Fractional Calculus in Hydrologic Modeling: A Numerical Perspective

David A. Benson

Hydrological Science and Engineering, Colorado School of Mines, Golden, CO, 80401, USA. (dbenson@mines.edu)

Mark M. Meerschaert

Department of Statistics and Probability, Michigan State University, East Lansing, MI, USA. (mcubed@stt.msu.edu)

Jordan Revielle

Hydrological Science and Engineering, Colorado School of Mines, Golden, CO, 80401, USA. (Now at Ward Petroleum Corporation, Fort Collins, CO 80521) (jkrevielle@gmail.com)

Abstract

Fractional derivatives can be viewed either as a handy extension of classical calculus or, more deeply, as mathematical operators defined by natural phenomena. This follows the view that the diffusion equation is defined as the governing equation of a Brownian motion. In this paper, we emphasize that fractional derivatives come from the governing equations of stable Lévy motion, and that fractional integration is the corresponding inverse operator. Fractional integration, and its multi-dimensional extensions derived in this way, are intimately tied to fractional Brownian (and Lévy) motions and noises. By following these general principles, we discuss the Eulerian and Lagrangian numerical solutions to fractional partial differential equations, and Eulerian methods for stochastic integrals. These numerical approximations illuminate the essential nature of the fractional calculus.

Keywords: Fractional Calculus, fractional Brownian motion, Mobile/Immobile, Subordination
PACS: 02.50.Ey, 02.50.Ga, 02.70.Ns, 05.10.Gg

1. Introduction

The term "fractional calculus" refers to the generalization of integer-order derivatives and integrals to rational order. This topic was first broached by L'Hopital and Leibniz after the latter's co-invention of calculus in the 1700s (see the excellent history by [1]). In fact, the operators can be extended to

complex as well as real order, so the "fractional" label is a minor historical misnomer.

Fractional calculus was primarily a mathematical curiosity for centuries (see examples in [1] and [2]). For example, when Heaviside would take the "square root" of both sides of a diffusion equation, he was generating a 1/2-order time derivative. Some of the first physical applications were by geophysicists describing material somewhere between elastic (Hooke's linear relationship between stress and strain) and viscous (described by Newton's stress proportional to strain rate). In his work on this area starting in the 1960's, geophysicist Michele Caputo derived the fractional derivative that carries his name. Benoit Mandelbrot's work on fractional Brownian motion and geophysical time series starting in the 1960's implicitly used fractional-order integration.

In the 1990's, a resurgence of interest surrounded the application of fractional derivatives in the model equations of anomalous diffusion (see [3] for an extensive review). At the same time, an understanding of the importance of general non-locality in upscaled transport in heterogeneous aquifer material emerged [4, 5]. The non-locality is defined by operators that account for (integrate) the concentrations at previous times and/or large regions of space. These studies were based on the simple idea that the concentration change at some collection point (a plane or well) depended on contributions from potentially large distances upstream and/or the concentration loading history for some time in the past. Formally, the non-locality arises when the underlying velocity field is uncertain and correlation scales are significantly large compared to the scale of observation [6]. Upscaled descriptions of transport lose detailed velocity information that is transferred to the non-local operators.

One attempt to incorporate spatial non-locality in a tractable form assumed a set of weights that decayed as a power-law [7, 8, 9], which forms the definition of a fractional-order dispersion term. This formulation assumed that the concentration change at some point depended on upstream concentrations, and the dependence decayed like a power law of the distance. Temporal non-locality, in which concentration change at a point depends on the prior concentration "loading" is the basis for hydrologic applications of continuous time random walks (CTRW). The CTRW were shown to define temporal fractional derivatives when the weighting of prior concentration decayed like a power-law (see the extensive review by *Metzler and Klafter* [3]). A few years later, the formal link between two-state (mobile/immobile) multi-rate mass transfer equations [10, 11] and temporally fractional-order models was made [12, 13]. This accounts for solute loading into relatively impermeable material that slowly releases the solute after the bulk of a plume has passed.

Forays into fractional calculus in multiple dimensions showed that the fractional derivatives could be extended in ways significantly different than classical cases. The derivative operators were defined by the underlying diffusion process of Lévy motion, which could have different scaling rates and skewness in different directions. The derivative operators inherit the different orders and descriptions of skewness in all directions. Because of the link between derivatives and integrals, these extensions can be transferred to any system that uses

fractional integrals. The most common hydrologic application of fractional integration is the generation of fractional Brownian motion as a representation of aquifer material with long-range correlation structure. Using the inverses of the newly defined fractional derivatives gave new tools to extend the classical fractional Brownian motion to more closely represent anisotropic aquifer structure [68].

Because the fractional derivative and integrals are defined as convolution operators, they are easy to implement using standard numerical techniques. In addition, because the fractional diffusion equations that generate the derivative operators are based on the motion of a single particle, the classical random walk particle tracking (RWPT) techniques are well-suited to solve the fractional advection-dispersion transport equations. We exploit the numerical implementations as a vehicle to define and solve to fractional-order differential and integral equations.

The paper is organized in three main sections dealing with fractional space derivatives (section 2), fractional time derivatives (section 3), and fractional integrals (section 5). Within the two derivative sections, we outline how the diffusion equation, and its fractional-order counterparts, are defined by the stochastic processes that they describe. We show how the equations naturally induce both their Eulerian (section 2.5) and Lagrangian (section 2.6) numerical approximations. In section 4 we briefly summarize how the fractional transport equations have been applied to contaminant transport problems in surface and subsurface hydrology. We then show in section 5 how the inverse of the fractional derivative operators define the fractional integrals in multiple dimensions, and how these integrals can be used to generate conditioned, multi-scaling, random aquifer facsimilies. We close with conclusions and recommendations for future work in section 6.

2. Markovian Diffusions and Fractional Space Derivatives

There are several forms of fractional derivatives that are distinguished by the domain over which they operate. Because they are non-local operators, they "look" for values from a certain distance ahead or behind for information. For spatial processes it may be correct to look ahead and/or behind (or at any angle) over all space. Temporal information is only used after some starting time, so the domain of interest is positive time only. We use these distinctions to explain the association of the different operators to different behaviors in diffusions based on random walks.

The starting point for all of the generalizations is classical Brownian motion. It is well known that Brownian motion $B(t)$ is the limit Markov (memoryless) process of finite-variance random walks with short-range correlation [14, 15]. This makes Brownian motion an attractive model for transport of passive tracers in surface and ground water: the exact nature of the individual motions is not particularly important in the long-term. The central limit theorem dictates that all finite-variance motions converge toward the Gaussian limit distribution. It is precisely this property that has made Brownian motion an attractive and

useful model of macrodispersion in aquifers. Even with non-Gaussian particle motions, the long term transport tends toward the Gaussian limit distribution (for perhaps the earliest experimental example see Taylor [16]).

If $B(t)$ denotes the location of a particle in one-dimensional space x at time t then the density of the location $p(x,t)$ is given by

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(\frac{-x^2}{4Dt}\right) \tag{1}$$

where D is half the variance of each motion size divided by the mean motion time. Throughout this paper, we will use Fourier $f(k) \equiv \int e^{-ikx} f(x)dx$ and Laplace $f(s) \equiv \int e^{-st} f(t)$ transforms, where it is understood for notational simplicity that $f(x) \Leftrightarrow f(k)$ and $f(t) \Leftrightarrow f(s)$ are transform pairs, not the same functions.

To connect the diffusion equation with Brownian motion, note that the Fourier transform (FT) of (1) is $p(k,t) = \exp(tD(ik)^2)$, with time derivative

$$\frac{dp(k,t)}{dt} = D(ik)^2 \exp(tD(ik)^2) = D(ik)^2 p(k,t) \tag{2}$$

A property of Fourier transforms of integer-order derivatives is that $(ik)^n f(k) \Leftrightarrow d^n f(x)/dx^n$, so that the inverse transform of the previous equation becomes

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2}{\partial x^2} p(k,t) \tag{3}$$

In a more general way that will be useful shortly, we can write the FT of the Brownian motion density as $p(k,t) = \exp(tA(k))$, where the function of the wavenumber $A(k) = D(ik)^2$, then following the same procedure the "inverse FT" of $A(k)$ defines the linear space operator in the Cauchy equations

$$\frac{dp(k,t)}{dt} = A(k)p(k,t), \tag{4}$$

with inverse FT

$$\begin{aligned} \frac{dp(x,t)}{dt} &= \int A(x)p(x-\xi,t)d\xi \\ &\equiv A_x p(x,t) \end{aligned} \tag{5}$$

where the $A_x()$ denotes the linear space operator defined by convolution with $A(x)$, the inverse FT of $A(k)$. Here we use the fact that the product of two functions $A(k)p(k,t)$ in Fourier space is a convolution in real space. This convolution, in turn, specifies an operation on the function $p(x,t)$ in real space. For example $(ik)^2 \Leftrightarrow d^2/dx^2$ represent the pair $A(k) \Leftrightarrow A_x$ for Brownian motion. This convolution machinery can be used to explain the diffusion equation for Brownian motion, because the function $(ik)^2$ is the (distributional) FT of the second derivative of the Dirac delta function. The Dirac delta function $\delta(x-a)$

for some constant shift a is a "generalized function" (also called a distribution) defined by

$$\int \delta(x-a)f(x)dx = f(a). \tag{6}$$

Its derivatives are defined via integration by parts:

$$\int \delta^{(n)}(\xi)f(x-\xi)d\xi = \int \delta(\xi)f^{(n)}(x-\xi)d\xi. \tag{7}$$

Because the values of $f(x)$ for $x \neq a$ do not affect the integral (6), we might say that

$$\delta(x-a) = \begin{cases} \infty & \text{if } x = a \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $\int \delta(x)dx = 1$, so that the infinity at $x = a$ is tamed by integration. Another intuitive definition of the Dirac function is that it is the limit of a Gaussian density function with mean a as the variance tends toward zero, i.e., the Dirac delta is like the probability density "function" of the constant number a.

Taking $f(x) = e^{-ikx}$ in equation (6) shows that the FT of $\delta(x-0)$ equals 1. Then the FT of $\delta''(x)$ is $(ik)^2 \times 1$, so that multiplying the FT by $(ik)^2$ is equivalent to convolution with $\delta''(x)$. Therefore, Brownian motion, by virtue of the FT of its density function, defines the diffusion equation. This is the sole connection between the diffusion equation and Brownian motion. The notion that a concentration gradient "drives" a diffusion by physical means was dispelled by Einstein [17] and Crank [18] in their seminal work. The extension of the probability distribution for a single particle, $p(x,t)$ to concentration for a large number of particles, requires independence of their motion and eliminates particles of one species acting upon each other as a driving force. This is also called the infinitely dilute approximation. The concept that the random motion of a single particle defines the diffusion equation, in which the flux happens to be proportional to concentration gradient, rather than the picture that a molecule moves *in response* to that gradient, is central to our further development, and is eloquently described by Crank [18]:

> If it were possible to watch individual molecules of iodine, and this can be done effectively by replacing them by particles small enough to share the molecular motions but just large enough to be visible under that microscope, it would be found that the motion of each molecule is a random one. In a dilute solution each molecule of iodine behaves independently of the others, which it seldom meets, and each is constantly undergoing collision with solvent molecules, as a result of which collisions it moves sometimes toward a region of higher, sometimes of lower, concentration, having no preferred motion towards one or the other. The motion of a single molecule can be described in terms of the familiar 'random walk' picture, and whilst it is

possible to calculate the mean-square distance travelled in a given interval of time it is not possible to say in what direction a given molecule will move in that time.

This picture of random molecular motions, in which no molecule has a preferred direction of motion, has to be reconciled with the fact that a transfer of iodine molecules from the region of higher to that of lower concentration is nevertheless observed. Consider a horizontal section in the solution and two thin, equal elements of volume one just below and one just above the section. Though it is not possible to say which way any particular iodine molecule will move in a given interval of time, it can be said that on the average a definite fraction of molecules in the lower element of volume will cross the section from below, and the same fraction of molecules in the upper element will cross the section from above, in a given time. Thus, simply because there are more iodine molecules in the lower element than in the upper one, there is a net transfer from the lower to the upper side of the section as a result of random molecular motions.

In the 1920s, Paul Lévy discovered the class of processes that correspond to the limits of all random walks (in 1-d) by easing the requirement of finite variance in the classical central limit theorem. When the probability of the individual motions have power-law tails $P(|W| > x) \sim Cx^{-\alpha}$ for some constant C and $0 < \alpha < 2$, the rescaled sum of these walks converges to a Lévy motion with FT

$$p(k,t) = \exp[-tD(a(ik)^\alpha + (1-a)(-ik)^\alpha)] \tag{9}$$

so that (9) admits the same form as (4) but with $A(k) = Da(ik)^\alpha + D(1-a)(-ik)^\alpha$. The (distributional) inverse transform of $(ik)^\alpha$ is the power law $x^{-1-\alpha}/\Gamma(-\alpha)$ for $x > 0$, and the inverse transform of $(-ik)^\alpha$ is $(-x)^{-1-\alpha}/\Gamma(-\alpha)$ for $x < 0$. The skewness parameter has a range $0 \le a \le 1$. The probability increase/decrease rate equation

$$\frac{dp(k,t)}{dt} = p(k,t)[-tD(a(ik)^\alpha + (1-a)(-ik)^\alpha)]$$

implies that particles may jump long distances. This can be seen in a long form of the real-space equation

$$\begin{aligned}\frac{dp(x,t)}{dt} &= \frac{Da}{\Gamma(-\alpha)} \int_{-\infty}^{x} (x-\xi)^{-1-\alpha} p(\xi,t) d\xi \\ &\quad + \frac{D(1-a)}{\Gamma(-\alpha)} \int_{x}^{\infty} (-x+\xi)^{-1-\alpha} p(\xi,t) d\xi.\end{aligned} \tag{10}$$

Strictly speaking, these are convolutions of generalized functions like the Dirac delta function, because the power law $x^{-1-\alpha}$ is not integrable at $x = 0$. Assuming a well-behaved function p that goes to zero at $\pm\infty$, the intergrals can

be regularized (tamed) into convergent forms using the integration by parts formula n times, where $n-1<\alpha<n$, to obtain converging convolution integrals involving $x^{-1-\alpha+n}$.

These convolutions with forward and backward power laws define two specific types of fractional-order derivatives, denoted in the diffusion equation

$$\frac{\partial p(x,t)}{\partial t} = Da\frac{\partial^\alpha p(x,t)}{\partial x^\alpha} + D(1-a)\frac{\partial^\alpha p(x,t)}{\partial(-x)^\alpha}. \tag{11}$$

The forward direction fractional derivative (as well as a fractional time derivative defined later) is "causal" in that the derivative at some point depends on values to the left on the real line. The backward fractional derivative generally only pertains to space functions, because it is not causal; it depends on values to the right. While this sounds counterintuitive, the backward derivative models backward jumps; therefore, the probability change at some point depends on probability that a particle starts a jump *from* a forward location.

The transition from integer to fractional derivatives is most easily understood in terms of Fourier transforms: Recall the FT pair $d^n f(x)/dx^n \Leftrightarrow (ik)^n f(k)$, and substitute a real-valued α for n. Some variations on this definition (described later) recognize the fact that for time derivatives, $t=0$ defines a boundary that has some influence on the convolution, and must be treated properly. The main point we wish to emphasize here is that the same Fourier symbol $A(k) = Da(ik)^\alpha + D(1-a)(-ik)^\alpha$ determines both the fractional derivative, and the corresponding stable Lévy motion.

2.1. 1-D Fractional Derivative: Numerics

Before venturing into the territory of multiple dimensions, it is instructive at this point to examine the convolution operator (4) in relation to the classical integer derivatives and their numerical approximations. The convolution specifies that the change in probability (and concentration) is due to the sum of concentrations everywhere weighted by the function in the convolution. The Dirac delta $\delta(x)$ and its derivatives are zero everywhere except where $x=0$. This property defines a "local" operator. In a numerical implementation of convolution, one takes a finite domain Ω and discretizes it into N partitions of size Δx. Convolution with a delta function $\delta(x)$ is represented by a weighted sum with zero weights everywhere, except for a value of 1 at $x=0$. Call the weights w_i, then $w_0=1, w_{i\neq 0}=0$. The derivative of $\delta(x)$ is represented by the numerical "slope" on either side of the impulse: the immediate rise $1/\Delta x$ and fall $-1/\Delta x$. A second derivative is the slope of that function: $1/\Delta x^2, -2/\Delta x^2$, and $1/\Delta x^2$. The second derivative at some grid location is a convolution of these weights with some function $f(x)$ discretized at the same points:

$$\frac{d^2 f(x)}{dx^2} \approx \sum_{l=0}^{\infty} w_l f(x-l\Delta x) = \frac{f(x)-2f(x-\Delta x)+f(x-2\Delta x)}{\Delta x^2} \tag{12}$$

Any integer derivative can be represented by a discrete convolution, using weights from the binomial coefficients:

$$\frac{d^n f(x)}{dx^n} = (\Delta x)^{-n} \sum_{l=0}^{N} (-1)^l \binom{n}{l} f(x - l\Delta x). \tag{13}$$

The n^{th} derivative has $n+1$ terms, and for stability, the weights are typically shifted to the right by the greatest integer less than $n/2$, denoted $[n/2]$, so that the formula for all integer-order derivatives becomes

$$\frac{d^n f(x)}{dx^n} = (\Delta x)^{-n} \sum_{l=0}^{N} (-1)^l \binom{n}{l} f(x - (l - [n/2])\Delta x). \tag{14}$$

Grünwald recognized in the 1800's that the formula for integer finite differences was easily extended to the fractional-order case. The binomial coefficients of any order α can be defined using the formula

$$\binom{\alpha}{l} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha - l + 1) l!}, \tag{15}$$

so the finite difference approximation formula for a fractional derivative of a function at point x (or approximate fractional integral for $\alpha < 0$) can be written

$$\begin{aligned} \triangle_+^\alpha f(x) &= \Delta x^{-\alpha} \sum_{l=0}^{N} (-1)^l \binom{\alpha}{l} f(x - l\Delta x) \\ &= \Delta x^{-\alpha} \sum_{l=0}^{N} w_l f(x - l\Delta x), \end{aligned} \tag{16}$$

where the Grünwald weights

$$w_l = (-1)^l \binom{\alpha}{l} = \frac{(-1)^l \Gamma(\alpha+1)}{\Gamma(\alpha - l + 1) l!} \tag{17}$$

are illustrated in Figure 1. Only a few orders are shown, but the weight functions smoothly interpolate between all orders, including when the derivative order is negative (indicating fractional integration). Note that for direct comparison, the derivative weights for positive α are not shifted to the left in Figure 1. While at their core, the fractional derivatives are defined by convolution with a power law, the discrete weights are not monotonic because the derivatives have a rise at the origin and subsequent fall. The same non-monotonic behavior is seen in the discrete version of any integer-order derivative as well.

Equation (16) corresponds to the positive fractional derivative, so the weights apply to points to the left of x. The negative fractional derivative defined by the FT multiplier $(-ik)^\alpha$ has a Grünwald approximation

$$\triangle_-^\alpha f(x) = \Delta x^{-\alpha} \sum_{l=0}^{N} (-1)^l \binom{\alpha}{l} f(x + l\Delta x) \tag{18}$$

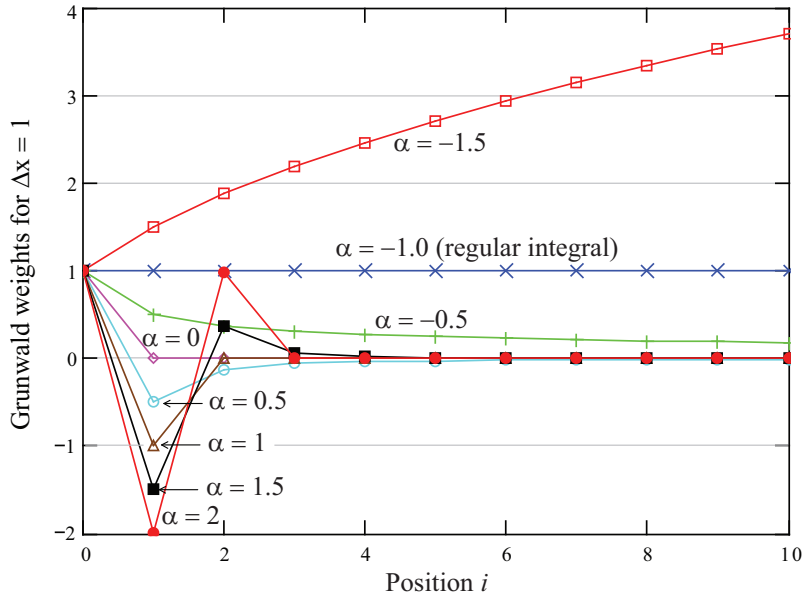


Figure 1: **Grunwald (convolution) weights for finite differences of orders 2 through -1.5 in 0.5 intervals. Connecting lines are used to guide the eye. Δx is set to 1.**

which depends on points to the right of x. It is called the negative, or backward, fractional derivative because it corresponds to particle jumps in the negative direction. Similar to the integer order case, for stability the weights shift by one position to the right for the positive direction derivative for $1 < \alpha < 2$, and shift one to the left for the negative direction derivative [19, 20]. It can be shown [**?** , Proposition 2.1] that these discrete Grünwald approximations converge to the integral convolutions in (10) as $\Delta x \to 0$ just as in the integer-order case.

Questions are often asked about what the fractional derivatives "mean" in terms of continuous functions, and answers are hard to deliver. But at this point it becomes apparent what the integer and fractional derivatives "mean" when related back to the random walks that generate the equations. The derivatives are accounting for mass transfer due to moving particles. Brownian motion is composed, by construction, of vanishingly small jumps, so one must look immediately to the left and right to see which particles might arrive at some point in a fixed time interval and change the concentration. Looking farther than Δx, which goes to zero, is pointless because those particles cannot make it to the current location. The second derivative, a local operator, is well-suited to describe this process. Heavy-tailed random walks embody a significant probability that particles from some distance can, in the rescaled random walk, make it to the current location, and the Grünwald weights account for mass accumulation at any point due to distant random walkers. Which derivative pertains, traditional or fractional, depends on the specifics of the random walk.

Adding a constant-in-time drift to the random walk changes the probability density of the random walker $p(k,t) = \exp(A(k)t)$ by adding a term $-v(ik)$ to $A(k)$. Bearing in mind the FT relation $d^n f(x)/dx^n \Leftrightarrow (ik)^n f(k)$ with $n = 1$, it is evident that this adds the advection term $-v\,\partial/\partial x$ to the space operator in the diffusion equation. Again we wish to emphasize that the same Fourier symbol $A(k) = -v(ik) + Da(ik)^\alpha + D(1-a)(-ik)^\alpha$ determines both the fractional advection/dispersion operator, and the corresponding Lévy process, including the case when $\alpha = 2$.

2.2. Lévy motion and Fractional Derivatives in Several Dimensions

The extension to multiple dimensions follows the same general approach. The limits of Markovian random walks define the derivative operators in multiple dimensions [**?** 22]. There are several scenarios. First, define the random walk jump magnitudes independent of direction by $P(R > r) \sim r^{-\alpha}$, and the probability of moving in any direction in d-dimensions by the random direction vector θ with probability measure $M(d\theta)$ on the unit sphere. If the random direction has a probability density $m(\theta)$, then the notation $M(d\theta) = m(\theta)d\theta$; otherwise, the discrete measure can be constructed by a sum of Dirac delta function terms (analogous to the probability mass function of a discrete random variable). The direction measure $M(d\theta)$ is often called the mixing measure. The random walk with these jumps converges [**?** , Theorem 6.21] to a Lévy motion with FT

$$p(\boldsymbol{k},t) = \exp\left[-t\langle i\boldsymbol{k},\boldsymbol{v}\rangle + Dt\int_{\|\theta\|=1}(\langle i\boldsymbol{k},\theta\rangle)^\alpha M(d\theta)\right], \tag{19}$$

where $\langle \boldsymbol{x},\boldsymbol{y}\rangle$ denotes the inner product of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. This model recovers the one dimensional case because the unit vectors would be $\theta = \pm 1$, the mixing measure $M(d\theta) = a\delta(\theta-1)d\theta + (1-a)\delta(\theta+1)d\theta$, the distribution of the forward and backward jumps is $M(+1) = a$ and $M(-1) = 1-a$, and the integral reduces to two terms: $a(ik)^\alpha + (1-a)(-ik)^\alpha$. The measure $M(d\theta)$ in multiple dimensions can be made as simple as a few Dirac delta functions on the coordinate axes or more elaborate to depict the superposition of flow directions [23].

Take time derivatives and invert the FT to get

$$\frac{\partial}{\partial t}p(\boldsymbol{x},t) = -\boldsymbol{v}\cdot\nabla p(\boldsymbol{x},t) + D\nabla_M^\alpha p(\boldsymbol{x},t). \tag{20}$$

The rightmost operator is an extended form of the original fractional Laplacian by Riesz [21], because it is a completely general mixture of fractional directional derivatives (explained in detail below). The point source solution $p(\boldsymbol{x},t)$ has Fourier transform $p(\boldsymbol{k},t) = \exp(tA(\boldsymbol{k}))$, where

$$A(\boldsymbol{k}) = -\boldsymbol{v}\cdot i\boldsymbol{k} + D\int_{\|\theta\|=1}(\langle i\boldsymbol{k},\theta\rangle)^\alpha M(d\theta).$$

Recall that the directional derivative is the inner product

$$\langle\theta,\nabla f(\boldsymbol{x})\rangle = \sum\theta_j\frac{\partial}{\partial x_j}f(\boldsymbol{x}) = \frac{d}{ds}g(s)$$

at $s = 0$ where $g(s) = f(\boldsymbol{x} + s\theta)$. Its FT is $\langle i\boldsymbol{k}, \theta\rangle f(\boldsymbol{k})$. Using the definition of a scalar positive fractional derivative (now in the radial coordinate r):

$$\frac{d^\alpha}{dr^\alpha} g(s) = \frac{1}{\Gamma(-\alpha)} \int_0^\infty r^{-1-\alpha} g(s-r) dr \tag{21}$$

The fractional directional derivative is this derivative evaluated at $s = 0$, and each directional derivative is weighted by its probability in every direction to get

$$\nabla_M^\alpha f(\boldsymbol{x}) = \frac{1}{\Gamma(-\alpha)} \int_{\|\theta\|=1} \int_0^\infty r^{-1-\alpha} f(\boldsymbol{x} - r\theta) dr M(d\theta). \tag{22}$$

The inner integral has FT $\langle i\boldsymbol{k}, \theta\rangle^\alpha$. So the fractional Laplacian is a mixture of directional fractional derivatives, i.e., a mixture of convolutions with a power law, the mixture defined by a directional probability measure (i.e., a density for continuous random variables). By virtue of (19), the fractional derivative once again is defined by the underlying Markovian Lévy motion.

The Grünwald finite difference formula can be directly applied to approximate certain cases of the mixing measure. For example, if there is only weight along the coordinate axes, the shifted Grünwald weights (17) can be used directly. The outer integral of (22) reduces to a sum along the components of $\boldsymbol{x}$. For a numerical solution using this idea, see [24]. If there is weight in-between the axes, the integer node locations no longer exist for all directions (i.e., nodes lying along the 45° direction are at a distance of $\sqrt{2}$ times the number of nodes away from the origin), hence the Grünwald weights have to be interpolated. For distances r larger than about 4 nodes in the range $1 < \alpha < 2$, the Grünwald weights closely follow the power law $r^{-1-\alpha}/\Gamma(-\alpha)$ (Fig. 2). A mathematical procedure for approximating the general fractional Laplacian $\nabla_M^\alpha f(\boldsymbol{x})$ with weight off the coordinate axes was detailed in [25], but numerical codes have yet to be implemented.

2.3. Operator Scaling and the Anisotropic Laplacian

There is no reason to expect that the power law index dictating the magnitude of large jumps must be the same in different directions. Using methods similar to the central limit theorem in multiple dimensions shows that up to d different power laws can persist in d-dimensions. One can construct a suitable random walk using matrix powers. Suppose the random variable R is characterized by $P(R > r) \sim r^{-1}$. One could transform this into an isotropic heavy-tailed random magnitude by taking its scalar power $R^{1/\alpha}$, which has a tail that now decays with $r^{-\alpha}$, or one can generate a jump with different tail parameters in different directions by taking the matrix power $R^{\boldsymbol{H}}$, where $\boldsymbol{H}$ is a $d \times d$ matrix. Taking the power of a martix (besides the obvious integer cases) is calculated analogous to the scalar power formulas $x^p = e^{p \log x}$ for real powers of positive real numbers. For matrix powers we have $R^{\boldsymbol{H}} = \exp(\boldsymbol{H} \log R)$ which expands using the matrix exponential $\exp(\boldsymbol{H}) = \boldsymbol{I} + \boldsymbol{H} + \boldsymbol{H}^2/2! + \cdots$ where $\boldsymbol{I}$ is the identity matrix. For reference, we use the symbol $\boldsymbol{H}$ because of the relationship to the classical Hurst coefficient (more on this in the next section). The

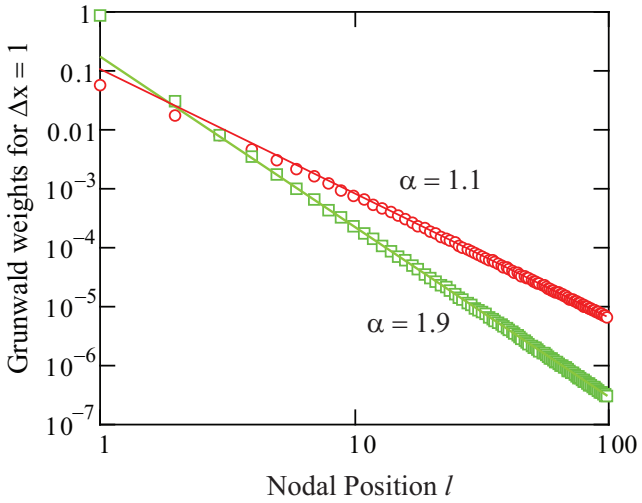


Figure 2: **Log-log plot of Grünwald (convolution) weights at integer node distances (symbols) and power law $x^{-1-\alpha}/\Gamma(-\alpha)$ interpolation (lines). $\alpha = 1.1$ and 1.9 are shown.**

matrix power creates larger or smaller exponents for the jump magnitudes in the eigenvector directions of $\boldsymbol{H}$. The direction of each jump is given once again by a random unit vector with distribution $M(d\theta)$. Adding up these jumps and rescaling appropriately—analogous to taking the scaling limit of a random walk to create Brownian motion—results in an operator Lévy motion [13, 26, 27]:

$$Z(t) = \sum_{i=1}^{[t/dt]} X_i = \sum_{i=1}^{[t/dt]} R_i^{\boldsymbol{H}} \cdot \theta_i, \tag{23}$$

where R_i and θ_i are independent.

As in the isotropic case (when $\boldsymbol{H} = \frac{1}{\alpha}\boldsymbol{I}$), the exponent of the random walk jumps is directly related to the order of the fractional derivatives that describe them. In the case of the matrix rescaled jumps, the order of the derivatives can be considered matrix-order as well. To illustrate the effect of the matrix scaling, consider a simple 2-d case where the two eigenvectors of $\boldsymbol{H}$ are orthogonal (or in other words, the primary directions of growth are perpendicular). Then the operator stable exponent dictates independent jumps:

$$R^{\boldsymbol{H}} = \begin{bmatrix} R^{H_1} & 0 \\ 0 & R^{H_2} \end{bmatrix} = \begin{bmatrix} R^{1/\alpha_1} & 0 \\ 0 & R^{1/\alpha_2} \end{bmatrix}.$$

Because $P(R > r) = r^{-1}$ for large values, the jump length probabilities on the k^{th} eigenvector of $\boldsymbol{H}$ fall off as $P(R^{1/\alpha_k} > r) = r^{-\alpha_k}$. The jump length probabilities for trajectories off the eigenvectors decay like a mixture of power laws. For an example, if we also restrict motion directions to the forward x- and

y-directions, then the corresponding fractional dispersion equation would take the form

$$\frac{\partial p(x,y,t)}{\partial t} = D_1 \frac{\partial^{\alpha_1} p(x,y,t)}{\partial x^{\alpha_1}} + D_2 \frac{\partial^{\alpha_2} p(x,y,t)}{\partial y^{\alpha_2}}, \tag{24}$$

and the Fourier symbol of this process $A(\boldsymbol{k}) = D_1(ik_1)^{\alpha_1} + D_2(ik_2)^{\alpha_2}$ also uniquely determines the underlying Lévy process: $p(\boldsymbol{k},t) = \exp(-tA(\boldsymbol{k}))$.

In general, the random walk (23) converges to an operator-Lévy motion with governing equation [**?** 27, 28]

$$\frac{\partial}{\partial t} p(\boldsymbol{x},t) = -\boldsymbol{v} \cdot \nabla p(\boldsymbol{x},t) + \nabla \cdot D_F \nabla p(\boldsymbol{x},t) + D \nabla_M^{\boldsymbol{A}} p(\boldsymbol{x},t), \tag{25}$$

where $\boldsymbol{A}$ is the inverse of $\boldsymbol{H}$. As opposed to (20), there can be a Fickian dispersion term in this equation, with D_F equal to $1/2$ times the covariance matrix of particle jumps per time. The multidimensional isotropic equation (20) assumes that motions in all directions have the same tail parameter, so that infinite-variance jumps (with $\alpha < 2$) occur in all directions. When jumps have different tail probabilities in different directions, there is room for Brownian motion in one direction and Lévy motion in another, hence the additional term in the anisotropic-order equation (25). In any given direction, either the Fickian dispersion term, or the fractional dispersion term, is zero, because only one of two possibilities (light-tailed or heavy-tailed random walk jumps) can apply. If all eigenvalues of $\boldsymbol{H}$ are greater than $1/2$, then jumps in all directions are heavy tailed, $D_F = 0$, and the Fickian portion disappears. Physically, this means that heavy-tailed jumps overwhelm thin-tailed ones. If all jumps in all directions are thin tailed, then the fractional dispersion term disappears. As mentioned previously, the matrix $\boldsymbol{H}$ is a scaling matrix that describes plume growth rates in all directions. In this way it is related to the classical Hurst coefficient, because the point source (Green's function) solution to equation (25) with $\boldsymbol{v} = 0$ is self-similar with a rescaling of time and space according to

$$p(\boldsymbol{x}, ct) = \|c^{-\boldsymbol{H}}\| p(c^{-\boldsymbol{H}} \boldsymbol{x}, t) \tag{26}$$

where $\|\cdot\|$ is the matrix determinant. Note that this includes the Fickian case where H is a scalar equal to $1/\alpha = 1/2$.

A very flexible Eulerian numerical solution to equation (25) could be achieved, along the lines laid out in [25]. The operator $\nabla_M^{\boldsymbol{A}}$ is defined by a convolution, see [28]. Then the finite-difference solution is a series of convolutions, each representing a time step. A similar methodology was used in [29] to create operator scaling conductivity fields. See Subsection 2.5 for additional discussion.

2.4. Divergence - Integer and Otherwise

Up to this point, we have assumed that the mean advective drift velocity $\boldsymbol{v}$ and the strength of the dispersion D have been homogeneous in space. For the mean drift this means that the divergence of the flux $\nabla \cdot \boldsymbol{v}p$ distributes like $\boldsymbol{v} \cdot \nabla p + p \nabla \cdot \boldsymbol{v}$. For divergence-free (incompressible) flow or first-order stationary processses, the second term is zero and there is no change to our

previous development. On the other hand, the traditional dispersion term in (25) can be viewed as the divergence of the particle flux. How can we view the fractional dispersion operator in (25) in terms of divergence (conservation of mass) and particle flux? In other words, what happens when the strength of dispersion varies in space? It turns out [30, 31] that the fractional dispersion may be derived in several different ways. If one starts from the microscopic expression of particle motion (i.e., the Ito or Langevin equations), the fractional Laplacian can be distributed in several ways.

To illustrate, in the case of scalar order α in multiple dimensions, one may take a classical integer divergence of a fractional dispersion ($\nabla \cdot D\nabla_M^{\alpha-1}$), or a fractional divergence of a classical integer-order flux ($\nabla_M^{\alpha-1} \cdot D\nabla$). Here we follow the typical abuse of notation: the generalized fractional Laplacian ∇_M^α is a scalar-valued operator that reduces to the Laplacian $\Delta = \nabla^2 = \nabla \cdot \nabla$ when $\alpha = 2$ and M is uniform, while the generalized fractional gradient $\nabla_M^{\alpha-1}$ is a vector-valued operator that reduces to the gradient ∇ when $\alpha = 2$ and M is concentrated on the positive coordinate axes. If the local dispersion coefficient is a constant, these are equivalent. The differences in the case of space-variable dispersivity $D = D(\boldsymbol{x})$ are subtle and small in many cases, but when the dispersion coefficient D has strong fluctuations, the difference can be significant. For illustration of the numerical methods in Section 2.6, we will concentrate here on the equation

$$\frac{\partial}{\partial t}p(\boldsymbol{x},t) = -\boldsymbol{v}\cdot\nabla p(\boldsymbol{x},t) + \nabla\cdot D_F\nabla p(\boldsymbol{x},t) + \nabla_M^{\boldsymbol{A}-\boldsymbol{I}} D(\boldsymbol{x})\nabla p(\boldsymbol{x},t), \qquad (27)$$

including the simpler forms when $\boldsymbol{A} - \boldsymbol{I}$ is the scalar $\alpha - 1$ in one or more dimensions. This formulation uses a fractional version of the conservation of mass equation: it implies that the change in probability, and by analogy concentration, is due to upstream differences in local advective flux. The magnitude of the local fractional dispersion coefficient $D(\boldsymbol{x})$, a scalar, is a measure of the difference between local mean velocity and the fluctuations of velocity [30, 31].

2.5. Simulating Spatial Fractional Derivatives: Eulerian Methods

As mentioned above, traditional finite difference methods can be thought of as discrete convolution formulas that lead to matrix equations [32, 33]. The local operators lead to sparse and banded matrix equations. Fractional-order equations are conceptually similar, except that the matrix of weights on other nodes is fuller, up to 100% full when the measure $M(d\theta)$ is non-zero everywhere. Because fuller matrices are typically solved iteratively, the fuller matrices should not pose tremendous numerical challenges. Many researchers are concentrating on efficient simulation of the fractional derivative operators (e.g., [34, 35] and references therein). To date, however, finite difference solutions for multidimensional fractional derivatives have been concentrated on the coordinate axes [33].

Fractional derivatives are linear operators; therefore, classical methods using finite elements can be adapted to solve fractional partial differential equations

[36, 37]. The finite element method hinges on the action of the linear operator on the chosen basis functions. That is why Dirac delta functions are commonly chosen as bases for the traditional integer order equations [38]. Roop's method uses polynomials for the bases, which, when properly chosen, simplify the calculation and implementation of the fractional derivative on the basis polynomials. This general procedure can be accelerated substantially, to the point where the fractional methods are not much more time-consuming to solve than the integer cases [35, 39]. To date, we are unaware of this method being applied to hydrological problems.

Eulerian approximations for fractional advection dispersion equations have been proven stable (when properly implemented) with a truncation error of order $(\Delta x)^2$, so they are nearly as robust as proven methods for classical diffusion. These methods may also suffer less from the truncation error associated with the advection term. The well-known phenomenon of numerical dispersion arises from simulation of the hyperbolic portion of the advection-dispersion operator, because the first truncated term is of the form $\Delta x\ v\ d^2/dx^2$. Therefore, traditional finite difference methods must keep the grid Peclet number $\Delta x\ v/D$ reasonably small. The fractional dispersion has heavier tails and greater spreading rates than the pseudo-Fickian numerical dispersion, so the constraints on the grid spacing may be reduced, although this has not, to our knowledge, been explored in detail.

2.6. Lagrangian (Particle) Methods

Particle-tracking methods became popular as a way to eliminate numerical dispersion, because each particle follows a characteristic curve (i.e., is an exact solution of the hyperbolic advection term [40]). Important research followed [41, 42, 43] concerning the solution of the expanded dispersion term $\nabla\cdot(D\nabla) = (\nabla\cdot D)\cdot\nabla + D\nabla^2$, primarily because geologic material may have very large, or even infinite gradients in the dispersion coefficient at sharp interfaces. These works highlighted the process involved in establishing the link between a (nonlinear) Langevin equation of instantaneous motion, the governing equation of that motion, and the link to the advection-dispersion equation that was the pre-supposed goal of the simulation.

In a series of papers [44, 31, 45], *Zhang* and coworkers defined the Langevin equations for motions that correspond to the multi-dimensional fractional ADEs. In particular, they showed the subtle differences in the random walks that correspond to the operators $D\nabla_M^{\boldsymbol{A}}$, $\nabla D\nabla_M^{\boldsymbol{A-I}}$, and $\nabla_M^{\boldsymbol{A-I}} D\nabla$, including the cases when $\boldsymbol{A}$ reduces to a scalar α and also in 1-d. The solutions are derived using the finding that $\nabla D\nabla_M^{\boldsymbol{A-I}}$, and $-\nabla_{\bar{M}}^{\boldsymbol{A-I}} D\nabla$ are adjoint operators for $\bar{M}(d\theta) = M(d(-\theta))$. For practical purposes, when variations of D are small, the differences in the solutions between these fomulas are reasonably small. The recognition and addition of the heavy-tailed dispersion in any case is the first-order effect.

To simulate the multi-scaling jumps, one distributes the initial condition, and subsequent sources of mass, into N particles, each of which follow a random walk approximation of (23) with finite time step Δt. A mass-weighted

histogram of particle positions gives the concentration. An Euler approximation of local advection $X(t+\Delta t) = X(t) + \boldsymbol{v}\Delta t$ or exact analytic methods [46] are used for the deterministic drift. For simplicity, we illustrate the case where the mixing measure $M(d\theta)$ is concentrated on the eigenvalue coordinates of the scaling matrix $\boldsymbol{H}$. For the random dispersion, as well as the effect of heterogeneous strength $D(\boldsymbol{x})$, one simply generates independent jumps in each eigenvector direction. This is conceptually similar to generating independent standard Gaussian longitudinal and transverse jumps to simulate classical dispersion. In the heavy-tailed case, the jump length of the particle along the eigenvector belonging to the k^{th} eigenvalue $1/\alpha_k$ of $\boldsymbol{H}$ can be calculated by generating the following random number [44, 45]

$$R^{1/\alpha_k} = D(\boldsymbol{x})^{\frac{1}{\alpha_k}} dL_{\alpha_k}(t) + \Theta \left| \frac{\partial D}{\partial x_k} \right|^{\frac{1}{\alpha_k - 1}} dL_{\alpha_k - 1}(t), \tag{28}$$

where k represents the direction of the k^{th} eigenvector of $\boldsymbol{H}$, $\Theta = \text{sign}(\partial D/\partial x_k)$, and $dL_\alpha(t)$ and $dL_{\alpha-1}(t)$ denote independent random noises underlying α-order and $(\alpha-1)$-order Lévy motions, respectively. These are generated by taking $dt^{1/\alpha}$ times a standard, maximally-skewed α_k-stable random variables with distribution $S_{\alpha_k}(\sigma = 1, \beta = +1, \mu = 0)$. The stable random variables can be generated exactly using the modified Chambers-Mallows-Stuck (CMS) method (for details, see the Appendix). The fractional dispersivity $D(\boldsymbol{x})$ must be first-order differentiable, so sharp interfaces are ruled out.

Generating Lévy-stable random variables is somewhat computationally expensive, so one can generate random vectors $R^{\boldsymbol{H}} \cdot \theta$ in the domain of attraction of the stables. This concept is similar in principle to classical random walk codes that use a Uniform $[-\sqrt{3}, \sqrt{3}]$ random variable as a substitute for a standard Gaussian: After as few as ten motions, the random walks with these jumps are indistinguishable from Brownian motion. The Langevin equation can be approximated using more easily-generated zero-mean random variables ξ with power law tails (Appendix). Once the heavy-tailed random variables ξ are generated and scaled as substitutes for the stable dL in (28), the jump contribution R^{1/α_k} in each eigenvector is specified. As for direction, if the mixing measure has a known or assumed distribution function $F(z) = P(\theta \leq z)$ on the unit circle, then the typical method using the inverse function on a Uniform [0,1] variable is used (as did [47]). Generate U, a Uniform [0,1] variable, and the direction vector $\theta = F^{-1}(U)$. Otherwise, the measure $M(d\theta)$ is discretized in m classes and summed to make the cumulative measure $M(\theta)$ via $P_m = \sum_{l=1}^{m} M(d\theta_l)$. Then the random direction vector in each case is $\theta = \theta_m$ if $P_{m-1} < U \leq P_M$.

Now represent the vector θ in terms of the unit eigenvectors (e_k) of $\boldsymbol{H}$: $\theta = \lambda_1 e_1 + \lambda_2 e_2 + \ldots$, and the final particle motion $R^{\boldsymbol{H}}\theta$ is given by the vector $\sum \lambda_k R^{1/\alpha_k}$. Several applications to field data are shown in Section 4.

3. Fractional Time Derivatives

Fractional time derivatives are formulated to respect causality, i.e., so that future events cannot affect the past. Therefore, the fractional time derivatives

are convolutions with a power law that is directional in time. The influence of events t units in the past decays with a power law. The fractional time derivatives take two different forms, based on the treatment of the initial condition. For clarity, we review both forms briefly here. A straightforward extension of the forward direction space derivative assumes that the function vanishes on $t < 0$:

$$\frac{d^\gamma f(t)}{dt^\gamma} = \frac{d^n}{dt^n}\frac{d^{\gamma-n} f(t)}{dt^{\gamma-n}} = \frac{d^n}{dt^n}\frac{d^{-(n-\gamma)} f(t)}{dt^{-(n-\gamma)}} = \frac{d^n}{dt^n}\int_0^t \frac{(t-y)^{n-\gamma-1}}{\Gamma(n-\gamma)} f(y)dy \tag{29}$$

This is called the Riemann-Liouville fractional derivative. Using the $R-L$ subscript for this formula and taking Laplace transforms, one finds

$$\mathcal{L}\left[\frac{d^\alpha f(t)}{dt^\alpha}\right]_{R-L} = s^\gamma \tilde{f}(s) + \sum_{k=0}^{n-1} s^k \frac{d^{\gamma-1-k}}{dt^{\gamma-1-k}} f(t)\big|_{t=0}. \tag{30}$$

For many applications, $0 < \gamma < 1$, so $n = 1$, the summation disappears and the γ^{th} derivative represents multiplication in Laplace space by the quantity s^γ, where s is the Laplace parameter. In other cases, the terms in the sum will disappear for most well-behaved functions because the fractional derivatives involve an integral from 0 to t evaluated at $t = 0$. Recall that the traditional derivative of integer order has a Laplace transform that involves values of the function and its lower order derivatives at time $t = 0$. Generalizing on this formula, Caputo [48] defined a new kind of fractional time derivative such that:

$$\mathcal{L}\left[\frac{d^\gamma f(t)}{dt^\gamma}\right]_C = s^\gamma \tilde{f}(s) + \sum_{k=0}^{n-1} s^{\gamma-1-k} \frac{d^k}{dt^k} f(t)\big|_{t=0} \tag{31}$$

where $n-1 < \gamma < n$. Factor out the term $s^{\gamma-n}$ and we see that the Caputo derivative (labelled with a subscript "C") is a convolution of a power law with the n^{th} integer derivative of a function:

$$\begin{aligned}\mathcal{L}\left[\frac{d^\gamma f(t)}{dt^\gamma}\right]_C &= s^{\gamma-n}\left(s^n \tilde{f}(s) + \sum_{k=0}^{n-1} s^{n-1-k}\frac{d^k}{dt^k} f(t=0)\right) \\ &= s^{\gamma-n}\mathcal{L}\left[\frac{d^n f(t)}{dt^n}\right]\end{aligned} \tag{32}$$

An inverse Laplace transform reveals the Caputo derivative in real space:

$$\left[\frac{d^\alpha f(t)}{dt^\alpha}\right]_C = \frac{t^{n-\alpha-1}}{\Gamma(n-\alpha)} \star \frac{d^n f(y)}{dy^n} = \int_0^t \frac{(t-y)^{n-\alpha-1}}{\Gamma(n-\alpha)} \frac{d^n f(y)}{dy^n} dy \tag{33}$$

In the usual case, where the terms under the sum in (30) vanish, the two types of derivatives are related by:

$$\left[\frac{d^\gamma f(t)}{dt^\gamma}\right]_{R-L} = \left[\frac{d^\gamma f(t)}{dt^\gamma}\right]_C + \sum_{k=0}^{n-1} \frac{t^{k-\gamma}}{\Gamma(k+1-\gamma)} \frac{d^k}{dt^k} f(t=0) \tag{34}$$

If $0 < \alpha < 1$, then $n = 1$, and

$$\left[\frac{d^\alpha f(t)}{dt^\alpha}\right]_{R-L} = \left[\frac{d^\alpha f(t)}{dt^\alpha}\right]_C + \frac{t^{-\alpha} f(0)}{\Gamma(1-\alpha)} \tag{35}$$

3.1. Fractional Time Derivatives and Random Walks

The classical random walk is typically defined by motions that are divided by equal duration "waits," and the passage to a continuous (Markov) motion process requires a subdivision of the motions into smaller and smaller independent jumps. *Montroll and Weiss* [49] defined a process, called a continuous time random walk (CTRW), in which the waiting times between particle jumps could have any distribution. A closely related topic called "subordination" was previously explored for continuous time Markov processes by Bochner [50] and Feller [51]. Because the original motion processes we are interested in are Markovian diffusions, we follow their development. The Markov particle motion process, whether Brownian motion or the many Lévy motion extensions in the previous sections, denoted $X(t)$, has density $p(x,t)$ governed by the Cauchy equation:

$$\frac{\partial}{\partial t} p(x,t) = A_x p(x,t). \tag{36}$$

The point source solution of equation (36) has FT $p(k,t) = e^{tA(k)}$. If the time a particle spends in motion during the epoch $[0,t]$ is a random variable $U(t)$, the resulting random particle location becomes $A(U(t))$. We will assume for simplicity that the amount of time actually spent in motion at any time is a continuous random variable with probability density $h(u,t)$. The density that now describes a particle's whereabouts, which we denote $q(x,t)$, is given by conditioning over all possible probabilities of the operational time u for the clock time t:

$$q(x,t) = \int_0^\infty p(x,u) h(u,t) du. \tag{37}$$

An explicit solution may be computed from the integral (37) if the density $h(u,t)$ of the operational time $U(t)$ at any clock time t can be found. We take two tacks: one that gives the governing equation of the limits of traditional CTRW, and another that gives the solution for a two-phase system in which the particles transfer between mobile and immobile phases.

3.1.1. Uncoupled CTRW

A CTRW is built on the model of each motion being separated by a single random waiting time W. If the waiting time and the subsequent motion are independent, the CTRW is called uncoupled. The jump sizes are often taken to be the limit of a large number of jumps (e.g., Gaussian as reflected in a second-order space derivative), so that the same large number of waiting times can be assumed to pass to their limit as well. It is straightforward to sum the waiting times, but more difficult to figure the inverse, which is the time spent in motion. The sum of the waiting times $T(n) = \sum_{i=1}^n W_i$ gives the time of the

n^{th} jump, while the operational time $U(t)$ relates to the number of jumps $N(t)$ that have occurred by time t. These random variables are inverses related by $\{N(t) \geq n\} = \{T(n) \leq t\}$. In the limit, the sum of random waiting times $T(n)$ converges to another Markov process $G(t)$, and this inverse relation becomes $P(U(t) \geq u) = P(G(u) \leq t)$. The density functions for $U(t)$ and $G(u)$, denoted $h(u,t)$ and $l(t,u)$ are then related by [52]:

$$h(u,t) = \frac{d}{du}\left(1 - \int_0^t l(\tau,u)d\tau\right). \tag{38}$$

If the individual waiting times W have a heavy tail, so $P(W > t) \sim Ct^{-\gamma}$ for some constant C, then similar to the limit of random walks in space, the density of the Lévy process $G(t)$ has Laplace transform $\mathcal{L}[l(t,u)] = l(s,u) = e^{-u\beta s^\gamma}$, where β is a scale parameter depending only on γ and C (Appendix). Taking Laplace transforms $t \mapsto s$ in (38) leads to

$$h(u,s) = \frac{d}{du}(-l(s,u)/s) = \beta s^{\gamma-1}e^{-u\beta s^\gamma}. \tag{39}$$

Take Laplace $t \mapsto s$ and Fourier $x \mapsto k$ transforms in (37) and use equation (39) along with $p(k,u) = e^{uA(k)}$ to get:

$$q(k,s) = \int_0^\infty p(k,u)h(u,s)du = \int_0^\infty e^{uA(k)}\ \beta s^{\gamma-1}e^{-u\beta s^\gamma}du. \tag{40}$$

Using $1/b = \int e^{-bu}du$, we have

$$q(k,s) = \frac{\beta s^{\gamma-1}}{\beta s^\gamma - A(k)} = \frac{s^{\gamma-1}}{s^\gamma - A(k)/\beta}. \tag{41}$$

Now invert the FT and LT (one at a time) to get the fractional-order limit equation for CTRW:

$$\frac{d^\gamma}{dt^\gamma}q(x,t) = \frac{1}{\beta}A_x q(x,t); \quad q(x,t=0) = \delta(x), \tag{42}$$

where we have used the Caputo fractional derivative (33). Note that the parameters in the space operator A_x (velocity and dispersion) are reduced by the factor β.

We are not aware of a similar development, using subordination, for coupled CTRW, i.e., the case where the size of particle jumps depends on the size of the preceding waiting time. This possibility was developed for CTRW by Scher and Lax [53]. The coupled CTRW have different rates of growth of moments of the Green function relative to the uncoupled CTRW [53, 54, 55]. Regarding the relationship of fractional calculus to coupled CTRW, certain functional forms of coupling can lead to more exotic governing equations with coupled space-time fractional derivatives, like $(d/dt + vd/dx)^\alpha$, see [19].

3.1.2. Mobile/Immobile Particles

Suppose that, between each waiting time in an immobile phase, the particle participates in the motion process for exponential random times [52]. The rescaled limit of the waiting times follows the same procedure as the classical CTRW above, but while the particle is in the mobile phase, the clock time and the operational time are ticking away at the same rate [52]. This shifts the limit of the waiting time density by adding $u = t$, which multiplies the Laplace transform by e^{us}: $\mathcal{L}[l(t,u)] = l(u,s) = e^{-us}e^{-u\beta s^{\gamma}}$. Now the operational time density is calculated as

$$h(u,s) = \frac{d}{du}\left(-l(u,s)/s\right) = (1+\beta s^{\gamma-1})e^{-u(s+\beta s^{\gamma})}. \tag{43}$$

The governing equation is calculated as before, by taking the FLT of (37) and substituting densities of $p(k,u)$ and $h(u,s)$:

$$q(k,s) = \frac{1+\beta s^{\gamma-1}}{s+\beta s^{\gamma} - A(k)} \tag{44}$$

Now invert the FLT for the real space governing equation of the limit of 2-state Mobile/Immobile processes:

$$\frac{d}{dt}q(x,t) + \beta\frac{d^{\gamma}f(t)}{dt^{\gamma}}q(x,t) = A_x q(x,t); \quad q(x,t=0) = \delta(x), \tag{45}$$

where once again we have used the Caputo fractional derivative.

It is not quite as simple to equate probability to concentration in this case. The single particle exists alternately, in two different states: mobile, while it actively participates in the motion process, and immobile between mobile epochs. The total probability of particle whereabouts is the sum of the mobile and immobile location probabilities. The FLT (44) has two terms in the numerator. These correspond, in a continuum sense, to the portion of the particles in the mobile and immobile phases, respectively. By combining these, equations (44) and (45) represent total resident concentration. This can be shown by considering the continuum multirate mass transfer with infinite mean, power-law random waiting times, which have the following three equations for concentration in 1) the "total" phase C_T, 2) the mobile phase C_M and 3) the immobile phase C_I [13, 52]:

$$\begin{aligned}
\frac{\partial C_T}{\partial t} + \beta\frac{\partial^{\gamma}C_T}{\partial t^{\gamma}} &= A_x C_T \\
\frac{\partial C_M}{\partial t} + \beta\frac{\partial^{\gamma}C_M}{\partial t^{\gamma}} &= A_x C_M - \frac{C_M(x,t=0)\beta t^{-\gamma}}{\Gamma(1-\gamma)} \\
\frac{\partial C_I}{\partial t} + \beta\frac{\partial^{\gamma}C_I}{\partial t^{\gamma}} &= A_x C_I + \frac{C_M(x,t=0)t^{-\gamma}}{\Gamma(1-\gamma)}
\end{aligned} \tag{46}$$

The total concentration $C_T = \theta_M C_M + \theta_I C_I$ where θ_M and θ_I are mobile and immobile porosities and β [$\mathrm{T}^{\gamma-1}$] is defined here as capacity coefficient (see

[11] for a translation of the many forms of mobile and immobile porosity and sorption). Equations (46) assume that all solute begins in the mobile phase: $C_I(x,0) = 0$. Taking the FLT of these equations and comparing to (44), it is convenient to define the two components of $q(k,t)$ by

$$\begin{aligned} q_M(k,s) &= \frac{1}{s + \beta s^\gamma - A(k)} \\ q_I(k,s) &= \frac{\beta s^{\gamma-1}}{s + \beta s^\gamma - A(k)} \end{aligned} \tag{47}$$

which partitions the particle density location function into mobile and immobile contributions because $q_M + q_I = q$. As a result, $\theta_M C_M(x,t) = q_M(x,t)$ is the mobile concentration, and $\theta_I C_I(x,t) = q_I(x,t)$ is the immobile concentration. The functions q_M and q_I are the portions of the total probabilities for a particle to occur at some point, hence these are related to the concentration in total (solid and liquid) aquifer material and must be adjusted by mobile and immobile porosities.

This gives a simple method to simulate the fractal multi-rate mass transfer process in a particle tracking routine [52, 56]: The mobile times are exponential with mean $1/(\beta\lambda)$, where in this case λ is a tuning parameter to ensure enough transitions between mobile and immobile by the time of interest. If the particles' locations are desired at some time t, choose λ so that the mean mobile step is approximately one tenth of this: $\lambda > 10/(t\beta)$. For waiting times, use either the chopped (Appendix) or shifted Pareto following $P(W > t) = S^\gamma(t+S)^{-\gamma}$, generated by $W = SU^{-1/\gamma} - S$, where S is the shift (which controls the scale of the waits) and U is a Uniform [0,1] random variable. To derive the proper shift S, start by setting $\lambda = 1$, then by the Appendix in [52], we need W to be in the domain of attraction of a standard [81] stable law. By [**?** , Theorem 3.37, Proposition 5.8], $S = (\Gamma(1-\gamma)\cos(\pi\gamma/2))^{-1/\gamma}$. Changing the value of λ does not change the overall solution but makes each mobile sojourn shorter, hence places more alternating mobile/immobile phase changes in any given time step. There are λ times as many mobile and immobile episodes compared to $\lambda = 1$, so to get the same overall process, each W is multiplied by $\lambda^{-1/\gamma}$ because for stable random variables $\lambda^{-1/\gamma}\sum_{i=1}^{[\lambda n]} W_i$ has the same distribution as $\sum_{i=1}^n W_i$. Therefore for any λ, the shift $S = (\lambda\Gamma(1-\gamma)\cos(\pi\gamma/2))^{-1/\gamma}$.

4. Fractional ADEs and Field-Scale Tracer Test Data

The first field application of the spatial fractional equation to tracer test data was the Cape Cod bromide plume, motivated by the apparent power-law growth of the plume's dispersivity. This model successfully replaced the time-variable dispersivity in the traditional ADE by a constant parameter in the fractional ADE [7]. That model used a symmetric mix of forward and backward spatial fractional derivatives ($a = 0.5$ in (11)). The use of a symmetric fractional ADE was criticized by *Zhang et al.* [58]. They maintain that a more proper way to account for solute spreading behind the mean is by trapping in low permeability

zones, rather than long journeys upstream. *Zhang et al.* [58] modeled the Cape Cod data with a forward space and time fractional ADE. The forward space derivative term models the leading plume edge, and the fractional time derivative term models the trailing edge. That method yielded improved fits, although the relative homogeneity in the hydraulic conductivity ($\mathrm{VAR}(\ln[K]) = 0.26$) at Cape Cod makes heavy tailed particle jumps less important. This is reflected in the fitted value $\alpha = 1.6$ in that model: Because the space fractional ADE reduces to the traditional ADE at $\alpha = 2$, the effect of setting $\alpha = 1.6$ is a small increase, relative to other sites, in the heavy leading plume edge.

Benson et al. [8] then examined the bromide and tritium plumes at the MADE site. Their analysis was done in 1-d using the maximum concentrations along the "core" of the plume. The higher $\ln(K)$ variance, recently measured to be on the order of seven [59], made the one-sided space-fractional ADE an attractive model, and some simple analyses of the K statistics allowed an *a priori* estimate of all equation parameters. The fitted value of $\alpha = 1.1$ indicates a heavy leading plume edge, reflecting a highly heterogeneous K field. The space-fractional ADE produced a good fit to normalized concentration snapshots (Fig. 3). *Schumer et al.* [13] applied an MRMT equation with $\gamma = 0.33$ and $\beta = 0.08\ \mathrm{d}^{-0.67}$ to explain the bromide plume zeroth spatial moment (total mobile mass) decline. The fit of the zeroth moment was improved over single-rate methods [60]. *Zhang et al.* [58] used this estimate to fit the MADE plumes in 1-d, using (11) in (45), and found that both spatial and temporal non-locality were important. A further analysis of the centered second moment of the MADE plume showed different growth rates in the longitudinal versus transverse directions [27], indicating the need for a multiscaling fractional dispersion term. *Zhang et al.* [23] used a 3-zone model with longitudinal derivative of order $\alpha = 1.1$ and transverse of order $\alpha = 1.5$ and included the fractal mobile/immobile parameters previously reported [13]. For the mixing measure, they assumed a braided stream network and derived the proportion of overlapping sinuous channels that point in any direction [23]. Their particle-tracking simulation (Fig. 4) is a reasonably faithful representation of the plume with comparatively few parameters—especially, a constant mean velocity $\boldsymbol{v}$. In particular, the spatial discretization (of, say, K) is vastly reduced. The non-local fractional derivatives are designed to replace finer-scaled velocity information and allow much coarser discretization. This concept has been demonstrated on an intensively studied 30.5×30.5 cm sandstone slab, where an analytic solution of a fractional PDE captures the important features of a plume that over 8,000 measured K values fail to reproduce when used in the classical ADE [56].

In related research, *Harman et al.* [61] examined water transport through hillslopes by assuming that, unabated, a parcel of water flows according to piston (wave equation) flow. But the parcels of water may be trapped in the heterogeneous K field for random, heavy-tailed amounts of time, giving a fractional-in-time wave-type equation. The solutions of the equation match numerical solutions of water flow through mildly to strongly heterogeneous hillslope material.

Bradley et al. [62] and *Ganti et al.* [63, 64] looked at experimental and

theoretical evidence for heavy-tailed transport and fractional ADEs for sand and gravel bedload transport in rivers. *Foufoula-Giorgiou et al.* [65] extended this concept to overland sediment transport and the evolution of landscapes. The transport equations take the exact forms presented herein. These processes are reviewed by *Schumer et al.* [66].

5. Fractional Integration: fBm Random Fields and Extensions

An integral is also called an antiderivative, with good reason. It is designed to be the inverse operator of a derivative. This is simply illustrated by the Fourier transform relation $\int f'(x)dx = f(x) \Leftrightarrow (ik)^{-1}(ik)f(k) = f(k)$. As we showed in previous sections, fractional derivatives in several dimensions can apply a different order of fractional differentiation in each coordinate, using the Fourier symbols of different Lévy motions. We have called this Fourier picture $A(\boldsymbol{k})$, and we have observed that multiplication by $A(\boldsymbol{k})$ in Fourier space defines the multidimensional fractional derivative. It follows that the fractional antiderivative corresponds to division in Fourier space by $A(\boldsymbol{k})$, in all of its forms presented above.

Fractional Brownian motion (fBm) in 1-D was originally defined as a weighted sum of prior values of white noise, where the weights fall off like a power law. Although the forward-direction fractional integral of white noise $B(t)$ diverges:

$$\tilde{B}_H(x) = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{x} (x-y)^{H-1/2} B(dy), \tag{48}$$

the difference of two fractional integrals $B_H(t) = \tilde{B}_H(t) - \tilde{B}_H(0)$ is a legitimate stochastic integral that converges [67]. Define for compact notation $E = H+1/2$ where the Hurst scaling index $0 < H < 1$, and equation (48) has the form of a fractional integral of order $1/2 < E < 3/2$. If the random measure $B(dy)$ were replaced by $f(y)dy$ for a suitable non-random function, then the fractional integral (48) would be the inverse FT of $(ik)^{-E}f(k)$. The same kind of finite sum approximation that is used to approximate the fractional derivatives is also possible with the stochastic integral (48), and this idea can be used to efficiently simulate an unconditioned fractional Brownian motion, as outlined in [72, 68]. Briefly, the stochastic integral (48) becomes a discrete convolution of a sequence of iid normal random variables $B(\Delta y)$ with the Grünwald weights corresponding to fractional integration (Fig. 1), evaluated by taking the FT of both, multiplying, and inverting. This uses the fact that the FT of a convolution is a product.

For space-functions, the (causal) positive fractional integral can be sensibly extended to a symmetrically weighted sum of positive and negative fractional integrals. The first multi-dimensional fractional Brownian motions were constructed by taking a power of the wave vector defined by a Fourier multiplier $|\boldsymbol{k}|^{-E}$ where $E = H + d/2$. A form of anisotropy can be implemeted by specifying a simple stretching in orthogonal directions by using fractional integration with Fourier symbol $|\lambda \cdot \boldsymbol{k}|^{-E}$ where the vector λ controls the correlation length

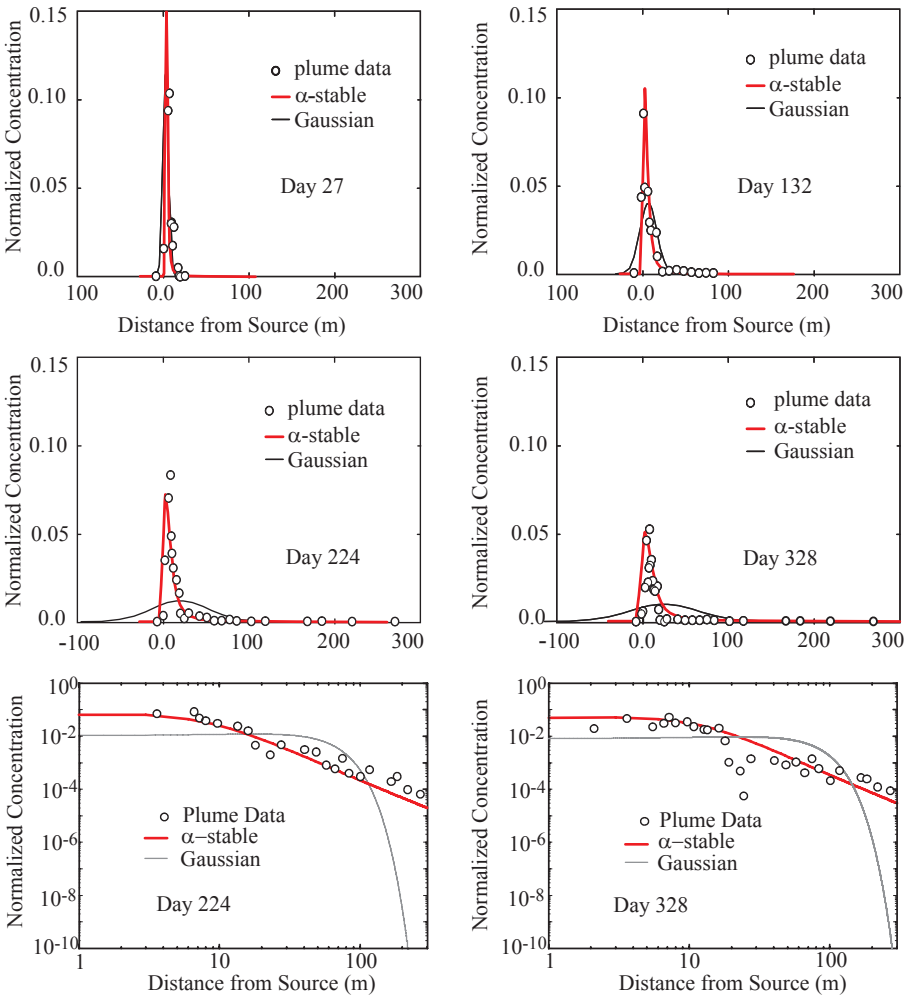


Figure 3: **Two spatial snapshots of the MADE tritium plume "core." Concentrations normalized to unit total mass. Bottom two plots are log-log scales. After [8].**

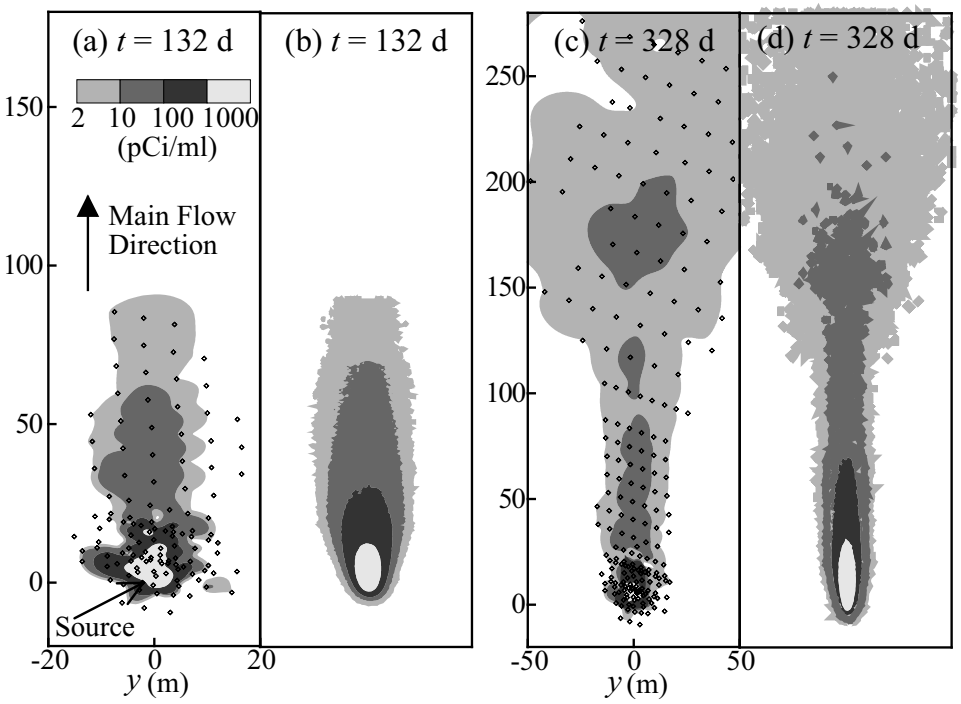


Figure 4: **(a,c) 2-d map-view snapshots of the MADE tritium plume and (b,d) particle tracking simulations. The simulations use a single velocity value and three zones of different dispersion strength and mixing measure. After [23].**

of the increments [69, 70, 71, 72]. Similarly, one could define a fixed distance and measure the value of the correlation. This corresponds to a radial mixing measure as defined above for the fractional derivatives. The stretching by vector λ just described represents an elliptical set of weights in the mixing measure [68]. Once again the mixing measure is a completely user-defined probability distribution on the unit sphere, so that correlation of the increments can be restricted to any set number of directions (Fig. 5). All of the fields constructed in this manner have the same Hurst index in every coordinate, but a different correlation length (or strength). A log-log plot of the correlation of increments in any direction would have the same power law slope, but different magnitudes.

However, many naturally occurring fields (e.g., K fields in alluvial aquifers) exhibit a different form of anisotropy, in which the Hurst index is different in each coordinate [68, 72, 73, 74, 75, 76, 77, 78]. These fields can be described using a multi-scaling fractional derivative, whose Fourier symbol $A(\boldsymbol{k})$ can be explicitly computed from the Lévy representation of the corresponding operator stable Lévy process [57, 28]. This Fourier symbol is characterized by its matrix-scaling property $c^E A(\boldsymbol{k}) = A(c^{\boldsymbol{Q}}\boldsymbol{k})$. In other words, the function is scale invariant only when stretched different amounts in different directions. The corresponding multi-scaling fractional integration has Fourier multiplier $\psi(\boldsymbol{k}) = A(\boldsymbol{k})^{-1}$, so that $\psi(c^{\boldsymbol{Q}}\boldsymbol{k}) = c^{-E}\psi(\boldsymbol{k})$. To make the parametrization unique, we require $\text{trace}(\boldsymbol{Q}) = d$, the number of dimensions. Then the matrix $\boldsymbol{Q}$ codes deviations from the overall order of fractional integration, and isotropic scaling has $\boldsymbol{Q} = \boldsymbol{I}$. The multi-scaling random field $B_\psi(\boldsymbol{x})$ constructed using this filter can be simulated in exactly the same way as a fractional Brownian motion, using the discrete Fourier transform in d dimensions. In the isotropic case with $E = H + d/2$, this random field scales according to $B_\psi(c\boldsymbol{x}) = c^H B_\psi(\boldsymbol{x})$ [68, 29, 79], consistent with isotropic fBm. Including the possibility of anisotropic scaling by giving $\boldsymbol{Q}$ different eigenvalues, we have the general scaling relationship:

$$B_\psi(c^{\boldsymbol{Q}}\boldsymbol{x}) = c^H B_\psi(\boldsymbol{x}) \tag{49}$$

If $\boldsymbol{Q}\boldsymbol{k}_j = q_j\boldsymbol{k}_j$, then the coordinate process $B_\psi(x_j)$ is an fBm with Hurst index H/q_j. Larger variations in the eigenvalues of the matrix $\boldsymbol{Q}$ describe more strongly anisotropic fields, with a different Hurst index in each coordinate.

5.1. Conditioned Random Fields: Numerics

Creating an unconditioned operator-scaling Gaussian random fields $B_\psi(\boldsymbol{x})$ is a simple matter, once the user has defined the mixing measure $M(d\theta)$ on the unit sphere (e.g., Fig. 5), and the possibly unique Hurst index H in each coordinate (Fig. 6). These unconditioned fields may be constructed several ways, each using the fact that a convolution is taking place. One may either construct the function $\psi(\boldsymbol{x})$ and take its fast Fourier transform (FFT) or directly construct $\psi(\boldsymbol{k})$. Similarly, one may either construct a same-sized white noise field of uncorrelated Gaussian random variables and take the FFT, or construct the FT of white noise directly using the spectral representation of an uncorrelated Gaussian field [80, 81]. The product of $\psi(\boldsymbol{k})$ and $B(\boldsymbol{k})$ is taken and inverse

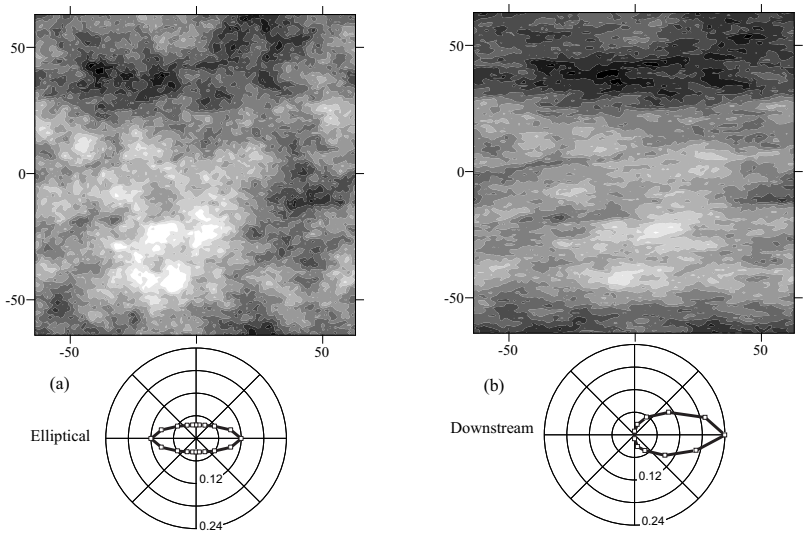


Figure 5: **Isotropically scaling fBm fields generated with identical input noise $B(\boldsymbol{x})$ and an single scalar $H = 0.3$. (a) Elliptical weights, and (b) downstream weighted mixing measure estimated from braided stream photograph. After [68].**

transformed. Only a portion of the field is retained due to periodicity of the FFT routine.

In real-world applications, the hydraulic conductivity (K) field is the most important control on the motion of conservative solutes. Furthermore, it is very common that some K data is collected, and any interpolation based on random field generation should honor these measurements. Creating a conditional field is more complicated (and significantly more time consuming) because the convolution algorithm changes any points specified before the convolution. Based on the research of Revielle [82], we recommend two useful conditioning methods: The first, based on the discussion in Feller [51], is called orthographic projection; the second, an adaptation of a method proposed by Journel and Huijbregts [83], is called random bridging.

5.1.1. Conditioning By Orthographic Projection

The orthographic projection method relies on the conditional probability distribution of any unknown point, based on a set of known points. Similar to existing sequential simulation methods [84, 85], initially the conditioning points comprise the known list, and the points to be simulated comprise the unknown list. The known list is used to sequentially create conditional probability distributions for each point within the unknown list. Once an unknown point has been simulated, it is added to the known list, and can be used to simulate subsequent unknown points. This process of simulating unknown points, and placing them on the known list, is continued until each point has been estimated and

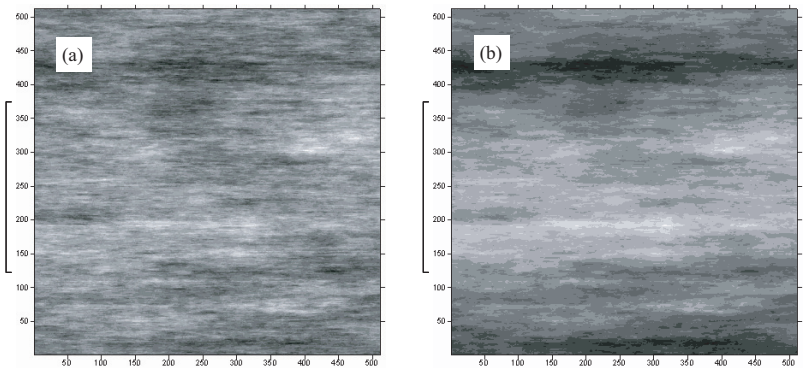


Figure 6: **Operator-scaling fBm fields generated with identical input noise $B(\boldsymbol{x})$ using the same H in the horizontal direction and different H in the vertical: (a) H vertical = 0.4 and (b) H vertical = 0.8. After [68].**

placed into the known list.

Without loss of generality we will consider the creation of a zero-mean random field. Make a discrete approximation $X_i = \sum_j \psi(j)B(i-j) \approx B_\psi(\boldsymbol{x}_i)$ of the stochastic integral that defines the random field, where the filter $\psi(j)$ comes from inverting the Fourier symbol $\psi(\boldsymbol{k})$, and $B(i)$ is an uncorrelated Gaussian sequence with mean 0 and variance σ_B^2. Then the covariance is

$$E[X_j X_i] = E\left[\sum_k \psi(k)B(j-k)\sum_l \psi(l)B(i-l)\right] \tag{50}$$

where $E[\,\cdot\,]$ is the expectation. Because $E[B(i)B(j)] = 0$ for $i \neq j$, the only nonzero terms in the sum occur when $j-k = i-l$, i.e., when $l = i-j+k$. Then the covariance reduces to:

$$\begin{aligned} E\left[X_j X_i\right] &= E\left[\sum_k \psi(k)B(j-k)\psi(i-j+k)B(j-k)\right] \\ &= \sigma_B^2 \sum_k \psi(k)\psi(i-j+k) \end{aligned} \tag{51}$$

This expression, a discrete convolution, is used to determine the covariance of two points. For larger fields, the covariance can be calculated efficiently using a discrete FFT to evaluate the convolution. Since the field has stationary increments, on a regular grid, the computation need be performed only once. These covariance values can then be used to simulate unknown points in the orthographic projection. Given known points $X_1, X_2, ..., X_{n-1}$, to simulate an unknown point x_n, compute the covariance matrix

$$M_n = \begin{bmatrix} E[X_1X_1] & \dots & E[X_1X_n] \\ \vdots & & \vdots \\ E[X_nX_1] & \dots & E[X_nX_n] \end{bmatrix} \tag{52}$$

and the inverse covariance matrix $Q = [q_{ij}] = M_n^{-1}$. Then simulate the unknown point by drawing a random Gaussian variate with mean and variance

$$\mu_n = a_1 X_1 + ... + a_{n-1} X_{n-1} \tag{53}$$

$$\sigma_n^2 = 1/q_{nn} \tag{54}$$

where $a_i = -q_{in}/q_{nn}$ and q_{ij} is the i, j entry of the inverse covariance matrix Q (see *F*eller [III.6] [51]). Once this unknown point is simulated, it is added to the known list, and can be used to simulate the remaining unknown points. As the list of known points grows, a subset of points must be selected from the known list, otherwise inverting the M_n matrix becomes inefficient. As a result, the known points included within each M_n must be carefully selected, in order to capture all important correlations present, yet remain efficient. Note also that this algorithm is very closely related to simple kriging: the mean value at a point is similar to a kriged surface, except that the "known" point list can (and almost certainly will) contain points that were not measured but were previously simulated.

There are several issues concerning this algorithm that influence the speed of calculation and characteristics of the output field. The first is choosing the order of points to be created. *Painter* [85] suggests moving sequentially through the field. [82] found that combining the nearest neighbor search and a random spiral search for more distant points was efficient and accurate. *Painter* [86] discusses the potential drawbacks of the various methods, and we find that the appearence of these field is very sensitive to the details of the order of simulation.

The procedure used to determine which known points are used to create each unknown point is also extremely important to the accuracy and efficiency of the output field. First, the number of points used to simulate each unknown point must be determined. Due to the computational inefficiency of using the entire known list to condition each unknown point, a selected portion of the known points is required. Known points must be selected which capture the correlations present at all scales, but also keep the size of the covariance matrix M_n manageable. A range of n from 20 to 50 was found to produce both accurate and efficient realizations and is consistent with previous work [85, 86, 87]. To select a representative set of $X_1, X_2, ..., X_{n-1}$ in 1-d, the only decision variable is the lag distance from the point of interest X_n. Therefore, to capture the correlations present at all scales in 1-d, the distance from each known point to X_n is calculated. When an isotropic kernel is used, only the absolute distance is required. When the kernel is anisotropic in multiple-dimensions, it is important to select points not only based on distance but also on orientation to the point of interest.

5.1.2. Conditioning By Random Bridging

The method of random bridging is similar to the conditional simulation method of Journel and Huijbregts [83]. On the surface it is more computationally demanding than orthographic projection because it requires kriging two mean surfaces — a procedure "built in" to the orthographic projection method.

However, when creating multiple realizations, the relative workload decreases for the random bridging method. For this reason, we typically prefer this method. We also find that it is insensitive to various numerical details such as the selection of conditioning points. MATLAB implementations of both methods in 3-d are available from the author.

Initially, the method begins with a set of conditioning values Z_i at locations X_i, and an unconditional realization, $B_\psi(\boldsymbol{x})$. The unconditional realization is then conditioned by "molding" the noise from the unconditional surface to a "mean" surface interpolated through the conditioning points. This method requires finding an apparent mean or trend surface between conditioning points, and a similar trend surface within the unconditioned realization.

Determining the deviation from a trend surface in the unconditional realization can be implemented by first interpolating the trend between the points $B_\psi(X_i)$ of the unconditional realization. This interpolated surface is called $b(\boldsymbol{x})$, and is best created using an unbiased process similar to kriging [83]. In our case, the covariance matrix is a convolution of the function $\psi(j)$ with itself. Using the same interpolation routine, the interpolated surface (call it $z(\boldsymbol{x})$) between the conditioning points Z_i is created. Finally, the conditioned field is created by taking the difference between the unconditioned field and its trend surface, and adding this to the trend surface from the measured points:

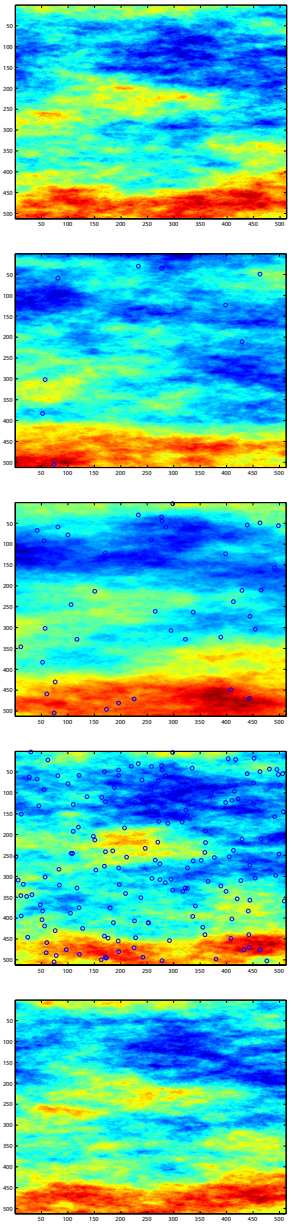
$$B_\psi(\boldsymbol{x})' = z(\boldsymbol{x}) + [B_\psi(\boldsymbol{x}) - b(\boldsymbol{x})] \tag{55}$$

5.2. Conditional osfBm

The orthographic projection and random bridging algorithms remain the same in multiple dimensions, as far as using an inverted covariance matrix to estimate μ_n and σ_n^2 at each unknown point. The only change occurs in how points are selected to construct M_n in each algorithm. The method of searching through a field and selecting the best points to create each M_n has significant consequence on the efficiency of creating conditional fields in both algorithms. Due to the slow power-law covariance decay, every point within the field has an effect on every other. Ideally, every previously simulated point would be used to simulate the next in order to account for the infinite correlation length. However, as the field size grows, a computational ceiling is quickly reached. Therefore, a subset of known points must be chosen to create each M_n. We have implemented a search based on sorting points based on their value of covariance with the point being simulated (or kriged). Examples of 2-d fields conditioned using the two algorithms are shown in Figure 7. It is clear that, with a knowledge of the correlation structure, defined by the filter $\psi(j)$, and a small number of conditioning points, faithful conditioned fields can be created that maintain the fractal structure at all scales.

5.3. Fractional Differencing of fBm to determine H

One of the handy features of the Grünwald weights is that it is very easy to perform fractional differences as well as fractional sums. Because fBm is a fractional integral of uncorrelated Gaussian noise, performing the "correct"



a) Training image from which conditioning data is taken

b) 10 conditioning points (circles), representing 0.0038% by volume.

c) 41 conditioning points (circles), representing 0.0156% by volume.

d) 164 conditioning points (circles), representing 0.063% by volume.

e) 655 conditioning points (not shown), representing 0.25% by volume.

Figure 7: **Conditional osfBm realizations created using $H = 0.4, q_1 = 0.9, q_2 = 1.1$; therefore, $H_{horizontal} = 0.44$ and $H_{vertical} = 0.36$. The mixing measure was fixed delta functions with weights of unity on the vertical axes and 3 on the horizontal axes. The data used to condition each field is from the training field shown at the top. Each figure shows the locations of conditioning points using blue circles, except for (e), in which the 655 points are omitted for clarity.**

order fractional derivative on fBm will give back an uncorrelated signal. This can aid in estimating the Hurst coefficient. In the discrete data case, following [88, 89], use (16) and (17) on a spatial series of hydraulic conductivity K_n. We took the 93×93 K values measured on a slab of Massillon sandstone (see also [90, 91, 56] and analyzed each of the 93 columns of data and separately the 93 rows of data, using a differencing distance of $N = 10$. Prior to differencing, the data in both columns and rows is highly correlated (Fig. 8, top plot). Here we show only the column data. Applying progressively higher-order fractional differences reduces correlation, until too high an order d induces statistically significant anticorrelation at lag one (Fig. 8, bottom plot). The differenced data also show more Normality (Fig. 9), although some heavier-tailed K data still exists at all levels of differencing.

6. Conclusions and Recommendations

This paper represents a survey of fractional calculus methods in hydrology with a few twists. First, we use the limit Markovian diffusions to define the full suite of fractional differential operators. This follows the logic that Brownian motion is an extremely useful model of diffusion and dispersion, in part because it represents a limit distribution, but also because it generates a solvable governing equation. When the basic tenets of Brownian motion are violated, some extended models are often similarly useful. When individual random motions follow a distribution with a power law tail, the motions converge to Lévy motion, which generates a diffusion equation with fractional space derivatives. In d-dimensions, up to d unique fractional derivative orders (including the classical second) may coexist. These orders dictate the rate of plume growth. If solute migrates into relatively immobile phase(s) and is released at a power law rate, then the diffusion equation's first derivative is either replaced, or joined, by a time-fractional derivative. This modifies the plume growth rate and simulates the power-law decay often seen in breakthrough curves of conservative tracers.

The fractional derivative operators are linear and admit Eulerian approximations similar to classical finite differences, but with fuller matrices. Lagrangian techniques may also be used in the same vein as classical random walk particle tracking codes are used to simulate classical diffusion. In all cases the simulations are more taxing, but the non-local fractional derivatives are designed to replace finer-scaled velocity information (e.g., [56]) and should require much coarser discretization—in some cases analytic solutions capture the important features of real solute plumes in aquifers.

Another approach to simulating transport in heterogeneous media is built on the creation of aquifer facsimilies. An attractive model is based on fractional Brownian motion because of the heterogeneity present at all scales and long-range correlation. Because fractional Brownian motion is built using fractional integrals of Gaussian noise, we are able to extend the classical isotropic fBm to have different Hurst coefficients in different directions. The fields are created by using the inverse operation specified by the multi-dimensional fractional derivative operators that we described previously. These fields can be conditioned

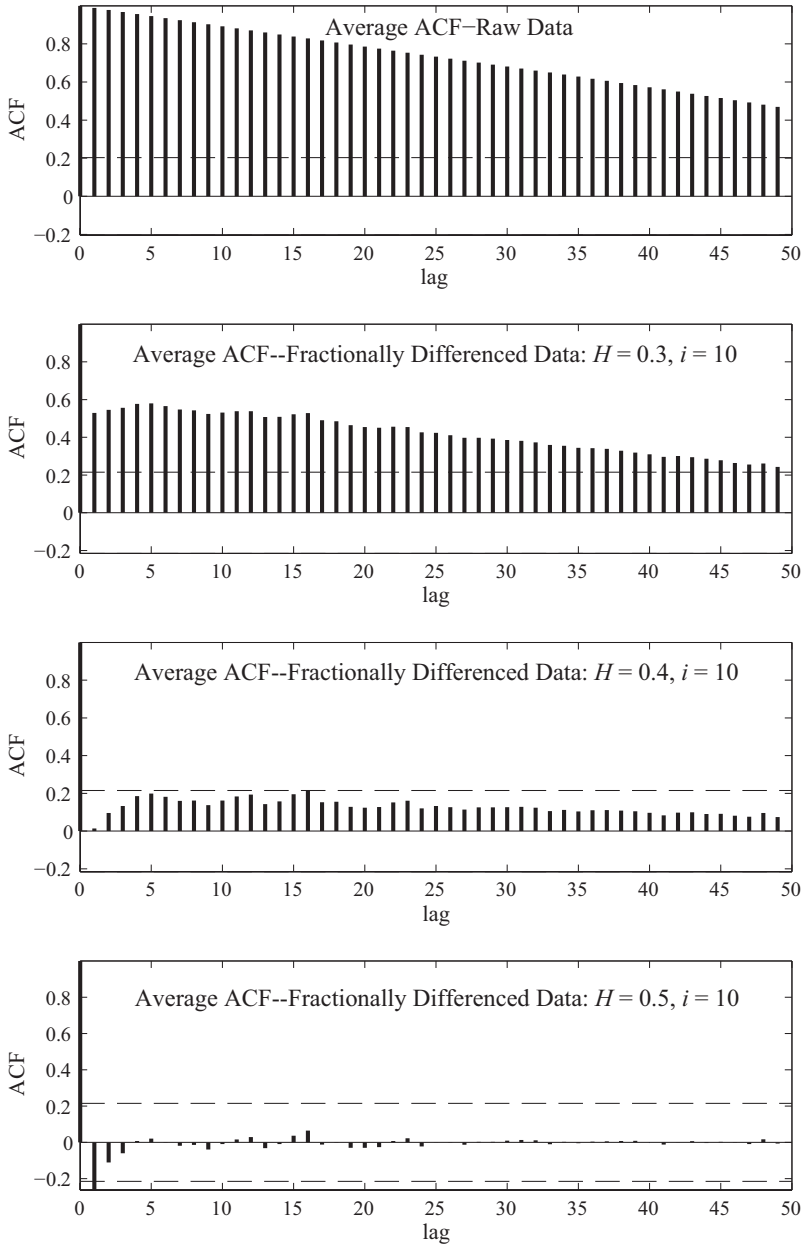


Figure 8: Autocorrelation function (ACF) for columns of $\log_{10}(K)$ data after fractional differencing. The ACF are calculated for each of the 93 columns and averaged. The 95% confidence interval is shown by the dashed lines. Differencing on the order of 0.9 to 1.0 ($H = 0.4$ to 0.5) reduces autocorrelation at all lags to non-significant levels. After [56].

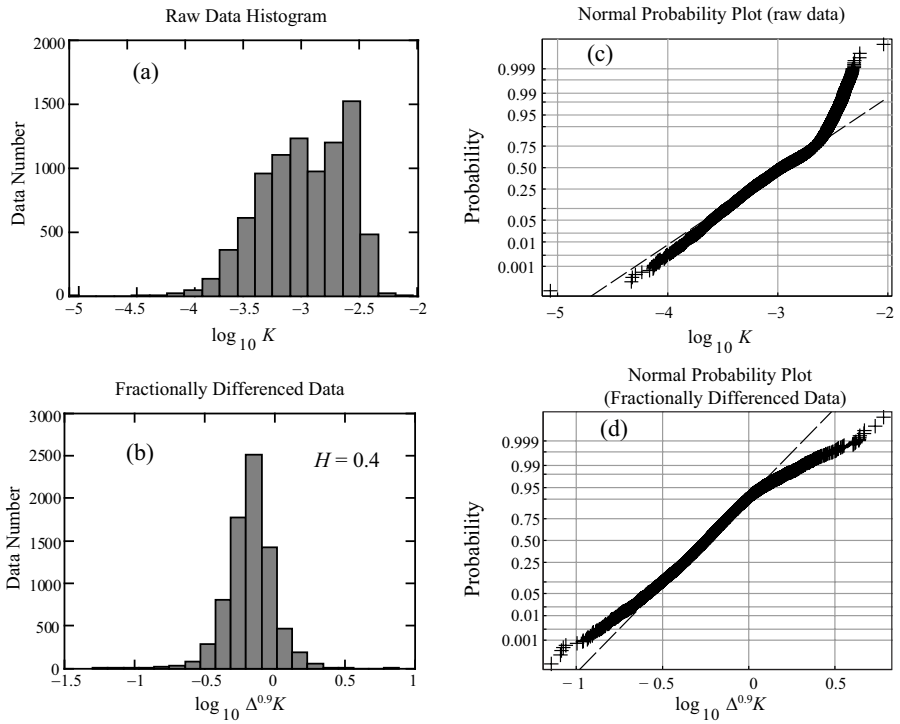


Figure 9: Histograms and Normal probability plots for original and differenced data from slab columns. (a) and (c): The raw $\log_{10}(K)$ data (K in cm/s) is significantly nonGaussian. (b) and (d): Following 0.9-order differencing (H=0.4), the data is more Normal, although somewhat heavy-tailed. After [56].

by measured data to create faithful representations of aquifer material—if it has the long-range dependence inherent in fBm—with the added flexibility of variable Hurst coefficients and user-defined weights in all directions. An open question is whether a link exists between these fractal K fields and the possible fractional PDE that describes the transport of a passive scalar within.

7. Acknowledgments

We thank T. Ginn and two anonymous reviewers for thoughtful analyses of our paper, and the editors D. A. Barry and C. Miller for inviting us to contribute to the 35^{th} anniversary issue of *Advances in Water Resources*. DAB was supported by NSF grants EAR–9980489, DMS–0139943, DMS–0417972, DMS–0539176, EAR–0749035, and USDOE Basic Energy Sciences grant DE-FG02-07ER15841. MMM was partially supported by NSF grants DMS-1025486, DMS-0803360, EAR-0823965, and NIH grant R01-EB012079-01. Any opinions, findings, conclusions, or recommendations do not necessary reflect the views of the NSF, DOE, or NIH.

Appendix A. Generating Heavy-Tailed Random Variables

There are several methods to generate approximately stable random walk variables ξ, for example the shifted Pareto $P(\xi > r) = s^\alpha (r+s)^{-\alpha}$. However, the density is too peaked at the origin and convergence is slow [92]. One can choose the modified Chambers-Mallows-Stuck (CMS) method to generate exactly the stable random variables [93], but calculations are slowed by numerous sine and cosine calls. Instead we recommend jumps R from a "chopped" Pareto distribution (e.g., Fig. A.10d).

Here we review the CMS method for max-skewed stable variables. Following [93], for $\alpha \neq 1$, generate V distributed uniformly on $(-\pi/2, \pi/2)$ and an independent exponential random variable W with mean 1. Then ξ is standard max skewed α-stable generated by:

$$\xi = \left(\cos\left(\frac{\pi\alpha}{2}\right) \right)^{-1/(2\alpha)} \cdot \frac{\sin[\alpha(V+\pi/2)]}{[\cos(V)]^{1/\alpha}} \cdot \left[\frac{\cos[V - \alpha(V+\pi/2)]}{W} \right]^{(1-\alpha)/\alpha} \tag{A.1}$$

A more computationally efficient approximation can be had. Following [45] use the distribution

$$P(\xi < r) = \begin{cases} m(r-p) & \text{if } p < r < \phi \\ 1 - cr^{-\alpha} & r > \phi, \end{cases} \tag{A.2}$$

and take $c = 1/(\Gamma(1-\alpha)\cos(\pi\alpha/2))$ to approximate standard stable jumps (with scale 1, center 0, skewness 1 in the parametrization of *Samorodnitsky and Taqqu* [81]) when $0 < \alpha < 2, \alpha \neq 1$, see (7.19)–(7.21) in [57]. To ensure the same slope at the cutoff $r = \phi$, set $m = \alpha c \phi^{-1-\alpha}$. To ensure continuity at the cutoff $r = \phi$, equate

$$1 - c\phi^{-\alpha} = \alpha c \phi^{(-1-\alpha)}(\phi - p). \tag{A.3}$$

For $1 < \alpha < 2$, *Zhang et al.* [45] recommend setting $p = -2.5$, so the cutoff ϕ is solved by finding the root of the last equation. The chopped Pareto random variable ξ can be generated by picking a Uniform [0,1] random variable U, and setting:

$$\xi = \begin{cases} U/m + p & \text{if } U < 1 - c\phi^{-\alpha} \\ (c/(1-U))^{1/\alpha} & \text{otherwise.} \end{cases} \tag{A.4}$$

For $\alpha > 1$, this random variable has mean

$$\mu = \frac{\alpha c \phi^{-1-\alpha}}{2}(\phi^2 - p^2) - \frac{\alpha c}{1-\alpha}\phi^{1-\alpha}$$

so that $\xi - \mu$ gives a zero-mean random walk jump.

We find that this method is approximately two to three times faster than the modified CMS method. When the random walk simulations break a particle's motion into at least ten jumps, then the sum of jumps converge nicely to the exact generation method given by the modified CMS method (Fig. A.10 a,b,c). In the infinite mean case $\alpha < 1$, we find that setting $p = 0$ and adding an empirical constant ($\approx 0.3(1-\alpha)^{-1.2}$) speeds convergence considerably (Figure A.10 d).

References

[1] K. B. Oldham, J. Spanier, The Fractional Calculus, Academic Press, 1974.

[2] F. Mainardi, Fractional Calculus and Waves in Linear Viscoelasticity, World Scientic, Singapore, 2010.

[3] R. Metzler, J. Klafter, The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics, J. Phys. A 37 (31) (2004) R161.

[4] J. H. Cushman, T. R. Ginn, Nonlocal dispersion in media with continuously evolving scales of heterogeneity, Transp. Por. Media 13(1) (1993) 123–138.

[5] S. P. Neuman, Eulerian-Lagrangian theory of transport in space-time non-stationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximation, Water Resour. Res. 29 (3) (1993) 633–645.

[6] E. Morales-Casique, S. P. Neuman, A. Guadagnini, Non-local and localized analyses of non-reactive solute transport in bounded randomly heterogeneous porous media: Theoretical framework, Adv. Water Resour. 29 (8) (2006) 1238–1255.

[7] D. A. Benson, S. W. Wheatcraft, M. M. Meerschaert, Application of a fractional advection-dispersion equation, Water Resour. Res. 36(6) (2000) 1403–1412. doi:10.1029/2000WR900031.

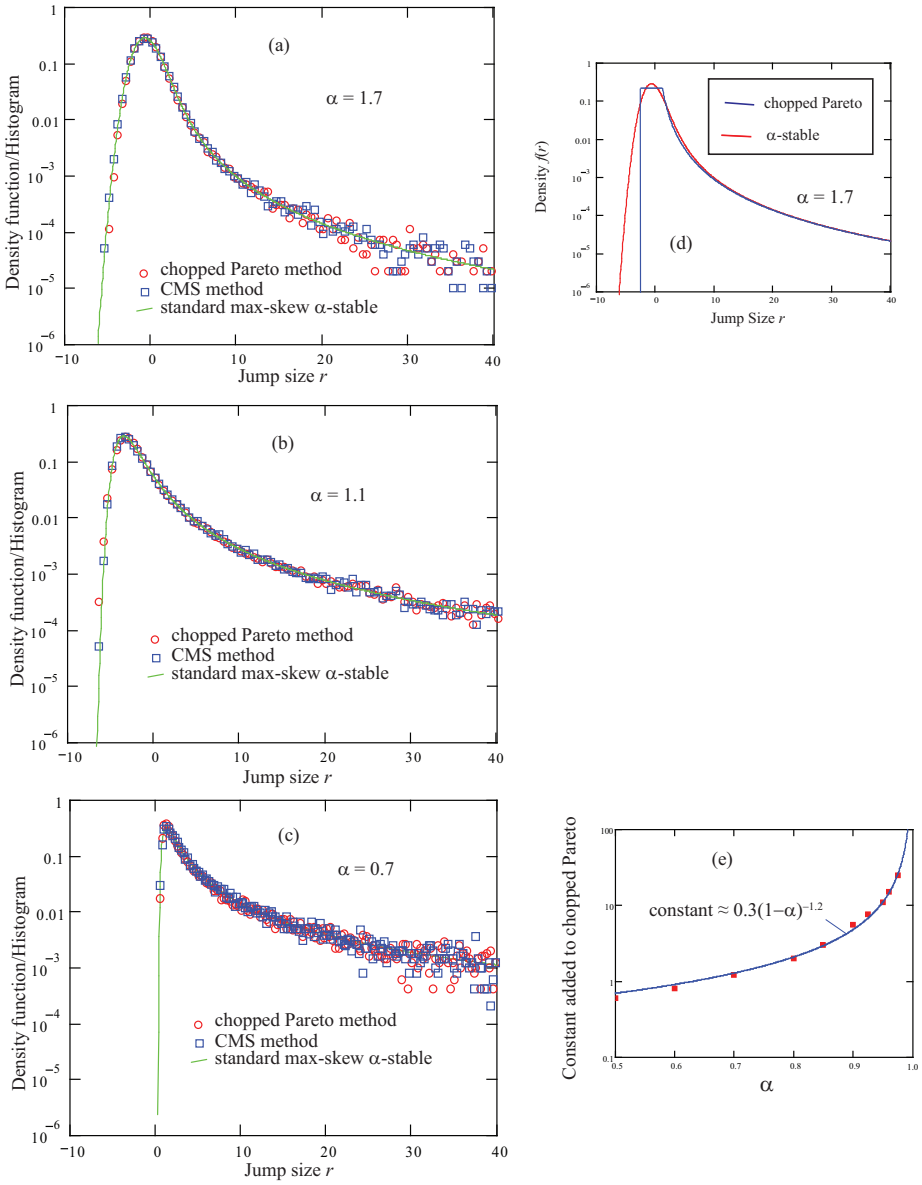


Figure A.10: (a,b,c) Histograms (symbols) versus standard, maximally-skewed α-stable density functions (curves). The histograms are for 20,000 variables generated by the modified CMS method [93] and rescaled sums of 10 jumps generated by the chopped Pareto method. (d) Plots of the standard, maximally-skewed α-stable and the chopped Pareto density functions for index $\alpha = 1.7$. (e) Plot of the empirical constant (symbols) added to chopped Pareto random variables for $\alpha < 1$ to speed convergence and a fitted function.

[8] D. Benson, R. Schumer, M. Meerschaert, S. Wheatcraft, Fractional dispersion, Lévy motion, and the MADE tracer tests, Transp. Porous Media 42 (1-2) (2001) 211–240. doi:10.1023/A:1006733002131.

[9] J. H. Cushman, T. R. Ginn, Fractional advection-dispersion equation: A classical mass balance with convolution-fickian flux, Water Resour. Res. 36 (12) (2000) 3763–3766.

[10] J. Carrera, X. Snchez-Vila, I. Benet, A. Medina, G. Galarza, J. Guimer, On matrix diffusion: formulations, solution methods and qualitative effects, Hydrogeology Journal 6 (1998) 178–190, 10.1007/s100400050143.
URL http://dx.doi.org/10.1007/s100400050143

[11] R. Haggerty, S. M. Gorelick, Multiple–rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, Water Resour. Res. 31 (10) (1995) 2383–2400.

[12] M. Dentz, B. Berkowitz, Transport behavior of a passive solute in continuous time random walks and multirate mass transfer, Water Resour. Res. 39 (5) (2003) 1111. doi:10.1029/2001WR001163.

[13] R. Schumer, D. A. Benson, M. M. Meerschaert, B. Baeumer, Fractal mobile/immobile solute transport, Water Resour. Res. 39 (2003) 1296. doi:10.1029/2003WR002141.

[14] R. Bhattacharya, V. Gupta, A theoretical explanation of solute dispersion in saturated porous media at the Darcy scale, Water Resour. Res. 19 (4) (1983) 938–944. doi:10.1029/WR019i004p00938.

[15] J. Klafter, I. M. Sokolov, First Steps in Random Walks: From Tools to Applications, Oxford University Press, 2011.

[16] G. Taylor, Dispersion of soluble matter in solvent flowing slowly through a tube, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 219 (1137) (1953) 186–203. doi:10.1098/rspa.1953.0139.

[17] A. Einstein, On the theory of the Brownian movement, Annalen der Physik 4 (1906) 371–381.

[18] J. Crank, Mathematics of Diffusion, Oxford University Press, 1956.

[19] M. Meerschaert, D. Benson, H. Scheffler, P. Becker-Kern, Governing equations and solutions of anomalous random walk limits, Phys. Rev. E 66 (6, Part 1). doi:10.1103/PhysRevE.66.060102.

[20] M. M. Meerschaert, C. Tadjeran, Finite difference approximations for fractional advection-dispersion flow equations, J. Comput. Appl. Math. 172 (1) (2004) 65 – 77. doi:10.1016/j.cam.2004.01.033.

[21] S. G. Samko, A. A. Klibas, O. I. Marichev, Fractional Integrals and Derivatives: Theory and Applications, Gordon and Breach Science Publishers, 1993.

[22] M. Meerschaert, D. Benson, B. Bäumer, Multidimensional advection and fractional dispersion, Phys. Rev. E 59 (5, Part A) (1999) 5026–5028.

[23] Y. Zhang, D. A. Benson, Lagrangian simulation of multidimensional anomalous transport at the MADE site, Geophys. Res. Lett. 35 (2008) L07403. doi:10.1029/2008GL033222.

[24] C. Tadjeran, M. M. Meerschaert, A second-order accurate numerical method for the two-dimensional fractional diffusion equation, J. Comput. Phys. 220 (2) (2007) 813–823.

[25] B. Baeumer, M. M. Meerschaert, J. Mortensen, Space-time fractional derivative operators, Proc. Amer. Math. Soc. 133 (2005) 2273–2282.

[26] W. Hudson, J. D. Mason, Operator-self-similar processes in a finite-dimensional space, Trans. Amer. Math. Soc. 273 (1982) 281–297.

[27] M. Meerschaert, D. Benson, B. Baeumer, Operator Lévy motion and multiscaling anomalous diffusion, Phys. Rev. E 63 (2, Part 1).

[28] R. Schumer, D. A. Benson, M. M. Meerschaert, B. Baeumer, Multiscaling fractional advection-dispersion equations and their solutions, Water Resour. Res. 39 (2003) 1022. doi:10.1029/2001WR001229.

[29] N. D. Monnig, D. A. Benson, M. M. Meerschaert, Ensemble solute transport in two–dimensional operator-scaling random fields, Water Resour. Res. 44 (2) (2008) W02434–. doi:10.1029/2007WR005998.

[30] M. Meerschaert, J. Mortensen, S. Wheatcraft, Fractional vector calculus for fractional advection-dispersion, PHYSICA A 367 (2006) 181–190. doi:10.1016/j.physa.2005.11.015.

[31] Y. Zhang, D. A. Benson, M. M. Meerschaert, E. M. LaBolle, Space-fractional advection-dispersion equations with variable parameters: Diverse formulas, numerical solutions, and application to the Macrodispersion Experiment site data, Water Resour. Res. 43 (5). doi:10.1029/2006WR004912.

[32] I. Pudlubny, Fractional Differential Equations. An Introduction to Fractional Derivatives, Fractional Differential Equations, Some Methods of Their Solution and Some of Their Applications, Academic Press, 1999.

[33] M. M. Meerschaert, H.-P. Scheffler, C. Tadjeran, Finite difference methods for two-dimensional fractional dispersion equation, J. Comput. Phys. 211 (1) (2006) 249 – 261. doi:10.1016/j.jcp.2005.05.017.

[34] K. Wang, H. Wang, A fast characteristic finite difference method for fractional advection-diffusion equations, Adv. Water Resour. 34 (7) (2011) 810–816. doi:10.1016/j.advwatres.2010.11.003.

[35] E. Hanert, A comparison of three Eulerian numerical methods for fractional-order transport models, Environ. Fluid Mech. 10 (2010) 7–20. doi:10.1007/s10652-009-9145-4.

[36] G. Fix, J. Roop, Least squares finite-element solution of a fractional order two-point boundary value problem, Comput. & Math. Appl. 48 (7-8) (2004) 1017–1033. doi:10.1016/j.camwa.2004.10.003.

[37] J. Roop, Computational aspects of FEM approximation of fractional advection dispersion equations on bounded domains in R-2, J. Comput. Appl. Math. 193 (1) (2006) 243–268. doi:10.1016/j.cam.2005.06.005.

[38] M. A. Celia, W. G. Gray, Numerical Methods for Differential Equations: Fundamental Concepts for Scientific and Engineering Applications, Prentice Hall, 1992.

[39] E. Hanert, On the numerical solution of spacetime fractional diffusion models, Computers & Fluids 46 (2011) 33–39.

[40] L. Konikow, J. Bredehoeft, Computer model of two-dimensional solute transport and dispersion in ground water, Tech. Rep. Techniques of Water-Resources Investigations, book 7, chap. C2., U.S. Geological Survey (1978).

[41] E. M. Labolle, G. E. Fogg, A. F. B. Tompson, Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods, Water Resour. Res. 32 (3) (1996) 583–593.

[42] E. M. LaBolle, J. Quastel, G. E. Fogg, Diffusion theory for transport in porous media: Transition-probability densities of diffusion processes corresponding to advection-dispersion equations, Water Resour. Res. 34(7) (1998) 1685–1693.

[43] E. M. LaBolle, J. Quastel, G. E. Fogg, , J. Gravner, Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients, Water Resour. Res. 36(3) (2000) 651–662.

[44] Y. Zhang, D. A. Benson, M. M. Meerschaert, H.-P. Scheffler, On using random walks to solve the space-fractional advection-dispersion equations, J. Stat. Phys. 123 (1) (2006) 89–110.

[45] Y. Zhang, D. A. Benson, M. M. Meerschaert, E. M. LaBolle, H.-P. Scheffler, Random walk approximation of fractional-order multiscaling anomalous diffusion, Phys. Rev. E 74 (2, Part 2). doi:10.1103/PhysRevE.74.026706.

[46] D. Pollock, Semianalytical computation of path lines for finite-difference models, GROUND WATER 26 (6) (1988) 743–750. doi:10.1111/j.1745-6584.1988.tb00425.x.

[47] D. M. Reeves, D. A. Benson, M. M. Meerschaert, H.-P. Scheffler, Transport of conservative solutes in simulated fracture networks: 2. ensemble solute transport and the correspondence to operator-stable limit distributions, Water Resour. Res. 44 (5) (2008) W05410. doi:10.1029/2008WR006858.

[48] M. Caputo, Linear models of dissipation whose q is almost frequescy independentm part ii., Geophys. J. Royal Astr. Soc. 13 (1967) 529–539.

[49] E. W. Montroll, G. H. Weiss, Random walks on lattices. ii, J. Math. Phys. 6 (2) (1965) 167–181.

[50] S. Bochner, Diffusion equations and stochastic processes, Proc. Natl. Acad. Sci. U.S.A. 85 (1949) 369–370.

[51] W. Feller, An Introduction to Probability Theory and Its Applications, vol. II WIley Series in Probability Theory and its Applications, Wiley, 1971.

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

[53] H. Scher, M. Lax, Stochastic transport in a disordered solid. i. theory, Phys. Rev. B 7 (10) (1973) 4491–4502. doi:10.1103/PhysRevB.7.4491.

[54] M. F. Shlesinger, J. Klafter, Y. M. Wong, Random walks with infinite spatial and temporal moments, J. of Stat. Phys. 27 (3) (1982-03-01) 499–512. doi:10.1007/BF01011089.

[55] M. Dentz, H. Scher, D. Holder, B. Berkowitz, Transport behavior of coupled continuous-time random walks, Phys. Rev. E 78 (4) (2008) 041110. doi:10.1103/PhysRevE.78.041110.

[56] E. Major, D. A. Benson, J. Revielle, H. Ibrahim, A. Dean, R. M. Maxwell, E. Poeter, Comparison of fickian and temporally non-local transport theories over many scales in an exhaustively sampled sandstone slab, submitted.

[57] M. M. Meerschaert, H.-P. Scheffler, Limit Distributions for Sums of Independent Random Vectors, Wiley Interscience, 2001.

[58] Y. Zhang, D. A. Benson, D. M. Reeves, Time and space nonlocalities underlying fractional-derivative models: Distinction and literature review of field applications, Adv. Water Resour. 32 (4) (2009) 561–581. doi:10.1016/j.advwatres.2009.01.008.

[59] M. Dogan, R. L. Van Dam, G. C. Bohling, J. Butler, James J., D. W. Hyndman, Hydrostratigraphic analysis of the MADE site with full-resolution GPR and direct-push hydraulic profiling, Geophys. Res. Lett. 38 (6) (2011) L06405.

[60] C. Harvey, S. M. Gorelick, Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the macrodispersion experiment (MADE) site?, Water Resour. Res. 36 (3) (2000) 637–650. doi:10.1029/1999WR900247.

[61] C. J. Harman, D. M. Reeves, B. Baeumer, M. Sivapalan, A subordinated kinematic wave equation for heavy-tailed flow responses from heterogeneous hillslopes, J. Geophys. Res. 115. doi:10.1029/2009JF001273.

[62] D. N. Bradley, G. E. Tucker, D. A. Benson, Fractional dispersion in a sand bed river, J. Geophys. Res. 115 (2010) F00A09–.

[63] V. Ganti, M. M. Meerschaert, E. Foufoula-Georgiou, E. Viparelli, G. Parker, Normal and anomalous diffusion of gravel tracer particles in rivers, J. Geophys. Res. 115 (2010) F00A12–.

[64] V. Ganti, K. M. Straub, E. Foufoula-Georgiou, C. Paola, Space-time dynamics of depositional systems: Experimental evidence and theoretical modeling of heavy-tailed statistics, J. Geophys. Res. 116 (F2) (2011) F02011–.

[65] E. Foufoula-Georgiou, V. Ganti, W. E. Dietrich, A nonlocal theory of sediment transport on hillslopes, J. Geophys. Res. 115 (2010) F00A16–.

[66] R. Schumer, M. M. Meerschaert, B. Baeumer, Fractional advection-dispersion equations for modeling transport at the earth surface, J. Geophys. Res. 114 (2009) F00A07–. doi:10.1029/2008JF001246.

[67] B. B. Mandelbrot, J. R. Wallis, Computer experiments with fractional gaussian noises. part 3, mathematical appendix, Water Resour. Res. 5 (1969) 260–267.

[68] D. A. Benson, M. M. Meerschaert, B. Baeumer, H.-P. Scheffler, Aquifer operator scaling and the effect on solute mixing and dispersion, Water Resour. Res. 42 (1) (2006) W01415–. doi:10.1029/2004WR003755.

[69] H. Rajaram, L. Gelhar, Plume-scale dependent dispersion in aquifers with a wide range of scales of heterogeneity, Water Resour. Res. 31 (10) (1995) 2469–2482.

[70] H. Zhan, S. W. Wheatcraft, Macrodispersivity tensor for nonreactive solute transport in isotropic and anisotropic fractal porous media: Analytical solutions, Water Resour. Res. 32 (1996) 3461–3474.

[71] V. Di Federico, S. P. Neuman, D. M. Tartakovsky, Anisotropy, lacunarity, and upscaled conductivity and its autocovariance in multiscale random fields with truncated power variograms, Water Resour. Res. 35 (1999) 2891–2908.

[72] F. J. Molz, H. H. Liu, J. Szulga, Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions, Water Resour. Res. 33 (10) (1997) 2273–2286.
URL http://dx.doi.org/10.1029/97WR01982

[73] T. A. Hewitt, Fractal distribution of reservoir heterogeneity and their influence on fluid transport, in: SPE Pap. 15386, Soc. Pet. Eng. 61st Annual Tech. Conf., 1986, new Orleans, LA.

[74] F. Molz, G. Boman, A fractal-based stochastic interpolation scheme in subsurface hydrology, Water Resour. Res. 29 (1993) 3769–3774.

[75] H. Liu, F. Molz, Multifractal analyses of hydraulic conductivity distributions, Water Resour. Res. 33 (1997) 2483–2488.

[76] A. Deshpande, P. B. Flemings, J. Huang, Quantifying lateral heterogeneities in fluvio-deltaic sediments using three-dimensional reflection seismic data: Offshore Gulf of Mexico, J. Geophys. Res. 102 (B7) (1997) 15385–15401. doi:10.1029/97JB01143.

[77] L. Tennekoon, M. C. Boufadel, D. Lavallee, J. Weaver, Multifractal anisotropic scaling of the hydraulic conductivity, Water Resour. Res. 39 (7) (2003) 1193–.
URL http://dx.doi.org/10.1029/2002WR001645

[78] J. W. Castle, F. J. Molz, S. Lu, C. L. Dinwiddie, Sedimentology and fractal-based analysis of permeability data, John Henry member, Straight Cliffs formation,(upper Cretaceous), Utah, U.S.A., J. Sed. Res. 74 (2004) 270–284.

[79] H. Biermé, M. M. Meerschaert, H.-P. Scheffler, Operator scaling stable random fields, Stoch. Proc. Appl. 117 (3) (2007) 312–332. doi:DOI: 10.1016/j.spa.2006.07.004.

[80] A. M. Yaglom, An Introduction to the Theory of Stationary Random Functions, Dover Publications, 1962.

[81] G. Samorodnitsky, M. S. Taqqu, Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance, Stochastic Modeling, Chapman Hall, London, 1994.

[82] J. Revielle, Incorporating Hydraulic Conductivity Measurements Into Fractal Aquifer Realizations and its Effect on Solute Transport, Master's thesis, Colorado School of Mines (2009).

[83] A. G. Journel, C. J. Huijbregts, Mining Geostatistics, Academic Press, San Diego, CA, 1978.

[84] C. V. Deutsch, A. G. Journel, GSLIB Geostatistical software library and user's guide, Oxford University Press, New York, 1993.

[85] S. Painter, Stochastic interpolation of aquifer properties using fractional Lévy motion, Water Resour. Res. 32 (5) (1996) 1323–1332.

[86] S. Painter, Numerical method for conditional simulation of Lévy random fields, Math. Geol. 30 (2) (1998) 163–179.

[87] S. Painter, Flexible scaling model for use in random field simulation of hydraulic conductivity, Water Resour. Res. 37 (5) (2001) 1155–1163.

[88] G. L. Isaacs, An iteration formula for fractional differences, Proc. London Math. Soc. s3-13 (1) (1963) 430–460.

[89] J. R. M. Hosking, Fractional differencing, Biometrika 68 (1981) 165–176.

[90] V. C. Tidwell, J. L. Wilson, Heterogeneity, permeability patterns, and permeability upscaling: Physical characterization of a block of Massillon sandstone exhibiting nested scales of hetergeneity, SPE Reservoir Eval. and Eng. 3 (4) (2000) 283–291.

[91] K. A. Klise, V. C. Tidwell, S. A. McKenna, Comparison of laboratory-scale solute transport visualization experiments with numerical simulation using cross-bedded sandstone, Adv. Water Resour. 31 (12) (2008) 1731–1741. doi:10.1016/j.advwatres.2008.08.013.

[92] D. Bolster, P. de Anna, D. A. Benson, A. M. Tartakovsky, Incomplete mixing and reactions with fractional dispersion, Adv. Water Resour. submitted.

[93] R. Weron, On the Chambers-Mallows-Stuck method for simulating skewed stable random variables, Stat. Prob. Lett. 28 (2) (1996) 165171.