AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020821

Key Points:
- The LATERS Markov model successfully predicts bimolecular reactive transport behavior
- Transport is modeled using the Lagrangian spatial Markov model
- Reactions are modeled via closure of volume-averaging equations

Supporting Information:
- Supporting Information S1

Correspondence to:
N. Sund,
nsund@dri.edu

Citation:
Sund, N., Porta, G., Bolster, D., & Parashar, R. (2017). A Lagrangian transport Eulerian reaction spatial (LATERS) Markov model for prediction of effective bimolecular reactive transport. *Water Resources Research*, *53*. https://doi.org/10.1002/2017WR020821

Received 24 MAR 2017
Accepted 7 OCT 2017
Accepted article online 13 OCT 2017

A Lagrangian Transport Eulerian Reaction Spatial (LATERS) Markov Model for Prediction of Effective Bimolecular Reactive Transport

Nicole Sund[1,2], Giovanni Porta[3], Diogo Bolster[2], and Rishi Parashar[1]

[1]Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, USA, [2]Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, IN, USA, [3]Department of Civil and Environmental Engineering, Polytechnic University of Milan, Milano, Italy

Abstract Prediction of effective transport for mixing-driven reactive systems at larger scales, requires accurate representation of mixing at small scales, which poses a significant upscaling challenge. Depending on the problem at hand, there can be benefits to using a Lagrangian framework, while in others an Eulerian might have advantages. Here we propose and test a novel hybrid model which attempts to leverage benefits of each. Specifically, our framework provides a Lagrangian closure required for a volume-averaging procedure of the advection diffusion reaction equation. This hybrid model is a LAgrangian Transport Eulerian Reaction Spatial Markov model (LATERS Markov model), which extends previous implementations of the Lagrangian Spatial Markov model and maps concentrations to an Eulerian grid to quantify closure terms required to calculate the volume-averaged reaction terms. The advantage of this approach is that the Spatial Markov model is known to provide accurate predictions of transport, particularly at preasymptotic early times, when assumptions required by traditional volume-averaging closures are least likely to hold; likewise, the Eulerian reaction method is efficient, because it does not require calculation of distances between particles. This manuscript introduces the LATERS Markov model and demonstrates by example its ability to accurately predict bimolecular reactive transport in a simple benchmark 2-D porous medium.

1. Introduction

Predicting chemical reactions in porous media can be challenging due to the complex and nonuniform nature of flows at pore scales (Dentz et al., 2011). One of the perhaps most obvious but critical features to recognize in any reactive system is that for reactions to occur, the reactants involved must physically come into contact with one another. Mixing is the physical process that enables this (De Simoni et al., 2007; Gramling et al., 2002; Le Borgne et al., 2010). When the flow is complex, prediction of mixing can be challenging, particularly if one desires to make predictions at scales larger than the pore scale, requiring an upscaled modeling framework. Development of such an upscaled framework, using traditional upscaling approaches can be challenging and lead to models that require assumptions so restrictive as to limit the models' utility (Battiato et al., 2009; Battiato & Tartakovsky, 2011). Thus, novel approaches to tackling such problems are needed.

In order to understand and ultimately accurately predict mixing-driven reactions, it is common to begin by studying the simple bimolecular reaction of the form $A+B \rightarrow C$. While this may seem excessively simple to anyone with detailed knowledge of complex subsurface geochemistry, this very reaction is the fundamental ingredient for many reactive transport models and Gillespie (2007) shows theoretically that understanding it is key to understand arbitrarily complex reaction systems. Therefore, in the context of porous media, this simple reactive setting has been widely studied in the literature, experimentally (e.g., de Anna et al., 2013a; Gramling et al., 2002; Kapoor et al., 1997; Tartakovsky et al., 2008), theoretically (e.g., Gramling et al., 2002; Kapoor et al., 1997) and numerically (e.g., Ding et al., 2012; Edery et al., 2010; Zhang et al., 2009).

A primary challenge in modeling bimolecular transport over scales of practical interest (e.g., Darcy scales) is in accurately accounting for the degree of mixing at the subscale (e.g., pore scales). Reactions in many instances of interest are limited by how quickly mixing can bring reactants together and even in very simple purely diffusive systems incorrect assumptions of complete mixing can dramatically over predict actual

rates of reaction (Bolster et al., 2012; de Anna et al., 2011; Kang & Redner, 1985; Tartakovsky et al., 2012). In porous media, experimental observations (Gramling et al., 2002; Kapoor et al., 1998; Raje & Kapoor, 2000) and numerical models (Ding et al., 2012; Edery et al., 2010; Tartakovsky et al., 2009) continually show that when the rate of reaction is fast compared to the rate at which solutes come together by diffusion, the overall rate of reaction scales slower than would be predicted by the thermodynamic reaction rate under the assumption of perfect mixing. Pore-scale mixing has a large effect on overall reaction rate and thus accurately representing it in an upscaled framework is key (Alhashmi et al., 2015).

One way around these limitations is to develop improved upscaled models that naturally account for subscale fluctuations in an accurate manner. To date various such models have been developed with the general goal of reducing the error caused by overprediction of mixing, and thus overprediction of overall reaction rates (Chiogna & Bellin, 2013; Herrera et al., 2017; Rubio et al., 2008; Sanchez-Vila et al., 2010). A review of many of these approaches, both Lagrangian and Eulerian is available in Dentz et al. (2011). Lagrangian or particle tracking models have been shown to be quite natural at capturing incomplete mixing effects, because while they can recover the correct continuum level equations (Bolster et al., 2016; Paster et al., 2013), the discrete nature of the particles results in inherently noisy concentration distributions, which can be related physically to true subscale concentration fluctuations (Fernàndez-Garcia & Sanchez-Vila, 2011; Paster et al., 2014). Additionally, the absence of numerical dispersion in particle-based methods is appealing, as it artificially enhances mixing effects, overpredicting reaction rates (Benson et al., 2017). While improvements are continuously being made, particle-based methods can still be expensive relative to their Eulerian counterparts and infrastructure for efficient large-scale computing of Eulerian models is to date more widespread (typically for reactions, Lagrangian methods tend to be $O(N^2)$ where N is the number of particles and Eulerian methods tend to be $O(h^n)$ where h is the grid scale and n depends on the solution method). As with all methods, there are benefits and weaknesses to each, summarized nicely in an intermodel comparison paper by Boso et al. (2013).

One particular upscaled Eulerian model of bimolecular reactive transport was developed in Porta et al. (2012b), which results from volume averaging of the advection diffusion reaction equation. This framework very naturally captures incomplete mixing effects, which mathematically are represented by the product of subscale fluctuations of concentrations, by making some simplifying assumptions and defining a closure problem for such terms. While the proposed model works well at capturing certain incomplete mixing effects that more traditional approaches might miss, the closure problem is quite complex, involving nonlocal convolutions in space and time (Hansen et al., 2014; Porta et al., 2016), which, while calculable numerically, can be computationally expensive, challenging to compute accurately, and difficult to physically interpret.

On the Lagrangian side, a promising model for upscaling conservative transport is the Spatial Markov model (SMM) (Le Borgne et al., 2008a, 2008b), which has been very successful at upscaling transport across a diverse and complex set of flows and geometries (Bolster et al., 2014; de Anna et al., 2013b; Kang et al., 2011, 2015; Le Borgne et al., 2011; Sund et al., 2015a). It has also been adapted to model reactive transport in a pore-scale system with a nonuniform spatially dependent first-order reaction rate by Sund et al. (2015b). However, in all of these cases, it has only ever been used to predict mean transport and average concentrations, which does not provide sufficient information for accurate prediction of mixing-driven reactions.

Given the SMM's proven accuracy and flexibility, it seems a natural choice of model to extend to modeling of bimolecular transport. Likewise, given the theoretically rigorous Eulerian approach of volume averaging, it seems appealing to work in this framework. Here we propose that by appropriately modifying the SMM, we can indeed use it to predict mixing processes, which in turn allows us to use it to more efficiently calculate the closure terms in the volume-averaging approach of Porta et al. (2012b). Indeed, the SMM has been shown to predict global mixing measures, such as the scalar dissipation rate and dilution index, very accurately using a trajectory-based downscaling approach (Sund et al., 2017). Here we take a slightly different approach, based on the ideas of Porta et al. (2015), who mapped the complex velocity distribution in a porous medium on to an equivalent shear velocity distribution, using this as the basis for a successful upscaling procedure. Our current proposed model is referred to as the Lagrangian Transport Eulerian Reaction Spatial Markov model, or the LATERS Markov model. Our goal is a hybrid approach that capitalizes on the respective strengths of both Lagrangian and Eulerian approaches. This approach has the advantage of

naturally embedding nonlocal terms. To demonstrate the feasibility of this model, we apply it to a benchmark problem, described in the following section.

2. Reactive Transport Setting

The porous medium and setup that we focus on is identical to that of Porta et al. (2013), which reports high-resolution numerical simulations of flow and reactive transport at the pore scale for the purposes of testing an upscaled model based on volume averaging. The domain is made up of a periodic array of cylinders and has porosity $\phi \approx .34$. A subset of the domain and a unit cell is depicted in Figure 1. The specific reactive transport problem that we focus on is a displacement setup, where the domain is initially filled uniformly with species A at concentration $\hat{C}_0$; then chemical species B is injected through the inlet (left hand boundary) of the domain at the same concentration. As species A and B come into contact, they can react together to form the product C ($A+B \rightarrow C$ takes place) at some rate k as prescribed by the law of mass action. For completeness, we briefly describe the setup here, but note that the model, results and data are exactly those of Porta et al. (2013) and further details are available in that paper.

2.1. Pore-Scale Setting

Flow is considered to be low Reynolds number and governed by the Stokes equations,

$$\hat{\nabla}\hat{p} = \hat{\mu}\hat{\nabla}^2\hat{\mathbf{u}}, \quad \hat{\nabla}\cdot\hat{\mathbf{u}} = 0 \tag{1}$$

where $\hat{\mathbf{u}} = [\hat{u} \quad \hat{v}]^\top$ is the fluid velocity, $\hat{p}$ is the pressure of the fluid, and $\hat{\mu}$ is the viscosity of the fluid. Here hats denote dimensional quantities. The fluid-solid boundary is no slip ($\hat{\mathbf{u}}=0$). The external boundaries of the cell are periodic: $\hat{\mathbf{u}}(\hat{x}=0) = \hat{\mathbf{u}}(\hat{x}=\hat{l}), \hat{\mathbf{u}}(\hat{y}=0) = \hat{\mathbf{u}}\left(\hat{y}=\frac{5}{3}\hat{l}\right)$, where $\hat{l}$ is the distance between cylinder centers in the longitudinal direction and $\frac{5}{3}\hat{l}$ is the distance between cylinder centers in the transverse direction. In order to induce flow, a unit pressure drop is imposed from the left to right boundaries of the unit cell. In the transverse direction, there is no pressure drop and thus no mean flow. The velocity field is then rescaled

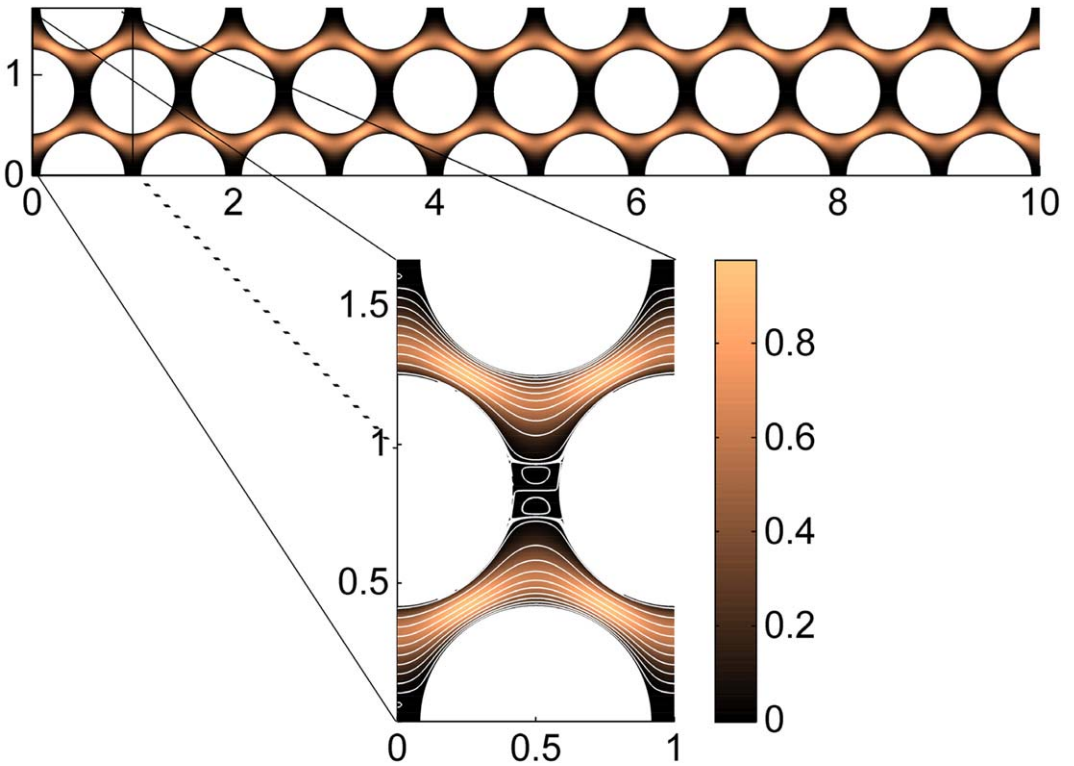

Figure 1. (top) Beginning section of the transport domain shaded by magnitude of longitudinal velocity ($\hat{u}$). (bottom) A zoomed in region of the domain that makes up a unit cell. The color contours show the magnitude of longitudinal velocity and streamlines of the flow are shown in white.

as needed to match the Péclet number, which can be done due to the Stokes flow assumption. Figure 1 shows the transport domain shaded by magnitude of longitudinal velocity ($\hat{u}$) and a zoomed in region of the domain that makes up a unit cell along with streamlines of the flow.

Transport of each species is governed by the advection diffusion reaction equation,

$$\frac{\partial \hat{C}_i}{\partial \hat{t}} + \hat{\nabla} \cdot \left(\hat{\mathbf{u}} \hat{C}_i\right) = \hat{\nabla} \cdot \left(\hat{D} \hat{\nabla} \hat{C}_i\right) - \hat{r}_{AB} \quad i = A, B \tag{2}$$

$$\frac{\partial \hat{C}_C}{\partial \hat{t}} + \hat{\nabla} \cdot \left(\hat{\mathbf{u}} \hat{C}_C\right) = \hat{\nabla} \cdot \left(\hat{D} \hat{\nabla} \hat{C}_C\right) + \hat{r}_{AB}, \tag{3}$$

where $\hat{C}_i$ is the concentration of reactive species i, $\hat{\mathbf{u}}$ is the pore-scale velocity field, and $\hat{D}$ is the molecular diffusion coefficient, which is assumed equal and constant for all species. The reaction rate is determined by the law of mass action such that $\hat{r}_{AB} = \hat{k} \hat{C}_A \hat{C}_B$, where $\hat{k}$ is the kinetic reaction rate constant. This reaction term is what couples the transport equations for each of the considered species.

A dimensionless form of (2) is created by introducing the following quantities,

$$\begin{aligned} C_i &= \frac{\hat{C}_i}{\hat{C}_0}, \quad \mathbf{u} = \frac{\hat{\mathbf{u}}}{\langle \hat{u} \rangle}, \quad t = \hat{t} \frac{\langle \hat{u} \rangle}{\hat{l}}, \\ \nabla &= \hat{\nabla} \hat{l}, \quad x = \frac{\hat{x}}{\hat{l}}, \quad y = \frac{\hat{y}}{\hat{l}}, \\ Pe &= \frac{\langle \hat{u} \rangle \hat{l}}{\hat{D}}, \quad Da = \frac{\hat{k} \hat{C}_0 \hat{l}^2}{\hat{D}}, \end{aligned} \tag{4}$$

where $\hat{C}_0$ is the initial concentration of species A and $\langle \hat{u} \rangle$ is the longitudinal component of velocity averaged over the fluid within the unit cell. The initial concentration $\hat{C}_0 = 1 \frac{\text{mol}}{\text{m}^3}$, Péclet number $Pe = 96$, and Damkohler number $Da = 1{,}038$, which are the values used in Porta et al. (2013). Let us also define a conservative component $C_D = C_B - C_A + 1$; note the 1 is not required, but added to force the concentration C_D to be positive. With these we recast our problem as,

$$\frac{\partial C_D}{\partial t} + \nabla \cdot (\mathbf{u} C_D) = \frac{1}{Pe} \nabla^2 C_D \tag{5}$$

$$\frac{\partial C_B}{\partial t} + \nabla \cdot (\mathbf{u} C_B) = \frac{1}{Pe} \nabla^2 C_B - \frac{Da}{Pe} C_B (C_B - C_D + 1). \tag{6}$$

The initial conditions are,

$$C_D(\mathbf{x}, 0) = 2\ H(-x), \quad C_B(\mathbf{x}, 0) = H(-x), \quad \mathbf{x} \in \Omega_l, \tag{7}$$

where H is the Heaviside step function and Ω_l is the fluid-filled portion of the domain. The boundary is split into various regions, Γ_{ls} is the liquid-solid boundary, $\Gamma_{top}, \Gamma_{bot}, \Gamma_{in}$, and Γ_{out} are the top, bottom, left, and right boundaries, respectively. The boundary conditions are zero concentration gradient (no diffusive flux) in the normal direction with respect to Γ_{ls}, $\Gamma_{top}, \Gamma_{bot}$, and Γ_{out} and constant flux along Γ_{in},

$$\mathbf{n} \cdot \nabla C_i(\mathbf{x}, t) = 0 \quad i = B, D \quad \mathbf{x} \in \Gamma_{ls}, \Gamma_{top}, \Gamma_{bot}, \Gamma_{out} \tag{8}$$

$$\mathbf{u} \cdot \mathbf{n} C_i(\mathbf{x}, t) - \frac{1}{Pe} \mathbf{n} \cdot \nabla C_i(\mathbf{x}, t) = \frac{\dot{m}_i}{A(0)} \quad \mathbf{x} \in \Gamma_{in}. \tag{9}$$

where $\mathbf{n}$ is the inward unit vector normal to the boundary, $\dot{m}_i$ is the rate of mass injection of species i through the inlet boundary which has area $A(0)$, which for this setup is the length of the fluid-filled portion of the domain along inlet $x = 0$. Note that as velocity in the normal direction is equal to zero on boundaries Γ_{ls}, $\Gamma_{top}, \Gamma_{bot}$ these boundaries are associated with zero solute flux. The full domain is made as large as needed so that Γ_{out} is never reached by the invading species. These equations were solved numerically in Porta et al. (2013) using a reactive random walk particle tracking method, the details of which are provided in Porta et al. (2012a).

Henceforth, the results of this simulation are referred to as "benchmarks"; these are the results we treat as the ground truth which we aim to capture with our proposed upscaled LATERS Markov model.

2.2. Volume-Averaged Bimolecular Reactive Transport

So far we have focused on the microscale equations, that is, the equations which resolve the full flow and transport at the pore scale, but if we wish to predict transport over larger scales, it is desirable to work in an upscaled framework. As noted above, we propose to do so in a hybrid form, which partially relies on volume averaging. Therefore, let us decompose quantities into average and fluctuation terms. Quantities are averaged over the volume of the fluid within a unit cell; quantity $*$ averaged in this way is defined as,

$$\langle * \rangle = \frac{1}{V_f} \int_{V_f} * \quad dV, \tag{10}$$

where V_f is the averaging volume. Then the fluctuation of the quantity $*$ is defined as,

$$*' = * - \langle * \rangle. \tag{11}$$

Volume averaging of (5) and (6) leads to a nonlocal transport formulation which we hope to solve with the LATERS Markov model,

$$\frac{\partial \langle C_D \rangle \phi}{\partial t} + \underbrace{\phi \langle \mathbf{u} \rangle \cdot \nabla \langle C_D \rangle}_{\text{advection}} = \underbrace{\frac{1}{Pe} \left[\phi \nabla^2 \langle C_D \rangle + \nabla \phi \nabla \langle C_D \rangle \right.}_{\text{diffusion}} + \underbrace{\left. \nabla \cdot \left(\frac{\phi}{V_f} \int_{\Gamma_{ls}} C_D' \mathbf{n} dA \right) \right]}_{\text{diffusion}} - \underbrace{\nabla \cdot \langle \mathbf{u}' C_D' \rangle}_{\text{dispersion}}, \tag{12}$$

$$\begin{aligned} \frac{\partial \langle C_B \rangle \phi}{\partial t} + \underbrace{\phi \langle \mathbf{u} \rangle \cdot \nabla \langle C_B \rangle}_{\text{advection}} = & \underbrace{\frac{1}{Pe} \left[\phi \nabla^2 \langle C_B \rangle + \nabla \phi \nabla \langle C_B \rangle \right.}_{\text{diffusion}} + \underbrace{\left. \nabla \cdot \left(\frac{\phi}{V_f} \int_{\Gamma_{ls}} C_B' \mathbf{n} dA \right) \right]}_{\text{diffusion}} - \underbrace{\nabla \cdot \langle \mathbf{u}' C_B' \rangle}_{\text{dispersion}} \\ & - \underbrace{\phi \frac{Da}{Pe} \left[\langle C_B \rangle \left(\langle C_B \rangle - \langle C_D \rangle + 1 \right) + \langle C_B' \left(C_B' - C_D' \right) \rangle \right]}_{\text{reaction}} \end{aligned} \tag{13}$$

where ϕ is the porosity of the unit cell. For $Pe \gg 1$ and $Da \gg Pe$ the following relationships can be proposed,

$$C_B' = M C_D' \qquad M = \frac{\langle C_B \rangle}{2\langle C_B \rangle - \langle C_D \rangle + 1} \tag{14}$$

Details of the volume-averaging procedure, as well as details on assumptions and restrictions, are available in Porta et al. (2012b). This yields a continuum-scale integro-differential system of equations, which is nonlocal in time. Solution of these equations via a closure variable, used in the volume-averaging method, can be cumbersome and resource intensive. In the following we assume that the Spatial Markov model naturally accounts for the advective, diffusive, and dispersive terms in these equations, including the nonlocal in time terms that arise in volume averaging, which are particularly relevant to model dispersion in the early times, i.e., for the preasymptotic (non-Fickian) dispersive regime. Our primary goal here is to extend the Spatial Markov model to account for the reactive term. By substituting (14) in the last term appearing in (13), the reactive term can be expressed as

$$r = \phi \frac{Da}{Pe} \left[\langle C_B \rangle \left(\langle C_B \rangle - \langle C_D \rangle + 1 \right) + M(M-1) \langle C_D'^2 \rangle \right]. \tag{15}$$

For our purposes, we are concerned with solving for all of the volume-averaged quantities. As noted earlier, the Spatial Markov model has already been shown to accurately model average concentrations of conservative species, so it should model $\langle C_D \rangle$ well. The primary challenge in our upscaling approach is then to adequately model the fluctuation term $\langle C_D'^2 \rangle$ appearing in (15). This term physically reflects the degree of incomplete mixing at the pore scale. If the system were perfectly mixed this would be zero. If the Spatial Markov model can accurately represent this term, then it can adequately represent mixing and thus it is suited to modeling mixing-driven reactions. Thus, the goal of this work is to augment the Spatial Markov model in such a way as to accurately model this term and thus accurately predict concentrations of reactive species.

3. The LATERS Markov Model

Here we describe the development of the LATERS Markov model, how it is parameterized, and how it is used to predict average concentrations and also to reconstruct subscale concentration distributions. We first recount the parameterization and the procedure followed for approximating longitudinal transport of a conservative species, which is in line with previous implementations of the Spatial Markov Model. Then, we describe how the information obtained from the transport parameterization is employed to model reaction, which is the key innovative element of this contribution. This requires an approximation of subscale fluctuations within each cell, which is performed here through a novel closure procedure. The approximation of subscale fluctuations of concentrations is finally employed to approximate the longitudinal upscaled reactive transport problem.

3.1. Parameterization of the LATERS Markov Model Via Direct Simulation of Transport

As with many Spatial Markov modeling approaches, the model is parameterized by investing computational effort into running high-resolution particle tracking simulations over a pair of unit cells and then using information from these high-resolution runs to parameterize the model and predict transport over much larger scales. This saves significant computational expense, as simulating transport over only two unit cells is significantly cheaper than simulating transport over the full domain. Furthermore, even though we are parameterizing the model for use in reactive transport modelling, we only simulate conservative transport at high resolution.

To obtain the flow field, equation (1) is solved with a finite element method with triangular elements and linear basis functions using the Galerkin method. Boundary conditions are those described in section 2.1 and the velocity field is rescaled to obtain $Pe = 96$. The velocity field is interpolated at all spatial locations through the P^1 finite element basis functions.

For the particle tracking, we use a random walk method, where particles of equal mass follow Langevin equation,

$$x_{n+1} = x_n + u\Delta t + \sqrt{\frac{2\Delta t}{Pe}}\xi_n \tag{16}$$

$$y_{n+1} = y_n + v\Delta t + \sqrt{\frac{2\Delta t}{Pe}}\eta_n \quad n = 0, 1, 2, \ldots, \tag{17}$$

where u and v are the dimensionless x and y components of the velocity field $\mathbf{u}$, $\{\xi\}$ and $\{\eta\}$ are independent identically distributed samples from the standard Gaussian distribution (with mean 0 and variance 1), and Δt is a fixed time step. The initial condition is flux weighted along the inlet $x = 0$. Fluid-solid boundaries are no flux, which are treated as elastic reflections for particles. Top and bottom external boundaries are periodic.

We simulate transport by this method across two unit cells in order to create distributions to be used in the LATERS Markov model. We simulate the transport of 10^6 particles and for each particle we store three pieces of information:

1. $y_0 =$ the particle's initial y position at the inlet
2. $\tau_1 =$ the amount of time it takes the particle to traverse the first unit cell
3. $\tau_2 =$ the amount of time it takes the particle to traverse the second unit cell

From this information, we obtain a joint distribution $f(y_0, \tau_1, \tau_2)$, which is used to create the LATERS Markov model. Note that compared to previous implementations of the SMM we store one piece of information that is typically not gathered, which is y_0. This is required for the purposes of predicting subscale fluctuations, i.e., provides all the information that is required to obtain a continuum scale solution of (13), as will be explained in the following sections.

3.2. Longitudinal Transport of A Conservative Solute

Longitudinal transport of the conservative component, defined as D in the previous sections, is modeled through a Lagrangian Spatial Markov model, consistent with previous contributions (Le Borgne et al., 2008a, 2011). We recall here the main features of the involved modeling procedure.

We aim to model transport along direction x, longitudinal to the average flow velocity for some time window $t \in [0, T_{end}]$. We also introduce a discrete set of time levels t_j with $j=1,\ldots,N_t$, with $t_1=0$ and $t_{N_t}=T_{end}$, which we employ to discretize the solution in time. We set these time levels such that the time step $dt=t_{j+1}-t_j$ is constant along the simulation.

The mean concentration of the conservative component D is solved using the standard Spatial Markov Model, in which the one-dimensional distribution of solute mass along the x axis is discretized into particles (p_i) of equal mass, each starting at $(x^{(0)}=0, t^{(0)})$. First we assign the mass to each particle used for computation of the conservative component of concentration. This is done so that the mass flux is the same as that in the pore-scale simulation,

$$m_D(p_i)=\frac{\dot{m}dt}{n_p}, \tag{18}$$

where $m_D(p_i)$ is the mass assigned to particles used to approximate the transport of the conservative component D, $\dot{m}$ is the rate of mass injection of the conservative component through the boundary in the pore-scale simulation, dt is the time step in the LATERS Markov model simulation, and n_p is the number of particles injected per time step. The simulation time step dt is chosen based on the smallest amount of time it took any particle to traverse the first unit cell in the simulation we used to parameterize the model such that,

$$dt=\frac{\min(\tau_1)}{10}. \tag{19}$$

Now, we can march our simulation through time t. In order to simulate the constant flux boundary condition in equation (9), we inject a fixed number of particles (n_p) at each time step t_j throughout the simulation, starting at time $t=0$. Each particle enters the domain at some random time throughout the time step of length dt, so that each particle's starting time is given by,

$$t^{(0)}=t_j+Udt, \tag{20}$$

where U is a uniform random number between 0 and 1. Each particle's subsequent motion is governed by Langevin equation,

$$x^{(n+1)}=x^{(n)}+1 \tag{21}$$

$$t^{(n+1)}=t^{(n)}+\tau^{(n+1)} \quad n=0,1,2,\ldots, \tag{22}$$

where $\tau^{(n+1)}$ is a sampled from the distribution $f(\tau)$, which we create from the joint distribution of τ_1 and τ_2 obtained from the pore-scale simulation described in section 3.1,

$$f(\tau)=\begin{cases} f(\tau_1) & n=0 \\ f(\tau^{(n+1)}|\tau^{(n)}) & n=1,2,\ldots \end{cases} \tag{23}$$

We model $f(\tau^{(n+1)}|\tau^{(n)})$ using a spatially homogeneous Markov chain. The Markov chain is created by discretizing the travel time distribution $f(\tau_1)$ into 20 equiprobable states and recording the cutoff times that split up the states. A travel time τ is in state i if $lim_i \leq \tau < lim_{i+1}$ where lim_i is the cutoff time for state i and $lim_1 \equiv 0$. Using the joint distribution of τ_2 and τ_1 we then define a transition matrix (T_{ji}), which represents a discrete form of $f(\tau_2|\tau_1)$,

$$T_{ji}=P(\tau_2 \in \text{state } j|\tau_1 \in \text{state } i) \approx f(\tau_2|\tau_1). \tag{24}$$

By modeling $f(\tau^{(n+1)}|\tau^{(n)})$ as a spatially homogeneous Markov chain, we are assuming that $f(\tau^{(n+1)}|\tau^{(n)})=f(\tau_2|\tau_1)$, and use the transition matrix T_{ji} for all following steps. Figure 2 shows the entire Markov chain, which consists of $f(\tau_1)$ and T_{ji}. Because the Péclet number here is relatively low, the transition matrix looks homogeneous in parts, implying that particles may transition between states readily. There are, however, a few states that particles have a high probability of remaining in, most notably the higher numbered states, which correspond to large travel times. This finding is ubiquitous has been explored in detail in Bolster et al. (2014). Further details on the creation of the Markov chain used in the Spatial Markov model are available in Le Borgne et al. (2008a, 2011) and Bolster et al. (2014).

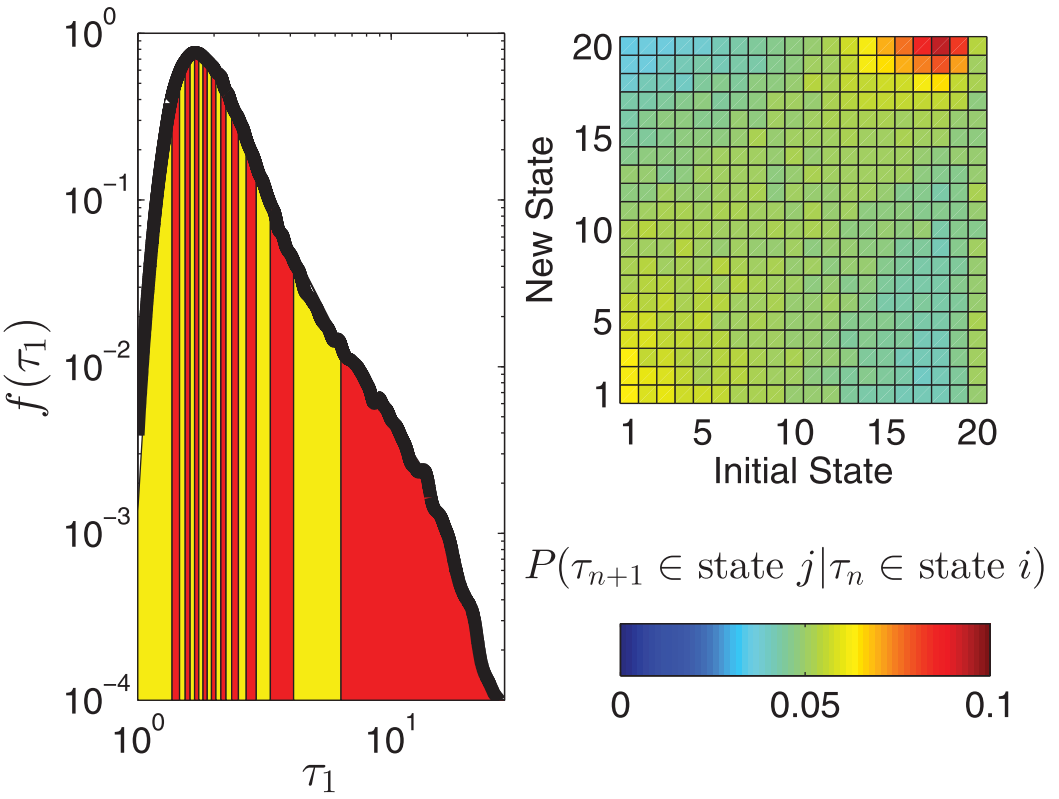

Figure 2. (left) Travel time distribution split up into equiprobable states ($f(\tau_1)$). (right) Transition matrix (T_{ji}), which shows the probability of a particle transitioning between travel time states.

3.3. Closure Approximation

To have (15) as a closed equation requires an approximation of not only the average concentration of the conservative component $\langle C_D \rangle$ but also the average intensity of the fluctuations $\langle C_D'^2 \rangle$ in each cell. This is in practice equivalent to predict mixing of the concentration of species D within the pore-space. Here we employ a closure procedure to approximate this quantity, i.e., our goal is to recover from the parameters of the LATERS effective (upscaled) model an estimate of $\langle C_D'^2 \rangle$.

To calculate this term, we resort to a mapping of the actual physical unit cell Ω to a reference cell $\widetilde{\Omega}$, following a conceptual idea proposed in Porta et al. (2015), who map the heterogeneous velocity field in a complex porous medium on to an equivalent monotonic shear flow. To this end, we formulate the following assumptions:

1. Reference unit cell $\widetilde{\Omega}$ is defined such that it has the same total dimension as the original cell, i.e., as the rectangle $[0, 1] \times [0, Y]$, but it is defined as a an equivalent shear flow configuration, i.e., as a planar flow between two parallel plates. Therefore, the fluid is distributed evenly throughout the cell, rather than confined between the large cylinders in the actual unit cell
2. Particles travel along straight streamlines in $\widetilde{\Omega}$, each with different longitudinal velocity, depending on their transverse location. This allows us to represent pore-scale fluctuations of velocity and concentration by assigning to each individual particle a different velocity. This is obtained via the characterization of travel time distribution and particle location obtained from the approach described in section 3.1, which accounts for both advective and diffusive contributions to space-time particle transitions within cell Ω.

The rationale behind these assumptions is to extract salient features from pore-scale information available from the parameterization described in section 3.1, consistent with the discussion in Porta et al. (2015). Our goal is to select meaningful pore-scale information necessary to obtain a characterization of pore-scale mixing. This corresponds in practice to quantitatively estimating $\langle C_D'^2 \rangle$. In contrast with the technique proposed in Sund et al. (2017), we do not perform a direct downscaling of the particle position to the actual pore-scale geometry. The use of the reference cell $\widetilde{\Omega}$ allows us to avoid issues which may arise from the mapping

of particle positions into the real unit cell due to the presence of liquid-solid boundaries, as discussed in Sund et al. (2017). Our numerical closure procedure has the two following steps:

1. For each particle p_i assign a position along the longitudinal and transversal direction within $\widetilde{\Omega}$, labeled $\widetilde{x}(p_i), \widetilde{y}(p_i)$
2. Estimate $\langle C_D'^2 \rangle$ at time t_j from the spatial distribution of particles within $\widetilde{\Omega}$. The two steps are separately introduced in the following subsections.

3.3.1. Approximation of Particle Location in $\widetilde{\Omega}$

From the Spatial Markov model, we know at any time level t_j which unit cell each particle p_i is in and how long it will take the particle to travel through that unit cell, i.e., the particle is somewhere within cell $n+1$ if $t^{(n)}(p_i) \leq t_j < t^{(n+1)}(p_i)$. Consistent with the simplified assumptions introduced above, we consider the particle's velocity within the unit cell as constant. That is, the particle travels at a constant effective velocity of,

$$U_{eff} = \frac{1}{\tau^{(n+1)}}, \tag{25}$$

over the $n+1$st unit cell. The amount of time that the particle has spent in that unit cell at time t_j is $t_j - t^{(n)}$. Thus,

$$\widetilde{x}[p_i(t_j)] = n + \frac{t_j - t^{(n)}}{\tau^{(n+1)}} \quad t^{(n)} \leq t_j < t^{(n+1)}. \tag{26}$$

Note that even though we consider a constant transition velocity through the cell for each individual particle p_i, variability of pore-scale velocities is maintained in our model by assigning U_{eff} consistent with the characterization of pore-scale transition times, obtained as described in section 3.1. Therefore, the distribution of velocities associated with an ensemble of particles p_i within a given cell and at a given time will honor the distribution of longitudinal effective velocities obtained from pore-scale information. Also note that the effective velocity U_{eff} includes information on both advective and diffusive longitudinal displacements.

We focus now on the estimation of the transversal position $\widetilde{y}[p_i(t_j)]$. We start from the quantity y_0, which is introduced section 3.1, corresponding to the starting position of a particle along the inlet portion of the cell. We can then introduce the conditional distribution of y_0 given the state k, defined as

$$P_0(y_0, k) = P(y_0^{(n)}(p_i) | \tau^{(n)} \in \text{ state } k) \tag{27}$$

To obtain this probability distribution numerically, we employ a discretization of the fluid portion Γ_l of the inlet boundary Γ_{in} into a number N_y of intervals $\mathcal{I}_m$. Let us define the length of the fluid portion of the inlet Γ_l as L_0 and the length of the full inlet boundary Γ_{in} as Y. We define a set of discretization intervals $\mathcal{I}_m = (y_{0,m-1}, y_{0,m})$ with $m = 1 \ldots N_y$, with $y_{0,0} = 0$ and $y_{0,N_y} = L_0$. The discretized probability distribution of the inlet positions y_0 for each state k is then defined as

$$P_0(\mathcal{I}_m, k) = P(y_0^{(n)}(p_i) \in \mathcal{I}_m | \tau^{(n)} \in \text{ state } k) \tag{28}$$

This distribution is shown in the left plot of Figure 3, upon setting $N_y = 10$. Due to the symmetry of the velocity field, the distribution of y_0 positions for each state is approximately symmetric. It can be seen that particles in the lower numbered states (those with the smallest travel times) have a higher probability of starting near at locations $y_0 \approx 0.25$ and $y_0 \approx 1.25$ where the velocity is higher. By contrast, particles in the higher numbered states (those with the largest travel times) have a higher probability of starting near the top or bottom of the unit cell, where the velocity is lower. There is a large section in the middle of the unit cell ($\frac{5}{12} < y_0 < 1.2$) where no particles can begin, due the presence of a solid cylinder. We note however that the distribution $P_0(\mathcal{I}_m, k)$ embeds the boundary condition employed at the stage of model parameterization, i.e., flux-weighted boundary condition in the case considered here. For this reason, we define the following rescaling,

$$P_{0,vel}(\mathcal{I}_m, k) = P_0(\mathcal{I}_m, k) \frac{\int_{\mathcal{I}_m} \frac{1}{u(x=0,y)} dy}{\int_{\Gamma_{in}} \frac{1}{u(x=0,y)} dy}, \tag{29}$$

which is shown in the middle plot of Figure 3. Notice that if the flux over an interval $\mathcal{I}_m$ is less than or equal to zero, we run into numerical problems of negative or infinite probabilities. To remedy this issue, we

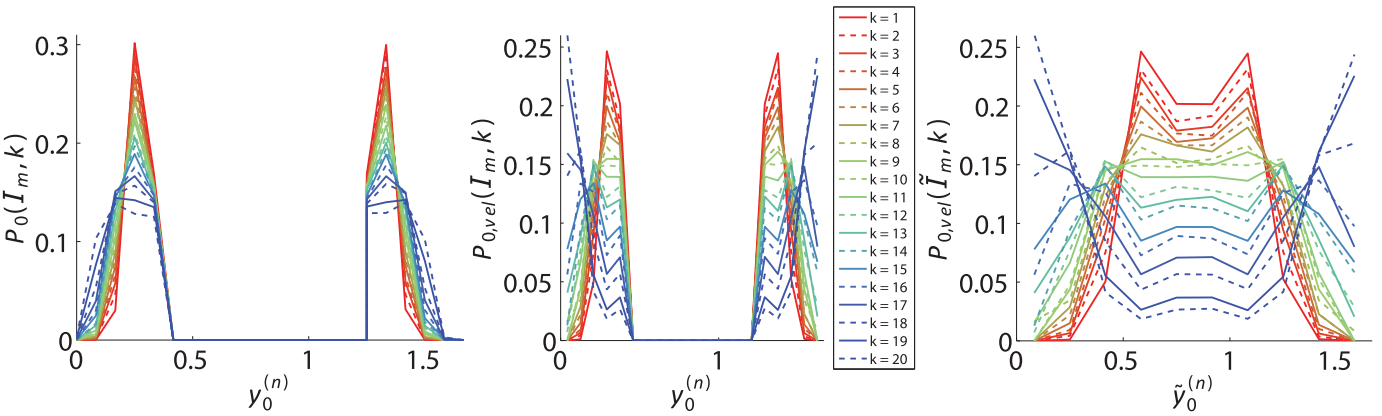

Figure 3. Probability of starting y position, $y_0^{(n)}$, given each travel time state. On the left is the original distribution ($P_0(\mathcal{I}_m,k)$). In the middle is the distribution rescaled by the inverse of the flux ($P_{0,vel}(\mathcal{I}_m,k)$). On the right is the distribution rescaled for the reference cell ($P_{0,vel}(\tilde{\mathcal{I}}_m,k)$).

ensured that the number of intervals was small enough that this would not be the case. Another option would have been to create intervals of varying lengths.

This correction is consistent with the discussion provided in Dentz et al. (2016), regarding the conceptual differences between Eulerian and Lagrangian approaches in terms of the effective velocity of transported solute mass. As noted in section 3.1, in order to match the Eulerian constant flux boundary condition, the distribution of y_0 assigned to Lagrangian particles is flux weighted in the parameterization stage. To understand why, consider two locations on the inlet, one of which has twice the flux of the other. Because twice as much fluid is flowing through this location, it requires injection of twice as many particles as the other in order to maintain a constant flux of concentration. To convert back to the Eulerian perspective, we need to correct for this flux weighting when we map the particle positions, because we are not concerned with the probability that a particle is injected at a location along the inlet, but rather with the probability that a particle was injected at a location along the inlet and has yet to flow out of the unit cell. In other words, if we neglect the correction in (29) the longitudinal position imposed in our Eulerian mapping would be forced to follow a flux-weighted transversal distribution at all spatial locations and times.

Upon relying on the input information introduced above, we pursue the following procedure to determine $\tilde{y}[p_i(t_j)]$:

1. at time t_j and given the state k associated with particle p_i sample the distribution $P_{0,vel}(\mathcal{I}_m,k)$ to determine $y_0[p_i(t_j)]$
2. map the location $y_0[p_i(t_j)]$ to the unit cell $\widetilde{\Omega}$ through

$$\tilde{y}_0[p_i(t_j)]=m\frac{Y}{N_y}+\frac{Y(y_0-y_{0,m})}{L_0} \quad y_{0,m-1}<y_0\le y_{0,m}. \tag{30}$$

3. consistent with the assumed shear flow configuration assign $\tilde{y}[p_i(t_j)]=\tilde{y}_0[p_i(t_j)]$.

The right plot of Figure 3 shows the distributions $P_{0,vel}(\mathcal{I}_m,k)$ mapped onto $\tilde{y}_0$. The cyan loop of Figure 4 summarizes these Lagrangian steps of the algorithm.

3.3.2. Approximation of $\langle C_D'^2\rangle$

Once the positions $\tilde{x}[p_i(t_j)],\tilde{y}[p_i(t_j)]$ are computed for all particles, we need to create a spatial grid within the reference cell $\widetilde{\Omega}$, in order to compute a spatial distribution of the particles' concentration and thereby the quantity $\langle C_D'^2\rangle$, which is the final goal of our closure procedure. The benchmark simulation, to which we will compare our results, discretizes the fluid portion of the unit cell into 371 squares. In order to keep roughly the same discretization, we discretize the reference cell $\widetilde{\Omega}$ into n_{vox} = 10x37 rectangular voxels. To ensure our mapped cell $\widetilde{\Omega}$ contains the same amount of fluid as the unit cell Ω, we normalize the concentration in each voxel by the porosity ϕ of the cell, i.e., we set the volume of each voxel, $V_{vox}=\phi V_F/n_{vox}$, where

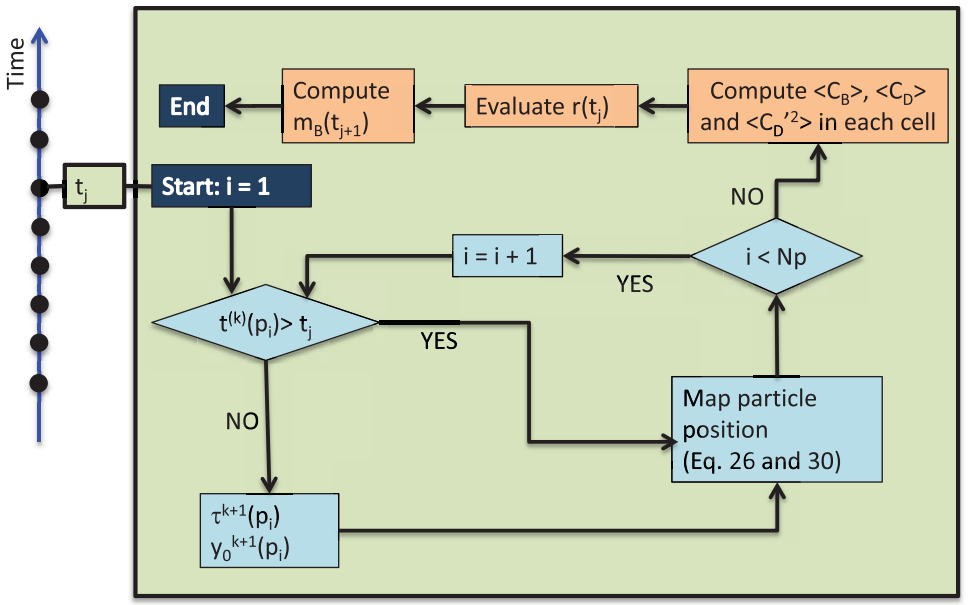

Figure 4. Definition sketch of the proposed reactive transport algorithm applied for each LATERS time level t_j. (Lagrangian cyan loop) For each particle p_i, we assign the longitudinal position $\tilde{x}$ according to the travel time, while the transversal position $\tilde{y}$ is randomly chosen from the distribution $P_{0,vel}(\tilde{\mathcal{I}}_m, k)$. (Eulerian orange steps) Once all particles are mapped to the reference cell, we compute concentration moments in each cell, evaluate the reactive term, and update the mass of each *B* particle.

V_F is the total amount of fluid contained in the unit cell Ω. This, in effect is giving each voxel an internal porosity equal to the porosity of the full original unit cell. Now the predicted concentration of species *D* within each voxel of the domain can be found by adding up the number of particles in each cell of mass $m_D(p_i)$, and dividing by the volume of fluid in the voxel, that is,

$$\widetilde{C}_D^w = \frac{1}{V_{vox}} \sum_{k=1}^{n_p^w} m_D(p_i), \tag{31}$$

where $\widetilde{C}_D^w$ is the concentration of species *D* in voxel *w*, n_p^w is the number of particles in voxel *w*, and $m_D(p_i)$ is the solute mass carried by particle *i*.

We can now obtain the quantities $\langle C_D \rangle$, and determine the quantity $\langle C_D'^2 \rangle$ through the discrete version of (10),

$$\langle C_D \rangle = \frac{1}{n_{vox}} \sum_{w=1}^{n_{vox}} \widetilde{C}_D^w, \tag{32}$$

$$\langle C_D'^2 \rangle = \frac{1}{n_{vox}} \sum_{w=1}^{n_{vox}} (\widetilde{C}_D^w - \langle C_D \rangle)^2, \tag{33}$$

where $\widetilde{C}_s^w$ is the concentration of species *s*, in voxel *w* of the unit cell $\widetilde{\Omega}$.

3.4. Reactive Transport Modeling Procedure

In this section, we describe how the modeling tools introduced so far are combined together to come up with an algorithm providing an upscaled solution strategy for the reactive transport setting introduced in section 2.1.

To model longitudinal transport of a reactive species, our goal is now to predict the reactive term of the volume-averaged reactive transport equations (equation (13)). To this end, we introduce a reactive species, *B*, and compute concentrations of both the reactive and conservative components. Introduction of reactive

species B only requires the assignment of appropriate mass to each individual particle in the random walk simulation described above. Each particle p_i in the LATERS model carries mass m_B of species B and a mass m_D of the D component. While the latter is constant, the former changes as the simulation progresses due to reactions. The mass of all particles used to compute the concentration of B as they are introduced in the system is,

$$m_B(p_i)(t=t^{(0)})=\frac{m_D(p_i)}{2}. \tag{34}$$

Recall that we added 1 to the concentration of D so that it would be positive. This means that as B particles are injected into the domain (i.e., prior to reaction) the mass of the conservative component associated with each particle p_i is twice that of the reactive component, B.

We now detail our numerical methodology to approximate the longitudinal profiles of volume-averaged concentrations of all chemical species at a given time level t_j. To do so let us assume that all quantities related to the time t_{j-1} are known. We perform the following steps (which are also graphically reported in the orange steps in Figure 4):

1. Approximate the position of the particles at time t_j (cyan-filled loop in Figure 4). This is obtained as detailed in section 3.3.
2. Compute $\langle C_D\rangle$, $\langle C_B\rangle$ in each cell along x and determine the quantity $\langle C_D'^2\rangle$ using (33).
3. Compute $r(t_j)$ by replacing the volume-averaged quantities approximated using by (32)-(33) into definition (15).
4. Reduce the mass of B particles appropriately. This is done by first computing the change in average concentration that must occur,

$$\phi\Delta\langle C_B\rangle=-r(t_j)dt, \tag{35}$$

where $\Delta\langle C_B\rangle=\langle C_B(t_{j+1})\rangle-\langle C_B(t_j)\rangle$ is the change in concentration of B over the time step dt. We then calculate the new mass m_B of the particles in the cell as,

$$m_B(p_i)(t_{j+1})=\frac{\langle C_B(t_{j+1})\rangle V_F}{np}, \tag{36}$$

where np is the number of particles in the cell and i is the index of a particle within the cell. It is not necessary for particles within the cell to have equal mass to each other, but this eliminates possible numerical issues (particles with negative mass).

Steps 2–4 in the above list are represented by orange-filled items in Figure 4.

Average concentrations of reactive species A and the reaction product C can be computed for each time t_j from $\langle C_B\rangle$ and $\langle C_D\rangle$ as,

$$\langle C_A\rangle=\langle C_B\rangle-\langle C_D\rangle+1 \tag{37}$$

$$\langle C_C\rangle=\frac{\langle C_D\rangle}{2}-\langle C_B\rangle. \tag{38}$$

4. Results

4.1. Predicting Concentrations of Reactive Species

Here we compare the predictions of the LATERS Markov model to the results of the high-resolution pore-scale simulation of bimolecular reactive transport, described in section 2. Dimensionless time $t_D=\frac{t}{Pe}$ will be used to denote intercomparison times so that $t_D=1$ corresponds to the Taylor or diffusion time scale $\hat{t}_D=\hat{l}^2\hat{D}$.

Practically speaking, the most relevant metrics for evaluation of our model are how well we predict the average concentration of the reactive components, $\langle C_A\rangle$, $\langle C_B\rangle$, and the reaction product, $\langle C_C\rangle$ relative to the benchmark measured values. Thus, we focus on these first. Figure 5 shows these comparisons. The

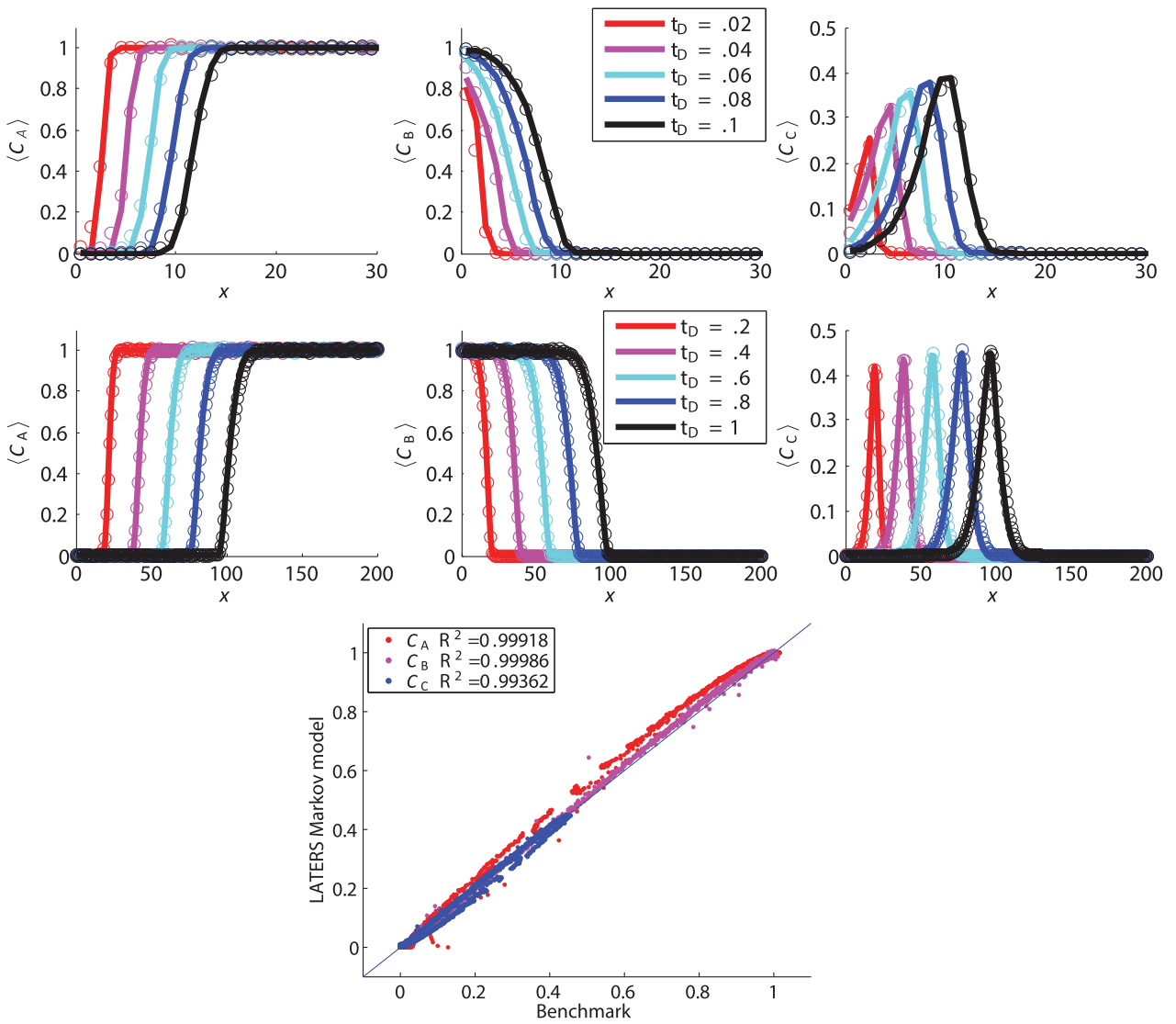

Figure 5. (top and middle) Average concentration of reactive species *A* and *B* and product species *C* at various times. Circles indicate benchmark values and solid lines are the predictions of the LATERS Markov model. (bottom) Scatter plot of average concentration of reactive species *A* and *B* and product species *C* for all cells in the domain and all times.

predictions of average concentrations of the reactive species and the product are qualitatively and quantitatively good. This can be seen more clearly from the scatter plot in the bottom plot, which shows the predicted value of $\langle C_i \rangle$ versus the benchmark value at all times (from.02 to 1 at intervals of.02) and at every cell in the domain, where the coefficient of determination, $R^2 > .99$ for all species. Focusing on some specific aspects that are key to capture we take a closer look at peak concentrations (e.g., Chiogna & Bellin, 2013; Gramling et al., 2002). Figure 6 directly compares the peak values of the concentration of species *C*, which demonstrates that indeed the peaks are well matched by the LATERS Markov model. We also include the results of a LATERS Markov model which assumes complete mixing. This is done by running the LATERS Markov model and setting $\langle C'^2 \rangle = 0$ in equation (15) and can be seen as the green squares in Figure 6 which persistently overpredict the peaks.

4.2. Successes and Failures of Intermediate Steps

The previous subsection highlights the success of our proposed LATERS model in achieving our end objective to predict reactive transport with our proposed upscaled model. We now discuss in detail all the intermediate steps which lead to the final results, discussing successes and limitations of the proposed

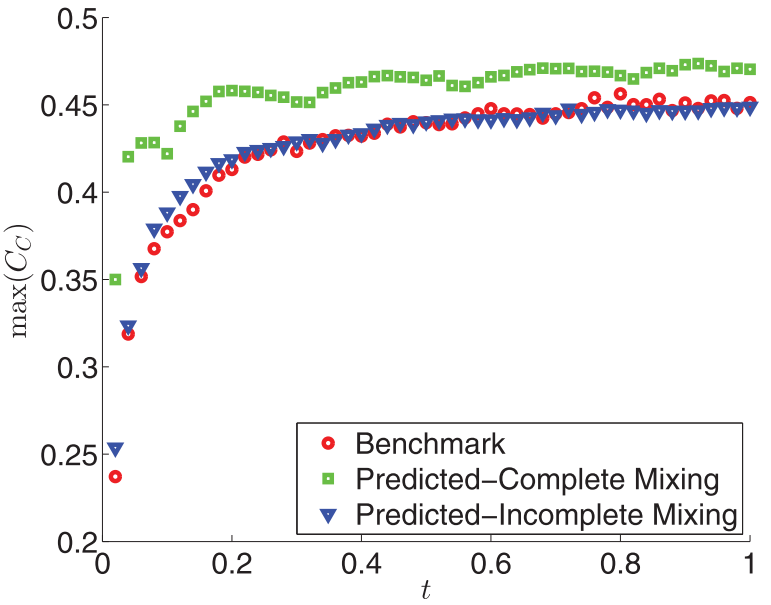

Figure 6. Peak concentration of product species *C* at various times.

approach. The first metric we compare is the average of the conservative component of concentration, $\langle C_D \rangle$. We expect the LATERS Markov model to accurately predict $\langle C_D \rangle$, because for average longitudinal transport the LATERS Markov model is equivalent to the standard Spatial Markov model, which as noted in the Introduction has had great success in reproducing conservative transport in many settings. Figure 7 shows the comparison of benchmark and predicted values of $\langle C_D \rangle$ at various times. Not surprisingly, the predicted values of $\langle C_D \rangle$ closely match the benchmark values. This is evident from the scatter plot in the right plot, where again $R^2 > .99$.

The next metric for comparison is the variance of the conservative component, $\langle C_D'^2 \rangle$. Recall that $\langle C_D'^2 \rangle$ is a key contributor to the amount of reaction that takes place, so it is important that the LATERS Markov model can predict this value accurately. The key to accurately predicting $\langle C_D'^2 \rangle$, is in reasonably mapping particle positions. We made many simplifying assumptions when going through this step, so here we check how much of an effect these assumptions had on the model's predictive capability.

Figure 8 shows the spatial distribution of benchmark and predicted values of $\langle C_D'^2 \rangle$ at various times with the reactive region (where both C_A and C_B are nonnegligible, having values of at least 10^{-4}) highlighted. For

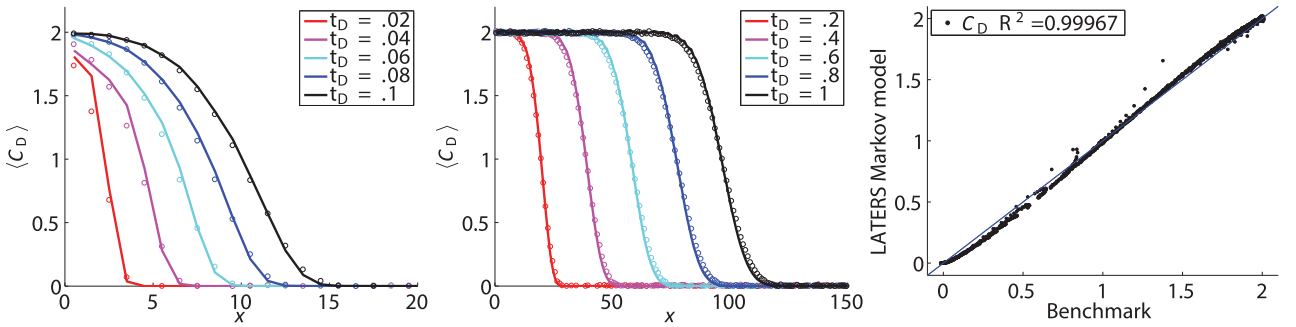

Figure 7. (left and middle) Average concentration of conservative component *D* at various times. Circles indicate benchmark values and solid lines are the predictions of the LATERS Markov model. (right) Scatter plot of average concentration of conservative component *D* for all cells in the domain and all times.

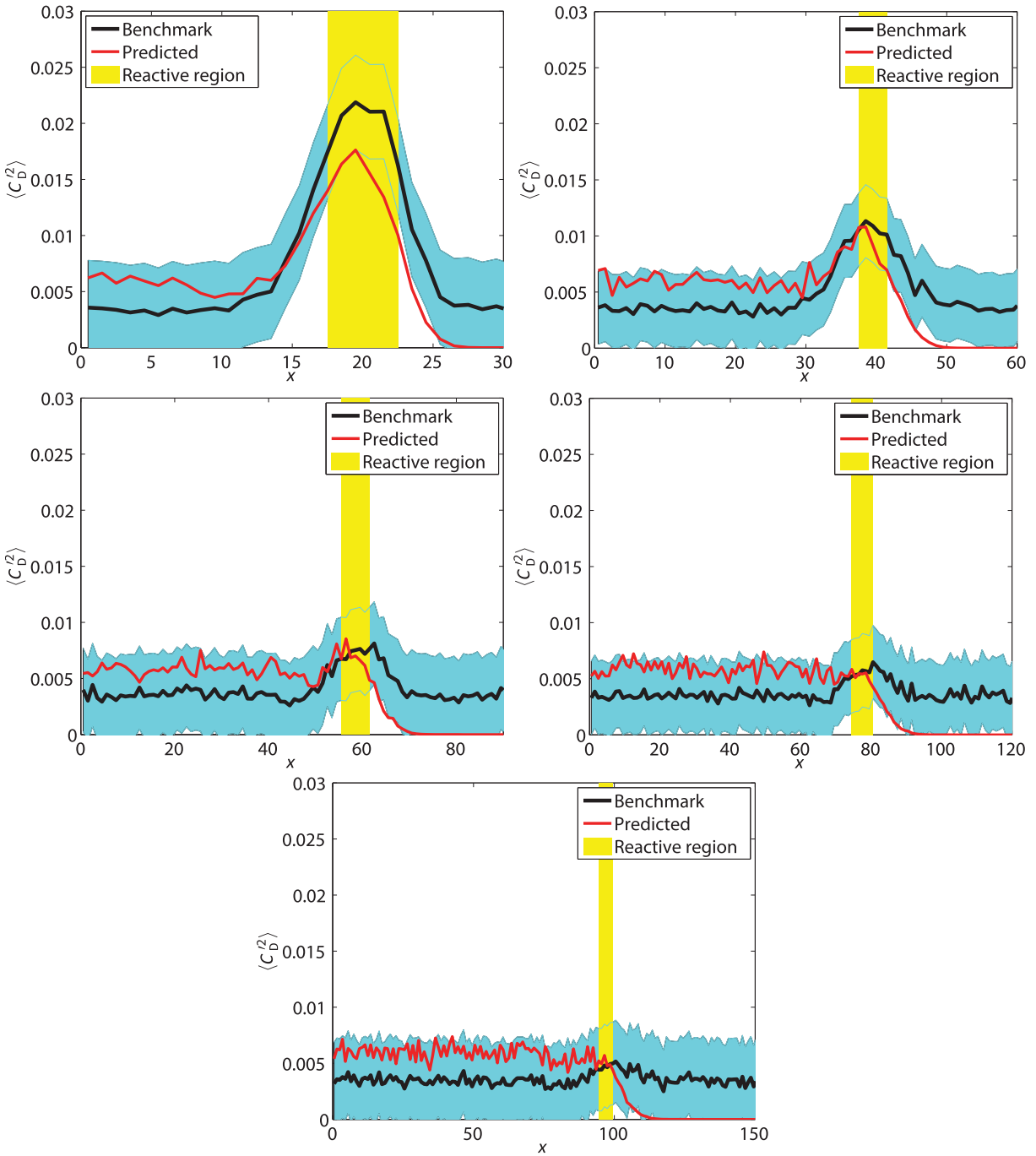

Figure 8. Variance of concentration of conservative component D at various times. The figures correspond to $t = .2, .4, .6, .8$, and 1.

the most part, the predicted values of $\langle C_D'^2 \rangle$ match closely to the benchmark values, while the predicted values consistently overpredict $\langle C_D'^2 \rangle$ behind the concentration front location.

The cyan-shaded region on Figure 8 shows the range of acceptable error. A certain amount of numerical background noise occurs in the calculation of $\langle C_D'^2 \rangle$ in both the benchmark and the LATERS Markov models due to the fact that we are using a limited number of discrete particles on a discrete grid. To quantify this amount in the benchmark simulation, we consider a region of the domain far past the front of species B. Recall that the domain is initially filled with a uniform concentration of species A, thus, at large

distances where species B has yet to reach, the concentration should be uniform ($\langle C_D'^2 \rangle = 0$). Since in this region of the domain the only cause for $\langle C_D'^2 \rangle > 0$ is numerical noise, we calculate $\langle C_D'^2 \rangle$ at the last cell of the domain, which species B does not reach. We consider error in the prediction of the LATERS Markov model to be acceptable if it is within this amount of the calculated benchmark value of $\langle C_D'^2 \rangle$.

The predictions deviate outside the acceptable error range by a small amount for $t = .2$, which leads to a slight overprediction of the peak of the product species shown in Figure 6 and after the reactive region, where they consistently underpredict $\langle C_D'^2 \rangle$. The latter is due to the fact that the LATERS Markov model does not simulate A particles, and thus predicts no variance where no B species is present. Note that this has no impact on the reactive simulation, i.e., if there are no B particles present, the reactive term is zero (because $\langle C_B \rangle = 0$), so our prediction of $\langle C_D'^2 \rangle$ at these locations is irrelevant. The region of importance is highlighted explicitly by the yellow strip in Figure 8. Outside of this region any mismatches have no consequence on upscaling of the reaction.

Next, in order to determine how quantitatively good the model predictions are, we consider two error measures, the mean absolute error (MAE(s)) and the relative error in total mass (ϵ_s), defined as follows,

$$\mathrm{MAE}(s) = \frac{1}{n_{cells}} \sum_{j=1}^{n_{cells}} |\langle C_{s_{pred}} \rangle^j - \langle C_{s_{act}} \rangle^j| \tag{39}$$

$$\epsilon_s = \frac{\sum_{j=1}^{n_{cells}} \langle C_{s_{pred}} \rangle^j - \langle C_{s_{act}} \rangle^j}{\sum_{j=1}^{n_{cells}} \langle C_{s_{act}} \rangle^j}, \tag{40}$$

where $n_{cells} = 600$ is the total number of unit cells in the domain such that the full domain extends far beyond the concentration front, $\langle C_{s_{act}} \rangle^j$ is the average concentration of species i in the benchmark simulation, and $\langle C_{s_{pred}} \rangle^j$ is the average concentration of species i predicted by the LATERS Markov model in the jth unit cell. Figure 9 shows these error measures. Note that ϵ_D should be identically zero, because D is a conservative component, however, due to the fact that we cannot create a constant injection with discrete particles, there is a small amount of error. Notice that the error in the mass of the product species changes sign throughout time: at the very earliest time we are overpredicting the amount of product produced and then we proceed to underpredict and finally at $t = .8$ we go back to overpredicting. At first, this may seem strange, but it can be explained by careful consideration of Figures 6 and 8. Up until $t \approx .2$, the peak concentration of product species C is overpredicted, because the variance of the conservative species $\langle C_D'^2 \rangle$ is underpredicted. This overreaction causes the overlap between the reactive species A and B to be small behind the concentration front, thus leading to underprediction of reactions behind the front. Later, around $t = .6$, the peak concentration of

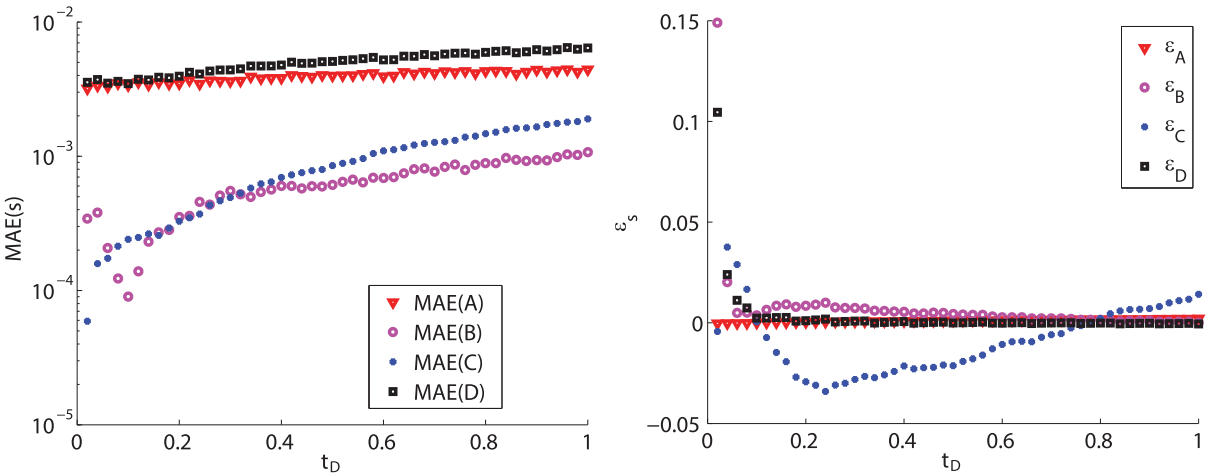

Figure 9. (left) Mean absolute error of predictions (MAE(s)) of average concentrations of all components throughout time. (right) Error in predictions of total mass (ϵ_s) of all components throughout time.

product species C is underpredicted, because the variance of the conservative species $\langle C_D'^2 \rangle$ is overpredicted in the reactive region (even though it is still within the acceptable range). This lower reaction causes the overlap between reactive species A and B to be large behind the concentration front, thus leading to overprediction of reactions behind the front. While the relative errors in total mass of the reactive and product species are greater than those of the conservative component, they only go above 5% at the very earliest time measured, when very little mass is present in the system. The vast bulk are below 1% for all but the product species.

Error convergence analyses to look at discretization effects were also conducted, exploring the influence of both particle numbers as well as number of travel time bins used. For the interested reader, these results are available in the supporting information (Bolster et al., 2014; Le Borgne et al., 2011; Sund et al., 2015b, 2016).

5. Conclusions

The LATERS Markov model proposed here was able to accurately predict mean concentrations of not only conservative species, but also reactive species involved in a benchmark bimolecular mixing-driven reactive system. The key idea of this approach is to combine Lagrangian transport models with the Eulerian derivation of effective reaction rate obtained through volume averaging. Our proposed model provides a numerical methodology to approximate the closure formulation associated with formal upscaling of bimolecular reactive transport, under advection-dominated transport conditions and when local reaction rates are associated with small time scales if compared to transport (i.e., large Pe and Da). The model in its current form is restricted to a periodic flow domain. It is also subject to the same assumptions and constraints as the Spatial Markov model and Volume Averaging method (i.e., spatially stationarity steady velocity and multiple length and time-scale constraints).

The LATERS Markov model is built using pore-scale simulation of conservative solute transport across two unit cells, like many previous implementations of the spatial Markov model. Starting from this information, we devise a methodology to reconstruct local fluctuations of a conservative component, which are the key to modeling mixing-driven bimolecular transport. To this end, we resort to a closure approach, which relies on a simplified unit cell. This latter is characterized on the basis of pore-scale information. Our approach combines advantages of the Lagrangian and Eulerian modeling perspective: the Lagrangian transport method naturally accounts for nonlocal transport terms, while the Eulerian upscaled reaction model allows for the application of formally derived upscaled reaction rate expressions.

We test the LATERS Markov model by comparing the results with those obtained from the pore-scale simulation of mixing-driven bimolecular reaction within a simple two-dimensional porous medium. The LATERS Markov model, within this context, appears able to predict the space-time evolution of the variance of local concentration of a conservative solute, indicating that the proposed procedure may be well suited to predict mixing. Accurate predictions are also obtained in predicting the space-time evolution of the concentrations of the two reactants and of the product. Most noteworthy, a close match is obtained in the early time mixing regime (preasymptotic regime), which is the regime where it is most challenging to make accurate predictions by means of standard Eulerian upscaling formulations.

While this initial study shows the potential promise of the LATERS Markov model, the considered flow domain is relatively simple. As such future research should focus on application to more complex benchmarks, that include larger-scale heterogeneity, both at pore (e.g., Porta et al., 2016) and Darcy scales (e.g., Herrera et al., 2017), in particular extending the methodology to nonperiodic systems, where for example the Spatial Markov model has been shown to work well for conservative transport (e.g., Kang et al., 2017; Le Borgne et al., 2008a; Massoudieh et al., 2017). Additionally a full and detailed comparison between the LATERS Markov model and other upscaled reactive transport models that exist is warranted. For example, while we have weak evidence suggesting the LATERS Markov model can run faster than previous purely Eulerian volume averaging approaches, none of the considered codes were optimized or written with efficiency and benchmarking in mind and as such rigorous comparison studies may be interesting.

Acknowledgments
We would like to express thanks for financial support via Army Office of Research grant W911NF1310082 and National Science Foundation grants EAR-1351623 and EAR-1417264.
G. Porta acknowledges funding from MIUR (Italian ministry of Education, University and Research, Water JPI, WaterWorks 2014, Project: WE-NEED-Water NEEDs, availability, quality and sustainability). The numerical data collected by the methods described in this manuscript and used in the figures, along with the MATLAB codes used to generate them are available upon request to N. Sund.

References

Alhashmi, Z., Blunt, M., & Bijeljic, B. (2015). Predictions of dynamic changes in reaction rates as a consequence of incomplete mixing using pore scale reactive transport modeling on images of porous media. *Journal of Contaminant Hydrology*, *179*, 171–181.
Battiato, I., & Tartakovsky, D. (2011). Applicability regimes for macroscopic models of reactive transport in porous media. *Journal of Contaminant Hydrology*, *120*, 18–26.
Battiato, I., Tartakovsky, D. M., Tartakovsky, A. M., & Scheibe, T. (2009). On breakdown of macroscopic models of mixing-controlled heterogeneous reactions in porous media. *Advances in Water Resources*, *32*(11), 1664–1673.
Benson, D. A., Aquino, T., Bolster, D., Engdahl, N., Henri, C. V., & Fernàndez-Garcia, D. (2017). A comparison of Eulerian and Lagrangian transport and non-linear reaction algorithms. *Advances in Water Resources*, *99*, 15–37.
Bolster, D., de Anna, P., Benson, D. A., & Tartakovsky, A. M. (2012). Incomplete mixing and reactions with fractional dispersion. *Advances in Water Resources*, *37*, 86–93.
Bolster, D., Méheust, Y., Borgne, T. L., 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.
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
Boso, F., Bellin, A., & Dumbser, M. (2013). Numerical simulations of solute transport in highly heterogeneous formations: A comparison of alternative numerical schemes. *Advances in Water Resources*, *52*, 178–189.
Chiogna, G., & Bellin, A. (2013). Analytical solution for reactive solute transport considering incomplete mixing within a reference elementary volume. *Water Resources Research*, *49*, 2589–2600. https://doi.org/10.1002/wrcr.20200
de Anna, P., Jimenez-Martinez, J., Tabuteau, H., Turuban, R., Borgne, T. L., Derrien, M., & Meheust, Y. (2013a). Mixing and reaction kinetics in porous media: An experimental pore scale quantification. *Environmental Science & Technology*, *48*(1), 508–516.
de Anna, P., Le Borgne, T., Dentz, M., Bolster, D., & Davy, P. (2011). Anomalous kinetics in diffusion limited reactions linked to non-Gaussian concentration pdf. *Journal of Chemical Physics*, *135*, 174104.
de Anna, P., Le Borgne, T., Dentz, M., Tartakovsky, A. M., Bolster, D., & Davy, P. (2013b). Flow intermittency, dispersion, and correlated continuous time random walks in porous media. *Physical Review Letters*, *110*, 184–502. https://doi.org/10.1103/PhysRevLett.110.184502
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), 074004.
Dentz, M., Le Borgne, T., Englert, A., & Bijeljic, B. (2011). Mixing, spreading and reaction in heterogeneous media: A brief review. *Journal of Contaminant Hydrology*, *120*, 1–17.
De Simoni, M., Sanchez-Vila, X., Carrera, J., & Saaltink, M. (2007). A mixing ratios-based formulation for multicomponent reactive transport. *Water Resources Research*, *43*, W07419. https://doi.org/10.1029/2006WR005256
Ding, D., Benson, D., Paster, A., & Bolster, D. (2012). Modeling bimolecular reactions and transport in porous media via particle tracking. *Advances in Water Resources*, *53*, 56–65. https://doi.org/10.1016/j.advwatres.2012.11.001
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
Fernàndez-Garcia, D., & Sanchez-Vila, X. (2011). Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions. *Journal of Contaminant Hydrology*, *120*, 99–114.
Gillespie, D. T. (2007). Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, *58*, 35–55.
Gramling, C., Harvey, C., & Meigs, L. (2002). Reactive transport in porous media: A comparison of model prediction with laboratory visualization. *Environmental Science & Technology*, *36*(11), 2508–2514. https://doi.org/10.1021/es0157144
Hansen, S. K., Scher, H., & Berkowitz, B. (2014). First-principles derivation of reactive transport modeling parameters for particle tracking and pde approaches. *Advances in Water Resources*, *69*, 146–158.
Herrera, P., Cortínez, J., & Valocchi, A. (2017). Lagrangian scheme to model subgrid-scale mixing and spreading in heterogeneous porous media. *Water Resources Research*, *53*, 3302–3318. https://doi.org/10.1002/2016WR019994
Kang, K., & Redner, S. (1985). Fluctuation-dominated kinetics in diffusion-controlled reactions. *Physical Review A*, *32*(1), 435.
Kang, P. K., Dentz, M., Borgne, T., Lee, L. S., & Juanes, R. (2017). Anomalous transport in disordered fracture networks: Spatial Markov model for dispersion with variable injection modes. *Advances in Water Resources*, *106*, 80–94.
Kang, P. K., Dentz, M., Le Borgne, T., & Juanes, R. (2011). Spatial Markov model of anomalous transport through random lattice networks. *Physical Review Letters*, *107*, 180–602. https://doi.org/10.1103/PhysRevLett.107.180602
Kang, P. K., Le Borgne, T., Dentz, M., Bour, O., & Juanes, R. (2015). 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
Kapoor, V., Gelhar, L. W., & Miralles-Wilhelm, F. (1997). Bimolecular second-order reactions in spatially varying flows: Segregation induced scale-dependent transformation rates. *Water Resources Research*, *33*(4), 527–536.
Kapoor, V., Jafvert, C. T., & Lyn, D. A. (1998). Experimental study of a bimolecular reaction in Poiseuille flow. *Water Resources Research*, *34*(8), 1997–2004.
Le Borgne, T., Bolster, D., Dentz, M., 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*, 090601. 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 Letters*, *78*, 026308. https://doi.org/10.1103/PhysRevE.78.026308
Massoudieh, A., Dentz, M., & Alikhani, J. (2017). A spatial Markov model for the evolution of the joint distribution of groundwater age, arrival time, and velocity in heterogeneous media. *Water Resources Research*, *53*, 5495–5515. https://doi.org/10.1002/2017WR020578
Paster, A., Bolster, D., & Benson, D. (2013). Particle tracking and the diffusion-reaction equation. *Water Resources Research*, *49*, 1–6. https://doi.org/10.1029/2012WR012444
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.
Porta, G., Bijeljic, B., Blunt, M., & Guadagnini, A. (2015). Continuum-scale characterization of solute transport based on pore-scale velocity distributions. *Geophysical Research Letters*, *42*, 7537–7545. https:doi.org/10.1002/2015GL065423

Porta, G., Ceriotti, G., & Thovert, J.-F. (2016). Comparative assessment of continuum-scale models of bimolecular reactive transport in porous media under pre-asymptotic conditions. *Journal of Contaminant Hydrology*, *185*, 1–13.

Porta, G., Thovert, J.-F., Riva, M., Guadagnini, A., & Adler, P. (2012a). Microscale simulation and numerical upscaling of a reactive flow in a plane channel. *Physical Review E*, *86*(3), 036102.

Porta, G. M., Chaynikov, S., Thovert, J.-F., Riva, M., Guadagnini, A., & Adler, P. M. (2013). Numerical investigation of pore and continuum scale formulations of bimolecular reactive transport in porous media. *Advances in Water Resources*, *62*, 243–253.

Porta, G. M., Riva, M., & Guadagnini, A. (2012b). Upscaling solute transport in porous media in the presence of an irreversible bimolecular reaction. *Advances in Water Resources*, *35*, 151–162.

Raje, D., & Kapoor, V. (2000). Experimental study of bimolecular reaction kinetics in porous media. *Environmental Science & Technology*, *34*(7), 1234–1239.

Rubio, A., Zalts, A., & El Hasi, C. (2008). Numerical solution of the advection–reaction–diffusion equation at different scales. *Environmental Modelling & Software*, *23*(1), 90–95.

Sanchez-Vila, X., Fernàndez-Garcia, D., & Guadagnini, A. (2010). Interpretation of column experiments of transport of solutes undergoing an irreversible bimolecular reaction using a continuum approximation. *Water Resources Research*, *46*, W12510. https://doi.org/10.1029/2010WR009539

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.

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*, 043107. https://doi.org/10.1103/PhysRevE.94.043107

Sund, N. L., Bolster, D., & Dawson, C. (2015b). Upscaling transport of a reacting solute through a periodically converging–diverging channel at pre-asymptotic times. *Journal of Contaminant Hydrology*, *182*, 1–15.

Sund, N. L., Porta, G. M., & Bolster, D. (2017). 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.

Tartakovsky, A. M., de Anna, P., Le Borgne, T., Balter, A., & Bolster, D. (2012). Effect of spatial concentration fluctuations on non-linear reactions in diffusion-reaction systems. *Water Resources Research*, *48*, W02526. https://doi.org/10.1029/2011WR010720

Tartakovsky, A. M., Redden, G., Lichtner, P. C., Scheibe, T. D., & Meakin, P. (2008). Mixing-induced precipitation: Experimental study and multiscale numerical analysis. *Water Resources Research*, *44*, W06S04. https://doi.org/10.1029/2006WR005725

Tartakovsky, A. M., Tartakovsky, G. D., & Scheibe, T. D. (2009). Effects of incomplete mixing on multicomponent reactive transport. *Advances in Water Resources*, *32*(11), 1674–1679.

Zhang, Y., Benson, D. A., & Reeves, D. M. (2009). Time and space nonlocalities underlying fractional-derivative models: Distinction and literature review of field applications. *Advances in Water Resources*, *32*(4), 561–581.