Random-walk Modeling of Reactive Transport in Porous Media With a Reduced-order Chemical Basis of Conservative Components

Guillem Sole-Mari[1,2,3], Michael J. Schmidt[4], Diogo Bolster[4], Daniel Fernàndez-Garcia[2,3]

[1]Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

[2]Department of Civil and Environmental Engineering (DECA), Universitat Politècnica de Catalunya, Barcelona, Spain

[3]Hydrogeology Group (UPC-CSIC), Barcelona, Spain

[4]Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, South Bend IN, USA

Key Points:

- A Lagrangian (particle tracking) method that employs conservative chemical components is presented and used to model reactive transport.
- This method keeps particle number constant in time, avoiding both changing resolution and increasing computational expense.
- This method simplifies the interaction between equilibrium and kinetic reactions through probabilities of reaction with effective stoichiometry.

Corresponding author: Guillem Sole-Mari, guillem.sole.mari@outlook.com

Abstract

In this work we employ a reduced-order basis of conservative chemical components to model reactive transport using a Lagrangian (particle tracking) method. While this practice is well-understood in the Eulerian (grid-based) context, its adaptation to a Lagrangian context requires a novel reformulation of particle transport properties. Because the number of conservative-species particles need not change during simulation, spatial resolution stays constant in time, and there is no increase in computational expense due to increasing numbers of product particles. Additionally, this treatment simplifies the interaction between equilibrium and kinetic reactions and allows the use of species-dependent transport operators at the same time. We apply this method to model a suite of simple test problems that include equilibrium and kinetic reactions, and results exhibit excellent match with base-case Eulerian results. Finally, we apply the new method to model a 2D problem concerning the mobilisation of cadmium by a CO_2 leak, showing the potential applicability of the proposed methodology.

1 Introduction

Chemical reactions are ubiquitous in hydrologic systems and play a controlling role in the small and large scale behavior of many systems of practical interest. However, modeling hydrogeological reactive transport is a computationally-intensive exercise that typically requires simulating the transport of numerous species and calculating the complicated, coupled geochemical reactions occurring between and among all of the species. These can occur over vastly different temporal and spatial scales in highly heterogeneous settings, making for a challenging problem in subsurface hydrology [*Dentz et al.*, 2011; *Sanchez-Vila and Fernàndez-Garcia*, 2016; *Benson et al.*, 2017; *Valocchi et al.*, 2019]. Any resulting mathematical models can involve a very large number of tightly coupled nonlinear equations, where some reactions can provide immensely restrictive conditions in terms of the resources required to calculate them (e.g. time step, grid resolution).

For this reason, it is common practice to form a reduced-order basis of chemically conservative components in order to reduce the degrees of freedom of the chemistry calculations and also reduce the number of calculations that are required for transport [*Saaltink et al.*, 1998; *De Simoni et al.*, 2005, 2007; *Molins et al.*, 2004; *Kräutle and Knabner*, 2005; *Bolster et al.*, 2010; *Gramling et al.*, 2002]. Such an approach requires the assumption that the

transport operator is the same for all species, restricting its general applicability to consider species with for example equal dispersion and/or retardation coefficients. This practice is already well-established for use in Eulerian, or grid-based methods (e.g., finite-difference, -volume, -element) and has been widely applied in that context [e.g. *Saaltink et al.*, 1998; *Molins et al.*, 2004]. However, this strategy has to date not been applied with Lagrangian, or particle tracking (PT), methods because of difficulties in formulating transport properties for numerical particles that carry multiple species of reactant within one conservative component. In recent years PT methods have grown in popularity within the hydrogeological reactive transport community, so we address this shortcoming in the present work. In doing this, we also show how to deal with species-dependent transport operators through chemically conservative components.

PT methods are attractive for numerous reasons, including the fact that they solve advection exactly and do not introduce spurious numerical diffusion, which would lead to artificially high degrees of mixing, compromising calculation of mixing driven reactions [*Benson et al.*, 2017]. In the context of reactive transport there are currently two main "flavors" of such PT methods. The first type are the random-walk particle tracking (RWPT) methods that simulate diffusion by Brownian motion-based random walks [*Benson and Meerschaert*, 2008; *Paster et al.*, 2014; *Rahbaralam et al.*, 2015; *Bolster et al.*, 2016; *Schmidt et al.*, 2017; *Sole-Mari et al.*, 2017; *Sole-Mari and Fernàndez-Garcia*, 2018]. These methods are popular for their ease of implementation, natural parallelism of transport calculations [*Rizzo et al.*, 2019], and their attractive speed to accuracy trade-offs, as compared to corresponding Eulerian methods [*Benson et al.*, 2017].

The other broad class is mass-transfer particle tracking (MTPT) methods that simulate diffusion by inter-particle mass transfers in addition to random-walks [e.g., *Benson and Bolster*, 2016; *Engdahl et al.*, 2017; *Schmidt et al.*, 2019a,b, 2020]. The key difference between these two methods lays in the concept of particle: while in RWPT particles represent only the solute, in MTPT they carry a fixed amount of the solvent (fluid) and variable solute concentrations. Advantages of MTPT methods are that they allow the separate simulation of solute mixing and non-mixed spreading of a plume [*Benson et al.*, 2019], they correctly simulate the entropy of a solute plume [*Benson et al.*, 2020], allow for significant parallel speedup [*Engdahl et al.*, 2019], and they have recently been linked to smoothed-particle hydrodynamics (SPH) methods [*Sole-Mari et al.*, 2019a; *Herrera et al.*, 2009, 2017; *Monaghan*, 2012], which have a rich and rigorous mathematical foundation.

The approach described in this manuscript belongs to the former category of RWPT methods. As such, each particle carries a mass, and concentrations are computed through spatial interpolation. But in this case the mass tracked by Lagrangian particles, and consequently the interpolated concentrations, correspond to conservative components which are invariant under the relevant equilibrium reactive system. By building the Lagrangian methodology around this simple idea we find that it is possible to simulate equilibrium reactive transport, featuring species with different transport properties (for instance, mobile/immobile species, species-dependent dispersion coefficients), by adapting the motion of each particle to the local chemistry. That is, the interaction of a particle with its neighbors may change its velocity and the extent of its Brownian motion, which may be seen as a generalized, chemically dynamic version of the classical concept of retardation coefficient. This presents distinct advantages because one may avoid continuous injection and destruction of particles due to equilibrium reactions, meaning that both resolution and computational load remain stable through the course of a simulation. Moreover, one could use such a model to track Lagrangian statistics of conservative components in a system: e.g., the trajectories of "total calcium" particles, or the residence time of "alkalinity" particles. Furthermore, for chemical systems which combine both equilibrium and kinetic reactions, the interaction is straightforwardly implemented through effective probabilities of kinetic reaction imposed on the component particles.

The manuscript is structured as follows. First, we describe how to solve chemical equilibrium in a Lagrangian framework, using a reduced-order conservative-component chemical basis (Section 2). We then outline, in Section 3, how this reformulation, in terms of components, interacts with the various other active physical and numerical processes, which includes the primary advance of our work: calculating the transport properties of a given particle, based on the chemical species it carries, such that it captures the proper transport dynamics (Section 3.1). At the end of the section we step through how the algorithm is implemented (Section 3.3). Next, we apply our new conservative-component particle tracking to some simple problems to validate the approach (Section 4), and we also use the method to model a more realistic problem involving the mobilisation of heavy metals due to a carbon dioxide (CO_2) leak (Section 5).

2 Chemical Equilibrium of a Lagrangian Particle Cloud

In this section, we describe a method to solve chemical equilibrium given information on particle positions that represent the mass of conservative components. We first outline the method for mapping a chemically-reactive transport system containing equilibrium reactions to and from a reduced-order system composed of only conservative (non-reactive) components [*Saaltink et al.*, 1998]. Next, we describe our Lagrangian (particle tracking) transport model that employs density estimation via kernel methods [*Sole-Mari et al.*, 2017; *Sole-Mari and Fernàndez-Garcia*, 2018; *Sole-Mari et al.*, 2019b]. Finally, we discuss how, given the solution to the conservative transport problem, we may recover the solution to the reactive system by solving the remaining equilibrium reactions at the given times of observation.

2.1 Chemically Conservative Components

Let us consider a system with N_C chemical species $\left\{C_1, \ldots, C_{N_C}\right\}$, with associated concentrations $\mathbf{c} = \left[c_1, \ldots, c_{N_C}\right]^{\mathrm{T}}$, subjected to N_S equilibrium reactions. Mass conservation in these reactions is described through the stoichiometry, or reaction, matrix $\mathbf{S}$ (of dimension $N_C \times N_S$, with each column representing a reaction's stoichiometry). This matrix relates the rate of mass variation of the species due to equilibrium reactions to the vector of equilibrium reaction rates $\mathbf{s} = \left[s_1, \ldots, s_{N_S}\right]^{\mathrm{T}}$, such that $\mathbf{Ss}$ is the vector of mass variation rates per unit of fluid volume due to equilibrium reactions. Then, the evolution of the concentrations in time and space can be formulated in matrix form as

$$\frac{\partial \phi \boldsymbol{\rho}}{\partial t} = \mathcal{L}\left(\mathbf{c}\right) + \phi \mathbf{S}\mathbf{s} + \phi \boldsymbol{\psi}, \tag{1}$$

where $\mathcal{L}\left(\mathbf{c}\right) \equiv \left[\mathcal{L}_1\left(c_1\right), \ldots, \mathcal{L}_{N_C}\left(c_{N_C}\right)\right]^{\mathrm{T}}$ is a vector of linear transport operators (typically advection-dispersion operators, although others are possible), $\boldsymbol{\rho}$ is a vector that contains the total mass of the species per unit fluid volume, $\mathbf{c}$ is a vector of the concentrations of species in the aqueous phase, and $\boldsymbol{\psi}$ is a source/sink term. The aqueous concentration of the species is related to the total mass density as $c_j = \omega_j \rho_j$, where ω_j is the mass fraction of the species C_j in the aqueous phase. This parameter can have a value of 1 if the species is mobile and 0 if it is immobile, but it can also have a value between 0 and 1 to account for sorption or similar processes not included in the equilibrium equations. Importantly, we note that, in contrast with previous formulations for decoupling reactive transport problems, the transport operators $\mathcal{L}_j\left(c_j\right)$ are species dependent.

We wish to describe the transport in equilibrium by $N_{\mathcal{U}} = N_C - N_{\mathcal{S}}$ chemically conservative components in lieu of the total N_C species. To do this, we must define a $N_{\mathcal{U}} \times N_C$ full-ranked matrix $\mathbf{U}$, referred to as the component matrix, such that

$$\mathbf{u} = \mathbf{U}\boldsymbol{\rho}, \tag{2}$$

where the component vector $\mathbf{u} = \left[u_1, \ldots, u_{N_{\mathcal{U}}}\right]^{\mathrm{T}}$ contains the concentrations of the $N_{\mathcal{U}}$ components $\left\{\mathcal{U}_1, \ldots, \mathcal{U}_{N_{\mathcal{U}}}\right\}$. Matrix $\mathbf{U}$ relates all the components (rows) and the species (columns). For the components to eliminate the reaction term

$$\frac{\partial \phi \mathbf{u}}{\partial t} = \mathbf{U}\frac{\partial \phi \boldsymbol{\rho}}{\partial t} = \mathbf{U}\mathcal{L}\left(\mathbf{c}\right) + \phi \mathbf{U}\boldsymbol{\psi}. \tag{3}$$

It follows directly from (3) and (1) that the component matrix $\mathbf{U}$ must satisfy

$$\mathbf{US} = \mathbf{0}. \tag{4}$$

In general, the matrix $\mathbf{U}$ satisfying (4) is not unique; however, a valid $\mathbf{U}$ may be generated by various standard linear algebra techniques, including Gauss-Jordan elimination [*Saaltink et al.*, 1998].

First we split the chemical species vector $\boldsymbol{\rho}$ between the first $N_{\mathcal{U}}$ (primary) species $\boldsymbol{\rho}_1$ and the remaining $N_{\mathcal{S}}$ (secondary) species $\boldsymbol{\rho}_2$. The stoichiometry matrix is split accordingly such that

$$\boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\rho}_1 \\ \boldsymbol{\rho}_2 \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix}. \tag{5}$$

Then the solution for $\mathbf{U}$ obtained by Gauss-Jordan elimination, denoted here as $\hat{\mathbf{U}}$, is

$$\hat{\mathbf{U}} = \left[\mathbf{1}_{N_{\mathcal{U}}}\ \mathbf{S}_1^*\right], \tag{6}$$

where $\mathbf{1}_{N_{\mathcal{U}}}$ is the $N_{\mathcal{U}} \times N_{\mathcal{U}}$ identity matrix, and

$$\mathbf{S}_1^* := -\mathbf{S}_1\mathbf{S}_2^{-1}. \tag{7}$$

Above, we may always form $\mathbf{S_2}$ such that it is invertible (see Appendix A). We note that $\hat{\mathbf{U}}$ can be modified by means of linear combinations of its rows

$$\mathbf{U} = \mathbf{L}\hat{\mathbf{U}}, \tag{8}$$

where $\mathbf{L}$ is any full-ranked $N_{\mathcal{U}} \times N_{\mathcal{U}}$ matrix, without changing the rank of the component matrix, and keeping identity (4) true. In the general case, this means we should be able to construct a nonnegative $\mathbf{U}$ such that

$$U_{ij} \geq 0, \ \forall i, j, \tag{9}$$

though we must note that at this time we do not have a rigorous proof for this claim (and see §4.3 for an example where for practical considerations we deliberately choose not to have it).

2.2 Lagrangian Representation of Components and Density Estimation

We may simulate the conservative transport problem, given in (3), by employing a Lagrangian, or particle tracking method. To do this, we discretize the "mass" for each component $\mathcal{U}_i$ into N_i numerical particles. Each particle has position $\mathbf{X}_i^p(t) \in \mathbb{R}^d, \ p = 1, \ldots, N_i, \ d = 1, 2, 3$, and carries a component mass m_i (here assumed all equal for simplicity). The component concentrations from particles that have undergone advective-dispersive transport, can then be estimated via kernel density estimation (KDE) as

$$\phi u_i(\mathbf{x}) \coloneqq \sum_{p=1}^{N_i} m_p W(\mathbf{x} - \mathbf{X}_p; \mathbf{h}_p), \tag{10}$$

where $W(\mathbf{x}; \mathbf{h})$ is a kernel function. Here it is chosen to be a multi-Gaussian, defined as

$$W(\mathbf{x}; \mathbf{h}) = \prod_{k=1}^{d} \frac{1}{\sqrt{2\pi} h_k} \exp\left(-\frac{x_k^2}{2h_k^2}\right), \tag{11}$$

in which $\mathbf{h} \equiv [h_1, \ldots, h_d]^{\mathrm{T}}$ is a vector of kernel bandwidths; $\mathbf{h}_p$ need not be equal for every particle of component p. This KDE method has been used previously to construct concentrations from Lagrangian particle clouds [*Fernàndez-Garcia and Sanchez-Vila*, 2011; *Sole-Mari and Fernàndez-Garcia*, 2018], and here we use the recently-developed technique of *Sole-Mari et al.* [2019b] to determine the optimal bandwidth for the kernel. We choose this implementation, but note that the specific choice of density estimator is not a crucial aspect of the methodology.

2.3 Solving the Equilibrium Equations

Once $\mathbf{u}$ has been estimated, expression (2) provides $N_C - N_S$ equations for determining the total mass density of the species $\boldsymbol{\rho}$. The remaining N_S equations are provided by the law of mass action:

$$\mathbf{S}^{\mathrm{T}} \log \mathbf{a} = \log \mathbf{k}, \tag{12}$$

where $\mathbf{a}$ is a vector that contains the activities of all species, and $\mathbf{k} = \left[k_1, \ldots, k_{N_S}\right]^{\mathrm{T}}$ is the vector of equilibrium constants. In general, if the total mass of species j is mobile or partially mobile (distributed in both the aqueous and solid phases), i.e., $\omega_j > 0$, the activity coefficient a_j is a nonlinear function of the aqueous concentration, and we have $a_j = c_j \gamma_j(c_j)$, where $c_j = \omega_j \rho_j$ and γ_j is the activity coefficient of species j. The activity coefficient can be calculated for instance by the Debye-Hückel equation. If species j is immobile ($\omega_j = 0$) and its mass density cannot be assumed to be constant (e.g., a surface complex species), then the activity coefficient is $a_j = \rho_j \gamma_j(\rho_j)$. For a pure mineral the activity is one as long as the mineral is present. The approach can also handle mineral exhaustion. When the solution for ρ_j reaches zero during the simulation, the hypothesis that the mineral is present is not fulfilled anymore. At this point, in order to consider that equilibrium is not reached, the algorithm eliminates the corresponding chemical reaction in the mass action law (12) and sets the activity and the total mass density of the mineral species to zero (i.e., $a_j = \rho_j = 0$) before speciation. Note that by doing this the number of chemical components is not changed.

Given some known value of $\mathbf{u}$, the chemical speciation can be determined by solving the system of equations formed by (2) and (12). To do this, it is convenient to express the primary species' mass densities as a function of the secondary species' mass densities by rearranging (2) as

$$\boldsymbol{\rho}_1(\boldsymbol{\rho}_2) = \mathbf{L}^{-1}\mathbf{u} - \mathbf{S}_1^{*}\boldsymbol{\rho}_2, \tag{13}$$

This way, we can rewrite the mass action law as

$$\mathbf{S}_1^{\mathrm{T}} \log \mathbf{a}_1(\boldsymbol{\rho}_1(\boldsymbol{\rho}_2)) + \mathbf{S}_2^{\mathrm{T}} \log \mathbf{a}_2(\boldsymbol{\rho}_2) = \log \mathbf{k}. \tag{14}$$

From this, the problem can be seen as finding the root of:

$$\mathbf{F}(\mathbf{z}) = \mathbf{S}_1^{\mathrm{T}} \log \mathbf{a}_1(\boldsymbol{\rho}_1(\mathbf{z})) + \mathbf{S}_2^{\mathrm{T}} \log \mathbf{a}_2(\mathbf{z}) - \log \mathbf{k}, \tag{15}$$

where $\mathbf{z} = \log \boldsymbol{\rho}_2$. In some very simple cases, an analytical solution exists for this problem. Otherwise, the system may be solved numerically with the application of a nonlinear solver

(e.g. Newton-Raphson iteration). Additionally, packaged geochemical solvers are available that may be applied to such a problem [e.g., *Parkhurst and Wissmeier*, 2015; *Leal*, 2015; *Steefel*, 2009]. We provide a simple Newton-Raphson iterative scheme in appendix B for chemical systems with $\mathbf{a} = \boldsymbol{\rho}$.

3 Interactions of Chemical Components with Physical Processes

The practice of reformulating our chemical system in what can be considered a reduced-order basis, composed of conservative components, has implications for how the various physical processes are simulated by a particle tracking method. This is because individual particles now carry conservative components, composed of multiple chemical species, that may have different transport or kinetic properties. In this section, we discuss how this chemical equilibrium formulation interacts with transport by advection and dispersion (Section 3.1) and also how kinetic reactions may be accounted for (Section 3.2). We then give a summary of how a user would employ the conservative-component particle tracking (CCPT) model (Section 3.3).

3.1 Interaction with Advection and Dispersion

In this section we derive the nonlinear random walk algorithm to correctly account for the effects of chemical speciation on component particles transport. We now explicitly define the advection-dispersion operator $\mathcal{L}_j$ introduced in (1) as

$$\mathcal{L}_j\left(c_j\right) \coloneqq -\nabla \cdot \left(\mathbf{q}_j c_j\right) + \nabla\nabla\colon \left(\phi \mathbf{D}_j c_j\right), \tag{16}$$

$$\mathbf{q}_j = \mathbf{q} + \nabla \cdot \left(\phi \mathbf{D}_j\right), \tag{17}$$

where $\mathbf{q}$ is the Darcy flow velocity, and $\mathbf{D}_j$ is the dispersion tensor for species C_j. Using this definition, let us rewrite expression (3) for a specific component $\mathcal{U}_i$, excluding for now the source term (see [*Sole-Mari et al.*, 2019b] for details on how this may be handled), as

$$\frac{\partial \phi u_i}{\partial t} = -\sum_{j=1}^{N_C} U_{ij} \nabla \cdot \left(\omega_j \mathbf{q}_j \rho_j\right) + \sum_{j=1}^{N_C} U_{ij} \nabla\nabla\colon \left(\omega_j \phi \mathbf{D}_j \rho_j\right). \tag{18}$$

Let us now define P_{ij} as the probability that a particle of the ith component is found in the form of the jth species. If $\mathbf{U}$ is nonnegative, i.e., if (9) holds, P_{ij} can be determined by

$$P_{ij} \coloneqq \frac{U_{ij}\rho_j}{u_i}. \tag{19}$$

From this, we can write the following governing equation for u_i,

$$\frac{\partial \phi u_i}{\partial t} = -\sum_{j=1}^{N_C} \nabla \cdot \left(P_{ij}\omega_j \mathbf{q}_j u_i\right) + \sum_{j=1}^{N_C} \nabla\nabla\colon \left(P_{ij}\omega_j \phi \mathbf{D}_j u_i\right). \tag{20}$$

We want to reproduce the combined transport equation (20) by means of a random walk particle tracking (RWPT) technique, with particles representing the distribution in space of the mass of a component $\mathcal{U}_i$. A general expression for the Itô integration of the Langevin equation [e.g., *Salamon et al.*, 2006] that describes the random motion of a particle during time-step $[t, t + \Delta t]$, is

$$\mathbf{X}_i^p\,(t + \Delta t) = \mathbf{X}_i^p\,(t) + \mathbf{A}_i\left(\mathbf{X}_i^p\,(t)\right)\Delta t + \mathbf{B}_i\left(\mathbf{X}_i^p\,(t)\right)\boldsymbol{\xi}\sqrt{\Delta t} \tag{21}$$

where $\mathbf{X}_i^p$ is the position of the pth particle of component $\mathcal{U}_i$, $\mathbf{A}_i$ is a $d \times 1$ vector (where d is the number of spatial dimensions) that defines the drift, $\mathbf{B}_i$ is a $d \times d$ matrix that defines the dispersion, and $\boldsymbol{\xi}$ is a $d \times 1$ vector of uncorrelated random numbers drawn, independently for each time step, from a standard normal distribution $\mathcal{N}(0, 1)$. Under the repeated application of (21), the density of particles f_i will obey the Fokker-Planck equation [*Risken*, 1989]:

$$\frac{\partial f_i}{\partial t} = -\nabla \cdot (\mathbf{A}_i f_i) + \frac{1}{2}\nabla\nabla\colon \left(\mathbf{B}_i \mathbf{B}_i^{\mathrm{T}} f_i\right). \tag{22}$$

Equations (20) and (22) can be made equivalent if the particle density is proportional to ϕu_i and we make the following substitutions

$$\mathbf{A}_i = \frac{1}{\phi}\left(\overline{\omega}_i \mathbf{q} + \overline{\mathbf{d}}_i\right), \qquad \mathbf{B}_i \mathbf{B}_i^{\mathrm{T}} = 2\overline{\mathbf{D}}_i, \tag{23}$$

where

$$\overline{\omega}_i \coloneqq \sum_{j=1}^{N_C} P_{ij}\omega_j, \qquad \overline{\mathbf{d}}_i \coloneqq \sum_{j=1}^{N_C} P_{ij}\omega_j \nabla \cdot \left(\phi \mathbf{D}_j\right), \qquad \overline{\mathbf{D}}_i \coloneqq \sum_{j=1}^{N_C} P_{ij}\omega_j \mathbf{D}_j. \tag{24}$$

Based on this, we see that the reduced-order system constitutes a system of coupled nonlinear Fokker-Planck equations. By implementing the random walk algorithm in (21) using parameters $\mathbf{A}_i$ and $\mathbf{B}_i$ from (23), evaluated at every pth particle's position, the particle masses will reproduce the equilibrium ADRE of every component $\mathcal{U}_i$ as given in (20). Hence, this straightforward adaptation of the transport properties allows us to correctly simulate the transport of component particles as they undergo different equilibrium states. Computation of space-varying $\mathbf{q}$, $\mathbf{D}_j$ and $\nabla \cdot \left(\phi \mathbf{D}_j\right)$ can be conducted using existing interpolation schemes [e.g., *Salamon et al.*, 2006]. It is worth highlighting the inherently Lagrangian nature of the proposed modeling interpretation of equilibrium reactive transport, with conservative component particles whose speciation is probabilistic, therefore displaying an effective transport behavior.

A common assumption, but not explicitly required in our scheme, is that all species share the same dispersion tensor at the Darcy scale, i.e., $\mathbf{D}_j = \mathbf{D},\ \forall j$. In this case $\mathbf{d}_i$ and $\mathbf{D}_i$ simplify to

$$\mathbf{d}_i = \overline{\omega}_i \nabla \cdot (\phi \mathbf{D}), \qquad \mathbf{D}_i = \overline{\omega}_i \mathbf{D}. \tag{25}$$

We note that, for mobile species, the quantity $1/\omega_i$ can be used to simulate a time-varying (linear or nonlinear) retardation factor on component $\mathcal{U}_i$. Note also that this model may also explicitly handle Langmuir adsorption in a straightforward fashion by considering available adsorption "sites" as a species and cation exchange capacity (CEC) as a component. In this case, the CEC component would not move (i.e., $\omega_j = 0,\ \forall j$), and so would not need to be carried by particles but only be represented in space according to the desired distribution. See the application to heavy metal mobilisation in Section 5 for an example of this.

When ω_i is equal to 1 and $\mathbf{D}_i$ is the same for all species, $\mathbf{A}_i$ and $\mathbf{B}_i$ become deterministic and we recover the standard random walk method for solving the advection-dispersion equation without reaction sink/source terms. That is to say that the governing equation of the component concentrations $\mathbf{u}$ given by (20) simplifies to $N_{\mathcal{U}}$ linear transport equations without reaction sink/source terms. This specific case is the only scenario considered for decoupling reactive transport in Eulerian methods [e.g., *Saaltink et al.*, 1998; *De Simoni et al.*, 2005].

3.2 Interaction with Kinetic Reactions

The reactive system presented in the previous sections considers equilibrium reactions only. Kinetic reactions are typically simulated in random walks as a decoupled process via an operator-splitting approach [e.g., *Benson and Meerschaert*, 2008; *Paster et al.*, 2013; *Sole-Mari et al.*, 2017; *Sole-Mari and Fernàndez-Garcia*, 2018]. In this section we describe one such approach to incorporate kinetic reactions that interact with the equilibrium component transport described in the previous sections. Let us consider now that there is a subset of $N_{\mathcal{T}}$ (tertiary) species that are affected only by kinetic reactions. We denote the mass density of these species as $\boldsymbol{\rho}_3$, and we rewrite the decomposition (5) of $\boldsymbol{\rho}$ and $\mathbf{S}$ as

$$\boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\rho}_1 \\ \boldsymbol{\rho}_2 \\ \boldsymbol{\rho}_3 \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \mathbf{0} \end{bmatrix}. \tag{26}$$

We consider each tertiary species as an independent component. Hence, we note the component vector and the component matrix of the equilibrium components (built as in Section 2.1) as $\mathbf{u}_{\mathcal{S}}$ and $\mathbf{U}_{\mathcal{S}}$, respectively, and redefine $\mathbf{u}$ and $\mathbf{U}$ as:

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_{\mathcal{S}} \\ \boldsymbol{\rho}_3 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} \mathbf{U}_{\mathcal{S}} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{N_{\mathcal{T}}} \end{bmatrix}. \tag{27}$$

Now let us consider the $N_C \times N_{\mathcal{R}}$ stoichiometry matrix $\mathbf{R}$ for the $N_{\mathcal{R}}$ kinetic reactions, and the corresponding reaction rates $\mathbf{r}(\mathbf{a}) = \left[r_1, \ldots, r_{N_{\mathcal{R}}}\right]^T$, such that the mass variation rates per unit of fluid volume due to kinetic reactions alone are given by the product $\mathbf{Rr}$. Equation (1) can then be extended to accommodate kinetic reactions:

$$\frac{\partial \phi \boldsymbol{\rho}}{\partial t} = \mathcal{L}(\mathbf{c}) + \phi \mathbf{S}\mathbf{s} + \phi \mathbf{R}\mathbf{r}(\mathbf{a}) + \phi \boldsymbol{\psi}, \tag{28}$$

and multiplying both sides by $\mathbf{U}$, as in (3), we have

$$\frac{\partial \phi \mathbf{u}}{\partial t} = \mathbf{U}\mathcal{L}(\mathbf{c}) + \phi \mathbf{U}\mathbf{R}\mathbf{r}(\mathbf{a}) + \phi \mathbf{U}\boldsymbol{\psi}. \tag{29}$$

The new term $\phi \mathbf{U}\mathbf{R}\mathbf{r}$ in (29) indicates that kinetic reactions, unlike equilibrium reactions, can modify the concentrations of components, according to the modified stoichiometry

matrix

$$\mathbf{R}^* = \mathbf{UR}, \tag{30}$$

and to the kinetic reaction rates $\mathbf{r}$ (which are a function of activities $\mathbf{a}$). Following previous works on kinetic reactions [*Sole-Mari and Fernàndez-Garcia*, 2018], the probability of reaction P_r of a particle in a time-step will be determined by

$$P_\mathrm{r}\left(\mathcal{U}_i \mid \Delta t\right) = \frac{\Delta u_i}{u_i} \simeq \frac{\left(\mathrm{d}u_i/\mathrm{d}t\right)\Delta t}{u_i} = \frac{\Delta t}{u_i}\sum_{k=1}^{N_\mathcal{R}} R_k^* r_k\left(\mathbf{a}\right), \tag{31}$$

where Δu_i and $\mathrm{d}u_i/\mathrm{d}t$ are the variation and the time-derivative of u_i at the particle position, respectively, considering reactions alone. Note that expression (31) considers a joint probability of reaction involving all kinetic reactions. In the case that there is more than one nonzero term in the summation, the ratios of each kth term of the summation with respect to the total are used as probabilities to determine, upon reaction, which is the actual reaction that the particle undergoes. When several reactants are involved in a kinetic reaction, it is generally preferable to determine the reaction occurrence for just one of them and, upon reaction, impose the stoichiometry on the others [*Sole-Mari et al.*, 2017].

The method by which we simulate kinetic reactions is akin to a particle-killing approach [e.g., *Benson and Meerschaert*, 2008]. In this method, a particle's probability of reaction, P_r in (31), is compared to a random number draw $Z \sim \mathcal{U}(0,1)$, and if $P_\mathrm{r} > Z$, the mass of kinetic reactant under consideration is removed from the particle. A user could also compute reactions in the "particle-number-conserving" style [e.g., *Bolster et al.*, 2016], in which a reactant's mass is reduced in proportion to P_r rather than eliminated entirely. This would be a relatively minor adaptation with no theoretical barrier to its use.

3.3 Summary of Conservative-component Particle Tracking Method

In this section, we provide a brief, high-level summary of how a conservative-component particle tracking (CCPT) simulation would be conducted. The basic steps of this process are depicted in the flowchart of Figure 1, and in the remainder of this section we refer to a step of this flowchart by its lettered denotation.

The beginning of such a simulation presupposes a system of N_C chemical species undergoing transport and interacting via $N_\mathcal{S}$ reaction channels, including any initial/boundary conditions and/or sources/sinks that are defined in terms of those species (step (a)). The next

step (step (b)) is to define the reduced-order basis of conservative components by solving for **U** in (4), which provides a mapping between chemical species and conservative components via (2). This allows the modeler to represent the relevant initial/boundary conditions and/or sources/sinks in terms of components (step (c)).

At this point, time stepping may begin. First, we map the component masses to a grid to generate concentrations (step (d)), and equilibrium reactions are calculated to get the particles' chemical speciation, as given in Section 2.3 (step (e)). At this point, we may optionally "sample" concentrations by recording the current results, for plotting or otherwise (step (f)). After speciation, these species concentrations are mapped back to the particles (step (g)). We now use these species concentrations to calculate kinetic reactions, if they are included in the model (step (h)), according to (31), and employing the defined method for simulating the reactions, as discussed in Section 3.2. As well, we use the species concentrations to calculate the particles' new transport parameters (step (i)), and move the particles accordingly (step (j)).

Finally, we impose any sources/sinks and/or boundary conditions (step (k)) and advance in time. Steps (d)-(k) are repeated until final time is reached, at which point the simulation is completed, and the final results may be depicted in terms of components or chemical species (step (l)).

4 Four Simple Test Examples

To test the presented methodology, we implement it on three variants of a simple 1D advection-dispersion-reaction problem. The chemical system involves 5 species subjected to the equilibrium reaction

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C}, \tag{32}$$

with equilibrium constant $k = 0.02$ mmol/L; and the irreversible kinetic reaction

$$\mathrm{A} + \mathrm{D} \rightarrow \mathrm{E}, \tag{33}$$

with reaction rate $s = \nu c_{\mathrm{A}} c_{\mathrm{D}}$, where $\nu = 0.04$ L/mmol/d is the reaction rate coefficient. We assume all activities to be equal to the concentrations, i.e., $\gamma_j = 1,\ \forall j$. In this simple case, the solution to the equilibrium equations can be determined analytically, which contributes to a better understanding of the proposed methodology.

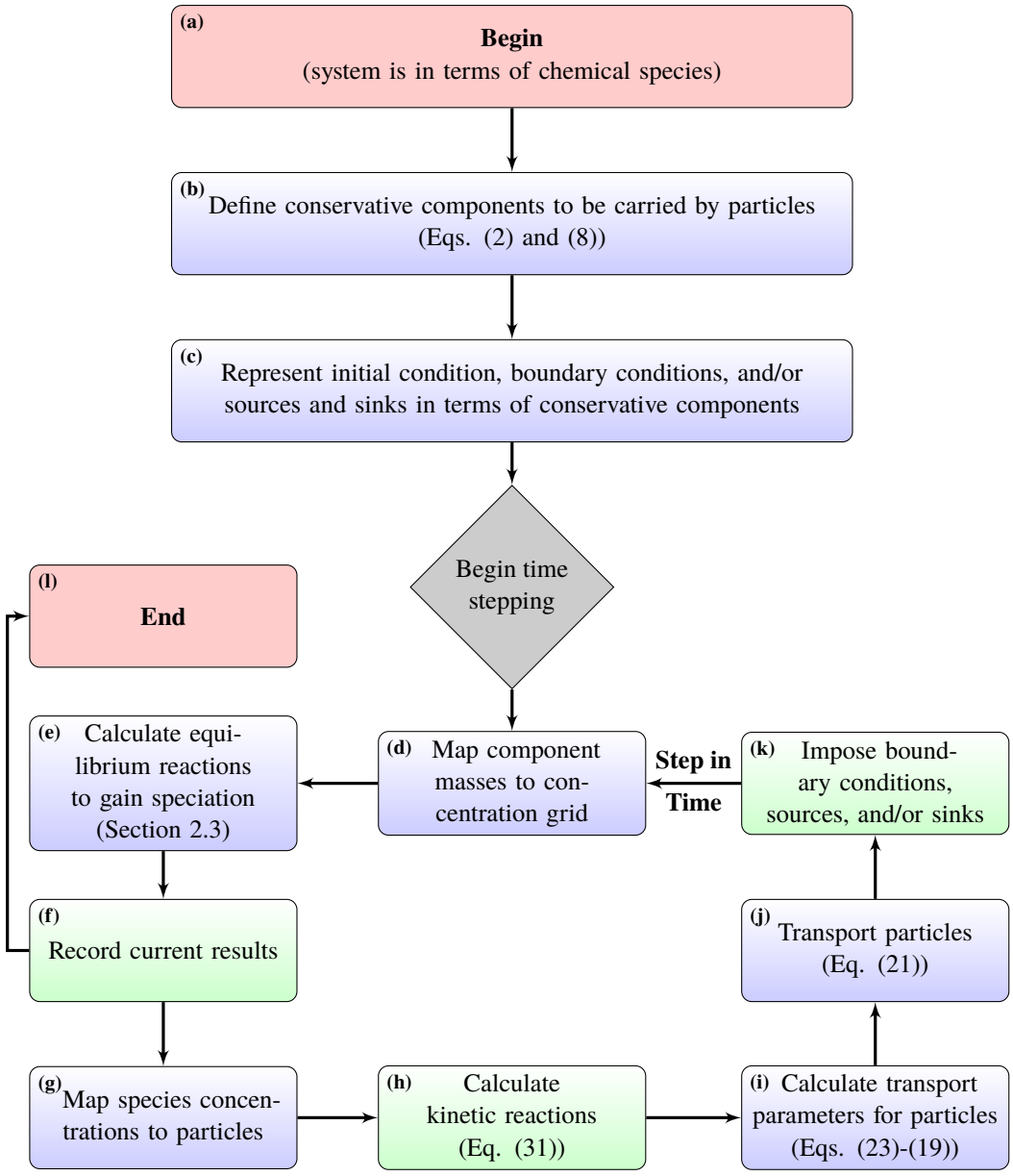

Figure 1. Flowchart depicting the order of operations for the conservative-component particle tracking reactive transport algorithm (described in Sections 2.2-3.2), in which transport is simulated by particles carrying conservative components (according to the process given in Section 2.1). Green boxes represent optional steps, in that kinetic reactions (box (h)) may not be included in the model, recording the results (box (f)) would likely be done infrequently, and all aspects of box (k) may not be included and may not be included in the final time step.

In all three cases we simulate an initial Dirac-like, instantaneous injection of species A, B and D (Figure 2) that advect through the domain, eventually mix due to dispersion, and react. The Darcy velocity is $q = 0.05$ m/d, the dispersion coefficient is $D = 0.04$ m^2/d, and the porosity is $\phi = 0.25$. These values are constant in space and time. We run the simulations for a total time of 600 days, and we sample the concentrations at the midway point of 300 days. We compare our CCPT results to an analogous Eulerian finite-difference model, with a third-order upwind scheme for advection to reduce numerical diffusion. We note that, in the cases that include precipitation/dissolution, we neglect any possible changes in the hydraulic properties of the porous medium due to these processes, as it is not germane to the proof-of-concept we seek to demonstrate.

4.1 All Mobile

In the first case, we consider that all three equilibrium species are mobile ($\omega_j = 1,\ \forall j$). Then we have the following stoichiometry matrix and component matrix

$$\mathbf{S} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{34}$$

Employing (2) and (12), for equilibrium species A, B and C, we may write

$$\begin{aligned} \rho_{\mathrm{A}} + \rho_{\mathrm{C}} &= u_{\mathrm{A}'}, \\ \rho_{\mathrm{B}} + \rho_{\mathrm{C}} &= u_{\mathrm{B}'}, \\ -\log \rho_{\mathrm{A}} - \log \rho_{\mathrm{B}} + \log \rho_{\mathrm{C}} &= \log k, \end{aligned} \tag{35}$$

where A′ and B′ are the labels of the two equilibrium components. This system of equations leads to a second-order polynomial equation with the following solutions for species concentrations

$$\begin{aligned} \rho_{\mathrm{A}} &= \frac{1}{2}\left(u_{\mathrm{A}'} - u_{\mathrm{B}'} - k^{-1}\right) + \sqrt{\frac{1}{4}\left(u_{\mathrm{A}'} - u_{\mathrm{B}'} - k^{-1}\right)^2 + u_{\mathrm{A}'}k^{-1}}, \\ \rho_{\mathrm{B}} &= \frac{1}{2}\left(u_{\mathrm{B}'} - u_{\mathrm{A}'} - k^{-1}\right) + \sqrt{\frac{1}{4}\left(u_{\mathrm{B}'} - u_{\mathrm{A}'} - k^{-1}\right)^2 + u_{\mathrm{B}'}k^{-1}}, \\ \rho_{\mathrm{C}} &= \frac{1}{2}\left(u_{\mathrm{A}'} + u_{\mathrm{B}'} + k^{-1}\right) - \sqrt{\frac{1}{4}\left(u_{\mathrm{A}'} + u_{\mathrm{B}'} + k^{-1}\right)^2 + u_{\mathrm{A}'}u_{\mathrm{B}'}}. \end{aligned} \tag{36}$$

The kinetic reaction stoichiometry matrix, expressed in terms of components, obtained via (30), is

$$\mathbf{R}^* = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \tag{37}$$

i.e., the reaction (33) can be written in terms of components as

$$\mathrm{A}' + \mathrm{D} \rightarrow \mathrm{E}. \tag{38}$$

Results of this numerical experiment are shown in Figure 2(a1) and (b1), where we see near-exact match between our CCPT results and the base-case Eulerian results for both presented times. The quality of match is no different for plotting the conservative components in 2(a1) or the full chemical speciation in Figure 2(b1).

4.2 Species-dependent dispersion and retardation

Here we maintain the chemistry of the previous example, and introduce an increased dispersion coefficient for species B ($D_{\mathrm{B}} = 0.16$), as well as a retardation ($\omega_{\mathrm{B}} = 0.5$). This example shows the method's ability to deal with species-dependent mobilities and dispersion coefficients with conservative components. Additionally, in order to make this case comparable to 4.1, since aqueous concentrations of species B will be half of the total mass density due to retardation, we set the kinetic reaction rate to $\nu = 0.08$ L/mmol/d.

Like in the previous example, there is near-exact match between the CCPT and the FD results at both presented times. The combination of retardation and multi-component reaction results in a markedly asymmetric plume of species B, as well as a bimodal late-time distribution of component B′. When B′ coincides with A′, both partially speciate as C, which in turn (*i*) reduces the rate of kinetic reaction between A and D and (*ii*) allows part of component B′ to move along with the rest of components without retardation. But as part of the latter starts lagging behind because of retardation B, more C transforms back to A and B.

This combination of behaviors does not necessarily represent any particular biogeochemical system, but it compellingly shows the ability of the CCPT method to deal with species-dependent transport through chemically conservative components, one of the main advantages and novel features of our proposed Lagrangian approach.

4.3 Precipitate C

In this case we introduce the variation that chemical species C is a solid precipitate. Therefore, it has zero mobility and unitary activity (as long as it is present). This modifies the system of equilibrium equations (35), which can be rewritten as:

$$\begin{aligned} \rho_{\mathrm{A}} + \rho_{\mathrm{C}} &= u_{\mathrm{A}'}, \\ \rho_{\mathrm{B}} + \rho_{\mathrm{C}} &= u_{\mathrm{B}'}, \\ -\log \rho_{\mathrm{A}} - \log \rho_{\mathrm{B}} &= \log k, \end{aligned} \tag{39}$$

and the solution, assuming $\rho_{\mathrm{C}} > 0$, is

$$\begin{aligned} \rho_{\mathrm{A}} &= \frac{1}{2}\left(u_{\mathrm{A}'} - u_{\mathrm{B}'}\right) + \sqrt{\frac{1}{4}\left(u_{\mathrm{A}'} - u_{\mathrm{B}'}\right)^2 + k^{-1}}, \\ \rho_{\mathrm{B}} &= \frac{1}{2}\left(u_{\mathrm{B}'} - u_{\mathrm{A}'}\right) + \sqrt{\frac{1}{4}\left(u_{\mathrm{B}'} - u_{\mathrm{A}'}\right)^2 + k^{-1}}, \\ \rho_{\mathrm{C}} &= \frac{1}{2}\left(u_{\mathrm{A}'} + u_{\mathrm{B}'}\right) - \sqrt{\frac{1}{4}\left(u_{\mathrm{A}'} - u_{\mathrm{B}'}\right)^2 + k^{-1}}. \end{aligned} \tag{40}$$

Note that ρ_{C} in (40) is not physically a concentration (since C is not a solute), but we can still determine its value as mass of C per unit volume of fluid. If we obtain $\rho_{\mathrm{C}} \leq 0$, the hypothesis that C is present is not fulfilled (i.e., the equilibrium is not reached). In this case, we need to remove the third equation in (39) and impose $\rho_{\mathrm{C}} = 0$. Then the solution is simply

$$\begin{aligned} \rho_{\mathrm{A}} &= u_{\mathrm{A}'}, \\ \rho_{\mathrm{B}} &= u_{\mathrm{B}'}, \\ \rho_{\mathrm{C}} &= 0. \end{aligned} \tag{41}$$

Since the component matrix $\mathbf{U}$ did not change with respect to the previous case, the kinetic reaction in terms of components is still defined by (37) and (38).

In Figure 2 (a3) and (b3) we see the results of this numerical experiment, and, as in the previous section, we see excellent agreement between our CCPT results and the Eulerian results we use for verification.

4.4 Ubiquitous Precipitate C

In this example, we further assume that there is an underlying high density of the solid precipitate C such that it is always present. The nature of this condition requires some nuanced computational considerations. Because the activity of C is invariably constant, the

reaction (32) can be effectively replaced by

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \varnothing. \tag{42}$$

Therefore, we have

$$\mathbf{S} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{43}$$

and equations (2) and (12) for species A and B reduce to

$$\begin{aligned} \rho_{\mathrm{A}} - \rho_{\mathrm{B}} &= u, \\ -\log \rho_{\mathrm{A}} - \log \rho_{\mathrm{B}} &= \log k. \end{aligned} \tag{44}$$

The solution of (44) is

$$\begin{aligned} \rho_{\mathrm{A}} &= \frac{1}{2}u + \sqrt{\frac{1}{4}u^2 + k^{-1}}, \\ \rho_{\mathrm{B}} &= -\frac{1}{2}u + \sqrt{\frac{1}{4}u^2 + k^{-1}}. \end{aligned} \tag{45}$$

This simple chemical system is studied in-depth by *De Simoni et al.* [2005], in the context of mixing of natural waters at different equilibria. Regarding the kinetic reaction (33), the modified stoichiometry matrix, given by (30), is in this case

$$\mathbf{R}^* = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \tag{46}$$

i.e., the kinetic reaction in terms of components is

$$\mathcal{U} + \mathrm{D} \rightarrow \mathrm{E}. \tag{47}$$

Note that choosing not to include C particles and instead simulating the equilibrium reaction given in (42) means that one may not construct a strictly non-negative component matrix $\mathbf{U}$ (i.e., it will not fulfill (9)). As such, there is one row (row 1 for the $\mathbf{U}$ we construct in (43)) that contains negative entries. This means that "negative-mass" component-particles will be required to represent the species corresponding to the negative entry in that row, in this case, reactant B. This can be seen in Figure 2 (a3) wherein a negative Dirac point-source of $\mathcal{U}$ is placed at the initial location corresponding to the initial condition of B in Figure 2 (b3). This complicates the physical interpretation of such negative-mass particles, including

their transport behavior and the computation of kinetic reactions; however, negative-mass particles are transported in the same manner as positive-mass particles, from a computational standpoint and are increasingly common in other random walk applications, such as for modeling systems with discontinuous dispersion coefficients [*Oukili et al.*, 2019].

The results of the simulation of this system can be seen in 2 (a3) and (b3), and we once again see that the CCPT results show near-perfect match to the base-case Eulerian solutions.

5 An Illustrative Implementation Example: Mobilisation of Cadmium Following a CO_2 Leak

In this section we use the proposed methodology to study a hypothetical problem of cadmium mobilisation promoted by the leakage of dissolved CO_2 in a physically and chemically heterogeneous saturated porous medium. For this model, we consider a 50 m by 100 m 2D domain, in which steady-state Darcy flow has been solved for the hydarulic conductivity field depicted in the upper plot of Figure 3. The domain is impermeable at the top and bottom boundaries, and a fixed head is imposed at the left and right boundaries generating a mean domain-wise rightward velocity $\overline{v} = 0.0747$ m/d. Dispersion is anisotropic, characterized by a longitudinal dispersivity $\alpha_\ell = 0.2$ m and a transverse dispersivity $\alpha_t = \alpha_\ell/10$. The porosity has a uniform value $\phi = 0.25$. Sorption sites are considered to only be present where $\log K \leq 0$ and proportional to the quantity $-\log K$; this is represented by cation-exchange capacity (CEC, mmol/L) in the lower plot of Figure 3. Over the course of the simulation, cadmium ions $\left(Cd^{2+}\right)$ are injected into the aquifer during days 0-500, mobilising the already-present calcium ions $\left(Ca^{2+}\right)$ away from the sorption sites, effectively slowing the travel of the cadmium plume. After some setting time, CO_2 is then released over days 2000-2050, increasing the aqueous calcium concentrations and leading to a decrease in the tendency of aqueous cadmium ions to sorb. A movie depicting the entire process can be found in the Supplementary Material.

From a chemical standpoint, we consider the following equilibrium reactions

$$\begin{aligned} Ca^{2+} + X_2Cd &\rightleftharpoons Cd^{2+} + X_2Ca, \\ Ca^{2+} + CO_3^{2-} &\rightleftharpoons CaCO_3, \\ CO_3^{2-} + 2H^+ &\rightleftharpoons H_2O + CO_2, \\ H^+ + OH^- &\rightleftharpoons H_2O, \\ CO_3^{2-} + H^+ &\rightleftharpoons HCO_3^-, \end{aligned} \tag{48}$$

with the associated equilibrium constants

$$\log \boldsymbol{k}^{\mathrm{T}} = \begin{bmatrix} 0.00 & 8.48 & 16.67 & 13.99 & 10.33 \end{bmatrix}. \tag{49}$$

Note that, in (48), species X_2 is an arbitrary anion to which calcium or cadmium could bind and sorb into a solid phase. However, as in Section 4.3, we consider both water (H_2O) and calcite ($CaCO_3$) to be ubiquitous, and thus, to simplify the simulation, we effectively replace the second and fourth equations, above, with

$$\begin{aligned} \mathrm{Ca}^{2+} + \mathrm{CO}_3^{2-} &\rightleftharpoons \varnothing, \\ \mathrm{H}^{+} + \mathrm{OH}^{-} &\rightleftharpoons \varnothing. \end{aligned} \tag{50}$$

This leads to the stoichiometry matrix

$$\mathbf{S}^{\mathrm{T}} = \begin{array}{c} \begin{matrix} \mathrm{Ca}^{2+} & \mathrm{Cd}^{2+} & \mathrm{CO}_2 & \mathrm{X_2Ca} & \mathrm{CO}_3^{2-} & \mathrm{HCO}_3^{-} & \mathrm{H}^{+} & \mathrm{OH}^{-} & \mathrm{X_2Cd} \end{matrix} \\ \begin{bmatrix} -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & 0 \end{bmatrix} \end{array}, \tag{51}$$

and, from this, we may compute the component matrix

$$\mathbf{U} = \begin{array}{c} \begin{matrix} \mathrm{Ca}^{2+} & \mathrm{Cd}^{2+} & \mathrm{CO}_2 & \mathrm{X_2Ca} & \mathrm{CO}_3^{2-} & \mathrm{HCO}_3^{-} & \mathrm{H}^{+} & \mathrm{OH}^{-} & \mathrm{X_2Cd} \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & -1 & -0.5 & 0.5 & -0.5 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0.5 & 0.5 & -0.5 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{l} \textbf{Ions} \\ \textbf{Total Cadmium} \\ \textbf{Acidity} \\ \textbf{CEC} \end{array}. \tag{52}$$

Above, the rows of $\mathbf{U}$ are labeled with the physical interpretations for a component (row) that naturally arise from the chosen component definitions. Here, we implemented a simple Newton-Raphson iterative scheme for solving the speciation (see Appendix B).

While we have no analytical solutions for this system, we analyze two time-dependent quantities to determine whether we have captured the proper dynamics of the system. We achieve this by comparing results of a simulation that include the previously-mentioned CO_2 release to results that do not include a CO_2 release, and we depict these both in Figure 4. In Figure 4(a), we see a plot of mean mobility of cadmium with respect to time (days)–where

mean mobility is defined as the particle average of the ratio between solute and fluid velocities, which is given by ω_i (see (23)). There, we see the same behavior for $t \in [0, 2000]$ d. Then, we see the expected behavior that, in the case with CO_2 release, there is a significant spike in mobile Cd^{2+} concentration occurring shortly after the acidifying CO_2 release. In Figure 4(b), we depict the breakthrough concentration of Cd^{2+} ions (mmol/L) at the outlet ($x_1 = 100$ m) versus time, measured in days. In this case, we see the cadmium breakthrough steadily decreasing with time until the release of CO_2, and, after a short time, the results of the two simulations diverge in that there is a significant increase in cadmium measured at the system outlet, caused by the CO_2-carrying water mobilizing calcium ions and displacing cadmium.

6 Conclusions

Simulating a reactive transport system can become computationally demanding, or infeasible, when the requisite number of chemical species becomes large. To address this concern, many modelers choose to redefine the system in terms of a reduced-order basis of conservative components. Previous to this work, such an approach had yet to be satisfactorily adapted for Lagrangian, or particle tracking, methods. The reason for this is that a subset of the conservative components are necessarily made up of multiple chemical species, which can have different transport properties. With varying chemical species "formulae" stored on conservative-component-carrying particles, one must account for this on the way a particle effectively moves.

In this manuscript, we solve the above issue by presenting a method that computes an effective drift and dispersion tensor for a particle, based on the proportions of chemical species that it carries and those species' respective mobility and species-dependent transport properties. Also, we discuss how established particle methods may be used to simulate kinetic reactions that occur in concert with equilibrium reactions. We describe how a conservative-component particle tracking (CCPT) model would be simulated. Finally, we apply the CCPT method to model some simple benchmark test problems that include both equilibrium and kinetic reactions. All results exhibit ideal match with base-case Eulerian results. Finally, we apply the CCPT method to model a 2D problem concerning the mobilisation of cadmium by a CO_2 leak, showing the potential applicability of the proposed methodology.

To summarize, the major contributions of this research are as follows:

- PT methods may now reduce associated computational burden of reactive transport models by reformulating the system into a reduced-order conservative-component basis.
- The reduced-order formulation can easily handle species-dependent transport operators and it is therefore not forced to consider species with equal dispersion and/or retardation coefficients.
- Transport properties are rigorously defined for these component-carrying particles based on their chemical speciation at a given time.
- Because the components carried by particles are conservative, equilibrium reactions can be simulated without altering the number (or mass) of particles in a system. This means:
  - resolution does not decrease with time, as reactant species particles need not be removed due to reaction.
  - there is no need to add particles of product species, a practice that can increase computational load. This has the added value of eliminating the non-trivial question of where to place new product-species particles.
- Interaction between equilibrium and kinetic reactions is simplified through probabilities of kinetic reaction with effective stoichiometry.

A Justification for the existence of an invertible block in the stoichiometry matrix

We consider the proposition that, given the $N_C \times N_S$ matrix

$$\mathbf{A} \coloneqq \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix}, \tag{A.1}$$

in which $\mathbf{S}_1$ is $N_\mathcal{U} \times N_S$ and $\mathbf{S}_2$ is $N_S \times N_S$, where N_C is necessarily greater than N_S, then $\mathbf{S}_2$ is always invertible.

In order for $\mathbf{S}_2$ to be invertible, we must have $\text{rank}(\mathbf{S}_2) = N_S$. As long as $\text{rank}(\mathbf{S}) = N_S$, and because the column rank is equal to the the row rank for any matrix, we can always construct an $N_S \times N_S$ block, namely $\mathbf{S}_2$, that is invertible. However, if $\mathbf{S}$ does not have full column rank, that means that there are linearly-dependent columns (reactions), meaning that those reactions can be given in terms of linear combinations of other reactions. In this case, we may remove any linearly-dependent column of $\mathbf{S}$, and if necessary, they can be recovered later as linear combinations of other reactions. As such, we may remove those reactions from

consideration by eliminating the linearly-dependent columns of $\mathbf{S}$ to form $\tilde{\mathbf{S}}$, and if necessary, they can be recovered later as linear combinations of other reactions. Then, $\text{rank}(\tilde{\mathbf{S}}) = N_{\mathcal{S}}$, and, as before, we may form an invertible block $\mathbf{S}_2$.

B Speciation for Dilute Solutions

In this appendix we present a simple Newton-Raphson iterative scheme for solving speciation when the activity of the species is equal to the mass density of the species, i.e., $\mathbf{a} = \boldsymbol{\rho}$. In this case, from (14), we have that the problem of speciation can be rewritten as finding the root of

$$\mathbf{F}(\mathbf{z}) = \mathbf{S}_1^{\mathrm{T}} \ln \boldsymbol{\rho}_1(\mathbf{z}) + \mathbf{S}_2^{\mathrm{T}} \mathbf{z} - \ln \mathbf{k}, \tag{B.1}$$

where $\mathbf{z} = \ln \boldsymbol{\rho}_2$, and the relationship between primary and secondary species (13) is rewritten as

$$\boldsymbol{\rho}_1(\mathbf{z}) = \mathbf{L}^{-1}\mathbf{u} - \mathbf{S}_1^{*} \exp(\mathbf{z}). \tag{B.2}$$

The Jacobian of $\mathbf{F}$ is

$$\mathbf{J}(\mathbf{z}) \coloneqq \frac{\partial \mathbf{F}}{\partial \mathbf{z}} = -\mathbf{S}_1^{\mathrm{T}} \left([\exp(\mathbf{z}^{\mathrm{T}}) \boldsymbol{\rho}_1^{-1}(\mathbf{z})] \odot \mathbf{S}_1^{*} \right) + \mathbf{S}_2^{\mathrm{T}}, \tag{B.3}$$

where $\boldsymbol{\rho}_1^{-1}$ is a vector whose entries are the inverse of those in $\boldsymbol{\rho}_1$, and $\odot$ is the Hadamard (entry-wise) product. The Newton-Raphson iterative scheme is then

$$\mathbf{z}^{k+1} = \mathbf{z}^{k} - \mathbf{J}^{-1}(\mathbf{z}^{k}) \mathbf{F}(\mathbf{z}^{k}). \tag{B.4}$$

Acknowledgments

This work was supported by the US Army Research Office under Contract/Grant number W911NF-18-1-0338. This work was partially supported by the Spanish Ministry of Economy and Competitiveness through projects WE-NEED (PCIN-2015-248) and INDEMNE (CGL2015-69768-R). All of the data presented in this paper was created using the BAKS code available at https://zenodo.org/record/3358178.

References

Benson, D. A., and D. Bolster (2016), Arbitrarily complex chemical reactions on particles, *Water Resources Research*, *52*(11), 9190–9200, doi:10.1002/2016WR019368.

Benson, D. A., and M. M. Meerschaert (2008), Simulation of chemical reaction via particle tracking: Diffusion-limited versus thermodynamic rate-limited regimes, *Water Resources Research*, *44*(12), doi:10.1029/2008WR007111.

Benson, D. A., T. Aquino, D. Bolster, N. Engdahl, C. V. Henri, and D. Fernàndez-Garcia (2017), A comparison of Eulerian and Lagrangian transport and non-linear reaction algorithms, *Advances in Water Resources*, *99*, 15–37, doi:10.1016/j.advwatres.2016.11.003.

Benson, D. A., S. Pankavich, and D. Bolster (2019), On the separate treatment of mixing and spreading by the reactive-particle-tracking algorithm: An example of accurate upscaling of reactive Poiseuille flow, *Advances in Water Resources*, *123*, 40 – 53, doi:https://doi.org/10.1016/j.advwatres.2018.11.001.

Benson, D. A., S. Pankavich, M. J. Schmidt, and G. Sole-Mari (2020), Entropy: 1) the former trouble with particle-tracking simulation, and 2) a measure of computational information penalty, *Advances in Water Resources*, p. 103509, doi:https://doi.org/10.1016/j.advwatres.2020.103509.

Bolster, D., D. A. Benson, T. Le Borgne, and M. Dentz (2010), Anomalous mixing and reaction induced by superdiffusive nonlocal transport, *Phys. Rev. E*, *82*, 021,119, doi:10.1103/PhysRevE.82.021119.

Bolster, D., A. Paster, and D. A. Benson (2016), A particle number conserving Lagrangian method for mixing-driven reactive transport, *Water Resources Research*, *52*(2), 1518–1527, doi:10.1002/2015WR018310.

De Simoni, M., J. Carrera, X. Sánchez-Vila, and A. Guadagnini (2005), A procedure for the solution of multicomponent reactive transport problems, *Water Resources Research*, *41*(11), doi:10.1029/2005WR004056.

De Simoni, M., X. Sanchez-Vila, J. Carrera, and M. W. Saaltink (2007), A mixing ratios-based formulation for multicomponent reactive transport, *Water Resources Research*, *43*(7), doi:10.1029/2006WR005256.

Dentz, M., T. Le Borgne, A. Englert, and B. Bijeljic (2011), Mixing, spreading and reaction in heterogeneous media: A brief review, doi:10.1016/j.jconhyd.2010.05.002.

Engdahl, N. B., D. A. Benson, and D. Bolster (2017), Lagrangian simulation of mixing and reactions in complex geochemical systems, *Water Resources Research*, *53*(4), 3513–3522,

569 doi:10.1002/2017WR020362.

570 Engdahl, N. B., M. J. Schmidt, and D. A. Benson (2019), Accelerating and parallelizing La-
571 grangian simulations of mixing-limited reactive transport, *Water Resources Research*, *55*.

572 Fernàndez-Garcia, D., and X. Sanchez-Vila (2011), Optimal reconstruction of concentra-
573 tions, gradients and reaction rates from particle distributions, *Journal of Contaminant Hy-*
574 *drology*, *120-121*(C), 99–114, doi:10.1016/j.jconhyd.2010.05.001.

575 Gramling, C. M., C. F. Harvey, and L. C. Meigs (2002), Reactive transport in porous media:
576 A comparison of model prediction with laboratory visualization, *Environmental Science &*
577 *Technology*, *36*(11), 2508–2514, doi:10.1021/es0157144.

578 Herrera, P. A., M. Massabó, and R. D. Beckie (2009), A meshless method to simulate solute
579 transport in heterogeneous porous media, *Advances in Water Resources*, *32*(3), 413–429,
580 doi:10.1016/j.advwatres.2008.12.005.

581 Herrera, P. A., J. M. Cortínez, and A. J. Valocchi (2017), Lagrangian scheme to model
582 subgrid-scale mixing and spreading in heterogeneous porous media, *Water Resources Re-*
583 *search*, *53*(4), 3302–3318, doi:10.1002/2016WR019994.

584 Kräutle, S., and P. Knabner (2005), A new numerical reduction scheme for fully coupled
585 multicomponent transport-reaction problems in porous media, *Water Resources Research*,
586 *41*(9), doi:10.1029/2004WR003624.

587 Leal, A. (2015), Reaktoro: A unified framework for modeling chemically reactive systems.

588 Molins, S., J. Carrera, C. Ayora, and M. W. Saaltink (2004), A formulation for decoupling
589 components in reactive transport problems, *Water Resources Research*, *40*(10), doi:10.
590 1029/2003WR002970.

591 Monaghan, J. (2012), Smoothed particle hydrodynamics and its diverse applications, *Annual*
592 *Review of Fluid Mechanics*, *44*(1), 323–346, doi:10.1146/annurev-fluid-120710-101220.

593 Oukili, H., R. Ababou, G. Debenest, and B. Noetinger (2019), Random walks with negative
594 particles for discontinuous diffusion and porosity, *Journal of Computational Physics*, *396*,
595 687 – 701, doi:https://doi.org/10.1016/j.jcp.2019.07.006.

596 Parkhurst, D. L., and L. Wissmeier (2015), PhreeqcRM: A reaction module for transport sim-
597 ulators based on the geochemical model PHREEQC, *Advances in Water Resources*, *83*,
598 176–189, doi:10.1016/j.advwatres.2015.06.001.

599 Paster, A., D. Bolster, and D. A. Benson (2013), Particle tracking and the diffusion-reaction
600 equation, *Water Resources Research*, *49*(1), 1–6, doi:10.1029/2012WR012444.

Paster, A., D. Bolster, and D. A. Benson (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, doi:10.1016/j.jcp.2014.01.020.

Rahbaralam, M., D. Fernàndez-Garcia, and X. Sanchez-Vila (2015), Do we really need a large number of particles to simulate bimolecular reactive transport with random walk methods? A kernel density estimation approach, *Journal of Computational Physics*, *303*, 95–104, doi:10.1016/j.jcp.2015.09.030.

Risken, H. (1989), *The Fokker-Planck equation. Methods of solution and applications*.

Rizzo, C. B., A. Nakano, and F. P. de Barros (2019), PAR2: Parallel random walk particle tracking method for solute transport in porous media, *Computer Physics Communications*, *239*, 265 – 271, doi:https://doi.org/10.1016/j.cpc.2019.01.013.

Saaltink, M. W., C. Ayora, and J. Carrera (1998), A mathematical formulation for reactive transport that eliminates mineral concentrations, *Water Resources Research*, *34*(7), 1649–1656, doi:10.1029/98WR00552.

Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2006), A review and numerical assessment of the random walk particle tracking method, *Journal of Contaminant Hydrology*, *87*(3), 277 – 305, doi:https://doi.org/10.1016/j.jconhyd.2006.05.005.

Sanchez-Vila, X., and D. Fernàndez-Garcia (2016), Debates-Stochastic subsurface hydrology from theory to practice: Why stochastic modeling has not yet permeated into practitioners?, *Water Resources Research*, *52*(12), 9246–9258, doi:10.1002/2016WR019302.

Schmidt, M. J., S. Pankavich, and D. A. Benson (2017), A Kernel-based Lagrangian method for imperfectly-mixed chemical reactions, *Journal of Computational Physics*, *336*, 288–307, doi:10.1016/j.jcp.2017.02.012.

Schmidt, M. J., S. D. Pankavich, A. Navarre-Sitchler, N. B. Engdahl, D. Bolster, and D. A. Benson (2019a), Reactive particle-tracking solutions to a benchmark problem on heavy metal cycling in lake sediments, *Submitted*, doi:https://arxiv.org/abs/1908.09818.

Schmidt, M. J., S. D. Pankavich, A. Navarre-Sitchler, and D. A. Benson (2019b), A Lagrangian method for reactive transport with solid/aqueous chemical phase interaction, *Journal of Computational Physics: X*, p. 100021, doi:https://doi.org/10.1016/j.jcpx.2019.100021.

Schmidt, M. J., N. B. Engdahl, S. D. Pankavich, and D. Bolster (2020), A mass-transfer particle-tracking method for simulating transport with discontinuous diffusion coefficients,

Advances in Water Resources, p. 103577, doi:https://doi.org/10.1016/j.advwatres.2020.103577.

Sole-Mari, G., and D. Fernàndez-Garcia (2018), Lagrangian modeling of reactive transport in heterogeneous porous media with an automatic locally adaptive particle support volume, *Water Resources Research*, *54*(10), 8309–8331, doi:10.1029/2018WR023033.

Sole-Mari, G., D. Fernàndez-Garcia, P. Rodríguez-Escales, and X. Sanchez-Vila (2017), A KDE-Based Random Walk Method for Modeling Reactive Transport With Complex Kinetics in Porous Media, *Water Resources Research*, doi:10.1002/2017WR021064.

Sole-Mari, G., M. J. Schmidt, S. D. Pankavich, and D. A. Benson (2019a), Numerical equivalence between SPH and probabilistic mass transfer methods for Lagrangian simulation of dispersion, *Advances in Water Resources*, doi:https://doi.org/10.1016/j.advwatres.2019.02.009.

Sole-Mari, G., D. Bolster, D. Fernàndez-Garcia, and X. Sanchez-Vila (2019b), Particle density estimation with grid-projected and boundary-corrected adaptive kernels, *Advances in Water Resources*, *131*, 103,382, doi:https://doi.org/10.1016/j.advwatres.2019.103382.

Steefel, C. I. (2009), *CrunchFlow Software for Modeling Multicomponent Reactive Flow and Transport CrunchFlow CRUNCHFLOW, Software for Modeling Multicomponent Reactive Flow and Transport, USER'S MANUAL*, Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA.

Valocchi, A. J., D. Bolster, and C. J. Werth (2019), Mixing-limited reactions in porous media, *Transport in Porous Media*, *130*(1), 157–182, doi:10.1007/s11242-018-1204-1.

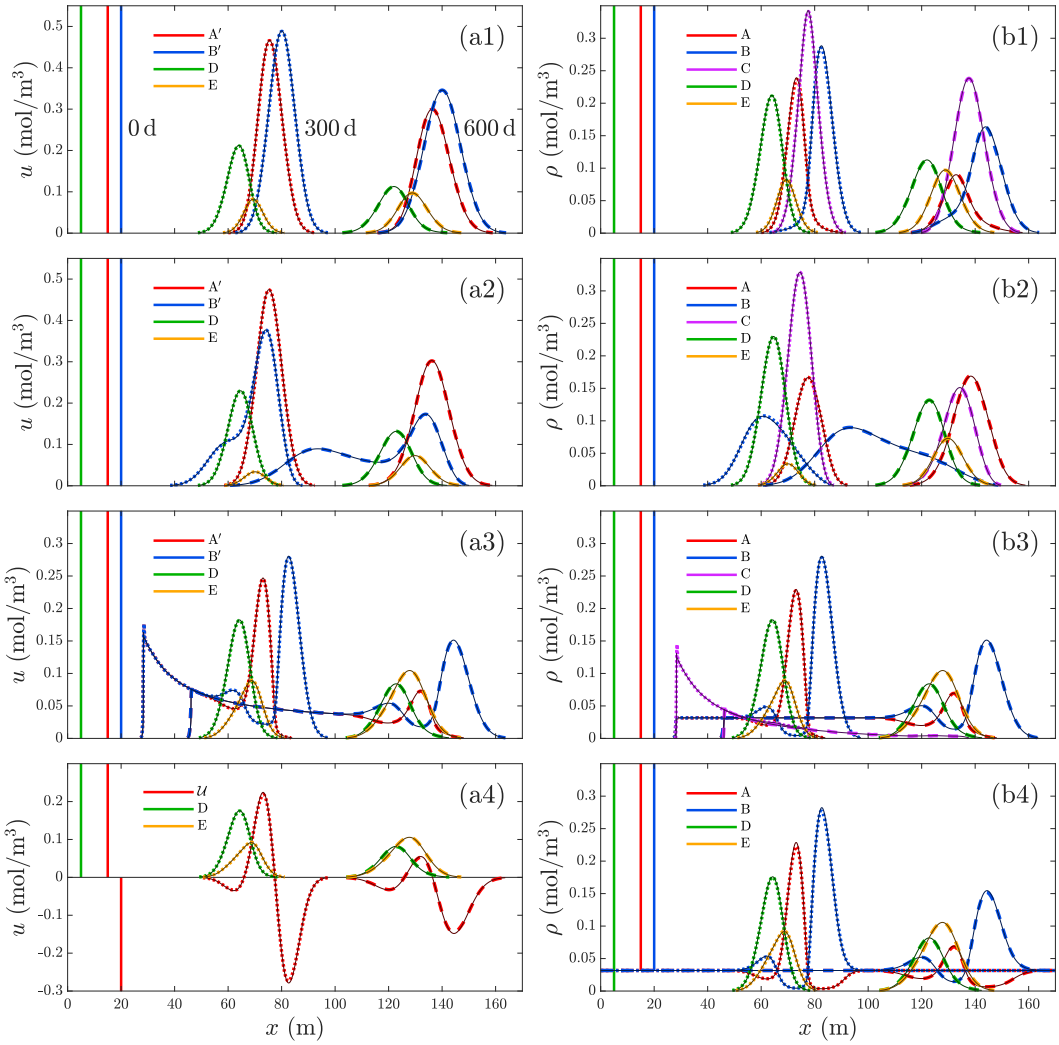

Figure 2. Results for 1D advection-dispersion-reaction simulations. Left plots (a1)-(a4) depict results in terms of conservative components, and right plots (b1)-(b4) depict chemical species. Initial Dirac-like conditions are depicted for $t = 0$ days (solid vertical bars), results are sampled at 300 days (dotted curves), and final results are shown for 600 days (dashed curves). Four cases are considered: (Top) All mobile case (Section 4.1); (Middle-Top) Species-dependent mobility and dispersion case (Section 4.2); (Middle-Bottom) Case including a solid-species precipitate (Section 4.3); (Bottom) Case in which there is an always-present amount of solid precipitate (Section 4.4). In all cases, the conservative-component particle tracking simulation results (colored curves) are plotted against trusted Eulerian results (gray curves) for comparison, and we see near-perfect agreement between the two models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

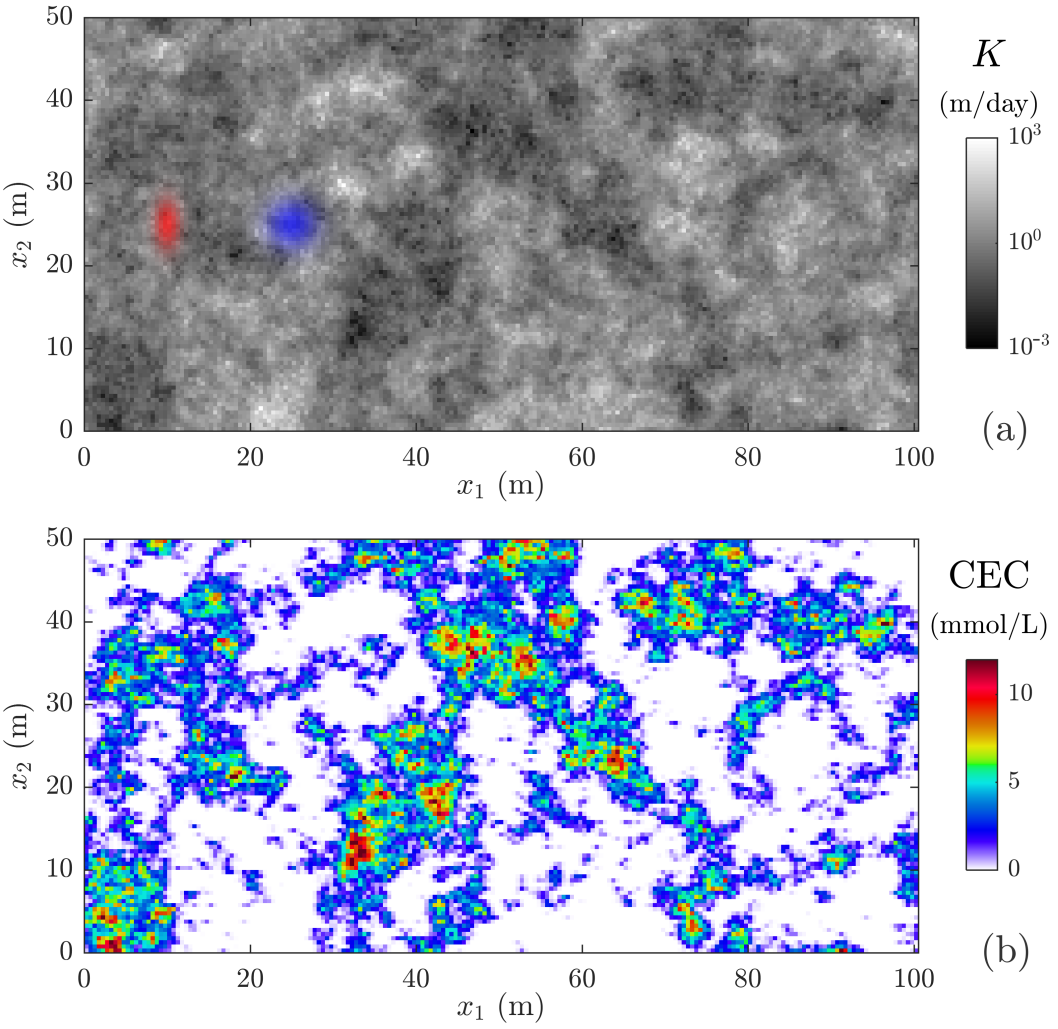

Figure 3. Depiction of Cadmium mobilisation simulation. (Top) Location of release of Cd^{2+} into the aquifer from days 0-500 days (blue) and location of release of CO_2 from days 2000-2050 (red), superimposed upon the steady-state flow field (greyscale gradient representing hydraulic conductivity). (Bottom) Cation-exchange capacity (CEC) distribution in space, which is proportional to $-\log K$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

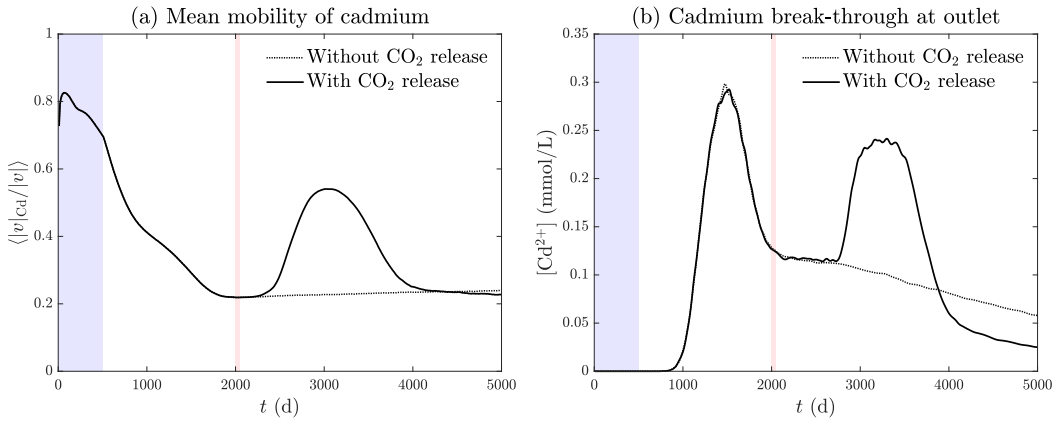

Figure 4. (a) Mean mobility of cadmium over time without CO_2 release (dotted curve) and with CO_2 release (solid curve). (b) Cadmium break-through at outlet ($x_1 = 100$ m) over time without CO_2 release (dotted curve) and with CO_2 release (solid curve). In both plots, the blue region represents the time period over which Cd^{2+} is released, and the red region represents the time period over which CO_2 is released. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)