Upscaling transport of a sorbing solute in disordered non periodic porous domains

Thomas Sherman[a], Emanuela Bianchi Janetti[b], Gaël Raymond Guédon[c], Giovanni Porta[b], Diogo Bolster[a]

[a]*Department of Civil & Environmental Engineering & Earth Sciences, University of Notre Dame*
[b]*Dipartimento di Ingegneria Civile ed Ambientale, Politecnico di Milano*
[c]*Dipartimento di Energia, Politecnico di Milano*

Abstract

Spatial Markov random walk models (SMM) have been demonstrated to accurately predict conservative solute transport across a wide range of hydrogeological systems, with recent advances enabling the SMM to model systems with linear kinetic reactive processes. However, the proposed reactive SMM's applicability is limited to systems that can be partitioned into a series of identical periodic cells where travel times across cells are highly correlated to the solute's entrance position at the cell inlet. In real geologic settings, the spatial layout and size of grains varies through space, decorrelating the relationship between travel time and transverse position. Here, we generalize previous SMM implementations and implement a Bernoulli CTRW, where transport behavior can be captured in disordered and non periodic porous media. We validate our upscaled model predictions with results from direct numerical simulation of transport in a 2D porous column that cannot be partitioned into identical periodic elements. We parameterize our model based on a subset of simulation statistics and explore how model accuracy changes due to our sampling method. We find that for a given sampling area, model predictions improve and uncertainty across realizations decreases when many small sampling slices distributed uniformly and randomly throughout the domain are used for parameterization, as opposed to fewer large continuous sampling areas. This finding yields important insights for optimizing efficiency of the upscaled transport model parameterization and can guide field sampling of geological structures as well as multiscale investigation of laboratory observations.

Keywords: Reactive Transport, Upscaled Modeling, Statistical Sampling

1. Introduction

In natural subsurface porous media settings, the spatial layout and broad distribution of solid-phase grain geometries and irregularity of grain shapes gives rise to a complex flow setting. Specifically, the fluid velocity field is highly heterogeneous and characterized by high velocity preferential channels that rapidly transport dissolved solute over larger distances as well as grain cavities and dead end pores that result in stagnant waters that delays downstream transport. The velocity distribution can span several orders of magnitude, which ultimately manifests as anomalous i.e., Non-Fickian, behavior at typical experimental observation scales [1, 2, 3, 4]. Understanding solute transport becomes even more complex when mass can exchange between the fluid and solid phase via chemical reactions, i.e. when solute adsorbs and desorbs to the solid grain surface, as demonstrated by [5, 6, 7, 8]. Such chemical interactions further delay transport relative to the case of a non reactive solute and impact larger scale transport behavior [9]. Capturing such complexity and anomalous behavior in mathematical modelling frameworks is challenging, but remains critical for many scientific and engineering applications, including contaminant transport in aquifers [10], biological membranes [11], and packed bed reactors [12].

High fidelity direct numerical simulations (DNS) are a common method for simulating flow and transport through porous media [13, 14, 15, 16]. In DNS, the flow field, porous medium geometry, and evolution of a solute plume as it traverses space-time are all fully resolved. An advantage of DNS is that physical mechanisms that drive transport can be studied separately and are well described, providing an in depth understanding of the hydro-geologic system and associated transport processes. However, such simulations demand large computational resources and so applying DNS at the field scale is practically impossible. Hence for studies at the field scale, upscaled models, which effectively represent physical processes without explicitly resolving them, provide an attractive alternative.

A variety of upscaled models have emerged to capture anomalous transport in subsurface hydrological systems, including among others multi-rate

Email address: tsherma3@nd.edu (Thomas Sherman)

mass transfer (MRMT) [17], fractional advection dispersion [18], continuous time random walk (CTRW) [19]and time domain random walk (TDRW) models [20]. Here, we consider the CTRW modelling framework. In a CTRW, a solute plume is typically conceptualized as a large ensemble of particles who transition through time and space by sampling a waiting time/velocity distribution. The particle's velocity series can be modeled as a spatial Markov process, where at each step a particle traverses a fixed streamwise distance and its associated travel time is modeled with an effective travel time distribution parameter. Successive model jumps are assumed independent when the jump distance exceeds a velocity correlation length scale [21, 19] and correlated at smaller scales [22, 23, 24]. When jumps are sufficiently small such that correlation effects are important, the velocity correlation structure is often imposed by assuming a velocity transition matrix that is stationary in space [22, 25, 26] or assuming velocity transitions via a Bernoulli process [27, 28, 29, 30, 31, 32]. CTRW models have demonstrated success in a variety of porous and fractured media studies [33, 24, 34], but often are applied in synthetic settings where the input correlation parameters can be more easily calculated.

To date, the majority of CTRW studies have focused on conservative transport. Recently, CTRW models that consider reactive transport have also emerged [35, 36, 37, 38, 39, 40]. Here, we focus on upscaling reactive transport where solute mass can exchange between the fluid and solid grain boundaries via a chemical reaction. We build upon a previous study where the SMM with adsorption and desorption was considered in an idealized periodic wavy channel [39]. In this modelling approach, a particle's inlet position at the periodic channel's throat informs the particle's random travel time across the cell, the particle's outlet position (which informs the next step in the Markov chain), and the number of times a particle strikes the boundary, which subsequently controls adsorption-desorption processes. The approach was successful in the idealized wavy channel considered, because the travel time statistics in each periodic cell are identical. Hence, the proposed model by [39] can be fully parameterized through direct numerical simulation of transport through one periodic cell. In more complex porous media geometries, the spatially varying distribution of grains means the relationship between a particle's trajectory statistics and a particle's transverse position are no longer spatially stationary and thus transport behavior is further complicated in more realistic heterogeneous geologic settings. The adsorption-desorption SMM therefore must be generalized, if to be applied

in realistic non periodic domains.

This study has two primary objectives. 1) We extend the adsorption-desorption SMM so that upscaled particle trajectories are not dependent on transverse position. To do so, we introduce a Bernoulli SMM and relate a tortuosity parameter, which describes particle motion in the transverse direction, to the number of boundary hits. When tortuosity is high, particles have increased pathline travel distances and are therefore more likely to contact the grain boundaries and react. 2) The second objective is to investigate the impact of sampled statistics on model parameterization and uncertainty. SMM parameterization has to date been achieved empirically through direct numerical simulations [22, 15, 41, 42], inverse modelling algorithms [43], and parametric analytical Markov models [44, 27, 45]. Here we focus on parameterization via numerical simulations, as these allow transport to be quantified in great detail and enable a full exploration of the parameter space. Specifically, we sample the studied domain uniformly, but randomly and parameterize the CTRW; then we observe how for a given sampling area, the spatial distribution and size of sample slices influences CTRW predictions. We find that CTRW models parameterized via many small sampling slices distributed through space improves model prediction capability and reduces uncertainty in predictions. Here the studied domain is a stochastically generated 2D porous media column where particle trajectories vary spatially. Although still highly idealized, the spatial heterogeneity of the porous medium structure makes this system considerably more complicated than previously studied periodic domains and the proposed CTRW framework provides a step forward towards modelling adsorption-desorption transport in such complex systems.

2. Pore-scale Setting and Observables

This section describes methodology employed to model pore-scale solute transport, as well as providing the definitions for the observables that are later used in sections 3 and 4 within upscaling. Our pore-scale simulations rely on an artificially generated porous geometry and the related flow field (for more details see Appendix Appendix A.1-Appendix A.2).

2.1. Pore-Scale Transport Simulation

We model the evolution of the solute plume through the porous media column with sorbing boundaries as:

$$\frac{\partial c(\mathbf{x},t)}{\partial t} + \nabla \cdot [\mathbf{u}(\mathbf{x})c(\mathbf{x},t)] = \nabla \cdot [D\nabla c(\mathbf{x},\mathbf{t})] \quad \forall\ \mathbf{x} \in \Gamma_{fluid} \tag{1}$$

$$\frac{\partial S(\mathbf{x},t)}{\partial t} = -\lambda S(\mathbf{x},t) + \alpha c(\mathbf{x},\mathbf{t}) = -D\frac{\partial c}{\partial \hat{n}} \quad \forall\ \mathbf{x} \in \Sigma_{surface} \tag{2}$$

The solute concentration field $c(\mathbf{x},t)$ evolves through pore space Γ_{fluid} and time due to the fluid velocity $\mathbf{u}(\mathbf{x})$ and molecular diffusion D, i.e. solute transport through the fluid is governed by the advection-diffusion equation. The fluid velocity is here computed numerically as detailed in Appendix A.2. Solute concentration on the solid surface $\Sigma_{surface}$, labeled $S(\mathbf{x},t)$, changes as solute mass exchanges between the fluid and solid phases, with desorption rate λ and adsorption rate α. Hence the rate of mass exchange between the solid and fluid phases is simply the difference between the adsorption and desorption rates, which mass conservation requires to be equivalent to the diffusive flux of solute at the boundary $-D\frac{\partial c}{\partial \hat{n}}$, where $\hat{n}$ is the unit normal at the boundary. In the numerical experiments considered here, solute enters the column inlet via a flux weighted, pulse injection, i.e. the solute concentration injected at the inlet is proportional to the velocity.

We solve equations (1,2) with a Lagrangian approach. The solute plume is discretized into many particles who transition through space and time according to a Langevin equation:

$$x_i^{n+1} = x_i^n + u_i\Delta t + \eta_i\sqrt{2D\Delta t} \tag{3}$$

$$y_i^{n+1} = y_i^n + v_i\Delta t + \xi_i\sqrt{2D\Delta t} \tag{4}$$

$$t_i^{n+1} = t_i^n + \Delta t + T_i \tag{5}$$

where particle i has position $[x,y]$ and velocity $[u,v]$ at model step n. Diffusion enables particles to deviate from velocity streamlines, which here is modeled as a Brownian process; η, ξ are independent, identically distributed random variables sampled from a standard normal distribution (zero mean and unit variance). Additionally, diffusion enables particles to strike the solid phase boundary. When such interaction occurs, a particle attaches to the solid phase with sorbing probability P_{sorb}. If sorption occurs, the particle remains attached for time T, where T is a waiting time sampled from an exponential distribution with mean λ. If no sorption occurs, which is always the case for conservative transport, elastic boundary conditions

and a zero waiting time $T = 0$ are imposed. We elaborate on details related to numerical implementation of adsorption-desorption processes in the following section 2.1.1.

2.1.1. Adsorption-Desorption Processes

We model adsorption and desorption processes following [46]. Every instance a particle interacts with the solid phase boundary via diffusion, it sorbs to the boundary with a probability P_{sorb}. To leading order, for sufficiently small time step, P_{sorb} is given by

$$P_{sorb} = \alpha\sqrt{\frac{\pi\Delta t}{D}} \tag{6}$$

where α is a sorption rate, D is a diffusion coefficient, and Δt is the simulation time step size. When a particle contacts the solid phase, we draw a random number U_i from a uniform distribution $U(0,1)$ and the particle sorbs if $U_i < P_{sorb}$ and elastically reflects if $U_i \geq P_{sorb}$.

If a particle adsorbs to the boundary, its waiting time T_i is then sampled from an exponential function with mean λ, i.e. $\psi(T) = \lambda\exp(-\lambda T)$:

$$T_i = \begin{cases} \psi(T) & U_i < P_{sorb} \\ 0 & U_i \geq P_{sorb} \end{cases} \tag{7}$$

We characterize the importance of adsorption-desorption processes through the porous media column with dimensionless Damköhler numbers, Da_a, Da_d, which compare adsorptive and desorptive timescales with diffusive time scales. Here, high fidelity simulations are run for a range of Da_a and Da_d spanning several orders of magnitude $O(10 - 10^3)$, allowing us to study the evolution of the solute plume as adsorption (desorption) processes become increasingly important. Respectively, we define the adsorptive and desorptive Damköhler numbers as

$$Da_a = \frac{\ell_c\alpha\sqrt{\pi}}{D}, \quad Da_d = \frac{\ell_c^2\lambda}{D}. \tag{8}$$

Here ℓ_c is a characteristic length quantifying the average distance between the fluid phase with a solid boundary, which is determined by the domain geometry. Here $\ell_c = 3.8 \times 10^{-5}$m.

Similarly, we investigate how reactive transport behavior changes as diffusion strength increases. We quantify the relative strength of diffusion with a Péclet number, which gives the ratio of diffusive to advective transport

time scales. The Péclet number Pe depends on the mean fluid velocity (accounting for x and y directions) $||\bar{u}||$, D, and ℓ_c and is given as

$$Pe = \frac{||\bar{u}||\ell_c}{D} \tag{9}$$

We tune D, α, λ in order to change dimensionless quantities.

2.2. Transport Statistics

One of the objectives of this study is to parameterize an upscaled model that accurately predicts transport through the porous column, while minimizing the sampling area from which upscaled parameters are derived. To do so, we sample Lagrangian statistics along particle trajectories from the high resolution simulations and quantify how particle ensemble statistics evolve through space and time in detail. Then we parameterize the upscaled model with only a subset of the observed data. Specifically, we install control planes perpendicular to the mean imposed pressure gradient and measure transport statistics at the first crossing of each control plane. Control planes are spaced over a distance 2.04×10^{-5} m, which is sufficiently ($O(10)$ times) smaller than a velocity correlation length scale, such that a particle's velocity evolution can be faithfully portrayed.

Here, the solute plume is comprised of many particles ($\mathbf{b}$) with a collective mass M; the ensemble of particles is denoted Ω_b. Below we detail the employed parameters, i.e. the probability density functions (pdfs) of particle velocities, travel times, paths' tortuosity and the velocity correlation function.

2.2.1. Velocity pdf

At time t, a particle's velocity depends on its position $\mathbf{x}$; we define a particle's instantaneous Lagrangian velocity $\mathbf{v}_x$ as the fluid velocity at $\mathbf{x}$, $\mathbf{v}_x(\mathbf{x}; \mathbf{b}) = \mathbf{u}(\mathbf{x}, \mathbf{b})$. Lagrangian velocities are sampled for particle first crossings of control planes, meaning all points on a given control plane have the same x coordinate. For a control plane located at $x = x_i$ we construct the Lagrangian velocity distribution, $\psi(v, x_i)$

$$\psi(v, x_i) = \frac{1}{M} \int_{\Omega_b} d\mathbf{b} \delta[v - \mathbf{v}_x(x_i; \mathbf{b})] \tag{10}$$

Note here that the $\psi(v, x_i)$ distribution is spatially dependent as it depends on x_i.

2.2.2. Travel Time pdf

Closely related to velocity is the travel time distribution. Here travel time over two control planes refers to the elapsed time between first particle breakthrough at each control plane. Define the first arrival time τ_{cp} of a particle at a control plane at $x = x_i$ as

$$\tau_{cp}(x_i, \mathbf{b}) = t[\Lambda(x_i), \mathbf{b}], \qquad \Lambda(x_i) = \inf[x | x_i(\mathbf{b}) \geq x_i]. \tag{11}$$

The first arrival time of a particle is simply the first instance in which its x position exceeds x_i. The distribution of first arrival times at a control plane, denoted in this study as a breakthrough curve, is defined

$$\psi(t; x_i) = \frac{1}{M} \int_{\Omega_b} d\mathbf{b} \delta[t - \tau_{cp}(x_i; \mathbf{b})] \tag{12}$$

Tracking particle breakthrough times enables us to calculate the elapsed travel time $\Delta\tau_{cp}(x_{i,j})$ between any two control planes at x_i and x_j ($x_j < x_i$) as $\Delta\tau_{cp} = \tau_{cp}(x_i; \mathbf{b}) - \tau_{cp}(x_j; \mathbf{b})$. Then for any spatial slice $x_{i,j}$ spanning between x_j and x_i we define the distribution of travel times for the ensemble of particles:

$$\psi(\Delta\tau_{cp}(x_{i,j})) = \frac{1}{M} \int_{\Omega_b} d\mathbf{b} \delta[\Delta\tau_{cp} - \Delta\tau_{cp}(x_{i,j}; \mathbf{b})] \tag{13}$$

Note when $x_j = 0$, the travel time and breakthrough time distributions for x_i are equivalent, $\psi(\Delta\tau_{cp}(x_{i,0})) = \psi(t; x_i)$.

2.2.3. Tortuosity pdf

The irregularity and distribution of the solid phase creates complex flow features resulting in streamlines not aligned with the primary pressure gradient, i.e. streamlines are not straight, but meander also in the transverse direction. Furthermore, diffusion allows solute to sample many streamlines. As a result, a particle's total travel distance is greater than the linear distance in the streamwise direction. Tortuosity quantifies the ratio of the total pathline distance to the linear streamwise travel distance, and this is closely correlated with travel time.

The total travel distance Δs between successive model steps, n and $n+1$, in the high resolution simulation is estimated with a Euclidean distance, $\Delta\mathbf{s}_{n,n+1} = \|\mathbf{x}^{n+1} - \mathbf{x}^n\|$. It follows that the total pathline distance $\mathbf{s}_{n,n+k}$ from model step n to $n + k$ is the summation of distances for each model

step, $\mathbf{s}_{n,n+k} = \sum_{i=0}^{k-1} \Delta\mathbf{s}_{n+i,n+i+1}$. Then a particle that first crosses a control plane at $x = x_j$ at model step n and first crosses a control plane $x = x_i$ at model step $n+k$ has a corresponding travel distance between control planes of $s(x_{i,j}) = s_{n,n+k}$.

Note in the literature, there are various definitions of tortuosity that have been used to study different subsurface properties (see [47]). Here, we adopt a pathline-dependent effective tortuosity that aligns with Lagrangian observations and can naturally be included as an upscaled model parameter. A particle's effective tortuosity $\chi(x_{i,j})$ between control planes x_i and x_j is the total pathline travel distance $s(x_{i,j})$ divided by the linear distance between control planes $\Delta x_{i,j} = |x_i - x_j|$:

$$\chi(x_{i,j}) = \frac{s(x_{i,j})}{\Delta x_{i,j}}. \tag{14}$$

The effective tortuosity between any two control planes at x_i and x_j is defined as

$$\psi(\chi_{i,j}) = \frac{1}{M} \int_{\Omega_b} d\mathbf{b}\delta[\chi_{i,j} - \chi_{i,j}(\mathbf{b})] \tag{15}$$

For any slice spanning control planes x_i and x_j, we can define a local mean effective tortuosity $\langle\chi_{i,j}\rangle^L$, where angle brackets denote an arithmetic average across the particle ensemble.

2.2.4. Velocity Correlation

Recall, the travel time and tortuosity distribution are defined over a slice spanning two control planes. We space control planes relative to a velocity correlation length scale, which is calculated from velocity correlation function

$$C_{vv}(s) = \frac{1}{N} \sum_{n=1}^{N} \int_0^\infty ds' \frac{v'(s')v'(s'+s)}{\sigma_{vv}^2}, \tag{16}$$

where v' refers to a velocity fluctuation about a mean, σ_{vv}^2 is the velocity variance, s is the spatial lag distance, and N is the number of particles (allowing for the average across the particle ensemble).

The velocity correlation length scale l_c is then the integral of the correlation function:

$$l_c = \int_0^\infty C_{vv}(s)ds \tag{17}$$

The velocity correlation length scales are $1.51, 1.89, 1.87 \times 10^{-4}$ m for $Pe = 100, 1000$, and ∞, respectively, demonstrating that velocity correlation features more rapidly diminish with increasing diffusive strength. The entire

porous media column is approximately $O(100)$, $O(10)$ times the velocity correlation scale along the longitudinal and transverse directions, respectively.

3. Flow and Transport Observations

In this section we investigate the spatial variation of Eulerian and Lagrangian flow and transport properties. We partition our domain into non-overlapping, equally-sized slices and measure the spatial evolution of transport statistics. Control planes are spaced at distances of 0.1, 1, and 10 times the correlation length scale, l_c. Here we highlight distributions for control planes spaced 0.1 l_c and $Pe = 1000$.

3.1. Eulerian Observations

The top three rows in Figure 1 show Eulerian characteristics of the column, i.e., the distribution of grains, porosity, and the Eulerian velocity field. The phase field displays a disordered structure. The grain geometries (shown in blue) are highly irregular and each grain has a unique shape and size. The porosity ϕ in each sample slice, defined as the fluid area divided by the entire sample area $\phi = A_{fluid}/A_{sample}$ varies through space with values ranging within [0.23, 0.97]; the entire column has a mean porosity of 0.64.

The fluid velocity field is characterized by a broad distribution spanning over 10 orders of magnitude, where high velocity preferential pathways connect and cut through the surrounding slow moving waters. Relatively high velocity channels form when the pore space gap between grains constricts, due to the incompressibilty constraint. For this reason high velocity channels form in areas of low porosity. The relatively small longitudinal sample size of the sample facilitates the occurrence of channelling [16], whose occurrence is documented in three-dimensional media [48].

3.2. Lagrangian Observations

Here, we show how Lagrangian particle statistics are influenced by the underlying flow field structure.

3.2.1. Spatial Evolution of Lagrangian pdfs

Figure 1 displays the spatial evolution of the travel time, tortuosity, and velocity pdfs for control planes spaced by distance $0.1l_c$. We immediately observe that the Eulerian flow field structure directly influences Lagrangian statistics. The transverse position breakthrough position distribution (TBPD) through space highlights where particles cross each control

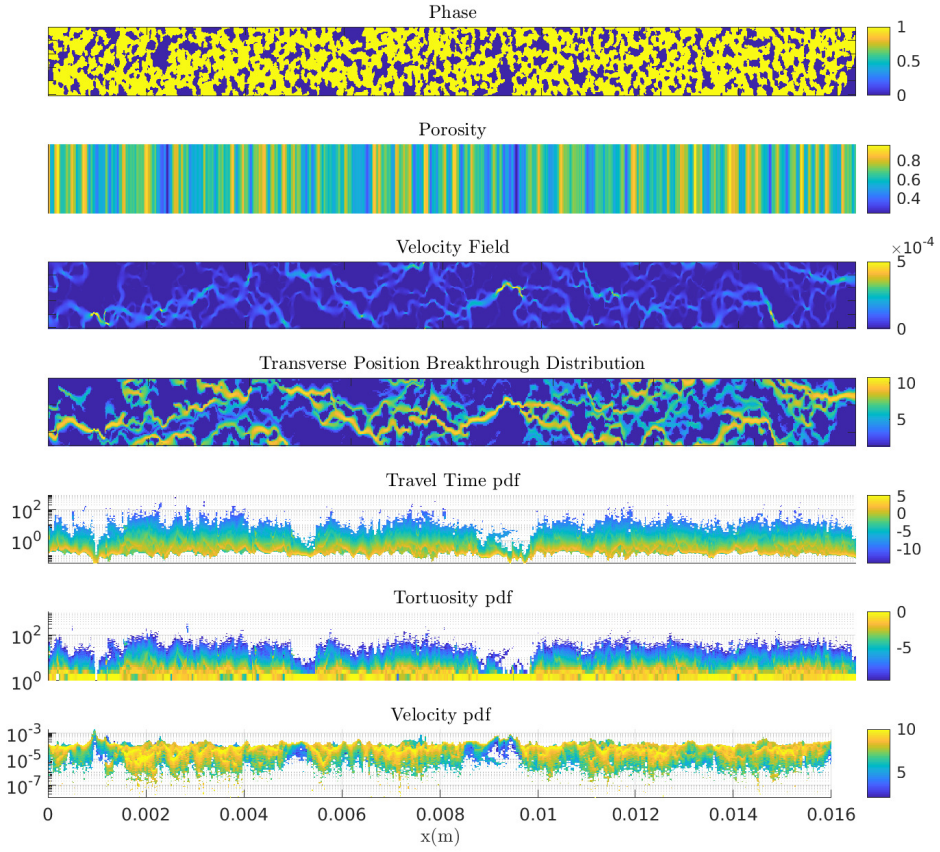

Figure 1: Eulerian and Lagrangian statistics of the 2D porous media column. Subfigure 1 shows the fluid (yellow) and solid (blue) phases. Subfigure 2 displays the porosity through space. Subfigure 3 shows the modulus of flow velocity. Subfigure 4 shows the transverse breakthrough position through space with colors in log probability. Subfigures 5,6, and 7 show the spatial evolution of the Lagrangian travel time, tortuosity, and velocity distributions with the colors corresponding to log probability. Distributions are sampled with spatial frequency $0.1l_c$, subfigures 4-7 refer to transport of a nonreactive tracer upon setting $Pe = 1000$.

plane in the transverse direction; colors correspond to log probabilities. The largest TBPD values correspond to regions of high fluid velocity, demonstrating that mass is channelized in high velocity pathways. As a result of such channelization, the Lagrangian velocity pdf transitions through space as observed in [32, 45].

When the velocity field displays strong dominant channeling features, such as near $x = 0.001, 0.009$ the statistics of the travel time, tortuosity, and velocity homogenize, with the width of the pdfs decreasing. Such behavior is expected because mass preferentially enters fast channels, meaning all channelized mass shares similar trajectories. These regions are made up of relatively fast velocity zones, which results in faster mean travel times. Consequently, the Lagrangian velocity and travel time distribution are negatively correlated, i.e. as mean velocity increases, mean travel time decreases because particles are traveling faster. Similarly, the mean tortuosity decreases with increasing channelization, as channelization is associated with high velocity pathways. The fastest channels tend to be aligned with x, meaning particles that enter such channels will have local tortuosity values closer to 1. In regions of slow flow, diffusion becomes increasingly important for particle trajectories, which increases tortuosity values on average. The largest tortuosity values are $O(100)$, meaning the total particle travel distance is two orders of magnitude greater than the x linear distance. These large values can occur near pore cavities, where the velocity field can advect particles counter to the primary pressure gradient and cause particles to recirculate through slow waters. Hence particles travel relatively long distances at slow velocities, which results in large travel times that manifest in late breakthough tailing behavior.

3.2.2. Solid-Fluid Phase Interaction

In addition to the aforementioned Lagrangian statistics, we also track the number of times a particle hits a boundary from inlet to outlet. Recall that adsorption to the solid phase boundary only occurs when diffusion brings a particle into contact with the solid phase. If a particle adsorbs, then it waits according to an exponential distribution, which delays transport (see section 2.1.1). The number of times a particle hits the solid phase is therefore important for the reactive case.

Figure 2 shows the joint pdf between number of hits and tortuosity; colors correspond to log probabilities. Here hits and tortuosity are measured over control planes with spacing $0.1l_c$. The mean number of hits tends to increase with tortuosity, demonstrating that particles are more likely to interact with the solid phase as particle travel distance increases. This has important implications for adsorption/desorption of particles, as it implies that slow particles (with high tortuosity) preferentially adsorb and are delayed. In contrast, fast particles preferentially transport via high velocity channels and therefore have less interaction and are less likely to adsorb to

the solid boundary on average.

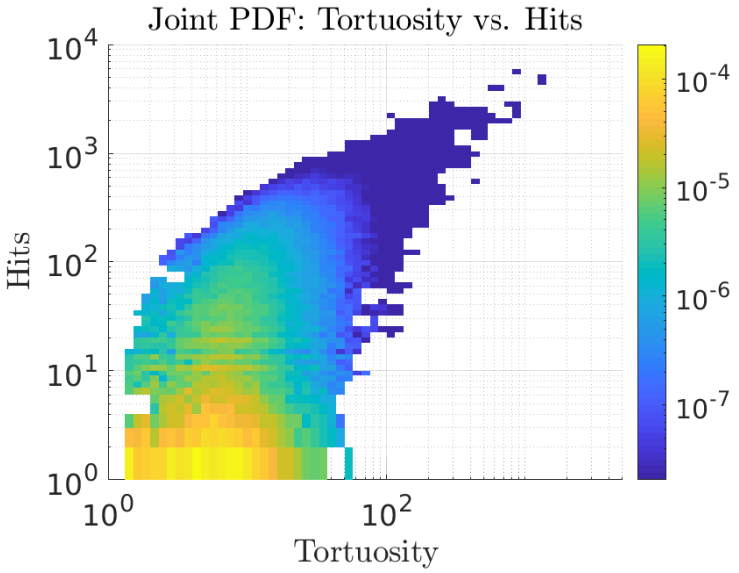

Figure 2: The joint pdf of tortuosity and the number of times a particle hits a solid boundary. Tortuosity statistics are shown for control planes spaced $\ell = .1l_c$. Colors show log probabilities. The majority of probability mass is concentrated at less than 10 hits. However, if a particle hits the boundary a large number of times $hits > 100$, then the tortuosity is very large, i.e. if particles travel very long distances, they are more likely to hit the boundary many times.

4. Upscaled Stochastic Modeling: Bernoulli CTRW

This study's objective is to predict transport at the column scale through an upscaled model parameterized using a limited portion of the 2D porous media column. We also investigate how the spatial distribution and area of sample slices influences upscaled model parameterization and prediction. These objectives are studied in the context of a continuous time random walk (CTRW), which is a Lagrangian, stochastic framework that naturally aligns with particle tracking methods applied in the high resolution numerical simulations described in section 2.1.

In a CTRW framework, the solute plume is conceptualized as many discrete tracer particles that transition through time and space according to probabilistic rules. As formulated in this study, particles traverse a fixed distance ℓ at every model step with an associated travel time τ, which

is sampled from a travel time distribution $\psi(\tau)$. To upscale adsorption-desorption processes, at every CTRW step we model the number of times a particle hits the solid boundary and each hit has an associated waiting time sampled from a distribution $\psi(\tau_{sorb})$, here set as an exponential distribution. Upscaled particle trajectories follow a Langevin equation:

$$x_i^{n+1} = x_i^n + \ell \tag{18}$$

$$t_i^{n+1} = t_i^n + \tau_i^{n+1} + \sum_{j=1}^{N_{hits}^k} \tau_{sorb}^j \tag{19}$$

At every model step, particles jump a fixed distance ℓ in time τ, which is randomly sampled from a travel time distribution $\psi(\tau)$. The cumulative waiting time $\sum_{j=1}^{N_{hits}^k} \tau_{sorb}^j$ delays particle transport due to adsorption/desorption and equals zero for the conservative case. Further details on implementation are discussed in the following sections.

4.1. Implementation of CTRW Correlation Structures

4.1.1. Velocity Correlation

It has been well documented that strong correlation properties exists in porous and fractured media, which may significantly impact solute transport behavior at typical scales of interest [22, 14, 25, 15, 28]. As a result, for scales less than the velocity correlation length, fast (slow) particles preferentially remain fast (slow). Thus when a CTRW jump distance is significantly less than the velocity correlation length scale $\ell << l_c$, velocity correlation needs to be captured to faithfully portray transport behavior.

The Bernoulli CTRW model, which we will apply in this study, is a spatial Markov model (SMM) that accounts for velocity correlation in a parsimonious, simplistic framework and was first introduced by Dentz et al. [27]. In the Bernoulli CTRW framework, a particle's velocity transitions via a Bernoulli process; that is the particle persists with the same travel time as the previous model step with probability P and samples from a global distribution otherwise:

$$\tau_i^{n+1} = \begin{cases} \tau_i^n & P \\ \psi(\tau) & 1-P \end{cases} \tag{20}$$

We assume that particle velocity transitions at a constant spatial rate, $1/l_c$, with $P = \exp(-\ell/l_c)$, which is consistent with past studies [29, 30].

4.1.2. Tortuosity-Hit Correlation and Adsorption-Desorption

Figure 2 shows a clear correlation structure between particle tortuosity and the number of times a particle strikes the solid phase boundary. Upscaling the number of boundary hits is important for estimating retardation effects due to adsorption-desorption. To impose such correlation effects, we sample from a joint tortuosity-hit distribution to upscale boundary hits. We sample DNS Lagrangian statistics of the porous media column and in each sampling slice store a travel time, tortuosity, number of hits tri-pair for each particle trajectory. At every model step, each particle samples a travel time and its associated tortuosity value. In the case of reactive transport, the sampled tortuosity conditionally informs the number of hits; specifically we implement hits by sampling a joint distribution $\psi(\chi|hits)$, which is created from the tri-pair sampled data.

The waiting time due to adsorption-desorption is calculated from the number of hits. We assume knowledge of Da_a, Da_d and make them identical to the values prescribed in the pore-scale numerical simulations. The cumulative waiting time for a single particle at one CTRW step is $\sum_{j=1}^{N_{hits}} \tau_{sorb}^{j}$. For every hit, a particle adsorbs with probability P_{sorb}. P_{sorb} is based on Da_a and is calculated with equation (6). If the particle adsorbs, then we sample a waiting time from an exponential distribution, $\psi(\tau_{sorb}) = \lambda \exp(-\lambda \tau_{sorb})$, with λ corresponding to Da_d. Hence, sampling travel time, tortuosity, and boundary hits for each particle trajectory is sufficient to parameterize an upscaled reactive CTRW model.

4.2. Bernoulli CTRW Parameterization

In this study, the 2D porous media column is partitioned into equally sized slices with x-length of size ℓ and with y height equivalent to column height. We select n slices which are used for CTRW parameterization. The cumulative sampling area a is smaller than the entire column area. The slice locations are randomly determined by sampling a uniform distribution, i.e., every spatial position in the domain has equal probability of being selected. In each slice, the ensemble of travel times τ, tortuosity χ, and number of boundary hits, statistics are stored. This provides sufficient data to construct $\psi(\tau)$ and $\psi(\chi|hits)$, and apply the adsorption-desorption Bernoulli CTRW framework.

4.2.1. Sampling Sensitivity

A primary objective of this study is to complete a sensitivity analysis relating sampling statistics to CTRW predictions. To do so, we parameter-

ize CTRW models with different sampling slice sizes, ranging from $O(.1)$ to $O(10)$ times the velocity correlation length scale. Furthermore, we investigate how increasing the sampling area used for parameterization statistics affects CTRW predictions.

Figure 3 provides a conceptual visualization of the method for sampling the domain. In the top and bottom sub-figures a fixed sampling area a is selected, highlighted in red. The sampling area in the top subfigure is comprised of one large slice, while the sampling area in the bottom subfigure consists of many smaller slices. The spatial distribution of slices is randomly selected from a uniform distribution. Simulation statistics corresponding to the randomly selected samples parameterize our Bernoulli CTRW model. We investigate whether the size and spatial distribution of the slices influences CTRW predictions. Since the slice locations are randomly sampled, we run 100 realizations for each sampling area - slice size combination; each realization has a different set of sampling slice locations. This enables us to quantify both the accuracy and sensitivity of CTRW predictions over an ensemble for a given sampling area - slice size combination parameter.

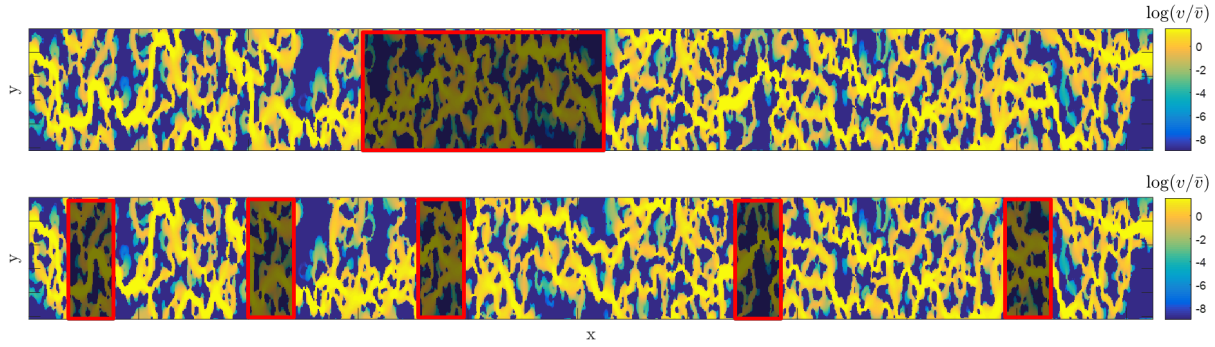

Figure 3: CTRW models are parameterized via sampling the porous media column and tracking Lagrangian statistics in direct numerical simulations. The above figure provides a conceptual illustration of the different sampling techniques. We partition a given sampling area into n slices of length ℓ and simulate transport in each slice to parameterize the CTRW. We test slices of size $.1, .2, .4, 1, 10l_c$ and sample $1/8, 1/4, 1/2$ the total domain area. Model prediction and uncertainty vs sampling slice size is explored.

5. Results

In this section, transport predictions of the proposed CTRW are compared with pore-scale simulations for different Pe, Da_a, Da_d combinations.

First, we present the conservative transport case and test model sensitivity to input parameters. Then we test CTRW model performance under varying adsorption-desorption rates. The section concludes by examining how diffusion influences both reactive and conservative transport behavior.

5.1. CTRW Predictions and Uncertainty for Conservative Transport

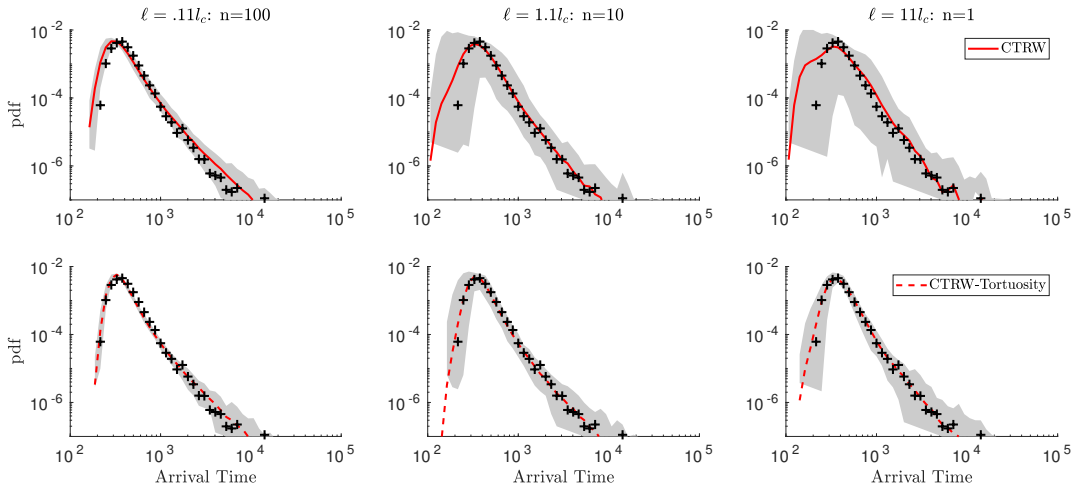

Figure 4: Bernoulli CTRW model predictions are compared with DNS observations (black crosses). CTRW models are parameterized via sampling $a = 1/8$ the entire column and n sampling slices varying in size $\ell = .1l_c, 1l_c, 10l_c$. For each sample slice size, we run 100 CTRW realizations, i.e., in each realization we sample the domain uniformly. The ensemble envelope of realizations is shown in gray shade and the mean model prediction is denoted with the red solid line. The top row shows raw Bernoulli CTRW predictions and the bottom row corresponds to corrected predictions made via a spatially dependent mean tortuosity parameter. Results are shown for $Pe = 1000$ and time units are in seconds.

We parameterize a CTRW model and predict first passage time distributions at the porous media column outlet. The CTRW performance is tested for different sample slice size - cumulative sampling area combinations; here we choose sampling slice sizes of $\ell = 0.1l_c, 0.2l_c, 0.4l_c, 1l_c, 10l_c$. The cumulative sampling area a has size $1/8, 2/8, 4/8$ the area of the total column area.

The top row of Figure 4 shows the entire ensemble distribution (gray) and mean (red) CTRW outlet breakthrough curve predictions for $a = 1/8$ and slice sizes $\ell = 0.1l_c, 1l_c, 10l_c$ for conservative transport. For each slice

size we run 100 CTRW realizations, with each realization sampling the domain randomly but uniformly. The mean breakthrough curves accurately predict tailing behavior for all slice sizes, however early and peak arrival are only captured when the sample slices are smaller than the velocity correlation length scale. Additionally, the spread across CTRW realizations decreases as the sample slice size decreases, as demonstrated by the reduced gray shaded area. In the right subfigure, one large continuous slice with size 1/8 the length of the column is sampled to parameterize the CTRW. Predictions are least accurate and uncertainty is greatest for this case. Recall that the mean travel time distribution evolves through space, and so sampling one region of the domain may not be representative of the global column statistics. When the sampled statistics are not representative, CTRW predictions will fail. These results suggest that for a given sampling area, we can reduce uncertainty across realizations by taking many small slices that are spatially distributed. Such a sampling method yields statistics more representative of the global column system at large.

The solid lines in the left panel of Figure 5 display the mean squared error of conservative mean breakthrough curve predictions for CTRW models parameterized with different sample slice lengths. Colors denote the total sampling area; $a = 1/8, 2/8, 4/8$ for black, red, blue lines, respectively. As expected, model prediction accuracy improves as the sampling area increases because the sampled statistics become more representative of the global statistics and thus spatial heterogenities in Lagrangian statistics are better captured. For all sampling areas, the largest CTRW error occurs when the parameterization slice size is $O(10)$ times larger than l_c. Note that when the slice size is much greater than the velocity correlation length scale $\ell >> l_c$ the probability of a particle velocity persisting over successive model steps goes to zero and the so the Bernoulli CTRW effectively reduces to an uncorrelated random walk. The most accurate predictions always occur when sample slices are smaller than l_c. This result demonstrates the importance of parameterizing models with representative statistics; an uncorrelated CTRW model most closely resembles using statistics from one continuous slice. However, even in this idealized porous media system, parameterizing the CTRW with sample sizes larger than the correlation scale provide less accurate predictions on average, suggesting that correlated CTRW models may be practically more suitable for prediction purposes, especially in this context where computational costs between sampling methods for a fixed sampling area are approximately equal.

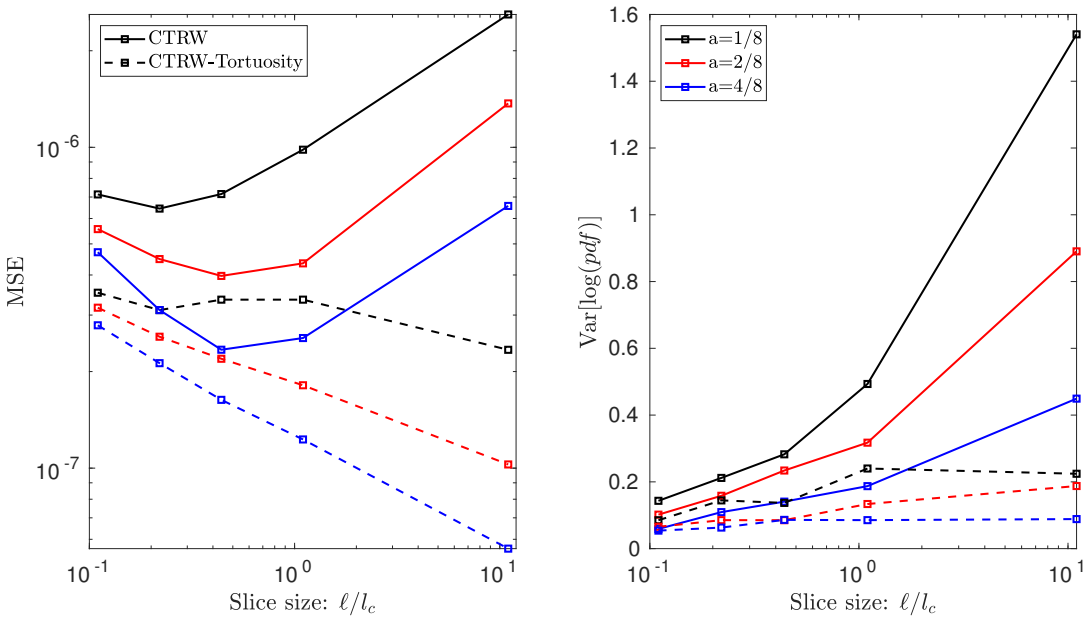

Figure 5: Bernoulli CTRW prediction accuracy is quantified via the mean squared error (MSE) of the ensemble average (left subfigure) and output uncertainty is quantified by the variance of log pdf values across all realizations averaged over all times (right subfigure). Model predictions are assessed for different sample slice sizes $\ell = .1l_c, .2l_c, 0.4l_c, 1l_c$ and for cases with sampling area $a = 1/8, 2/8, 4/8$ the entire domain area. Solid lines indicate the Bernoulli CTRW model and dashed lines denote the tortuosity corrected model. Results are shown for $Pe = 1000$.

In addition to reduced accuracy, parameterizing the CTRW with a slice size greater than the velocity correlation length scale $\ell >> l_c$ generates increased uncertainty across realizations. We quantify uncertainty with the time averaged variance of log arrival time pdf values. The solid lines in the left panel in Figure 5 displays the variance vs. slice size length for different sampling areas. As the sampling area increases, the uncertainty in prediction across CTRW realizations decreases because sampled statistics become more representative of global statistics and better capture spatial heterogeneities that impact transport behavior. On the other hand, the ensemble variance decreases as the sampling slice size decreases. This suggests that when the sampling volume needed for representative statistics is unknown, obtaining statistics from many small slices randomly distributed throughout the domain will reduce prediction uncertainty. Using data from fewer larger sample slices does not fully capture the entire spectrum of spatial

heterogeneity, and thus generates increased variability across the ensemble of predictions.

5.2. Accounting for Spatial Heterogeneity via Tortuosity

The spatial variability of Lagrangian and Eulerian statistics makes it challenging to characterize the upscaled model by sampling only a portion of the porous media column. In this section, we explore methods that correct the previously sampled distributions to account for spatial variations in the flow field structure, thereby improving mean model predictions and reducing the spread across the ensemble. We do so by relating tortuosity and travel time. Next in section 5.2.1, we assume full knowledge of the mean tortuosity field to correct the sampled data; we acknowledge that such information is currently practically unobtainable without relying on full simulation of the domain. Thus this section serves merely as a proof of concept. Then, in section 5.2.2, we evenly space sampling areas throughout the entire column and find that this method, on average, better captures spatial heterogeneity of transport characteristics than sampling randomly.

5.2.1. Correcting Sampled Data via Tortuosity: a Proof of Concept

The spatial evolution of the travel time and tortuosity distributions and their associated mean values are positively correlated, i.e. large travel distances are more likely to have large travel times. Here we investigate how correcting for spatial variations in mean travel distance influences upscaled model prediction. To do so, we assume full knowledge of the mean tortuosity field: we divide the porous media column into slices with length ℓ and each slice has a mean tortuoisty value $\langle\chi(x)\rangle^G$. Then as done previously, we sample a portion of the domain and construct a travel time distribution $\psi(\tau)$ from high resolution simulations. We store an additional piece of information, the mean tortuosity for each local sample slice $\langle\chi\rangle^L$: hence, we now sample a travel time, mean tortuosity pair $\psi(\tau, \langle\chi\rangle^L)$. Our sampled travel time in the CTRW is then scaled according to the current location of the particle, reflecting that areas of the domain where particles travel a farther (shorter) distance should have larger (smaller) travel times. The scaled CTRW time equation becomes:

$$t_i^{n+1} = t_i^n + \tau_i^{n+1} \frac{\langle\chi(x)\rangle_{n+1}^G}{\langle\chi(x)\rangle^L} \tag{21}$$

Here our scaling factor $\frac{\langle\chi(x)\rangle_{n+1}^G}{\langle\chi(x)\rangle^L}$ is simply the tortuosity corresponding to the particle's current x position over the tortuosity of the sampled travel time

slice. If the sampled mean tortuosity is smaller than the mean tortuosity of the particle's current position, then the time is increased by the correction factor to account for the fact that the particle will likely experience a higher tortuosity due to its current position in the column.

The bottom row of Figure 4 compares the mean CTRW predictions with simulated breakthrough curves and displays the spread across realizations when a tortuosity correction is applied using equation (21). For all sample area sizes, the tortuosity correction improves mean model performance, most notably for early arrival times. More importantly, the spread across realizations also decreases with the tortuosity correction. Hence the correction, both improves model accuracy and decreases uncertainty by accounting for the spatial heterogenity of tortuosity through the column.

The dashed lines in Figure 5 shows the mean model error vs. slice size as well as the realization variance vs. slice size. For all sampling area - slice size combinations, the tortuosity correction reduces model error and the prediction uncertainty. knowledge of the the spatial heterogeneity of tortuosity has a marked impact on the model results. Similar to the trends observed when no correction is applied, the model uncertainty increases with increasing slice size, although the rate of such increase is largely reduced in this corrected model. However, a different trend emerges for mean squared error; with the tortuosity correction, the mean CTRW error decreases with increasing slice size. When sample slice sizes are larger than the velocity correlation length scale $\ell > l_c$, such correlation effects are built into the measured statistics and do not need to be explicitly represented in the upscaled model. Hence for large slice sizes, only the larger scale spatial transitions in Lagrangian statistics, caused by heterogeneities of the fluid flow structure, need to be captured to upscale transport, which is explicitly done in this method by scaling travel times based on the tortuosity field.

5.2.2. Evenly Distributed Sample Slices

Knowledge of the full Lagrangian statistics is practically unobtainable in most non-insilico or field settings. We investigate here strategies to obtain a representative sample of Lagrangian statistics, by constraining the model on partial information of the tortuosity field. Our first attempt was to reconstruct the full log-tortuosity field with conditional kriging, i.e. using sample slices as conditioning points. The results of these numerical experiments are here omitted since the model performances were similar to those obtained with random sampling and no tortuosity correction (i.e., shown in Figure 4, top row). Here, we then simply distribute sample slices, such that

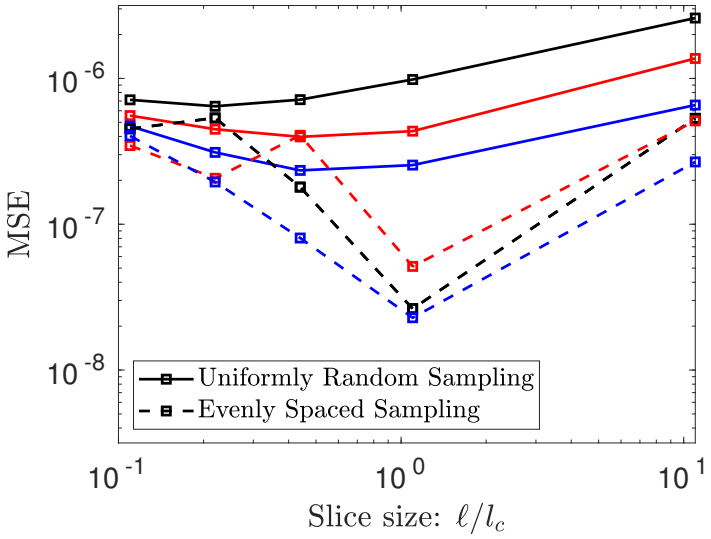

Figure 6: The MSE of the ensemble average are given for conservative transport and $Pe = 1000$ for the two sampling methods; 1) sampling the porous media uniformly random (solid lines) and 2) equi-distant spaced sampling areas (dashed lines). MSE values are shown for different sampling areas $a = 1/8, 2/8, 4/8$, black, red, blue, respectively and for different sample slice sizes $\ell = 0.1l_c, 0.2l_c, 0.4l_c, 1l_c, 10l_c$.

they span the entire domain and are spaced by equal distance.

Again a fraction a of the the porous media domain is sampled and divided into n slices of equal size a/n. The center of each sample slice is set such that it is equi-distant from neighbor slices, with distance between centers equal to L/n, where L is the column length. We compare CTRW prediction accuracy when parameterized with the evenly spaced samples with predictions via uniformly distributed samples for various sample area - slice size combinations (Figure 6). Solid lines show the MSE with the uniformly random method for sample area fractions $a = 1/8, 2/8, 4/8$ and dashed lines correspond to the evenly spaced sampling method. For all slice lengths ℓ, sampling the porous media with equi-distant spaced sample slices improves model predictions, especially as sample slice size increases. This is because the equi-spaced sampled tortuosity distribution is more representative with respect to its global (generally unknown) distribution than when random sampling is performed. Note howerver, that results obtained in this case rely on a single realization and may be subject to the particular case analyzed here, i.e. might not be generally exportable to any setting.

5.3. Adsorption-Desorption Results

The Bernoulli CTRW model is extended to include adsorption-desorption processes for particle-boundary interactions. We test reactive CTRW performance for a wide range of Da_a, Da_d values and combinations. Recall that adsorption-desorption is only possible if a particle strikes the solid boundary. Thus predicting the number of particle hits is important for accurately capturing transport of the reactive solute plume.

5.3.1. Upscaling Hits: Particle - Boundary Interactions

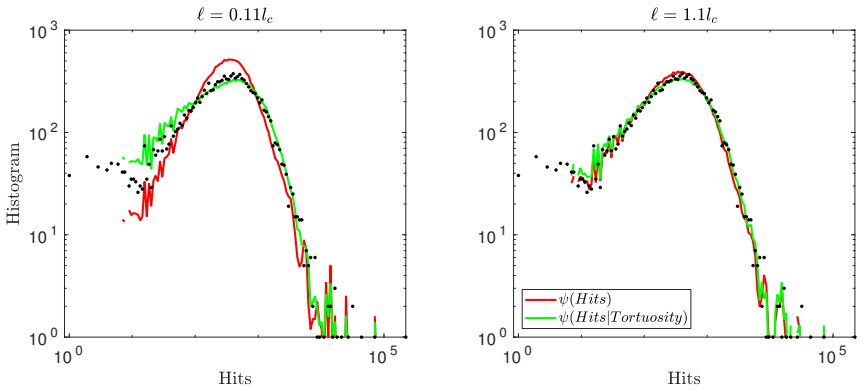

Figure 7: We predict the number of total hits from column inlet to outlet with the Bernoulli CTRW for two sampling slice sizes $\ell = .1l_c, 1l_c$ for $Pe = 1000$ and sampling area $a = 2/8$ the column area. The hit distribution is predicted for a Bernoulli model when hits are sampled from a random distribution $\psi(Hits)$ (red) and when hits are sampled from a conditional distribution $\psi(Hits|Tortuosity)$ (green).

For every sample slice we track tortuosity, travel time, and particle boundary hits, and use these values to estimate the global distribution from which the CTRW is parameterized. We test two methods to upscale the number of particle hits.

1. At every model step, particles sample randomly from a hit distribution $\psi(Hits)$, which is the aggregate of hit distributions for each sample slice and independent from all other variables.
2. At every model step, we sample from a hit distribution dependent on tortuosity $\psi(Hits|Tortuosity)$, i.e., we sample a tortuosity value which then informs the number of hits using a realization dependent joint hits-tortuosity distribution similar to the one shown in Figure 2.

We estimate the total number of times a particle hits a grain boundary from inlet to outlet with the two upscaled methods and compare with high fidelity pore-scale simulations.

Figure 7 compares the distribution of total hits for the two upscaled methods vs. simulation data for two sample slice sizes. The total hit distribution spans several orders of magnitude. For sample slices of size $\ell = .1l_c$ the peak of the distribution is overestimated and low hit values are underestimated when the Bernoulli model samples $\psi(Hits)$. However, conditioning the hits on particle tortuosity enables the upscaled model to faithfully represent the particle - boundary interactions. This indicates that over small scales, the relationship between hits and tortuosity should be considered in upscaled models, because particles with very tortuous paths are more likely to interact with boundaries. For the sample slice of size $\ell = l_c$ both upscaling methods accurately capture the total hit distribution, suggesting the number of hits and tortuosity correlation structure homogenizes after particles have traveled sufficient distance in the longitudinal direction.

5.3.2. Upscaling Transport with Sorption-Desorption

We predict breakthrough curves at the column outlet with an adsorption-desorption Bernoulli CTRW and compare with high fidelity simulations. Simulations are tested for all combinations of $Da_a = 10, 10^2, 10^3$ and $Da_d = 10, 10^2, 10^3$; additionally we consider extreme conditions where particles sorb with probability 1 if they contact the solid boundary $P_{sorb} = 1$ and have large waiting times $Da_d = 10^5$. This extreme case allows us to test the veracity of the proposed upscaled CTRW model.

Both adsorption and desorption rates influence column scale particle transport. A particle that hits the solid boundary N_{hits} time has an expected delay in transport $E[T_{delay}] = N_{hits}P_{sorb}\lambda$. Hence transport is delayed as P_{sorb} (controlled by adsorption rate) and λ (controlled by desorption rate) increase. Figure 8 displays predicted and simulated breakthrough curves for selected Da_a, Da_d combinations and $Pe = 1000$. The left and right columns in Figure 8 correspond to sample slice sizes of length $\ell = .1l_c, 1l_c$, respectively. The solid line corresponds to CTRW predictions without correcting for spatial variations in the tortuosity field, and the dashed lines indicates a tortuosity correction. When the Da_a, Da_d are relatively low $O(100)$, breakthrough curves remain relatively similar to the conservative case. As Da_a, Da_d increase, the mean arrival breakthrough time increases and tailing is enhanced. Early time breakthrough is not significantly delayed by adsorption-desorption processes, as these early particles are transported

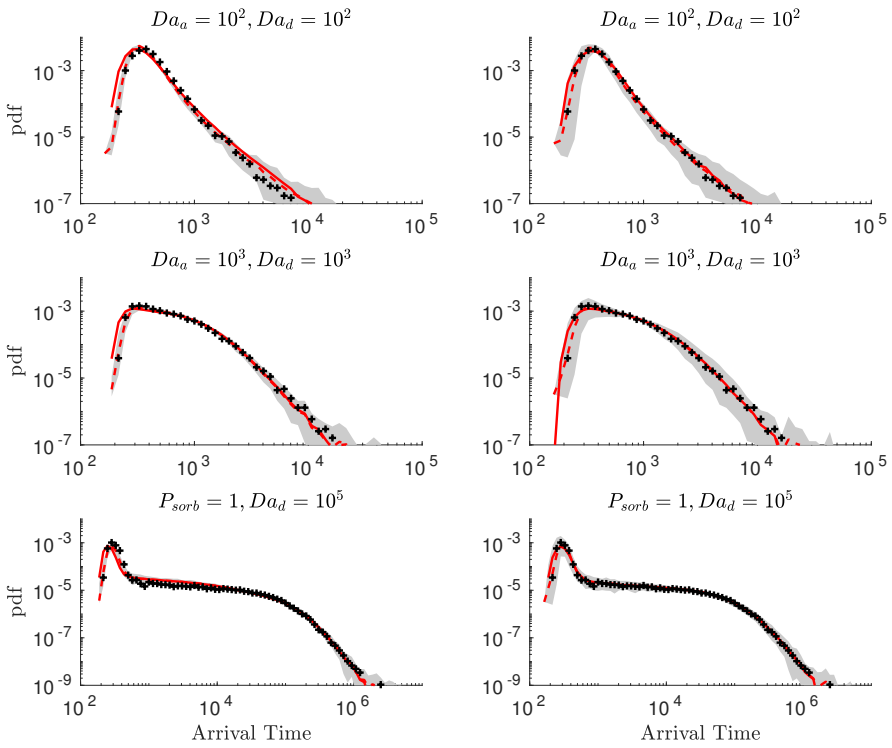

Figure 8: We predict BTCs at the domain outlet for different adsorption-desorption rates. Solid lines are mean CTRW BTC curves, dashed lines are mean CTRW-tortuosity curves. The left column has sampling size $\ell = .1l_c$ and right column has sampling size $\ell = l_c$. The hit distribution in the CTRW model is sampled from joint tortuosity-hit pdf. The gray shaded region shows the ensemble envelope for 100 realizations. Time units are in seconds.

via fast channels where interaction with the solid boundary is limited or even zero. Additionally, we consider a case with $P_{sorb} = 1$ and $Da_d = 10^5$, meaning every time a particle hits a boundary it is guaranteed to adsorb and will likely remain attached to the boundary for a relatively large time. In this case, there is a distinct peak at early times followed by enhanced tailing. The distinct peak corresponds to particles that have little or zero interaction with the boundary, while particles that hit the boundary have longer waiting times; the largest breakthough times are delayed from $O(10^4)$ in the conservative case to $O(10^6)$.

The reactive Bernoulli CTRW model captures breakthrough curves across the $Da_a - -Da_d$ parameter space. Trends are similar to those observed in

the conservative case. The spread across CTRW realizations increases when sample slice size ℓ increases. Additionally, the tortuosity correction (dashed lines) improves breakthrough prediction, especially at early times. Recall that early travel times are those particles in fast channels who are less likely to interact with the solid boundary, and thus the tortuosity correction has the same effect as in the conservative case. However, as the adsorption-desorption processes play an increasing role, i.e. Da_a, Da_d increase, the tortuosity correction becomes less significant because the reactive boundary processes dominate retardation.

5.4. Impact of Péclet Number on Reactive Transport

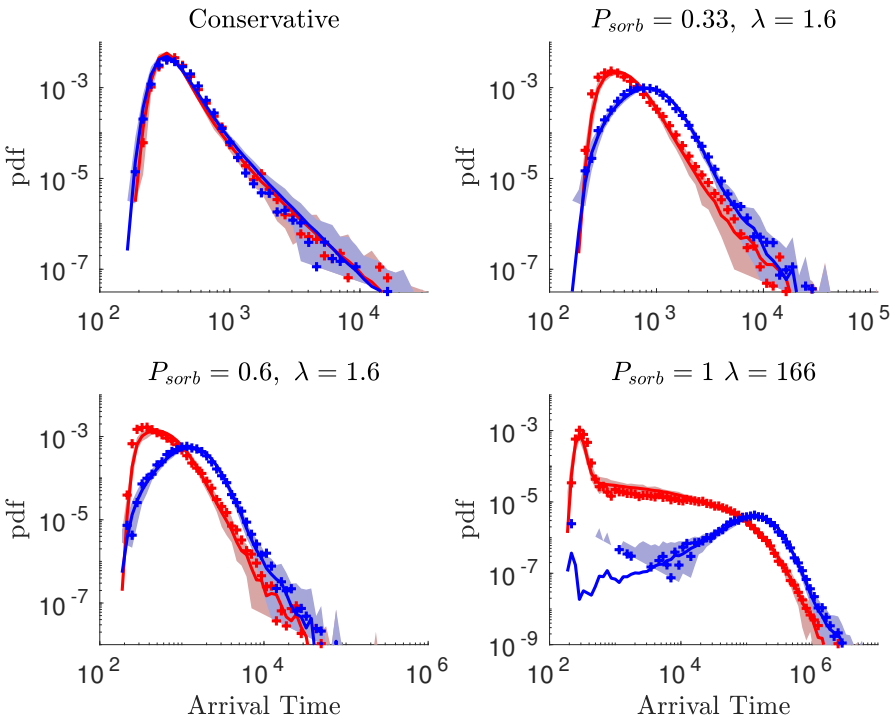

Figure 9: Bernoulli CTRW predictions at the outlet (solid lines) are compared to DNS simulations (crosses) for the $Pe = 100$ (blue) and $Pe = 1000$ (red). Breakthrough curves are shown for different combinations of sorption probabilities and mean waiting times. The gray shaded areas display the ensemble of CTRW predictions for 100 realizations. Time units are in seconds.

We test the influence of diffusion on reactive transport by comparing breakthrough curves for $Pe = 100, 1000$. Recall that diffusion is the mech-

anism which enables particles to interact with solid boundary, hence when diffusion strength increases the number of times a particle hits the boundary increases on average; the mean number of hits from inlet to outlet is 871, 281 for $Pe = 100, 1000$, respectively. Figure 9 shows outlet breakthrough curves for the $Pe = 100$ (blue) and $Pe = 1000$ (red) cases under different adsorption and desorption rates.

In the conservative case, diffusion does not have a significant impact on particle breakthrough times and so breakthrough curves appear similar across Pe. As P_{sorb} and waiting time λ increase, the peak breakthrough concentration is delayed and tailing is enhanced because particles now have increased likelihood of sorbing to grain boundaries. In the extreme case, where $P_{sorb} = 1$ and $\lambda = 166$, the peak concentration is significantly delayed for the $Pe = 100$ case. Increasing diffusion strength, increases the likelihood of hitting the boundary, which significantly delays transport when the adsorption rate is high and desorption rate is low.

The reactive Bernoulli CTRW model's robustness is tested by varying Pe (Figure 9). For all cases considered, the CTRW model (solid lines) captures reactive transport through the porous media column. Here the CTRW model is parameterized with sample slice size $\ell = 0.1 l_c$ and sample area fraction $a = 1/8$. Again, we run 100 CTRW realizations, with each realization sampling the domain uniformly random. The spread across realizations remains relatively small for this sample slice. Hence parameterizing the reactive Bernoulli CTRW with slices smaller than the correlation length scales can accurately predict transport under a wide variety of conditions.

6. Discussion and Conclusions

We run direct numerical simulation of transport through a 2D porous column with adsorption-desorption. The spatial distribution of grains and irregularity of grain geometry results in a fluid velocity field that is complex and spatially variant. Consequently, Lagrangian transport statistics are spatially dependent (non-stationary), and capturing such features in an upscaled modelling framework is challenging. We use direct numerical simulations to explore reactive transport through the column, which guides development of a more general spatial Markov model with adsorption-desorption. The simulations show:

1. The velocity, travel time, and tortuosity distributions are highly intertwined. High velocity channels are mainly aligned with the pri-

mary pressure gradient, meaning particles transported in such channels travel a shorter distance, resulting in relatively fast travel times.

2. High velocity regions channelize mass which homogenizes corresponding Lagrangian statistics, as a significant portion of mass is forced through the small high velocity region.
3. Tortuosity and the number of times particles hit the solid grain boundary exhibit a strong correlation structure, such that more hits are likely with high tortuosity values. In the context of adsorption-desorption, this means that slow velocity particles preferentially sorb-desorb, further delaying downstream transport.

These correlation structures are parameterized in a reactive Bernoulli CTRW model. In the proposed framework, we assume that velocity transitions at a constant spatial rate, which can be upscaled as a Bernoulli process. In the case of conservative transport, the Bernoulli CTRW accurately predicts column outlet breakthrough curves, which is consistent with past studies [28, 29, 30]. We extend the Bernoulli framework to include adsorption-desorption processes. For scales where velocity correlation is important, tortuosity and the number of times a particle strikes the boundary also exhibit a correlation structure that must be captured in the upscaled model for accurate prediction of adsorption-desorption to the grain surfaces. We do so by sampling a tortuosity travel time pair at each model step and then sampling a particle hit number conditioned on the tortuosity. This method enables the hit distribution and adsorption-desorption effects on transport to be incorporated in the upscaled framework. Failing to include such correlation inaccurately represents the hit distribution of particles, which subsequently propagates as error for adsorption-desorption transport predictions.

The proposed reactive Bernoulli CTRW model faithfully portrays transport influenced by adsorption-desorption processes under a range of Damköhler and Péclet numbers. This modelling framework has been generalized from previously proposed reactive SMM frameworks [39] via a tortuosity dependent, hit parameter. Previous models needed generalization as they would fail in the 2D porous media considered here, because it cannot be partitioned into identical periodic elements, which is an inherent assumption of the model proposed by [39]. We show that under certain adsorption-desorption conditions, conservative breakthrough curves are delayed and their shapes significantly altered. Accounting for adsorption-desorption via a retardation coefficient, as has been commonly assumed, simply scales the breakthrough

curves and would not predict such shape shifts. Hence, there is great utility in upscaling the underlying correlation structures of adsorption-desorption processes, as we have done here. Although our proposed model was only tested in an idealized 2D setting, it provides a framework which can be expanded to aid in our understanding of more complicated 3D porous media geometries, where for example pore-scale sampling could be obtained from high-resolution imaging of geological media or from three-dimensional artificially generated pore structures. Application of our approach to these systems is envisaged in future research efforts.

We parameterize our upscaled CTRW models by sampling statistics from a portion of the column and explore model sensitivity to sample slice size and total sampling area. When particle sample slices exceed the velocity correlation length scale, both velocity correlation via the Bernoulli parameter and the tortuosity-hit correlation structure decorrelate and no longer need to be accounted for in an upscaling framework. If this is the case, one might naturally question the use of a more complicated correlated model, as a simpler framework could be substituted by considering statistics over a larger sampling area. We address this issue by parameterizing the CTRW model with different sampling slice size-sampling area combinations. For each combination, samples used for parameterization are randomly distributed throughout the domain and 100 realizations are considered, with each realization having different samples and thus different statistics for model parameterizations. For a given sampling area, mean CTRW breakthrough curve predictions are most accurate when sample sizes are less than the velocity correlation length scale, meaning velocity and tortuosity-hit correlational effects are important for these sample sizes. More importantly, sampling many small slices distributed through the domain significantly reduces the uncertainty of CTRW predictions. This suggests that spatially distributing samples throughout the domain better captures global system statistics than sampling a few large continuous slices, which agrees with the observations that show Lagrangian statistics change significantly through space.

The spatial evolution of tortuosity and travel time are highly correlated and non-stationary. Since only a portion of the porous column is sampled to parameterize the CTRW models, the sampled tortuosity and travel times may not be representative of the global sample. In a numerical experiment we assume full knowledge of the mean tortuosity field and leverage this knowledge to correct for spatial heterogeneity of statistics. Specifically

we scale the travel time by a linear coefficient dependent on the sampled tortuoisty and particle's current spatial position. In essence, this scaling increases (decreases) travel time for regions where particles must travel a farther (shorter) distance than the mean travel distance. With this intuitive and simple correction, the CTRW model has higher prediction accuracy and smaller uncertainty. However, full knowledge of the tortuosity field is currently not obtainable without simulation of the entire domain, and so this spatial correction serves as a proof of concept that upscaled models can be significantly improved if such spatial knowledge is included in input parameters.

This leaves a common unresolved issue in upscaled transport modeling studies; in order to upscale transport we often first need to simulate transport. Input upscaled modelling parameters must reflect the spatial heterogenity of Lagrangian statistics for accurate model predictions. However, accounting for that spatial heterogentiy without sampling the entire domain is challenging. Note that in this study geo-statistical interpolation methods, including conditional kriging techniques were applied to infer global Lagrangian statistics from sampled areas (results not shown). However, these methods failed to capture the spatial evolution of statistics in sufficient detail to improve model prediction and model predictions were only improved once the global distribution was sufficiently sampled via a large sampling fraction a. We found however that sampling areas spaced equi-distant across the entire domain yield more representative samples than random uniform sampling. Lagrangian statistics are correlated through space and the equi-distant sampling method evenly spreads out sampling areas, thereby increasing the probability that the the spatial heterogentiy from column inlet to outlet is represented. However, even with this method a sufficient sample area fraction must be considered, or the global statistics will not be captured.

Acknowledgments

This material is based upon work supported by, or in part by, the US Army Research Office under Contract/Grant number W911NF-18-1-0338. TS is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1841556.

Appendix A. Geometry and Flow field

We detail here below the methodology followed to generate the porous media column and numerically compute the single-phase flow field.

Appendix A.1. Porous Media Column: Generation

The synthetic porous media column is generated using a variant of the method introduced by Smolarkiewicz and Winter [49, 50]. It follows four successive steps as described here. While three-dimensional structures are possible, we focus here on the two-dimensional version of the method.

1. A two-dimensional regular grid with uniform spacing h is generated and is populated with random values $u(\mathbf{x})$ ($\mathbf{x} = (x, y)$, being the vector of grid node coordinates) sampled from a continuous uniform distribution on the closed interval $[0, 1]$ (Figure A.10).

2. This random field is convolved with a symmetric flattened Gaussian kernel,

$$k(\mathbf{x}, \sigma, \theta) = \min\left[\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \frac{\theta}{2\pi\sigma^2}\right], \tag{A.1}$$

   with variance σ^2 and flattening factor θ, to generate the isotropic correlated random topography (Figure A.10),

$$T(\mathbf{x}) = \int_{\mathbf{R}^2} k\left(\mathbf{x}-\mathbf{y}, \beta, \theta\right) u\left(\mathbf{y}\right) d\mathbf{y}. \tag{A.2}$$

   The correlation length in the topography is controlled by the value of β while the roughness of the topography is determined by the value of θ. After the topography is generated, it is normalized to the interval $[0, 1]$. Note that the convolution is computed in the frequency domain. As a consequence, $T(\mathbf{x})$ is double-periodic in space.

3. A level threshold, $\gamma \in [0, 1]$, is applied to given $T(\mathbf{x})$ to map values onto a phase indicator function I, where $I(\mathbf{x}) = 1$ in the fluid nodes and $I(\mathbf{x}) = 0$ otherwise, i.e.,

$$I(\mathbf{x}) = \begin{cases} 0, & \text{if } T(\mathbf{x}) \geq \gamma \\ 1, & \text{if } T(\mathbf{x}) < \gamma \end{cases}. \tag{A.3}$$

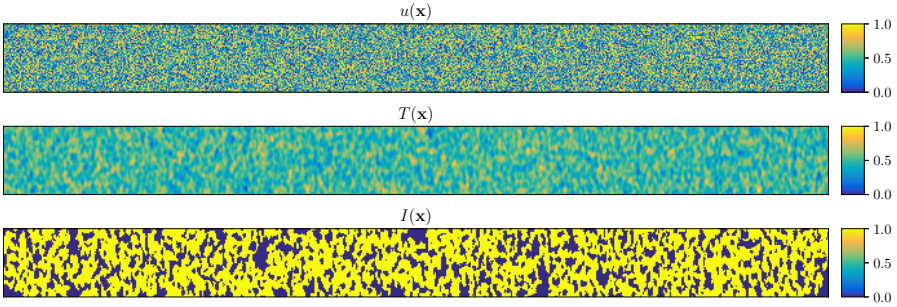

Figure A.10: Generation of the 2D porous media column. Subfigure 1 shows the random field achieved at step 1. Subfigure 2 displays the topography obtained at step 2 after convolution of the random field with a flattened Gaussian kernel. Subfigure 3 shows the phase indicator function resulting from thresholding of the topography.

Table A.1: Generation parameters and geometric observables of the porous media column analyzed.

Number of pixels [-]	Column size [mm^2]	h [μm]	σ [μm]	θ [-]	γ [-]	Porosity [-]	β [μm]
10240×512	20.48×1.024	2.0	24.0	0.5	0.55	63.937%	38.0

The resulting two-dimensional field $I(\mathbf{x})$ is representative of a porous media structure with fluid and solid nodes identified by the value of $I(\mathbf{x})$ (Figure A.10). Intuitively, as γ increases so does the volume of the fluid space in the porous medium.

4. The given $I(\mathbf{x})$ is finally processed to remove isolated (disconnected) regions that would prevent convergence of the fluid flow solver.

The main difference here with respect to the original method is the use of a flattened Gaussian kernel instead of a Gaussian kernel. This modification permits us to control the roughness of the fluid-solid boundaries of the porous medium. Here, we generate a porous media column with the generation parameters and geometric observables listed in Table A.1. We measure the degree of spatial correlation of the fluid space, i.e., the integral scale of the phase indicator function, β, by calculating the empirical variogram of I and by characterizing it through fitting of a spherical model (see, e.g., [51] for additional details). We note that the column analyzed here has about 539 integral scales in the x direction and about 27 in the y direction.

Appendix A.2. Flow

Single-phase fluid flow in the porous media column is directly solved on the fluid domain and delimited by the phase indicator function by using a finite volume method. We solve the steady-state Navier-Stokes equations for incompressible, isothermal and Newtonian fluid:

$$\nabla \cdot \mathbf{u}(\mathbf{x}) = 0 \tag{A.4}$$

$$\nabla \cdot [\mathbf{u}(\mathbf{x})\mathbf{u}(\mathbf{x})] = -\frac{1}{\rho}\nabla p(\mathbf{x}) + \nu \nabla \cdot \nabla \mathbf{u}(\mathbf{x}) \tag{A.5}$$

Here, $\mathbf{u}(\mathbf{x})$ denotes the velocity vector, p is pressure, ρ is fluid density and ν is fluid kinematic viscosity. The OpenFOAM® open-source CFD library, release v1712 [52], is used to solve the target system of equations through the built-in solver simpleFoam. The computational grid is created using an in-house mesh generator that converts the pixels of the porous media column (i.e., the phase indicator function field) into square grid cells. Periodic boundary conditions are applied in the x and y directions. To force flow in the x direction, we add a pressure drop between the inlet and outlet faces in the x direction. No-slip conditions are implemented at the fluid-solid boundaries (walls). After steady-state is reached, the flow rate is such that Stokes flow conditions are satisfied. The measured permeability of the porous media column is 4.55×10^{-11} m^2.

Appendix B. Mean Tortusoity Mean Travel Time Correlation

We demonstrate the strong correlation between mean tortuosity and the mean travel time. The mean tortuosity field changes through space and is leveraged to account for spatial heterogeneity of the flow field and improve Bernoulli CTRW predictions in section 5.2.1.

The column is partitioned into non-overlapping slices of equal size. Lagrangian statistics in each slice are are arithmetically averaged and the spatial evolution of mean values are observed. Figure B.11 shows the mean tortuosity and mean travel time through space. Values have been normalized by the maximum mean value for each respective distribution. The mean tortuosity and mean travel times are nearly perfectly correlated with $r^2 = .99$. As the mean travel distance of particle increases, the mean travel time increases. The distance particles travel is related to the flow field structure, which is controlled by the distribution and spatial position of grains. This observation is exciting from an upscaled modeling perspective,

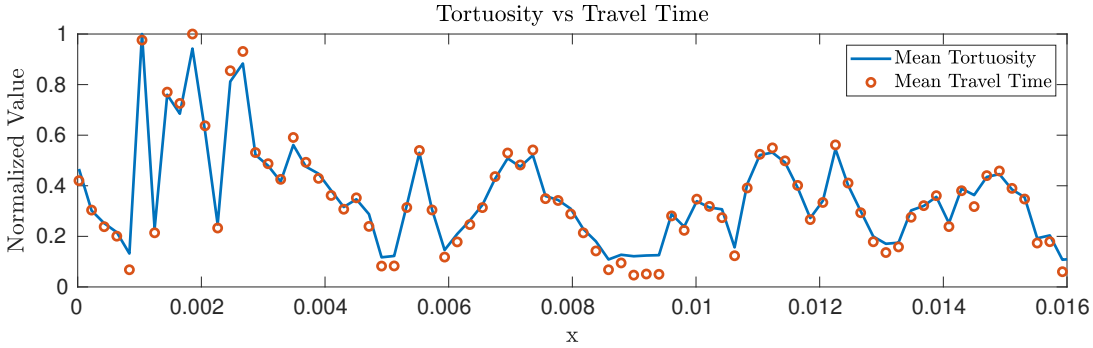

Figure B.11: The spatial evolution of the mean travel time and mean tortuosity for $Pe = 1000$. Values are normalized so that maximum equals unity. Mean values are calculated with spacing of $\ell = 1l_c$. The mean travel time and mean tortuoisty are highly correlated, i.e. the domain regions where particles travel a longer distance on average have an increased associated mean travel time.

as it implies that if the spatial evolution of the tortuosity field is uncovered, then the evolution of the travel time distribution may also be estimated by linear scaling. The mean tortuosity values provide the global distibution $\langle \chi(x) \rangle_{n+1}^{G}$, which is incorporated into the Bernoulli CTRW framework to scale travel times based on the local column's grain distributions; see section section 5.2.1.

[1] M. N. Goltz, P. V. Roberts, Interpreting organic solute transport data from a field experiment using physical nonequilibrium models, Journal of Contaminant Hydrology 1 (1) (1986) 77 – 93, transport and Transformations of Organic Contaminants.

[2] C. Harvey, S. M. Gorelick, Ratelimited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (made) site?, Water Resources Research 36 (3) (2000) 637–650.

[3] B. Zinn, C. F. Harvey, When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate gaussian hydraulic conductivity fields, Water Resources Research 39 (3).

[4] E. Major, D. A. Benson, J. Revielle, H. Ibrahim, A. Dean, R. M. Maxwell, E. Poeter, M. Dogan, Comparison of fickian and temporally nonlocal transport theories over many scales in an exhaustively sampled sandstone slab, Water Resources Research 47 (2011) W10519.

[5] P. V. Roberts, M. N. Goltz, D. M. Mackay, A natural gradient experiment on solute transport in a sand aquifer: 3. retardation estimates and mass balances for organic solutes, Water Resources Research 22 (13) (1986) 2047–2058.

[6] M. Dentz, A. Castro, Effective transport dynamics in porous media with heterogeneous retardation properties, Geophysical Research Letters 36 (3) (2009) L03403.

[7] M. Maghrebi, I. Jankovic, R. M. Allen-King, A. J. Rabideau, I. Kalinovich, G. S. Weissmann, Impacts of transport mechanisms and plume history on tailing of sorb-

ing plumes in heterogeneous porous formations, Advances in water resources 73 (2014) 123–133.
[8] M. Maghrebi, I. Jankovic, G. S. Weissmann, L. S. Matott, R. M. Allen-King, A. J. Rabideau, Contaminant tailing in highly heterogeneous porous formations: Sensitivity on model selection and material properties, Journal of Hydrology 531 (2015) 149–160.
[9] V. Cvetkovic, G. Dagan, Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, Journal of Fluid Mechanics 265 (1994) 189215.
[10] B. Rathi, H. Neidhardt, M. Berg, A. Siade, H. Prommer, Processes governing arsenic retardation on p leistocene sediments: Adsorption experiments and model-based analysis, Water Resources Research 53 (5) (2017) 4344–4360.
[11] B. D. Wood, K. Radakovich, F. Golfier, Effective reaction at a fluid–solid interface: applications to biotransformation in porous media, Advances in water resources 30 (6-7) (2007) 1630–1647.
[12] C. da Luz, S. M. d. A. G. Ulson, A. A. U. de Souza, A. Dervanoski, A. d. O. S. Moraes, B. D. Wood, et al., A multiscale model for carbon adsorption of btx compounds: Comparison of volume averaging theory and experimental measurements, Chemical Engineering Science 184 (2018) 285–308.
[13] M. J. Blunt, B. Bijeljic, H. Dong, O. Gharbi, S. Iglauer, P. Mostaghimi, A. Paluszny, C. Pentland, Pore-scale imaging and modelling, Advances in Water Resources 51 (2013) 197–216.
[14] P. De Anna, T. Le Borgne, M. Dentz, A. M. Tartakovsky, D. Bolster, P. Davy, Flow intermittency, dispersion, and correlated continuous time random walks in porous media, Physical review letters 110 (18) (2013) 184502.
[15] P. K. Kang, P. Anna, J. P. Nunes, B. Bijeljic, M. J. Blunt, R. Juanes, Pore-scale intermittent velocity structure underpinning anomalous transport through 3-d porous media, Geophysical Research Letters 41 (17) (2014) 6184–6190.
[16] G. Ceriotti, A. Russian, D. Bolster, G. Porta, A double-continuum transport model for segregated porous media: Derivation and sensitivity analysis-driven calibration, Advances in Water Resources 128 (2019) 206–217.
[17] R. Haggerty, S. M. Gorelick, Multiple-rate mass transfer for modeling diffusion and, Water Resources Research 31 (10) (1995) 2383–2400.
[18] D. A. Benson, S. W. Wheatcraft, M. M. Meerschaert, The fractional-order governing equation of lévy motion, Water Resources Research 36 (6) (2000) 1413–1423.
[19] B. Berkowitz, A. Cortis, M. Dentz, H. Scher, Modeling non-Fickian transport in geological formations as a continuous time random walk, Reviews of Geophysics 44 (2) (2006) RG2003.
[20] B. Noetinger, D. Roubinet, A. Russian, T. Le Borgne, F. Delay, M. Dentz, J.-R. De Dreuzy, P. Gouze, Random walk methods for modeling hydrodynamic transport in porous and fractured media from pore to reservoir scale, Transp. Porous Media (2016) 1–41.
[21] B. Berkowitz, H. Scher, Anomalous transport in random fracture networks, Physical Review Letters 79 (20) (1997) 4038–4041.
[22] T. Le Borgne, M. Dentz, J. Carrera, Lagrangian statistical model for transport in highly heterogeneous velocity fields, Physical review letters 101 (9) (2008) 090601.

[23] T. Le Borgne, M. Dentz, J. Carrera, Spatial markov processes for modeling lagrangian particle dynamics in heterogeneous porous media, Physical Review E 78 (2) (2008) 026308.
[24] P. K. Kang, T. Le Borgne, M. Dentz, O. Bour, R. Juanes, Impact of velocity correlation and distribution on transport in fractured media: field evidence and theoretical model, Water Resources Research 51 (2015) 940–959.
[25] D. Bolster, Y. Méheust, T. Le Borgne, J. Bouquain, P. Davy, Modeling preasymptotic transport in flows with significant inertial and trapping effects–the importance of velocity correlations and a spatial markov model, Advances Water Research 70 (2014) 89–103.
[26] N. L. Sund, D. Bolster, C. Dawson, Upscaling transport of a reacting solute through a peridocially converging–diverging channel at pre-asymptotic times, Journal of contaminant hydrology 182 (2015) 1–15.
[27] M. Dentz, P. K. Kang, A. Comolli, T. Le Borgne, D. R. Lester, Continuous time random walks for the evolution of lagrangian velocities, Physical Review Fluids 1 (7) (2016) 074004.
[28] V. L. Morales, M. Dentz, M. Willmann, M. Holzner, Stochastic dynamics of intermittent pore-scale particle motion in three-dimensional porous media: Experiments and theory, Geophysical Research Letters 44 (18) (2017) 9361–9371.
[29] J. Hyman, M. Dentz, A. Hagberg, P. K. Kang, Linking structural and transport properties in three-dimensional fracture networks, Journal of Geophysical Research: Solid Earth (2019) 1185–1204.
[30] P. K. Kang, Q. Lei, M. Dentz, R. Juanes, Stress-induced anomalous transport in natural fracture networks, Water Resources Research (2019) 4163–4185.
[31] A. Puyguiraud, P. Gouze, M. Dentz, Upscaling of anomalous pore-scale dispersion, Transport in Porous Media 837–855.
[32] A. Puyguiraud, P. Gouze, M. Dentz, Stochastic dynamics of lagrangian pore-scale velocities in three-dimensional porous media, Water Resources Research 55 (2) (2019) 1196–1217.
[33] T. Le Borgne, D. Bolster, M. Dentz, P. Anna, A. Tartakovsky, Effective pore-scale dispersion upscaling with a correlated continuous time random walk approach, Water Resources Research 47 (12) (2011) W12538.
[34] T. Sherman, A. Foster, D. Bolster, K. Singha, Predicting downstream concentration histories from upstream data in column experiments, Water Resources Research 9684–9694.
[35] Y. Edery, H. Scher, B. Berkowitz, Particle tracking model of bimolecular reactive transport in porous media, Water Resources Research 46 (7) (2010) W07524.
[36] S. K. Hansen, H. Scher, B. Berkowitz, First-principles derivation of reactive transport modeling parameters for particle tracking and pde approaches, Advances in Water Resources 69 (2014) 146 – 158.
[37] T. Aquino, M. Dentz, Chemical continuous time random walks, Physical review letters 119 (23) (2017) 230601.
[38] N. Sund, G. Porta, D. Bolster, R. Parashar, A lagrangian transport eulerian reaction spatial (laters) markov model for prediction of effective bimolecular reactive transport, Water Resources Research 53 (11) (2017) 9040–9058.
[39] T. Sherman, A. Paster, G. Porta, D. Bolster, A spatial markov model for upscaling

transport of adsorbing-desorbing solutes, Journal of contaminant hydrology 222 (2019) 31–40.

[40] E. Wright, N. Sund, D. Richter, G. Porta, D. Bolster, Upscaling mixing in highly heterogeneous porous media via a spatial markov model, Water 11 (1) (2019) 53.

[41] N. L. Sund, D. Bolster, D. A. Benson, Testing the limits of the spatial markov model for upscaling transport: The role of nonmonotonic effective velocity autocorrelations, Physical Review E 94 (4) (2016) 043107.

[42] S. Most, D. Bolster, B. Bijeljic, W. Nowak, Trajectories as training images to simulate advective-diffusive, non-fickian transport, Water Resources Research 55 (2019) 3465–3480.

[43] T. Sherman, A. Fakhari, S. Miller, K. Singha, D. Bolster, Parameterizing the spatial markov model from breakthrough curve data alone, Water Resources Research 53 (12) (2017) 10888–10898.

[44] P. K. Kang, M. Dentz, T. Le Borgne, R. Juanes, Anomalous transport on regular fracture networks: Impact of conductivity heterogeneity and mixing at fracture intersections, Physical Review E 92 (2) (2015) 022148.

[45] V. Hakoun, A. Comolli, M. Dentz, Upscaling and prediction of lagrangian velocity dynamics in heterogeneous porous media, Water Resources Research 55 (5) (2019) 3976–3996.

[46] G. Boccardo, I. M. Sokolov, A. Paster, An improved scheme for a robin boundary condition in discrete-time random walk algorithms, Journal of Computational Physics 374 (2018) 1152–1165.

[47] B. Ghanbarian, A. G. Hunt, R. P. Ewing, M. Sahimi, Tortuosity in porous media: a critical review, Soil science society of America journal 77 (5) (2013) 1461–1477.

[48] M. Siena, O. Iliev, T. Prill, M. Riva, A. Guadagnini, Identification of channeling in pore-scale flows, Geophysical Research Letters 46 (6) (2019) 3270–3278.

[49] P. K. Smolarkiewicz, C. L. Winter, Pores resolving simulation of Darcy flows, Journal of Computational Physics 229 (9) (2010) 3121–3133.

[50] J. D. Hyman, C. L. Winter, Stochastic generation of explicit pore structures by thresholding Gaussian random fields, Journal of Computational Physics 277 (2014) 16–31.

[51] G. R. Guédon, J. D. Hyman, F. Inzoli, M. Riva, A. Guadagnini, Influence of capillary end effects on steady-state relative permeability estimates from direct pore-scale simulations, Physics of Fluids 29 (12) (2017) 123104.

[52] OpenCFD Limited, OpenFOAM The Open Source CFD Toolbox, user guide version v1712, `https://www.openfoam.com/`, accessed: 2019-09-29 (2017).