Simulating Solute Transport in Porous or Fractured Formations Using Random Walk Particle Tracking: A Review

Frédérick Delay,* Philippe Ackerer, and Charles Danquigny

ABSTRACT

Since the first attempts some 20 yr ago in the field of hydrology, random walk (RW) particle tracking as applied to solute transport has experienced profound changes. Concepts and mathematical techniques have improved to the point that numerically difficult problems (e.g., advection-dominated transport in highly heterogeneous media, or reactive transport) are now much easier to address. Random walk methods have never been widely used for multiphase flow, probably because numerical dispersion is not a major problem for modeling exercises at large scales. However, vadose zone hydrologic studies often point out very strong variations in fluid velocity over relatively short distances. Random walk methods may be well suited for such studies, a possibility which motivated us to write this review. We first give a comprehensive discussion of the theoretical context of the method. The Fokker–Planck–Kolmogorov equation (FPKE) is established for solute transport, as well as the ordinary Langevin equation and its simplifications for transport of small particles (e.g., colloids). Next, numerical methods are developed for the motion of particles in space. An important section is subsequently dedicated to recent RW concepts in the time domain, and to their application to anomalous (non-Fickian) transport and inverse problems. Adaptations of RW to transport with solute–solid reactions are also provided, as well as several numerical recipes for resolving a few computational difficulties with the RW method. We purposely did not include any comparisons with Eulerian and Lagrangian approaches. These approaches are discussed at length in several references cited in this review. We note, however, that today's computing capabilities provide new incentives to using RW methods for problems where Eulerian methods are potentially unstable or hampered by numerical diffusion.

There is an increasing need for accurate numerical solutions to solute transport problems in porous and/or fractured formations. This need is predicated by increased concerns about soil and water quality, the long-term impact of industrial and agricultural contaminants on the environment, and the feasibility of underground repository sites for nuclear or other wastes. While considerable efforts during the past several decades have resulted in vastly improved numerical techniques for solving subsurface transport problems, accurate simulations of seemingly straightforward advection–dispersion problems often remain a challenge when transport is advection dominated. This problem becomes even more challenging when attempts are made to accurately account for the enormous heterogeneity of soil and groundwater systems (e.g., de Marsily et al., 1998, 2004).

Traditional Eulerian approaches to solute transport generally require very fine discretization of the transport domain to overcome recurrent problems of unstable numerical solutions and/or artificial diffusion. Solving such problems generally imposes a heavy computational burden. If the discretization is too coarse, oscillations and numerical diffusion of standard Eulerian methods may yield poor or even incorrect solutions, particularly when nonlinear processes such as multiphase flow or adsorption–desorption are considered in conjunction with standard advective–dispersive transport. This problem proved to be crucial in safety calculations of underground repository sites for nuclear wastes. To address these problems, several alternative Eulerian–Lagrangian schemes based on particle tracking have been developed, such as the method of characteristics (Konikow et al., 1996). Particle tracking has also been used for numerical integration in more complicated schemes, such as the Eulerian–Lagrangian localized adjoint method (ELLAM) (Russell and Celia, 2002; Younès, 2004).

We provide a review of the principles of fully Lagrangian schemes based on RW particle tracking for solving advection–dispersion problems. Lagrangian approaches have long been used widely in physics for a variety of numerical applications. In the hydrologic sciences they were first applied in the late 1970s and early 1980s to transport in saturated alluvial formations (e.g., Ahlstrom et al., 1977; Prickett et al., 1981). Despite the encouraging results during the last two decades for simulating transport in saturated media, RW methods have been used only very rarely for unsaturated media. This is surprising since no inherent incompatibility exists between the RW method and unsaturated flow. On the contrary, soils are often characterized by strong contrasts in local flow velocities when a moisture front propagates during variably saturated flow. While Eulerian methods often experience numerical difficulties when simulating such velocity contrasts over short distances, Lagrangian models are much easier and more accurately to implement, as we shall demonstrate later.

In attempts to focus our review primarily on RW methods, we decided not to provide detailed comparisons between discrete Eulerian approaches (such as finite volume and finite element methods) and Lagrangian methods. Abundant literature already exists on such comparisons (we also provide several useful references). An extensive comparison between Eulerian and Lagrangian approaches would have required a much longer review. For these same reasons, we will not discuss all

Abbreviations: ADE, advection–dispersion equation; CTRW, continuous time random walk; ELLAM, Eulerian Lagrangian localized adjoint method; EPT, enhanced particle tracking; FPKE, Fokker–Planck–Kolmogorov equation; LEA, local equilibrium assumption; RW, random walk; TDRW, time domain random walk.

F. Delay, Univ. of Poitiers, Earth Sciences Building, 40 Av. du Recteur Pineau, 86022 Poitiers, Cedex, France; P. Ackerer and C. Danquigny, Fluid and Solid Mechanics Institute of Strasbourg, 2 rue Boussingault, 67000 Strasbourg, France. Received 9 Sept. 2004. Review and Analysis. ©Corresponding author (fred.delay@hydrasa.univ-poitiers.fr).

Published in Vadose Zone Journal 4:360–379 (2005).
doi:10.2136/vzj2004.0125
© Soil Science Society of America
677 S. Segoe Rd., Madison, WI 53711 USA
available RW methods, such as Boltzmann lattice gas approaches, the use of particles for numerical thermodynamics, stochastic properties of perfect gases, and molecular simulations in theoretical chemistry. Our study is aimed at solving macroscopic solute transport problems in porous or fractured media involving advection–dispersion as the principal mechanism. However, some discussions will be devoted to the ordinary Langevin equation which, as shown below, is well suited to mimic transport of solid particles subject to interaction forces in a moving fluid. This additional discussion seems justified here in that it may help the formulation of models for colloidal transport in porous media, and because the Langevin equation under some assumptions can be simplified into a slightly modified form of the advection–dispersion equation (ADE).

The main topics to be reviewed are as follows. The first section provides a comprehensive overview of the FPKE and the ordinary Langevin equations, and their relationship with the advection–dispersion equation for solute transport. The principles of stochastic physics on which these equations rely are simplified as best as possible to make this section readable without much background. A second section explains how to construct “explicit” solutions by randomly moving particles in space (classical RW particle tracking). Special emphasis is placed on algorithms that preserve solute mass balance at interfaces showing sharp contrasts in dispersion and on algorithms that accurately account for flow field heterogeneity. A third section presents recent concepts of random walk in the time domain. Two approaches are discussed: (i) a continuous time RW approach that calculates the probability density of particle residence times as a function of the fluid velocity distribution and (ii) a time domain RW method which moves particles explicitly between fixed sites of a network, but calculates the process that calculates the probability density of particle residence of duration \( t \) as a function of the fluid velocity distribution and lag-time \( t \). Introduce now \( P(x, t) \), the probability density for a particle to be at location \( x \) at time \( t \). \( P(x, t) \) is equivalent to \( \psi(x, x_0, t) \), that is, the probability density of a transition \( x_0 \rightarrow x \) within duration \( t \). If \( t \) is large enough for the transition \( x_0 \rightarrow x \) to be performed in several steps \( x_0 \rightarrow x_1, \ldots, x_{k-1} \rightarrow x_k, x_k \rightarrow x \) because of the Markovian nature of this process, \( \psi(x, x_0, t) \) only depends on all possible locations \( x_i \) experienced before the last step \( x_k \rightarrow x \). Thus, any reference to the initial location \( x_0 \) may be dropped without loss of generality, and \( \psi(x, x_0, t) \) becomes \( P(x, t) \). A classical assumption of first-order derivatives allows us to write

\[
\frac{\partial \psi(x, x_0, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{\psi(x, x_0, t + \Delta t) - \psi(x, x_0, t)}{\Delta t}
\]

or with the equivalence between \( P( ) \) and \( \psi( ) \):

\[
\frac{\partial P(x, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{\psi(x, x_0, t + \Delta t) - P(x, t)}{\Delta t}
\]

An expression for \( \psi(x, x_0, t + \Delta t) \) for small values of \( \Delta t \) can be derived using the property of the Markovian process that

\[
\psi(x, x_0, t + \Delta t) = \int_{\Omega} \psi(y, x_0, t) \psi(y, x, \Delta t) dy
\]

which is the probability density to be at \( y \) at time \( t \) and to make a jump \( y \rightarrow x \) within duration \( \Delta t \). This probability is summed over all possible locations \( y \) of the domain \( \Omega \). Note that Eq. [3] is often referred to as the Chapman–Kolmogorov equation for random particles. Introducing Eq. [3] in [2] yields

\[
\frac{\partial P(x, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\Omega} [P(y, t) \psi(y, x, \Delta t) dy - P(x, t)]
\]

When \( \Delta t \) tends to zero, no transition \( y \rightarrow x \) may occur except for \( x = y \). Thus, we can write

\[
\lim_{\Delta t \to 0} \psi(x, y, \Delta t) = \delta(x - y)
\]

in which \( \delta( ) \) is the Dirac-delta function \([\delta(0) = 1, \delta(x \neq 0) = 0]\). This property allows us to expand \( \psi(x, y, \Delta t) \) as a series with respect to the derivatives of the \( \delta( ) \) function (Zaslavsky, 2002):

\[
\psi(x, y, \Delta t) = \delta(x - y) + \sum_{i=1}^{\infty} \frac{1}{i!} M(y, \Delta t) \delta_i(x - y)
\]

where \( \delta_i( ) \) is the \( i \)-th order derivative of \( \delta(x) \) with respect
to $\chi$, and $M_i(y, \Delta t)$ is the $i$th order moment of the transitions starting at $y$ and being of duration $\Delta t$:

$$M_i(y, \Delta t) = \int \psi(y, x, \Delta t)(x - y)dx$$ \[7\]

Note that the exponent $i$ in Eq. [7] refers to elevation to a power. To the first-order, $M_1(y, \Delta t)$ is the vector (e components) of the mean size of transitions starting at location $y$ at any time $\tau$ and lasting $\Delta t$. To the second-order, $M_2(y, \Delta t)$ is the covariance matrix (e components) between the components along the $e$ spatial directions of the above transitions. Replacing $\psi(y, x, \Delta t)$ in Eq. [4] by its expansion Eq. [6] up to the second order yields

$$\frac{dP(x, t)}{dt} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ \frac{P(y, t)\delta(x - y)dy}{\Omega} + P(y, t)M_1(y, \Delta t)\delta'(x - y)dy + \frac{1}{2} \Omega P(y, t)M_2(y, \Delta t)\delta''(x - y)dy - P(x, t) \right]$$ \[8\]

Knowing that the first and fourth terms of the right-hand side cancel, and using integration by parts for the second and third terms, leads to

$$\frac{dP(x, t)}{dt} = -\lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ \frac{\partial}{\partial x} \left[ P(x, t)M_1(x, \Delta t) \right] \right]$$

$$+ \frac{1}{2} \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ \frac{\partial^2}{\partial x^2} \left[ P(x, t)M_2(x, \Delta t) \right] \right]$$ \[9\]

The so-called Kolmogorov assumption states that the following limits exist:

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} M_1(x, \Delta t) = A(x); \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} M_2(x, \Delta t) = B(x);$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} M_3(x, \Delta t) = 0 \quad (n > 2)$$ \[10\]

The third condition (i.e., moments of orders above 2 for transitions starting from $x$ and of duration $\Delta t$ cancel out) justifies the limited development of Eq. [9] up to the second order. The vector $A(x)$ and tensor $B(x)$ have dimensions of $(L T^{-1})$ and $(L^2 T^{-1})$, respectively. Their common interpretation for the motion of particles by successive jumps is that $A(x)$ corresponds to the mean of the jump velocity, while $B(x)$ is the statistical dispersion of this velocity around its mean. Introducing Kolmogorov assumption (Eq. [10]) in Eq. [9] yields the so-called FPKE for transport:

$$\frac{dP(x, t)}{dt} = -\frac{\partial}{\partial x} \left[ A(x)P(x, t) \right]$$

$$+ \frac{1}{2} \left[ \frac{\partial^2}{\partial x^2} \left[ B(x)P(x, t) \right] \right]$$ \[11\]

We show below that this equation has an explicit solution that for the simplest case consists of moving particles that undergo successive jumps of mean size $A(x)\Delta t$ and covariance $B(x)\Delta t$, with $\Delta t$ being the time step of the jumps.

The classical ADE for solute transport in a porous medium is given by

$$\frac{\partial C(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left[ u(x, t)C(x, t) \right] + \frac{\partial}{\partial x} \left[ D(x, t) \frac{\partial C(x, t)}{\partial x} \right]$$ \[12\]

in which $C(x, t) (M L^{-3})$ is the volumetric concentration of the solute at location $x$ and time $t$, $u(x, t) (L T^{-1})$ is the mean fluid velocity vector, and $D(x, t) (L^2 T^{-1})$ is the dispersion tensor. The similarity of the ADE and FPKE should be evident by noting that

$$C(x, t) = P(x, t); \quad u(x, t) = A(x);$$

$$D(x, t) = \frac{1}{2} B(x)$$ \[13\]

Note in Eq. [13] that $A(x)$ and $B(x)$ are independent of time but compared with possible time-dependent values $u(x, t)$ and $D(x, t)$. This apparent discrepancy stems from the stationary assumption for $\psi(\cdot)$ and could mean that FPKE is not suited for transport under, for example, transient flow conditions. In fact, for practical and numerical applications, the stationary assumption can be relaxed provided transition times between successive particle locations are small compared with characteristic times of the flow field variations. Strict equivalence between the ADE and FPKE using Eq. [13] holds only if the dispersion tensor $D(x, t)$ is constant in space; that is, $\partial(D\partial C/\partial x)/\partial x = \partial(D\partial C/\partial x^2)/\partial x = D\partial^2 C/\partial x^2$. While this may be the case for a homogeneous medium subject to flow with a uniform velocity $u$, such a condition is not frequently encountered in natural environments. However, the ADE given by Eq. [12] may be rewritten as

$$\frac{\partial C(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left[ u(x, t) + \frac{\partial D(x, t)}{\partial x} \right] C(x, t)$$

$$+ \frac{\partial^2}{\partial x^2} \left[ D(x, t)C(x, t) \right]$$ \[14\]

In this form, and assuming the equivalence $A(x) = u(x, t) + \partial D(x, t)/\partial x$ instead of $A(x) = u(x, t)$, the ADE is similar to the FPKE. The importance of this modified velocity $u(x, t) + \partial D(x, t)/\partial x$ will be discussed in more detail in a section below. The modified velocity is key in random walk applications to avoid mass-balance discrepancies at interfaces with sharp dispersion contrasts. Ackerer and Kinzelbach (1985) and Uffink (1985) were first in the hydrologic community to raise this problem. Ever since, much literature has been dedicated to algorithms able to solve the nonphysical excess of mass in low-dispersion areas when FPKE-based random particles are used without the modified velocity. We will revisit this topic below. While the form of $A(x) = u(x, t) + \partial D(x, t) / \partial x$ may appear surprising at first, we note that $A(x)$ has the physical meaning of an estimate of the mean displacement per unit time, whereas $u(x, t)$ is a unit flux of particles through a certain section (Uffink, 1990).
This difference between $A(x)$ and $u(x, t)$ was also shown by Kitanidis (1988, 1994) using integration by parts of the ADE given by Eq. [12]. Note that when the porosity $\phi$ of the medium is variable in space and/or time, or when dealing with transport in unsaturated soils having a water content $\theta$, the concentration $C(x, t)$ in the ADE must be replaced by $\phi(x, t)C(x, t)$ or $\theta(x, t)C(x, t)$. The similarity with the FPKE is then ensured by using $\phi(x, t)$ $C(x, t)$ $= P(x, t)$. The physical meaning is that $\phi C V$ is the solute mass in a bulk volume $V$ of porous medium, while the mass conservation principle has a meaning equivalent to the property of probability density functions which states that their integral should be equal to one.

The Ordinary Langevin Equation

The ordinary Langevin equation of motion has been widely used to model the transport of immiscible gas bubbles or drops in fluids (Ramarao and Tien, 1992) and of atmospheric aerosols (Gupta and Peters, 1985). Its application to transport problems in natural porous media is less common. One reason is that its solution is cumbersome when dealing with complex velocity fields since the problem is addressed at the pore scale and requires, as a preliminary step, a solution of the Navier–Stokes equation. Explicit solutions for the flow field are still available for simplified systems such as the “sphere in cell” model (Happel, 1958) or periodic stacking of spherical beads. These models are useful, for instance, to simulate the interaction between suspended matter and the solid phase of the porous medium (Elimelech and O’Melia, 1990), and may find many applications involving colloidal transport in soils. Nevertheless, the Langevin approach to transport requires, in theory, very small time steps for the best accuracy. While our aim here is not to discuss in detail potential applications for which the Langevin equation is relevant, it is included in our discussion for three reasons. First, the Langevin equation is not well known in current literature on groundwater. Second, its complete solution is based on a Lagrangian approach for moving particles in space and time. Third, the equation can be simplified with certain assumptions into the classical RW approach, which is the main topic of this review.

The ordinary Langevin equation is based on the conservation of momentum $(m v)$ of a particle of mass $m$ moving at velocity $v$ in a fluid. In the following we assume that the Euclidean space is three-dimensional. Thus, except when specified, the variables involved are vector quantities depending on their space and time coordinates. The notation as a simple variable without reference to space coordinates is used here for the sake of simplicity. A classical form of the Langevin equation for transport of particles in porous media can be written as follows:

$$\frac{dv(t)}{dt} = -\beta[v(t) - u(t)] + \frac{F(t)}{m} + A(t) \tag{15}$$

where $v(t)$ is the particle velocity (L T$^{-1}$), $u(t)$ is the fluid velocity (L T$^{-1}$), while $F(t)$ (M L T$^{-2}$) combines interaction forces between both the particle and the solid phase and between the particles themselves. Additionally, $A(t)$ (L T$^{-2}$) is the acceleration due to Brownian motion stemming from collisions between the particle and fluid molecules, $m$ is the mass of the particle, and $\beta = 6\pi \eta r/m$ is the scalar Stokes coefficient of the drag force (T$^{-1}$) in which $\eta$ is the dynamic viscosity of the fluid (M L$^{-1}$ T$^{-1}$), and $r$ the radius of the particle. Given the form of Eq. [15], the particle velocity is of the following general form:

$$v(t) = K(t)\exp(-\beta t) \tag{16}$$

where $K(t)$ is a vector of dimension (L T$^{-1}$). Reintroducing Eq. [16] in [15] gives an equation for $K(t)$:

$$\frac{dK(t)}{dt} \exp(-\beta t) = \beta u(t) + \frac{F(t)}{m} + A(t) \tag{17}$$

For a time step defined by the relative times $0 \rightarrow t$, one obtains

$$K(t) - K_0 = \int_0^t [\beta u(\tau) + \frac{F(\tau)}{m} + A(\tau)] \exp(\beta \tau) d\tau \tag{18}$$

in which $K_0 = v_0$ (i.e., the value given by Eq. [16] at $t = 0$). Substituting Eq. [18] into [16] gives the following expression for the particle velocity:

$$v(t) = v_0 \exp(-\beta t) + \exp(-\beta t) \int_0^t \beta u(\tau) + \frac{F(\tau)}{m} + A(\tau) \exp(\beta \tau) d\tau \tag{19}$$

If the time step $0 \rightarrow t$ is small enough such that both the fluid velocities $u(\tau)$ and the external forces $F(\tau)$ can be assumed constant, Eq. [19] simplifies to yield a recursive algorithm for calculating the particle velocity at relative time $t$, provided its value is known at $t = 0$:

$$v(t) = v_0 \exp(-\beta t) + \frac{u_0 + \frac{F_0}{\beta m}[1 - \exp(-\beta t)]}{\beta} + \exp(-\beta t) \int_0^t A(\tau) \exp(\beta \tau) d\tau \tag{20}$$

The third term of the right-hand side of Eq. [20]; that is,

$$\xi(t) = \int_0^t A(\tau) \exp(\beta \tau) d\tau \tag{21}$$

is random and expresses the Brownian component of the particle velocity. Its main characteristics will be discussed below.

If the particle velocity is available at successive time steps, then the particle trajectory can be calculated by integration of $v$ over $0 \rightarrow t$:

$$r(t) = r_0 + \int_0^t v(\tau) d\tau \tag{21}$$

where $r(t)$ is the location of the particle at relative time $t$. Introducing Eq. [20] in [21] yields
\[ r(t) = r_0 + \frac{v_0}{\beta} \int_0^t \left[ 1 - \exp(-\beta \tau) \right] d\tau + \left( u_0 + \frac{F_0}{\beta m} \right) \left[ t + \frac{1}{\beta} \left( \exp(-\beta t) - 1 \right) \right] \\
+ \frac{1}{\beta} \int_0^t A(\tau) \left( 1 - \exp[\beta(\tau - t)] \right) d\tau \]

Finally, the particle location at relative time \( t \) is

\[ r(t) = r_0 + \frac{v_0}{\beta} \left[ 1 - \exp(-\beta t) \right] + \left( u_0 + \frac{F_0}{\beta m} \right) \left[ t + \frac{1}{\beta} \left( \exp(-\beta t) - 1 \right) \right] + \frac{1}{\beta} \int_0^t A(\tau) \left( 1 - \exp[\beta(\tau - t)] \right) d\tau \]

The last term, that is,

\[ \xi(t) = \frac{1}{\beta} \int_0^t A(\tau) \left( 1 - \exp[\beta(\tau - t)] \right) d\tau \]

is also a random vector corresponding to the Brownian component of the particle displacement during \( 0 \rightarrow t \). Einstein's work between 1905 and 1908 on Brownian motion (e.g., as documented in a book edited by Furth, 1956) has shown that the acceleration \( A(t) \) is a vector of independent components \( A_x(t) (\lambda = x, y, z \) directions) obeying Gaussian distributions and having the following properties:

\[ \langle A_x(t) \rangle = 0; \quad \langle A_x(t) A_y(t) \rangle = 2q \delta(\lambda - \gamma) \delta(t - \tau) \]

in which \( \langle \rangle \) is the mathematical expectation, \( \delta \) the Dirac-delta function as before, and \( q = kT\beta/m \), where \( k \) is the Boltzmann constant \( (\text{ML}^2\text{T}^{-1}\text{K}^{-1}) \) and \( T \) is absolute temperature (K). Using these properties of Brownian acceleration, Chandrasekhar (1943) showed that the vectors \( \xi_0(t) \) and \( \xi(t) \) are correlated and obey multi-Gaussian distributions. It can be shown that more generally the vector quantities in three dimensions, that is,

\[ W(t) = \int_0^t A(\tau) \psi(t - \tau) d\tau \quad \text{and} \quad X(t) = \int_0^t A(\tau) \phi(t - \tau) d\tau \]

have Gaussian probability density functions of the form:

\[ P_w = \frac{1}{2\pi q} \exp \left[ -\frac{1}{2} \int_0^t W(t)^2 W(t)^2 \right] \]

\[ P_w = \frac{1}{8\pi^2(fg - h^2)^2} \]

\[ f = 2q \int_0^t \psi(t - \tau) d\tau \]

\[ g = 2q \int_0^t \psi(t - \tau) d\tau \]

in which \( T \) indicates the transposition operator. Given the algebraic forms of \( \xi(t) \) and \( \xi_0(t) \) (see Eq. [20] and [24]) and Eq. [26], the first- and second-order moments of these distributions can be calculated as

\[ \langle \xi_x(t) \rangle = \langle \xi_0(t) \rangle = 0; \quad \lambda, \gamma = x, y, z \]

\[ \langle \xi_x(t) \xi_y(t) \rangle = \frac{q}{\beta} \int_0^t \left[ 1 - \exp(-2\beta t) \right] \delta(\lambda - \gamma) \quad \lambda, \gamma = x, y, z \]

\[ \langle \xi_x(t) \xi_y(t) \rangle = \frac{q}{\beta} \left[ 1 - \exp(-2\beta t) \right] \delta(\lambda - \gamma) \]

where \( \xi_\lambda (i = 1, 2; \lambda = x, y, z) \) are components of the vector \( \xi(t) \) along direction \( \lambda \).

Equations [20], [24], and [27] define a Lagrangian method for solving the ordinary Langevin equation. As stated above, the calculations are cumbersome since they require an accurate description of the flow field at the scale. Also, the assumption of having a constant fluid velocity and constant external forces during a time step forces the time steps to be small (e.g., on the order of \( 10^{-4} \) to \( 10^{-6} \) s for transport of stable clay suspensions in sand columns, as shown by Compere et al., 2001). In many cases, however, the above Lagrangian approach can be simplified to produce a slightly modified ADE. Assume that the time step \( 0 \rightarrow t \) is such that \( t \gg 1/\beta \), and hence \( \beta t \gg 1 \) (common values of \( \beta \) for clay particles of size 0.1 to 1 \( \mu \text{m} \) in water are in the range of \( 10^6 \) to \( 10^8 \text{ s}^{-1} \)). Equations [20] and [24] then simplify to

\[ v(t) = u_0 + \frac{F_0}{\beta m} + \xi(t) \]

\[ r(t) = r_0 + \frac{v_0}{\beta} \left[ t + \left( u_0 + \frac{F_0}{\beta m} \right) \right] + \xi(t) \]

Knowing that \( u_0 + F_0/\beta m \) has the magnitude of \( v_0 \), the term \( v_0/\beta \) is negligible compared with \( (u_0 + F_0/\beta m) t \). Moreover, \( \langle \xi(t) \rangle = 0 \) (see Eq. [27]), which allows Eq. [28] to be rewritten in the form

\[ \langle v \rangle = u_0 + \frac{F_0}{\beta m} \]

\[ r(t) = r_0 + \langle v \rangle t + \xi(t) \]

The term \( \xi(t) \) in Eq. [29] is known to be a vector with independent Gaussian components, while simplifications of Eq. [27] stemming for \( \beta t \gg 1 \) lead to \( \langle \xi_\lambda(t) \rangle = 2q/t \beta^2 \).

With \( D = kT/(6\pi \mu R) \) \( (L^2 \text{T}^{-1}) \) designating the Stokes–Einstein diffusion coefficient, we have:

\[ \langle \xi_\lambda(t) \rangle = 2Dt \]
and $1/\beta m = D/kT$. Equation [29] can now be rewritten as follows:

$$\langle \mathbf{v} \rangle = \mathbf{u}_0 + \frac{\mathbf{F}_0 D}{kT}$$

$$\mathbf{r}(t) = \mathbf{r}_0 + \langle \mathbf{v} \rangle t + \mathbf{z} \sqrt{2Dt} \quad [30]$$

where $\mathbf{z}$ is a vector having independent random components drawn from normal deviates of zero mean and unit variance. As further discussed below, Eq. [30] expresses a classical random walk procedure for advective–dispersive transport with a unit mass flux, $\mathbf{J}$ (M L$^{-2}$ T$^{-1}$): $\mathbf{J} = [\mathbf{u} + (\mathbf{F} D / kT)] \mathbf{C} - D \nabla \mathbf{C}$, which is the usual form when external forces have a nonnegligible influence on solute transport.

The same simplifications of the ordinary Langevin equation can be implemented in another manner, and probably with a more physical basis. Assume that the particles are very small and very light in weight, such that inertia in the Langevin equation can be neglected (the extreme would be particles with properties of fluid molecules, i.e., a “perfect” solute). Setting $d\mathbf{v}/dt$ in Eq. [15] to zero and assuming constant fluid velocities and external forces during a time step $0 \rightarrow t$ yields

$$\mathbf{v}(t) = \mathbf{u}_0 + \frac{\mathbf{F}_0}{\beta m} t + \frac{1}{\beta} \mathbf{A}(t)$$

Since $\langle \mathbf{A}(t) \rangle = 0$ (see Eq. [25]), the mean velocity of the particles becomes $\langle \mathbf{v} \rangle = \mathbf{u}_0 + \mathbf{F}_0 / \beta m$. Calculation of the particle displacement from Eq. [31] is now straightforward to give

$$\mathbf{r}(t) = \mathbf{r}_0 + \left( \mathbf{u}_0 + \frac{\mathbf{F}_0}{\beta m} t \right) t + \frac{1}{\beta} \mathbf{A}(t)$$

Applying the general Eq. [26] to the term $\xi(t) = 1/\beta f_0^t \mathbf{A}^\prime(t) dt$ results in

$$P_{\xi} = \frac{1}{(4\pi D t)^{3/2}} \exp \left( -\frac{\xi(t)^T \xi(t)}{4Dt} \right)$$

which is a tri-Gaussian distribution of independent components $\xi_i(t)$ (L) with zero mean and a variance of $2qt/\beta^2$. Using the earlier definition of the Stokes–Einstein diffusion coefficient $D$ in the expressions for $\langle \mathbf{v} \rangle$ and $\mathbf{r}(t)$ leads again to the classical random walk method given by Eq. [30].

NUMERICAL IMPLEMENTATION OF RANDOM WALK PARTICLE TRACKING IN SPACE

We have already shown that both the FPKE given by Eq. [11] and the simplified form of the Langevin equation given by Eq. [30] relate the transport mechanism to the motion of particles with well-established probability distributions for transitions between two successive locations. Basically, these equations are valid from the Darcy scale (say, a few ten thousand pores) up to large scales (from meters to kilometers). On the other hand, a full Langevin approach (Eq. [20], [24], and [27]) to transport in porous media is only applicable at the pore scale but proceeds in the same way, that is, moving particles with known transition probabilities. For the sake of simplicity we focus the following on mimicking the FKPE with particles. Such an approach assumes that the location of a particle evolves in time as follows:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + t \mathbf{u} \Delta t + \left[ \mathbf{B}(x) \mathbf{u} \Delta t \right] \mathbf{Z}$$

where $\mathbf{x}(t)$ is the vector of $e$ coordinates defining the location of the particle at time $t$, $\mathbf{u}$ is the dimension of Euclidean space, and $\mathbf{A}(x)$ and $\mathbf{B}(x)$ are the vector and the tensor given by Eq. [10]. As discussed earlier, $\mathbf{A}(x) \Delta t$ is the mean displacement of the particle during time step $\Delta t$ and $\mathbf{B}(x) \Delta t$ its covariance, while $\mathbf{Z}$ is a vector of independent random numbers drawn from a normal deviate (with zero mean and unit variance). As for the Langevin equation as discussed above, the multi-Gaussian distribution of stochastic jumps $\mathbf{B}(x) \Delta t^{1/2} \mathbf{Z}$ is inspired by the work of Einstein on Brownian motion and molecular diffusion. Since the vector $\mathbf{Z}$ is drawn independently for each particle at each time step, the particle moves randomly while following a general direction as dictated by vector $\mathbf{A}(x) \Delta t$. This feature renders the name random walk to the method.

The first numerical implementations of RW in subsurface hydrology for solving the ADE were performed by Ahlstrom et al. (1977) and Prickett et al. (1981). They used a simplified RW based on the equivalence of FPKE and ADE as given by Eq. [13]. This yields the algorithm

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \mathbf{u}[\mathbf{x}(t)] \Delta t + \left[ 2\mathbf{D}[\mathbf{x}(t)] \Delta t \right]^{1/2} \mathbf{Z}$$

in which $\mathbf{u}[\mathbf{x}(t)]$ and $\mathbf{D}[\mathbf{x}(t)]$ are the mean pore-water velocity vector and the dispersion tensor, respectively, at location $\mathbf{x}(t)$. This standard equation is often referred to as the discrete form of the Itô (1951) stochastic differential equation. An alternative way of computing the dispersive term is given by Stratonovich (1966), where dispersion is taken at an intermediate location, that is:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \mathbf{u}[\mathbf{x}(t + \Delta t/2)] \Delta t + \left[ 2\mathbf{D}[\mathbf{x}(t + \Delta t/2)] \Delta t \right]^{1/2} \mathbf{Z}$$

The Advection Step

Because of the random displacement of the particles, the velocity components must be known everywhere. The velocity field is usually obtained with a numerical flow model based on a finite volume or finite element method, and calculated on edges or faces of elements or at the element centroids. Various interpolation schemes to obtain the velocity at any location were compared by Goode (1990), Semma (1994), and LaBolle et al. (1996). Schafer-Perini and Wilson (1991) suggested that the interpolation scheme should preserve a null divergence of the velocity field. When a two-dimensional finite volume scheme with rectangular elements was used to compute the heads, they showed that fluid mass balance would be preserved only when the following interpolator within one element is used

$$u_x = a + b(x - x_0); \quad u_y = c + d(y - y_0)$$

where $x$ and $y$ now are the two spatial coordinates, and
The constants $b$ and $d$ correspond to the velocity gradient into the element along the $x$ and $y$ directions. Note that the method can be extended to three-dimensional flow and parallelepiped elements by simply adding the velocity in the $z$ direction at $z_0 = (i, j, k - 1/2)$ and the velocity gradient along $z$. For a particle initially located at $(x_0, y_0)$, integration of Eq. [37] gives its new location:

\[
\begin{align*}
x - x_0 &= \frac{a + b(x_t - x_0)}{b} \exp(b\Delta t) - \frac{a}{b} \\
y - y_0 &= \frac{c + d(y_t - y_0)}{d} \exp(d\Delta t) - \frac{c}{d}
\end{align*}
\]  

[39]

after time step $\Delta t$ (Pollock, 1988; Schafer-Perini and Wilson, 1991). Of course, the algorithm needs to be slightly modified when the time step $\Delta t$ is fixed and the particle presumably leaves the element within the time step. The time step is then split such that when the particle reaches an element boundary, the exit point coordinates are stored, the constants for the velocity interpolation are updated with values calculated from the characteristics of the neighboring element, and the rest of the time step is allotted to motion of the particle within the neighboring element (e.g., see Pollock, 1988, for details of the algorithm). This interpolation is consistent with the scheme used in the mixed finite element method for rectangular elements (e.g., Hoteit et al., 2002), except that the scheme with the mixed finite element is given a priori. The method is therefore self-consistent, and no additional assumptions are required for interpolating the velocities, unlike with finite volume and finite difference methods.

**The Dispersion Step**

As shown by many authors (e.g., Ackerer and Kinzelbach, 1985; Uffink, 1987; Kinzelbach, 1988; Tompson and Gelhar, 1990; Kinzelbach and Uffink, 1991) and as discussed above when we established the equivalence between FPKE and ADE, the RW algorithm used to mimic the ADE requires a correction term that depends on the derivatives of the dispersion coefficients when the dispersion tensor varies significantly in space. Equivalence between the mean displacement of the FPKE and the modified velocity of the ADE (see Eq. [14]) allows algorithm [35] to be rewritten as

\[
x(t + \Delta t) = x(t) + \left[ u(x(t)) + \frac{\partial D[x(t)]}{\partial x} \right] \Delta t \\
+ (2D\delta[x(t)]\Delta t)^{1/2} Z
\]  

[40]
where the coefficient $\beta$ serves to modify the location of the source to account for the fact that diffusion is calculated for a homogeneous medium $\gamma$, whereas in reality diffusion occurs first in medium $\lambda$. Considerations on continuity at the interface (see Appendix) enable $R$ and $\beta$ to be calculated as

$$R = \frac{\sqrt{D_x} - \sqrt{D_y}}{\sqrt{D_x} + \sqrt{D_y}}; \quad \beta = \frac{\sqrt{D_x}}{\sqrt{D_y}}$$  \hspace{1cm} [43]

Semra et al. (1993) suggested the following numerical procedure to reflect particles. When a particle may cross the interface during a certain time step $\Delta t$, its displacement is split into two parts. During the first time step $\Delta t$, the particle moves to the interface, and during the second time step $\Delta t (=\Delta t - \Delta t_i)$ the particle moves from the interface to either layer $\lambda$ or layer $\gamma$. The displacement is then computed using the transport properties of the medium into which the particle entered. To preserve mass balance when the particle is at the interface, the probabilities that the particle enters layers $\lambda$ and $\gamma$ are $P_\lambda = \sqrt{D_x}/(\sqrt{D_x} + \sqrt{D_y})$ and $P_\gamma = 1 - P_\lambda = \sqrt{D_y}/(\sqrt{D_x} + \sqrt{D_y})$, respectively. This algorithm must be used whether the particle originates from one side of the interface or the other one. It can be shown that this splitting is equivalent to the reflection principle suggested by Uffink (see Appendix).

An alternative to reflection methods is the interpolation technique. LaBolle et al. (1996) suggested to interpolate the velocities in the dispersion tensor ($\mathbf{D} = \alpha \mathbf{u}$, with $\alpha$ the dispersivity) to smooth the dispersion tensor across the interface, and hence to dampen or even completely eliminate the discontinuities. However, as stated by the authors, an unbiased solution to transport requires not only very small time steps when the particle reaches the discontinuity, but also a fine spatial discretization for the interpolation scheme. Particle displacement during each time step must be significantly smaller than the width of the transition area over which the interpolation is performed. The size of this transition area is unknown a priori, but should be large enough to depict the change in the dispersion coefficient and small enough to avoid perturbations of the transport properties on either side of the discontinuity. The interpolation technique of LaBolle et al. (1996) hence leads to small time steps and high computational costs. A smoother interpolation was used by Bunzl (2002), who suggested the following equation:

$$D = D_\lambda + \frac{D_\gamma - D_\lambda}{1 + \exp[(x_i - x)/w]}$$  \hspace{1cm} [44]

where $D_\lambda$ and $D_\gamma$ are the dispersion coefficients on both sides of the interface, $x_i$ is the location of the interface and $w$ is the width of the transition zone.

We note here that LaBolle et al. (2000) suggested still another method for handling abrupt changes in dispersion. That method is based neither on reflection nor on strict interpolation. They changed the RW Eq. [35] for this purpose into

$$x(t + \Delta t) = x(t) + u[x(t)]\Delta t + [2D(x + \delta x)\Delta t]^{0.5}Z$$  \hspace{1cm} [45]

where

$$\delta x = Z\sqrt{D[x(t)]\Delta t}$$  \hspace{1cm} [46]

LaBolle et al. (2000) suggested to evaluate dispersion at location $x + \delta x$, where the increment $\delta x$ is the dispersive step that would have been calculated in $x(t)$ without correction. Their scheme was found to compare successfully with analytical solutions (LaBolle et al., 2000; Hassan and Mohamed, 2003).

An Alternative Using Cellular Automata

Delay et al. (1996) proposed a numerical alternative for solving RW in space with particles managed as variables over a regular grid. For simplicity we assume here that a two-dimensional domain is discretized into square cells of size $\Delta x$ on a side. Our reasoning is the same for three-dimensional transport. Each cell contains a numerical variable representing the number of particles that can move by advection and dispersion during a certain time step $\Delta t$. If the RW algorithm with a uniform distribution of the dispersion jumps (Uffink, 1985) is applied to a particle located at $(x, y)$ at time $t$, its position at $t + \Delta t$ is

$$x(t + \Delta t) = x(t) + u_x \Delta t + \delta x;$$

$$y(t + \Delta t) = y(t) + u_y \Delta t + \delta y$$  \hspace{1cm} [47]

where $(\delta x, \delta y)$ represents a random displacement within a dispersion rectangle of size $2(6D_x\Delta t)$, centered at $[x(t) + u_x\Delta t, y(t) + u_y\Delta t]$ and with its principal axis along the direction of the fluid velocity $u$. We used $D_x$ and $D_y$ for the longitudinal and transverse dispersion coefficients, respectively. Let us assume that all particles in cell $j$ are assembled at its center $(x, y)$ and represented by a single variable $n_j$; their number in cell $j$. The jump $t \rightarrow (t + \Delta t)$ will spread the particles uniformly over cells $k$ that overlap the dispersion rectangle defined above. Each cell $k$ receives a number of particles corresponding to the number $n_i$ of particles in cell $j$ multiplied by the ratio of the area of cell $k$ within the dispersion rectangle to the total area of the dispersion rectangle. With this method, all advection–dispersion jumps during time step $\Delta t$ are performed sequentially, with $j$ varying from 1 to the number of cells in the domain.

A classical scenario is as follows: the jump of particles from cell $\lambda$ is performed before the jumps from cell $\gamma$ while cell $\lambda$ spreads its particles over $\gamma$. To avoid counting these particles in the jumps of cell $\gamma$, which would result in the anomaly of having two displacements during one time step, the number of dispersed particles that arrive in a cell are first stored and added to form an intermediate variable. When all jumps have been completed, the particles of the intermediate variable are assigned cell by cell to the variable representing the mean concentration that will move during the next time step. This enhanced particle tracking (EPT) scheme updates at each time step $N$ variable ($N$ being the number of cells) instead of moving particles individually and is therefore faster than classical RW methods. Moreover, as long as the velocity field remains constant in time, the jumps from each cell keep the same settings of the particle distribution. These settings in each cell (i.e., the neighbor-
Their time and space position. For a jump at velocity computation is improved with particles holding both media (e.g., de Dreuzy et al., 2001a, 2001b, for fractured few jumps are necessary for high velocity situations. These scale for the connectivity and the flow properties of such numerous jumps for low velocity areas, whereas only a few jumps may become very time-consuming. Several theoretical studies have shown the absence of a homogenization which the velocity field is depicted. This results in numerous jumps for low velocity areas, whereas only a few jumps are necessary for high velocity situations. The computation is improved with particles holding both their time and space position. For a jump at velocity \( u_j \), the local time step \( \delta t \) is adjusted so that the jump \( u_j \delta t \) is constant for any \( u_j \) (e.g., \( u_j \delta t = \Delta x/2 \)). After the jump (to which is also added a classical random motion due to dispersion), the location of the particles and their resting time are updated.

With the EPT method using dispersion rectangles, a constant displacement scheme may be generated as follows. An elementary time step \( \delta t \) is defined such that for the maximum velocity of the flow field, \( u_{max} \delta t \) is for instance equal to \( \Delta x/2 \). Suppose that a cell \( j \) with velocity \( u_j \) requires \( n_j \) elementary time steps for a jump \( u_j \delta t = \Delta x/2 \). A series of \( n_j \) intermediate variables is assigned to the cell and managed by means of a “first in–first out” sequence or queue. For each elementary time step \( \delta t \), the standard variable corresponding to the moving particles is spread over the other cells within a dispersion rectangle which size is given by Eq. [47] for a time step \( n_j \delta t \). The outlet of the queue is then flushed into the standard variable, while the ranking of the particles in the queue are decreased such that particles that arrive from other cells by advection–dispersion are stored at the inlet of the queue. Proceeding in this manner enables the particles to be stored during \( n_j \delta t \) in the cell, and then to be moved by advection over an almost constant distance irrespective of the velocity in the cell. As compared with the algorithm of Wen and Gomez-Hernandez (1996), more jumps may be required since the EPT moves particles at each elementary time step \( \delta t \). However, the aforementioned principle of moving sets of particles with a table that stores the particle distribution remains unchanged and saves considerable computational time. The EPT method is rapid, even with a constant displacement scheme, while the use of constant jumps avoids the atavistic problems of mass conservation at the interfaces between layers with contrasting dispersion properties.

A similar approach for moving groups of particles was described by Vamos et al. (2003). The particles belonging to one cell are for this purpose first gathered at its center. They are subsequently moved to a neighboring cell following a Bernoulli distribution, depending on the advective fluxes between the starting cell and its neighbors. As stated by the authors, random fluctuations are damped, but the scheme is equivalent to standard finite differences for rectangular cells, and therefore often generates artificial numerical diffusion.

**APPROACHES TO RANDOM WALK IN THE TIME DOMAIN**

In the preceding section we showed that classical RW methods are based on the motion of particles in space. They are displaced during each time step over the domain according to local properties of the flow field (see e.g., Eq. [40]). In very heterogeneous media such as fractured rocks and macroporous soils where flow occurs in both the matrix and the fractures, RW calculations may become very time-consuming. Several theoretical studies have shown the absence of a homogenization scale for the connectivity and the flow properties of such media (e.g., de Dreuzy et al., 2001a, 2001b, for fractured networks). Very heterogeneous media are often modeled as discrete networks involving tens of thousands to even a few millions bonds (Bour and Davy, 1997; Adler and Thovert, 1999; Rivard and Delay, 2004). In addition, flow velocities may span several orders of magnitude, which imposes very small time steps for accurate sampling of the flow field using a classical RW. A large number of jumps are required in low velocity areas to move the particles significantly, which may cause the calculations to become inefficient in terms of computational costs. These constraints have motivated the development of new approaches based on random particles managed in the time domain. There are two reasons for this. One is to provide a framework for describing non-Gaussian transport often observed in field and laboratory studies (Scher et al., 2002; Bromly and Hinz, 2004; Cortis and Berkowitz, 2004). A second reason is to produce efficient and rapid methods for calculating transport problems, for example for wide bond networks, while accounting for processes other than pure advection.

One assumption of the use of time domain approaches applied to subsurface hydrology is that the flow field can be represented by a set of spatially distributed sites connected by bonds that ensure the transition of particles between sites. This is an efficient abstraction for fracture networks (e.g., Dershowitz and Fidelibus, 1999) and for heterogeneous porous media, which may be viewed as regular percolation networks with varying hydraulic conductivities in the bonds (e.g., Sahimi and Mukhopadhyay, 1996). The particles are also assumed to follow a
Markovian process in that there is no correlation between successive transitions from one site to the other.

The Continuous Time Random Walk

The continuous time random walk (CTRW) theory for lattices was first developed by Montroll and Weiss (1965), Montroll and Scher (1973), and Scher and Lax (1973) for problems involving electron hop conductivities in semiconductors. The theory was recently enhanced to anomalous transport in fracture networks (Berkowitz and Scher, 1998) and more generally to groundwater hydrology in heterogeneous media (Berkowitz et al., 2001; Berkowitz and Scher, 2001). Contrary to classical RW moving particles through a given domain, CTRW is an up-scaled approach that does not use particles explicitly. The theory is based on calculations of \( P(s, t) \), the probability density for a particle to be at site \( s \) at time \( t \). The aim is to depict the macroscopic behavior of the system, with the first abstraction considering that the real domain can be averaged by sites located on a regular lattice with \( N \) sites on a side and involving periodic boundary conditions. To avoid any loss of generality, the jumps are not limited to adjacent sites, which means that particles can move from site \( s' \) to site \( s \) with a distance \( s - s' \) spanning all sizes available in the lattice. This is an important feature when CTRW is applied to random fracture networks where the distance between adjacent sites at fracture intersections may vary from almost zero to the size of the network.

If \( R(s, t) \) designates the probability density for a particle to arrive at site \( s \) and time \( t \), and \( \Psi(d, t) \) is the probability density of making a jump of size \( d \) of duration \( t \), the Markovian process followed by the particles allows us to write

\[
R(s, t) = \sum_{s'} \left[ \Psi(s - s', t - \tau) R(s', \tau) d\tau + \delta(s - 0) + \delta(t - 0) \right] \tag{48}
\]

which is the probability of just arriving at \( s' \) at time \( \tau \) multiplied by the probability of moving from \( s' \) to \( s \) within time \( t - \tau \). This local probability is extended to all possible site \( s' \) and each time \( \tau \) between 0 \( \rightarrow \) \( t \). The \( \delta(s - 0) \) and \( \delta(t - 0) \) Dirac-delta functions represent the initial conditions to ensure that the probability for a particle to be in 0 at time 0 is one. We emphasize that the form of Eq. [48] does not presume any particular transport process in the bonds. Potentially several processes can occur simultaneously, but they need to be merged in the probability \( \Psi(d, t) \). In other words, small-scale transport details are supposed to be encapsulated in \( \Psi(\cdot) \) before averaging them when calculating \( P(\cdot) \).

The probability \( P(s, t) \) can be written as:

\[
P(s, t) = \int_0^t \left[ \phi(t - \tau) R(s, \tau) d\tau \right], \tag{49}
\]

which is the probability of arriving at site \( s \) at time \( \tau \) between 0 \( \rightarrow \) \( t \), multiplied by the probability \( \phi(t - \tau) \) of staying in place during \( t - \tau \). The variable \( \phi(t) \) in Eq. [49] may be expressed as the probability that no jump occurs during \( t \):

\[
\phi(t) = 1 - \int_0^t \sum_d \Psi(d, \tau) d\tau, \quad \text{i.e.,}\]

\[
\phi(t) = 1 - \int_0^t \Psi(\tau) d\tau \quad \text{with} \quad \Psi(\tau) = \sum_d \Psi(d, \tau) \tag{50}
\]

Note that \( \Psi(\tau) \) is the probability of doing a jump of any size and of duration \( \tau \). Equations [48] and [49] are convolution products in time and space. Their calculation can be simplified if Laplace–Fourier transforms are used. For two functions \( f(t) \) and \( g(s) \), the Laplace and discrete Fourier transforms are defined as

\[
f_L(p) = L[f(t)]_p = \int_0 \left[ e^{-pt} f(t) \right] dt \]

\[
g_L(w) = F_L[g(s)]_w = \sum_s e^{-iw} g(s) \tag{51}
\]

The Laplace and Fourier transforms of convolution products are the simple algebraic products of the Laplace and Fourier transforms. Therefore, in using the properties \( L[\delta(t)] = 1 \) and \( F_L[\delta(s)] = 1 \), Eq. [48] yields

\[
R_{FL}(w, p) = 1 + \Psi_{FL}(w, p) R_{FL}(w, p), \quad \text{i.e.,}
\]

\[
R_{FL}(w, p) = \frac{1}{1 - \Psi_{FL}(w, p)} \tag{52}
\]

Integration Eq. [50] by parts and taking the Laplace transform of \( \phi(t) \) leads to

\[
\phi_L(p) = \frac{1}{p} \left[ 1 - \Psi_L(p) \right] \tag{53}
\]

where \( \Psi_L(p) \) is the Laplace transform of \( \Psi(\tau) = \Sigma_d \Psi(d, \tau) \). The probability density \( P(s, t) \) given by Eq. [49] can be calculated as the inverse Fourier transform of \( P_L(w, t) \) as follows:

\[
P(s, t) = \frac{1}{N^2} \sum_w e^{iw} g_L(w, t) \]

\[
g_L(w, t) = L^{-1}\left[ R_{FL}(w, p) \phi_L(p) \right]_p \]

\[
= L^{-1}\left[ \frac{1 - \Psi_L(p)}{p[1 - \Psi_{FL}(w, p)]} \right]_p \tag{54}
\]

in which \( L^{-1} \) is the inverse Laplace transform, \( N \) the number of sites, and \( e \) the dimension of Euclidean space. \( P(s, t) \) expresses the evolution in time of the concentration plume at location \( s \). Another probability density relevant to solute transport is that of the first arrival time at a location \( s \). This probability is equivalent to a breakthrough curve that would be sampled at the outlet \( s \) of a laboratory column in response to a Dirac injection at the inlet. Under these conditions and for transitions \( 0 \rightarrow s \), we have

\[
R(s, t) = \int_0^t F(s, \tau) R(0, t - \tau) d\tau \tag{55}
\]

where \( F(s, t) \) is the probability of arriving for the first time at \( s \) at time \( t \). Note here that Dirac-delta function \( \delta(\cdot) \) are not required for initial conditions since they are already in \( R(0, t) \) (see Eq. [48]). Equation [55] expresses the fact that a particle just arriving at \( s \) at \( t \) may have visited the site for a first time at \( \tau \), and then may have
moved back and forth during \( t - \tau \). In the Laplace domain, one obtains

\[
F_\tau(s, p) = \frac{R_\tau(s, p)}{R_\tau(0, p)} \quad [56]
\]

Assuming that the Laplace and Fourier inversions can be calculated either numerically or analytically, the central notion of the CTRW theory is the probability density \( \Psi(d, t) \) of doing a jump of size \( d \) during time \( t \). As stated above, \( \Psi(t) \) is assumed to enclose small-scale details of transport mechanisms in the real domain. For transport by advection in two-dimensional fracture networks, Berkowitz and Scher (1998) decomposed this probability as follows:

\[
\Psi(d, t) \propto K\Phi(u)p(du)f(u) \quad [57]
\]

where \( K \) is a normalization constant, \( \Phi(u) \) is the probability density of the fluid velocity \( u \) in the bonds, and \( p(du) \) is the conditional probability of a transition of size \( d \) knowing that it occurs at velocity \( u \). In other words, since \( t = du/\rho \), \( p(du) \) is the probability of a transition of size \( d \) knowing that it occurs within time \( t \). Also, \( f(u) \) in Eq. [57] is the probability for a particle to experience the velocity \( u \) in the network. Assuming perfect mixing of mass fluxes at bond intersections, this probability at the scale of the network is the ratio of the total flow rate through bonds with velocity \( u \) to the total flow rate in the network.

Berkowitz and Scher (1998) performed calculations of flow through several realizations of two-dimensional random fracture networks having exponential law length distributions [i.e., the number of fractures of length \( l \) are given by \( n(l) \propto \exp(-l/\lambda) \)]. They found the best fits for \( p(du) \) and \( \Phi(u) \) to be

\[
\Phi(u) \propto u^{1+\beta}\exp(-u/\lambda_0); \quad p(du) \propto d^{0.5}\exp(-d/\lambda_0) \quad [58]
\]

where \( \lambda_0, \lambda_0 \), and \( \beta \) are fitting parameters. Whereas \( \lambda_0 \), and \( \lambda_0 \) are the characteristic distance and velocity that normalize the probability densities, \( \beta \) is a key parameter that expresses the typical length of fracture networks to exhibit strong asymmetry in the velocity distribution. With \( u = du/\rho \), Eq. [57] becomes

\[
\Psi(d, t) \propto C_\tau^{d^{1+\beta}\exp\left[-d\left(\frac{1}{\lambda_0 u_0} + \frac{1}{\lambda_0}\right)\right]} \quad [59]
\]

Introducing Eq. [59] into Eq. [54] and [56] with a value of \( \beta \) in the range 0.5 to 0.9, Berkowitz and Scher (1998) demonstrated the persistent nature of anomalous (non-Gaussian) behavior of solute plumes and breakthrough curves, even for long residence times. For \( \beta \) values of approximately 0.5, the plumes were highly asymmetric with a peak remaining close to the injection point and a flat forward front due to only a few particles experiencing very rapid transitions. On the other hand, the breakthrough curves showed flat tails for large residence times. For \( \beta \) values of about 0.8 to 0.9, the peak moved away from the injection point, with a forward tail in space becoming thinner and shorter than the backward tail. The centers of mass of the plumes along the main flow direction \( \bar{x}(t) \), and their standard deviations \( \sigma_x(t) \), were shown to remain proportional to \( t^{0.5} \). Thus, the ratio \( \bar{x}(t)/\sigma_x(t) \) remained almost constant whereas Gaussian transport should have resulted in \( \bar{x}(t)/\sigma_x(t) = t^{0.5} \). This difference is the signature of anomalous transport. The CTRW hence appears to be a promising theory for upscaling the transport properties of very heterogeneous media. In the initial work by Berkowitz and Scher (1998), advection was considered to be the only transport process in the bonds. However, as stated above, the method remains versatile and can handle additional processes provided their effects can be modeled and included into the probability density \( \Psi(d, t) \) (e.g., Margolin et al., 2003).

**Time Domain Random Walk**

Another method for managing the transport of particles in time was proposed by Banton et al. (1997). Their time domain random walk method (TDRW) was generalized by Delay and Bodin (2001) to provide an efficient and rapid method for explicitly calculating transport problems over bond networks, while accounting for several processes in the bonds such as advection, dispersion, matrix diffusion, radioactive decay, and adsorption assuming instantaneous local equilibrium. We will focus here on methods for solving advection–dispersion in a bond subject to matrix diffusion. The medium is again viewed as a set of sites connected by one-dimensional bonds, but without additional simplifications since TDRW is based on the explicit displacement of particles between adjacent sites and the calculation of the transition times in each bond within one step.

Recall the Fokker–Planck–Kolmogorov formalism for one-dimensional advection–dispersion equation in free water:

\[
\frac{\partial C}{\partial t} + \frac{\partial}{\partial x}\left(\frac{u + \frac{D}{\rho}}{\partial x}C\right) - \frac{\partial^2}{\partial x^2}(DC) = 0 \quad [60]
\]

where \( C \) is the solute concentration (M L\(^{-1}\)), \( u \) the fluid velocity (L T\(^{-1}\)), and \( D \) is the dispersion coefficient (L\(^2\) T\(^{-1}\)). In the Lagrangian framework, the material derivative of any scalar \( \lambda \) corresponds to its partial derivative with respect to time. Thus, along a streamline at velocity \( u \), the material derivative can be written as

\[
\frac{\partial \lambda}{\partial x} = \frac{d\lambda}{dt} = \frac{\partial \lambda}{\partial t} \frac{d\rho}{dt} \quad [61]
\]

Applying this change of variables to Eq. [60] yields

\[
\frac{\partial C}{\partial x} + \frac{1}{u} \frac{\partial}{\partial t}\left(\frac{u + \frac{D}{\rho}}{\partial x}C\right) - \frac{1}{u^2} \frac{\partial^2}{\partial x^2}(DC) = 0 \quad [62]
\]

Equation [62] may be reinterpreted within a stochastic context of random moving particles. Assuming a constant dispersion in the bond (\( aD/\partial x = 0 \) in Eq. [62]), this interpretation views particles moving through a bond of length \( d \) characterized by a transition time distribution between the inlet and the outlet of the bond with mean \( m_i \), and a variance \( \sigma^2_t \) as follows:

\[
m_i = \frac{d}{u}; \quad \sigma^2_t = \frac{2Dd}{u^2} \quad [63]
\]

The time distribution is assumed lognormal (as reported also from numerical experiments involving pure diffu-
sion by James and Chryssikopoulos, 2001), which can be verified as follows. In a one-dimensional semi-infinite homogeneous medium, the probability density function of travel times for advective-dispersive transport over distance \( d \) is

\[
f(d, t) = \frac{d}{(4\pi D t^2)^{3/2}} \exp\left(-\frac{(d - ut)^2}{4Dt}\right) \tag{64}
\]

The mean and variance of travel times calculated with Eq. [64] are \( dtu \) and \( 2Ddtu^2 \), respectively. The skewness coefficient \( \gamma_1 \) and the kurtosis coefficient \( \gamma_2 \) are given by

\[
\gamma_1 = \frac{(\mu_3)}{(\mu_2)^{3/2}} = \frac{18}{Pe}, \quad \gamma_2 = \frac{\mu_4}{(\mu_2)^2} - 3 = \frac{30}{Pe} \tag{65}
\]

in which \( Pe = u/dD \) is the bond-Peclet number, and

\[
\mu_3 = \int_0^\infty f(d, t)(t - \bar{t})^3 dt, \quad \mu_4 = \int_0^\infty f(d, t) (t - \bar{t})^4 dt
\]

On the other hand, a lognormal distribution of mean \( m = dtu \) and variance \( \sigma^2 = 2Ddtu^2 \) yields

\[
\gamma_1 = \frac{\sigma^2}{m} \left(3 + \frac{\sigma^2}{m^2}\right) = \frac{18}{Pe} + \frac{24}{Pe^2} + \frac{8}{Pe^3};
\]

\[
\gamma_2 = \frac{\sigma^2}{m^2} \left[16m^6 + 15m^4\sigma^2 + 6m^2(\sigma^2)^2 + (\sigma^2)^3\right]
\]

\[
= \frac{32}{Pe} + \frac{60}{Pe^2} + \frac{48}{Pe^3} + \frac{16}{Pe^4} \tag{66}
\]

The lognormal assumption is therefore accurate to the first order in \( 1/Pe \), with numerical tests showing that Eq. [66] holds for bond \( Pe > 5 \), which is the case for most transport problems involving fractures. Numerical tests were performed over random two- and three-dimensional lattices with power-law length distributions for the bonds. When a nonnegligible fraction of bonds has Peclet numbers <5, this fraction was found to enclose very small bonds with high fluid velocities. These bonds do not influence large-scale solute transport in terms of breakthrough curve behavior and the very slight discrepancies from TDRW are negligible. Note, however, that these small bonds may have an important effect on the network connectivity, and removing them from the network would be a severe error. The transition time \( t \) of a particle between the inlet and the outlet of a bond of length \( d \) can be calculated as

\[
t = \exp(m_{log} + z\sigma_{log}); \quad m_{log} = \log \left[\frac{m_0}{\sqrt{1 + (\sigma^2/m^2)}}\right];
\]

\[
\sigma_{log} = \log[1 + (\sigma^2/m^2)] \tag{67}
\]

where \( m_0 \) is the mean and \( \sigma^2 \) the variance (Eq. [63]), and \( z \) is a random number drawn from a normal deviate of mean zero and variance unity. Note that Reimus and James (2002) proposed an exact analytical solution for the transition time distribution over a space step \( \Delta x \). Their solution is based on a classical infinite series solution typical for diffusion in finite media (e.g., Carslaw and Jaeger, 1959). While this solution is very close to a lognormal distribution advocated by Delay and Bodin (2001) and James and Chryssikopoulos (2001), the probability density of transition times is best be tabulated to make the algorithm computationally efficient. However, comparisons with the algorithm of Eq. [67] have shown the latter much faster (Delay, 2004, unpublished computations). A time distribution could also be drawn directly from the inversion of Eq. [64], that is, extract a time \( t \) from the right-hand side of Eq. [64] with a left-hand side equal to a random number uniformly distributed between 0 and 1. Unfortunately, the inversion is not straightforward, time-consuming and may result in partial indetermination. A transport model at the scale of the entire network is built by following the particles and incrementing their travel times according to Eq. [64] using the bond properties \( (d, u, D) \) between adjacent connected sites. Mixing rules such as those proposed by Park and Lee (1999) may be used to calculate the transition probabilities from bond to bond at their intersections.

Matrix diffusion can be accounted for by calculating a diffusion time spent by the particle in the matrix attached to each bond. Consistent with an analytical solution proposed by Malowszewski and Zuber (1985) for advection in a single fracture with simultaneous diffusion in an infinite matrix, let \( t_{dif} \) be the equivalent diffusion time of a particle in the matrix. The probability density \( (T^{-1}) \) of transition times by advection and diffusion is

\[
P_{+\text{dif}}(d, t) \propto a \frac{\mu_0}{\sqrt{\pi}} \left(\frac{t}{t_0} - 1\right)^{-1} \exp\left[-\frac{a^2t_0}{(t/t_0)^2}\right]
\]

\[
t > t_0; \quad t_0 = \frac{d}{u}; \quad a = \frac{\phi_0 D_e}{2b} \tag{68}
\]

in which \( d \) is the length of the bond, \( 2b \) bond aperture (L), \( u \) fluid velocity (L T\(^{-1}\)), \( \phi_0 \) the matrix porosity, and \( D_e \) the apparent diffusion coefficient in the matrix (L\(^2\) T\(^{-1}\)). The time spent by a particle in the matrix because of diffusion can be derived from the cumulative probability density of Eq. [68]; that is,

\[
\int_{t_0}^t P_{+\text{dif}}(d, t) dt = \int_{t_0}^t P_{+\text{dif}}(t) dt = \text{erfc}\left[a(t_0)/(t - t_0)^2\right]
\]

Since \( t_0 = dtu \) is the advection time in the bond, the time \( t_{dif} \) is

\[
t_{dif} = t - t_0 = \left[\frac{at_0}{\text{erfc}^{-1}(\omega)}\right]^2 \tag{69}
\]

where \( \omega \) is drawn from a uniform distribution between 0 and 1. The time \( t_{dif} \) cannot be added to the time by advection–dispersion without modifying Eq. [67] since matrix diffusion delays the particles as compared with what is expected for transport only in the fracture at velocity \( u \). Assume that \( u_{ap} = u/R \) (R > 1) is the apparent velocity in the bond. The operator in the time domain of Eq. [67] must be corrected for the fluid velocity to generate transport by advection at velocity \( u \), despite the presence of an apparent velocity, \( u_{ap} \). This is done by assuming that advection occurs at a velocity \( u^* = u/R \), which then compensates for the delay stemming from \( u_{ap} \), while keeping the dispersion operator unchanged. The mean of the distribution \( m \) given by Eq. [63] is therefore re-
placed by \( m^* = d\mu^* = d\mu R \). The retardation factor \( R \) is defined as the ratio \( t_b/t_i \), where \( t_i \) is a characteristic time by advection–plus-diffusion. Given the probability density of Eq. \([68]\), \( t_i \) is calculated as

\[
t_i = \int \left[ P_{x,\omega}(x, t) \right] dt / \int \left[ P_{x,\omega}(x, t) \right] dt
\]

where \( t_0 + \tau \) is the time beyond which matrix diffusion has no longer any influence on advection–dispersion in the bond. With these assumptions, the retardation factor becomes

\[
R = 1 + 2\sqrt{\frac{\omega}{\tau}} \left[ a_0 - \frac{\sqrt{\omega}}{\tau} \right] \sqrt{\frac{\omega}{\tau}} \exp \left( \frac{a_0 \sqrt{\omega}}{\tau} \right) \text{erf} \left( \frac{a_0 \sqrt{\omega}}{\tau} \right)
\]

Delay and Bodin (2001) showed that \( \tau \) could be given by \( \sigma = (2Dd\mu^*)^{1/2} \); that is, the standard deviation of transition times by advection–dispersion in the bond.

The advantages of the TDRW method were shown by means of comparisons with analytical solutions at the scale of both a single fracture and small synthetic networks (Bodin et al., 2003). Because the TDRW method calculates in one step the transition time of a particle in each bond of a network, the approach is much faster than classical RW methods. Delay and Porel (2003) showed another advantage of random moving particles managed in time in inverse problems are involved. Optimization techniques commonly are based on gradient-like techniques. The aim is to seek an optimal set of parameters that provides the best match between a simulated transport scenario and available measured data. Deviations between observation and simulation are measured by means of a vector \( \xi(n) \) of sampled errors \( \xi = C^{\text{sim}}(x, t_b) - C^{\text{obs}}(x, t_b) \). Let \( \mathbf{p}^k \) be a vector at iteration \( k \) of the \( m \) parameters to be sought (e.g., velocities and dispersion coefficients from subareas of the modeled domain). At iteration \( k \), the optimization process tries to minimize a closeness criterion between simulation and observation of the form:

\[
F(\mathbf{p}^k) = \frac{1}{2} \xi^T \mathbf{W} \cdot \xi^k
\]

where \( T \) is the transposition operator and \( \mathbf{W} \) an \( n \times n \) matrix. Equation \([71]\) is often referred to as a weighted least-square objective function, in which the matrix \( \mathbf{W} \) enables one to weigh each term \( \xi_i^2 \) and \( \xi_i \), of the objective function. In the simplest case, \( \mathbf{W} \) is merely a diagonal matrix or identity matrix. Since \( F \) depends on \( \mathbf{p}^k \), this function can be expanded for \( \mathbf{p}^{k+1} \) as follows:

\[
F^{k+1} = F^k + (\mathbf{p}^{k+1} - \mathbf{p}^k)^T \cdot \nabla F^k + \frac{1}{2} (\mathbf{p}^{k+1} - \mathbf{p}^k)^T \cdot \nabla^2 F^k \cdot (\mathbf{p}^{k+1} - \mathbf{p}^k)
\]

in which \( \nabla F \) is the gradient vector (\( \partial F/\partial \mathbf{p} ; i = 1 \ldots m \)), and \( \nabla^2 F \) the Hessian matrix (\( \partial^2 F/\partial \mathbf{p} \partial \mathbf{p} ; i, j = 1 \ldots m \)). Algorithms based on gradient methods assume that \( \nabla F^{k+1} = 0 \); that is, \( F^{k+1} \) reaches a minimum at \( \mathbf{p}^{k+1} \), which on taking the derivative of Eq. \([72]\) yields

\[
\mathbf{p}^{k+1} - \mathbf{p}^k = - (\nabla^2 F)^{-1} \cdot \nabla F^k
\]

It has been shown (e.g., Tarantola, 1987) that the more accurate the Jacobian, the more local minima are avoided and the more the inverse problem converges rapidly. Thus, the best way to proceed is to obtain analytical expressions for the terms \( \partial^2 C/\partial \mathbf{p} \). This is possible with RW in time. Let \( t_i \) be the transition time in a bond as given by Eq. \([67]\). The derivative of this transition time with respect to a parameter \( p \) is given by

\[
\frac{\partial t_i}{\partial p} = t_i \frac{\partial \log t_i}{\partial p} = t_i \left( \frac{\partial m_{\text{out}}}{\partial p} + z \frac{\partial a_{\text{out}}}{\partial p} \right)
\]

where \( m_{\text{out}} \) and \( a_{\text{out}} \) are the mean and standard deviation of \( \log t_i \) as defined in Eq. \([63]\) and \([67]\), respectively. The derivatives \( \partial m_{\text{out}} / \partial p \) and \( \partial a_{\text{out}} / \partial p \) can be obtained analytically (e.g., see details for heterogeneous diffusion in Delay and Porel, 2003), and hence it is not necessary to redraw the Gaussian random number \( z \). The calculation is therefore very fast and \( t_i \), as well as its derivatives with respect to all parameters \( p \), are calculated simultaneously. Assume now that the particle after several transitions is a fraction of mass of the simulated concentration \( C^k \) at time \( t_i \). Since the successive transitions of the particles are independent, the derivative of its arrival time \( t \) at any location is the sum of all transition time derivatives: \( \partial t_i / \partial p = \sum \partial t_i / \partial p \). If \( \text{Nb} \) particles contribute to the concentration \( C_n \), with each particle having its own arrival time \( r_n \), where \( n \) is the particle index, then the average derivative of \( t_i \) may be written as

\[
\langle \partial t_i / \partial p \rangle = \frac{1}{\text{Nb}} \sum_{n=1}^{\text{Nb}} \partial t_i / \partial p
\]

Knowing this average derivative, one can approximate the term of the Jacobian matrix by

\[
\frac{\partial C_n}{\partial p} = \frac{\partial C_n}{\partial t_i} \langle \partial t_i / \partial p \rangle ; \quad \text{with} \quad \frac{\partial C_n}{\partial t_i} = C_{n+1} - C_n
\]

This approximation is accurate, even when a finite difference scheme is used for \( \partial C_n/\partial t_i \); this is because \( C_n \) values are simulated concentrations available for small lag times \( t_{n+1} - t_n \). This analytical calculation of the Jacobian matrix for inversion problems has proved to be very powerful, as compared with approximations of the Jacobian terms by means of perturbations \( \partial C/\partial p = (C(p + \delta) - C(p))/\delta \). The inversion converges more rapidly and there is no need to recalibrate \( m \) times the direct problem at each iteration of convergence. To our knowledge, nothing similar has been attempted with classical RW approaches in space. Nevertheless, the same concept could be employed in space, except for using \( \partial C/\partial p = \partial C/\partial p \).
with the average derivative of the motion of the particles \( \frac{d}{dt} \phi \) obtained analytically from derivatives of the RW equations in space. Note, however, that \( \frac{d}{dt} \phi \) is the derivative of the pathway lengths experienced by the particles and not of their location at time \( t \). The question remains whether or not these lengths are really needed to address the sensitivity of the model to the parameters (which is the physical notion of the Jacobian matrix). The intuitive answer is positive, but additional studies and numerical tests may be needed to obtain more definite answers.

## MODELING REACTIVE TRANSPORT WITH RANDOM WALK

### Instantaneous Linear and Nonlinear Processes

Instantaneous processes often grouped together under the so-called local equilibrium assumption (LEA) commonly comprise simplified models for mimicking adsorption of a solute onto the solid matrix (e.g., Weber et al., 1991). The corresponding mathematical formulation is given by

\[
\frac{\partial C}{\partial t} + \frac{\partial}{\partial x} (D \cdot \nabla C) - u \cdot \nabla C = \frac{\partial}{\partial x} (\frac{\phi}{K_c}CN) \tag{77}
\]

where \( C \) is the solute concentration in the fluid (\( \text{M} \cdot \text{L}^{-3} \)), \( S \) the adsorbed mass per mass of solid (\( \text{M} \cdot \text{M}^{-1} \)), \( \rho \) the dry bulk density of the porous material (\( \text{M} \cdot \text{L}^{-3} \)), \( \phi \) the porosity, \( D \) the dispersion tensor (\( \text{L}^2 \cdot \text{T}^{-1} \)), \( u \) the pore velocity vector (\( \text{L} \cdot \text{T}^{-1} \)), \( K_c \) the sorption coefficient (\( (\text{L}^2 \cdot \text{M}^{-1}) \)), while \( n \) relates to sorption nonlinearity. The expression for \( S \) in Eq. [77] describes the Freundlich isotherm often used in subsurface solute transport studies (e.g., Sposito, 1984; Weber et al., 1991; Kasteel et al., 2002). Equation [77] can be rewritten in the form:

\[
\frac{\partial RC}{\partial t} = \nabla \cdot (D \cdot \nabla C) - u \cdot \nabla C; \quad R = 1 + \frac{\rho}{\phi} K_c C^{n-1} \tag{78}
\]

where \( R \) is the retardation factor. Linear (i.e., \( n = 1 \)) reversible adsorption is easily simulated with RW using the retarded velocities and dispersion tensors \( u/R \) and \( D/R \). Note that the mass \( m \) of a particle located at \( x(t) \) includes then both the solute mass in mobile water (\( \approx m \)) and the adsorbed mass on the solid (\( \approx (R - 1)m \)).

Random walk has also been extended to nonlinear adsorption (\( n \neq 1 \)), in which case \( R \) depends on \( C \) as indicated by Eq. [78] (e.g., Tompson, 1993; Bosma et al., 1996; Abulaban et al., 1998). Solving solute transport with nonlinear adsorption requires at each time step the calculation of the concentration at the nodes of a regular grid superimposed onto the domain. Bagtzoglou et al. (1992) calculated the concentrations by projecting the particles on the grid

\[
C(X_i, t) \approx \sum_{i=1}^{N} m_i W_i [x(t) - X_i] \tag{79}
\]

where \( m_i \) is the mass of particle \( i \) located at \( x(t) \), \( N \) the total number of particles in the system, \( X_i \) the centroid of a grid cell, and \( W_i \) a projection function. \( W_i \) is usually a “box” function equal to \( 1/V \) inside a cubic cell centered on \( X_i \) (a square mesh in two dimensions), and zero otherwise, with \( V \) being the cell volume (mesh surface in two dimensions). The function \( W_i \) can also be a “chapeau” function to provide smoother results (see Bagtzoglou et al., 1992, for details). Irrespective of the form of \( W_i \), a sequential noniterative algorithm is typically invoked for nonlinear sorption as follows. First, the retardation factors are calculated over the grid using concentrations at time \( t \). Second, the particles are moved with the classical RW method based on Eq. [40], but with retarded velocities and dispersion tensors using retardation factors as calculated in Step 1. This algorithm for Freundlich sorption was used by Tompson (1993) and Bosma et al. (1996), and extended to Langmuir isotherms by Tompson (1993).

Since iterative procedures should provide more accurate results for nonlinear sorption, Abulaban et al. (1998) used the iterative coupling scheme suggested by Herzer and Kinzelbach (1989). This iterative procedure consists of the following steps:

1. Retardation factors are computed using concentrations at time \( t \).
2. All particles are moved with the retarded velocities and dispersion characteristics, while the set of random numbers used to perform the dispersion jumps is stored.
3. The particles at their new location are next mapped into concentrations, and the retardation factors are recalculated according to the new concentrations.
4. For each cell, the retardation factor is averaged between that of the initial field (at time \( t \)) and that of the final field (at time \( t + \Delta t \)).
5. The particles are moved again starting from their initial position at time \( t \). To reduce undesirable fluctuations of the retardation factor due to the random dispersion jumps, the stored set of random numbers generated at time \( t \) is used again for this updated motion.

This general algorithm can be iterated on as many times as needed to ensure convergence (i.e., until a stable concentration field between two successive attempts is reached). Numerical tests have shown that this iterative procedure is less prone to errors than noniterative schemes. The procedure is also less dependent on the initial conditions, the total number of particles, the size of the time step, and the nonlinearity of the isotherm (Abulaban et al., 1998).

### Kinetics-Controlled Linear Processes

Simple kinetics (first-order decay reactions) can be easily incorporated in the RW method. As shown by Kinzelbach (1987), a very efficient way is to modify for this purpose the solute mass assigned to each particle as follows:

\[
m_i(t) = m_i(0) \exp(-\lambda t) \tag{80}
\]

where \( m_i(0) \) is the initial mass assigned to particle \( i \) and \( \lambda \) the reaction rate coefficient.

Linear physical nonequilibrium processes are widely
encountered in fractured media (sometimes simulated along with matrix diffusion), in macroporous unsaturated soils (solute exchange between mobile and immobile water), and in heterogeneous saturated porous media. These processes are usually modeled with a set of equations that are mathematically similar to those for nonequilibrium sorption (e.g., Nkedi-Kizza et al., 1984):

\[
\frac{\partial C}{\partial t} + \frac{\rho \partial S}{\partial t} = \nabla \cdot (D \cdot \nabla C) - u \cdot \nabla C
\]

\[
\frac{\partial S}{\partial t} = k(K_iC - S)
\]

where \( C, S, D, u, \rho, \) and \( \phi \) refer to the same quantities as in Eq. [77]; \( K_i \) is the distribution coefficient \((L^3 \cdot M)\); and \( k \) the mass transfer coefficient \((T^{-1})\). The following notation is used here in our discussion of possible solutions of Eq. [81]: \( \Omega_m \) holds for the moving aqueous phase (mobile water or fracture) and \( \Omega_s \) for the nonmoving phase (solid phase, immobile water, or rock matrix), \( K_i = kK_i \) is the forward rate coefficient, and \( K_o = k \) is the backward rate coefficient for nonequilibrium sorption. Finally, we define an indicator variable associated with each particle: \( I_j(t) = 0 \) for the \( j \)th particle located in \( \Omega_m \) and \( I_j(t) = 1 \) for the \( j \)th particle located in \( \Omega_s \).

One of the first RW solutions of Eq. [81] was proposed by Kinzelbach (1987). The solution assumes that each particle located in \( \Omega_m \) is displaced using Eq. [40], while each particle in \( \Omega_s \) stays immobile. At the end of the time step, each particle can move to the other phase according to the following rules:

- if \( I_j(t) = 0 \) \( I_j(t + \Delta t) = \begin{cases} 0 & \text{for } \omega > P_{M \rightarrow S} \\ 1 & \text{for } \omega \leq P_{M \rightarrow S} \end{cases} \)
- if \( I_j(t) = 1 \) \( I_j(t + \Delta t) = \begin{cases} 0 & \text{for } \omega > P_{S \rightarrow M} \\ 1 & \text{for } \omega \leq P_{S \rightarrow M} \end{cases} \)

with \( P_{M \rightarrow S} = K_o\Delta t \), \( P_{S \rightarrow M} = K_o\Delta t \)

in which \( \omega \) is a random number drawn from a uniform distribution between 0 and 1. The time step must be very small so that the transition probabilities \( P_{M \rightarrow S} \) and \( P_{S \rightarrow M} \) remain smaller than one. Conceptually, the approach assumes that a particle cannot move to another phase during \( \Delta t \). This is another reason for keeping \( \Delta t \) small. The numerical algorithm is as follows:

- A particle located in \( \Omega_m \) is transported during \( \Delta t \) using Eq. [40]. Next a random number \( \omega \) is drawn: if \( \omega \leq P_{M \rightarrow S} \), the particle moves toward \( \Omega_s \), if not, the particle stays in \( \Omega_m \).
- The spatial location of a particle initially located in \( \Omega_s \) does not change during \( \Delta t \), but its change of phase is governed by a motion to \( \Omega_m \) if and only if \( \omega \leq P_{S \rightarrow M} \).

Given the assumption that a particle does not move to another phase during \( \Delta t \), Delay et al. (1996) provided an estimate of the number of particles in each phase at \( t + \Delta t \) by analytical integration of the phase transfer equation (the first-order kinetics model given by Eq. [81]). From a purely numerical point of view, larger time steps can be used.

Valocchi and Quinodoz (1989) suggested another procedure to let the particle make several jumps from one phase to the other during a single time step \( \Delta t \). The time spent by a particle within one phase \( \Omega_m \) or \( \Omega_s \) is given by

\[
t_m = -\frac{\log \omega_m}{K_i} \quad \text{and} \quad t_s = -\frac{\log \omega_s}{K_o}
\]

where \( \omega_m \) and \( \omega_s \) are random numbers uniformly distributed over \([0,1] \), \( \omega_m \) and \( \omega_s \) represents the probability that a particle remains in its original phase after a given time \( t_m \) and \( t_s \). The random numbers \( \omega_m \) and \( \omega_s \) are drawn until the sum of the time spent in \( \Omega_m \) and \( \Omega_s \) is equal to \( \Delta t \). The particle is then moved by advection–dispersion during the time spent in \( \Omega_m \) or \( \Omega_s \). At the end of the time step, the particle can stay in its initial phase or move to the next phase according to probabilities defined as follows (Parzen, 1962):

\[
p_{M \rightarrow S} = \frac{K_i}{K_i + K_b} [1 - \exp\{-(K_i + K_s)\Delta t\}]
\]

\[
p_{M \rightarrow M} = 1 - p_{M \rightarrow S}
\]

\[
p_{S \rightarrow M} = \frac{K_b}{K_i + K_b} [1 - \exp\{-(K_i + K_s)\Delta t\}]
\]

\[
p_{S \rightarrow S} = 1 - p_{S \rightarrow M}
\]

in which \( p_{M \rightarrow S} \) is the probability for a particle in Phase A to move to Phase B or to stay in A if A = B. It must be noted that for small values of \( (K_i + K_s)\Delta t \), Eq. [84] tends to Eq. [82]. The same phase transition probabilities were found by Michalak and Kitanidis (2000) using a development based on zero-order moments of the solute distribution in both phases. The algorithm given by Eq. [83] and [84] has been applied to various test cases, including random \( K_i \) and/or \( K_o \) parameters (Hassan et al., 1997, 1998).

To reduce the computational costs, Andrievic and Foufoula-Georgiou (1991) directly estimated the time spent by a particle in \( \Omega_m \) using

\[
\Delta t_m = \Delta t + z\sigma(\Delta t_m)
\]

in which \( z \) is a random normal deviate. The mean \( \Delta t_m \) and standard deviation \( \sigma(\Delta t_m) \) of the residence time distribution in \( \Omega_m \) are calculated numerically at the beginning of the time step \( \Delta t \). In the same way, Valocchi and Quinodoz (1989) estimated the fraction of the time step spent by a particle in phase \( \Omega_m \) for each initial or final phase combination. They referred to the work by Keller and Giddings (1960) to define the phase transition probabilities:

\[
f_{M \rightarrow M}(\tau) = \Delta t \sqrt{\frac{K_iK_b}{1 - \tau}} \exp\{-(1 - \tau)K_s + \tau K_i\} B_1(\chi)
\]

\[
f_{M \rightarrow S}(\tau) = \Delta t K_b \exp\{-(1 - \tau)K_b + \tau K_i\} B_1(\chi)
\]

\[
f_{S \rightarrow M}(\tau) = \frac{K_b}{K_i} f_{M \rightarrow S}(\tau)
\]

\[
f_{S \rightarrow S}(\tau) = \frac{1 - \tau}{\tau} f_{M \rightarrow M}(\tau)
\]

with

\[
\chi = 2\Delta t \sqrt{K_iK_b} \tau(1 - \tau)
\]
where \( f_{A\to B}(\tau) \) are the probability distribution functions of phase transitions \( A \to B \) for the fraction of time \( \tau \) spent by a particle in \( A \), and \( B_a \) and \( B_r \) are modified Bessel functions of the first kind and of zero- and first-order, respectively. These probability functions are estimated numerically for \( \tau \in [0, 1] \), by assuming a random initial distribution of the solute between the phases. At each time step, the final phase reached by each particle is determined by the transition phase probabilities given by Eq. [84]. Each particle is displaced using a modified RW algorithm:

\[
x(t + \Delta t) = x(t) + M_1(x) + [M_2(x)]^{1/2}Z \tag{87}
\]

where \( M_1(x) \) is the first-order spatial moment, \( M_2(x) \) is the centered second-order spatial moment, and \( Z \) a random normal deviate (with a mean of 0 and a variance of 1). The moments \( M_1(x) \) and \( M_2(x) \) have four different forms depending on the final and initial phase locations; that is, the particle is located in phase \( A \) at time \( t \) and in phase \( B \) at time \( t + \Delta t \) with all possible combinations \( A \to B, A \to A, B \to A, B \to B \). Michalak and Kitanidis (2000) performed numerical comparisons between their algorithm, the method suggested by Kinzelbach (1987) as described earlier, and the above two methods described by Valocchi and Quinodoz (1989). Their algorithm was found to require less computation time without loss of accuracy.

We note that while several other methods have been developed, they generally apply only to more restricted conditions. These methods include algorithms suggested by Selroos and Cvetkovic (1992) for a uniform sorption coefficient and Mishra et al. (1999) for situations where local dispersion can be neglected.

Transfer probability functions have also been studied and used to simulate fracture–matrix interactions. If applied to fracture–matrix systems, this concept can be extended to any dual-porosity or dual-permeability system involving preferential flow paths in conjunction with low conductivity areas (e.g., soils containing macropores and matrix porosity). Liu et al. (2000) modeled the presence of a matrix and fractures in terms of two interacting continua within each cell of the discretized domain. They estimated the transfer probability functions by evaluating the net mass flux between the two continua. As stated by Liu et al. (2000), and confirmed by Hassan (2002), this approach requires smooth concentration distribution in the matrix (i.e., relatively large diffusion coefficients). Still another approach for matrix–fracture interactions was suggested by Yamashita and Kimura (1990), who estimated the apparent increase in residence time in a fracture due to matrix diffusion. Their main idea was to select an analytical solution to the phase transfer equation that describes the normalized concentration ratio \( C/C_0 \) exiting the fracture and diffusing into the matrix, with \( C_0 \) being the constant concentration continuously injected at the inlet of the fracture. Such analytical solutions can be written in the general form:

\[
C/C_0 = f(t, t_r, p) \tag{88}
\]

where \( t_r \) is the residence time in the fracture if no diffusion occurs, and \( p \) is a set of parameters depending on the geometry and properties of both the fracture and the matrix. To account for matrix diffusion, a random number \( \omega \) uniformly distributed between \([0, 1]\) is drawn and the residence time \( t \) including matrix diffusion is obtained by inverting \( \omega = f(t, t_r, p) \). Tsang and Tsang (2001) and Tsang and Doughty (2003) later extended this approach to blocks of finite dimension and more complex systems. Note that the principle of these approaches, which analyze the influence of matrix-diffusion on advection–dispersion and calculate the time spent by matrix diffusion, is very similar to the approach developed by Delay and Bodin (2001) for the TDRW as we discussed earlier.

A FEW NUMERICAL RECIPES

Compared with Eulerian methods, the three main advantages of RW methods are as follows:

1. Advection-dominated transport can be simulated without introducing numerical dispersion (i.e., nearly hyperbolic problems can be considered).
2. They do not require any space discretization if velocities are known everywhere.
3. Because of their intuitive nature, the algorithms are quite easy to implement and run on parallel computers.

Random fluctuations in computed concentrations constitute the main drawback of the method. These fluctuations can be minimized by using a large number of particles. However, because fluctuations are proportional to the square root of the number of particles, improvements in the results are not of the same order as the increase in computational costs.

Numerical accuracy is also affected when relatively large time steps are used. With large time steps, a particle may overlook local flow characteristics when it moves in one step over a large heterogeneous velocity area. To overcome this problem, the time step is usually defined on the basis of the Courant criterion, which is the ratio between \( u\Delta t \) and a characteristic length of a grid cell (Tompson, 1993). Since heterogeneities are usually defined using a constant value per grid cell, the time step must be chosen such that each particle is able to correctly sample the transport properties of each cell. Except for specific algorithms (e.g., see the section above on alternative approaches using cellular automata), the time step may become very small and lead to cumber-some calculations. To significantly reduce the CPU time of the computation, Uffink (1985, 1990) suggested the use of a random number \( \omega \) drawn from a uniform distribution between \((-1, 1)\), whose calculation may be \( 10 \) to \( 40 \) times faster than using of a normally distributed number. Equation [35] then becomes
\[ x(t + \Delta t) = x(t) + u[x(t)]\Delta t + [6D[x(t)]\Delta t]^{0.5} \omega \]  

The CPU time required to generate a random number may vary within a factor of 40 according to the method used (Press et al., 1992). Generators based on linear congruent methods are very fast but not free from sequential correlation between successive sets of random numbers. These sequential correlations reduce the theoretical variance that should be obtained for the random dispersive jumps. Press et al. (1992) gave detailed guidelines on how to select valuable random number generators.

Several types of boundary conditions can be implemented with the RW particle tracking method. No-flux conditions are treated by bouncing the particles against a particular boundary. Dirichlet conditions are described by injecting a prescribed number of particles along the boundary. Injecting particles adds mass at a given time to the system. Thus, the number of particles must be defined according to the mass of each particle and the water flux across the boundary to obtain the concentration value prescribed by the Dirichlet condition. For time-varying boundary conditions, an initial injection time can be allotted to each particle to avoid additional fluctuations. For example, the injection time \( t_j \) of particle \( j \) within a range \([t, t + \Delta t]\) may be written as

\[ t_j = t + \omega \Delta t \]  

where \( \omega \) is a random number uniformly distributed between 0 and 1. On the other hand, if the concentration injected at a boundary is invariant in time, using the integral of a Dirac type boundary can be very efficient and may significantly reduce the random fluctuations. In practice, transport is simulated for a Dirac (pulse) injection, and the number of particles at \( x \) and time \( t \) responding to continuous injection is simply calculated as

\[ N_j^\delta(x, t) = \sum_{i=1}^{N_j} N_j^\delta(x, i\Delta t) \quad \text{with } t = N_j\Delta t \]  

in which \( N_j^\delta(x, i\Delta t) \) is the number of particles at \( x \) and time \( i\Delta t \) computed for the Dirac injection. Of course, this method is not applicable to nonlinear transport since the evolution of concentrations (number of particles) in space and time also depends on local concentrations present in the system. These local concentrations are different between the responses to a pulse or a continuous injection. We note that total flux boundary conditions (in which both advection and dispersion fluxes are prescribed) cannot be considered using particle tracking.

The calculation of concentrations requires the definition of a reference volume to “transform” particles (with their mass and potentially located anywhere in the domain because of random jumps) into a mean mass per unit volume around a specified location. This integration over a volume may generate artificial diffusion when the concentrations are used for further calculations. This mapping from particle distribution to concentrations has been studied by Bagtzoglou et al. (1992). As suggested earlier in our discussion, the least biased mapping is obtained when concentrations are calculated by projecting the particles onto a grid according to Eq. [79].

**CONCLUSIONS**

This review was performed to provide an assessment of recent advances in RW particle tracking methods for simulating solute transport in porous and/or fractured media. While our review is certainly not exhaustive, the many references included in this text reflect the fact that particle tracking has motivated much research and remains a critical component of advanced numerical methods and their applications. As stated in our introduction, we did not intend to compare Eulerian and Lagrangian methods or to come up with long lists of advantages and drawbacks for different types of case studies. Our view in fact was very simple: one must be pragmatic and able to move from one technique to the other depending on the expected difficulties. No easy and standard recipe exists for selecting the best method for a particular problem, except perhaps for the general statement that Lagrangian methods are well suited for problems involving very sharp contrasts in hydrodynamic properties. The RW computations themselves may still be somewhat cumbersome. By contrast, while the same problem may be solved more rapidly using Eulerian approaches, these approaches often lead to inaccurate solutions. But even these statements can be questioned. For example, Lagrangian methods have evolved to the point where time domain approaches are now very rapid also. However, as compared with classical random walk methods in the space domain, time domain approaches have not yet advanced enough for nonlinear problems, which still require a considerable amount of time for estimating local concentrations at prescribed times. Eulerian methods have progressed also. Recent concepts such as the Eulerian Lagrangian Localized Adjoint Method (ELLAM) or other techniques using distorted and/or moving grids have borrowed several Lagrangian features to enable the solution of advection-dominated problems while preserving computational efficiency.

Finally, the best suggestion may be to try both the Lagrangian and Eulerian approaches for the same problem, and to compare results in terms of both accuracy, ease of use, and computational efficiency. We hope to have shown in this review that RW methods have enormous potential. We believe that RW particle tracking may well pioneer promising new lines of research in subsurface flow and transport, especially when applied to mass transfer in unsaturated soils. The vadose zone is the interface between the very transient lower atmosphere and less transient underlying groundwater systems, while being affected more directly also by human activities. This is what makes the vadose zone a very complex entity in terms of the prevailing mass transfer process, including the coupling of solute transport with a large number of physical and chemical processes and reactions. We fully acknowledge that additional efforts are needed to adapt RW to such problems as multiphase and/or multicomponent reactive transport. A true potential, however, exists for addressing such problems. This
progress may well occur along similar lines as for use of RW to provide efficient solutions for transport in flow domains with high velocity contrasts, and for nonlinear reactive transport. Rapid improvements are more likely when the method becomes more widely used in quantitative vadose zone flow and transport studies. As such, we hope that this review and the abundant literature referenced herein will give new incentives to development of accurate and computationally efficient subsurface solute transport models.

APPENDIX

Following Uffink (1990), we consider a one-dimensional diffusion process and an interface located at $x = 0$ between two layers $\lambda$ and $\gamma$. The $x$ coordinate is taken positive in layer $\lambda$ and negative in layer $\gamma$. The particle is assumed to be at $x_0$ in layer $\lambda$. Let $D_\lambda$ and $D_\gamma$ be the dispersion coefficients of layers $\lambda$ and $\gamma$, respectively, with $D_\lambda > D_\gamma$. The probability density function of the particle displacement in layer $\lambda$ is given by

$$P_\lambda \propto \frac{1}{2\pi D_\lambda t^{3/2}} \exp \left[ -\frac{(x - x_0)^2}{2D_\lambda t} \right]$$

$$+ \frac{R}{2\pi D_\lambda t^{3/2}} \exp \left[ -\frac{(x + x_0)^2}{2D_\lambda t} \right]$$

where $R$ is the fraction of the mass flux that is reflected and $(1 - R)$ the fraction of the mass flux that crosses the interface. For layer $\gamma$, the probability density function of the particle displacement is

$$P_\gamma \propto \frac{(1 - R)}{2\pi D_\gamma t^{3/2}} \exp \left[ -\frac{(x + \beta x_0)^2}{2D_\gamma t} \right]$$

where $\beta x_0$ is the modified location of the source, assuming that layer $\lambda$ has the same dispersion property as layer $\gamma$. At the interface, the probability density functions are

$$P_\lambda \propto \frac{1 + R}{2\pi D_\lambda t^{3/2}} \exp \left[ -\frac{x_0^2}{2D_\lambda t} \right]$$

$$P_\gamma \propto \frac{1 - R}{2\pi D_\gamma t^{3/2}} \exp \left[ -\frac{(\beta x_0)^2}{2D_\gamma t} \right]$$

To ensure continuity at the interface, $P_\lambda = P_\gamma$ at any time $t$ and therefore

$$R = \frac{\sqrt{D_\lambda} - \sqrt{D_\gamma}}{\sqrt{D_\lambda} + \sqrt{D_\gamma}}; \quad \beta = \sqrt{\frac{D_\gamma}{D_\lambda}}$$


$$P_\lambda \propto \frac{2\sqrt{D_\lambda}}{(\sqrt{D_\lambda} + \sqrt{D_\gamma})^{3/2}} \frac{1}{2\pi D_\lambda t^{3/2}} \exp \left[ -\frac{x_0^2}{2D_\lambda t} \right]$$

$$P_\gamma \propto \frac{2\sqrt{D_\gamma}}{(\sqrt{D_\lambda} + \sqrt{D_\gamma})^{3/2}} \frac{1}{2\pi D_\gamma t^{3/2}} \exp \left[ -\frac{(\beta x_0)^2}{2D_\gamma t} \right]$$

Thus, for a particle located at the interface (i.e., $x_0 = 0$), the probability for this particle to move toward layer $\lambda$ is $P_\lambda = \sqrt{D_\lambda}/(\sqrt{D_\lambda} + \sqrt{D_\gamma})$ and to move toward layer $\gamma$ is $P_\gamma = 1 - P_\lambda = \sqrt{D_\gamma}/(\sqrt{D_\lambda} + \sqrt{D_\gamma})$.

ACKNOWLEDGMENTS

The authors are grateful to the French National Program for Research in Hydrology that funded this work. Helpful comments from two anonymous reviewers and discussions with Dr. R. van Genuchten are also acknowledged.

REFERENCES


Rivard, C., and F. Delay. 2004. Simulations of solute transport in...


