Mechanisms, upscaling, and prediction of anomalous dispersion in heterogeneous porous media

Alessandro Comolli, Vivien Hakoun, Marco Dentz

Mechanisms, upscaling and prediction of anomalous dispersion in heterogeneous porous media

Alessandro Comolli, Vivien Hakoun, Marco Dentz

[1]Spanish National Research Council (IDAEA-CSIC), 08034 Barcelona, Spain

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019WR024919

Corresponding author: Marco Dentz, marco.dentz@csic.es

Abstract

We study the upscaling and prediction of large scale solute dispersion in heterogeneous porous media with focus on preasymptotic or anomalous features such as tailing in breakthrough curves and spatial concentration profiles as well as non-linear evolution of the spatial variance of the concentration distribution. Spatial heterogeneity in the hydraulic medium properties is represented in a stochastic modeling approach. Direct numerical Monte Carlo simulations of flow and advective particle motion combined with a Markov model for streamwise particle velocities give insight in the mechanisms of preasymptotic and asymptotic solute transport in terms of the statistical signatures of the medium and flow heterogeneity. Based on the representation of equidistantly sampled particle velocities as a Markov process, we derive an upscaled continuous time random walk approach that can be conditioned on the flow velocities and thus hydraulic conductivity in the injection region. In this modeling framework, we identify the Eulerian velocity distribution, advective tortuosity and the correlation length of particle velocities as the key quantities for large scale transport prediction. Thus, the upscaled model predicts the spatial concentration profiles, their first and second centered moments as well as the breakthrough curves obtained from direct numerical Monte Carlo simulations in spatially heterogeneous conductivity fields. The presented approach allows to relate the medium and flow properties to large scale preasymptotic and asymptotic solute dispersion.

Key Points:

- We derive a continuous time random walk approach for transport upscaling in heterogeneous media
- The approach is predictive through relation of its key parameters with medium and flow properties
- The derived framework links classical stochastic approaches and continuous time random walk models

1 Introduction

The dispersion of dissolved substances in the spatially varying flow through heterogeneous media has been the focus of intense research in the last 4 decades [*Dagan*, 1989; *Gelhar*, 1993; *Rubin*, 2003] due its central importance for the understanding and prediction of large scale solute transport in environmental and industrial applications ranging from groundwater management and remediation [*Domenico and Schwartz*, 1997] to underground storage and energy production [*Poinssot and Geckeis*, 2012; *Niemi et al.*, 2017]. Spatial heterogeneity in the hydraulic medium properties implies spatial fluctuations in the groundwater flow, which leads to macroscopic transport behaviors that are different from the ones observed on the local scale. Spatial variability in hydraulic conductivity is a singular feature for porous media transport because it may vary over orders of magnitude ranging from ~ 1 m/s for coarse sand and gravel to $\sim 10^{-12}$ for clays and granite [*Bear*, 1972; *Sanchez-Vila et al.*, 2006a].

The impact of spatial heterogeneity on large scale solute dispersion has been quantified in a systematic way using stochastic modeling, which derives the large scale transport behaviors as the ensemble average over the behaviors in individual medium realizations that share the same geostatistical characteristics [*Dagan*, 1989; *Gelhar*, 1993]. For moderately heterogeneous porous media, *Gelhar and Axness* [1983] used perturbation theory in the variance of the logarithm of hydraulic conductivity to derive an expression for the longitudinal macrodispersion coefficient in terms of the mean flow velocity, the variance and correlation length of log-hydraulic conductivity. This expression

is a central result for transport upscaling in heterogeneous porous media because it relates macroscopic solute transport to the hydraulic medium properties and flow conditions. It describes solute dispersion at times that are large compared to the characteristic advection time over the correlation scale for moderately heterogeneous media. At preasymptotic times, however, solute dispersion evolves in time from local towards macrodispersion [*Dagan*, 1984]. With increasing heterogeneity, the asymptotic macrodispersion coefficients show a non-linear dependence on the variance of the logarithm of hydraulic conductivity [*Fiori et al.*, 2003; *de Dreuzy et al.*, 2007; *Gotovac et al.*, 2009]. Furthermore, the time scales to reach the asymptotic behavior increase with increasing heterogeneity [*de Dreuzy et al.*, 2007].

In fact, for strong spatial heterogeneity, transport exhibits non-Fickian features such as broad preasymptotic dispersion regimes characterized by non-linear growth of the dispersion coefficients, heavy-tailed breakthrough curves and non-Gaussian spatial concentration distributions [*Adams and Gelhar*, 1992; *Zheng et al.*, 2011; *Haggerty et al.*, 2000; *Levy and Berkowitz*, 2003; *Kang et al.*, 2015; *Haslauer et al.*, 2012]. For many practical situations, dispersion is indeed preasymptotic because the time scales to reach the asymptotic behavior are given by the advection or diffusion times over characteristic heterogeneity length scales, which, depending on the underlying disorder may be of the order of years. Anomalous or preasymptotic anomalous or preasymptotic large scale transport has been modeled by a series of approaches including spatio-temporal non-local advection-dispersion equations, fractional advection-dispersion equations, as well as multirate mass transfer, continuous time random walk (CTRW) and time-domain random walk (TDRW) approaches [*Berkowitz et al.*, 2006; *Frippiat and Holeyman*, 2008; *Neuman and Tartakovsky*, 2008; *Dentz et al.*, 2011; *Noetinger et al.*, 2016]. While stochastic perturbation theory provides a direct link between large scale transport and the medium and flow properties, this is often not the case for non-Fickian transport approaches. In the frameworks of time-domain and continuous time random walks, *Fiori et al.* [2007], *Cvetkovic et al.* [2014], *Edery et al.* [2014], and *Tyukhova et al.* [2016] investigated the relation between the hydraulic conductivity distribution and advective travel times for heterogeneous porous media. *Edery et al.* [2016], *Kang et al.* [2017] and *Hyman et al.* [2019] studied the relation of structural and hydraulic properties to large scale transport for 2 and 3 dimensional discrete fracture networks.

Here we employ a particle based stochastic approach for transport and dispersion upscaling. Stochastic particle approaches have quantified the impact of velocity fluctuations on large scale transport based on stochastic representations of the series of particle velocities either in time [*Dagan*, 1989; *Meyer and Tchelepi*, 2010] or in space [*Berkowitz and Scher*, 1997; *Benke and Painter*, 2003; *Berkowitz et al.*, 2006; *Fiori et al.*, 2007; *Le Borgne et al.*, 2007, 2008; *Gotovac et al.*, 2009; *Dentz et al.*, 2016]. Here, we choose the latter approach, which renders the stochastic particle motion a continuous time or time-domain random walk. This view accounts for the fact that particle velocities vary on the characteristic heterogeneity length scale, while the time scales of variability can be broadly distributed for a broad distribution of flow velocities. In fact, time series of particle velocities in steady porous media flows exhibit intermittent patterns [*de Anna et al.*, 2013; *Kang et al.*, 2014; *Holzner et al.*, 2015; *Morales et al.*, 2017], which reflects the long persistence times in regions of low flow velocity [*Dentz et al.*, 2016]. Using a stochastic modeling approach for the medium heterogeneity, *Hakoun et al.* [2019] provide a thorough statistical analysis of the particle velocities sampled equidistantly along trajectories and their evolution from different initial conditions, this means different initial velocity distributions. In that work, the authors model equidistant velocity series as continuous Markov processes which evolve as a function of streamline length. The velocity processes are parameterized by the Eulerian velocity distribution, a flow attribute, and the velocity correlation scale.

In this paper, we study the upscaling of solute dispersion from the Darcy to the regional scale using a continuous time random walk approach that is informed by the velocity models presented in *Hakoun et al.* [2019]. The aim is to derive a framework that allows to systematically link the statistical medium and flow properties to large scale preasymptotic transport behaviors in the spirit of the perturbation theory result for the longitudinal macrodispersion coefficient. To this end, we consider spatial and temporal transport signatures in terms of the average concentration profiles projected on the mean flow direction, the evolution of the mean and variance of the concentration distribution, as well as breakthrough curves. We focus also on the impact of the injection conditions and heterogeneity in the injection domain on large scale transport [*Le Borgne et al.*, 2007; *Hyman et al.*, 2015; *Dagan*, 2017; *Kang et al.*, 2017; *Morales et al.*, 2017; *Puyguiraud et al.*, 2019] and its prediction in an upscaled transport approach. The predictions of the derived large scale transport model are compared to direct numerical Monte Carlo simulation of flow and particle motion in 2-dimensional heterogeneous hydraulic conductivity fields, which also provides a further test for the robustness of the velocity models presented in *Hakoun et al.* [2019].

The paper is organized as follows. The next section introduces in the stochastic flow and transport problem in heterogeneous porous media and explains the numerical setup. Section 3 presents the upscaling methodology and the stochastic particle model. Section 4 discusses the transport behaviors determined by the numerical Monte Carlo simulations and their prediction in terms of the stochastic particle model.

2 Basics

In the following, we present the equations governing flow and advective transport in heterogeneous porous media, the stochastic modeling approach as well as some basic relations between Eulerian and Lagrangian velocity statistics.

2.1 Flow

At the Darcy scale, porous media flow is ruled by the Darcy equation [*Bear*, 1972]

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x})\,, \tag{1}$$

where $\mathbf{x} = (x_1, x_2)^\top$ denotes the position vector, $K(\mathbf{x})$ is the hydraulic conductivity and $h(\mathbf{x})$ is the hydraulic head. In the following, we consider incompressible flow, which implies $\nabla \cdot \mathbf{q}(\mathbf{x}) = 0$ and thus the groundwater flow equation

$$\nabla K(\mathbf{x}) \cdot \nabla h(\mathbf{x}) + K(\mathbf{x})\nabla^2 h(\mathbf{x}) = 0. \tag{2}$$

Spatial variability in hydraulic conductivity is quantified by a stochastic approach, which models $K(\mathbf{x})$ as a stationary and ergodic random field [*Rubin*, 2003]. Hydraulic conductivity here is obtained through the map $K(\mathbf{x}) = F\{Y(\mathbf{x})\}$ from the stationary and ergodic multi-Gaussian random field $Y(\mathbf{x})$. The map is given by

$$F(y) = P_K^{-1}[\Phi(y)], \tag{3}$$

where $P_K(k)$ is the cumulative distribution function (CDF) of hydraulic conductivity, $P_K^{-1}(u)$ its inverse, and $\Phi(y)$ is CDF of the Gaussian distribution. The Gaussian random field $Y(\mathbf{x})$ is characterized by zero mean $\langle Y(\mathbf{x})\rangle = 0$ and the exponential covariance function

$$\langle Y(\mathbf{x})Y(\mathbf{x}')\rangle = \sigma_Y^2 \exp(-|\mathbf{x} - \mathbf{x}'|/\lambda) \tag{4}$$

where λ is the correlation length and σ_Y^2 the variance of $Y(\mathbf{x})$. The ensemble average is denoted by angular brackets. We consider two distributions of point values for the hy-

draulic conductivity $K(\mathbf{x})$, namely the lognormal and the truncated-Gamma distributions

$$p_K(k) = \frac{\exp\left\{-\frac{[\ln(k)-\mu]^2}{2\sigma_f^2}\right\}}{k\sqrt{2\pi\sigma_f^2}} \tag{5a}$$

$$p_K(k) = \frac{N(\alpha, k_c, k_0)}{k_c\Gamma(\alpha)}\left(\frac{k}{k_c}\right)^{\alpha-1}\exp\left(-\frac{k}{k_c}-\frac{k_0}{k}\right), \tag{5b}$$

where μ and σ_f^2 are the mean and variance of $f(\mathbf{x}) = \ln[K(\mathbf{x})]$, k_0 and k_c are the lower and upper cut-offs in the truncated Gamma distribution, α is the shape parameter and N the normalization factor which is given numerically. The choice of the lognormal distribution is motivated by the fact that it is in a certain sense a standard disorder model for large scale heterogeneous porous media and was applied for the field scale medium characterization [*Rehfeldt et al.*, 1992; *Gelhar*, 1993] as well as in many modeling works [*Gómez-Hernández and Gorelick*, 1989; *Jankovic et al.*, 2003; *de Dreuzy et al.*, 2007; *Gotovac et al.*, 2009]. Some authors, however, have challenged the Gaussianity of $\ln K$ for heterogeneous sedimentary media [*Painter*, 1996; *Sanchez-Vila et al.*, 2006b]. For example, *Haslauer et al.* [2012] have shown that the Borden aquifer heterogeneity is best described by a distribution characterized by power-law tailing towards low K values and exponential decay for large hydraulic conductivities. The truncated Gamma distribution that we use in this work is characterized by similar features. We refer to the truncated Gamma distribution in the following as Gamma distribution. The lognormal $K(\mathbf{x})$ is obtained from $Y(\mathbf{x})$ through the exponential map

$$K(\mathbf{x}) = \exp\left[\mu + Y(\mathbf{x})\right]. \tag{6}$$

For the Gamma distributed conductivity field, there is no analytical map $F(k)$ because there is no closed form expression for $P_K^{-1}(u)$. Thus, $K(\mathbf{x})$ is obtained from $Y(\mathbf{x})$ numerically through inverse sampling. The degree of heterogeneity of $K(\mathbf{x})$ is measured by the variance σ_f^2 of $f(\mathbf{x})$. For the lognormal K-field, $\sigma_f^2 = \sigma_Y^2$. For the Gamma K-fields, σ_f^2 increases with decreasing α for a fixed mean $\langle K(\mathbf{x})\rangle$.

The stochasticity of $K(\mathbf{x})$ is mapped onto the Eulerian velocity field $\mathbf{q}(\mathbf{x})$ through the Darcy equation (1). The mean hydraulic gradient $\langle\nabla h(\mathbf{x})\rangle$ is aligned with the 1-direction of the coordinate system. Thus, the mean flow velocity is $\langle\mathbf{q}(\mathbf{x})\rangle = \mathbf{e}_1\langle q\rangle$. It defines together with the correlation length λ the characteristic time scale $\tau_c = \lambda/\langle q\rangle$. We denote the magnitude of $\mathbf{q}(\mathbf{x})$ by $v_e(\mathbf{x}) = |\mathbf{q}(\mathbf{x})|$, where $|\cdot|$ denotes the Euclidean norm. The velocity statistics are characterized here by the probability density function (PDF) of $v_e(\mathbf{x})$, which is obtained by spatial sampling as

$$p_e(v) = \lim_{V\to\infty}\frac{1}{V}\int_\Omega d\mathbf{x}\,\delta[v - v_e(\mathbf{x})], \tag{7}$$

with Ω is the sampling domain, V its volume. Due to ergodicity, spatial sampling of $v_e(\mathbf{x})$ is equivalent to ensemble sampling and thus $p_e(v) = \langle\delta[v - v_e(\mathbf{x})]\rangle$.

The numerical generation of the hydraulic conductivity fields $K(\mathbf{x})$, solution of the flow equation (2) and the setup of the numerical flow simulations are detailed in Appendix A: . Figures 1a and b illustrate realizations of Gamma-distributed and lognormal K fields, Figures 1c and d the spatial distribution of the corresponding velocity magnitude $v_e(\mathbf{x})$ and the particle distributions evolving from a line source perpendicular to the mean flow direction.

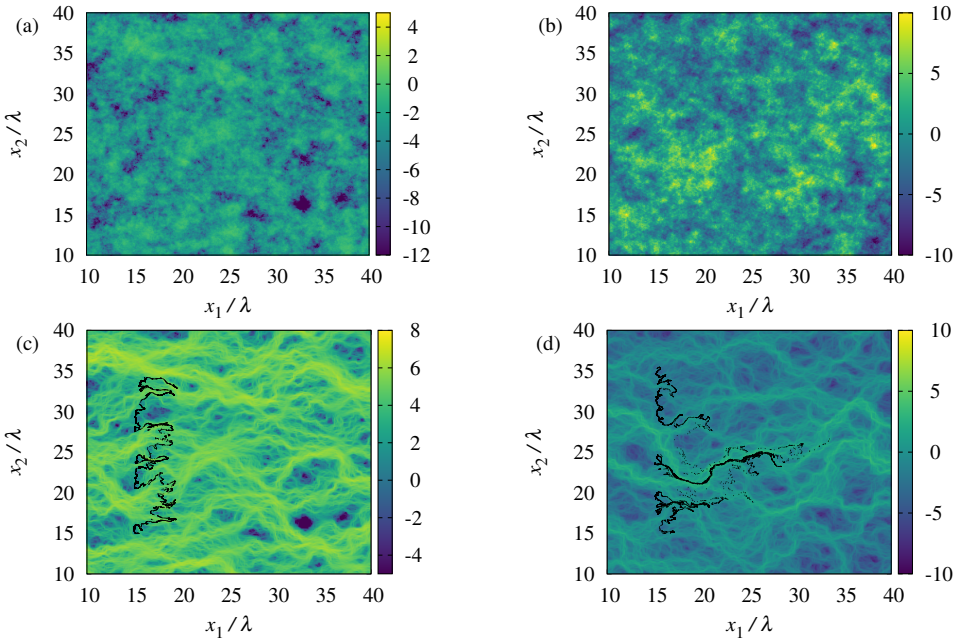

Figure 1. Map of $f(\mathbf{x}) = \ln[K(\mathbf{x})]$ for (a) Gamma-distributed $K(\mathbf{x})$ with $\alpha = 1/2$, $k_c = 5$ m/s and $k_0 = 10^{-11}$ m/s, which gives $\sigma_f^2 = 4.5$, and (b) lognormal $K(\mathbf{x})$ with $\sigma_f^2 = 7$. Map of $\ln(v_e(\mathbf{x})/\langle v_e \rangle)$ for (c) Gamma and (d) lognormal $K(\mathbf{x})$. The black dots denotes the distribution of particles after time $t = 2.5\tau_c$ that are injected at time $t = 0$ along a line located at $x_1 = 15\lambda$.

2.2 Transport

We consider purely advective transport. Thus, the motion of a particle that starts at $\mathbf{x}(t = 0, \mathbf{a}) = \mathbf{a}$ is governed by the advection equation

$$\frac{d\mathbf{x}(t, \mathbf{a})}{dt} = \mathbf{v}_t(t, \mathbf{a}), \tag{8}$$

where $\mathbf{v}_t(t, \mathbf{a}) = \mathbf{q}[\mathbf{x}(t, \mathbf{a})]/\phi$ is the Lagrangian particle velocity and ϕ is porosity, which we will assume constant. In particular, we set $\phi = 1$, which is equivalent to rescaling time. We refer to $\mathbf{v}_t(t, \mathbf{a})$ as t-Lagrangian velocity because it is measured as a function of time. The distance $s(t, \mathbf{a})$ that a particle travels along a streamline satisfies

$$\frac{ds(t, \mathbf{a})}{dt} = v_t(t, \mathbf{a}), \tag{9}$$

where $v_t(t, \mathbf{a}) = |\mathbf{q}[\mathbf{x}(t, \mathbf{a})]|$ is the t-Lagrangian velocity magnitude. We can express the equation of motion (8) as a function of distance s along the streamline by using the map (9) from $t \to s$, which gives

$$\frac{d\mathbf{x}(s, \mathbf{a})}{ds} = \frac{\mathbf{v}_s(s, \mathbf{a})}{v_s(s, \mathbf{a})} \qquad \frac{dt(s, \mathbf{a})}{ds} = \frac{1}{v_s(s, \mathbf{a})}, \tag{10}$$

where $\mathbf{v}_s(s, \mathbf{a}) = \mathbf{q}[\mathbf{x}(s, \mathbf{a})]$. Its magnitude is denoted by $v_s(s, \mathbf{a}) = |\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|$, to which we refer in the following as s-Lagrangian velocity. The distribution of initial particle positions is denoted by $\rho(\mathbf{a})$. Numerical particle tracking simulations to solve for the direct problem are detailed in Appendix A: . Figure 1b shows the distribution of particles that were initially uniformly distributed along a line. The tortuous shape is due to spatial velocity heterogeneity. Furthermore, particles are retained in regions of low flow velocities.

Under Lagrangian ergodicity and stationarity the velocity statistics sampled along streamlines and between an ensemble of particles are equivalent. The stationary s-Lagrangian velocity PDF $p_s(v)$ is related to the Eulerian velocity PDF $p_e(v)$ by flux-weighting [*Dentz et al.*, 2016; *Comolli and Dentz*, 2017]

$$p_s(v) = \frac{v p_e(v)}{\langle v_e \rangle} \tag{11}$$

In general the Lagrangian velocity statistics evolve with time t or distance s along streamline if the initial velocity distribution is not stationary. The next section provides a description of two Markov velocity models to account for these evolutions and the approach to quantify its impact on particle transport.

We consider here initial particle distributions along a line of length $L \gg \lambda$ characterized by different weighting. Specifically, we take $L = 100\lambda$ for lognormal and $L = 270\lambda$ for the Gamma fields. For a uniform initial particle distribution, $\rho(\mathbf{a}) = L^{-1}\delta(a_1)$ for $0 \leq a_2 \leq L$, the initial velocities $v_0(\mathbf{a}) = v_s(s = 0, \mathbf{a}) = v_t(t = 0, \mathbf{a})$ are distributed according to $p_0(v) = p_e(v)$. This represents the stationary initial condition for the t-Lagrangian velocity statistics [*Dentz et al.*, 2016]. For a flux-weighted initial distribution,

$$\rho(a) = \frac{v_e(\mathbf{a})}{\langle v_e \rangle}\delta(a_1), \tag{12}$$

the initial velocity distribution $p_0(v) = p_s(v)$ is given by the flux-weighted Eulerian velocity PDF (11). This means, the s-Lagrangian velocity PDF is stationary for the initial distribution (12). Furthermore, we consider particle injections into velocity intervals $v_\ell < v < v_u$, this means, we condition the injection points $\mathbf{a}$ in the line on the velocity magnitude as

$$\rho(\mathbf{a}) = \frac{\mathbb{I}(v_\ell < v_e(\mathbf{a}) < v_u)\delta(a_1)}{L\int_{v_\ell}^{v_u} dv p_e(v)} \tag{13}$$

where $\mathbb{I}(\cdot)$ denotes an indicator function which is 1 if the argument is true and 0 else. This condition allows to investigate the impact of the initial particle velocities on average transport in a systematic way. Note that for an unconditional point injection, $\rho(\mathbf{a}) = \delta(\mathbf{a}-\mathbf{a}_0)$, the ensemble average erases the dependence on the injection conditions in a single realization and is equivalent to the average over the initial particle positions for a uniform injection in a single medium realization. For the flux-weighted initial distribution and in general for initial distributions that are conditioned on velocity this is different. The average preasymptotic transport properties in general depend on the initial particle distribution [*Hyman et al.*, 2015; *Dentz et al.*, 2016; *Kang et al.*, 2017; *Fiori et al.*, 2017; *Zech et al.*, 2018].

We focus here on particle transport along the mean flow direction. The distribution of particles, here equivalent to the concentration distribution, is obtained through averaging over the particles in single realizations and over the ensemble of disorder realizations as

$$c(x_1, t) = \left\langle \int d\mathbf{a}\rho(\mathbf{a})\delta[x_1 - x_1(t|\mathbf{a})] \right\rangle. \tag{14}$$

The displacement mean and variance are given by

$$\mu(t) = \left\langle \int d\mathbf{a}\rho(\mathbf{a})x_1(t, \mathbf{a}) \right\rangle, \qquad \kappa(t) = \left\langle \int d\mathbf{a}\rho(\mathbf{a}) \left[x_1(t, \mathbf{a}) - \mu(t)\right]^2 \right\rangle. \tag{15}$$

The longitudinal dispersion coefficient $\mathcal{D}_L(t)$ describes the temporal rate of change of the displacement variance

$$D_L(t) = \frac{1}{2}\frac{d\kappa(t)}{dt}. \tag{16}$$

Temporal aspects of longitudinal solute transport are characterized in terms of the breakthrough curves at a control plane located at a longitudinal position x_1. The breakthrough time is defined by

$$\tau(x_1, \mathbf{a}) = \min[t | x_1(t, \mathbf{a}) \geq x_1]. \tag{17}$$

The breakthrough curve $f(t, x_1)$ is obtained by sampling arrival times in single realizations and averaging over the ensemble as

$$f(t, x_1) = \left\langle \int d\mathbf{a} \rho(\mathbf{a}) \delta[t - \tau(x_1, \mathbf{a})] \right\rangle. \tag{18}$$

In the following section, we discuss the upscaling methodology to quantify the behaviors of particle distributions and breakthrough curves and their dependence on the injection conditions.

3 Upscaling methodology

The upscaling methodology is based on the representation of the s-Lagrangian velocity magnitude as a Markov process [*Dentz et al.*, 2016; *Hakoun et al.*, 2019]. This means, the series of particle velocities $\{v_s(s, \mathbf{a})\}$ is modeled as a stationary stochastic process $\{v_s(s)\}$ which forms a Markov chain. One realization of $v_s(s)$ is considered the velocity series of a single particle. The average behavior is obtained by ensemble averaging over all realizations of $v_s(s)$. The Markov chain is fully characterized by the transition probability $r(v, s|v')$ to make a transition from v' to v after a streamwise distance s, and the steady-state s-Lagrangian velocity PDF $p_s(v)$, which is obtained from the Eulerian velocity PDF according to (11). In the following, we first discuss the time-domain random walk (TDRW) transport framework, which follows naturally from modeling $v_s(s)$ as a Markov chain. Then, we summarize two analytical models to quantify the stochastic evolution of $v_s(s)$.

3.1 Stochastic particle model

In the framework of a Markov model for the s-Lagrangian velocities, particle motion is described in terms of the distance s along the streamline and the particle time $t(s)$. Here, we focus on particle motion along the mean flow direction, which is aligned with the 1-direction of the coordinate system. The motion along the streamline is projected onto the 1-direction by using the advective tortuosity χ. Advective tortuosity is a measure for the ratio between the displacement s along the trajectory and the displacement $x_1(s)$ in mean flow direction in the limit of $s \to \infty$. In Appendix B: we derive

$$\chi = \frac{\langle v_e \rangle}{\langle q \rangle}, \tag{19}$$

where $\langle v_e \rangle = \langle |\mathbf{q}(\mathbf{x})| \rangle$ and $\langle q \rangle = \langle q_1(\mathbf{x}) \rangle$. Thus, for $s \gg \lambda$, longitudinal particle motion is described in terms of the position $x_1(s)$ and particle time $t(s)$ as

$$\frac{dx_1(s)}{ds} = \chi^{-1}, \qquad \frac{dt(s)}{ds} = \frac{1}{v_s(s)}. \tag{20}$$

In practice, this provides a good approximation at distances of the order of a few λ. The particle velocity $v_s(s)$ is sampled from the s-Lagrangian velocity PDF $p_s(v, s)$, which evolves

from the initial distribution $p_0(v)$ as

$$\hat{p}_s(v,s) = \int_0^\infty dv' r(v,s|v') p_0(v'). \tag{21}$$

This model belongs to the class of continuous time or time-domain random walks [*Painter and Cvetkovic*, 2005; *Berkowitz et al.*, 2006; *Noetinger et al.*, 2016] because it models particle motion through equidistant spatial steps of length ds/χ with a random transition time $ds/v_s(s)$, which is determined kinematically by the particle velocity $v_s(s)$. Note that the inverse velocity $\gamma(s) = 1/v_s(s)$ is also termed slowness [*Gotovac et al.*, 2009]. The steady state distribution $\psi_\gamma(\gamma)$ of slowness can be expressed in terms of $p_s(v)$ as

$$\psi_\gamma(\gamma) = \gamma^{-2} p_s(1/\gamma). \tag{22}$$

In this framework, the initial particle distribution $\rho(\mathbf{a})$ is translated to a distribution of initial particle velocities $p_0(v)$. For the uniform initial condition and $L \gg \lambda$, $p_0(v) = p_e(v)$ due to the ergodicity of the velocity field. For a the flux-weighted injection, it is $p_0(v) = p_s(v)$. For the injection mode conditioned on velocities in a certain velocity range, $p_0(v)$ is

$$p_0(v) = \frac{p_e(v)\mathbb{I}(v_\ell < v < v_u)}{\int_{v_\ell}^{v_u} dv p_e(v)}. \tag{23}$$

Furthermore, the particle position and velocity at a time t are given by $x_1(t) = x_1[s(t)]$ and $v_t(t) = v_s[s(t)]$, where $s(t) = \min[s|t(s) \le t]$. The particle distribution along the mean flow direction is given by

$$c(x_1,t) = \langle \delta\left[x_1 - s(t)/\chi\right]\rangle. \tag{24}$$

The angular brackets here denote the average over all particles. It is equivalent to the ensemble average. The displacement mean and variance are given by

$$\mu(t) = \chi^{-1}\langle s(t)\rangle, \qquad \kappa(t) = \chi^{-2}\langle [s(t) - \langle s(t)\rangle]^2\rangle. \tag{25}$$

The solute breakthrough curve is given in this framework in terms of the particle time $t(s)$ as

$$f(t,x_1) = \langle \delta\left[t - t(x_1\chi)\right]\rangle. \tag{26}$$

The upscaled transport dynamics can be expressed in terms of the joint distribution $p(x_1,v,t) = \langle \delta[x_1 - x_1(t)]\delta[v - v(t)]\rangle$ of particle position and velocity. This is similar in spirit to the quantification of Brownian motion, first in terms of a stochastic differential equation for the particle velocity, and then in terms of a Kramers equation [*Risken*, 1996]. In Appendix C: , we derive for $p(x_1,v,t)$ the integro-differential equation

$$\frac{\partial p(x_1,v,t)}{\partial t} + \frac{v}{\chi}\frac{\partial p(x_1,v,t)}{\partial x_1} = -v p(x_1,v,t) + \int_0^\infty dv' \frac{v' r(v,\Delta s|v')}{\Delta s} p(x_1,v',t). \tag{27}$$

The left side quantifies the advective translation of the distribution by the local velocity, the right side the transitions between velocities. This formulation of large scale transport in terms of a Boltzmann equation is further studied elsewhere.

The Markov chain $\{v_s(s)\}$ converges toward its steady state on a characteristic length scale ℓ_c, the correlation length of $v_s(s)$. For lag distances $\Delta s > \ell_c$, subsequent velocities may be assumed independent. Thus, coarse-graining of the governing equations (20)

on the correlation scale ℓ_c gives

$$x_1(s_{n+1}) = x_1(s_n) + \frac{\ell_c}{\chi}, \qquad t(s_{n+1}) = t(s_n) + \frac{\ell_c}{v_n}, \tag{28}$$

where the v_n are independent identical distributed random variables distributed as $p_s(v)$ for $n > 0$ and as $p_0(v)$ for $n = 0$. The time increment $\tau_n = \ell_c/v_n$ are distributed according to

$$\psi(t) = \frac{\ell_c}{t^2} p_s(\ell_c/t). \tag{29}$$

for $n > 0$. For $n = 0$ we define $\psi_0(t)$ analogously. The asymptotic transport behavior may be described by (28), which constitutes a CTRW or TDRW with independent velocity, or time increments. Under steady conditions, this means for $\psi_0(t) = \psi(t)$, the CTRW framework provides predictions for the asymptotic dispersion behavior based on the behavior of $\psi(t)$ [*Berkowitz et al.*, 2006]. For $\psi(t) \propto t^{-1-\beta}$ and $1 < \beta < 2$, the asymptotic behaviors of the displacement mean $\mu_x(t)$ and variance $\kappa_x(t)$ the scalings $\mu_x(t) \propto t$ and $\kappa_x(t) \propto t^{3-\beta}$, the breakthrough curves scale as $f(t, x) \propto t^{-1-\beta}$. For the CTRW model presented in this work, it is crucial to have a correct characterization of the velocity PDF. This is provided by two velocity Markov models, which were proposed in previous studies and are summarized in the next section.

3.2 Velocity transition model

In the following, we present two Markov models for the evolution of the s-Lagrangian velocities, the Bernoulli and the Ornstein-Uhlenbeck processes [*Dentz et al.*, 2016; *Morales et al.*, 2017; *Kang et al.*, 2017]. Both processes reproduce the basic features of evolution of the s-Lagrangian velocity PDF from a non-stationary initial condition toward the steady state. While the Ornstein-Uhlenbeck model captures different aspects of intermittency such as the distribution of velocity increments [*Morales et al.*, 2017] and the evolution of Lagrangian velocity statistics [*Hakoun et al.*, 2019], the Bernoulli model is appealing due to its simplicity. We compare the performance of both models in the prediction of Darcy scale transport. Both models are stationary and have as the only parameter the correlation scale ℓ_c. For completeness, in the following we summarize the Bernoulli and Ornstein-Uhlenbeck velocity models as given in [*Hakoun et al.*, 2019].

3.2.1 Bernoulli process

The Bernoulli process models the s-Lagrangian velocity series by the stochastic relaxation

$$v_s(s + \Delta s) = \xi(s) v(s) + [1 - \xi(s)] \nu(s), \tag{30}$$

where the $\nu(s)$ are distributed according to $p_s(v)$ and the $\xi(s)$ are independent identically distributed Bernoulli variables, which take the value 1 with probability $\exp(-\Delta s/\ell_c)$ and 0 with probability $1-\exp(-\Delta s/\ell_c)$. This means, after a distance Δs, a particle remains at the same velocity with probability $\exp(-\Delta s/\ell_c)$ and changes to a new random velocity sampled from $p_s(v)$ with probability $1-\exp(-\Delta s/\ell_c)$. The initial velocities are sampled from $p_0(v)$. The transition probability is given by [*Dentz et al.*, 2016]

$$r(v, s|v') = \exp(-s/\ell_c)\delta(v - v') + [1 - \exp(-s/\ell_c)] p_s(v). \tag{31}$$

For the numerical calculations, we choose $\Delta s = \ell_c/10$. We refer to the stochastic particle model combined with the Bernoulli velocity process in the following as *btdrw*.

3.2.2 Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process models the evolution of the normal scores $w(s)$ corresponding to the $v_s(s)$. They are defined by the transformation

$$w(s) = \Phi^{-1}\{P_s[v_s(s)]\}, \qquad v_s(s) = P_s^{-1}\{\Phi[w(s)]\}, \tag{32}$$

where $P_s(v)$ is the CDF of $p_s(v)$ and $\Phi(w)$ the CDF of the unit Gaussian distribution. The $w(s)$ evolve according to the Langevin equation [*Gardiner*, 1986]

$$\frac{dw(s)}{ds} = -\ell_c^{-1} w(s) + \sqrt{2\ell_c^{-1}}\eta(s), \tag{33}$$

where $\eta(s)$ is a Gaussian white noise characterized by 0 mean and correlation $\langle \eta(s)\eta(s')\rangle = \delta(s-s')$. The transition probability $r_w(w, s|w')$ of $w(s)$ is given by the conditional Gaussian PDF

$$r_w(w, s|w') = \frac{\exp\left(-\frac{[w-w'\exp(-s/\ell_c)]^2}{2[1-\exp(-2s/\ell_c)]}\right)}{\sqrt{2\pi\left[1-\exp(-2s/\ell_c)\right]}}. \tag{34}$$

The numerical implementation of (33) is based on the Euler scheme

$$w(s+\Delta s) = w(s)\left(1-\ell_c^{-1}\Delta s\right) + \sqrt{2\ell_c^{-1}\Delta s}\zeta(s), \tag{35}$$

where $\zeta(s)$ is a Gaussian random variable of zero mean and unit variance. For the numerical simulations reported in the following, we set $\Delta s = \ell_c/10^2$ for the calculation of the spatial profiles and breakthrough curves and $\Delta s = \ell_c/10$ for the calculation of the spatial moments. The stochastic particle model combined with the Ornstein-Uhlenbeck velocity process is in the following referred to as *outdrw*.

3.3 Discussion of model parameters

Some remarks on the determination of the key parameters are in order. The velocity correlation distance ℓ_c is of the order of the correlation length λ of the multi-Gaussian random field, from which the lognormal and Gamma-distributed hydraulic conductivity fields are derived. It varies only moderatley with increasing heterogeneity. Analytical expressions for ℓ_c are only known from perturbation theory for weakly heterogeneous fields [*Cvetkovic et al.*, 1996], who find $\ell_c = 8\lambda/3$. For strong heterogeneity, we currently rely on empirical expressions [*Cvetkovic et al.*, 1996; *Hakoun et al.*, 2019]. Thus, more insight is needed on the dependence of ℓ_c on the medium structure and hydraulic conditions. Also, the approach relies on the knowledge of the Eulerian velocity statistics and tortuosity, both flow attributes. Thus the presented approach is predictive in the sense that it allows for transport predictions based on the estimation of a transport-independent quantities, in the spirit of the perturbation theory expression for the longitudinal macrodispersion coefficient [*Gelhar and Axness*, 1983]. This is an important step towards a predictive approach that links medium and flow properties to transport. A direct link between hydraulic conductivity and flow velocity can be established for small values following exact analytical results for isolated inclusions [*Fiori et al.*, 2007; *Cvetkovic et al.*, 2014; *Tyukhova et al.*, 2016]. A quantitative link between the full velocity and conductivity spectra, however, is still missing. Note also that the stochastic particle model considers particle motion along streamlines, which is projected here on the mean flow direction in terms of the average advective tortuosity. Using this average tortuosity leads to accurate transport predictions at intermediate and asymptotic times, but may underestimate solute spreading at early times. Studying the distribution of tortuosity may give further insight in the impact of medium geometry and hydraulic conditions on large scale transport. Furthermore, the model predictions are of course affected by uncertainty in the model parameters such as Eulerian velocity statistics and correlation length, which however is the subject of a different study.

4 Transport behavior

We study here the transport behavior in the heterogeneous conductivity fields described in Section 2.1 and its upscaling using the methodology detailed in Section 3. The average transport behavior is obtained numerically as an ensemble average over 10^2 realizations for each random field. We consider lognormally distributed $K(\mathbf{x})$ characterized by σ_f^2 from 0.1 to 7, and Gamma-distributed fields characterized by $\alpha = 1/2$ with $\sigma_f^2 = 4.5$. The numerical Monte Carlo simulations use 10^2 realizations and between 10^4 and 10^7 particles per realization. The detailed numerical setup is given in Appendix A: . The setup of the stochastic particle model is discussed in Appendix D: .

4.1 Anomalous Dispersion

We study here the temporal evolution of the mean displacement and displacement variance defined in (15) for the lognormal and Gamma fields by direct numerical simulations and the stochastic particle model, and discuss the expected early and long time behaviors as well as estimates for the asymptotic longitudinal macrodispersion coefficient. For the simulation times under consideration, there is no quantitative difference between the predictions of the *outdrw* and the *btdrw* models. Thus, we display only the results of the stochastic particle model based on the Ornstein-Uhlenbeck velocity process.

4.1.1 Mean displacement

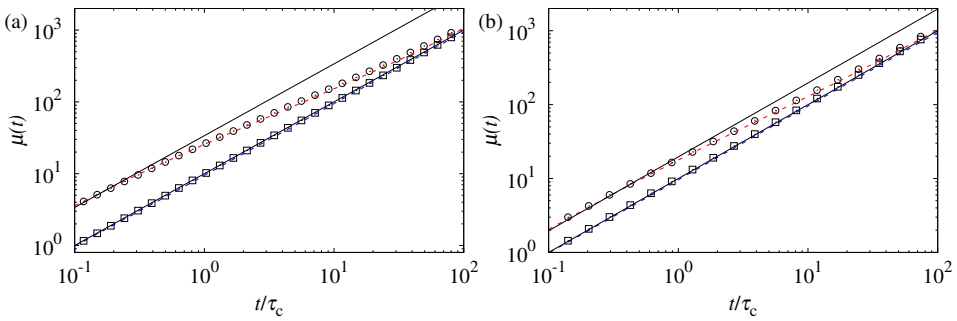

Figure 2. Mean displacements from Monte Carlo simulations in 10^2 realizations each with 10^4 particles for (squares) uniform and (circles) flux-weighted injection in (a) Gamma field with $\alpha = 1/2$ and (b) lognormal conductivity field with $\sigma_f^2 = 7$. The simulations based on the *outdrw* model use 10^4 particles and are marked by (flux-weighted) red dashed and (uniform) blue dashed lines. The solid black lines denote the linear early and late time behaviors.

At times smaller than the advection time scale $\tau_c = \lambda/\langle q \rangle$, the particle velocity is constant and equal to the initial velocity such that the mean displacement is

$$\mu(t) = \int_{\Omega_0} d\mathbf{a} \rho(\mathbf{a}) q_1(\mathbf{a}) t. \tag{36}$$

The stochastic particle model (20) predicts the early time behavior

$$\mu(t) = \chi^{-1} \int_0^\infty dv p_0(v) v t. \tag{37}$$

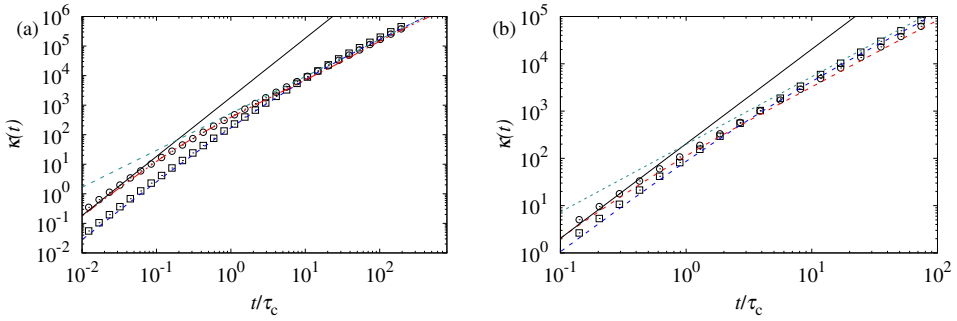

Figure 3. Displacement variances from Monte Carlo simulations in 10^2 realizations each with 10^4 particles for (squares) uniform and (circles) flux-weighted injection in (a) Gamma field with $\alpha = 1/2$ and (b) lognormal conductivity field with $\sigma_f^2 = 7$. The simulations based on the *outdrw* model use 10^4 particles and are marked by (flux-weighted) red dashed and (uniform) blue dashed lines. The solid black lines denote the ballistic early time behavior, the dashed green lines the (pre-) asymptotic behaviors.

Under stationary condition, this means for a uniform initial distribution, or equivalently for $p_0(v) = p_e(v)$, the particle displacement evolves as

$$\mu(t) = \langle q \rangle t = \chi^{-1} \langle v_e \rangle t, \tag{38}$$

which is equal to the asymptotic long time behavior under arbitrary initial conditions. Note that this is the behavior that is predicted by an unconditional stochastic model because the unconditional ensemble average erases any deviation from a uniform initial condition. Figure 2 shows the evolution of the mean displacements for the lognormal and Gamma fields from uniform and flux-weighted initial conditions. The *outdrw* model predicts quantitatively the full temporal evolution.

4.1.2 Displacement variance

We now focus on the displacement variance. At early times $t \ll \tau_c$, the displacement variance grows ballistically and is given by

$$\kappa(t) = \sigma_{q_1}^2 t^2, \tag{39}$$

where $\sigma_{q_1}^2$ is the variance of $q_1(\mathbf{a})$ in the initial plume

$$\sigma_{q_0}^2 = \int_{\Omega_0} d\mathbf{a} \rho(\mathbf{a}) q_1(\mathbf{a})^2 - \left[\int_{\Omega_0} d\mathbf{a} \rho(\mathbf{a}) q_1(\mathbf{a}) \right]^2 \tag{40}$$

The stochastic particle model (20) predicts

$$\kappa(t) = \chi^{-2} \sigma_{v_0}^2 t^2, \tag{41}$$

where $\sigma_{v_0}^2$ is the variance of $p_0(v)$. The ballistic early time behavior of the displacement variance is underestimated by the approximation (20) of $x_1(s)$ in terms of tortuosity, see the discussion in Appendix B: . The ballistic early time behavior is illustrated in Figure 3 for the Gamma and lognormal fields as well as the predictions of the stochastic particle models. We now consider the behavior for times $t \gg \tau_c$.

4.1.2.1 Gamma fields The Eulerian velocity PDF $p_e(v)$ behaves as a power-law for $v \ll \langle v_e \rangle$, which can be approximated by $p_e(v) \propto v^{\alpha-1}$ with $\alpha = 0.58$. This is illustrated in Figure D.1. Thus, the stationary s-Lagrangian velocity PDF behaves as

$$p_s(v) \propto v^{\beta-1} \tag{42}$$

with $\beta = 1.58$. As discussed in Section 3.1, asymptotically the stochastic particle model may be approximated by a CTRW characterized by independent random time increments whose distribution is given by (29). For the power-law velocity PDF (42), it behaves as $\psi(t) \propto t^{-1-\beta}$. Thus, the CTRW framework predicts for the asymptotic growth of $\kappa(t)$ the superlinear evolution [*Shlesinger*, 1974; *Dentz et al.*, 2004; *Berkowitz et al.*, 2006]

$$\kappa(t) \propto t^{3-\beta}. \tag{43}$$

This behavior is illustrated in Figure 3. The *outdrw* describes the full asymptotic variance evolution with a slight underestimation of the data from the numerical Monte Carlo simulation, which can be traced back again to the value of the correlation length ℓ_c, see also the discussion in the previous section. The model values increase with ℓ_c. We note that, while the power-law scaling is the same for both injection modes, there is a quantitative difference in the pre-factors. This has been predicted in *Dentz et al.* [2016] based on the Bernoulli velocity model. These authors provide explicit analytical solutions for the longitudinal dispersion coefficients $D_L(t)$ for Gamma-distributed $p_s(v)$ under uniform and flux-weighted initial conditions. The parallel dashed lines in Figure 3 show the corresponding $\kappa(t)$ using the parameters obtained from a Gamma fit to the $p_e(v)$ shown in Figure D.1, see also Appendix E: .

4.1.2.2 Lognormal fields For the lognormal conductivity field, we also observe a power-law like behavior that sets in after the ballistic regime. In fact, we find the power-law behavior $\kappa(t) \propto t^{5/4}$. Indeed, this is not a true power-law scaling law, but rather a preasymptotic cross-over behavior. This power-law like behavior arises in the time interval that corresponds to the velocity range over which the velocity distribution $p_e(v)$ shown in Figure D.1 may be fitted by the power-law $p_e(v) \propto v^{\alpha-1}$ with $\alpha = 3/4$, as discussed in Appendix D: . Thus, the CTRW framework predicts the behavior $\kappa(t) \propto t^{3-\beta}$, see (43), with $\beta = 1+\alpha$ over the time range corresponding to the velocity range in Figure D.1. The cross-over behavior is well predicted by the *outdrw* and the *btdrw* (not shown) approaches as show in Figure 3b.

4.2 Macrodispersion

In this section, we focus on the asymptotic long time dispersion behavior for the Lognormal hydraulic conductivity fields predicted by the proposed CTRW model. The motivation for studying the asymptotic regime is threefold. First, we want to emphasize that the displacement variance asymptotically grows linearly with time, and thus show that the observed power-law in the $\sigma_f^2 = 7$ case is indeed only a cross-over, as discussed above. Second, we show that the derived CTRW model provides estimates for macrodispersion, which are consistent with published results from direct numerical Monte-Carlo simulations and stochastic perturbation theory calculations. This demonstrates again the capabilities of the proposed continuous time random walk approach for the prediction of large scale asymptotic solute dispersion at a low computational cost, which is reduced by orders of magnitude compared to the full Monte-Carlo simulation.

Figure 4 shows the long time behaviors estimated from the *outdrw* and *btdrw* models. The asymptotic behavior of $\kappa(t)$ is linear and given by

$$\kappa(t) = 2D_L^\infty t, \tag{44}$$

where D_L^∞ is the asymptotic longitudinal macrodispersion coefficient. Unlike the velocity distribution for the Gamma fields, here the behavior of $p_e(v)$ in the limit $v \to 0$ does

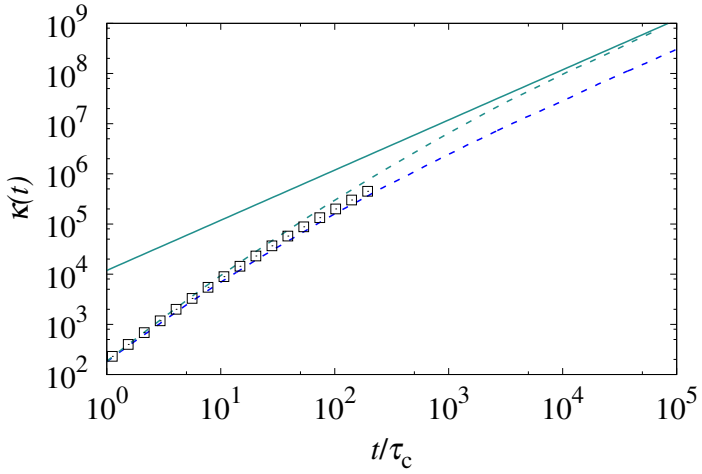

Figure 4. Displacement variances for the lognormal conductivity field with $\sigma_f^2 = 7$ obtained from (squares) Monte-Carlo simulations in 10^2 realizations each with 10^4 particles per realization, (green dashed line) the *btdrw* model and (blue dashed line) *outdrw* model for uniform injection. The solid line denotes the prediction of the analytical expression (45) for the asymptotic longitudinal dispersion coefficient.

not give rise to heavy-tailed transition time distributions $\psi(t)$. The Bernoulli velocity model gives an analytical expression for the asymptotic dispersion coefficient for the displacement along streamlines [*Dentz et al.*, 2016]. In Appendix E: we use this result in order to derive the following expression for the asymptotic longitudinal dispersion coefficient,

$$D_L^\infty = \frac{\langle v_e \rangle \ell_c}{\chi^2} \left(\frac{\langle v_e \rangle}{v_H} - 1 \right), \tag{45}$$

where v_H is the harmonic mean of the Eulerian velocity PDF $p_e(v)$. As shown in Figure 3, the Bernoulli model gives an asymptotic dispersion coefficient that is slightly larger than the one estimated by the Ornstein-Uhlenbeck model. *Fiori et al.* [2003] used a self-consistent approximation to derive an exponential dependence of $D_L^\infty \propto \exp(\sigma_f^2/2)$ of the asymptotic longitudinal macrodispersion coefficient for $\sigma_f^2 \gg 1$. This estimation is based on the assumption that velocity distribution at small values is proportional to the longnormal distribution of hydraulic conductivity. Expresssion (45) is consistent with this estimate, as can be seen by using a lognormal distribution for v_e. Then, at $\sigma_{\ln v}^2 \gg 1$, where $\sigma_{\ln v}^2 \propto \sigma_f^2$ is the variance of $\ln v_e$, we find an exponential dependence because $v_H \propto \exp(-\sigma_{\ln v}^2/2)$ and $\langle v_e \rangle / \chi = \langle q \rangle$.

We study now the longitudinal macrodispersion coefficients for different σ_f^2 based on the *outdrw* model. We choose to rely on this model because it provides an accurate prediction of the Lagrangian velocity statistics [*Hakoun et al.*, 2019]. Figure 5a shows the evolution of $D_L(t)$ for uniform injection condition, this means $p_0(v) = p_e(v)$. At short times, we observe a ballistic linear increase and a cross-over toward the asymptotic regime for times $t \gg \tau_c$. The time to reach the asymptotic regime depends on the variance σ_f^2 of the logarithm of hydraulic conductivity. While for $\sigma_f^2 = 0.1$ the asymptotic value is reached for times around $10\tau_c$, the time to reach the asymptotic regime is about $10^4\tau_c$ for $\sigma_f^2 = 7$. A similar observation was made by *de Dreuzy et al.* [2007] in direct numerical simulations, which reached times of about $10^3\tau_c$. This behavior can be attributed

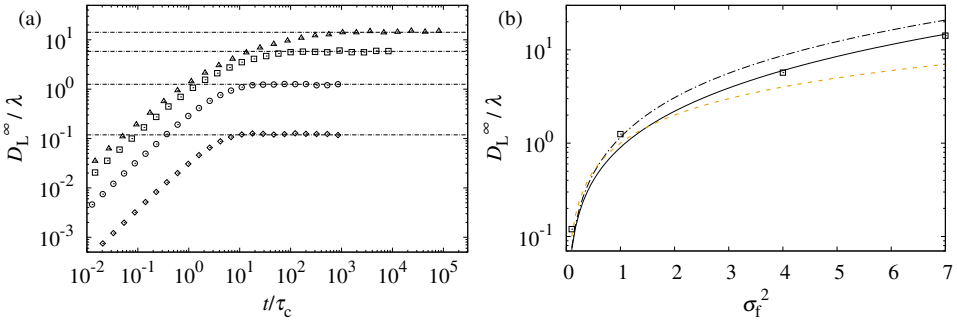

Figure 5. (a) Evolution of the longitudinal dispersion coefficient predicted by the Ornstein-Uhlenbeck model for (diamonds) $\sigma_f^2 = 0.1$, (circles) 1, (squares) 4 and (triangles) 7 for uniform particle injection. The horizontal lines denote the respective asymptotic values. (b) Asymptotic dispersion coefficients as a function of σ_f^2 from (squares) the Ornstein-Uhlenbeck model, (dashed line) Eq. (46), (solid line) Eq. (47) and (dash-dotted lines) Eq. (48).

on one hand to the fact that the correlation length ℓ_c is increasing with σ_f^2, which is weak, however. On the other hand, and more important is the broader distribution of velocities, which imply a broader spectrum of relaxation times toward the asymptotic regime.

Figure 5b shows the asymptotic values D_L^∞ obtained from the *outdrw* model, the estimates by *de Dreuzy et al.* [2007] and *Gotovac et al.* [2009] as well as first-order perturbation theory. For small disorder variance $\sigma_f^2 \ll 1$, stochastic perturbation theory gives for an exponential covariance model the expression [*Gelhar and Axness*, 1983]

$$D_L^\infty = \sigma_f^2 \lambda \langle q \rangle. \tag{46}$$

Expression (45) at small $\sigma_f^2 \ll 1$ is consistent with (46), see Appendix E: . For stronger disorder several authors have proposed expressions for the dependence of $\mathcal{D}$ on the disorder variance σ_f^2 for multi-lognormal hydraulic conductivity fields. *de Dreuzy et al.* [2007] found the following relation

$$D_L^\infty = \langle q \rangle \lambda \left(0.7\sigma_f^2 + 0.2\sigma_f^4\right). \tag{47}$$

Gotovac et al. [2009] propose the expression

$$D_L^\infty = \langle q \rangle \lambda \left(\frac{\sigma_f^2}{4} + \frac{\sigma_f^4}{5} + \frac{\sigma_f^6}{500}\right)\left[\frac{4}{3} + \frac{3}{2}\exp\left(-\frac{\sigma_f^2}{5}\right)\right]. \tag{48}$$

The value obtained for D_L^∞ for $\sigma_f^2 = 0.1$ is consistent with expressions (46)–(48). For $\sigma_f^2 = 1$, the *outdrw* estimate for D_L^∞ is on the curve (48), for $\sigma_f^2 = 4$ and 7, the *outdrw* estimate lies on the curve (47). We have shown that the proposed Markov model allows to make predictions for the non-perturbative behaviors and we have provided an explanation why the time scales for which the asymptotic time regime is reached increase with the heterogeneity strength, i.e. with σ_f^2. The results of the Markov model are fully consistent with classical stochastic perturbation theory, as well as with the known results for large σ_f^2 [*de Dreuzy et al.*, 2007; *Gotovac et al.*, 2009]. In this sense, the presented results close a gap between classical stochastic models and time domain and continuous time random walk models.

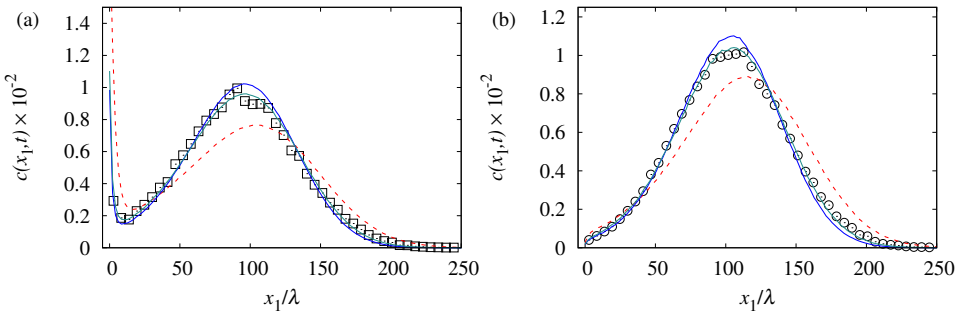

Figure 6. Concentration distributions at $t = 100\tau_c$ for the lognormal field with $\sigma_f^2 = 7$ for (a) uniform and (b) flux-weighted injection conditions. Symbols indicate the direct numerical simulations with 10^6 particles per each of the 10^2 realizations. The red dashed lines refer to the predictions of the *btdrw* with 10^7 particles, the solid blue lines to the *outdrw* predictions for $\ell_c = 3.5\lambda$ and the solid green lines for $\ell_c = 4\lambda$ using 10^7 particles.

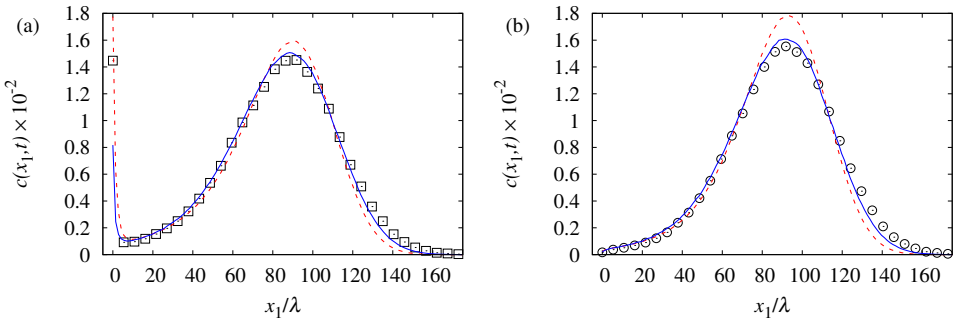

Figure 7. Concentration profiles at $t = 83\tau_c$ in Gamma K fields for uniform (left panel) and flux-weighted (right panel) injection modes. Symbols indicate the results from Monte Carlo direct numerical simulations (100 realizations, 10^6 particles/realization). Lines refer to numerical simulations of the TDRW model with 10^7 particles, using the Bernoulli (dashed red lines) and the Ornstein-Uhlenbeck (solid blue lines) models for the evolution of the velocity PDF.

4.3 Concentration distribution

We study the concentration distributions along the mean flow direction in direct numerical simulations and the predictions by the upscaled stochastic model based on the proposed velocity Markov processes for lognormal and Gamma conductivity fields.

Figure 6 shows the particle concentration for the lognormal fields obtained through direct numerical simulations and the predictions of the *btdrw* and *outdrw* models. As expected, for the uniform injection, we observe particle localization at the injection point because particles sample high and low flow velocities with equal probability. For the flux-weighted injection, there is no localization because particles sample preferentially higher flow velocities. The peaks move approximately with the same velocity. This result emphasizes the importance of conditioning on the initial data, see also the works by *Le Borgne et al.* [2007] and *Dagan* [2017] for Darcy scale heterogeneous media, *Hyman et al.* [2015] and *Kang et al.* [2017] for fractured media, *Morales et al.* [2017] and *Puyguiraud et al.* [2019] for por-scale transport. The *outdrw* predicts the concentration profiles under both injection conditions with a slight underestimation of the forward tail under both injection conditions. This slight mismatch may be caused by uncertainty in the correlation length ℓ_c given by the empirical expression (D.1). For a larger ℓ_c, more particles are found

in the backward and forward tails due to the increased persistence of low and high velocities in particular. This is illustrated in Figure 6 for $\ell_c = 4\lambda$, which provides a better match with the direct numerical simulations. Nevertheless, in the following, we will employ the values for ℓ_c given in Appendix D: . The *btdrw* predictions are qualitatively sound as they capture solute localization at the origin in the case of uniform injection and the lack of it for flux-weighted. The weight of the peak at zero and peak displacement are overestimated, however. This may be traced back to the fact that in the Bernoulli model velocity transitions occur at a constant velocity-independent rate, while it appears that high velocities converge slower and low velocities faster toward the steady state. Remarkably, Figure 6 b also shows a mismatch between the *btdrw* predictions and the results by direct numerical simulations, despite the fact that the velocity statistics are stationary under flux-weighted injection conditions. This discrepancy again can be traced back to the velocity transition probability between subsequent velocities, which is the same for all velocities, unlike in the Ornstein-Uhlenbeck process, for which the transition probability depends on the velocity value.

Figure 7 shows the concentration profiles for the Gamma conductivity fields. The basic features distinguishing the uniform and flux-weighted injection conditions are the same as for the lognormal case. Also, the prediction of the *outdrw* model compares well with the direct numerical simulations except for a slight mismatch in the forward tails, similar as for the lognormal fields. Here, the prediction of the *btdrw* model compares better with the direct simulation than for the lognormal field. The peak is slightly overestimated, the tail underestimated. This better match of the *btdrw* model may be due to the fact that here transport is dominated rather by the strong tailing of the steady velocity PDF toward low velocities than by correlation. This means the specific velocity evolution model is less important than accounting for the fact that the initial velocity distribution is non-stationary.

In order to emphasize further the importance of flow heterogeneity at the injection region, we condition the initial velocity to the knowledge of the local properties at the injection site. To do so we consider two different scenarios. First, we inject particles in regions of low velocities. This scenario is motivated by the MADE experiment [*Boggs et al.*, 1992; *Adams and Gelhar*, 1992; *Rehfeldt et al.*, 1992], where the tracer injection was performed in low permeability regions [*Harvey and Gorelick*, 2000]. In our simulations, injection of particles is performed along a line, but only the velocities belonging to the first 10 percentiles of the Eulerian velocity PDF are taken into account. Second, we consider a scenario, in which particles are injected into regions of high velocities, so that the initial velocities belongs to the highest 10 percentiles of the full Eulerian velocity PDF. The velocity ranges are highlighted in Figure D.1. Figure 8 shows the concentration profiles for injections in regions of low and high velocities obtained by the direct Monte Carlo simulations and the predictions by the *outdrw* approach for lognormal and Gamma conductivity fields. For injection in regions of low velocities, particles travel on average less distance than in the uniform and flux-weighted injection cases and a significant backward tail can be observed. The concentration distribution is characterized by a peak at the origin, which is due to retention by low velocities and a moving peak of mobile solute. For injection in high velocity zones, there is no particle retention at the injection point and a fast advance of the peak position. The *outdrw* predicts both scenarios qualitatively and quantitatively with a slight mismatch at the forward tails for the reasons discussed above.

4.4 Breakthrough curves

We now study the impact of medium heterogeneity and initial velocity distribution and its prediction in terms of the stochastic particle models for solute breakthrough curves measured at different control planes. The direct numerical simulations consider 10^2 disorder realizations with 10^7 particles per realization. The data from the direct nu-

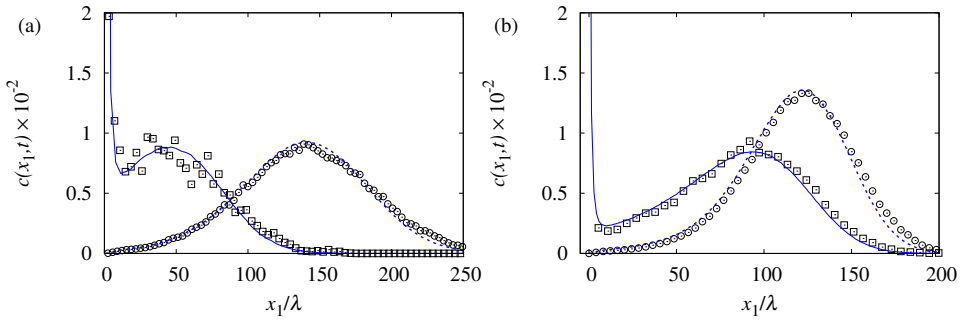

Figure 8. Concentration profiles obtained by Monte Carlo simulation for injection in (circles) low and (squares) high velocity regions using 10^7 particles in 10^2 medium realizations. The profiles for the lognormal K-distribution with $\sigma_f^2 = 7$ are for times $t = 80\tau_c$ and $t = 135\tau_c$ respectively. For the Gamma distribution with $\alpha = 1/2$, the profiles are at $t = 110\tau_c$. The direct numerical simulations are compared with the predictions of the corresponding *outdrw* model conditioned on (solid line) low and (dashed line) high flow velocities using 10^7 particles.

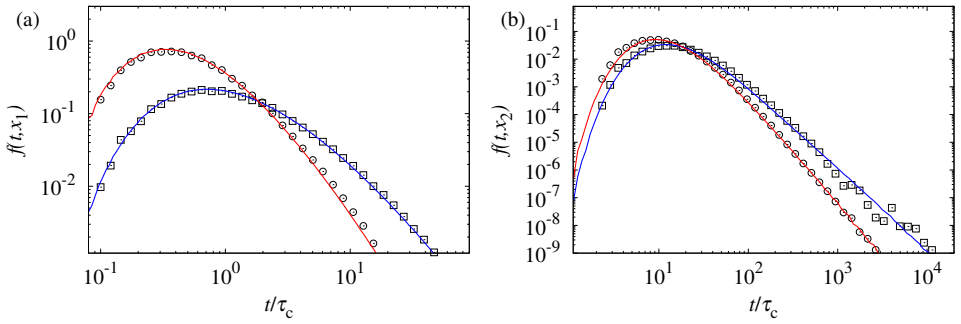

Figure 9. Breakthrough curves measured at (a) $x_1 = 2\lambda$ and (b) $x_1 = 20\lambda$ from (symbols) Monte Carlo simulations for (squares) uniform and (circles) flux-weighted injection for lognormal conductivity with $\sigma_f^2 = 7$. The solid lines indicate the corresponding predictions by the *outdrw* for 10^7 particles.

merical simulations is compared to the predictions of the stochastic particle model, which is solved numerically using 10^8 particles. As above, we consider lognormal conductivity fields with $\sigma_f^2 = 7$ and Gamma fields characterized by $\alpha = 1/2$. The initial particle distributions are uniform and flux-weighted. The stochastic particle model predicts the breakthrough time at a control plane at distance x_1 from the inlet as

$$t(x_1\chi) = \int_0^{x_1\chi} ds \frac{1}{v_s(s)}, \tag{49}$$

see (20). For the flux-weighted initial condition, the s-Lagrangian velocity PDF $p_s(v, s)$ is stationary and equal to $p_s(v) = vp_e(v)/\langle v_e\rangle$. Thus, under these conditions, the mean arrival time is

$$\langle t(x_1\chi)\rangle = \int_0^{x_1\chi} ds \int_0^{\infty} dv \frac{1}{v} \frac{vp_e(v)}{\langle v_e\rangle} = \frac{x_1\chi}{\langle v_e\rangle} \equiv \frac{x_1}{\langle q\rangle}. \tag{50}$$

Under non-stationary conditions, this is in general not the case and the mean breakthrough time evolves with distance x_1,

$$\langle t(x_1\chi)\rangle = \int_0^{x_1\chi} ds \int_0^{\infty} dv \frac{p_s(v, s)}{v}. \tag{51}$$

For the *btdrw* model, $p_s(v, s) = p_0(v)\exp(-s/\ell_c) + [1 - \exp(-s/\ell_c)]p_s(v)$. This gives for the mean breakthrough time

$$\langle t(x_1\chi)\rangle = \frac{x_1}{\langle q\rangle} + [1 - \exp(-x_1\chi/\ell_c)] \int_0^{\infty} dv \frac{\ell_c}{v}[p_0(v) - p_s(v)]. \tag{52}$$

For $x_1\chi \gg \ell_c$, the mean breakthrough time is in leading order $\langle t(x_1\chi)\rangle \approx x_1/\langle q\rangle$, this means the mean breakthrough times for different initial distributions converge at large distances. These results emphasize the impact of the initial velocity distribution on the breakthrough curves. Note that the breakthrough curve in this framework is given by (26). For a control plane located at a distance $x_1 < \ell_c$ smaller than the velocity correlation length, we may approximate the s-Lagrangian velocity distribution $p_s(v, s) \approx p_0(v)$ by the distribution of initial velocities. Thus, we obtain for the breakthrough curve at $x_1 < \ell_c$

$$f(t, x_1) = \langle \delta(t - x_1\chi/v_0)\rangle \equiv \frac{x_1\chi}{t^2} p_0(x_1\chi/t). \tag{53}$$

This implies for the uniform and flux-weighted initial particle distributions

$$f(t, x_1) = \frac{x_1\chi}{t^2} p_e(x_1\chi/t), \qquad f(t, x_1) = \frac{x_1^2\chi^2}{t^3\langle v_e\rangle} p_e(x_1\chi/t), \tag{54}$$

respectively, see also *Kang et al.* [2017]. Thus, at short distances the breakthrough curves for uniform injection show stronger tailing than for flux-weighted. In the following, we investigate these features and the impact of the heterogeneity distribution for the full breakthrough curves. The direct numerical data are compared to the predictions of the *outdrw*, which is based on an Ornstein-Uhlenbeck process for the normal scores of velocity.

Figures 9 and 10 show the breakthrough curves obtained by direct Monte Carlo simulations for the lognormal and Gamma fields, and the respective predictions by the *outdrw* model. The *outdrw* provides accurate predictions at short and longer distances from the inlet both for flux-weighted and uniform injection conditions. As discussed above,

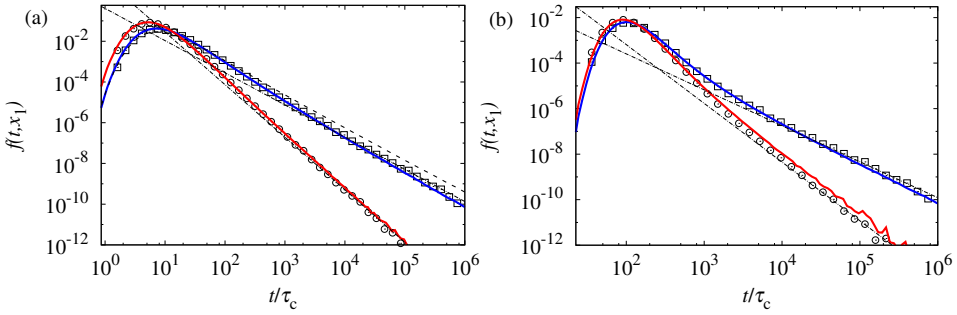

Figure 10. Breakthrough curves at measured at (a) $x_1 = 2\lambda$ and (b) $x_1 = 20\lambda$ from (symbols) Monte Carlo simulations for (squares) uniform and (circles) flux-weighted injection for Gamma conductivity with $\alpha = 1/2$. The solid lines indicate the corresponding predictions by the *outdrw* for 10^7 particles. The dashed line is expression (54) for uniform injection and the dash-dotted lines indicates the scalings $t^{-1-\alpha}$ and $t^{-1-\beta}$ with $\alpha = 0.58$ and $\beta = 1.58$.

the breakthrough curves show stronger tailing for the uniform injection conditions because it places more weight on the low flow velocities. For the lognormal field the breakthrough curves are converging with increasing distance x_1 from the inlet.

For the Gamma fields, we observe a clear power-law tailing, whose exponent is related to the behavior of the velocity PDF $p_0(v)$ and $p_s(v)$ at small velocities. As discussed in the previous section, $p_e(v) \propto v^{\alpha-1}$ with $\alpha = 0.58$. According to (54), this implies that $f(t, x_1) \propto t^{-1-\alpha}$ for uniform injection and $x_1 < \ell_c$. For the flux-weighted injection, we obtain accordingly $f(t, x_1) \propto t^{-2-\alpha}$. These behaviors are clearly shown in Figure 10a. Figure 10b shows that the bulk of the breakthrough curves converge at increasing distance x_1 and thus the mean breakthrough times, as discussed above. The difference in the exponent of the power-law tails, however, persists. In order to understand this, we note that the breakthrough curves can be seen as the convolution of the initial transition time distribution $\psi_0(t)$ with the transition time distributions $\psi(t)$ of the subsequent steps. Here, $\psi_0(t) \propto t^{-1-\alpha}$ with $\alpha = 0.58$ while $\psi(t) \propto t^{-1-\beta}$ with $\beta = 1 + \alpha = 1.58$. This convolution is dominated by $\psi_0(t)$, which means that the tail of the breakthrough curves here is dominated by the initial velocity distribution. For the lognormal conductivity field on the other hand, the initial velocity distribution only dominates at short distance, but does not persist at large distances x_1 from the inlet because the velocity distribution decays stronger than a power-law at small velocities.

5 Conclusions

We study the mechanisms and stochastic dynamics of advective transport in heterogeneous porous media with the aim of deriving an upscaled modeling approach to predict large scale solute dispersion. To this end, we conduct direct numerical Monte Carlo simulations in heterogeneous hydraulic conductivity fields, which are compared to a stochastic particle model which is based on the representation of equidistant s(pace)-Lagrangian velocities as Markov processes. We focus on two injection modes over a line that spans more than 100 correlation lengths λ perpendicular to the mean flow direction, which is close to ergodic injection conditions. As observed in the literature, the preasymptotic transport behavior depends critically on the injection mode and the heterogeneity in the injection region, which stresses the need for a predictive upscaled transport approach that can be conditioned on the initial data. The proposed continuous time random walk approach can be conditioned on the initial velocity distribution in a systematic way through the representation of the s-Lagrangian velocities as a Markov process. We consider two

Markov models of different complexity, a Bernoulli process for the actual velocities and a Ornstein-Uhlenbeck process for the normal scores of the particle velocities. The resulting CTRW models provide accurate predictions for the evolution of the concentration distribution, dispersion and breakthrough curves for different types of heterogeneity and different injection conditions.

The derived CTRW model can be implemented according to the following strategy. First, determine the flow statistics of the heterogeneous porous medium from numerical flow simulations in the hydraulic conductivity field with given geostatistical characteristics and determine the velocity statistics at the injection region as well as the advective tortuosity. Second, determine the velocity correlation length either from perturbation theory or from an empirical regression such as Eq. D.1 for the *outdrw* model or the similar expression given in [*Cvetkovic et al.*, 1996]. These quantities then fully parameterize the velocity transition model, the Bernoulli or the Ornstein-Uhlenbeck process, which is incorporated into the time domain random walk model that allows for accurate and fast predictions of the large scale transport behavior.

Apart from the practical value for the fast prediction of large scale transport, the stochastic particle model sheds some new light on the mechanisms of preasymptotic and asymptotic solute dispersion, specifically on the role of the velocity distribution versus correlation. This is manifested for the Gamma fields in the superlinear asymptotic growth of the spatial variance of the solute distribution is observed. This superdiffusive behavior originates in a broad distribution of low flow velocities. For the lognormal fields, the longitudinal macrodispersion coefficient evolves from a ballistic early time behavior through a broad crossover region towards a constant asymptotic long time value. The size of the crossover region and the time-scale to reach the asymptotic value depend strongly on the variance σ_f^2 of log-hydraulic conductivity. For increasing values of σ_f^2, the Eulerian velocity PDF becomes more tailed towards low velocities. As a result, the times for which particles are caught in low velocity zones, increase. This explains the increasing times to reach the asymptotic macrodispersion regime.

We expect the fundamental flow and transport mechanisms observed and quantified in this work to hold also for 3-dimensional heterogeneous porous media. The basic mechanism is the existence of velocities transitions on characteristic length scales along particle trajectories. This property requires essentially that the heterogeneous medium is stationary and can be characterized by finite correlation lengths. Thus, it is not restricted to 2-dimensional isotropic media. The dimensionality of space is expected to affect quantitatively both the tortuosity χ and the correlation scale ℓ_c. The extension of the stochastic particle model to 2 and 3 dimensions requires the statistical characterization of the direction angle [*Benke and Painter*, 2003], whose mean points in the mean flow direction with the magnitude $1/\chi$.

The proposed approach is valid for advection-dominated transport at high Péclet numbers, which at the field scale is rather the rule than an exception [*Dagan*, 1989; *Gelhar*, 1993; *Rubin*, 2003]. Nevertheless, the impact of local scale dispersion is expected to mitigate for example the localization of solute at the injection regions because particles may be released by dispersive mass transfer. Also, persistent superlinear growth of the variance, which is due to the persistence of low velocities will eventually tend towards a linear behavior because particle velocities decorrelate due to local dispersive mass transfer. In the current modeling framework, the impact of local-scale dispersion on large scale transport can be accounted for through a cut-off in the transition times over the characteristic heterogeneity length scales [*Dentz et al.*, 2004].

In conclusion, while some questions remain open, the proposed approach represents a step towards linking hydraulic medium properties and flow conditions to large scale transport in a predictive upscaled modeling framework.

Acknowledgments
The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC Grant Agreement No. 617511 (MHetScale). The data presented in the paper can be accessed at https://gitlab.com/marco.dentz/darcy-simulations-data.

A: Direct numerical simulations

Direct numerical simulations are of the Monte Carlo type. We consider 10^2 realizations of $K(\mathbf{x})$ of each random field under consideration. Monte Carlo simulations are carried out according to a methodology that consists of two steps. The first step consists in generating random fields of hydraulic conductivity and in solving Darcy equation (2) numerically to obtain the flow field. Secondly, we simulate transport through particle tracking, which uses 10^4 particles in each disorder simulation. The methodology is summarized in the following. More details can be found in *Hakoun et al.* [2019].

A.1 Field generation and flow solver

We generate multi-Gaussian fields $Y(\mathbf{x})$ of size $L_x \times L_y$ using the Random Fields package [*Schlather et al.*, 2015] of the R software environment for statistical computing. The resolution of all fields is $\Delta x = \Delta y = \lambda/10$. We then map the point values of the multi-Gaussian field onto the desired distribution through the transformation (3), which is implemented by inverse transform sampling [*Devroye*, 1986]. Table A.1 summarizes the values of the parameteres that we used in the simulations.

Next, we solve the Darcy equation (2) for incompressible flows within the generated $K(\mathbf{x})$ fields. We impose permeameter-like boundary conditions, i.e. a constant head gradient along the x direction and no-flow conditions at the bottom and top boundaries. The head gradient is set so that flow occurs from left to right. The inter-cell hydraulic conductivity is taken as the harmonic mean of the K of two adjacent cells. We use a flow solver based on the finite volume scheme adapted from *Aarnes et al.* [2007]. To avoid instabilities of the flow field due to border effects, we perform transport simulations by excluding an external frame of 20λ for the lognormal fields and of 15λ for the Gamma fields. Particle injection is thus performed along a line placed at $x_1 = 20\lambda$ and at $x_1 = 15\lambda$, respectively.

A.2 Particle tracking

The last step consists in simulating transport in the obtained Darcy flows by tracking particles at equal distances as they advect along their streamlines. The discretized trajectory reads

$$\mathbf{x}_{n+1} = \frac{\mathbf{q}[\mathbf{x}_n]\Delta s}{||\mathbf{q}(\mathbf{x}_n)||}, \qquad t_{n+1} = t_n + \frac{\Delta s}{||\mathbf{q}[\mathbf{x}_n]||}, \tag{A.1}$$

where Δs is a constant spatial increment. The process (A.1) is a time domain random walk [*Noetinger et al.*, 2016; *Russian et al.*, 2016]. Compared to classical random walk particle tracking, this process guarantees faster computations for our scenarios, since the number of steps does not depend on the local velocity. This is an asset for us, since our flow fields are characterized by very low velocities. The finite volume flow solver provides the values of the velocity components at the cell faces. To obtain the velocities within the cells, we use the bi-linear interpolation scheme proposed by *Pollock* [1988]

$$q_1(\mathbf{x}) = A_1(x_1 - x_{1,L}) + q_{1,L} \tag{A.2}$$
$$q_2(\mathbf{x}) = A_2(x_2 - x_{2,B}) + q_{2,B}, \tag{A.3}$$

Field	λ [m]	L_x	L_y	Δx, Δy	μ	σ_Y^2	α	k_0 [m/s]	k_c [m/s]
Lognormal	10	600λ	150λ	$\lambda/10$	0	0.1–7	-	-	-
Gamma	10	300λ	300λ	$\lambda/10$	-	-	1/2	10^{-11}	5

Table A.1. Parameters used for direct numerical simulations for the lognormal and Gamma fields considered in this study.

where

$$A_1 = (q_{1,R} - q_{1,L})/\Delta x_1 \tag{A.4}$$

$$A_2 = (q_{2,T} - q_{2,B})/\Delta x_2, \tag{A.5}$$

where $x_{1,L}$ and $x_{2,B}$ are the coordinates of the left and bottom face of the cell, respectively, $q_{1,L}$ and $q_{1,R}$ are the longitudinal components of the velocity at the left and right cell face, respectively, while $q_{2,B}$ and $q_{2,T}$ are the transverse components of the velocity at the bottom and top cell face, respectively.

B: Advective tortuosity

The advective tortuosity compares the length of the streamline to the linear distance. Advective tortuosity has been considered in the context of pore-scale flow [*Koponen et al.*, 1996; *Ghanbarian et al.*, 2013]. Here we determine it in the context of Darcy-scale flow with the aim of providing an approximation for the particle displacement in mean flow direction. To this end, we consider the 1-component $x_1(s, \mathbf{a})$ of the s-Lagrangian trajectory $\mathbf{x}(s, \mathbf{a})$ defined in Eq. (10),

$$\frac{dx_1(s, \mathbf{a})}{ds} = \frac{q_1[\mathbf{x}(s, \mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|} \equiv \omega(s, \mathbf{a}). \tag{B.1}$$

Thus, the linear distance as a function of streamwise distance is given by

$$x_1(s, \mathbf{a}) = \int_0^s ds' \frac{q_1[\mathbf{x}(s, \mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|}. \tag{B.2}$$

Thus, the ratio between linear and streamwise distance is given by

$$\frac{x_1(s, \mathbf{a})}{s} = \frac{1}{s} \int_0^s ds' \frac{q_1[\mathbf{x}(s, \mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|}, \tag{B.3}$$

this means, it is equal to the streamwise average of $q_1[\mathbf{x}(s, \mathbf{a})]/|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|$. Under ergodic conditions, the streamwise average is equal to the average over the ensemble of particles. In order to define the ensemble average here, we recall that for the streamwise sampling mode employed in the s-Lagrangian formulation, the flux-weighted injection represents the stationary initial condition [*Dentz et al.*, 2016]. Thus, we consider the ensemble average with respect to the flux-weighted injection condition (12). This gives

$$\lim_{s\to\infty} \frac{1}{s} \int_0^s ds' \frac{q_1[\mathbf{x}(s, \mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|} = \lim_{V_0\to\infty} \int_{\Omega_0} d\mathbf{a} \frac{|\mathbf{q}(\mathbf{a})|}{\langle |\mathbf{q}(\mathbf{a})| \rangle} \frac{q_1[\mathbf{x}(s, \mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s, \mathbf{a})]|} \equiv \chi^{-1}, \tag{B.4}$$

which is equal to the inverse of tortuosity. Thus, tortuosity here compares the streamwise distance s to linear distance $x_1(s, \mathbf{a})$ at large distances s.

B.1 Approximation of displacement in mean flow direction

It quantifies at the same time the average of $\omega(s, \mathbf{a})$ such that we can approximate

$$\frac{dx_1(s,a)}{ds} \approx \chi^{-1}. \tag{B.5}$$

This is valid for streamwise distances $s \gg \lambda$ because then

$$x_1(s,a) = s\left[\frac{1}{s}\int_0^s ds' \frac{q_1[\mathbf{x}(s,\mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s,\mathbf{a})]|}\right] \approx \frac{s}{\chi}. \tag{B.6}$$

Note that for $s \ll \lambda$

$$x_1(s,\mathbf{a}) = \chi^{-1}s + \omega_0' s, \tag{B.7}$$

where $\omega'(s,\mathbf{a}) = \omega(s,\mathbf{a}) - \chi^{-1}$. This implies that

$$x_1(t,\mathbf{a}) = \chi^{-1}v_0 t + \omega_0' v_0 t, \tag{B.8}$$

where $\omega_0' = \omega'(s=0,\mathbf{a})$. Thus, we obtain for the mean displacement

$$\mu(t) = \chi^{-1}\langle v_0\rangle t \tag{B.9}$$

and for the displacement variance

$$\kappa(t) = \chi^{-2}\sigma_{v_0}^2 t^2 + \sigma_{\omega_0}^2\sigma_{v_0}^2 t^2 \tag{B.10}$$

Thus, the approximation (B.5) estimates accurately the early time behavior of the mean displacement, but underestimates the ballistic early time behavior of the displacement variance.

B.2 Explicit analytical expression for tortuosity

We obtain an explicit analytical expression for χ by considering the ensemble average

$$\chi^{-1} = \lim_{V_0\to\infty}\frac{1}{V_0}\int_{\Omega_0} d\mathbf{a}\frac{|\mathbf{q}(\mathbf{a})|}{\langle|\mathbf{q}(\mathbf{a})|\rangle}\frac{q_1[\mathbf{x}(s,\mathbf{a})]}{|\mathbf{q}[\mathbf{x}(s,\mathbf{a})]|} \tag{B.11}$$

Thus, we perform the variable transform $\mathbf{a} \to \mathbf{x}$,

$$\chi^{-1} = \lim_{V_0\to\infty}\frac{1}{V_0}\int_{\Omega(s)} d\mathbf{x}\frac{1}{\mathbb{J}(s,\mathbf{a})}\frac{|\mathbf{q}(\mathbf{a})|}{\langle|\mathbf{q}(\mathbf{a})|\rangle}\frac{q_1(\mathbf{x})}{|\mathbf{q}(\mathbf{x})|}\bigg|_{\mathbf{a}=\mathbf{a}(\mathbf{x},s)} \tag{B.12}$$

which is characterized by the Jacobian determinant $\mathbb{J}(s,\mathbf{a})$, which is the determinant of the deformation gradient tensor

$$F_{ij}(s,\mathbf{a}) = \frac{\partial x_i(s,\mathbf{a})}{\partial a_j}. \tag{B.13}$$

The Jacobian satisfies [*Batchelor*, 2000, p. 75]

$$\frac{\partial\mathbb{J}(s,\mathbf{a})}{\partial s} = \mathbb{J}(s,\mathbf{a})\ \nabla\cdot\frac{\mathbf{q}(\mathbf{x})}{|\mathbf{q}(\mathbf{x})|}\bigg|_{\mathbf{x}=\mathbf{x}(s,\mathbf{a})}. \tag{B.14}$$

This expression can be expanded as

$$\frac{\partial\mathbb{J}(s,\mathbf{a})}{\partial s} = -\mathbb{J}(s,\mathbf{a})\ \mathbf{q}(\mathbf{x})\cdot\frac{\nabla|\mathbf{q}(\mathbf{x})|}{|\mathbf{q}(\mathbf{x})|^2}\bigg|_{\mathbf{x}=\mathbf{x}(s,\mathbf{a})}, \tag{B.15}$$

where we used $\nabla \cdot \mathbf{q}(\mathbf{x}) = 0$. In order to simplify the right side, we note that

$$\frac{d|\mathbf{q}[\mathbf{x}(s,a)]|}{ds} = \frac{d\mathbf{x}(s,a)}{ds} \cdot \nabla|\mathbf{q}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}(s,\mathbf{a})}. \tag{B.16}$$

By using (10), the latter can be written as

$$\frac{d|\mathbf{q}[\mathbf{x}(s,a)]|}{ds} = \mathbf{q}(\mathbf{x}) \cdot \left.\frac{\nabla|\mathbf{q}(\mathbf{x})}{|\mathbf{q}(\mathbf{x})|}\right|_{\mathbf{x}=\mathbf{x}(s,\mathbf{a})} \tag{B.17}$$

and therefore

$$\mathbf{q}(\mathbf{x}) \cdot \left.\frac{\nabla|\mathbf{q}(\mathbf{x})}{|\mathbf{q}(\mathbf{x})|^2}\right|_{\mathbf{x}=\mathbf{x}(s,\mathbf{a})} = \frac{1}{v_s(s,\mathbf{a})}\frac{dv_s(s,\mathbf{a})}{ds}, \tag{B.18}$$

where we set $v_s(s,\mathbf{a}) = |\mathbf{q}[\mathbf{x}(s,\mathbf{a})]|$. Thus, Eq. (B.15) can be written as

$$\frac{\partial \mathbb{J}(s,\mathbf{a})}{\partial s} = -\mathbb{J}(s,\mathbf{a})\frac{d\ln[v_s(s,\mathbf{a})]}{ds}. \tag{B.19}$$

This equation can be integrated by separation of variables, which gives

$$\mathbb{J}(s,\mathbf{a}) = \frac{v_s(0,\mathbf{a})}{v_s(s,\mathbf{a})} = \frac{|\mathbf{q}(\mathbf{a})|}{|\mathbf{q}[\mathbf{x}(s,\mathbf{a})]|}, \tag{B.20}$$

where we used that by definition $\mathbb{J}(s=0,\mathbf{a}) = 1$. Inserting this expression into (B.12) gives

$$\chi^{-1} = \lim_{V_0\to\infty}\frac{1}{V_0}\int\limits_{\Omega(s)} d\mathbf{x}\frac{q_1(\mathbf{x})}{\langle|\mathbf{q}(\mathbf{a})|\rangle} = \frac{\langle q_1(\mathbf{x})\rangle}{\langle|\mathbf{q}(\mathbf{x})|\rangle}. \tag{B.21}$$

C: Boltzmann equation

We derive here equation (27) for the joint distribution $p(x_1,v,t)$ of particle position and velocity. The joint distribution is defined by

$$p(x_1,v;t) = \int\limits_0^\infty ds\langle\delta[s-s(t)]\delta[x_1-x_1(s)]\delta[v-v(s)]\rangle, \tag{C.1}$$

We make use of the following property of Dirac Delta

$$\delta[s-s(t)] = \frac{\delta[t-t(s)]}{ds(t)/dt}, \tag{C.2}$$

which, in virtue of the equivalence $ds(t)/dt = v(t) = v_s[s(t)]$, reads as

$$\delta[s-s(t)] = \frac{\delta[t-t(s)]}{v_s(s)}. \tag{C.3}$$

By substituting the latter into Equation (C.1), we obtain

$$p(x_1,v;t) = \int\limits_0^\infty dsv^{-1}R(x_1,v,t;s) \tag{C.4}$$

where we defined

$$R(x_1,v,t;s) = \langle\delta[t-t(s)]\delta[x_1-x_1(s)]\delta[v-v(s)]\rangle. \tag{C.5}$$

By considering an increment of length Δs, we can write

$$R(x_1, v, t; s + \Delta s) = \int_0^\infty dv' r(v, \Delta s|v') R(x_1 - \Delta s/\chi, v', t - \Delta s/v'; s). \tag{C.6}$$

We consider the limit $\Delta s \to 0$ and perform the Taylor expansion

$$R(x_1 - \Delta s/\chi, v, t - \Delta s/v; s) = R(x_1, v, t; s) - \frac{\Delta s}{\chi}\frac{\partial}{\partial x_1} R(x_1, v, t; s) - \frac{\Delta s}{v}\frac{\partial}{\partial t} R(x_1, v, t; s) + ... \tag{C.7}$$

so that Equation (C.6) reduces to

$$R(x_1, v, t; s + \Delta s) = \int_0^\infty dv' r(v, \Delta s|v') R(x_1, v', t; s) - \int_0^\infty dv' r(v, \Delta s|v') \Delta s \left[\frac{1}{\chi}\frac{\partial}{\partial x_1} + \frac{1}{v'}\frac{\partial}{\partial t}\right] R(x_1, v', t; s). \tag{C.8}$$

We can further write

$$\frac{\partial R(x_1, v, t; s)}{\partial s} = \int_0^\infty dv' \frac{r(v, \Delta s|v')}{\Delta s} [R(x_1, v', t; s) - R(x_1, v, t; s)] - \left[\frac{1}{\chi}\frac{\partial}{\partial x} + \frac{1}{v}\frac{\partial}{\partial t}\right] R(x_1, v, t; s), \tag{C.9}$$

where we used that $r(v, 0|v') = \delta(v - v')$. Integration of Equation (C.9) over s and using expresion (C.4) gives (27).

D: Setup of stochastic particle model

The stochastic particle model requires as input the Eulerian velocity point distribution $p_e(v)$, advective tortuosity χ and the Lagrangian velocity correlation scale ℓ_c. Throughout Section 4, we focus on the preasymptotic transport behaviors in the Lognormal field for $\sigma_f^2 = 7$ and the Gamma field with $\beta = 1/2$. In Section 4.2, we use the *ouctrw* to model asymptotic longitudinal macrodispersion coefficients for $\sigma_f^2 = 0.1, 1, 4$ and 7.

The velocity correlation scale ℓ_c for the hydraulic conductivity fields under consideration are were studied in detail in *Hakoun et al.* [2019]. The correlation length is model dependent and increases with the variance of the log-conductivity $f(\mathbf{x})$. In the Ornstein-Uhlenbeck model, the correlation length for the lognormal fields is given by the empirical expression

$$\frac{\ell_c}{\lambda} = 0.181\sigma_f^2 + 2.221, \tag{D.1}$$

which is consistent with the expression provided in *Cvetkovic et al.* [1996]. The correlation length increases with increasing σ_f^2 because the streamline tortuosity increases and thus the streamline length. The correlation length for the Gamma field in the Ornstein-Uhlenbeck model is $\ell_c = 2.7\lambda$. In the Bernoulli model, the correlation length for the Lognormal field with $\sigma_f^2 = 7$ is $\ell_c = 5.7\lambda$ and for the Gamma field it is $\ell_c = 3.1\lambda$.

The PDFs are computed by sampling the Eulerian velocity magnitudes in every node of the numerical velocity field for all the 10^2 medium realizations, which gives a sample of about 10^9 velocity values. The histograms are computed using logarithmically

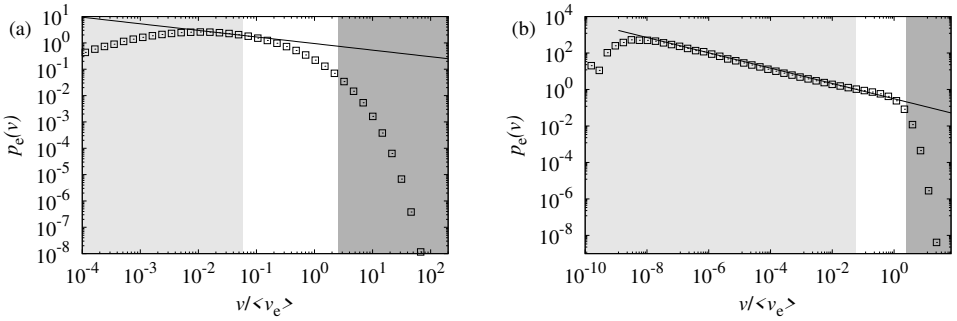

Figure D.1. Eulerian velocity PDF for the lognormal and Gamma conductivity fields. Colored areas indicate the velocity range corresponding to the first (light gray) and the last (dark gray) 10 percentiles. The solid lines represent the scalings $p_e(v) \sim v^{\alpha-1}$ with (a) $\alpha = 3/4$ and (b) $\alpha = 0.58$.

spaced bins between the minimum and the maximum velocities. The data is normalized by the total count and respective bin size. No smoothing is performed nor needed, because of the large amount of entries. We briefly describe some features of the Eulerian velocity PDF for the lognormal fields with $\sigma_f^2 = 7$ and the Gamma field for $\beta = 1/2$, which are shown in Figure D.1. The velocity PDFs for the Lognormal fields with $\sigma_f^2 = 0.1, 1$ and 4, which are used in Section 4.2 for the extrapolation of the longitudinal macrodispersion coefficients in the *ouctrw* are shown in Figure D.2. Figure D.1a shows that, in an intermediate velocity range of about 1 and a half orders of magnitude, the velocity PDF can be approximated with the power-law $v^{\alpha-1}$ with $\alpha = 3/4$. Although no real power-law arises, this approximation is useful to understand the preasymptotic dispersion behavior that we observe in the intermediate regime, which could be mistaken for superdiffusion, as we discuss below. Figure D.1b shows the velocity PDF corresponding to the Gamma field. Here, we observe a clear power-law tail as $v^{\alpha-1}$ toward low velocities over several orders of magnitude with $\alpha = 0.58$. For a power-law $p_e(v) \propto v^{\alpha-1}$, the steady s-Lagrangian velocity PDF behaves as $p_s(v) \propto v^{\beta-1}$ with $\beta = \alpha + 1$. This implies that the transition probability $\psi(t)$ given by Equation (29) scales as $\psi(t) \propto t^{-1-\beta}$. The velocity distribution for $\sigma_f^2 = 0.1$ can be approximated by the lognormal distribution

$$p_e(v) = \frac{1}{v} \frac{\exp\left[-\log(v/\langle v_e\rangle)^2/2\sigma^2_{\ln v}\right]}{\sqrt{2\pi\sigma^2_{\ln v}}}. \tag{D.2}$$

In fact, the Darcy equation indicates that

$$v_e(\mathbf{x}) = K(\mathbf{x})|\nabla h(\mathbf{x})|. \tag{D.3}$$

Thus, for a slowly varying head gradient $v_e(\mathbf{x}) \propto K(\mathbf{x})$. Furthermore, the mean and variance of $v_e(\mathbf{x})$ are given in first order perturbation theory by $v_e = \overline{q}$ and $\sigma_v^2 = \sigma_{q_1}^2 = 3\sigma_f^2/8$ [*Dagan*, 1989], and $\sigma^2_{\ln v} = \langle v_e\rangle^2\sigma_v^2$.

The advective tortuosities are determined from the mean velocity magnitude and mean streamwise velocity according to Equation (19). Thus, we obtain for the advective tortuosities of the lognormal fields $\chi = 1$ for $\sigma_f^2 = 0.1$, $\chi = 1.06$ for $\sigma_f^2 = 1$, $\chi = 1.2$ for $\sigma_f^2 = 4$ and $\chi = 1.3$ for $\sigma_f^2 = 7$. The tortuosity for the Gamma field is $\chi = 1.2$.

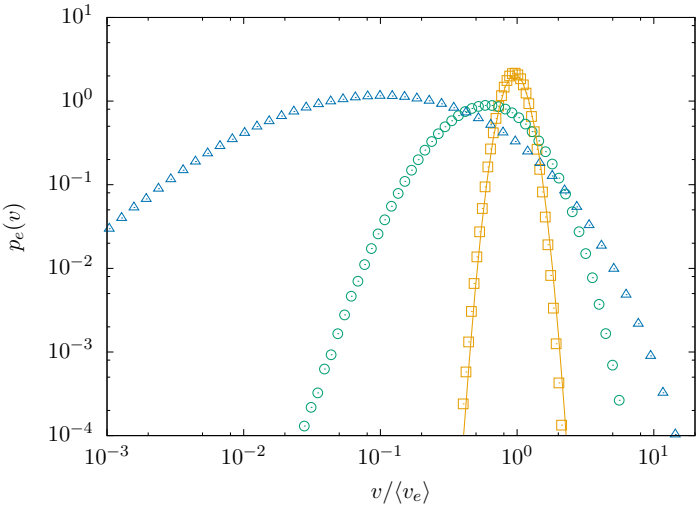

Figure D.2. The triangles show $p_e(v)$ for $\sigma_f^2 = 4$, the circles for $\sigma_f^2 = 1$ and the square for $\sigma_f^2 = 0.1$. The solid line denotes the lognormal distribution (D.2) with $\langle v_e\rangle = 1$ and $\sigma^2_{\ln v} = 3\sigma_f^2/8$.

E: Longitudinal dispersion in the Bernoulli model

We consider here the analytical expressions for dispersion coefficients in the Bernoulli model derived in *Dentz et al.* [2016]. These authors consider the dispersion of the particle position $s(t)$ along streamlines, which is given by

$$\mathcal{D}(t) = \frac{1}{2}\frac{d}{dt}\left[\langle s(t)^2\rangle - \langle s(t)\rangle^2\right]. \tag{E.1}$$

According to (20), $x_1(t) = s(t)/\chi$. Thus, the longitudinal dispersion coefficient $D_L(t)$ is given in terms of $\mathcal{D}(t)$ as

$$D_L(t) = \frac{\mathcal{D}(t)}{\chi^2}. \tag{E.2}$$

E.1 Anomalous Dispersion

Dentz et al. [2016] derive expressions for $\mathcal{D}(t)$ for the Gamma-distributed $p_e(v)$

$$p_e(v) = \frac{1}{v_c\Gamma(\alpha)}\left(\frac{v}{v_c}\right)^{\alpha-1}\exp\left(-\frac{v}{v_c}\right). \tag{E.3}$$

Under stationary conditions, this means for the uniform injection, they obtain for $t \gg \tau_v$

$$D_L(t) = \frac{\langle v_e\rangle \ell_c c\alpha}{\chi^2(1-\alpha)}(t/\tau_0)^{1-\alpha}, \tag{E.4}$$

where $\tau_0 = \ell_c/v_c$, $c = \Gamma(2-\alpha)/\Gamma(1-\alpha)$. For the non-stationary case, this means for the flux-weighted injection, $D_L(t)$ is given by

$$D_L(t) = \frac{\langle v_e\rangle^2 \tau_0^\alpha c\alpha}{\chi^2(1-\alpha)}t^{1-\alpha}. \tag{E.5}$$

This gives for the displacement variances

$$\kappa(t) = \begin{cases} 2c_\alpha\langle v_e\rangle \ell_c \tau_0^{\alpha-1}t^{2-\alpha} & \text{steady} \\ 2c_\alpha\langle v_e\rangle^2\tau_0^\alpha t^{2-\alpha} & \text{non-steady,}\end{cases} \tag{E.6}$$

where $c_\alpha = c\alpha/[(1-\alpha)(2-\alpha)\chi^2]$. Note that $\alpha = \beta - 1$.

E.2 Macrodispersion

Dentz et al. [2016] derived the following analytical expression for the dispersion of particle positions $s(t)$ along streamlines,

$$\mathcal{D} = \langle v_e \rangle \ell_c \left(\frac{\langle \tau_s^2 \rangle}{2\tau_v^2} - 1 \right), \tag{E.7}$$

where τ_s is distributed according to $\psi_s(t)$, which is given by

$$\psi_s(t) = \frac{1}{\tau_v} \int_0^\infty dv \exp(-tv/\ell_c) \frac{v p_s(v)}{\langle v_e \rangle} \tag{E.8}$$

where $\tau_v = \ell_c / \langle v_e \rangle$. Using relation (11) it can be written as

$$\psi_s(t) = \frac{1}{\tau_v} \int_0^\infty dv \exp(-tv/\ell_c) \frac{v^2 p_e(v)}{\langle v_e \rangle^2} \tag{E.9}$$

In order to obtain expression (45), we evaluate

$$\langle \tau_s^2 \rangle = \int_0^\infty dt t^2 \psi_s(t). \tag{E.10}$$

Using expression (E.9), we obtain

$$\langle \tau_s^2 \rangle = \frac{1}{\tau_v} \int_0^\infty dv \frac{2\ell_c^3}{v^3} \frac{v^2 p_e(v)}{\langle v_e \rangle^2} = \frac{2\ell_c^2}{\langle v_e \rangle} \int_0^\infty dv \frac{p_e(v)}{v} = \frac{2\ell_c^2}{\langle v_e \rangle v_H} = \frac{2\tau_v^2 \langle v_e \rangle}{v_H}, \tag{E.11}$$

where the harmonic mean of $p_e(v)$ is defined as

$$\frac{1}{v_H} = \int_0^\infty dv \frac{p_e(v)}{v}. \tag{E.12}$$

Inserting (E.11) into (E.7) and using (E.2) gives (45). We now show the consistence of this expression with stochastic perturbation theory. To this end, we first decompose v_e into its mean and fluctuation

$$v_e = \sqrt{q_1^2 2 + q_2^2} = \langle q \rangle + q_1', \tag{E.13}$$

where we only account for terms linear in the velocity fluctuation and note that $\langle q_2 \rangle = 0$. Using this decomposition in expression (45) and expanding up to first order in the velocity fluctuation gives

$$D_L^\infty \approx \langle q \rangle \ell_c \frac{\sigma_{q_1}^2}{\langle q \rangle^2}. \tag{E.14}$$

We note that in this approximation ℓ_c is equal to the correlation distance of the fluctuations q_1'. The velocity correlation ℓ_c and the velocity covariance are in leading order given by [*Dagan*, 1989; *Cvetkovic et al.*, 1996]

$$\ell_c = \frac{8}{3}\lambda, \qquad \sigma_{q_1}^2 = \frac{3}{8}\sigma_f^2. \tag{E.15}$$

Inserting this expressions into (E.14) gives expression (46) for the longitudinal macrodispersion coefficient in first order perturbation theory in σ_f^2.

References

Aarnes, J. E., T. Gimse, and K.-A. Lie (2007), An introduction to the numerics of flow in porous media using matlab, in *Geometric modelling, numerical simulation, and optimization*, pp. 265–306, Springer.

Adams, E. E., and L. W. Gelhar (1992), Field study of disperion in a heterogeneous aquifer 2. spatial moment analysis, *Water Resour. Res.*, *28*, 3293–3307.

Batchelor, G. K. (2000), *An Introduction to Fluid Dynamics*, Cambridge University Press.

Bear, J. (1972), *Dynamics of Fuids in Porous Media*, American Elsevier, New York.

Benke, R., and S. Painter (2003), Modeling conservative tracer transport in fracture networks with a hybrid approach based on the boltzmann transport equation, *Water Resour. Res.*, *39*, 1324, doi:10.1029/2003WR001966,.

Berkowitz, B., and H. Scher (1997), Anomalous transport in random fracture networks, *Phys. Rev. Lett.*, *79*(20), 4038–4041.

Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling non-fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, *44*, RG2003.

Boggs, J. M., S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt, and E. E. Adams (1992), Field study of dispersion in a heterogeneous aquifer 1. overview and site description, *Water Resour. Res.*, *28*, 3281–3291.

Comolli, A., and M. Dentz (2017), Anomalous dispersion in correlated porous media: a coupled continuous time random walk approach, *The European Physical Journal B*, *90*(9), 166.

Cvetkovic, V., H. Cheng, and X.-H. Wen (1996), Analysis of nonlinear effects on tracer migration in heterogeneous aquifers using Lagrangian travel time statistics, *Water Resour. Res.*, *32*(6), 1671–1680.

Cvetkovic, V., A. Fiori, and G. Dagan (2014), Solute transport in aquifers of arbitrary variability: A time-domain random walk formulation, *Water Resources Research*, *50*(7), 5759–5773.

Dagan, G. (1984), Solute transport in heterogenous porous formations, *J. Fluid Mech.*, *145*, 151–177.

Dagan, G. (1989), *Flow and transport in porous formations*, Springer, New York.

Dagan, G. (2017), Solute plumes mean velocity in aquifer transport: Impact of injection and detection modes, *Advances in Water Resources*, *106*, 6–10.

de Anna, P., T. Le Borgne, M. Dentz, A. M. Tartakovsky, D. Bolster, and P. Davy (2013), Flow intermittency, dispersion and correlated continuous time random walks in porous media, *Phys. Rev. Lett.*, *110*, 184,502.

de Dreuzy, J.-R., A. Beaudoin, and J. Erhel (2007), Asymptotic dispersion in 2d heterogeneous porous media determined by parallel numerical simulations, *Water Resources Research*, *43*(10).

Dentz, M., A. Cortis, H. Scher, and B. Berkowitz (2004), Time behavior of solute transport in heterogeneous media: transition from anomalous to normal transport, *Adv. Water Resour.*, *27*(2), 155–173.

Dentz, M., T. Le Borgne, A. Englert, and B. Bijeljic (2011), Mixing, spreading and reaction in heterogeenous media: a brief review, *J. Cont. Hydrol.*, *120-121*, 1–17.

Dentz, M., P. K. Kang, A. Comolli, T. Le Borgne, and D. R. Lester (2016), Continuous time random walks for the evolution of lagrangian velocities, *Physical Review Fluids*, *1*(7), 074,004.

Devroye, L. (1986), *Non-Uniform Random Variate Generation*, Springer, New York.

Domenico, P. A., and F. W. Schwartz (Eds.) (1997), *Physical and Chemical Hydrogeology*, Wiley.

Edery, Y., A. Guadagnini, H. Scher, and B. Berkowitz (2014), Origins of anomalous transport in heterogeenous media: Structural and dynamic control, *Water Resour*

Res, *50 (2)*, 1490 – 1505.

Edery, Y., S. Geiger, and B. Berkowitz (2016), Structural controls on anomalous transport in fractured porous rock, *Water Resources Research*, *52*(7), 5634–5643, doi:10.1002/2016wr018942.

Fiori, A., I. Janković, and G. Dagan (2003), Flow and transport in highly heterogeneous formations: 2. semianalytical results for isotropic media, *Water Resources Research*, *39*(9), doi:10.1029/2002wr001719.

Fiori, A., I. Jankovic, G. Dagan, and V. Cvetkovic (2007), Ergodic transport through aquifers of non-gaussian log conductivity distribution and occurence of anomalous behavior, *Water Resour. Res.*, *43*, W09,407.

Fiori, A., A. Zarlenga, I. Jankovic, and G. Dagan (2017), Solute transport in aquifers: The comeback of the advection dispersion equation and the first order approximation, *Advances in water resources*, *110*, 349–359.

Frippiat, C. C., and A. E. Holeyman (2008), A comparative review of upscaling methods for solute transport in heterogeneous porous media, *Journal of Hydrology*, *362*(1-2), 150–176, doi:10.1016/j.jhydrol.2008.08.015.

Gardiner, C. W. (1986), Handbook of stochastic methods for physics, chemistry and the natural sciences, *Applied Optics*, *25*, 3145.

Gelhar, L. W. (1993), *Stochastic subsurface hydrology*, Prentice Hall.

Gelhar, L. W., and C. L. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180.

Ghanbarian, B., A. Hunt, R. P. Ewing, and M. Sahimi (2013), Tortuosity in porous media: A critical review, *Soil Science Society of America Journal*, *77*, 1461.

Gómez-Hernández, J., and S. M. Gorelick (1989), Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge, *Water Resources Research*, *25*(3), 405–419.

Gotovac, H., V. Cvetkovic, and R. Andricevic (2009), Flow and travel time statistics in highly heterogeneous porous media, *Water resources research*, *45*(7).

Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late time behavior of tracer test breakthrough curves, *Water Resour. Res.*, *36*(12), 3467–3479.

Hakoun, V., A. Comolli, and M. Dentz (2019), Upscaling and prediction of lagrangian velocity dynamics in heterogeneous porous media, *Water Resour. Res.*, *55*, 10.1029/2018WR023,810.

Harvey, C., and S. M. Gorelick (2000), Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (made) site?, *Water Resources Research*, *36*(3), 637–650.

Haslauer, C., P. Guthke, A. Bárdossy, and E. Sudicky (2012), Effects of non-gaussian copula-based hydraulic conductivity fields on macrodispersion, *Water Resources Research*, *48*(7).

Holzner, M., V. L. Morales, M. Willmann, and M. Dentz (2015), Intermittent lagrangian velocities and accelerations in three-dimensional porous medium flow, *Phys. Rev. E*, *92*, 013,015.

Hyman, J. D., S. L. Painter, H. Viswanathan, N. Makedonska, and S. Karra (2015), Influence of injection mode on transport properties in kilometer-scale three-dimensional discrete fracture networks, *Water Resour. Res.*

Hyman, J. D., M. Dentz, A. Hagberg, and P. K. Kang (2019), Linking structural and transport properties in three-dimensional fracture networks, *Journal of Geophysical Research: Solid Earth*, doi:10.1029/2018jb016553.

Jankovic, I., A. Fiori, and G. Dagan (2003), Effective conductivity of an isotropic heterogeneous medium of lognormal conductivity distribution, *Multiscale Modeling & Simulation*, *1*(1), 40–56.

Kang, P. K., P. de Anna, J. Nunes, B. Bijeljic, M. J. Blunt, and R. Juanes (2014), Pore-scale intermittent velocity structure underpinning anomalous transport through 3-d porous media, *Geophys. Res. Lett.*, *41 (17)*, 6184–6190.

Kang, P. K., T. Le Borgne, M. Dentz, O. Bour, and R. Juanes (2015), Impact of velocity correlation and distribution on transport in fractured media: field evidence and theoretical model, *Water Resour. Res.*, *51*, 940–959.

Kang, P. K., M. Dentz, T. Le Borgne, S. Lee, and R. Juanes (2017), Anomalous transport in disordered fracture networks: Spatial markov model for dispersion with variable injection modes, *Adv. Water Resour.*, doi:10.1016/j.advwatres.2017.03.024.

Koponen, A., M. Kataja, and J. Timonen (1996), Tortuous flow in porous media, *Physical Review E*, *54*(1), 406.

Le Borgne, T., J. R. de Dreuzy, P. Davy, and O. Bour (2007), Characterization of the velocity field organization in heterogeneous media by conditional correlation, *Water Resour. Res.*, *43*, 2006WR004,875.

Le Borgne, T., M. Dentz, and J. Carrera (2008), Spatial markov processes for modeling lagrangian particle dynamics in heterogeneous porous media, *Phys. Rev. E*, *78*, 041,110.

Levy, M., and B. Berkowitz (2003), Measurement and analysis of non-fickian dispersion in heterogeneous porous media, *Journal of Contaminant Hydrology*, *64*(3-4), 203–226.

Meyer, D. W., and H. A. Tchelepi (2010), Particle-based transport model with Markovian velocity processes for tracer dispersion in highly heterogeneous porous media, *Water Resources Research*, *46*(11), doi:10.1029/2009wr008925.

Morales, V. L., M. Dentz, M. Willmann, and M. Holzner (2017), Stochastic dynamics of intermittent pore-scale particle motion in three-dimensional porous media: Experiments and theory, *Geophysical Research Letters*, *44*(18), 9361–9371.

Neuman, S. P., and D. M. Tartakovsky (2008), Perspective on theories of anomalous transport in heterogeneous media, *Adv. Water Resour.*, *32*, 670–680, doi:10.1016/j.advwatres.2008.08.005.

Niemi, A., J. Bear, and J. Bensabat (Eds.) (2017), *Geological Storage of CO2 in Deep Saline Formations*, Springer Netherlands.

Noetinger, B., D. Roubinet, A. Russian, T. Le Borgne, F. Delay, M. Dentz, J.-R. De Dreuzy, and P. Gouze (2016), Random walk methods for modeling hydrodynamic transport in porous and fractured media from pore to reservoir scale, *Transport in Porous Media*, pp. 1–41.

Painter, S. (1996), Evidence for non-gaussian scaling behavior in heterogeneous sedimentary formations, *Water Resources Research*, *32*(5), 1183–1195.

Painter, S., and V. Cvetkovic (2005), Upscaling discrete fracture network simulations: An alternative to continuum transport models, *Water Resources Research*, *41*(2).

Poinssot, C., and H. Geckeis (Eds.) (2012), *Radionuclide Behaviour in the Natural Environment*, Woodhead Publishing.

Pollock, D. (1988), Semianalytical computation of path lines for finite-difference models, *Ground Water*, *26 (6)*, 743–750.

Puyguiraud, A., P. Gouze, and M. Dentz (2019), Upscaling of anomalous pore-scale dispersion, *Transport in Porous Media*, *128*(2), 837–855, doi:10.1007/s11242-019-01273-3.

Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 3. geostatistical analysis of hydraulic conductivity, *Water Resources Research*, *28*(12), 3309–3324.

Risken, H. (1996), *The Fokker-Planck Equation*, Springer Heidelberg New York.

Rubin, Y. (2003), *Applied stochastic hydrogeology*, Oxford University Press, New York.

Russian, A., M. Dentz, and P. Gouze (2016), Time domain random walks for hydrodynamic transport in heterogeneous media, *Water Resources Research*.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera (2006a), Representative hydraulic conductivities in saturated groundwater flow, *Reviews of Geophysics*, *44*(3), n/a–n/a, doi:10.1029/2005RG000169.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera (2006b), Representative hydraulic conductivities in saturated groundwater flows, *Rev. Geophys.*, *44*, RG3002.

Schlather, M., A. Malinowski, P. J. Menck, M. Oesting, K. Strokorb, et al. (2015), Analysis, simulation and prediction of multivariate random fields with package randomfields, *Journal of Statistical Software*, *63*(8), 1–25.

Shlesinger, M. (1974), Asymptotic solutions of continuous-time random walks, *J. Stat. Phys.*, *10 (5)*, 421–434.

Tyukhova, A., M. Dentz, W. Kinzelbach, and M. Willmann (2016), Mechanisms of anomalous dispersion in flow through heterogeneous porous media, *Physical Review Fluids*, *1*(7), 074,002.

Zech, A., C. D'Angelo, S. Attinger, and A. Fiori (2018), Revisitation of the dipole tracer test for heterogeneous porous formations, *Advances in water resources*, *115*, 198–206.

Zheng, C., M. Bianchi, and S. M. Gorelick (2011), Lessons learned from 25 years of research at the made site, *Groundwater*, *49*(5), 649–662.