A KDE-based random walk method for modeling reactive transport with complex kinetics in porous media

Guillem Sole-Mari[1,2], Daniel Fernàndez-Garcia[1,2], Paula Rodríguez-Escales[1,2], Xavier Sanchez-Vila[1,2]

[1]Department of Civil and Environmental Engineering (DECA), Universitat Politècnica de Catalunya (UPC), Jordi Girona 1-3, 08034 Barcelona, Spain.

[2]Associated Unit: Hydrogeology Group (UPC-CSIC).

Corresponding author: Guillem Sole-Mari (guillem.sole.mari@upc.edu)

Key points

- A Random Walk Particle Tracking Method capable to simulate reactions with complex kinetics is presented
- Particles are equipped with optimal kernels to represent the uncertainty in the particle position driven by subsampling a large population
- The method is implemented in a column transport model and tested for four different reactive transport case examples

Abstract

In recent years a large body of literature has been devoted to study reactive transport of solutes in porous media based on pure Lagrangian formulations. Such approaches have also been extended to accommodate second-order bimolecular reactions, in which the reaction rate is proportional to the concentrations of the reactants. Rather, in some cases, chemical reactions involving two reactants follow more complicated rate laws. Some examples are (1) reaction rate laws written in terms of powers of concentrations, (2) redox reactions incorporating a limiting term (e.g. Michaelis-Menten), or (3) any reaction where the activity coefficients vary with the concentration of the reactants, just to name a few. We provide a methodology to account for complex kinetic bimolecular reactions in a fully Lagrangian framework where each particle represents a fraction of the total mass of a specific solute. The method, built as an extension to the second-order case, is based on the concept of optimal Kernel Density Estimator, which allows the concentrations to be written in terms of particle locations, hence transferring the concept of reaction rate to that of particle location distribution. By doing so, we can update the probability of particles reacting without the need to fully reconstruct the concentration maps. The performance and convergence of the method is tested for several illustrative examples that simulate the Advection-Dispersion-Reaction Equation in a 1D homogeneous column. Finally, a 2D example of application is presented evaluating the need of fully describing non-linear chemical kinetics in a randomly heterogeneous porous medium.

Index terms: Groundwater Transport (1832), Computational Hydrology (1805), Stochastic Hydrology (1869), Modeling (1847), Geochemical Modeling (1009)

Keywords: Reactive transport, complex kinetics, porous media, random walk, particle tracking, Kernel Density Estimators

1. Introduction

Random Walk Particle Tracking Methods (RWPTMs) offer a convenient Lagrangian numerical approach to simulate solute transport in porous media. RWPTMs have been demonstrated to be particularly efficient in dealing with aquifer heterogeneities and non-reactive transport involving a large variety of complex processes such as non-Fickian transport and multiple porosity systems [*Wen and Gómez-Hernández*, 1996; *LaBolle et al.*, 1996; *Sanchez-Vila and Solis-Delfin*, 1999; *Salamon et al.*, 2006a, 2006b; *Riva et al.*, 2008; *Delay and Bodin*, 2001; *Cvetkovic and Haggerty*, 2002; *Berkowitz et al.*, 2006; *Zhang and Benson*, 2008; *Dentz and Castro*, 2009; *Benson and Meerschaert*, 2009; *Tsang and Tsang*, 2001; *Huang et al.*, 2003; *Willmann et al.*, 2013; *Henri and Fernàndez-Garcia*, 2014, 2015]. This family of methods essentially consist of discretizing the solute mass (existing initially or injected through the boundaries with time) into a finite number of particles, each representing a fraction of the total mass, and then moving such particles according to simple relationships that represent the transport mechanisms considered (e.g., advection, dispersion or diffusion into stagnant zones). RWPTMs are mass conservative by construction, and avoid some of the inherent numerical difficulties associated with Eulerian approaches, i.e., numerical dispersion and oscillations [*Salamon et al.*, 2006a; *Benson et al.*, 2017].

However, several disadvantages have prevented the general use of RWPTMs in reactive transport problems with few limited exceptions. The main roadblock is that most chemical reactions are written in terms of concentrations (or chemical activities), which are not directly accessible at any given time, unless previously reconstructed from discrete particle information. At this stage, one needs to keep in mind that a naive reconstruction, such as the use of histograms, is an error prone process that can lead to spurious fluctuations [e.g., *Boso et al.*, 2013]. Consequently, as concentrations - and in some cases their gradients [e.g., *De Simoni et al*., 2007] - are reaction drivers, errors can propagate to reaction rates. Albeit recent works [*Fernàndez-Garcia and Sanchez-Vila*, 2011; *Pedretti and Fernàndez-Garcia*, 2013; *Schmidt et al*., 2017] have shown that the spurious fluctuations of the concentrations reconstructed from particles can be largely minimized by using a post-processing analysis based on kernels, modeling complex reactive transport problems with RWPTMs is still a challenge.

The focus of this paper is on kinetic chemical reactions. In this context, several methods have been proposed in the literature to simulate reactive transport with RWPTMs. Simple linear kinetic reactive transport problems such as first-order network reactions and slow sorption can easily be treated with transition probabilities, without having to estimate the concentrations during the course of the simulations [e.g., *Kinzelbach*, 1987; *Andricevic and Foufoula-Georgiou*, 1991; *Michalak and Kitanidis*, 2000; *Henri and Fernàndez-Garcia*, 2014, 2015]. Reconstruction here is an efficient post-processing tool with little drawbacks.

However, the incorporation of non-linear chemical reactions involving more than one chemical species into the RWPTM is remarkably cumbersome. In this case, one needs to either re-estimate solute concentrations at any given time step or to use particle proximity relationships. Both these approaches present important disadvantages, which have hindered the widespread use of RWPTMs – since the most common processes in geochemistry and biogeochemistry are complex, being non-linear, multi-species and affected by water-rock interaction. The first approach is a hybrid Lagrangian-Eulerian method by which reaction rates are determined from concentrations. Here, a compromise between CPU time and the back and forth transformation of particles to concentrations is necessary [*Tompson*, 1993; *Tompson et al*., 1996; *Cui et al*., 2014]; as aforementioned, this process is either error-prone or computationally expensive. The second approach is purely Lagrangian, and sophisticated search algorithms are needed to calculate proximity relationships [*Paster et al*., 2014]. Along this line, *Benson and Meerschaert* [2008] studied a simple bimolecular system ($\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C}$) with second-order kinetics, and found that the probability of reaction of two isolated particles depends on both thermodynamics and the probability of collocation of two particles. *Paster et al.* [2013, 2014] extended these concepts to higher dimensions, and *Ding and Benson* [2015] used this bimolecular type of reaction as a building block to simulate the Michaelis-Menten enzyme kinetic model. *Rahbaralam et al.* [2015] demonstrated that the support volume of particles in the probability of collocation can be determined by using an optimal kernel bandwidth approach. This method speeds up the algorithm and avoids incomplete mixing due to the use of a limited number of particles. A first field application of the *Benson and Meerschaert* [2008] method has been recently presented by *Ding et al.* [2017], who simulated the degradation of Carbon Tetrachloride at the Schoolcraft, MI site, under anaerobic conditions. All existing variations of this method share an important limitation: they can only reproduce second-order kinetics, with the exception of those complex reactions that can be modeled as a combination of first-order monomolecular reactions and second-order bimolecular reactions, such as the aforementioned Michaelis-Menten enzyme kinetic model.

In some other Lagrangian approaches such as SPH [e.g. *Tartakowsky and Meakin.* 2005; *Tartakovsky et al*., 2007; *Herrera et al.*, 2009, 2017] each particle represents a volume of fluid, so concentrations are directly attributed to particles and diffusion/dispersion is simulated by exchanging mass between particles. A similar approach was used by *Benson and Bolster* [2016] to propose a particle tracking method for the simulation of chemical reactions of arbitrary complexity, based on mass exchange between particles which could contain any variety of chemical compounds. *Engdahl et al.* [2017] recently generalized the capabilities of the method by coupling it to the reaction engine PhreeqcRM [*Parkhurst and Wissmeier*, 2015]. Each particle can be seen as a mobile bin containing a fixed volume of water, and reactions occur inside particles according to the particle-specific solute concentrations. Some limitations can be attributed to these kind of methods. For instance, one needs to artificially inject empty particles in places where solutes can potentially diffuse, or to add immobile particles and use very small time steps to represent linear sorption.

Most of these approaches to Lagrangian modeling of reactive transport use kernel functions to account for either dispersion or reaction between particles. Kernels have also been widely used in other fields of science like fluid mechanics [e.g., *Wu and Li*, 2007; *Yue et al.*, 2004], computer vision and image processing [e.g., *Chang and Ansari*, 2005; *Stoessel and Sagerer*, 2006; *Takeda et al.*, 2007], or 3D animation [e.g., *Ihmsen et al.*, 2011], just to name a few.

In this paper, we propose a new random walk particle tracking method capable of simulating different sorts of complex kinetic reactions occurring between two reactants (thus generalizing the existing methods to simulate second-order kinetics), while maintaining the classical interpretation of a particle (a fraction of the total mass of a given species). To simulate reactions, we determine the probability that any particle reacts based on particle interactions, the reaction rate law and the stoichiometry. The idea behind the proposed method is to equipped each particle with an optimal kernel function that defines the particle support [*Fernàndez-Garcia and Sanchez-Vila*, 2011; *Rahbaralam et al.*, 2015] from the beginning of the simulation. For convenience, complex reaction rates are expressed as the product of a second-order bimolecular reaction and a compensation function (g) that depends on the reactant concentrations. An approximate solution of the probability of reaction is then determined, providing a fully Lagrangian approach that does not entail any kind of spatial discretization. The probability of reaction is demonstrated to depend on the particle interaction, expressed as the volume integral of the product between particle kernel functions, and on the point-value of g at a weighted mid-position between the two particles.

We then show four example column transport (1D) applications to illustrate the performance and the convergence of the method as a function of the initial number of particles for different chemical systems. To achieve this, the random walk particle tracking solution is compared with a highly-discretized finite difference solution that is assumed to represent the exact solution. The four examples represent a wide sample of the most common problems in biogeochemistry: two examples of non-linear aqueous reactions and two examples of non-linear reactions considering the water-rock interaction. Finally, a 2D example of application is presented evaluating the need of fully describing non-linear chemical kinetics in a randomly heterogeneous porous medium.

Although the example applications are 1D or 2D reactive transport problems in stationary flow, the proposed method has no limitations regarding the number of spatial dimensions or the effect of variable velocity with time (full 4D).

2. Second-order kinetic reactions

In order to lay the groundwork for the implementation of arbitrarily complex kinetic reactions, we start by reviewing some concepts and then reformulating the mathematical expressions corresponding to second-order bimolecular reactions. Let us consider a simple bimolecular irreversible reaction $\alpha \mathrm{A} + \beta \mathrm{B} \rightarrow \gamma \mathrm{C}$ with a reaction rate proportional to the concentration of both reactants,

$$r(\mathbf{x}, t) = k_f \, c_{\mathrm{A}}(\mathbf{x}, t) \, c_{\mathrm{B}}(\mathbf{x}, t), \#(1)$$

where c_s is the concentration of the *s*th-species $\{s = \mathrm{A}, \mathrm{B}, \mathrm{C}\}$, k_f is the forward reaction coefficient, $\{\alpha, \beta, \gamma\}$ are the stoichiometric coefficients, and $r(\mathbf{x}, t)$ is the reaction rate at the $\mathbf{x}$ location and time t, defined as:

$$r(\mathbf{x}, t) = \frac{1}{\gamma}\frac{dc_{\mathrm{C}}}{dt} = -\frac{1}{\alpha}\frac{dc_{\mathrm{A}}}{dt} = -\frac{1}{\beta}\frac{dc_{\mathrm{B}}}{dt} \#(2)$$

We refer to chemical reactions that follow equation (1) as second-order kinetic reactions, also implying that the reaction is of first-order with respect to each reactant.

Although here we study an irreversible reaction, reversibility can be modeled as a combination of a forward reaction and a backward reaction. Further details are given at the end of section 3.

2.1. The particle pair annihilation method

Benson and Meerschaert [2008] found that this problem could be solved by simply analyzing how two isolated A and B particles react to form a C particle when $\alpha = \beta = 1$. Although the original expression was developed for a general application, here we present it incorporating explicitly the effect of porosity for the particular case of porous media. In one dimension, the probability of reaction of these two particles in a given time interval Δt is given by the expression,

$$P(\mathrm{A} \rightarrow \mathrm{C}, \Delta t) = \phi^{-1}\, k_f \Delta t\, m\, \frac{1}{\sqrt{4\pi h^2}} \exp\left(-\frac{(X_\mathrm{A} - X_\mathrm{B})^2}{4h^2}\right), \#(3)$$

which is obtained as the product of the probability that the two particles will occupy the same differential volume times the conditional probability that, upon collocation, the particles will react during the time step Δt. Equation (3) is written in terms of the particle mass m (or amount of substance, depending on how k_f is defined; thus, in this work the term *particle mass* is used in a general sense). Here, the mass of all particles is assumed equal to $m = \frac{\Omega\, \phi[\mathrm{A}]_0}{N_0}$, where Ω is the initial volume occupied by the injected particles, ϕ is porosity, $[\mathrm{A}]_0$ is the initial concentration of species A, and N_0 the number of A particles injected. Finally, $h = \sqrt{2D\Delta t}$ is the length of influence of one particle defined only in terms of local diffusion and/or dispersion.

Once the probability of reaction of two particles is calculated, chemical reactions in the random walk method can be incorporated by particle annihilation, i.e., when two particles react, they disappear. This means that the number of particles of the reactant species decreases as the simulation progresses, and numerical resolution problems may arise at low concentrations. This limitation was addressed by *Bolster et al*. [2016], who showed that a change in the particle mass is also a valid alternative to particle annihilation.

There is another strong limitation in the particle pair annihilation method. Chemical reactions depend on the activities of the reactants rather than on their concentrations. Thus, the aforementioned approach cannot reproduce second-order reactions correctly unless the ionic strength is not affected by the reaction or its effect on the activity coefficients is negligible. This is particularly relevant when modelling reactions that have an important impact on the ionic strength of the solution.

2.2. The optimal kernel approach

2.2.1 Representation of a particle

The RWPTM satisfies the transport equation in the limit when the number of particles approaches infinity. Considering that each *i*th particle associated with species *s* at time *t* is located at a point $\mathbf{X}_s^i$, and that no size is attributed to it, its spatial distribution can be expressed as a Dirac delta distribution and then the concentration of a given species can be written formally as,

$$c_s(\mathbf{x}, t) = \frac{1}{\phi(\mathbf{x})} \sum_{i=1}^{n_s} m_s^i \, E\left\{\delta\left(\mathbf{x} - \mathbf{X}_s^i(t)\right)\right\}, \#(4)$$

where m_s^i is the mass of the *i*th particle of species *s*, $\phi(\mathbf{x})$ is the location dependent porosity, and $E\{\cdot\}$ is the expectation operator over all particle realizations. The expectation of the Dirac delta function is the probability density function (pdf) of the particle position, $p_s^i(\mathbf{x}; t)$. In practice, simulations cannot use an infinite number of particles and the inference of $p_s^i(\mathbf{x};t)$ becomes the Achilles heel of all random walk methods. Typically, the concentration field is estimated by averaging the mass over a fixed support volume $V(\mathbf{x})$ centered at the $\mathbf{x}$ location. This can be achieved by counting the mass of particles in fixed bins or by projection functions [*Tompson and Gelhar*, 1990; *Tompson et al.*, 1996]. However, these methods suffer from the same problems as those associated with the estimation of pdfs through histograms, i.e., results depend on the discretization of the domain or the bin size.

An alternative approach was introduced by *Fernàndez-Garcia and Sanchez-Vila* [2011]. The method recognizes the uncertainty associated with subsampling an infinite number of particles by equipping each particle with a pdf (the kernel function). The estimation of concentrations can then be written as a direct extension of (4),

$$c_s(\mathbf{x},t) = \frac{1}{\phi(\mathbf{x})}\sum_{i=1}^{n_s} m_s^i \, W\left(\mathbf{x} - \mathbf{X}_s^i \,; \mathbf{H}_s\right), \#(5)$$

where $\mathbf{H}_s$ is the kernel bandwidth matrix associated to species s and $W(\mathbf{u}\,;\mathbf{H})$ is the scaled kernel function, for which several shapes have been suggested, the most common one being the Gaussian kernel function,

$$W(\mathbf{u};\mathbf{H}) = (2\pi)^{-\frac{d}{2}}|\mathbf{H}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\,\mathbf{u}^T\mathbf{H}^{-1}\mathbf{u}\right), \#(6)$$

where d is the space dimension. In the Gaussian kernel (6), the bandwidth matrix is the covariance matrix. Expression (5) is valid for an infinite domain or away from the domain boundaries. The particular treatment of boundaries is discussed in the subsequent sections. Note that the concentration of a given species at any given $\mathbf{x}$ location does not depend only on the subset of particles falling into an arbitrary bin, but on all existing particles associated with that species. Assuming that $\mathbf{H}_s = h_s^2\,\mathbf{I}_d$ (we will refer to this case later as the isotropic kernel) the optimal bandwidth h_s associated with a given species s (also denoted as particle support) can be determined based on the amount of particles n_s and their distribution in space, by minimizing the Asymptotical Mean Integrated Squared Error (A-$MISE$). This is a well-known procedure in statistics [e.g., *Silverman*, 1986; *Härdle*, 1991]. For a second-order kernel,

$$h_s = \left(\frac{d\,R(W)}{R(\nabla^2 p_s)\mu_2^2(W) n_s}\right)^{\frac{1}{d+4}}, \#(7)$$

where R is the L_2 norm of a function, μ_2 is the second moment, and p_s is the normalized concentration,

$$p_s(\mathbf{x}, t) = \frac{c_s(\mathbf{x}, t)}{\int_{\Omega^d} c_s(\mathbf{x}, t)\, d\mathbf{x}}, \#(8)$$

where Ω^d is the d-dimensional domain of the model. Note that, in this setup, the estimation of c_s is not explicit, i.e. the estimator (7) depends circularly on the estimation (5). Hence, one needs to either use an iterative method or make an assumption on the approximate shape of the particle plume. The former approach can be computationally intensive, whereas the latter can lead to a suboptimal bandwidth choice, hindering the convergence rate of the estimation with respect to the number of particles. We refer to *Engel et al*. [1994] for details on the calculation of h_s. Since p_s in RWPTMs changes over time, the kernel bandwidth matrix $\mathbf{H}_s$ is a time-dependent variable that not only accounts for local diffusion and/or dispersion but also for the spreading and stretching of each particle plume. This approach has been used in subsurface hydrology to reconstruct key variables associated with a wide variety of problems, e.g., reaction rates and mixing measures [*Fernàndez-Garcia and Sanchez-Vila*, 2011], power-law tailing in breakthrough curves [*Pedretti and Fernàndez-Garcia*, 2013], and human health risk estimates [*Siirila-Woodburn et al*., 2015].

2.2.2. The probability of reaction of a particle

This section derives the probability of reaction of a given particle for a second order reaction with arbitrary stoichiometric coefficients. For the derivation, we assume that the problem domain Ω^d is infinite, so expression (5) is valid at any location. At the end of section 3 it is discussed how the methodology can be adapted to simulate reactions near the boundaries of a finite domain. The chemical reaction is still represented by $\alpha\mathrm{A} + \beta\mathrm{B} \rightarrow \gamma\mathrm{C}$ and the reaction rate follows equation (1). The probability that a particle reacts in the time interval $[t, t + \Delta t]$ can be simply expressed as mass consumed per unit of mass,

$$P\left(\mathrm{A}^i \rightarrow \mathrm{C}^k, \Delta t\right) = -\frac{\Delta m_\mathrm{A}^i}{m_\mathrm{A}^i}, \#(9)$$

$$P\left(\mathrm{B}^j \rightarrow \mathrm{C}^k, \Delta t\right) = -\frac{\Delta m_\mathrm{B}^j}{m_\mathrm{B}^j}. \#(10)$$

Here, A^i refers to the *i*th-particle associated with species A, $P\left(\mathrm{A}^i \rightarrow \mathrm{C}^k, \Delta t\right)$ is the probability that A^i is transformed into a new particle C^k in the time interval Δt, and Δm_A^i is the increment of mass of the particle A^i due to the chemical reaction. This relationship was used by *Salamon et al*. [2007] and *Henri and Fernàndez-Garcia* [2014, 2015] to develop particle transition probabilities for modeling solute transport with multi-rate mass transfer and network reactions. From the definition of reaction rate given in (2), expressions (9) and (10) can be rewritten as:

$$P\left(\mathrm{A}^i \to \mathrm{C}^k, \Delta t\right) = \frac{\alpha}{m_{\mathrm{A}}^i} \int_t^{t+\Delta t} \int_{\Omega^d} \phi \, r_{\mathrm{A}}^i(\mathbf{x}, t') \; d\mathbf{x} \, dt' \approx \frac{\alpha}{m_{\mathrm{A}}^i} \, \Delta t \int_{\Omega^d} \phi \, r_{\mathrm{A}}^i(\mathbf{x}, t) \; d\mathbf{x} \, , \#(11)$$

$$P\left(\mathrm{B}^j \to \mathrm{C}^k, \Delta t\right) = \frac{\beta}{m_{\mathrm{B}}^j} \int_t^{t+\Delta t} \int_{\Omega^d} \phi \, r_{\mathrm{B}}^j(\mathbf{x}, t') \; d\mathbf{x} \, dt' \approx \frac{\beta}{m_{\mathrm{B}}^j} \, \Delta t \int_{\Omega^d} \phi \, r_{\mathrm{B}}^j(\mathbf{x}, t) \; d\mathbf{x} \, , \#(1$$

where $r_{\mathrm{A}}^i(\mathbf{x}, t)$ and $r_{\mathrm{B}}^j(\mathbf{x}, t)$ are particle reaction rates. The products $\alpha \, r_{\mathrm{A}}^i(\mathbf{x}, t)$ and $\beta \, r_{\mathrm{B}}^j(\mathbf{x}, t)$ define the amount of particle mass consumed per unit volume of liquid in a unit of time. The particle reaction rates can be derived as it follows. Substituting (5) into (1), it is possible to find an expression of the total chemical reaction rate as a function of particle kernel distributions,

$$r(\mathbf{x}, t) = \frac{k_f}{\phi^2} \sum_{i=1}^{n_{\mathrm{A}}} \sum_{j=1}^{n_{\mathrm{B}}} m_{\mathrm{A}}^i m_{\mathrm{B}}^j \; W\left(\mathbf{x} - \mathbf{X}_{\mathrm{A}}^i; \mathbf{H}_{\mathrm{A}}\right) W\left(\mathbf{x} - \mathbf{X}_{\mathrm{B}}^j; \mathbf{H}_{\mathrm{B}}\right) . \#(13)$$

The reaction rate of any particle A^i or B^j is determined, respectively, from the interaction of A^i with all existing B-particles and the interaction of B^j with all existing A-particles. Thus, the total reaction rate can be decomposed as

$$r(\mathbf{x}, t) = \sum_{i=1}^{n_{\mathrm{A}}} r_{\mathrm{A}}^i(\mathbf{x}, t) = \sum_{j=1}^{n_{\mathrm{B}}} r_{\mathrm{B}}^j(\mathbf{x}, t) \, , \#(14)$$

where

$$r_{\mathrm{A}}^i(\mathbf{x}, t) = \frac{k_f}{\phi^2} m_{\mathrm{A}}^i \sum_{j=1}^{n_{\mathrm{B}}} m_{\mathrm{B}}^j \; W\left(\mathbf{x} - \mathbf{X}_{\mathrm{A}}^i; \mathbf{H}_{\mathrm{A}}\right) W\left(\mathbf{x} - \mathbf{X}_{\mathrm{B}}^j; \mathbf{H}_{\mathrm{B}}\right), \#(15)$$

$$r_{\mathrm{B}}^j(\mathbf{x}, t) = \frac{k_f}{\phi^2} m_{\mathrm{B}}^j \sum_{i=1}^{n_A} m_{\mathrm{A}}^i \; W\left(\mathbf{x} - \mathbf{X}_{\mathrm{A}}^i; \mathbf{H}_{\mathrm{A}}\right) W\left(\mathbf{x} - \mathbf{X}_{\mathrm{B}}^j; \mathbf{H}_{\mathrm{B}}\right). \#(16)$$

Each term in the summation represents the interaction between two individual particles A^i and B^j. In the particular case of a Gaussian kernel function, the kernel product can be rewritten as

$$W\left(\mathbf{x}-\mathbf{X}_{\mathrm{A}}^{i};\mathbf{H}_{\mathrm{A}}\right)W\left(\mathbf{x}-\mathbf{X}_{\mathrm{B}}^{j};\mathbf{H}_{\mathrm{B}}\right)=W\left(\mathbf{x}-\mathbf{X}_{\mathrm{AB}}^{ij};\mathbf{H}_{\mathrm{AB}}\right)W\left(\mathbf{X}_{\mathrm{A}}^{i}-\mathbf{X}_{\mathrm{B}}^{j};\mathbf{H}_{\mathrm{A}}+\mathbf{H}_{\mathrm{B}}\right),\#(17)$$

where

$$\mathbf{H}_{\mathrm{AB}}=(\mathbf{H}_{\mathrm{A}}^{-1}+\mathbf{H}_{\mathrm{B}}^{-1})^{-1},\#(18)$$

$$\mathbf{X}_{\mathrm{AB}}^{ij}=\mathbf{H}_{\mathrm{AB}}\left(\mathbf{H}_{\mathrm{A}}^{-1}\mathbf{X}_{\mathrm{A}}^{i}\ +\mathbf{H}_{\mathrm{B}}^{-1}\mathbf{X}_{\mathrm{B}}^{j}\right),\#(19)$$

which means that the product of two Gaussian kernel density functions associated with particles A^i and B^j is proportional to another Gaussian kernel function centered at $\mathbf{X}_{\mathrm{AB}}^{ij}$ with a covariance matrix $\mathbf{H}_{\mathrm{AB}}$. Figure 1 illustrates this equivalence in one dimension. This indicates that the reaction between two individual particles is occurring mostly around $\mathbf{X}_{\mathrm{AB}}^{ij}$. The second kernel function on the right hand side of (17) is a constant scaling factor that only depends on the separation between particles.

In the case where $\mathbf{H}_{\mathrm{A}}$ and $\mathbf{H}_{\mathrm{B}}$ are isotropic ($\mathbf{H}_s=h_s^2\mathbf{I}_d$), then it derives from (18) that $\mathbf{H}_{\mathrm{AB}}$ is also isotropic ($\mathbf{H}_{\mathrm{AB}}=h_{\mathrm{AB}}^2\mathbf{I}_d$) and

$$h_{\mathrm{AB}}=\sqrt{\frac{h_{\mathrm{A}}^2h_{\mathrm{B}}^2}{h_{\mathrm{A}}^2+h_{\mathrm{B}}^2}}\#(20)$$

is proportional to the harmonic mean of the squares of $h_{\mathrm{A}},h_{\mathrm{B}}$. As aforementioned, $\mathbf{X}_{\mathrm{AB}}^{ij}$ is the position with maximum probability density of collocation of particles A^i and B^j; in the isotropic case, expression (19) can be rewritten so that $\mathbf{X}_{\mathrm{AB}}^{ij}$ is simply the mid-position of the particle pair weighted by their corresponding squared particle support, i.e.,

$$\mathbf{X}_{\mathrm{AB}}^{ij} = \frac{\mathbf{X}_{\mathrm{A}}^{i}\, h_{\mathrm{B}}^{2} + \mathbf{X}_{\mathrm{B}}^{j} h_{\mathrm{A}}^{2}}{h_{\mathrm{A}}^{2} + h_{\mathrm{B}}^{2}}. \#(21)$$

In order to integrate expressions (11) and (12), we assume a locally constant porosity over the kernel product support centered at $\mathbf{X}_{\mathrm{AB}}^{ij}$ and represented by $\mathbf{H}_{\mathrm{AB}}$. By substituting (15) and (16) into (11) and (12) respectively and integrating, we finally obtain that

$$P\left(\mathrm{A}^{i} \to \mathrm{C}^{k}, \Delta t\right) = \frac{\alpha\, k_f}{\phi(\mathbf{X}_{\mathrm{AB}}^{ij})} \Delta t \sum_{j=1}^{n_{\mathrm{B}}} m_{\mathrm{B}}^{j}\, W\left(\mathbf{X}_{\mathrm{A}}^{i} - \mathbf{X}_{\mathrm{B}}^{j}; \mathbf{H}_{\mathrm{A}} + \mathbf{H}_{\mathrm{B}}\right), \#(22)$$

$$P\left(\mathrm{B}^{j} \to \mathrm{C}^{k}, \Delta t\right) = \frac{\beta\, k_f}{\phi(\mathbf{X}_{\mathrm{AB}}^{ij})} \Delta t \sum_{i=1}^{n_{\mathrm{A}}} m_{\mathrm{A}}^{i}\, W\left(\mathbf{X}_{\mathrm{A}}^{i} - \mathbf{X}_{\mathrm{B}}^{j}; \mathbf{H}_{\mathrm{A}} + \mathbf{H}_{\mathrm{B}}\right). \#(23)$$

In the particular one-dimensional case where only one particle of each reactant is present, porosity ϕ is constant in space, $\alpha = \beta = 1$, $H_{\mathrm{A}} = H_{\mathrm{B}} = h^2$, and all particles share the same mass *m*, we have

$$P(\mathrm{A} \to \mathrm{C}, \Delta t) = P(\mathrm{B} \to \mathrm{C}, \Delta t) = \frac{k_f}{\phi} \Delta t\, m\, \frac{1}{\sqrt{4\pi h^2}} \exp\left(-\frac{(x_{\mathrm{A}} - x_{\mathrm{B}})^2}{4h^2}\right), \#(24)$$

and we directly recover the probability of reaction between two isolated particles obtained by *Benson and Meerschaert* [2008]. We note that h in (24) is not $h = \sqrt{2D\Delta t}$ but rather it is defined as an optimal kernel support that changes with time according to the number of particles remaining and the actual shape of the solute plume. We claim that this difference in the definition of h is very significant. *Benson and Meerschaert* [2008] simulate incomplete mixing by using a low number of uniform- randomly distributed particles, which limits the reaction rate after some time as the A-particles become isolated from the B-particles (described by the authors as "islands of particles"). Along the same line, *Paster et al.* [2013, 2014] derive a relationship between the initial particle density and the noise of the initial condition, suggesting that the simulation of smoother initial conditions requires a higher number of particles. In contrast, *Rahbalaram et al.* [2015] show that using the adaptive kernel makes it possible to highly reduce the dependence of the numerical solution on the number of particles. Another important difference between the two approaches becomes evident when more than one particle of each reactant is present. In this case, the probability of reaction of a particle given by (22) or (23) can be seen as the sum of independent particle pair interactions. This is only satisfied by the particle pair annihilation method in the limit when $\Delta t \rightarrow 0$. Otherwise, the reaction between two particles is not a disjoint event. Section 4 provides the details of the new particle tracking algorithm.

3. Extension to kinetic reactions with arbitrary reaction rate laws

The challenge in extending second-order reactions to arbitrary reaction rate laws resides in that now the total reaction rate cannot be simply split into combinations of kernel functions between particle pairs. Consequently, the rate at which two particles react depends also on all other surrounding particles. In this case, without any loss of generality, it is convenient to represent the total reaction rate as the product of a second-order reaction times g, a function of any arbitrary shape involving the reactants' concentrations, and denoted as compensation function,

$$r(\mathbf{x},t) = k_f c_{\mathrm{A}}(\mathbf{x},t) c_{\mathrm{B}}(\mathbf{x},t)\ g\big(c_{\mathrm{A}}(\mathbf{x},t), c_{\mathrm{B}}(\mathbf{x},t)\big). \#(25)$$

Applying $g = 1$ implies recovering (1). Substituting (5) into (25), and then decomposing as in (14) and substituting into (11) and (12), we now have,

$$P\left(\mathrm{A}^{i} \rightarrow \mathrm{C}^{k}, \Delta t\right) =$$

$$\frac{\alpha\, k_f}{\phi(\mathbf{X}_{\mathrm{AB}}^{ij})} \Delta t \sum_{j=1}^{n_{\mathrm{B}}} m_{\mathrm{B}}^{j}\, W\left(\mathbf{X}_{\mathrm{A}}^{i} - \mathbf{X}_{\mathrm{B}}^{j}; \mathbf{H}_{\mathrm{A}} + \mathbf{H}_{\mathrm{B}}\right) \int_{\Omega^d} W\left(\mathbf{x} - \mathbf{X}_{\mathrm{AB}}^{ij}; \mathbf{H}_{\mathrm{AB}}\right) g\left(c_{\mathrm{A}}(\mathbf{x}, t), c_{\mathrm{B}}(\mathbf{x}, t)\right) d\mathbf{x}, \#(26)$$

$$P\left(\mathrm{B}^{j} \rightarrow \mathrm{C}^{k}, \Delta t\right) =$$

$$\frac{\beta\, k_f}{\phi(\mathbf{X}_{\mathrm{AB}}^{ij})} \Delta t \sum_{j=1}^{n_{\mathrm{A}}} m_{\mathrm{A}}^{j}\, W\left(\mathbf{X}_{\mathrm{A}}^{i} - \mathbf{X}_{\mathrm{B}}^{j}; \mathbf{H}_{\mathrm{A}} + \mathbf{H}_{\mathrm{B}}\right) \int_{\Omega^d} W\left(\mathbf{x} - \mathbf{X}_{\mathrm{AB}}^{ij}; \mathbf{H}_{\mathrm{AB}}\right) g\left(c_{\mathrm{A}}(\mathbf{x}, t), c_{\mathrm{B}}(\mathbf{x}, t)\right) d\mathbf{x}, \#(27)$$

Because the compensation function $g(\mathbf{x}, t)$ depends on $\mathbf{x}$ in a complex manner, the integration of (26) and (27) is no longer direct. To overcome this problem, we approximate this integral by localizing the function $g(\mathbf{x}, t)$ about the point $\mathbf{x} = \mathbf{X}_{\mathrm{AB}}^{ij}$, i.e., at the centroid of the kernel product (see figure 1), using a truncated first-order Taylor series expansion (i.e., linearizing it in terms of location),

$$g(\mathbf{x}, t) \cong g\left(\mathbf{X}_{\mathrm{AB}}^{ij}, t\right) + \nabla g\left(\mathbf{X}_{\mathrm{AB}}^{ij}, t\right)^{T} \left(\mathbf{x} - \mathbf{X}_{\mathrm{AB}}^{ij}\right) \#(28)$$

The validity of this approximation is subjected to the significance of higher order terms of g over the kernel product domain represented by $\mathbf{H}_{\mathrm{AB}}$. Note that the truncation error will always converge towards zero with an increasing number of particles, namely, as $\mathbf{H}_{\mathrm{AB}}$ approaches the Dirac delta. Introducing (28) into (26) and (27), and given that the first moment of the kernel about its centroid equals zero, we obtain

$$P\left(\mathrm{A}^{i} \rightarrow \mathrm{C}^{k}, \Delta t\right) =$$

$$\frac{\alpha\, k_f}{\phi(\mathbf{X}_{AB}^{ij})} \Delta t \sum_{j=1}^{n_B} m_{\mathrm{B}}^{j}\, W\left(\mathbf{X}_{A}^{i} - \mathbf{X}_{B}^{j}; \mathbf{H}_{A} + \mathbf{H}_{B}\right) g\left(c_A\left(\mathbf{X}_{AB}^{ij}, t\right), c_B\left(\mathbf{X}_{AB}^{ij}, t\right)\right), \#(29)$$

$$P\left(B^j \rightarrow C^k, \Delta t\right) = \frac{\beta\ k_f}{\phi(\mathbf{X}_{AB}^{ij})} \Delta t \sum_{i=1}^{n_A} m_A^i\, W\left(\mathbf{X}_A^i - \mathbf{X}_B^j; \mathbf{H}_A + \mathbf{H}_B\right) g\left(c_A\left(\mathbf{X}_{AB}^{ij}, t\right), c_B\left(\mathbf{X}_{AB}^{ij}, t\right)\right). \#(30)$$

The evaluation of $g\left(c_{\mathrm{A}}\left(\mathbf{X}_{\mathrm{AB}}^{ij}, t\right), c_{\mathrm{B}}\left(\mathbf{X}_{\mathrm{AB}}^{ij}, t\right)\right)$ in (29) and (30) requires an approximate solution of the concentrations of the species A and B at the specific location $\mathbf{X}_{\mathrm{AB}}^{ij}$. One possibility is to estimate these concentrations directly using the kernel estimator given in (5). However, this would require an excessive amount of calculations. To minimize CPU time, here we estimated these concentrations as a linear interpolation of the concentrations obtained only at the particle positions, estimated a priori by (5). This is possible as long as $\mathbf{X}_{\mathrm{A}}$, $\mathbf{X}_{\mathrm{AB}}$ and $\mathbf{X}_{\mathrm{B}}$ are aligned, i.e., $\mathbf{H}_{\mathrm{A}}$ and $\mathbf{H}_{\mathrm{B}}$ are isotropic, a condition that is inherently true in one dimension. This approach constitutes a simplification, and therefore it introduces some error in the solution. In the subsequent sections, we show that this error is small for a relatively low number of particles injected.

In the case where the reaction is reversible, it can be solved by combination of a forward and a backward reaction probability [*Benson and Meerschaert*, 2008]. For example, if the backward reaction is a first-order decay, i.e.,

$$\frac{1}{\gamma} \frac{dc_{\mathrm{C}}}{dt} = k_f\ c_{\mathrm{A}}(\mathbf{x}, t)\ c_{\mathrm{B}}(\mathbf{x}, t) - k_b c_{\mathrm{C}}(\mathbf{x}, t), \#(31)$$

where k_b is the backward reaction coefficient, then the probability of backward reaction is simply,

$$P\left(\mathrm{C}^k \rightarrow \mathrm{A}^i + B^j, \Delta t\right) = \gamma\ k_b\ \Delta t, \#(32)$$

and the mass of the disappearing particle C^k has to be distributed between the generated particles A^i and B^j in proportion to their stoichiometric coefficients. This, just like the separate treatment of transport and reaction described in the following section, constitutes a split operator approach, which implies that the time step Δt should not be too large in order to avoid error and instabilities.

Expressions (29) and (30) were derived under the assumption that particles are not at close distance from the domain boundaries. Should this condition not be fulfilled, different methods exist in the literature to make KDE valid near domain boundaries. A simple one, in principle only valid for regular boundaries, is the reflection method [*Silverman*, 1986]: for every particle that is at close distance from a boundary (beneath some significance threshold) an identical virtual particle is placed as a reflection on the other side of that boundary. This complies with mass conservation inside the domain ($\int_{\Omega^d} c_s(\mathbf{x}, t)\, d\mathbf{x} = \sum_{i=1}^{n_s} m_s^i$), and also imposes a zero-gradient boundary condition. Then, the methodology that we describe in this paper can be used as long as the virtual particles are considered in the computation of the right hand side of (29) and/or (30).

4. The random walk algorithm

In the proposed method, reactive transport is solved in two stages, one corresponding to the chemical reactions, and another one to the standard advection-dispersion equation. This split operator approach is known in the literature as RT [*Simpson and Landman*, 2007]. Of course, other split operator approaches could also be implemented in a similar way. *Morshed and Kaluarachchi* [1995] show that operator splitting in non-linear reactive transport can have significant restrictions on the time step size to obtain accurate solutions. *Simpson and Landman* [2007] show that the error associated to operator splitting can be removed by using an alternating scheme provided Δt is sufficiently small. *Paster et al.* [2014] derive some practical criteria for the selection of the time step in a Lagrangian model of reactive transport with second order kinetics. In this work, the time step was simply chosen small enough in each example to reach convergence of the solution. Alternatively, an adapted time step can be estimated by fixing the maximum probability of reaction. This way, the time step is respectively small or large at stages where the reaction is fast or slow.

The procedure used in this work to simulate kinetic reactions based on the probabilities determined by (29) can be written as it follows: First, for each time step Δt, the probability of reaction of only one of the reactants (A or B) is estimated. For simplicity, and without any loss of generality, we will assume it to be the reactant A. Then, a uniform [0, 1] random number μ is drawn for each A-particle and compared to the corresponding probability of reaction, $P\left(\mathrm{A}^i \rightarrow \mathrm{C}^k, \Delta t\right)$. If $\mu \leq P\left(\mathrm{A}^i \rightarrow \mathrm{C}^k, \Delta t\right)$, it is considered that particle A^i does not react and the algorithm continues with the next A-particle. On the contrary, if $\mu > P\left(\mathrm{A}^i \rightarrow \mathrm{C}^k, \Delta t\right)$, the A-particle reacts with a number of nearby B-particles (the closest ones). To satisfy stoichiometry, the number of B-particles reacting with the A-particle, denoted here as n_r, is a positive integer value that should fulfill the following expression,

$$\alpha \sum_{j=1}^{n_r} m_{\mathrm{B}}^j = \beta\, m_{\mathrm{A}}^i. \#(33)$$

When the reaction occurs, one C-particle is injected at each $\mathbf{X}_{\mathrm{AB}}^{ij}$ position located between the reacting particle pairs $\{\mathrm{A}^i, \mathrm{B}^j\}$. These reacting particle pairs disappear after that. Again, by stoichiometry, the mass associated with each new C-particle should fulfill that

$$\sum_{k=1}^{n_r} m_{\mathrm{C}}^k = \frac{\gamma}{\alpha + \beta}\left(m_{\mathrm{A}}^i + \sum_{j=1}^{n_r} m_{\mathrm{B}}^j\right). \#(34)$$

If all particles associated with a given species share a constant mass, these expressions reduce to the following simple relationships,

$$\frac{m_{\mathrm{A}}}{m_{\mathrm{B}}} = \frac{\alpha}{\beta}\, n_r, \#(35)$$

$$m_{\mathrm{C}} = \frac{\gamma}{\beta}\, m_{\mathrm{B}}. \#(36)$$

To satisfy (35) when n_r is a real value, this expression simply requires to slightly modify the particle mass associated with the reactants prior to the beginning of the simulation. In the case of an instantaneous injection or to reproduce an initial condition, this will imply choosing an adequate ratio between the number of particles of each reactant. A valid alternative, not implemented in this work although perfectly compatible with the presented method, is to change the particle mass upon reaction [*Bolster et al.*, 2016], using (9) and (10) to determine the particle mass variation from the computed probability. Another alternative is to decide the reaction of particle pairs $\{\mathrm{A}, \mathrm{B}\}$ based on Bernoulli trials with a probability of failure determined by $f = n_r - F(n_r)$. Here, $F(x)$ is the floor function defined as the greatest integer less than or equal to x. However, in this case, stoichiometry is only fulfilled in an average sense. The latter approach is used in Example 1.

After this, following the standard random walk method, each particle is moved according to a drift term and a Brownian motion to respectively simulate advection and dispersion,

$$\mathbf{X}_s^i(t+\Delta t) = \mathbf{X}_s^i(t) + \mathbf{v}_s\left(\mathbf{X}_s^i(t)\right)\,\Delta t + \mathbf{E}_s\left(\mathbf{X}_s^i(t)\right)\boldsymbol{\xi}\sqrt{\Delta t}, \#(37)$$

where $\mathbf{X}_s^i(t)$ is the ith particle position associated with species s, $\mathbf{v}_s$ is the particle velocity associated with species s given by $\mathbf{v}_s = \frac{\mathbf{q}}{\phi R_s} + \frac{1}{\phi R_s}\nabla\cdot(\phi\mathbf{D}_s)$, R_s is the retardation factor associated with species s, $\mathbf{D}_s$ is the local hydrodynamic dispersion tensor associated with species s, $\mathbf{E}_s$ is the Brownian displacement matrix determined by $\mathbf{E}_s\mathbf{E}_s^T = 2\mathbf{D}_s/R_s$, and $\boldsymbol{\xi}$ is a vector of d standard normally distributed random numbers. Note that the method can directly support species-dependent properties such as effective particle velocity (affected by retardation) and dispersion. Note also that alternative equations to the advection-dispersion could be used in this step (e.g., Continuous Time Random Walks), as the processes of transport and reaction are fully decoupled.

5. Performance and convergence of the method

Four simple hypothetical case examples were solved using the proposed methodology to evaluate the performance and convergence of the method as a function of the number of injected particles. The selected problems illustrate a wide range of possible applications. For each problem, we simulate reactive transport in a one-dimensional column of unit (1 m^2) cross-section, with constant velocity, porosity, and dispersion, to emphasize only the relevance of the complex reactions. The parameter values adopted in each example are given in tables 1-4.

Simulations are performed in a Monte Carlo framework consisting of 100 random walk particle tracking realizations. Results are compared with those obtained from a very finely discretized finite difference solver for the ADRE with explicit time stepping and upwind differences in space for the advection term, which was checked for spatial and temporal convergence. The finite difference solution is assumed to represent the true solution. As explained in the previous section and although other approaches could be used, we assigned equal mass to all particles belonging to the same species so that stoichiometry is fulfilled exactly. Whenever possible, we imposed that the ratio of the reactant masses matches that of the stoichiometric coefficients, i.e., $n_r = 1$ in (35). The method was implemented in a Random Walk Particle Tracking code written in Matlab. At the start of each simulation, 5000 particles of each reactant were injected following Gaussian distributions in space characterized by the mean, the standard deviation and the total amount of substance indicated in tables 1-4. In all cases, the concentration of all compounds in the inflow is zero at all times.

The support of each species was estimated through (7) by assuming a Gaussian shape of the particle plume. This leads to a suboptimal approximation of the particle support volume written as [e.g. *Silverman*, 1986],

$$h_s = 1.06\,\sigma_{x,s}\,n_s^{-\frac{1}{5}}, \#(38)$$

where the index *s* denotes the chemical species, n_s is the number of particles of the *s*th species, and $\sigma_{x,s}$ is the standard deviation of the particle positions of the *s*th species at a given time.

5.1. Description of the chemical systems

Example 1. A Generic reaction with fractional exponents

In this first example, we consider a generic kinetic reaction with arbitrary stoichiometric coefficients, $\alpha\ \mathrm{A} + \beta\ \mathrm{B} \rightarrow \gamma\ \mathrm{C}$. The reaction rate is written as

$$r(x,t) = k_f\ c_{\mathrm{A}} c_{\mathrm{B}}\ g(c_{\mathrm{A}}, c_{\mathrm{B}}), \#(39)$$

where the compensation function g in this case is

$$g(c_{\mathrm{A}}, c_{\mathrm{B}}) = c_{\mathrm{A}}^{\theta_{\mathrm{A}}-1}\ c_{\mathrm{B}}^{\theta_{\mathrm{B}}-1}. \#(40)$$

Here, θ_{A} and θ_{B} are arbitrary real values, often (but not always) associated with the stoichiometric coefficients. To illustrate that any reaction with fractional exponents can be properly simulated, we chose $\theta_{\mathrm{A}} = \alpha = 2.3$ and $\theta_{\mathrm{B}} = \beta = 1.3$. The parameters adopted during the simulations are summarized in Table 1.

Example 2. Aerobic Michaelis-Menten degradation considering linear sorption of organic carbon

In this example we reproduce the aerobic biodegradation of an organic chemical compound dissolved in groundwater. The organic compound ($\mathrm{CH_2O}$) is subject to linear sorption, with a retardation factor $R = 3$. Microbial growth and decay is neglected, and the dissolved organic carbon is assumed to react with dissolved oxygen to form carbon dioxide and water, $\mathrm{CH_2O} + \mathrm{O_2} \rightarrow \mathrm{CO_2} + \mathrm{H_2O}$. The reaction rate follows the Michaelis-Menten kinetic model written here as

$$r(x,t) = k_f\ c_{\mathrm{CH_2O}}\ c_{\mathrm{O_2}}\ g\left(c_{\mathrm{CH_2O}}, c_{\mathrm{O_2}}\right), \#(41)$$

with function g being defined in this case as

$$g\left(c_{\mathrm{CH_2O}}, c_{\mathrm{O_2}}\right) = \frac{1}{k_{\mathrm{CH_2O}} + c_{\mathrm{CH_2O}}}\ \frac{1}{k_{\mathrm{O_2}} + c_{\mathrm{O_2}}}. \#(42)$$

$k_{\mathrm{CH_2O}}$ and $k_{\mathrm{O_2}}$ are the half-saturation constants associated with the dissolved organic carbon and oxygen, respectively.

The plume of oxygen rapidly migrates towards the organic chemical compound with an effective retardation of $R = 1$. We assumed that the carbon dioxide produced by the chemical reaction remains dissolved in groundwater as $\mathrm{CO_{2\,(aq)}}$. The degradation constant value and the half-saturation constant values are taken from *Rolle et al.* [2008] and *Nagy et al.* [2009]. The parameters adopted during the simulations are summarized in Table 2.

Example 3. Calcite precipitation

This example simulates the precipitation of calcium carbonate that takes place at the contact fringe of two moving solute plumes of $\mathrm{Ca^{2+}}$ and $\mathrm{CO_3^{2-}}$. Remarkably, we consider the effect of the nontrivial activity coefficients involved in the chemical reaction. We neglect the changes in the hydraulic properties of the porous medium resulting from precipitation. Back-dissolution is also omitted. The chemical reaction is formally written as $\mathrm{Ca^{2+}} + \mathrm{CO_3^{2-}} \rightarrow \mathrm{CaCO_{3(s)}}$. The rate of precipitation is represented by [e.g., *Nancollas*, 1979],

$$r(x,t) = k_{obs}(\Omega - 1), \#(43)$$

where k_{obs} is an observed or effective rate constant and Ω is the saturation state. We can rewrite this expression as:

$$r(x,t) = k' \mathrm{c}_{\mathrm{Ca^{2+}}}\, \mathrm{c}_{\mathrm{CO_3^{2-}}} \left(\gamma_{\mathrm{Ca^{2+}}}\, \gamma_{\mathrm{CO_3^{2-}}} - \frac{k_{eq}}{\mathrm{c}_{\mathrm{Ca^{2+}}}\, \mathrm{c}_{\mathrm{CO_3^{2-}}}} \right). \#(44)$$

Here, $\gamma_{Ca^{2+}}$, $\gamma_{CO_3^{2-}}$ are the activity coefficients of Ca^{2+} and CO_3^{2-}, respectively, k_{eq} is the equilibrium or solubility constant, and $k' = k_{obs}/k_{eq}$. From this, the compensation function associated with this chemical reaction is expressed as

$$g\left(\mathrm{c}_{\mathrm{Ca}^{2+}}, \mathrm{c}_{\mathrm{CO}_3^{2-}}\right) = \gamma_{Ca^{2+}}\gamma_{CO_3^{2-}} - \frac{k_{eq}}{\mathrm{c}_{\mathrm{Ca}^{2+}}\,\mathrm{c}_{\mathrm{CO}_3^{2-}}}. \#(45)$$

We assume that Ca^{2+} and CO_3^{2-} are the only ions with significant concentrations in the solution. Then, by using the extended Debye-Hückel formula, the activity coefficients $\gamma_{\mathrm{Ca}^{2+}}\gamma_{\mathrm{CO}_3^{2-}}$ are calculated as,

$$\log_{10}\left(\gamma_{\mathrm{Ca}^{2+}}\gamma_{\mathrm{CO}_3^{2-}}\right) = -4k_1\left(\frac{1}{\frac{1}{\sqrt{2\left(\mathrm{c}_{\mathrm{Ca}^{2+}} + \mathrm{c}_{\mathrm{CO}_3^{2-}}\right)}} + k_2\,\text{å}_{\mathrm{Ca}^{2+}}} + \frac{1}{\frac{1}{\sqrt{2\left(\mathrm{c}_{\mathrm{Ca}^{2+}} + \mathrm{c}_{\mathrm{CO}_3^{2-}}\right)}} + k_2\,\text{å}_{\mathrm{CO}_3^{2-}}}\right), \#(46)$$

where $k_1 = 0.018846\ \mathrm{m}^{3/2}/\mathrm{mol}^{1/2}$ and $k_2 = 0.103755\ \mathrm{m}^{3/2}/\mathrm{mol}^{1/2}\ nm$ for water at 25ºC (assuming that the density of water is $\rho_w = 1\ \mathrm{Kg/dm}^3$), and $\text{å}_{\mathrm{Ca}^{2+}}$, $\text{å}_{\mathrm{CO}_3^{2-}}$ are the hydrated radii of the respective ions [*Garrels and Christ*, 1965]. Values for k_{obs}, k_{eq} were taken from *van Breukelen* [2003] and *Appelo and Postma* [2005], respectively. The parameters adopted during the simulations are summarized in Table 3.

Example 4. Acidic dissolution of Fluorite:

This example describes the acidic dissolution of fluorite. The chemical reaction is $\mathrm{CaF_2} + 2\mathrm{H}^+ \rightarrow \mathrm{Ca}^{2+} + 2\mathrm{HF}^0$, and the dissolution rate is typically represented by [*Zhang et al*., 2006],

$$r(x,t) = k\,S_s\,(c_{\mathrm{H}^+}^2/c_{\mathrm{Ca}^{2+}})^{\alpha}, \#(47)$$

Where S_s is the mineral (CaF_2) surface per cubic meter of the porous medium, and k and α are experimental parameters. Zhang and coworkers found that at 25ºC $\log k$ ranged approximately between -2 and -4 for different experimental conditions, whereas α had values between 0.495 and 1.146. Here, we chose $\log k = -4$ and $\alpha = 0.8$, so that

$$r(x,t) = k\, S_s\, c_{\mathrm{H^+}}^{\theta_{\mathrm{H^+}}}\, c_{\mathrm{Ca^{2+}}}^{\theta_{\mathrm{Ca^{2+}}}}, \#(48)$$

where $\theta_{\mathrm{H^+}} = 1.6$ and $\theta_{\mathrm{Ca^{2+}}} = -0.8$. This kinetic model resembles that of the example 1, but with the presence of a negative exponent in the concentration of Ca^{2+}. We consider that Fluorite is everywhere in the system and in high amounts, and so S_s is a constant. Then the model has only one reactant and two products, although one of the products has an influence on the reaction rate. This means that, for this particular case, injection of the product particles is performed directly on the position of the reacting particle. We neglect the changes in the hydraulic properties of the porous medium resulting from dissolution. The chemical reaction can be embedded in (24) by defining that

$$g(\mathrm{c_{H^+}}, \mathrm{c_{Ca^{2+}}}) = c_{\mathrm{H^+}}^{\theta_{\mathrm{H^+}}-1}\, c_{\mathrm{Ca^{2+}}}^{\theta_{\mathrm{Ca^{2+}}}-1}, \#(49)$$

and $k\, S_s = k_f$. In this case, two overlapping plumes of H^+ and Ca^{2+} are injected at the same initial location (note that the reaction rate has an asymptote in case of total absence of Ca^{2+}). The parameters adopted during the simulations are summarized in Table 4.

5.2. Results

Figures 2-9 compare the random walk solution obtained at the end of the simulation time with the corresponding finite difference solution for each reactive transport problem. The random walk solution is presented in terms of the mean concentration of the different chemical species and its standard deviation (the shaded zone delimits ±1 standard deviation) obtained from 100 realizations. For completeness, the corresponding evolution of the total mass of the different chemical species remaining in the column during the simulation are also depicted in these figures. Considering that the reactive transport problems were simulated with only 5000 particles, a good match is obtained for all cases.

We note that larger deviations from the finite difference solution can be seen at the concentration peaks. This is mostly attributed to the fact that the suboptimal approximation of the particle support volume directly affects the calculation of the probabilities in (29) through the estimation of concentrations in the compensation function g. This effect is seen most significant when g deviates from zero-order (corresponding to second-order kinetic reactions, where there is no need for compensation).

The approximation (38) used to determine the particle support volume h_s is only valid for Gaussian distributions of the species' concentrations. This is particularly not satisfied for calcium ion in the acidic dissolution of Fluorite (see Figure 8). Hence, errors in the estimation of the resulting concentration map in this case example are slightly larger than in the others. In practice, the use of such an approximation of h_s (known as the rule-of-thumb in statistics), implies that more particles are needed to match the exact solution. Yet, the use of (7) may become computationally expensive in reactive transport problems otherwise.

Figure 10 shows the average relative error (ϵ_r) and the coefficient of variation (CV_r) of the increase in the total amount of substance at the end of the simulation, calculated over 100 realizations by comparison with the finite difference solution,

$$\epsilon_r = \frac{\langle M_{PT}\rangle - M_{FD}}{\Delta M_{FD}}, \#(50)$$

$$CV_r = \frac{\sqrt{\langle {M_{PT}}^2\rangle - \langle M_{PT}\rangle^2}}{|\Delta M_{FD}|}, \#(51)$$

where M_{PT} is the total mass of a given chemical compound obtained at the final simulation time, $\langle\cdot\rangle$ is the mean operator over all realizations, M_{FD} is the total mass of the chemical compound obtained with finite difference at the end of the simulation time, and ΔM_{FD} is the total mass variation of the chemical compound obtained with the finite difference method. The mean relative error ϵ_r represents the systematic error associated to one realisation, whereas CV_r accounts for its random variability. Note that the sum of the squares of these two parameters is the Mean Squared Error (MSE) of ΔM_{PT}, normalized by ${\Delta M_{FD}}^2$. Results show that the proposed random walk method converges towards the exact solution with an increasing number of particles. It also demonstrates that not many particles are needed to simulate non-linear chemical reactions with sufficient accuracy.

6. Importance of chemical kinetics in heterogeneous aquifers: An example

A two-dimensional implementation of the proposed method in a heterogeneous aquifer is given in this section. The objective of this example is to illustrate the application of the presented random walk approach in a more realistic setting. In doing this, we analyze the need of fully describing non-linear chemical kinetics in heterogeneous porous media.

We study a reactive transport problem in a 2D rectangular confined aquifer with dimensions of 100×50 m^2. The aquifer is characterized by a randomly generated log-normally distributed hydraulic conductivity field $Y = \ln K$ with a mean of $\langle Y \rangle = 3$ and a variance of $\sigma_Y^2 = 1$. The Y field follows an isotropic exponential covariance function model with integral scale of $I_Y = 5$ m. The other aquifer properties are assumed homogeneous. Groundwater flow is considered at steady-state and subject to constant head conditions at the lateral boundaries and impermeable conditions otherwise. As a result, the mean flow direction is oriented in the x-direction and characterized by a mean hydraulic gradient of 0.00622. The flow problem is solved with a finite-difference code, MODFLOW-2000 [*Harbaugh et al*., 2000], with a domain discretized into regular cells of size 0.5 m. The resulting cell face flows were used to compute the random walk particle velocities according to the hybrid interpolation method suggested by *LaBolle et al.* [1996].

The reactive transport problem is similar to the one defined in example 2 but considers a two-dimensional heterogeneous porous media. A schematic representation of the system is shown in Figure 11. A plume of dissolved organic matter, retarded with respect to groundwater by linear sorption, passes after some time through an oxygen plume. The chemical reaction follows the Michaelis-Menten kinetic model with the same formulation and parameter values as given in example 2. Table 2 shows the values of the parameters. The concentrations of the reactants are initially uniform in two separate rectangular areas depicted in Figure 11 and zero everywhere else in the domain. The concentration of all compounds in the inflow water is zero at all times.

The fast method of *Botev et al.* [2010] was used to determine the kernel bandwidth with a slight modification: the anisotropic kernel bandwidth matrix $\boldsymbol{H}$ obtained by this method was transformed into an isotropic bandwidth by matching the determinants, i.e., $h^2 = \det(\boldsymbol{H})$. As explained in section 3, the use of isotropic kernels facilitates the computation of the compensation function g at the $\mathbf{X}_{AB}^{ij}$ location by linear interpolation. However, the kernel obtained by this approach is suboptimal compared to the original method by *Botev et al.* [2010], and presumably produced a slower convergence with respect to the number of particles.

The convergence of the random walk solution was controlled by choosing a small enough time step and by performing a sensitivity analysis with respect to the number of particles. As expected, the convergence occurred for a higher number of particles compared to the 1D examples. Nevertheless, by using only 32,768 particles of each reactant, the estimated error in the total amount of product generated was below 1% as compared to the solution obtained with 131,072 particles. Figure 12 shows the three particle plumes at different moments of the simulation (for a better visualization, only a random subsample of 5,000 particles is shown), and the corresponding KDE reconstruction of the product concentrations.

The actual impact of the reaction kinetics on the problem solution depends on whether mixing or chemical kinetics is the limiting process. In order to illustrate this, we compare the evolution of the $\mathrm{CO_2}$ production with the following equivalent second-order reaction,

$$r(\mathbf{x}, t) = \frac{k_f}{k_{\mathrm{CH_2O}}\, k_{\mathrm{O_2}}}\, c_{\mathrm{CH_2O}}\, c_{\mathrm{O_2}},$$

for three different values of k_f ranging from five times smaller to five times higher than the value given in Table 2. Figure 13 shows that for a very fast reaction rate the process is mixing-limited (in this case mixing is driven by the difference in the retardation coefficients), and therefore the reaction kinetics do not have a significant effect on the results. These kind of reactions can be modeled as instantaneous (ref. xavi), as long as the mixing process is well represented by the transport model. On the other hand, in slow reactions, the reaction kinetics can make a very important difference in the results.

7. Conclusions

We have presented a new random walk particle tracking method to simulate reactions with complex kinetics involving two reactants. Reactive transport is solved in two stages: the first one corresponding to the chemical reactions, and the second one to the standard advection-dispersion equation. The method is based on the representation of particles by optimal kernel functions. This way, we derived the probability that a given particle reacts with any particle associated with other reactants. In the proposed methodology, complex kinetic reactions require linearizing a function of the local concentrations at the location of highest probability density of encounter between potentially reactive particle pairs. The implementation of the probability of reaction in random walk models has been achieved in this paper by particle annihilation, but other approaches such as particle mass variations can easily be incorporated.

In addition, two simple relationships should be satisfied to fulfill stoichiometry: one relating the mass of interacting particles with the stoichiometric coefficients, and another one relating the mass of particles produced from reactions with the stoichiometric coefficients. In practice, the first relationship requires a careful choice of the mass of the particles injected. The second relationship determines the mass of particles produced from the chemical reaction.

Several synthetic examples demonstrate the potential applicability of the method in a wide range of applications, ranging from reaction-rate laws with fractional exponents to acidic dissolution and precipitation systems with nontrivial activity coefficients. Results have shown that a good match with a finite difference solution is obtained with relatively few particles. The method has been demonstrated to converge to the solution with an increasing number of particles. This rate of convergence depends on the type of chemical reaction, i.e., on the shape of the compensation function g. Finally, a 2D example dealing with non-linear Michaelis-Menten biodegradation in a randomly heterogeneous aquifer is provided to illustrate the capabilities of the method in a more realistic setting.

Acknowledgements

Financial support was provided by the Spanish Government, through projects WE-NEED, PCIN-2015-248; ACWAPUR, PCIN-2015-239, and INDEMNE, CGL2015-69768-R (MINECO/FEDER). GS acknowledges financial support by AGAUR. The paper provides all the information needed to replicate the results. The codes and the output data from the simulations are freely available from the authors upon request.

References

Andricevic, R., and E. Foufoula-Georgiou (1991), Modeling kinetic non-equilibrium using the first two moments of the residence time distribution, *Stoch. Hydrol. Hydraul.*, *5*(2), 155–171, doi:10.1007/BF01543057.

Appelo, C. A. J., and D. Postma (2006), Geochemistry, Groundwater and Pollution, *Vadose Zo. J.*, *5*(1), 510, doi:10.2136/vzj2005.1110br.

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, *Adv. Water Resour.*, *99*, 15–37, doi:10.1016/j.advwatres.2016.11.003.

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

Benson, D. A., and M. M. Meerschaert (2009), A simple and efficient random walk solution of multi-rate mobile/immobile mass transport equations, *Adv. Water Resour.*, *32*(4), 532–539, doi:10.1016/j.advwatres.2009.01.002.

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

Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling Non-fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, *44*(2), doi:10.1029/2005RG000178.

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

Cui, Z., C. Welty, and R. M. Maxwell (2014), Modeling nitrogen transport and transformation in aquifers using a particle-tracking approach, *Comput. Geosci.*, *70*, 1–14, doi:10.1016/j.cageo.2014.05.005.

Cvetkovic, V., and R. Haggerty (2002), Transport with multiple-rate exchange in disordered media, *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, *65*(5), doi:10.1103/PhysRevE.65.051308.

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

Delay, F., and J. Bodin (2001), Time domain random walk method to simulate transport by advection-dispersion and matrix diffusion in fracture networks, *Geophys. Res. Lett.*, *28*(21), 4051–4054, doi:10.1029/2001GL013698.

Dentz, M., and A. Castro (2009), Effective transport dynamics in porous media with heterogeneous retardation properties, *Geophys. Res. Lett.*, *36*(3), doi:10.1029/2008GL036846.

Ding, D., and D. A. Benson (2015), Simulating biodegradation under mixing-limited conditions using Michaelis-Menten (Monod) kinetic expressions in a particle tracking model, *Adv. Water Resour.*, *76*, 109–119, doi:10.1016/j.advwatres.2014.12.007.

Ding, D., D. A. Benson, D. Fernàndez-Garcia, C. V. Henri, M. S. Phanikumar, and D. W. Hyndman (2017), Elimination of the reaction "scale effect": Application of the Lagrangian reactive particle-tracking method to simulate mixing-limited, field scale biodegradation at the Schoolcraft, Michigan site., *Water Resour. Res.*, Under review.

Engel, J., E. Herrmann, and T. Gasser (1994), An iterative bandwidth selector for kernel estimation of densities and their derivatives, *J. Nonparametr. Stat.*, *4*(1), 21–34, doi:10.1080/10485259408832598.

Fernàndez-Garcia, D., and X. Sanchez-Vila (2011), Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions, *J. Contam. Hydrol.*, *120–121*(C), 99–114, doi:10.1016/j.jconhyd.2010.05.001.

Garrels, R. M., and C. L. Christ (1965), *Solutions, minerals, and equilibria*, New York: Harper and Row.

Härdle, W. (1991), Kernel Density Estimation, in *Smoothing Techniques: With Implementation in S*, pp. 43–84, Springer New York, New York, NY.

Henri, C. V., and D. Fernàndez-Garcia (2014), Toward efficiency in heterogeneous multispecies reactive transport modeling: A particle-tracking solution for first-order network reactions, *Water Resour. Res.*, *50*(9), 7206–7230, doi:10.1002/2013WR014956.

Henri, C. V., and D. Fernàndez-Garcia (2015), A random walk solution for modeling solute transport with network reactions and multi-rate mass transfer in heterogeneous systems: Impact of biofilms, *Adv. Water Resour.*, *86*, 119–132, doi:10.1016/j.advwatres.2015.09.028.

Huang, H., A. E. Hassan, and B. X. Hu (2003), Monte Carlo study of conservative transport in heterogeneous dual-porosity media, in *Journal of Hydrology*, vol. 275, pp. 229–241.

Kinzelbach, W. (1988), The Random Walk Method in Pollutant Transport Simulation, in *Groundwater Flow and Quality Modelling*, edited by E. Custodio, A. Gurgui, and J. P. L. Ferreira, pp. 227–245, Springer Netherlands, Dordrecht.

LaBolle, E. M., G. E. Fogg, and A. F. B. Tompson (1996), Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods, *Water Resour. Res.*, *32*(3), 583–593, doi:10.1029/95WR03528.

Michalak, A. M., and P. K. Kitanidis (2000), Macroscopic behavior and random-walk particle tracking of kinetically sorbing solutes, *Water Resour. Res.*, *36*(8), 2133–2146, doi:10.1029/2000WR900109.

Nagy, A. M., G. Mourot, B. Marx, G. Schutz, and J. Ragot (2009), Model structure simplification of a biological reactor, *IFAC Proc. Vol.*, *42*(10), 257–262, doi:10.3182/20090706-3-FR-2004.00043.

Nancollas, G. H. (1979), The growth of crystals in solution, *Adv. Colloid Interface Sci.*, *10*(1), 215–252, doi:10.1016/0001-8686(79)87007-4.

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, *J. Comput. Phys.*, *263*, 91–112, doi:10.1016/j.jcp.2014.01.020.

Pedretti, D., and D. Fernàndez-Garcia (2013), An automatic locally-adaptive method to estimate heavily-tailed breakthrough curves from particle distributions, *Adv. Water Resour.*, *59*, 52–65, doi:10.1016/j.advwatres.2013.05.006.

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, *J. Comput. Phys.*, *303*, 95–104, doi:10.1016/j.jcp.2015.09.030.

Rahbaralam, M., D. Fernàndez-Garcia, 662 and X. Sanchez-Vila (2017), Modeling of non-linear adsorption with particle tracking and kernel density estimators, *Water Resour. Res.*, Under review.

Riva, M., A. Guadagnini, D. Fernandez-Garcia, X. Sanchez-Vila, and T. Ptak (2008), Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the Lauswiesen site, *J. Contam. Hydrol.*, *101*(1–4), 1–13, doi:10.1016/j.jconhyd.2008.07.004.

Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2007), Modeling tracer transport at the MADE site: The importance of heterogeneity, *Water Resour. Res.*, *43*(8), doi:10.1029/2006WR005522.

Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2006), Modeling mass transfer processes using random walk particle tracking, *Water Resour. Res.*, *42*(11), doi:10.1029/2006WR004927.

Salamon, P., D. Fernàndez-Garcia, and J. J. Gómez-Hernández (2006), A review and numerical assessment of the random walk particle tracking method, *J. Contam. Hydrol.*, *87*(3–4), 277–305, doi:10.1016/j.jconhyd.2006.05.005.

Sánchez-Vila, X., and J. Solís-Delfín (1999), Solute transport in heterogeneous media: The impact of anisotropy and non-ergodicity in risk assessment, *Stoch. Environ. Res. Risk Assess.*, *13*(5), 365–379, doi:10.1007/s004770050056.

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

Siirila-Woodburn, E. R., D. Fernàndez-Garcia, and X. Sanchez-Vila (2015), Improving the accuracy of risk prediction from particle-based breakthrough curves reconstructed with kernel density estimators, *Water Resour. Res.*, *51*(6), 4574–4591, doi:10.1002/2014WR016394.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*.

Simpson, M. J., and K. A. Landman (2007), Analysis of split operator methods applied to reactive transport with Monod kinetics, *Adv. Water Resour.*, *30*(9), 2026–2033, doi:10.1016/j.advwatres.2007.04.005.

Tompson, A. F. B. (1993), Numerical simulation of chemical migration in physically and chemically heterogeneous porous media, *Water Resour. Res.*, *29*(11), 3709–3726, doi:10.1029/93WR01526.

Tompson, A. F. B., and L. W. Gelhar (1990), Numerical simulation of solute transport in three???dimensional, randomly heterogeneous porous media, *Water Resour. Res.*, *26*(10), 2541–2562, doi:10.1029/WR026i010p02541.

Tompson, A. F. B., A. L. Schafer, and R. W. Smith (1996), Impacts of physical and chemical heterogeneity on cocontaminant transport in a sandy porous medium, *Water Resour. Res.*, *32*(4), 801–818, doi:10.1029/95WR03733.

Tsang, Y. W., and C. F. Tsang (2001), A particle-tracking method for advective transport in fractures with diffusion into finite matrix blocks, *Water Resour. Res.*, *37*(3), 831–835, doi:10.1029/2000WR900367.

van Breukelen, B. M. (2003), Natural Attenuation of Landfill Leachate:: a Combined Biogeochemical Process Analysis and Microbial Ecology Approach, Ipskamp.

Wen, X.-H., and J. J. Gómez-Hernández (1996), The Constant Displacement Scheme for Tracking Particles in Heterogeneous Aquifers, *Ground Water*, *34*(1), 135–142, doi:10.1111/j.1745-6584.1996.tb01873.x.

Willmann, M., G. W. Lanyon, P. Marschall, and W. Kinzelbach (2013), A new stochastic particle-tracking approach for fractured sedimentary formations, *Water Resour. Res.*, *49*(1), 352–359, doi:10.1029/2012WR012191.

Zhang, R., S. Hu, and X. Zhang (2006), Experimental Study of Dissolution Rates of Fluorite in HCl--H2O Solutions, *Aquat. Geochemistry*, *12*(2), 123–159, doi:10.1007/s10498-005-3658-3.

Zhang, Y., and D. A. Benson (2008), Lagrangian simulation of multidimensional anomalous transport at the MADE site, *Geophys. Res. Lett.*, *35*(7), doi:10.1029/2008GL033222.

Tables

*Table 1: Simulation parameter values used in example 1.**

	A	B	C
$(\mu_x, \sigma_x)_{initial}$	$(40, 6)$	$(50, 6)$	–
$m_{initial}$	1 mol	1 mol	0
α, β, γ	2.3	1.3	1
θ (eq. 37)	2.3	1.3	–
R	1	1	1
k_f	$6\ (\mathrm{mol/m^3})^{-2.6}\ \mathrm{day}^{-1}$		
q	$0.3\ \mathrm{m/day}$		
ϕ	0.25		
D	$0.4\ \mathrm{m^2/day}$		
τ	80 days		

* $(\mu_x, \sigma_x)_{initial}$ are the mean and standard deviation defining the initial normal distribution of solute particles in space, $m_{initial}$ is the total amount of substance at the start of the simulation, and τ is the total simulated time. The other variables are defined in the text.

*Table 2: Simulation parameter values used in example 2.**

	CH_2O	O_2	CO_2
$(\mu_x, \sigma_x)_{initial}$	$(70, 2)$	$(50, 8)$	–
$m_{initial}$	1 mol	1 mol	0
α, β, γ	1	1	1
K_{CH_2O}, K_{O_2} (eq.39)	$1.6667\ \mathrm{mol/m^3}$	$0.0156\ \mathrm{mol/m^3}$	–
R	3	1	1
k_f	$0.15\ (\mathrm{mol/m^3})\ \mathrm{day}^{-1}$		

q	$0.32\ \mathrm{m/day}$
ϕ	0.25
D	$0.5\ \mathrm{m^2/day}$
τ	$65\ \mathrm{days}$

* $(\mu_x, \sigma_x)_{initial}$ are the mean and standard deviation defining the initial normal distribution of solute particles in space, $m_{initial}$ is the total amount of substance at the start of the simulation, and τ is the total simulated time. The other variables are defined in the text.

*Table 3: Simulation parameter values used in example 3.**

	Ca^{2+}	CO_3^{2-}	$CaCO_3$
$(\mu_x, \sigma_x)_{initial}$	$(25, 5)$	$(35, 8)$	$-$
$m_{initial}$	$2.5\ \mathrm{mol}$	$2.5\ \mathrm{mol}$	0
α, β, γ	1	1	1
$\mathring{a}_{Ca^{2+}}, \mathring{a}_{CO_3^{2-}}$ (eq. 43)	$0.6\ \mathrm{nm}$	$0.5\ \mathrm{nm}$	$-$
R	1	1	∞
k_{obs}	$0.002\ (\mathrm{mol/m^3})^{-2.6}\ \mathrm{day^{-1}}$		
k_{eq}	$10^{-2.3}\ (\mathrm{mol/m^3})^2$		
q	$0.1\ \mathrm{m/day}$		
ϕ	0.25		
D	$0.15\ \mathrm{m^2/day}$		
τ	$100\ \mathrm{days}$		

* $(\mu_x, \sigma_x)_{initial}$ are the mean and standard deviation defining the initial normal distribution of solute particles in space, $m_{initial}$ is the total amount of substance at the start of the simulation, and τ is the total simulated time. The other variables are defined in the text.

*Table 4: Simulation parameter values used in example 4.**

	H^+	Ca^{2+}	HF^0
$(\mu_x, \sigma_x)_{initial}$	$(90, 5)$	$(90, 20)$	$-$
$m_{initial}$	$2\ \mathrm{mol}$	$4\ \mathrm{mol}$	0
α, β, γ	2	1	2

$\boldsymbol{\theta}$ (eq. 45)	$\mathbf{1.6}$	$\mathbf{-0.8}$	$-$
$\boldsymbol{R}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$
$\boldsymbol{k_f}$	$\mathbf{0.72\ (mol/m^3)^{0.2}\ day^{-1}}$		
$\boldsymbol{q}$	$\mathbf{0.5\ m/day}$		
$\boldsymbol{\phi}$	$\mathbf{0.25}$		
$\boldsymbol{D}$	$\mathbf{0.4\ m^2/day}$		
$\boldsymbol{\tau}$	$\mathbf{7\ days}$		

* $(\boldsymbol{\mu_x}, \boldsymbol{\sigma_x})_{\boldsymbol{initial}}$ are the mean and standard deviation defining the initial normal distribution of solute particles in space, $\boldsymbol{m_{initial}}$ is the total amount of substance at the start of the simulation, and τ is the total simulated time. The other variables are defined in the text.

Figures

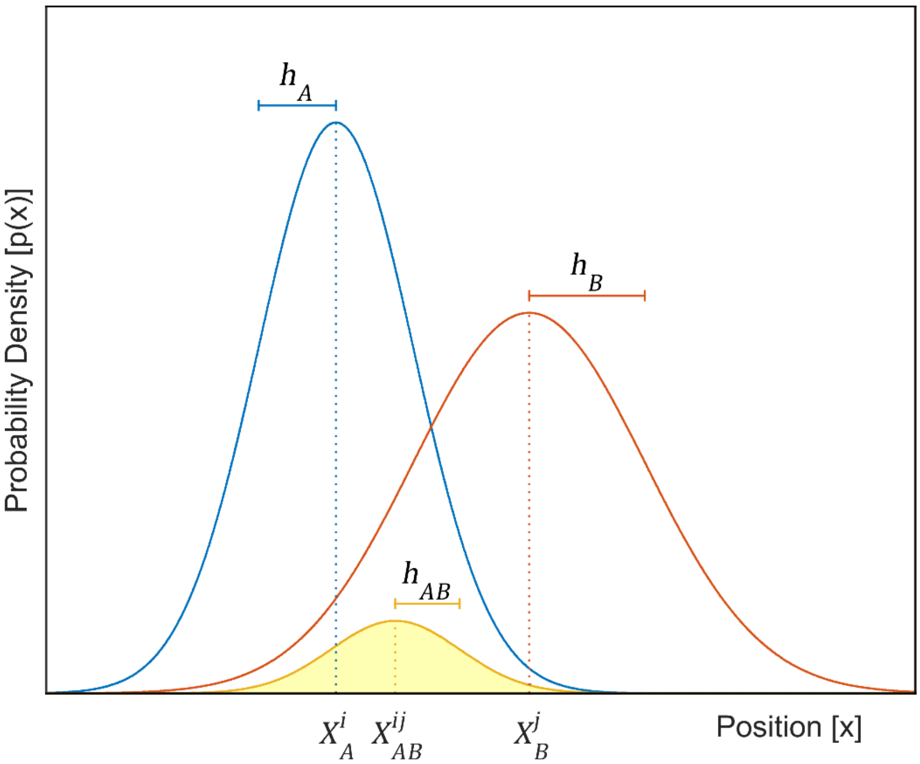

Figure 1: Schematic example of a product between two Gaussian pdf's in one dimension. The product (yellow) is another Gaussian function centered in X_{AB} and with

a standard deviation $h_{ab} = \sqrt{H_{AB}}$. *Its integral over x is the probability of collocation of the two particles.*

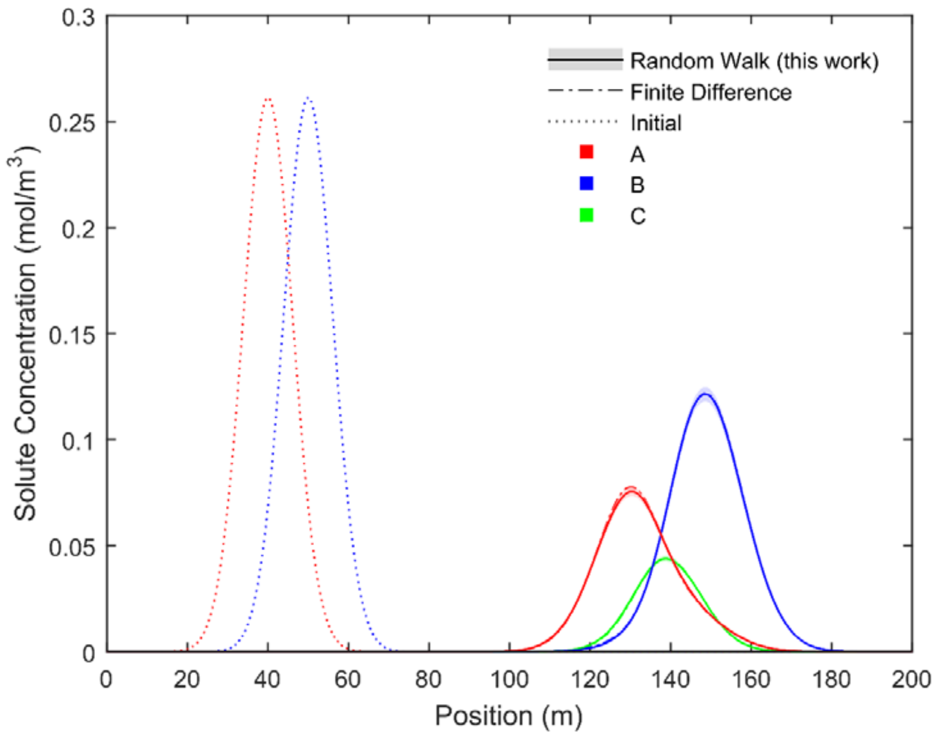

Figure 2: Solute concentrations in example 1, at the start of the simulation and after 80 days. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

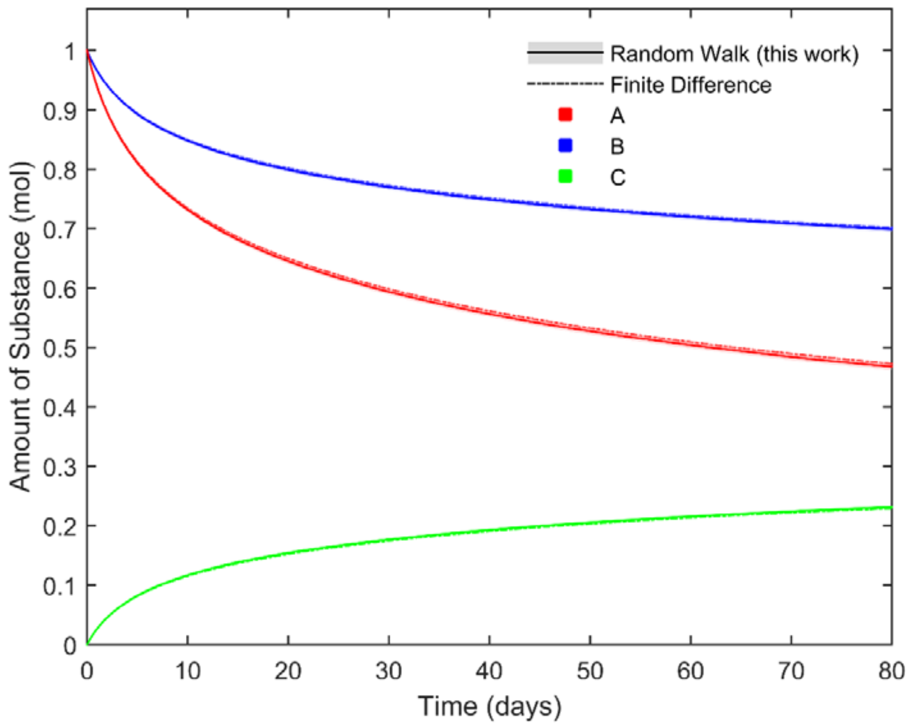

Figure 3: Evolution in time of the total amount of substance of each compound in example 1. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

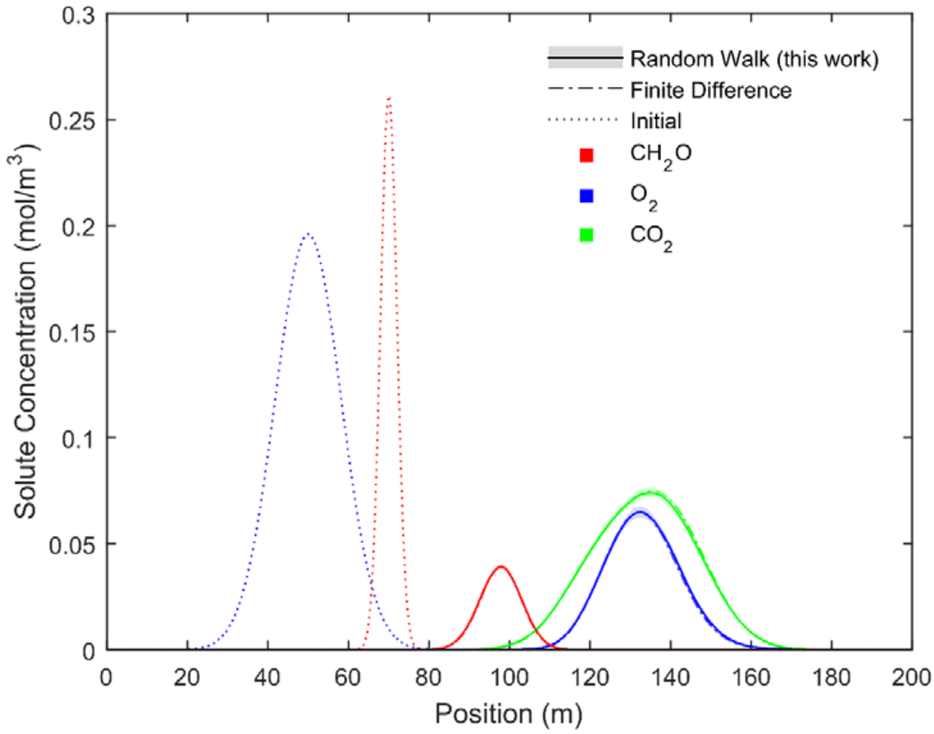

Figure 4: Solute concentrations in example 2, at the start of the simulation and after 65 days. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

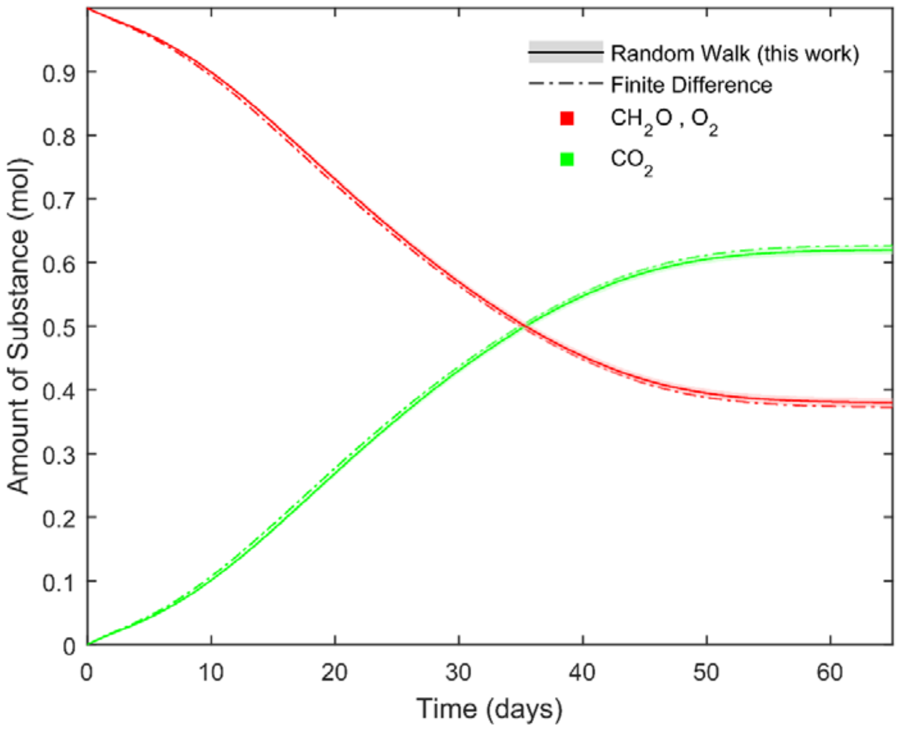

Figure 5: Evolution in time of the total amount of substance of each compound in example 2. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

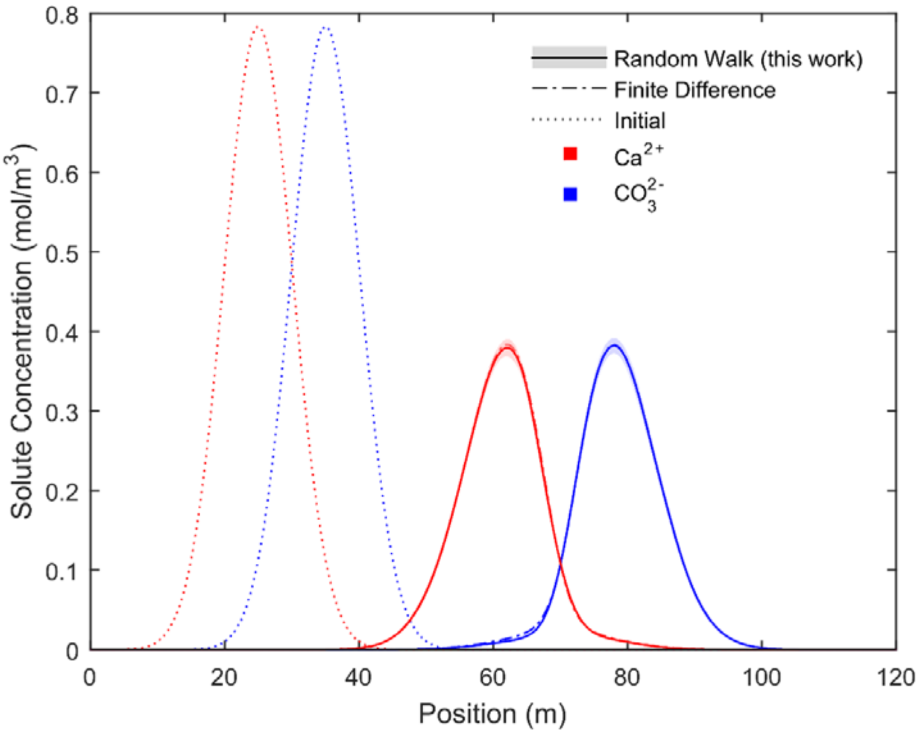

Figure 6: Solute concentrations in example 3, at the start of the simulation and after 100 days. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

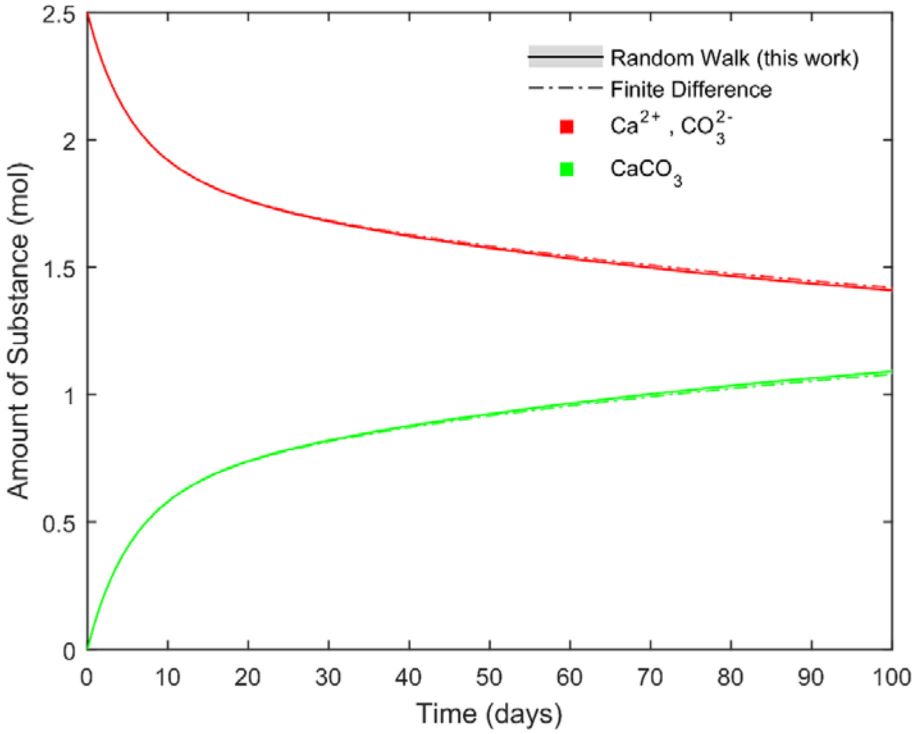

Figure 7: Evolution in time of the total amount of substance of each compound in example 3. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

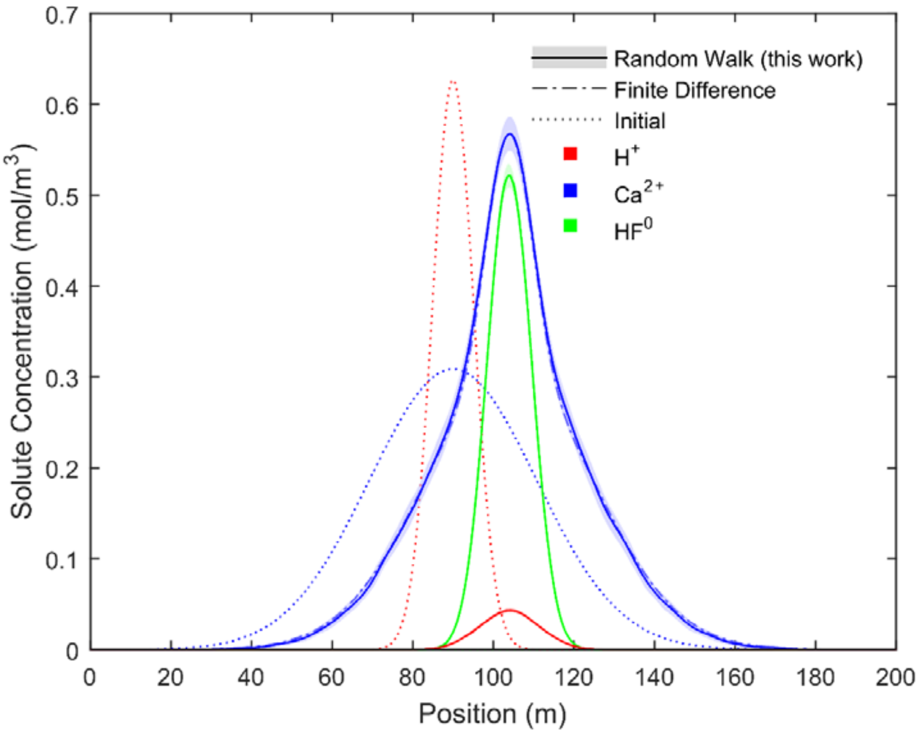

Figure 8: Solute concentrations in example 4, at the start of the simulation and after 7 days. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

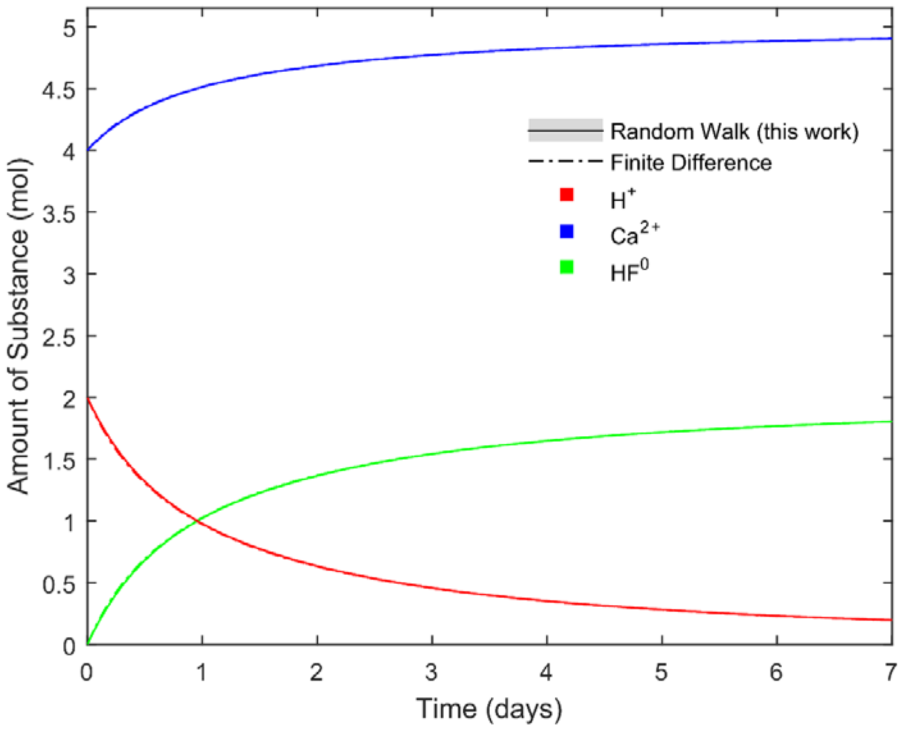

Figure 9: Evolution in time of the total amount of substance of each compound in example 4. The error zone around the Particle Tracking curves indicates ±1 standard deviation.

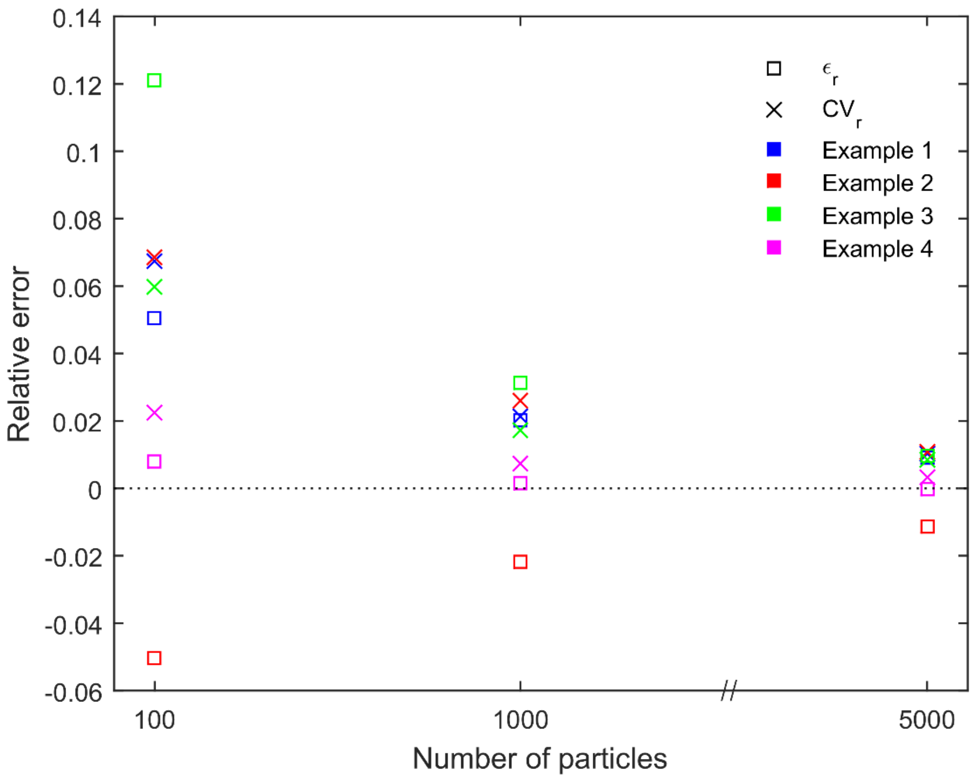

Figure 10: The two measured error components for different initial number of particles of the reactants.

923