispersion coefficients as functions*  
 627 *of the particle number  $N_p$  for  $L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ . In this figure as well as in the*  
 628 *following figures, the term advection in the legend refers to the pure advection case without*  
 629 *diffusion and the legend is the same for both graphs.*

630 *Figure 4: a) Asymptotic longitudinal dispersion coefficient and b) standard deviation of the*  
 631 *dispersion coefficient at a given time  $t_N=600$  as functions of the number of simulations for*  
 632  *$L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ . In a), the dispersion coefficient  $D_L$  for  $t_N=600$  has been added*  
 633 *for the pure advection cases.  $t_N=600$  is taken in the second half of the signal as the full signal*  
 634 *length is around  $t_N=1000$ .*

635 *Figure 5: a) Asymptotic transverse dispersion coefficient and b) standard deviation of the*  
 636 *dispersion coefficient at a given time  $t_N=600$  as functions of the number of simulations. Same*  
 637 *parameters as in figure 6.*

638 *Figure 6: Normalized longitudinal dispersion coefficient for single realizations  $D_L^i(t_N)$*   
 639 *(points), their averages  $D_L(t_N)$  over 100 realizations (lines) and the confidence interval at*  
 640 *95% on the dispersion coefficient (dashed line) ( $\sigma^2=1$  and 9, pure advection case). The*  
 641 *dashed-dotted line represents the normalized apparent dispersion coefficient*  
 642  *$D_{app}(t) = 0.5 \langle x^2(t) \rangle / t$ . Computational domain size are for  $\sigma^2=1$  (a)  $L_x=L_y=819,2 \lambda$  and*  
 643 *for  $\sigma^2=9$  (b)  $L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ .*

644 *Figure 7: Normalized transverse dispersion coefficient. Same parameters as in figure 6.*

645 *Figure 8: Velocity variance  $u_{xx}$  and  $u_{yy}$  as functions of  $\sigma^2$  obtained analytically in*  
 646 *Rubin [1990] and numerically in the present study and in Salandin and Fiorotto [1998].*

647 *Figure 9: Asymptotic longitudinal dispersion coefficient as a function of the correlation*  
 648 *length for the pure advection case. Lines are linear fit through 0. Same parameters as in*  
 649 *figure 6.*

650 *Figure 10: Longitudinal mean dispersion coefficient as a function of  $t_N$  for the pure advection*  
 651 *case (time in terms of correlation scales crossed by the plume). Dashed lines mark the*  
 652 *asymptotic coefficients. Same parameters as in figure 6.*

653 *Figure 11: Normalized longitudinal asymptotic effective dispersion coefficient  $D_{LA}$  as a*  
 654 *function of the variance of the log conductivity with pure advection. Vertical bars on data*  
 655 *points represent the standard deviation on each side of the data point.  $D_{LA}(av)$  and  $D_{LA}(fit)$*   
 656 *are obtained respectively by averaging and fitting by an exponential function. Theoretical*  
 657 *predictions [Gelhar, 1993] are represented by the line. The dashed curve stands for*  
 658  *$0.7 \sigma^2 + 0.2 \sigma^4$ . Same parameters as in figure 6.*

659 *Figure 12: Normalized transverse dispersion coefficient as a function of the normalized time*  
660 *in the pure advection case. Same parameters as in figure 6.*

661 *Figure 13: Normalized transverse asymptotic dispersion coefficient for the pure advection*  
662 *case.*

663 *Figure 14: Standard deviation of a) the longitudinal and b) transverse dispersion coefficients*  
664 *in the pure advection case.*

665 *Figure 15: Characteristic convergence time to the asymptotic regime  $t_{N0}$ .*

666 *Figure 16: Longitudinal dispersion coefficient as a function of normalized time for  $\sigma^2 \geq 4$ .*

	FCM	TP	TCM	$\sigma^2$	$\lambda/l_m$	$L_x/l_m$	$L_y/l_m$	$N_m$	MC	PT
[Rubin, 1990]	1 <sup>st</sup>	A	PT	0.79						300
[Bellin, et al., 1992]	full	A	PT	[0,1.6]	8	36	18	$4 \cdot 10^4$	MC · PT = 1500	
[Cvetkovic, et al., 1996]	full	A	PT	[0,4]	4	24	18	$7 \cdot 10^3$	500-1000	1
[Salandin and Fiorotto, 1998]	full	A	PT	[0.05,4]	2,4,8	64	64	$2.6 \cdot 10^5$	500	40
[Schwarze, et al., 2001]	1 <sup>st</sup>	AD	PT	[0.1,1]	50	5000			3200	1
[Hassan, et al., 2002]	full	A	PT	[0.25,2.25]	5	50	25	$3 \cdot 10^4$	2000-3000	
[Dentz, et al., 2002]	1 <sup>st</sup>	AD	PT	[0.1,2]	20	1500			2000	100
[Trefry, et al., 2003]	full	A $\alpha$	NS	[0.25,4]	8 2	256-1024	64-256	$10^6$	1	
<i>this study</i>	full	AD	PT	[0.25,8]	10	819-1638	819	$7 \cdot 10^7$ - $1.4 \cdot 10^8$	100	2000

Table 1: Characteristics of 2D flow and transport simulations. FCM stands for flow computation method. It can be 1<sup>st</sup> order when flow is obtained by first order approximation of the flow equation or full when flow is obtained by solving directly the full discretized flow equation. TP is the transport processes accounted for (A for advection, D for diffusion,  $\alpha$  for dispersion). TCM stands for transport computation method (PT for particle tracking, NS for numerical scheme).  $N_x$  and  $N_y$  are the number of correlation lengths within the domain respectively in the main direction of flow and perpendicularly to it.  $N_m$  is the total number of cells ( $N_m = L_x L_y / l_m^2$ ). MC realizations is the number of Monte-Carlo realizations per parameter set. PT trajectories is the total number of analyzed trajectories per realization when particle tracking is used.

	$D_{LA}(Pe)/D_{LA}(\text{advection})$		$D_{TA}(Pe)$	
	$Pe=10^2$	$Pe=10^3$	$Pe=10^2$	$Pe=10^3$
$\sigma^2=0.25$	1.08	1.06	0.01	0.00
$\sigma^2=1$	1.05	1.06	0.00	-0.01
$\sigma^2=2.25$	0.98	1.01	-0.06	0.05
$\sigma^2=4$	0.90	1.01	-0.06	-0.09
$\sigma^2=6.25$	0.75	0.96	0.20	0.14
$\sigma^2=9$	0.57	0.84	0.24	0.16

Table 2: Asymptotic longitudinal dispersion coefficient normalized by its pure advective counterpart  $D_{LA}(Pe)/D_{LA}(\text{advection})$  and asymptotic transversal dispersion coefficient as functions of  $\sigma^2$  and  $Pe$ . Dispersion coefficients are obtained with the averaging method.

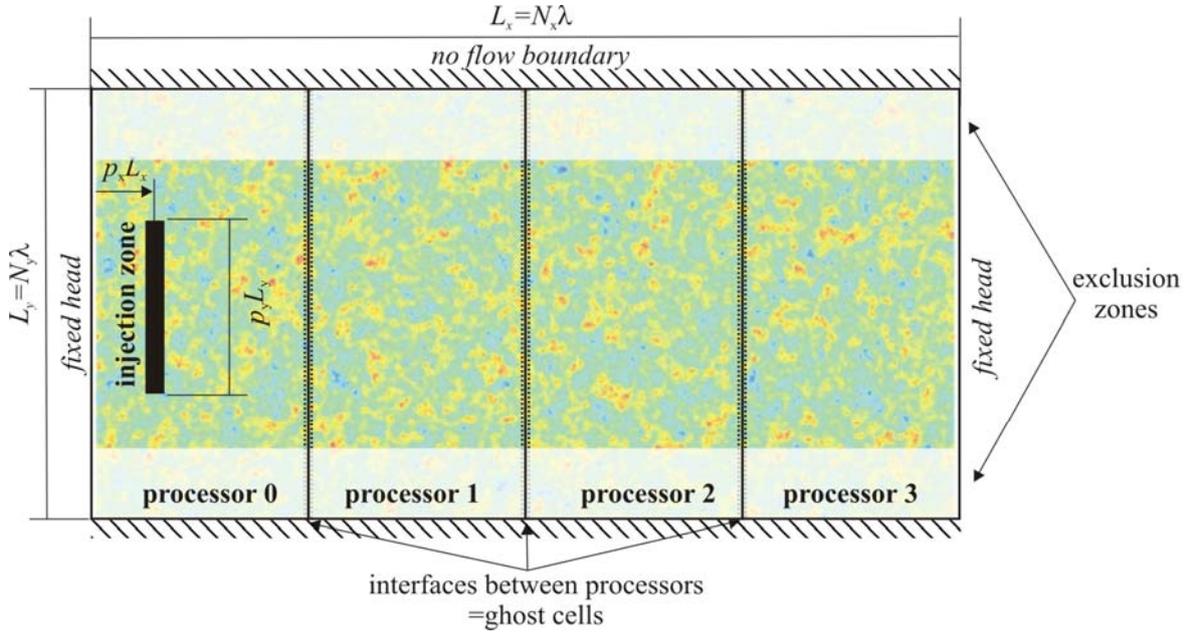


Figure 1: Permeability field stored on four processors, boundary conditions, injection and exclusion zones. The characteristics of the computational domain are  $L_x=2048.l_m=204,8.\lambda$ ,  $L_y=1024.l_m=102,4.\lambda$ ,  $\lambda=10.l_m$  and  $\sigma^2=2.25$  where  $\lambda$  is the correlation length and  $l_m$  is the grid cell size. Permeability is increasing from blue to red. Computational domains used for asymptotic dispersion determination where 4 to 8 times longer and larger than this one.

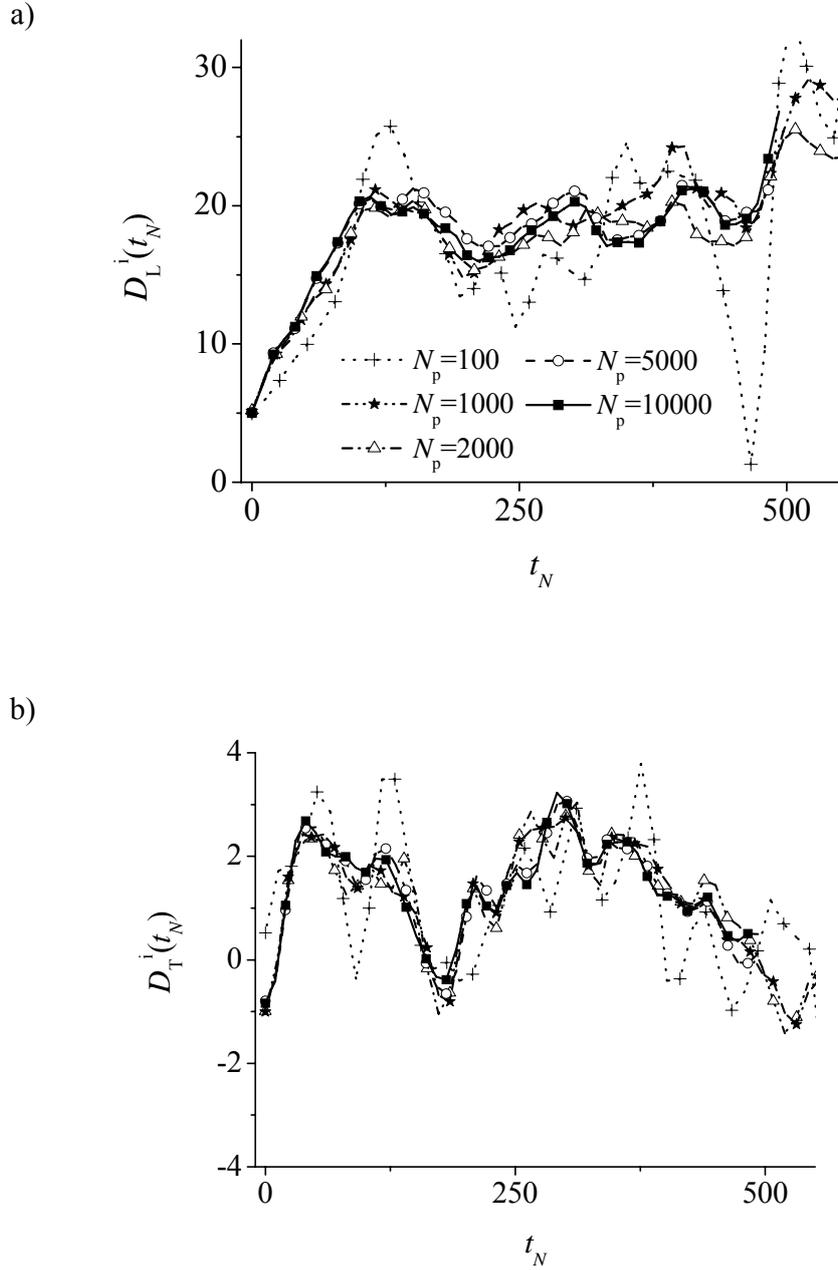
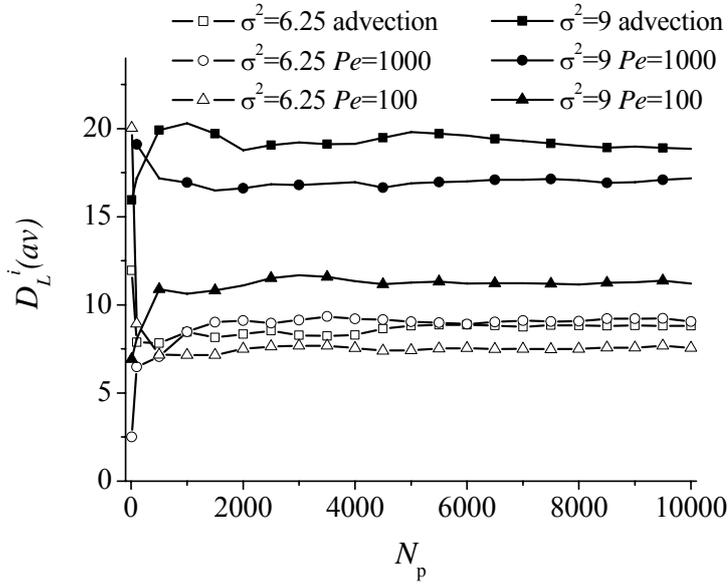


Figure 2 : a) Longitudinal and b) transverse dispersion coefficients as functions of time for increasing particle numbers with  $\sigma^2=9$ ,  $\lambda=10 l_m$  and  $L_x=819,2.\lambda$  and  $L_y=819,2.\lambda$  (pure advection case).

a)



b)

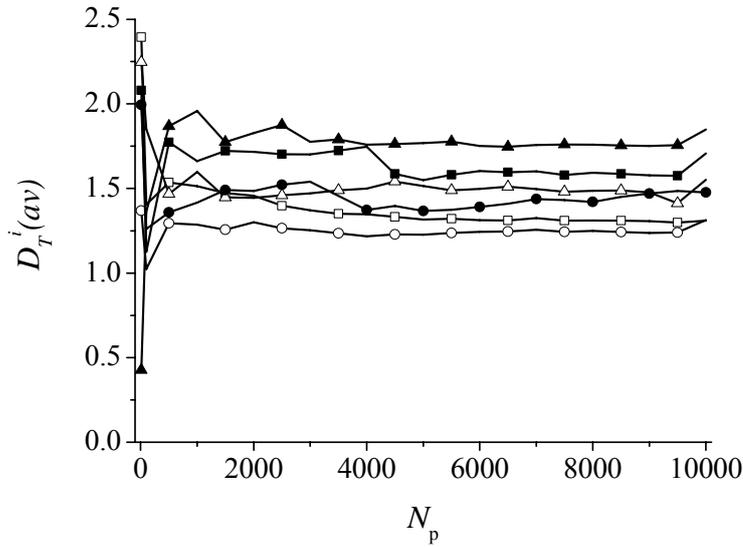


Figure 3: Asymptotic longitudinal (a) and transverse (b) dispersion coefficients as functions of the particle number  $N_p$  for  $L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ . In this figure as well as in the following figures, the term advection in the legend refers to the pure advection case without diffusion and the legend is the same for both graphs.

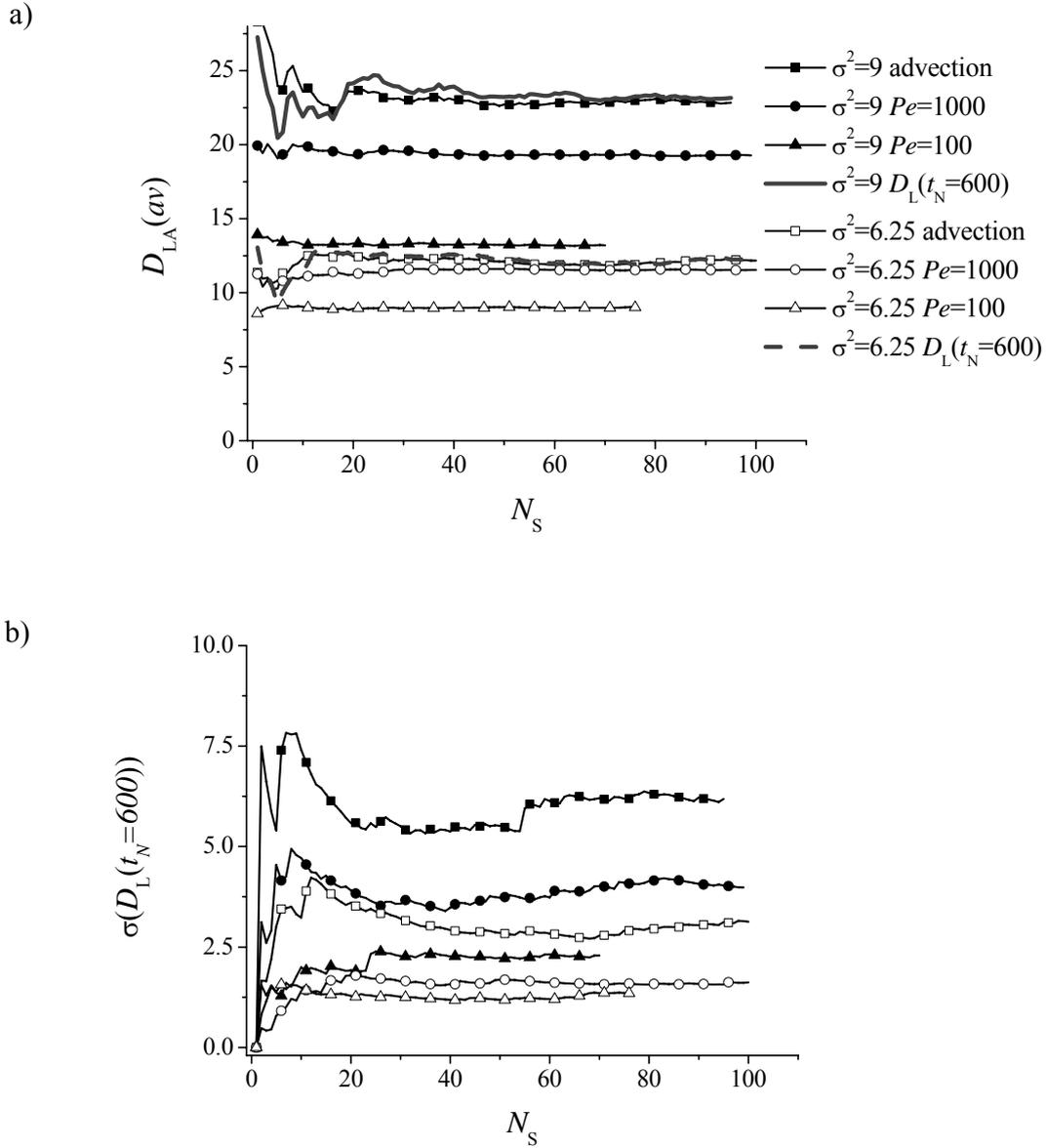
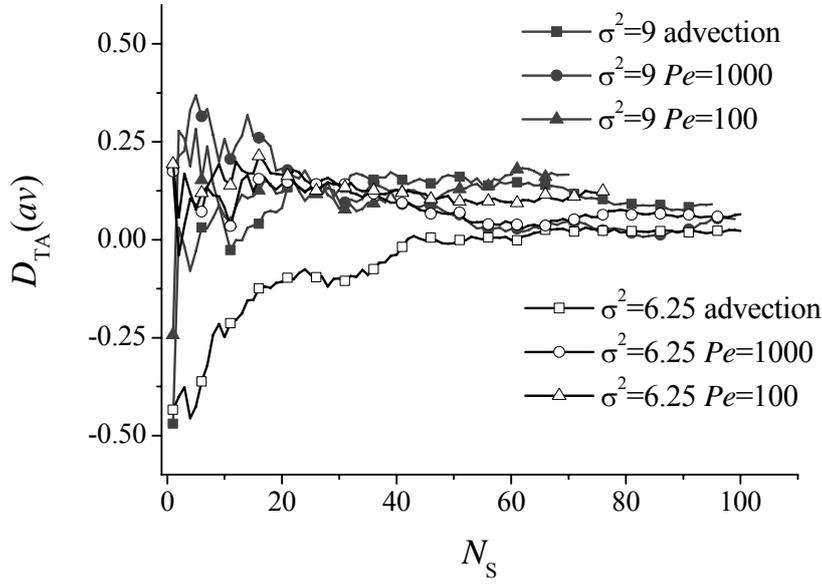


Figure 4: a) Asymptotic longitudinal dispersion coefficient and b) standard deviation of the dispersion coefficient at a given time  $t_N=600$  as functions of the number of simulations for  $L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ . In a), the dispersion coefficient  $D_L$  for  $t_N=600$  has been added for the pure advection cases.  $t_N=600$  is taken in the second half of the signal as the full signal length is around  $t_N=1000$ .

a)



b)

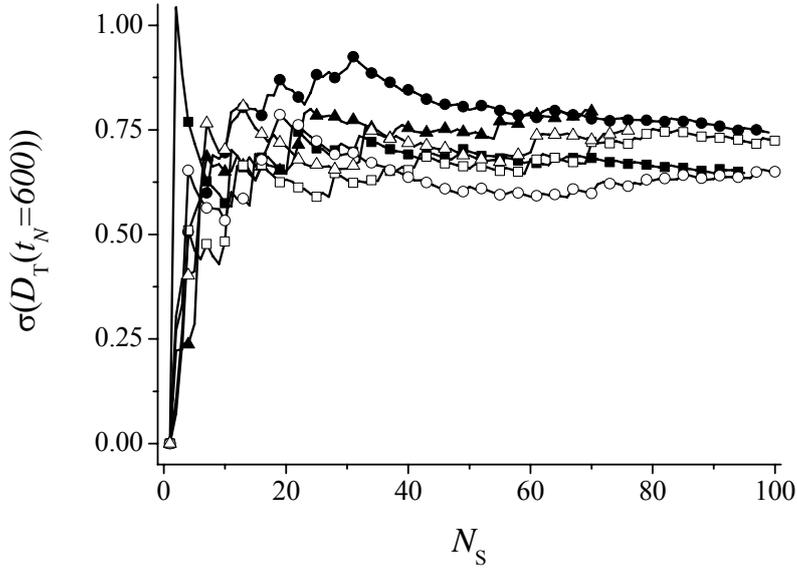


Figure 5: a) Asymptotic transverse dispersion coefficient and b) standard deviation of the dispersion coefficient at a given time  $t_N=600$  as functions of the number of simulations. Same parameters as in figure 6.

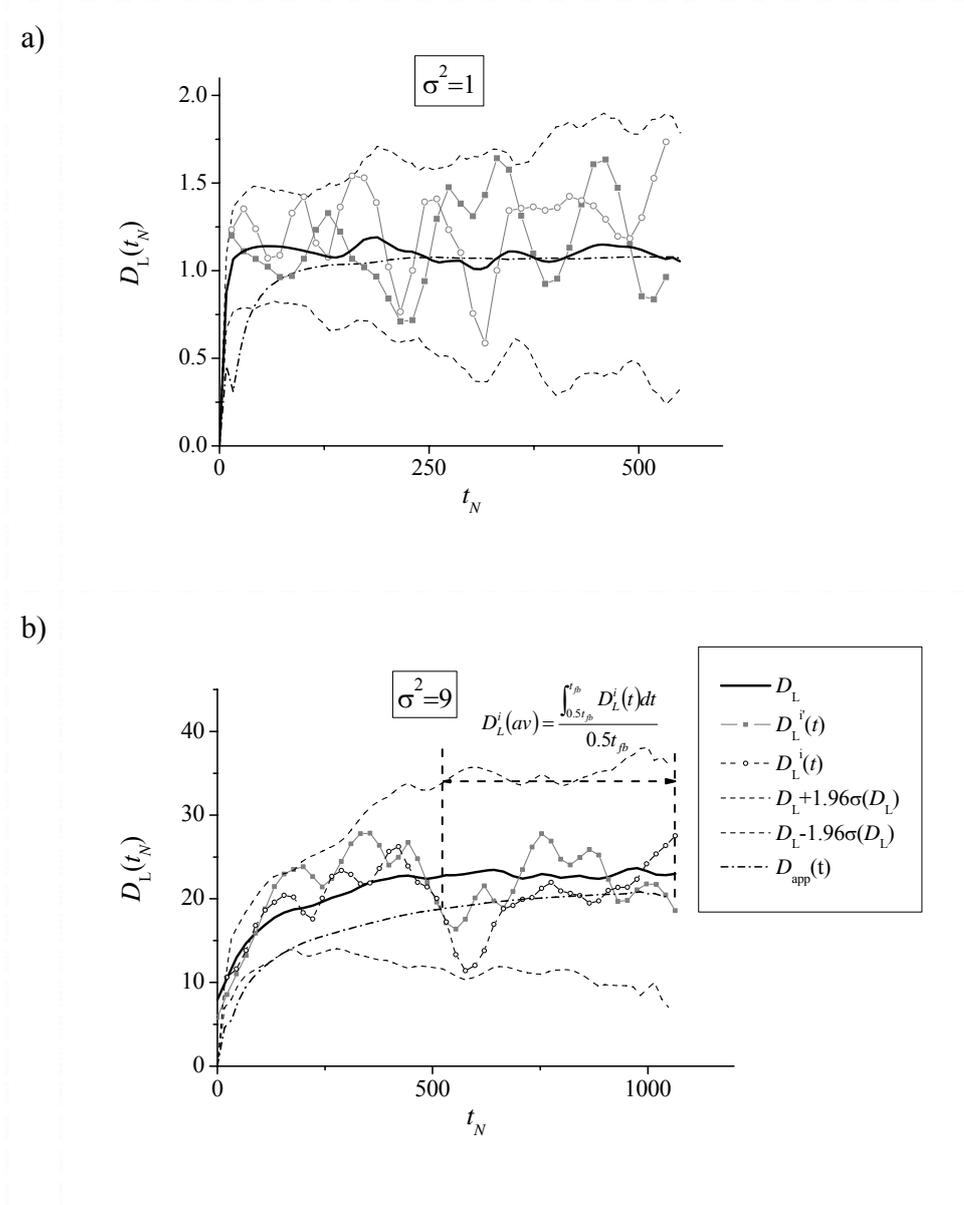


Figure 6: Normalized longitudinal dispersion coefficient for single realizations  $D_L^i(t_N)$  (points), their averages  $D_L(t_N)$  over 100 realizations (lines) and the confidence interval at 95% on the dispersion coefficient (dashed line) ( $\sigma^2=1$  and 9, pure advection case). The dashed-dotted line represents the normalized apparent dispersion coefficient  $D_{app}(t) = 0.5 \langle x^2(t) \rangle / t$ . Computational domain size are for  $\sigma^2=1$  (a)  $L_x=L_y=819,2 \lambda$  and for  $\sigma^2=9$  (b)  $L_x=1638,4 \lambda$  and  $L_y=819,2 \lambda$ .

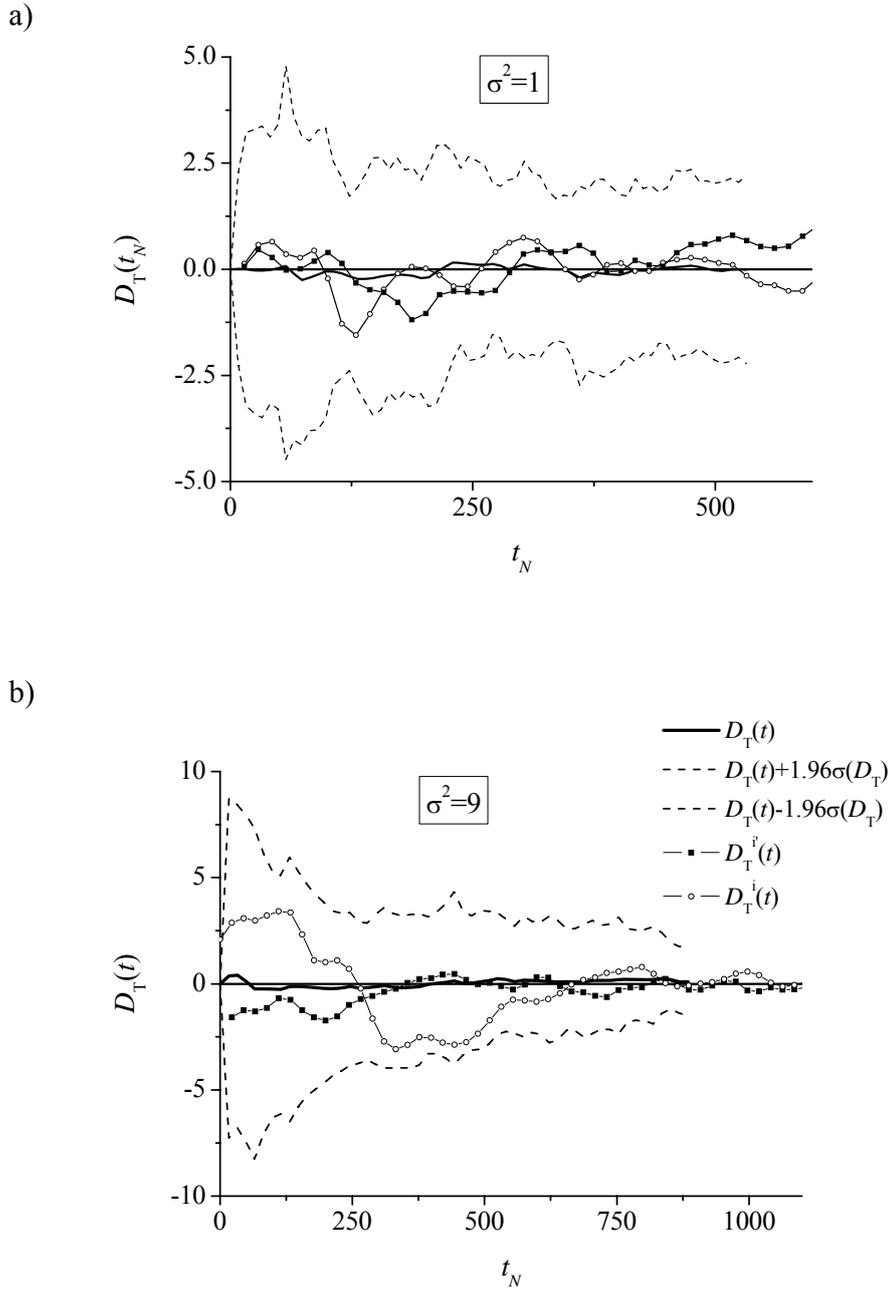


Figure 7: Normalized transverse dispersion coefficient. Same parameters as in figure 6.

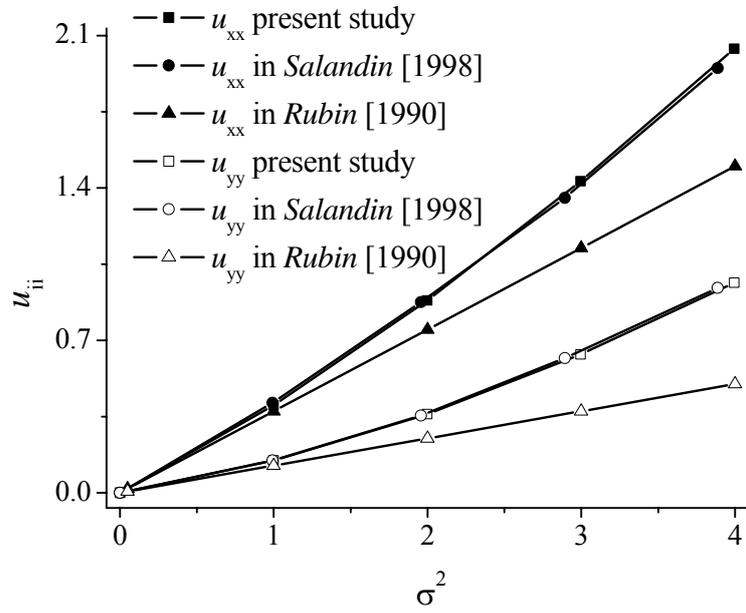


Figure 8: Velocity variance  $u_{xx}$  and  $u_{yy}$  as functions of  $\sigma^2$  obtained analytically in Rubin [1990] and numerically in the present study and in Salandin and Fiorotto [1998].

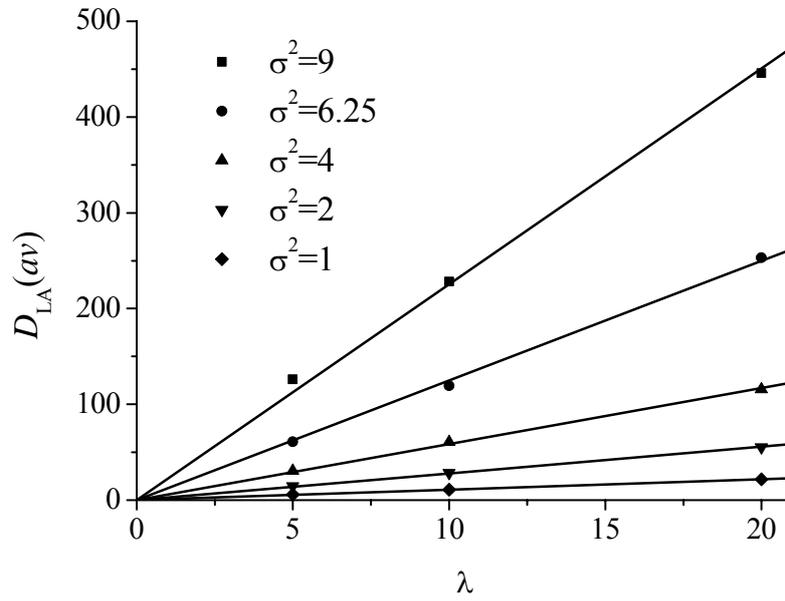


Figure 9 : Asymptotic longitudinal dispersion coefficient as a function of the correlation length for the pure advection case. Lines are linear fit through 0. Same parameters as in figure 6.

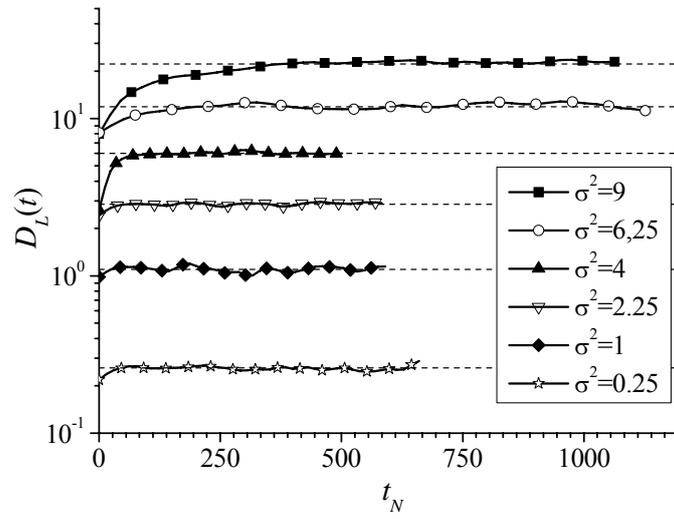


Figure 10: Longitudinal mean dispersion coefficient as a function of  $t_N$  for the pure advection case (time in terms of correlation scales crossed by the plume). Dashed lines mark the asymptotic coefficients. Same parameters as in figure 6.

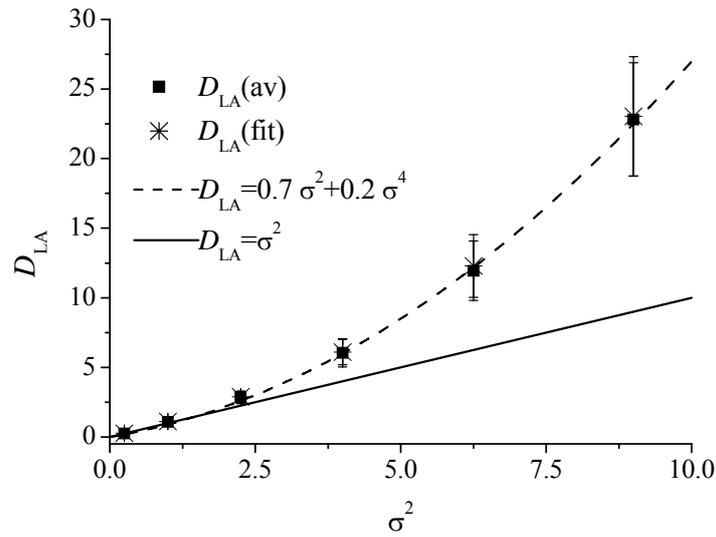


Figure 11: Normalized longitudinal asymptotic effective dispersion coefficient  $D_{LA}$  as a function of the variance of the log conductivity with pure advection. Vertical bars on data points represent the standard deviation on each side of the data point.  $D_{LA}(av)$  and  $D_{LA}(fit)$  are obtained respectively by averaging and fitting by an exponential function. Theoretical predictions [Gelhar, 1993] are represented by the line. The dashed curve stands for  $0.7\sigma^2 + 0.2\sigma^4$ . Same parameters as in figure 6.

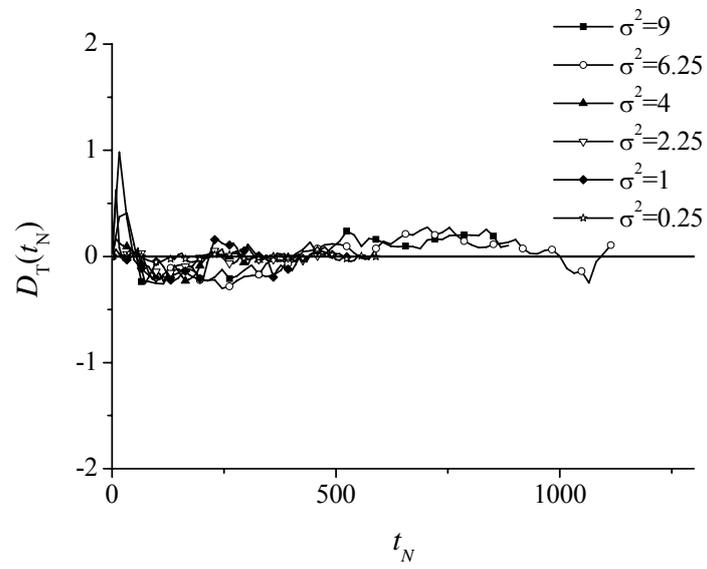


Figure 12 : Normalized transverse dispersion coefficient as a function of the normalized time in the pure advection case. Same parameters as in figure 6.

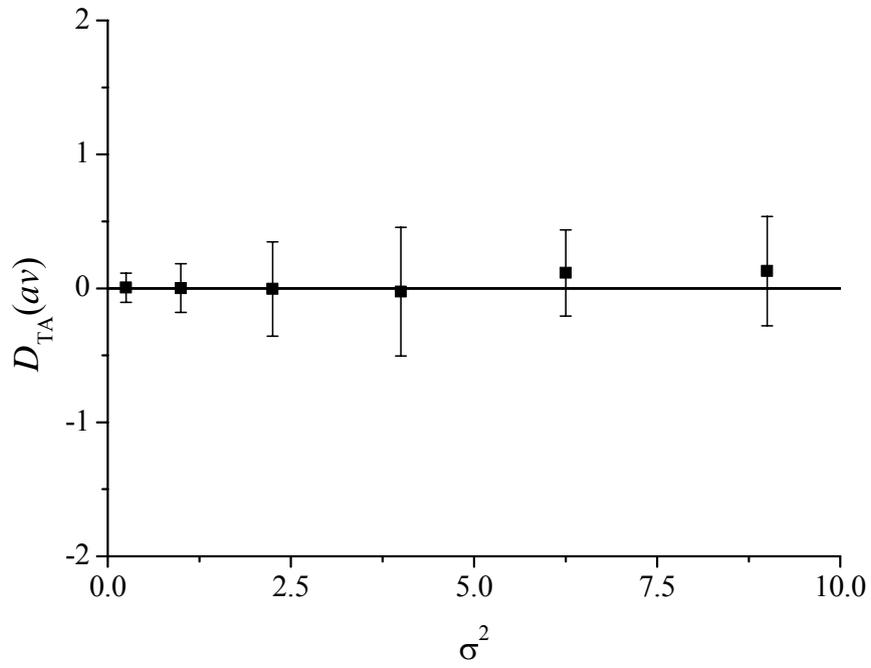


Figure 13 : Normalized transverse asymptotic dispersion coefficient for the pure advection case.

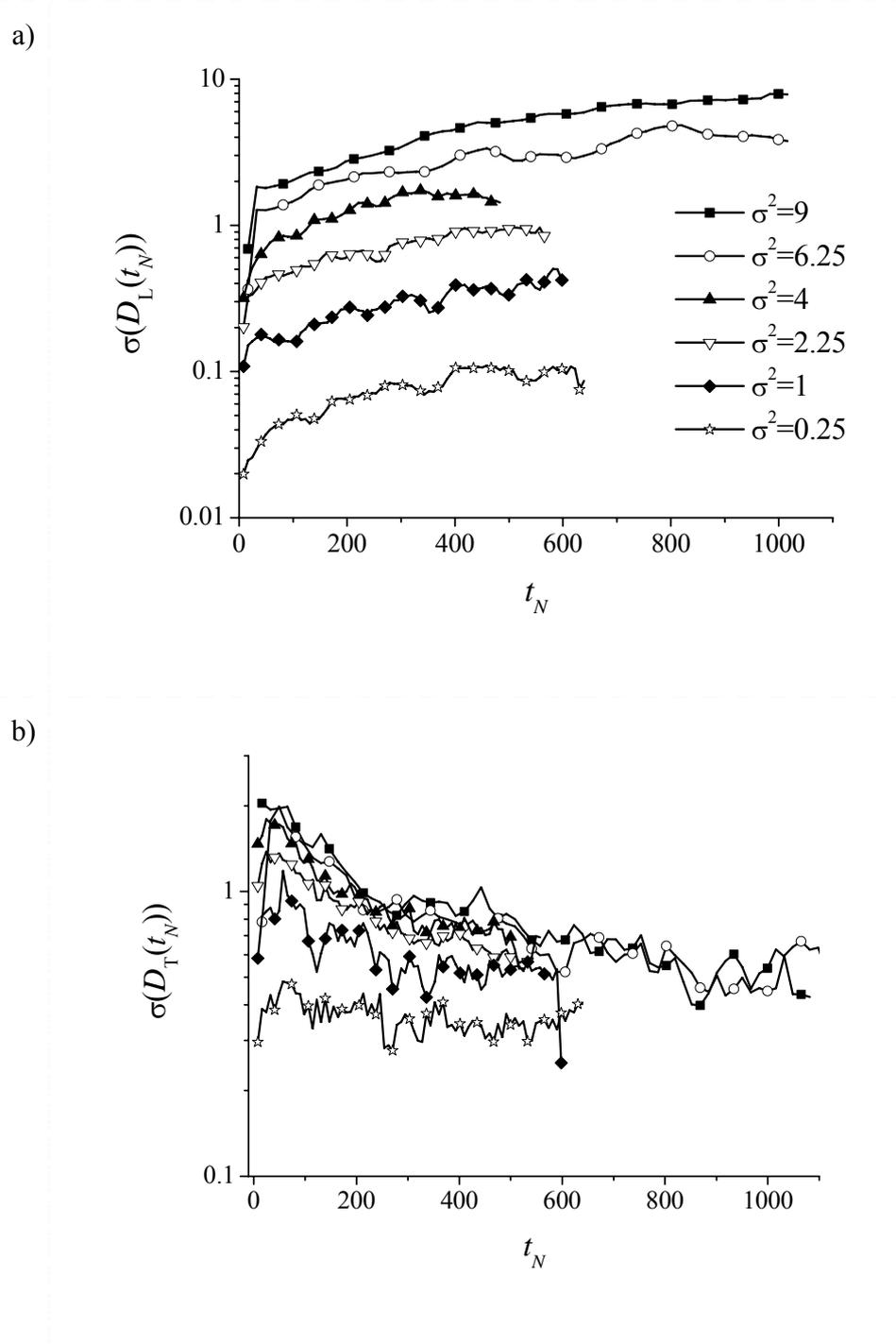


Figure 14 : Standard deviation of a) the longitudinal and b) transverse dispersion coefficients in the pure advection case.

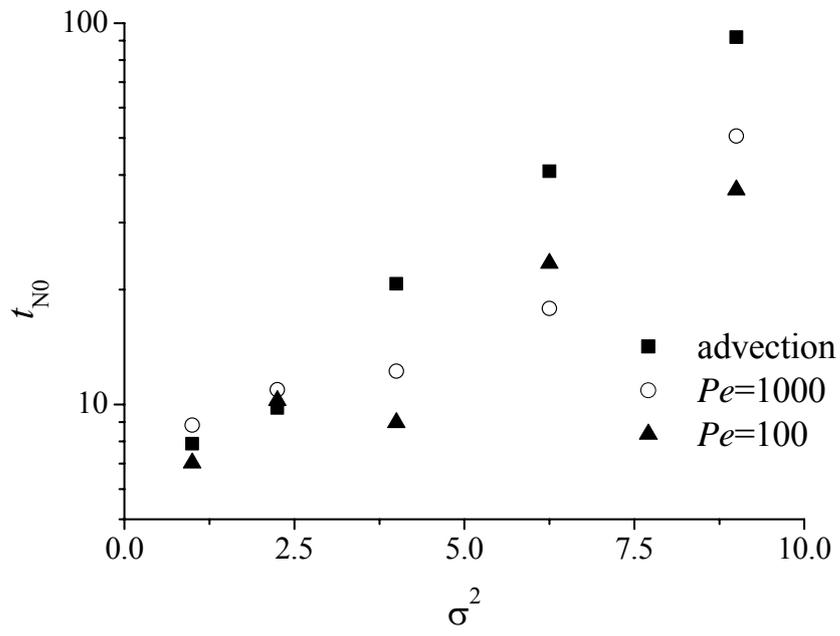


Figure 15 :Characteristic convergence time to the asymptotic regime  $t_{N0}$ .

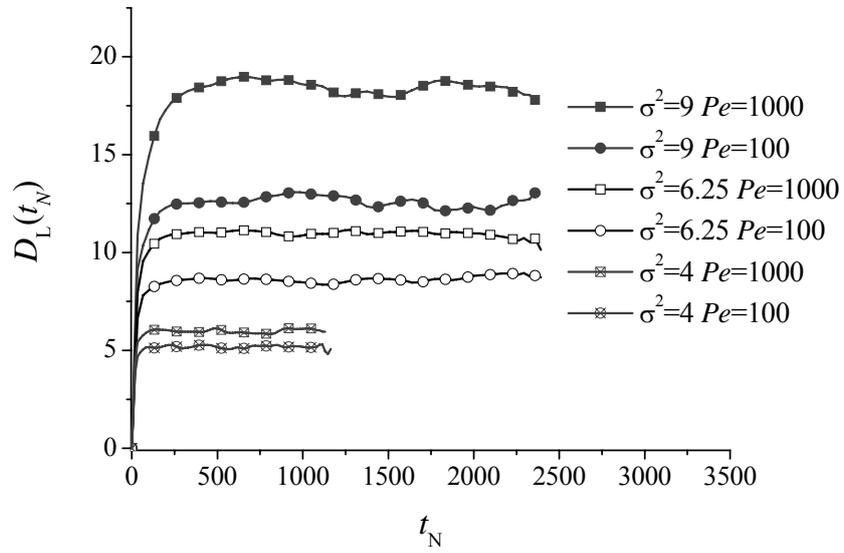


Figure 16: Longitudinal dispersion coefficient as a function of normalized time for  $\sigma^2 \geq 4$ .