Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR023552

Trajectories as Training Images to Simulate Advective-Diffusive, Non-Fickian Transport

Sebastian Most[1], Diogo Bolster[2], Branko Bijeljic[3], and Wolfgang Nowak[1]

[1]Department of Stochastic Simulation and Safety Research for Hydrosystems, University of Stuttgart, Stuttgart, Germany, [2]Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, IN, USA, [3]Department of Earth Science and Engineering, Imperial College London, London, UK

Key Points:
- We propose a data-driven spatial Markov model to effectively simulate transport in 3-D without common simplifications
- We capture process dependence by a systematic rearrangement of highly resolved training trajectories
- We simulate transport at the resolution of these highly resolved training trajectories

Supporting Information:
- Supporting Information S1
- Text S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4

Correspondence to:
S. Most,
Sebastian.Most@iws.uni-stuttgart.de

Citation:
Most, S., Bolster, D., Bijeljic, B., & Nowak, W. (2019). Trajectories as training images to simulate advective-diffusive, non-Fickian transport. *Water Resources Research*, *55*, 3465–3480. https://doi.org/10.1029/2018WR023552

Received 25 JUN 2018
Accepted 2 MAR 2019
Accepted article online 28 MAR 2019
Published online 30 APR 2019

©2019. American Geophysical Union. All Rights Reserved.

Abstract We propose a spatial Markov model to simulate transport in three-dimensional complex porous media flows. Our methodology is inspired by the concept of training images from geostatistics. Instead of using a training image we use highly resolved training trajectories obtained by high-resolution particle tracking, from which we sample increments in our random walk model. To reflect higher-order processes, subsequent increments are correlated. The approach can be split into three steps. First, we subdivide (*cut*) the training trajectories to form an archive of trajectory segments. Next, we recursively sample segments, where subsequent samples are chosen conditioned to the previous one to ensure continuity and smoothness of velocity (*conditional copy*). Finally, we merge (*paste*) consecutive segments together to generate simulated trajectories of arbitrary length. This training trajectory approach aims to overcome three common shortcomings of spatial Markov models: (1) We simulate finite-Péclet transport in three dimensions without commonly made simplifications (e.g., dimensionality reduction, and neglecting diffusion). (2) We do not parameterize dependence via a high-dimensional transition matrix. (3) We simulate transport at the resolution of the (highly resolved) training trajectories, which can be important for processes such as mixing and reaction. To validate our methodology, we apply it to simulate transport within a three-dimensional sandstone sample and compare predictions of a broad range of benchmark metrics against measurements from direct numerical simulations. We demonstrate that the training trajectories approach accurately represents three-dimensional particle motion, suggesting that this method can capture the governing dependence structure and simulate transport processes in full complexity.

1. Introduction

Non-Fickian or anomalous transport is very common in (fractured) porous media (e.g., Berkowitz et al., 2002, 2006; Bijeljic et al., 2004, 2013; Bolster & Dentz, 2012; Bolster et al., 2014; Dentz & Berkowitz, 2003; Dentz et al., 2016; Lester et al., 2013, 2016). Typically, it manifests itself as the nonlinear scaling of the second centered spatial moment of a plume and the associated anomalously early and late arrival times of solute that cannot be described by an independent normal (Gaussian) process with drift. Capturing such anomalous arrival times is important, for instance, for technical applications to ensure secure groundwater supply (e.g., Harvey et al., 2002) or in the field of water and health risk assessment (e.g., de Barros et al., 2013).

In advectively dominated transport settings, spatial velocity contrasts control the spreading of a particle plume (Le Borgne et al., 2008a) and local velocities control particle residence times within slow or fast zones (Dentz et al., 2016). Within these zones, a Lagrangian particle's velocity fluctuation is small compared to the total velocity variation of the entire velocity field (Bolster et al., 2014; De Anna et al., 2013; Dentz et al., 2016; Le Borgne et al., 2008b). We can interpret this piecewise constant velocity v as *memory of speed* that persists over some characteristic correlation length λ_v. Likewise, the direction α of a particle's motion persists (memory of direction) over a characteristic correlation length λ_α. Both these correlation lengths are specific to a given medium, reflecting details of the pore-scale geometry, which ultimately controls particle flow paths. Such memory effects prohibit the assumption that particle motion is a stochastically independent process, an assumption inherent to classical Fickian transport laws, as well as many state-of-the-art non-Fickian ones. Ideally, upscaled models should account for the influence of *memory of speed and direction* (Most et al., 2016; Meyer et al., 2013; Meyer & Bijeljic, 2016).

In Meyer et al. (2013) and Meyer and Bijeljic (2016) the effect of memory of speed and direction is captured without explicitly mimicking process memory. Transport is simulated by generating an ensemble of particle trajectories.The trajectories are generated by separating particle motion into a velocity and two angle processes (spherical coordinates), where the latter follows an Ornstein-Uhlenbeck process that controls the smoothness of the trajectory. The processes and their interplay are parameterized with a set of parameters that can be understood as the transport fingerprint. The mentioned studies are successfully applied to a variety of artificial and natural porous media. In this study we also propose a trajectory-based method but follow a different approach in merging them to preserve memory of speed and direction. Lester et al. (2013, 2016) highlight the importance of process memory, which results from advective-diffusive particle motion in complex porous media. Chaotic advection is strongly linked with memory of speed (i.e., the trapping of particles in stagnation points) and controls non-Fickian dilution and mixing processes. Our approach reflects chaotic advection up to the extent to which it is represented by the reference simulation. However, our approach here is fundamentally different from that of Lester et al. (2016), who proposed a stretching continuous time random walk.

To this end, we propose a data-driven Lagrangian approach to simulate non-Fickian transport in realistic porous media. We apply the method to the pore scale.

Spatial Markov models (SMMs; Bolster & Dentz, 2012; Kang et al., 2014; Sund et al., 2016) are a family of models that represent memory of speed and direction through the use of a transition matrix. While these models have been applied to a broad range of systems (e.g., Bolster et al., 2014; Kang et al., 2014, 2015a, 2016; Le Borgne et al., 2008b, 2008a; Sund et al., 2015b, 2015a, 2017b, 2017a), they typically rely on parameterizing a transition matrix, which can be difficult to do, although approaches applicable to real data have emerged recently (Kang et al., 2015a; Sherman et al., 2017). Another approach has involved sampling of particle trajectories (Sund et al., 2017b), which in turn can be used for mixed upscaling and downscaling models. These models can estimate nonlinear measures such as the dilution index or scalar dissipation and even predict mixing-driven reactions (Sund et al., 2017a). However, such applications are to date limited to synthetic and idealized periodic systems.

Translating such approaches to more realistic and complex geologies and geometries will be key to their future success. Here we propose a trajectory-based method to model transport and mixing in a realistic complex porous medium. Our method is largely inspired by ideas relating to *training images*, considered state of the art in the field of geostatistics (Hu et al., 2016). Training images extend the classical representation of spatial dependence beyond (linear) correlation and are capable of reflecting highly complex structures (Wen & Gómez-Hernández, 1996). Training images are chosen such that they contain all relevant information about the spatial dependence one wishes to impose. Then they are *cut*, *copied*, and *pasted* in a suitable random fashion to simulate realizations of random space functions.

In this study, instead of using an actual image, we replace the idea of a training image with an archive of fine-scale particle trajectories that contain all information about memory of speed and direction. These trajectories are obtained from highly resolved direct numerical simulations in a small sample of the medium of interest. The key idea to this approach is to randomly sample, cut, and then suitably rearrange segments of these trajectories to construct larger-scale random trajectories that naturally reflect the required dependence structure. We assume that all aspects of dependence are contained in the trajectories, and an appropriate rearrangement of trajectory segments reveals trajectories that resemble the original dependence structure. The ensemble of such trajectories represents the transport process. Unlike other SMM approaches, we no longer require a transition matrix. The manner in which we cut, copy, and paste the trajectories is conceptually similar to how training images are used to generate spatial random fields, and therefore, we call our method the training trajectory approach (TTA).

Continuous time random walk (CTRW) models are widely and often implemented to model anomalous transport processes in porous media (e.g., Berkowitz & Scher, 1998; Berkowitz, 2002; Bijeljic & Blunt, 2006; Dentz et al., 2004; Lester et al., 2016; Nœtinger & Estebenet, 2000). As most often classically implemented, a particle adopts a random transit time for a random particle displacement in each simulation step, where both—the transit times and the displacements—are modeled as an independent stochastic process (Berkowitz et al., 2006; Sokolov, 2012). For this to be representative of heterogeneous systems, each spatial displacement must be larger than the length scales over which velocities or transit times are statistically dependent (e.g., correlated). However, for the independence assumption to hold, the spatial resolution

of CTRW-based transport simulations is limited (from below) by a spatial scale defined by a correlation length λ_{ind} (Berkowitz et al., 2006; Bolster & Dentz, 2012; Bolster et al., 2014; Le Borgne et al., 2011; Schumer et al., 2001; Sund et al., 2017b). It is important to note that the representation of transport as a CTRW does not mean at all that correlation is neglected. CTRWs reflect process dependence below the scales defined as λ_{ind} in the appropriate choice of the transit time functions for a particle to leave a state. Yet, in many instances, the scale λ_{ind} exceeds scales we are interested in, so that we cannot meaningfully use an uncorrelated model (Cushman & O'Malley, 2015; Schumer et al., 2001). One can overcome such restrictions by incorporating dependence, for example, a SMM in which the subsequent transit time depends on the current one (Bolster & Dentz, 2012; Kang et al., 2014; Sund et al., 2016). Even though SMMs refine spatial resolution, there are still limitations on the smallest possible scale of which they are valid. Advective-diffusive transport in porous media follows as a Gauss-Markov process (with drift), which is characterized by an autocorrelation that decays exponentially (Doob, 1942; Sund et al., 2016; Wang & Stock, 1992) and hence the resolution of an SMM cannot be finer than one over which this holds. Another challenge with SMMs is parameterization of the transition matrix that describes dependence. To thoroughly describe the governing statistics of a three-dimensional transport problem, one must parameterize a nine-dimensional transition matrix to represent transitions in all Cartesian directions (e.g., *x*-*x*, *x*-*y*, *x*-*z*, *y*-*x* … ; Most et al., 2016). This is not trivial and, in most applications, the transport process is simplified, for example, by reduced dimensions (e.g., Le Borgne et al., 2008b) or diffusion is neglected (e.g., De Anna et al., 2013). While pure diffusion or pure advection simplifies modeling, the interplay between advection, diffusion and the complex pore space induces a dependence into particle that goes beyond plain linear correlation. The complex parameterization of this transition matrix might be why, to date, spatial Markov models are seldom applied to build effective three-dimensional models of finite-Péclet transport. As such, most applications to date are of reduced dimensional order, typically one dimension aligned with the principle direction of flow. In this case, parameterization of the transition matrix is substantially easier (while not trivial). However, we argue that the ultimate goal should be the simulation of three-dimensional particle motion. For this we propose an alternative parametrization of a SMM.

Our proposed training trajectories approach extends state-of-the-art SMMs as follows: (a) It is a model that can simulate upscaled three-dimensional finite-Péclet transport, incorporating high-order process memory of speed and direction. Upscaling usually refers to much larger averaging volumes (e.g., at the meter or even at the kilometer scale) than used in this study, where the averaging length is $8 \cdot 10^{-4}$m. In its current form, the proposed method is not applicable to the simulation of transport problems at the aquifer scale. The generalization of this proposed method to such large scales of practical interest requires further work. (b) Training trajectories do not require a high-dimensional transition matrix and (c) we can represent particle motion at scales below the validity scale of spatial Markov models as we use a segment of a real trajectory to describe the particle path between states instead of a linear interpolation (the average location) as used in other SMMs. This last point is relevant for phenomena that are driven by details at small scales, such as dilution (Kitanidis, 1994), mixing (Le Borgne et al., 2010), and mixing-limited reaction kinetics (Sund et al., 2017a). An accurate representation of such small-scale processes requires more than just the average location of a particle as they are nonlinear quantities and must be able to resolve subscale variability in concentrations accurately. (see Figure 2).

Therefore, we will test the performance of our approach not only against metrics that depend on mean concentrations or one-particle statistics such as arrival time distributions (e.g., breakthrough curves) but also against higher-order nonlinear metrics such as the dilution index (Kitanidis, 1994). We also analyze particle-pair statistics, for example, the temporal evolution of the separation distance of particle-pairs conditioned on their initial separation.

In the next section we give a detailed description of the proposed training trajectories approach. Then, we briefly describe the direct numerical simulation (DNS) at the pore scale from which we obtain the training trajectories and which we use as the reference simulation for testing of our proposed approach. We proceed with the results and discuss comparisons of the one-particle and particle-pair statistics. Finally, we draw conclusions.

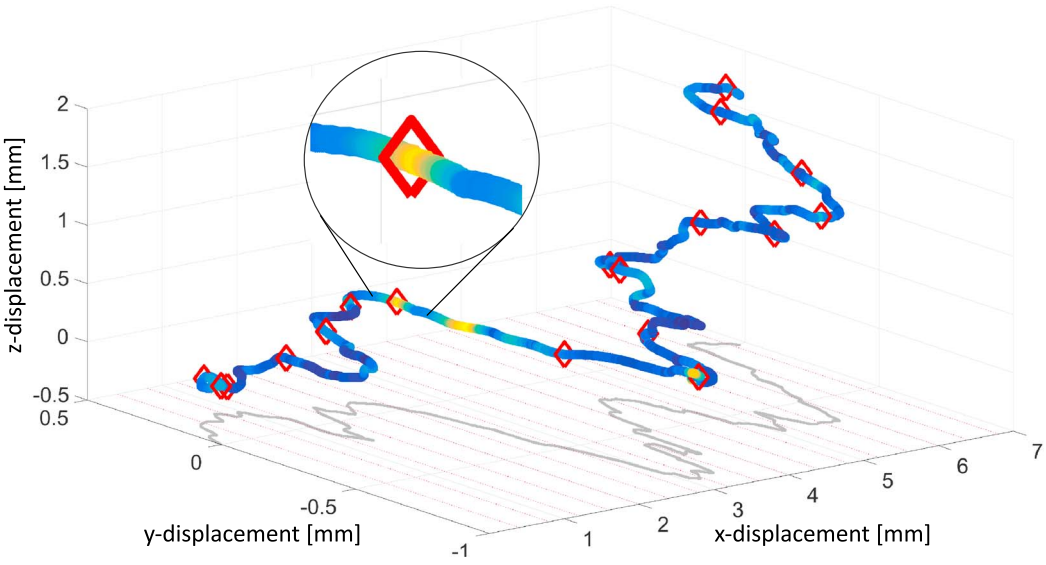

Figure 1. Resampled trajectory: The red diamonds subdivide the full trajectory into trajectory segments of length λ_v. The magnified area focuses on the intersection between two segments which is smooth in speed and direction. The red dotted lines on the ground indicate spatial intervals of length λ_v.

2. The Training Trajectory Approach

Advances in digital rock physics together with DNS of pore-scale flow and transport in porous media can simulate highly realistic particle trajectories (Bijeljic et al., 2011; Mostaghimi et al., 2012; Pereira Nunes et al., 2015). The details of the DNS are outlined in section 3 but, to summarize, we have to conduct thee major steps to produce trajectories: (i) discretize the pore geometry obtained from micro CT-imaging, (ii) solve the Navier-Stokes equations to obtain the flow field, and (iii) run a particle tracking random walk simulation. By tracking particle positions over time, we build what we refer to as our *training trajectories*. Full trajectories can then be divided into smaller trajectory segments of length λ_v (Figure 1). These segments are then used to build a database (or archive) from which the increments within our SMM model are sampled. The full training trajectories inherently contain the processes that form a "larger-scale" perspective. To take advantage of this information, we propose an SMM that cuts, copies, and pastes segments of the training trajectories in a manner that imposes continuity of velocity and direction, so that the resulting simulated trajectories inherit relevant properties and statistics from the full DNS-based trajectories, resulting in an effective model true to the small-scale physics.

A critical part of this model is the resampling approach (cut, conditional copy, and randomized paste) that we propose. To avoid unphysical features at the intersections of trajectory segments (e.g., sharp edges and infinite accelerations; see Figure 2, top) subsequent segments should be sampled conditional to the previous one, typically based on correlating consecutive velocities. This results in smoother transitions between simulation steps (Figure 2, bottom). Doing this with average velocities per segment, however, would not yet ensure total smoothness because the average velocity need not align with the velocity at the end of a specific segment. It is this final part of the segment that is important if we want to ensure physically realistic continuity of trajectory segments that are pasted together. Here we propose a methodology that reflects this.

We can subdivide our simulation scheme in three major steps: cut, conditional copy, and paste. First, from the full-length trajectories we cut out trajectory segments so that the series of appropriately merged segments can reflect the original process dependence (Figure 1 and more details given in section 2.1). Second, we choose (conditional copy) the upcoming trajectory segment such that we avoid sharp transitions in speed and direction between two segments (details in section 2.2). We ensure smooth transitions (i.e., smooth transitions in velocity and direction) by randomly selecting the upcoming segment (from the training trajectory archive) conditioned on the Cartesian velocity triplet of the current segment end. Third, we append (paste) the selected trajectory segment (details in section 2.3). We sequentially repeat this procedure many times

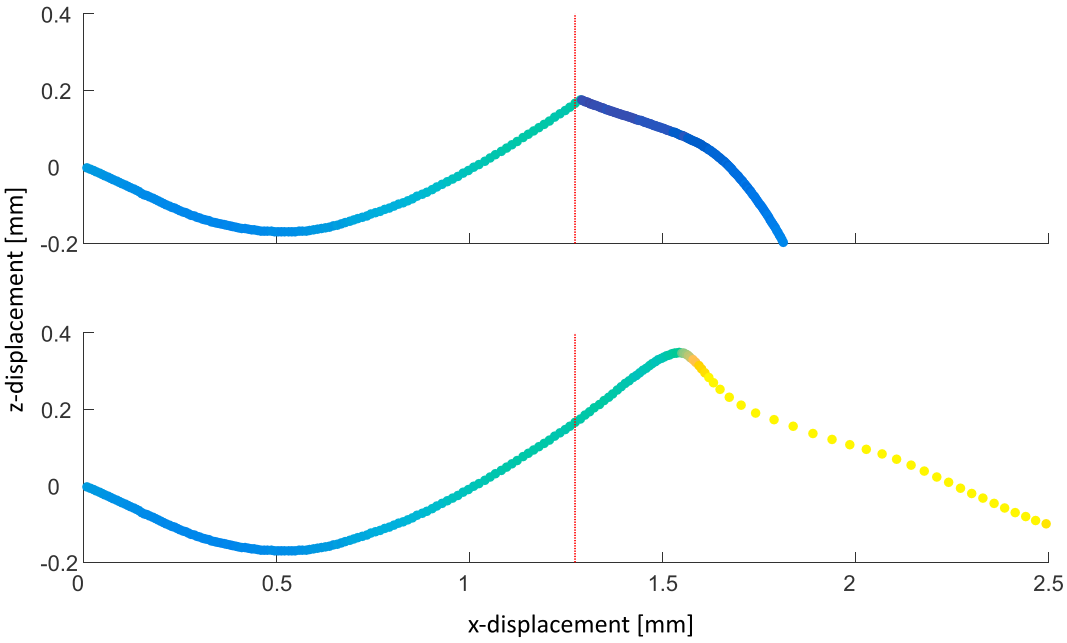

Figure 2. (top) Unconditioned transition between trajectory segments results in a sharp transition in speed and direction at the intersection (red dotted line). (bottom) Smooth transitions between trajectory segments generated by the TTA.

to generate long trajectories. For the generation of an ensemble of full trajectories, we repeat the procedure multiple times.

These three steps require us to identify two characteristic lengths to form a valid Markov process that generates physically reasonable trajectories: (a) the segment length λ_v and (b) the length of the segment end λ_α that is relevant for the conditional random selection of the next segment (Figure 3).

2.1. Step 1. Cutting: Identifying the Relevant Scales

For a valid spatial Markov process two assumptions must hold: (1) The spatial correlation structure must be statistically stationary and (2) the particles next to the velocity state depends solely on the current state (Renner et al., 2001a, 2001b; Strumik & Macek, 2008; Sund et al., 2017b). In this study we follow the methodology of Sund et al. (2016) and define the scale over which these two assumptions are valid by making use of a less commonly used metric—the autocorrelation K of the particles' velocity along the trajectory,

$$\mathrm{K}(j,i) = \frac{\mathrm{Cov}\left(v_i(\lambda), v_j(\lambda)\right)}{\sqrt{\mathrm{Cov}\left(v_i(\lambda), v_i(\lambda)\right) \cdot \mathrm{Cov}\left(v_j(\lambda), v_j(\lambda)\right)}} \tag{1}$$

where $\mathrm{Cov}\left(v_i(\lambda), v_j(\lambda)\right)$ is the covariance of $v_i(\lambda)$ and $v_j(\lambda)$. $v_i(\lambda)$, and $v_j(\lambda)$ are the velocities averaged along trajectory segments i and j with the length λ. As we assume statistical stationarity, $K(j,i) = K(|j-i|)$, all possible pairs of segments separated by $|j-i| \cdot \lambda$ for $|j-i| \in \mathbb{N}$ along the full-length trajectories have the same correlation. The velocity $v_i = \frac{\lambda}{\tau_i}$ with the transit time τ comes from the DNSs.

To make use of the autocorrelation as an indicator, we exploit that particle motion is a normal process controlled by the local velocity and diffusivity. The spatial coordination of the velocity field defines the number of simulation steps a particle stays within a certain velocity class. The autocorrelation of a particle's velocity $v(\lambda)$ is a function of the averaging length λ. For very small λ, we can observe an independent process ($\mathrm{K} \approx 0$) for all $i, j \in \mathbb{N}$ as particle motion on very small scales if dominated by diffusion. From there, the autocorrelation K increases with increasing λ as advection takes over process control and introduces memory into the velocity process. For even larger λ, the particle loses its memory and the particle motion can be considered again as independent (K converges back to zero; Sund et al., 2016). For a certain intermediate λ_v, however, advective-diffusive particle motion can be described as a Gauss-Markov process, if the autocorrelation $K(\lambda)$ decays exponentially (Doob, 1942; Sund et al., 2016; Wang & Stock, 1992), and hence, the resolution of a

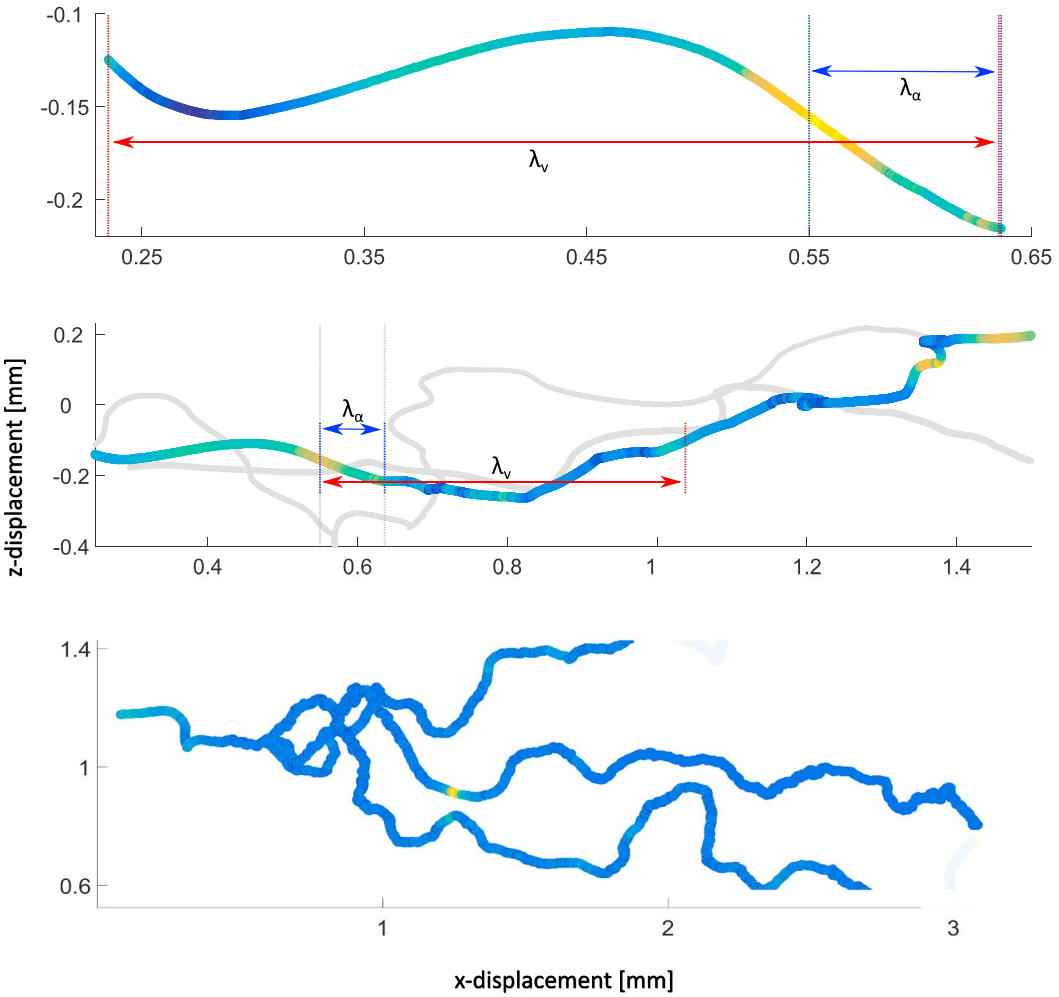

Figure 3. (top) Trajectory segment of length λ_v over which the velocity process forms a spatial Markov process. Subsegment length λ_α is the scale over which the component-wise velocity correlation is maximal. (middle) Sketch of the merging procedure between trajectory segments. At the juncture of two segments, we always exchange the first part (λ_α) of the next segment with the last part (λ_α) of the current segment. This ensures smoothness. The merging of all other trajectory segments (e.g., the gray trajectories) occurs in the same spatial intervals ($I = [n \cdot (\lambda_v - \lambda_\alpha) n \cdot \lambda_v]$). (bottom) Three trajectories evolving from the very same initial trajectory segment to show that the TTA does not simply copy the original trajectory based on the chosen initial segment.

spatial Markov process cannot be finer than one over which this holds. A Gauss-Markov process can be expressed as an autoregressive model with a one-step memory (AR(1)) in which the noise term is Gaussian. Hence, to express the velocity process as an AR(1), we require the autocorrelation function K between velocities averaged over a distance λ and separated by a distance $|j - i| \cdot \lambda$ to be exponential, such that

$$K(j,i)|_\lambda = K(|j-i|)|_\lambda = \exp\left(a_\lambda |j-i|\right) \tag{2}$$

where $a_\lambda \in \mathbb{R}^-$ can be related to the second largest eigenvalue of the (doubly stochastic) transition matrix $T(|j-i|)$. This eigenvalue describes the convergence of T toward the uniform stationary matrix (Kang et al., 2015a, 2015b). Parameter a_λ therefore uniquely parameterizes the transition matrix of the underlying Gauss-Markov process.

Now we need to relate $K(i,j)|_\lambda$ to $K(\lambda_v)|_{1,2}$ in order to define the scale λ_v over which we obtain the valid spatial Markov process. While $K(i,j)|_\lambda$ is the velocity autocorrelation in the proper sense, $K(\lambda)|_{1,2}$ is simply the velocity correlation between two adjacent trajectory segments of averaging length λ, but it is calculated the same way (see (1)). Sund et al. (2016) provide this link by analyzing $K(\lambda_v)|_{1,2}$ as a function of multiples of (a small) λ. They found that a_λ is in $\mathbb{R}^-$ and therefore parameterizes the transition matrix T only if

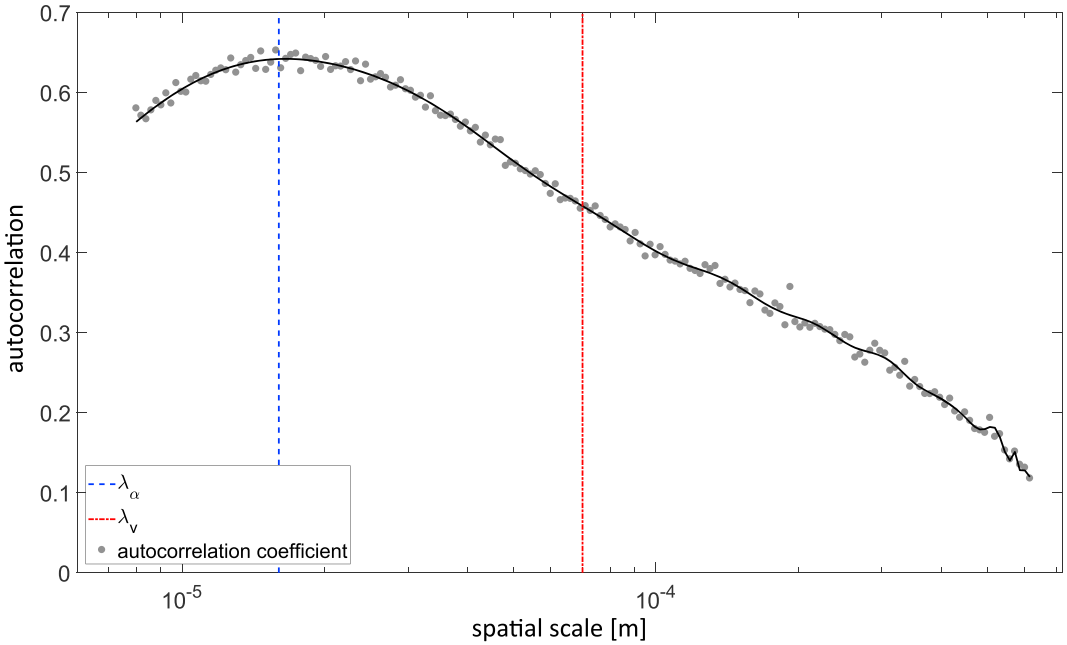

Figure 4. Evolution of the autocorrelation over space in longitudinal direction. The blue dashed line indicates the highest correlation and hence defines $\lambda_{\boldsymbol{\alpha}}$. The red dashed line lies in the range in which $K(\lambda)|_{1,2}$ decays monotonically with a positive second derivative and indicates our choice of λ_v. The full length of the trajectories is 2.69 mm and the λ_v lies between 1 and 2 typical grain sizes.

$K\left(\lambda_v\right)|_{1,2}$ decreases monotonically and with a positive second derivative. Hence, a valid λ_v must be chosen from the range over which the shape of the correlation function $K\left(\lambda_v\right)|_{1,2}$ is concave, as here the correlation function decreases monotonically and with a positive second derivative (see Figure 4). These are the fundamental requirements that we are subject to when selecting the averaging length λ_v. The choice of λ_v is not unique and the model accuracy, predominately regarding the late time behavior of the particle plume, appears to improve with larger λ_v. This effect of an increasing averaging length λ_v is shown in the supporting information and a more detailed discussion of how the choice of λ_v affects the overall simulation results is given in section 4.1.

2.2. Step 2. Conditional Copy: Identifying the Transition Probabilities Between Segments

A second characteristic length λ_α is essential to determine the transition probabilities between consecutive trajectory segments in our SMM approach. Under advective influence, particle trajectories become smooth in direction and velocity after a scale smaller than λ_v. This is the scale, $\lambda_{\boldsymbol{\alpha}}$, at which the Péclet number indicates that advection dominates over diffusion, but well before larger-scale dispersion becomes relevant. We identify this scale by determining the peak of the autocorrelation function in (1) (Sund et al., 2017b). We ensure smooth transitions in direction and velocity by merging trajectory segments only if the Cartesian velocities at the contact faces are similar in velocity and direction over that scale. Therefore, we average the velocities $\bar{v}_{\boldsymbol{\alpha}} \in \mathrm{R}^{d\times 1}$ component wise (x, y, and z directions; $d = 3$) over each segment end and beginning over length λ_α. Note, the averaging length λ_α in the main flow direction is a constant, and as a result the respective time ($\tau_{\boldsymbol{\alpha}}$) over which we average varies for each subsegment we analyze.

Process memory is induced by advection, as diffusion is a fully memoryless process. Therefore, we store only the advective components s_A^l of the trajectory start segments (averaged over length $\lambda_{\boldsymbol{\alpha}}$) in an archive with $l = 1 \ldots n_{\mathrm{arc}}$ entries. As mentioned before, we ensure smoothness by merging segments with similar Cartesian velocities $\bar{v}_{\boldsymbol{\alpha}}(s)$.

Hence, the transition probability $p(s_{\mathrm{A}}^l|s_{\mathrm{AD}})$ of appending trajectory segments with start s_{A}^s must be proportional to the similarity between the velocity triplet of the current advective-diffusive subsegment s_{AD} and all candidate subsegments s_{A}^l. Here, similarity will be defined based on the tolerance introduced by diffusion within the time it takes advection to cover the distance λ_α. An example for a low-transition probability $p(s_{\mathrm{A}}^l|s_{\mathrm{AD}})$ can be found in Figure 2 (top). The velocities averaged over λ_α at the left (s_{AD}) and the right (s_{A}^l) of

the red-dotted line results in two vectors that point in very different directions. The same averaging applied to another potential upcoming trajectory segment shown in Figure 2 (bottom) results in two vectors that point much more in the same direction which indicates a much higher transition probability $p(s_{\mathrm{A}}^{l}|s_{\mathrm{AD}})$.

To define the transition probability $p(s_{\mathrm{A}}^{l}|s_{\mathrm{AD}})$ based on the physics of our system, we must quantify the likelihood that the discrepancy between purely advective velocities in s_{α} and advective-diffusive velocities s_{AD} comes from diffusion alone which acts for a certain amount of time $\tau_{\boldsymbol{\alpha}} = \frac{\lambda_{\boldsymbol{\alpha}}}{\bar{v}_{\boldsymbol{\alpha}}}$ a particle needs to pass $\lambda_{\boldsymbol{\alpha}}$. We quantify this likelihood by exploiting that molecular diffusion follows a Gaussian process. Hence, we can quantify the likelihood of diffusion being the cause for the velocity discrepancies via the expression of the Gaussain likelihood:

$$p\left(s_{\mathrm{AD}}|s_{\mathrm{A}}^{l}\right) = \left(2\pi\sigma^{2}\right)^{-\frac{d}{2}} \cdot \exp\left(-\frac{1}{2\sigma^{2}}\sum_{i=1}^{d}\left(s_{\mathrm{AD,i}} - s_{\mathrm{A,i}}^{l}\right)^{2}\right) \tag{3}$$

where $\sigma(\tau_{\boldsymbol{\alpha}}) = \frac{d}{2}D_{\mathrm{m}}\tau_{\boldsymbol{\alpha}}$ describes the standard deviation of diffusive particle displacement in three dimensions ($d = 3$) over a time $\tau_{\boldsymbol{\alpha}}$ (Einstein, 1905).

We derive the transition probabilities $p\left(s_{\mathrm{A}}^{l}|s_{\mathrm{AD}}\right)$ between the current segment s_{AD} and all candidate segments s_{A}^{l} by applying Bayes rule

$$p\left(s_{\mathrm{A}}^{l}|s_{\mathrm{AD}}\right) \propto p\left(s_{\mathrm{AD}}|s_{\mathrm{A}}^{l}\right) \cdot p\left(s_{\mathrm{A}}^{l}\right) \tag{4}$$

where $p\left(s_{\mathrm{A}}^{l}\right)$ is the prior probability of s_{A}^{l} being the best-fitting upcoming trajectory segment. Here, s_{A}^{l} is the candidate segment (of length λ_{v}) defined by the lth purely advective (subscript A) velocity triplet in the archive where $l \in (1, n_{arc})$. s_{AD} is the start of the current segment that contains advection and diffusion indicated by the subscripted AD. With this expression, we condition all potential upcoming segments on the last part (of length λ_{α}) of segment s_{AD}. The prior probability $p\left(s_{\mathrm{A}}^{l}\right)$ is uniformly distributed over the archive as, before relating s_{AD} to s_{A}^{l}, all candidate subsegments from the training trajectory archive are equally likely. As a consequence and without loss of generality, we can erase $p\left(s_{\mathrm{A}}^{l}\right)$ from equation (4).

Altogether, we determine the transition probabilities of our spatial Markov process by applying Bayes rule together with the likelihood of a normal process. The transition procedure between consecutive trajectory segments can be understood as a classical conditional sampling in which the upcoming segment is drawn conditional to the velocity observed in the end part of the current trajectory segment (i.e., one-step memory). This random conditional sampling ensures that we do not simply copy trajectories starting from the initial trajectory segment as we show in Figure 3 (bottom).

2.3. Step 3. Paste: Generating the Simulated Trajectories

In the remaining step, we merge the trajectory segments. Here, we exchange the first part (λ_{α}) of the next segment with the last part (λ_{α}) of the current segment (Figure 3). By doing so, we generate smooth trajectories that inherently encode the full memory of speed and direction of the underlying pore-scale transport process.

3. Obtaining the Training Trajectories by Direct Numerical Simulation

To demonstrate our approach, a DNS of flow and transport through a sample of Doddington sandstone provides the necessary archive of training trajectories and also acts as a benchmark. The DNS incorporates detailed information about pore geometry obtained from high-resolution X-ray tomography (Bijeljic et al., 2013). The resulting 3-D image of the sample comprises 1,000 spatial voxels along each Cartesian coordinate. The voxel length of $d_{\mathrm{voxel}} = 2.6929 \cdot 10^{-6}$ m results in a total of 10^{9} voxels, and in a total image length of $2.6929 \cdot 10^{-3} m$ in each direction. The porosity of the sample is $\phi = 0.195$ (Alhashmi et al., 2016). Each voxel becomes a grid block in the flow simulation, where we solve the Navier-Stokes equations for incompressible flow as described in Raeini et al. (2012) and Bijeljic et al. (2013).

Next, we simulate transport with a random walk method (Bijeljic et al., 2011), tracking advective and diffusive particle displacements with constant time steps $\Delta t = 1 \cdot 10^{-4} s$. Particle positions are updated as

$$\boldsymbol{X}(t + \Delta t) = \boldsymbol{X}(t) + \Delta\boldsymbol{X}_{\mathrm{adv}} + \Delta\boldsymbol{X}_{\mathrm{diff}}. \tag{5}$$

In each time step, particles are moved advectively by displacement ΔX_{adv} along streamlines using a modified Pollock algorithm (Pereira Nunes et al., 2015). This is followed by a diffusive step $\Delta \boldsymbol{X}_{\text{diff}} = \sqrt{24 D_m \Delta t}(R - 0.5)$ with $D_m = 2.2 \cdot 10^{-9} \frac{m^2}{s}$ as the molecular diffusion coefficient and R a uniformly distributed random number. We assume that this is sufficiently accurate to neglect discretization error, which has been validated in previous studies (Bijeljic et al., 2004; Mostaghimi et al., 2012).

The transport regime we work with is characterized by the Péclet number $Pe = \frac{u_{\text{av}} \cdot L}{D_m}$ with characteristic length $L = 200.18 \cdot 10^{-6}$m given by the grain size (Alhashmi et al., 2016; Mostaghimi et al., 2012). For the reference transport simulation at $Pe = 100$, we set the volume averaged velocity to $u_{\text{av}} = 1.09896 \cdot 10^{-3} \frac{m}{s}$. We simulate 1,000 particles, following them until they exit the downstream image face. We inject the particles flux-weighted into the 30th image voxel layer to avoid backward directed diffusion across the upstream image face. A more detailed description of the flow and transport simulation is given in Most et al. (2016).

4. Results and Discussion

4.1. Velocity Autocorrelation

First, we determine the characteristic length λ_{v}, the length (in the main flow direction) over which we have to average the velocities along the trajectories so that the autocorrelation $\text{K}(j,i)|_{\lambda_v}$ decays exponentially. The measured correlation as a function of increasing averaging lengths λ for $Pe = 100$ is shown in Figure 4. For $\text{K}(j,i)|_{\lambda_v}$ to decay exponentially, the velocity correlation $K(\lambda)|_{1,2}$ between two adjacent trajectory segments averaged over λ at least has to decay monotonically with a positive second derivative (see section 2.1 or Sund et al., 2016). The correlation function $K(\lambda)$ appears to start being convex ($K''(\lambda) > 0$) at $\lambda \geq 6 \cdot 10^{-5}$m and is clearly convex at $\lambda_{\text{v}} = 8 \cdot 10^{-5}$m which represents 3% of the total sample length. We keep λ_{v} as small as possible in order to merge as many trajectory segments as possible (roughly 34).

Second, we determine the characteristic length λ_{α}, required to ensure smooth transitions (see 2.2) by identifying the interval over which the autocorrelation is maximum; here $\lambda_{\alpha} = 1.6 \cdot 10^{-5}$m. This represents 0.6% of the total sample length. Given these values we can now run our proposed training trajectory model.

In our advectively dominated scenario ($Pe = 100$), diffusion is the reason why we have a peak in the autocorrelation function at λ_{α}. The location of the peak λ_{α} and of the inflection point λ_{v} (of $K''(\lambda)|_{1,2}$) move in positive direction with increasing influence of diffusion. Therefore, the characteristic lengths λ_{α} and λ_{v} are shorter when derived from an in-average faster trajectory than when derived from a slow trajectory where diffusion has more time to act. Note, we derive λ_{α} and λ_{v} by averaging over all trajectories, and for this reason we potentially underestimate characteristic lengths for "slow" trajectories. In the supporting information, we show that a larger λ_v improves the accuracy of the late time arrivals. A larger λ_v better captures the influence of the diffusively dominated slow trajectories, while the influence of advectively dominated fast trajectories is already captured within the smaller λ_v.

Next we describe and compare several benchmark metrics to test and validate our proposed modeling approach.

4.2. Breakthrough Curves

We simulate the arrival time statistics at three control planes for a Péclet = 100 transport process. Breakthrough curves (BTCs) are measured at downstream distances corresponding to 30%, 60%, and 100% of the total simulation domain. For testing, we use 5,000 trajectories generated by the proposed training trajectory approach. For comparison, we produce 5,000 trajectories for which we select the consecutive trajectory segments randomly (*unconditional sampling*) from the archive of trajectory segments. The BTCs predicted by the training trajectories and by the DNS-based reference solution (1,000 trajectories) show good agreement at all three control planes (Figure 5). There are no major deviations between model and benchmark in either early or late arrivals, which are typically the aspects where strong non-Fickian transport signatures emerge, which are often the hardest to capture. The predicted BTCs deviate, as might be expected, from the benchmark when we select consecutive trajectory segments unconditionally.

To accurately predict arrival times at control planes in the longitudinal direction, our model has to reflect typical residence times in low-velocity (leads to tailing) and high-velocity zones (leads to early arrivals). Memory of speed is nothing other than these typical residence times that we aim to mimic. Hence, our BTC results indicate that we are realistically representing memory of speed with our training trajectory approach. The comparison between the conditional and unconditional TTA predictions versus the benchmark clearly

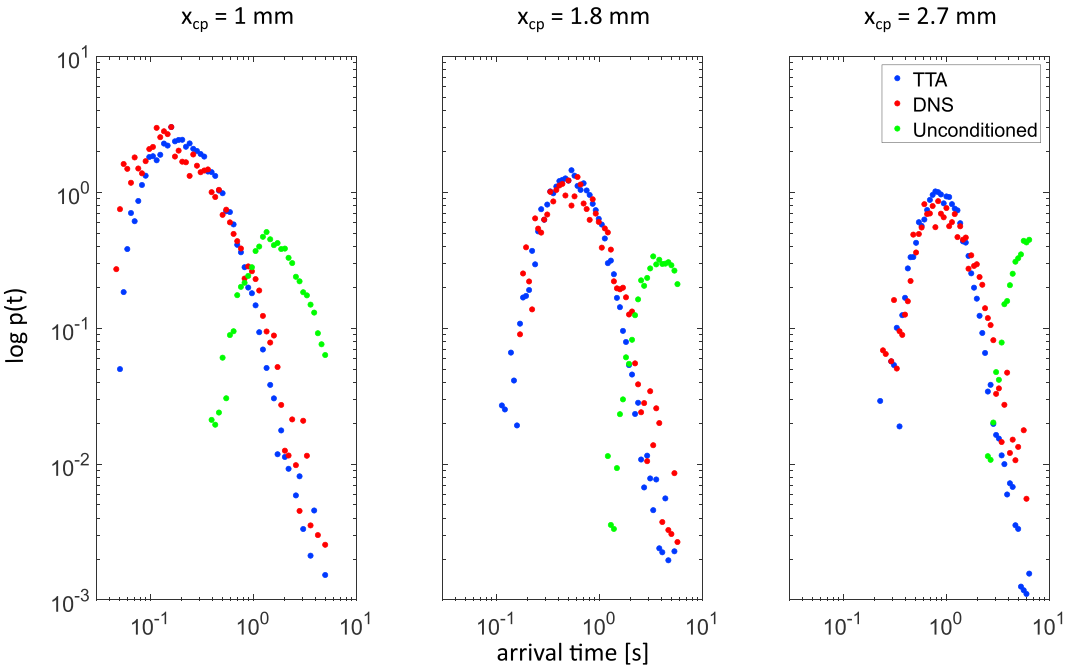

Figure 5. Comparison of the arrival time densities between the reference simulation (DNS, direct numerical simulation) and our training trajectories approach (TTA) at three control planes at 30%, 60%, and 100% of the simulation domain (≈2.7 mm).

suggest that the spatial Markov approach (i.e., conditional sampling) is the key step that introduces the required process memory to reflect the characteristic residence times in low- or high-velocity zones.

Even if the agreement of the arrival times is very good, we slightly underestimate the late-time arrivals. This is presumably an effect of underestimating λ_α, but especially λ_v, for the "slow" trajectories (see section 4.1) which control the late time behavior of the particle plume. To test this hypothesis we increase λ_v by a factor of 3 and 6 (constant λ_α), which clearly improves the match, particularly in the late time behaviors (see the supporting information). This suggests that perhaps λ_α and λ_v should not be derived from averaging over all trajectories but rather accounting for the slow ones in particular.

4.3. The Dilution Index

Matching BTCs indicate that the model does a good job of representing particle motion in direction of mean flow. Yet it tells us nothing about the ability of the model to upscale more complex nonlinear behaviors that depend on the full three-dimensional nature of the particle plume. We will now further test the TTA's ability to accurately model such behaviors.

To this end, we will compare the evolution of the dilution index (Kitanidis, 1994) over time as measured from the DNS and predicted by the TTA. The dilution index is a metric that describes how dilute a plume is, or in other words measures the total volume occupied. It is based on the concept of entropy that classically describes the disorder in a system. Dilution is controlled by spreading and mixing, both of which are tightly coupled (Le Borgne et al., 2010). We estimate the dilution index in discrete form. That means, we perform a histogram-type analysis where we count the number of particles in each cubic bin of 10^3 voxels which relates to a typical pore size. For the histogram-type estimation, the dilution index is defined as

$$E = \Delta V \cdot \exp\left(-\sum_{k=1}^{r} P_k \cdot \ln\left(P_k\right)\right) \tag{6}$$

where ΔV is the volume of a voxel and P_k is the fraction of particles in histogram bin k divided by the total number of particles. The dilution index is an excellent metric for testing the veracity of effective modeling approaches as it is highly sensitive to the full-dimensional nature of the plume and subtly reflects the complex interplays between spreading and mixing (Kitanidis, 1994), which most effective modeling approaches fail to capture.

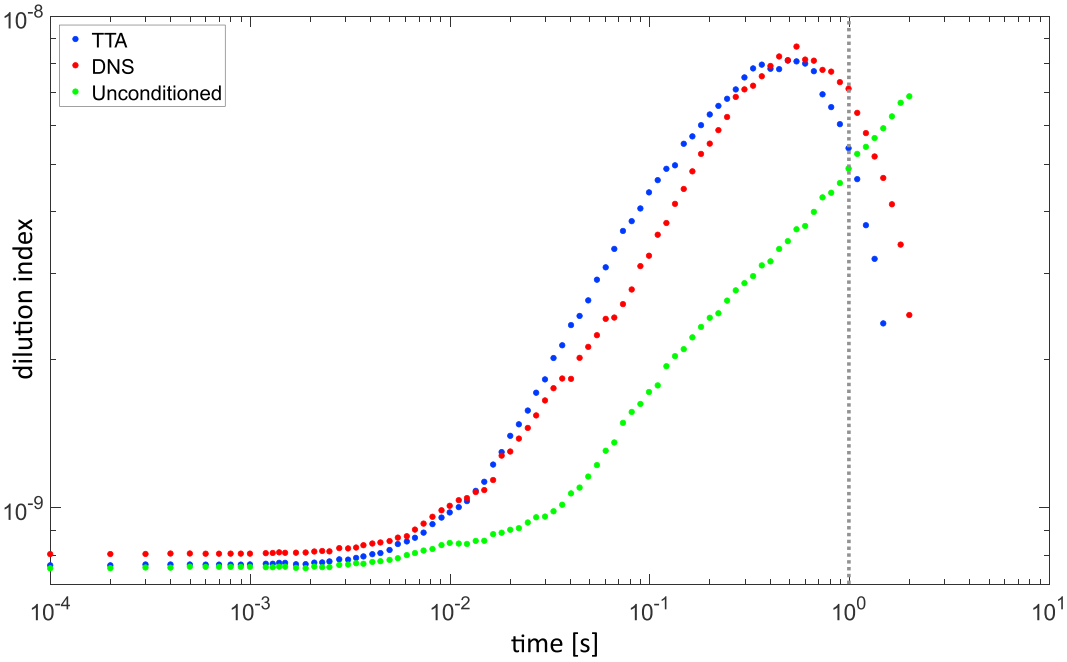

Figure 6. Temporal evolution of the dilution index derived by the reference simulation (DNS, direct numerical simulation) and our training trajectories approach (TTA).

In Figure 6 we plot the temporal evolution of the dilution index based on the TTA predictions, on its unconditioned counterpart, and on the DNS reference. In general, again, the agreement between the TTA and the DNS reference is close while we see clear a deviation when we merge trajectory segments unconditionally.

However, the agreement between TTA and DNS for this benchmark is less good than for the BTCs. The reason is that, in contrast to the DNS, the TTA does not account for reflection at no-flow boundaries and the accessible space (number of bins) for particles is therefore unconstrained. Also, the dilution index is a nonlinear measure that is very sensitive to low concentrations (P_k enters equation (6) with $\ln(P_k)$). Therefore, minor deviations between TTA- and DNS-based particle motion already lead to visible deviations in the dilution index. Again, as for the BTCs, the agreement is better for larger λ_v.

In Figure 6, the dilution index decreases after approximately 0.7 s. This seemingly unphysical behavior (reappearance of order) comes from particles that exit the simulation domain and are thus removed from the system. However, the fact that DNS and TTA, especially for larger λ_v where the Gauss-Markov assumption also holds for slow particles, show this behavior in a very similar way. This fact still demonstrates the veracity of the TTA approach and suggests that we capture the desired memory processes also in three dimensions where particle motion is controlled by memory of speed and direction plus diffusion. Again, our results indicate that conditional sampling is significant to correctly incorporate memory of speed and memory of direction, necessary to accurately predict full three-dimensional transport processes.

4.4. Particle Pair Statistics

Both BTCs and the dilution index are measures derived from the evolution of an ensemble of individual particles. Next, we look at the interaction between particles by analyzing the separation distance of two particles (i.e., particle pairs). This particle pair perspective of transport is important because it provides insights on how particles diverge and converge, which can play an important role in understanding mixing and reaction processes. Additionally, it is a further and robust test of our approach to capture nontrivial dynamics that go beyond just predicting mean behaviors.

Our key analysis metric is the probability density function $p(s^t|s^0)$ of the Euclidean separation distance s^t (at discrete time steps t) conditioned on the initial separation distances s^0. In the following, we will refer to $p(s^t|s^0)$ as the *fingerprint of dilution*. In Figure 7a, we plot an example of this metric for an individual particle separation of 1.0 mm $< s_0 <$ 1.1 mm. Each cross section orthogonal to the time axis (e.g., white dashed lines at $t = 0.05$, $t = 0.1$, and $t = 0.7$ s) represents a separation density $p(s^t|s_0)$ of particle pairs at these

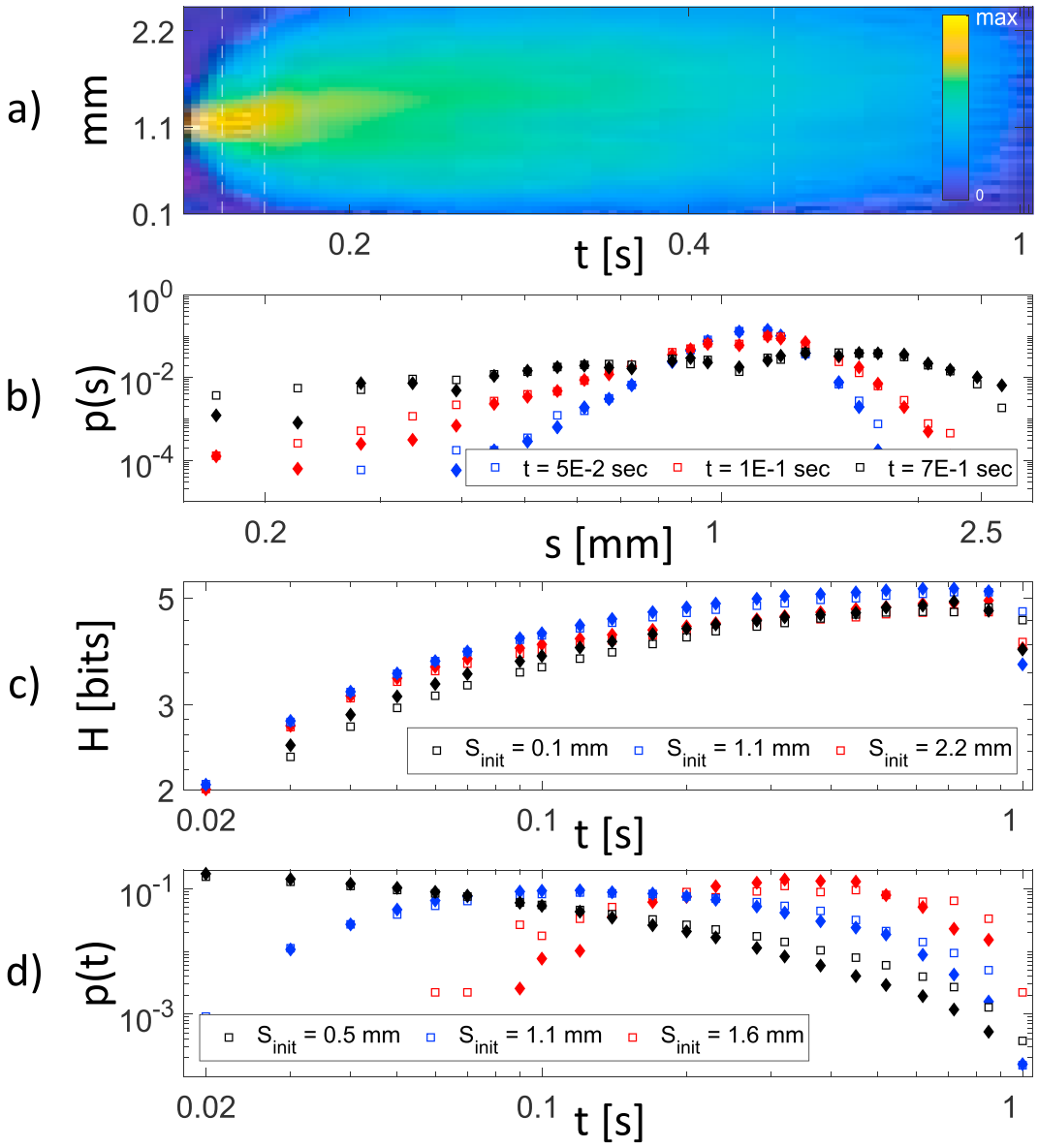

Figure 7. In general, the filled symbols represent the results for the training trajectories approach and the outlined symbol represent the results from the DNS. (a) Fingerprint of mixing: This figure shows the temporal evolution of separation distance density function $p(s^t|s^0)$ of particle pairs conditioned on an initial separation of $s^0 \approx 1.1$ mm. (b) Pair-wise separation-distance density $p(s^t|s^0)$ with initial separation $s^0 \approx 1.1$ mm at three different points in time. (c) Entropy over time as a measure for similarity between reference separation-distance density and the according pdf derived from the training trajectories. (d) Reaction potential over time given an initial pair-wise separation distance.

points in time. The horizontal line indicates the distance (s_{reac}) below which we assume potential reaction of particles, chosen as half of a typical pore size (0.13 mm) for our sample (Edery et al., 2014).

Figures 7b–7d highlights three important measures related to the fingerprint of dilution, comparing results measured from DNS to predictions made by the TTA. These three metrics are as follows:

1. The separation distance densities $p(s^t|s^0)$ shown at three different times for an initial separation distance 1 mm < s_0 < 1.1 mm, which as noted correspond to the white vertical lines in Figure 7a.
2. The evolution of entropy $H(s) = -\sum_{i=1}^{d} p(s^t|s^0) \cdot \log_2 p(s^t|s^0)$ over time for three initial separation distances $s^0 \approx 0.1$, $s^0 \approx 1.1$, and $s^0 \approx 2.2$ mm. We choose these relatively small and large separations to span a broad range of relevant separation distances.

3. The reaction probability defined here as the waiting time distribution of initially separated particles to fall below a fixed distance (here $s_{reac} = 0.13$ mm, as discussed above. Results are shown for three initial separation distances $s^0 \approx 0.5$, $s^0 \approx 1.1$, and $s^0 \approx 1.6$ mm.

Across all three metrics, the agreement between measurements from DNS simulations and predictions with the TTA are excellent, further solidifying the veracity of our proposed TTA, and suggesting that the TTA is capable of accurately simulating the rich dynamics of 3-D particle motion in a complex porous medium flow.

Specifically, the agreement with the separation distance density suggests that the TTA captures characteristic convergence and divergence (i.e., interaction) mechanisms of particles. The agreement on entropy suggests that characteristic diverging and converging mechanisms of particle pairs are not only well captured but also associated with mixing dynamics. The results here also suggest that the TTA captures the complex interaction of particles independent of the initial condition, given the broad range of initial separation distances. Reactions are strongly dictated by how close reacting particles can get, and the probability of reaction increases dramatically when particle distances fall below a given threshold (Edery et al., 2010; Paster et al., 2014). The agreement here between DNS and TTA suggests that the TTA may be well suited to modeling not only conservative transport but also reactive transport processes as well. This it may be coupled with reactive random walk approaches (e.g., Benson & Meerschaert, 2008; Bolster et al., 2016; Edery et al., 2010; Engdahl et al., 2017).

4.5. Challenges and Opportunities

All of the tests that we performed against our benchmarks demonstrate good agreement between TTA predictions and reference DNS measurements, suggesting promise for the TTA approach. However, the way we parameterize the proposed spatial Markov process, while novel, is certainly not comprehensive and still necessitates further work. In our view, remaining research questions for TTA are as follows:

Data Requirements. The approach as we have implemented it has been quite a demanding computational expense. It requires high-resolution simulation of flow and transport for the DNS simulation from which the trajectories are extracted. While the ultimate goal is to expend this computational cost to simulate transport in a single representative elementary volume to achieve upscaling over scales much larger than this, any ability to reduce computational demands at the representative elementary volume scale will be essential. In the context of previous SMM advances, such gains have been made by more efficient numerical models of small-scale processes (e.g., Sund et al., 2017b) or by using simplified analytical models (e.g., Dentz et al., 2016; Kang et al., 2016). At this stage we do not have a clear vision of what such approaches might be for our TTA approach. However, we are hopeful that similar advances are possible here and should be pursued in future research efforts.

Parameterization. We defined the scale over which we describe transport as a Gauss-Markov process by the correlation between the velocities averaged over consecutive trajectory segments. As we do that for all trajectories, the correlation function $K(\lambda)_{1|2}$ we obtain reveals the peak location λ_α for an "average" trajectory. The characteristic length scales λ_α and λ_v obtain like this, however, underestimate the lengths that are required to obtain optimal simulation results, presumably due to the slowest trajectories in the system. As shown, tripling this length scale leads to much better results. Thus, while our approach yields a decent first estimate, an improved and rigorous approach for obtaining these length scales should be sought.

Optimization. The conditional sampling procedure that finds the upcoming trajectory segment is currently computationally intensive and increases with the number of archive entries. Thus, it can become a computational bottleneck in effectively implementing the method. This step, however, can be accelerated by implementing more efficient search algorithms, for example, a k–d tree supported search algorithm that acts on equation (4). Currently, all trajectories have the same prior probability of being the upcoming segment. This is inefficient as the velocity triplet of many candidates deviates strongly from the current reference so that we could exclude them immediately before the conditional sampling through a nonnaive efficient search. This could improve the search on average from being an $O(n)$ to an $O(\log n)$ operation.

Validation. In this study, we have used the same sample to generate the trajectories for TTA parameterization and for the DNS validation. Admittedly, this makes the *validation* somewhat circular in nature, although this is not uncommon for many transport models, particularly in their first implementation. We openly admit that this is not ideal, but computational limitations restrict our efforts to go further. For now,

we accept this compromise, but future efforts must go further. Ideally, if staying at small scales, we would have two distinct sandstone samples (which are statistically similar) and use one to infer the trajectories and the other to validate our simulation, for example, via tracer tests or another DNS. Similarly, to test the true upscaling potential of this approach, it will be essential to use it to compare to larger-scale transport experiments through larger sandstone samples of such column experiments (e.g., Sherman et al., 2018) and ultimately field-scale data (e.g., Kang et al., 2016). For this purpose, however, we require a DNS reference performed on a sufficiently large rock sample to fill an archive that reflects the statistics of particle motion thoroughly. Limited storage capacities, however, do not currently allow for this.

Sufficient Statistics. Similar to the previous comments, it will also be important to analyze what properties (e.g., number, length, initial condition) characterize an ideal set of training trajectories? In the ergodic limit, one very long trajectory should be sufficient to fill the archive. Most likely, the number of trajectories is less important than their initial condition, especially when the domain is limited. It seems that, in our case, 1,000 trajectories are sufficient. But a thorough convergence study will be necessary to better validate this and identify potential error sources and uncertainties.

Generalization. Here, we have applied the TTA only to one 2.6 mm^3 Doddington sandstone, and the resulting velocity field is subject to the real pore geometry. While we see no obvious reasons why the TTA cannot be applied to other porous media than sandstones, for example, carbonates or unconsolidated media, this remains to be truly validated.

Practical Use. If we can overcome some of the identified hurdles and validate the TTA across a diverse range of complex flows, we believe that this approach holds great success in effectively and efficiently predicting large-scale transport in complex porous media flows. In particular, the fact that the TTA can provide a three-dimensional upscaled picture of transport is key to as many effective models of transport are of reduced dimensionality, which can limit their ability to predict nonlinear processes without further complex closures. This is relevant to better represent dilution or reaction process that occur at such fine scales.

Conclusions

In this study we have proposed a trajectory-based spatial Markov approach to simulate three-dimensional transport in porous media through a Doddington sandstone sample at $Pe = 100$. The proposed method borrows ideas from state-of-the-art non-Fickian spatial Markov transport models and blends them with ideas from the concepts of training images. Training images are commonly used in geostatistics and are known for their ability to model highly complex spatial dependence structures that go well beyond linear correlation. This feature is highly desirable in the context of upscaling pore-scale non-Fickian transport as this is often tied to process dependence beyond linear correlation (De Anna et al., 2013).

Instead of a training image, for our proposed method we use highly resolved particle trajectories, obtained from DNS of pore-scale transport. These trajectories inherently contain all relevant information about the dependence structure of the velocity process at pore scales. The key idea is to subdivide particle trajectories into trajectory segments (*cut*) and then sample these to build simulated trajectories. Key to our approach is that we only allow subsequent segments to add to the previous ones in a way that ensures smooth transitions of direction and speed at the intersections (*conditional copy*). By recursively merging (*paste*) trajectory segments, we can generate trajectories of arbitrary length. This way, we can generate trajectories that reflect the same dependence structure as the original trajectories. We use the trajectories to represent the dependence structure of the velocity process in a similar manner that training images are used to represent spatial dependence. Therefore, we call our method TTA.

With the TTA we aim to overcome three frequent drawbacks of other spatial Markov models:

1. We do not require simplifications that have often been invoked in other studies such as dimensionality reduction or neglecting of diffusion.
2. We do not require an explicit parameterization of process dependence, that is, a high-dimensional transition matrix. We incorporate process dependence directly by a merging procedure that ensures transitions are smooth in speed and direction.
3. We represent particle motion at the resolution of the training trajectory and hence can accurately make inferences at scales smaller than other spatial Markov models.

These benefits do come at a price as discussed in section 4.5, including in particular computational costs both in parameterizing and implementing the models. However, as advances in the SMM from the original work of Le Borgne et al. (2008a) have shown, we are hopeful that similar advances can be made to our TTA.

To test the TTA, we compared DNS measurements of a broad range of benchmark metrics against predictions made with our proposed TTA. First, we analyzed statistics of a particle plume including arrival time distributions and the temporal evolution of the dilution index. Additionally, we analyze a slew of particle pair statistics based on the evolution of separation distances of particle pairs over time. In all cases, the agreement between measurements and predictions was good, demonstrating the veracity of our TTA and demonstrating that the TTA can accurately capture the fully dimensional nature of transport along with the complex interplays between spreading and mixing dynamics. Based on our outcomes, we are confident that our method can accurately model the complex three-dimensional dynamics of a particle plume and the interaction between the particles within the plume. Given this, in future efforts it would be interesting to extend the TTA beyond the milimeter scale, to other kinds of porous media and especially to simulate reactive transport where it is known that upscaling efforts must be able to account for these complex dynamics and interactions.

Acknowledgments
The authors would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart. The data used can be found at this website (https://www.imperial.ac.uk/engineering/ departments/earth-science/research/research-group/perm/research/porescale -modelling/micro-ct-images-andnetworks/).

References

Alhashmi, Z., Blunt, M. J., & Bijeljic, B. (2016). The impact of pore structure heterogeneity, transport, and reaction conditions on fluid-fluid reaction rate studied on images of pore space. *Transport in Porous Media*, *115*(2), 215–237. https://doi.org/10.1007/s11242-016-0758-z

Benson, D. A., & Meerschaert, M. M. (2008). Simulation of chemical reaction via particle tracking: Diffusion-limited versus thermodynamic rate-limited regimes. *Water Resources Research*, *44*, W12201. https://doi.org/10.1029/2008WR007111

Berkowitz, B. (2002). Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources*, *25*(8–12), 861–884. https://doi.org/10.1016/S0309-1708(02)00042-8

Berkowitz, B., Cortis, A., Dentz, M., & Scher, H. (2006). Modeling non-Fickian transport in geological formations as a continuous time random walk. *Reviews of Geophysics*, *44*, W11416. https://doi.org/10.1029/2005RG000178

Berkowitz, B., Klafter, J., Metzler, R., & Scher, H. (2002). Physical pictures of transport in heterogeneous media: Advection-dispersion, random-walk, and fractional derivative formulations. *Water Resources Research*, *38*(10), 1191. https://doi.org/10.1029/2001WR001030

Berkowitz, B., & Scher, H. (1998). Theory of anomalous chemical transport in random fracture networks. *Physical Review E*, *57*(5), 5858–5869. https://doi.org/10.1103/PhysRevE.57.5858

Bijeljic, B., & Blunt, M. J. (2006). Pore-scale modeling and continuous time random walk analysis of dispersion in porous media. *Water Resources Research*, *42*, W01202. https://doi.org/10.1029/2005WR004578

Bijeljic, B., Mostaghimi, P., & Blunt, M. J. (2011). Signature of non-Fickian solute transport in complex heterogeneous porous media. *Physical review letters*, *107*(20), 204–502. https://doi.org/10.1103/PhysRevLett.107.204502

Bijeljic, B., Mostaghimi, P., & Blunt, M. J. (2013). Insights into non-Fickian solute transport in carbonates. *Water Resources Research*, *49*, 2714–2728. https://doi.org/10.1002/wrcr.20238

Bijeljic, B., Muggeridge, A. H., & Blunt, M. J. (2004). Pore-scale modeling of longitudinal dispersion. *Water Resources Research*, *40*, W11501. https://doi.org/10.1029/2004WR003567

Bolster, D., & Dentz, M. (2012). Anomalous dispersion in chemically heterogeneous media induced by long-range disorder correlation. *Journal of Fluid Mechanics*, *695*, 366–389. https://doi.org/10.1017/jfm.2012.25

Bolster, D., Méheust, Y., Le Borgne, T., Bouquain, J., & Davy, P. (2014). Modeling preasymptotic transport in flows with significant inertial and trapping effects—The importance of velocity correlations and a spatial Markov model. *Advances in Water Resources*, *70*, 89–103. https://doi.org/10.1016/j.advwatres.2014.04.014

Bolster, D., Paster, A., & Benson, D. A. (2016). A particle number conserving Lagrangian method for mixing-driven reactive transport. *Water Resources Research*, *52*, 1518–1527. https://doi.org/10.1002/2015WR018310

Cushman, J. H., & O'Malley, D. (2015). Fickian dispersion is anomalous. *Journal of Hydrology*, *531*, 161–167. https://doi.org/10.1016/ j.jhydrol.2015.06.036

De Anna, P., Le Borgne, T., Dentz, M., Tartakovsky, A. M., Bolster, D., & Davy, P. (2013). Flow intermittency, dispersion, and correlated continuous time random walks in porous media. *Physical review letters*, *110*(18), 184–502. https://doi.org/10.1103/PhysRevLett. 110. 184502

de Barros, F. P. J., Fernàndez-Garcia, D., Bolster, D., & Sanchez-Vila, X. (2013). A risk-based probabilistic framework to estimate the endpoint of remediation: Concentration rebound by rate-limited mass transfer. *Water Resources Research*, *49*, 1929–1942. https://doi.org/10.1002/wrcr.20171

Dentz, M., & Berkowitz, B. (2003). Transport behavior of a passive solute in continuous time random walks and multirate mass transfer. *Water Resources Research*, *39*(5), 1111. https://doi.org/10.1029/2001WR001163

Dentz, M., Cortis, A., Scher, H., & Berkowitz, B. (2004). Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport. *Advances in Water Resources*, *27*(2), 155–173. https://doi.org/10.1016/j.advwatres.2003.11.002

Dentz, M., Kang, P. K., Comolli, A., Le Borgne, T., & Lester, D. R. (2016). Continuous time random walks for the evolution of Lagrangian velocities. *Physical Review Fluids*, *1*(7), 74004.

Doob, J. L. (1942). The Brownian movement and stochastic equations. *Annals of Mathematics*, *43*(2), 351–369. https://doi.org/ 10.2307/1968873

Edery, Y., Guadagnini, A., Scher, H., & Berkowitz, B. (2014). Origins of anomalous transport in heterogeneous media: Structural and dynamic controls. *Water Resources Research*, *50*, 1490–1505. https://doi.org/10.1002/2013WR015111

Edery, Y., Scher, H., & Berkowitz, B. (2010). Particle tracking model of bimolecular reactive transport in porous media. *Water Resources Research*, *46*, W07524. https://doi.org/10.1029/2009WR009017

Einstein, A. (1905). Investigations on the theory of the Brownian movement. Ann. der Physik.

Engdahl, N. B., Benson, D. A., & Bolster, D. (2017). Lagrangian simulation of mixing and reactions in complex geochemical systems. *Water Resources Research*, *53*, 3513–3522. https://doi.org/10.1002/2017WR020362

Harvey, C. F., Swartz, C. H., Badruzzaman, A. B. M., Keon-Blute, N., Yu, W., Ali, M. A., et al. (2002). Arsenic mobility and groundwater extraction in Bangladesh. *Science*, *298*(5598), 1602–1606. https://doi.org/10.1126/science.1076978

Hu, L. Y., Chugunova, T., MultBolster, D., Paster, A., & Benson, D. A. (2016). A particle number conserving Lagrangian method for mixing-driven reactive transport. *Water Resources Research*, *52*, 1518–1527. https://doi.org/10.1002/2015WR018310

Kang, P. K., Anna, P., Nunes, J. P., Bijeljic, B., Blunt, M. J., & Juanes, R. (2014). Pore-scale intermittent velocity structure underpinning anomalous transport through 3-D porous media. *Geophysical Research Letters*, *41*, 6184–6190. https://doi.org/10.1002/2014GL061475

Kang, P. K., Brown, S., & Juanes, R. (2016). Emergence of anomalous transport in stressed rough fractures. *Earth and Planetary Science Letters*, *454*, 46–54. https://doi.org/10.1016/J.EPSL.2016.08.033

Kang, P. K., Dentz, M., Le Borgne, T., & Juanes, R. (2015b). Anomalous transport on regular fracture networks: Impact of conductivity heterogeneity and mixing at fracture intersections. *Physical Review E*, *92*(2), 22148.

Kang, P. K., Le Borgne, T., Dentz, M., Bour, O., & Juanes, R. (2015a). Impact of velocity correlation and distribution on transport in fractured media: Field evidence and theoretical model. *Water Resources Research*, *51*, 940–959. https://doi.org/10.1002/2014WR015799

Kitanidis, P. K. (1994). The concept of the dilution index. *Water Resources Research*, *30*, 2011–2026. https://doi.org/10.1029/94WR00762

Le Borgne, T., Bolster, D., Dentz, M., de Anna, P., & Tartakovsky, A. (2011). Effective pore-scale dispersion upscaling with a correlated continuous time random walk approach. *Water Resources Research*, *47*, W12538. https://doi.org/10.1029/2011WR010457

Le Borgne, T., Dentz, M., Bolster, D., Carrera, J., De Dreuzy, J.-R., & Davy, P. (2010). Non-Fickian mixing: Temporal evolution of the scalar dissipation rate in heterogeneous porous media. *Advances in Water Resources*, *33*(12), 1468–1475.

Le Borgne, T., Dentz, M., & Carrera, J. (2008a). Lagrangian statistical model for transport in highly heterogeneous velocity fields. *Physical review letters*, *101*(9), 90601. https://doi.org/10.1103/PhysRevLett.101.090601

Le Borgne, T., Dentz, M., & Carrera, J. (2008b). Spatial Markov processes for modeling Lagrangian particle dynamics in heterogeneous porous media. *Physical Review E*, *78*(2), 26308. https://doi.org/10.1103/PhysRevE.78.026308

Lester, D. R., Dentz, M., & Le Borgne, T. (2016). Chaotic mixing in three-dimensional porous media. *Journal of Fluid Mechanics*, *803*, 144–174. https://doi.org/10.1017/jfm.2016.486

Lester, D. R., Metcalfe, G., & Trefry, M. G. (2013). Is chaotic advection inherent to porous media flow? *Physical Review Letters*, *111*(17), 174101. https://doi.org/10.1103/PhysRevLett.111.174101

Meyer, D. W., & Bijeljic, B. (2016). Pore-scale dispersion: Bridging the gap between microscopic pore structure and the emerging macroscopic transport behavior. *Physical Review E*, *94*(1), 13107. https://doi.org/10.1103/PhysRevE.94.013107

Meyer, D. W., Tchelepi, H. A., & Jenny, P. (2013). A fast simulation method for uncertainty quantification of subsurface flow and transport. *Water Resources Research*, *49*, 2359–2379. https://doi.org/10.1002/wrcr.20240

Most, S., Bijeljic, B., & Nowak, W. (2016). Evolution and persistence of cross-directional statistical dependence during finite-Péclet transport through a real porous medium. *Water Resources Research*, *52*, 8920–8937. https://doi.org/10.1002/2016WR018969

Mostaghimi, P., Bijeljic, B., & Blunt, M. (2012). Simulation of flow and dispersion on pore-space images. *SPE Journal*, *17*(04), 1–131. https://doi.org/10.2118/135261-PA

Nœtinger, B., & Estebenet, T. (2000). Up-scaling of double porosity fractured media using continuous-time random walks methods. *Transport in Porous Media*, *39*(3), 315–337. https://doi.org/10.1023/A:1006639025910

Paster, A., Bolster, D., & Benson, D. A. (2014). Connecting the dots: Semi-analytical and random walk numerical solutions of the diffusion-reaction equation with stochastic initial conditions. *Journal of Computational Physics*, *263*, 91–112. https://doi.org/10.1016/ J.JCP.2014.01.020

Pereira Nunes, J. P., Bijeljic, B., & Blunt, M. J. (2015). Time-of-flight distributions and breakthrough curves in heterogeneous porous media using a pore-scale streamline tracing algorithm. *Transport in Porous Media*, *109*(2), 317–336. https://doi.org/10.1007/s11242-015-0520-y

Raeini, A. Q., Blunt, M. J., & Bijeljic, B. (2012). Modelling two-phase flow in porous media at the pore scale using the volume-of-fluid method. *Journal of Computational Physics*, *231*(17), 5653–5668. https://doi.org/10.1016/j.jcp.2012.04.011

Renner, C., Peinke, J., & Friedrich, R. (2001a). Evidence of Markov properties of high frequency exchange rate data. *Physica A: Statistical Mechanics and its Applications*, *298*(3-4), 499–520. https://doi.org/10.1016/S0378-4371(01)00269-2

Renner, C., Peinke, J., & Friedrich, R. (2001b). Experimental indications for Markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, *433*, 383–409. https://doi.org/10.1017/S0022112001003597

Schumer, R., Benson, D. A., Meerschaert, M. M., & Wheatcraft, S. W. (2001). Eulerian derivation of the fractional advection-dispersion equation. *Journal of Contaminant Hydrology*, *48*(1-2), 69–88. https://doi.org/10.1016/S0169-7722(00)00170-4

Sherman, T., Fakhari, A., Miller, S., Singha, K., & Bolster, D. (2017). Parameterizing the spatial Markov model from breakthrough curve data alone. *Water Resources Research*, *53*, 10,888–10,898. https://doi.org/10.1002/2017WR021810

Sherman, T., Foster, A., Bolster, D., & Singha, K. (2018). Predicting downstream concentration histories from upstream data in column experiments. *Water Resources Research*, *54*, 9684–9694. https://doi.org/10.1029/2018WR023420

Sokolov, I. M. (2012). Models of anomalous diffusion in crowded environments. *Soft Matter*, *8*(35), 9043–9052. https://doi.org/10.1039/C2SM25701G

Strumik, M., & Macek, W. M. (2008). Testing for Markovian character and modeling of intermittency in solar wind turbulence. *Physical Review E*, *78*(2), 26414. https://doi.org/10.1103/PhysRevE.78.026414

Sund, N. L., Bolster, D., & Benson, D. A. (2016). Testing the limits of the spatial Markov model for upscaling transport: The role of nonmonotonic effective velocity autocorrelations. *Physical Review E*, *94*(4), 43107.

Sund, N. L., Bolster, D., & Dawson, C. (2015b). Upscaling transport of a reacting solute through a peridocially converging-diverging channel at pre-asymptotic times. *Journal of Contaminant Hydrology*, *182*, 1–15. https://doi.org/10.1016/J.JCONHYD.2015.08.003

Sund, N., Bolster, D., Mattis, S., & Dawson, C. (2015a). Pre-asymptotic transport upscaling in inertial and unsteady flows through porous media. *Transport in Porous Media*, *109*(2), 411–432. https://doi.org/10.1007/s11242-015-0526-5

Sund, N. L., Porta, G. M., & Bolster, D. (2017b). Upscaling of dilution and mixing using a trajectory based spatial Markov random walk model in a periodic flow domain. *Advances in Water Resources*, *103*, 76–85. https://doi.org/10.1016/j.advwatres.2017.02.018

Sund, N., Porta, G., Bolster, D., & Parashar, R. (2017a). A Lagrangian Transport Eulerian Reaction Spatial (LATERS) Markov model for prediction of effective bimolecular reactive transport. *Water Resources Research*, *53*, 9040–9058. https://doi.org/10.1002/2017WR020821

Wang, L.-P., & Stock, D. E. (1992). Stochastic trajectory models for turbulent diffusion: Monte Carlo process versus Markov chains. *Atmospheric Environment. Part A General Topics*, *26*(9), 1599–1607. https://doi.org/10.1016/0960-1686(92)90060-X

Wen, X.-H., & Gómez-Hernández, J. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. *Journal of Hydrology*, *183*(1–2), ix–xxxii. https://doi.org/10.1016/S0022-1694(96)80030-8