Article

New Insights into the Fractional Order Diffusion Equation Using Entropy and Kurtosis

Carson Ingo [1,]*, **Richard L. Magin** [2] **and Todd B. Parrish** [3]

[1] C.J. Gorter Center for High Field MRI, Department of Radiology, Leiden University Medical Center, Albinusdreef 2, 2333ZA Leiden, The Netherlands

[2] Department of Bioengineering, University of Illinois at Chicago, 851 S. Morgan St, Chicago, 60607, IL, USA; E-Mail: rmagin@uic.edu

[3] Department of Radiology, Northwestern University, 737 Michigan Ave 16th Floor, Chicago, 60611 IL, USA; E-Mail: toddp@northwestern.edu

* Author to whom correspondence should be addressed; E-Mail: c.j.ingo@lumc.nl; Tel./Fax: +31-71-526-5473 / +31-71-524-8256.

External Editor: J. A. Tenreiro Machado

Abstract: Fractional order derivative operators offer a concise description to model multi-scale, heterogeneous and non-local systems. Specifically, in magnetic resonance imaging, there has been recent work to apply fractional order derivatives to model the non-Gaussian diffusion signal, which is ubiquitous in the movement of water protons within biological tissue. To provide a new perspective for establishing the utility of fractional order models, we apply entropy for the case of anomalous diffusion governed by a fractional order diffusion equation generalized in space and in time. This fractional order representation, in the form of the Mittag–Leffler function, gives an entropy minimum for the integer case of Gaussian diffusion and greater values of spectral entropy for non-integer values of the space and time derivatives. Furthermore, we consider kurtosis, defined as the normalized fourth moment, as another probabilistic description of the fractional time derivative. Finally, we demonstrate the implementation of anomalous diffusion, entropy and kurtosis measurements in diffusion weighted magnetic resonance imaging in the brain of a chronic ischemic stroke patient.

1. Introduction

The fundamental concept in continuous time random walk (CTRW) theory is to extend the diffusion equation, such that the fractional order partial derivatives can be used as the governing mathematical operators to describe the diffusion propagator, $P(x,t)$:

$$\frac{\partial^\alpha P(x,t)}{\partial t^\alpha} = D_{\alpha,\beta}\frac{\partial^\beta P(x,t)}{\partial |x|^\beta}, \tag{1}$$

that is, where $\partial^\alpha/\partial t^\alpha$ represents the Caputo fractional derivative in time for $0 < \alpha \leq 1$, $\partial^\beta/\partial|x|^\beta$ represents the order of the Riesz fractional derivative in space for $1 < \beta \leq 2$ and $D_{\alpha,\beta}$ is the generalized diffusion constant (distance$^\beta$/time$^\alpha$) [1–3].

The justification for making use of the fractional derivative operators is to provide a mathematical means to interpolate from homogeneous and relatively simple systems that exhibit local, Gaussian behavior to heterogeneous and relatively complex systems that exhibit non-local, power-law behavior [2–8]. In the CTRW context, the order of the fractional operators, α and β, provides a description of a random walker's likelihood to have broader distributions of waiting times and jump lengths, respectively, in comparison to classical Brownian motion. Because anomalous diffusion behavior implies a different probability model, one which postulates that the random walker sometimes halts between steps and sometimes takes larger or smaller steps, the likelihood, or weighting, placed on jumps and waiting times, which asymptotically follow simple inverse power laws: $t^{-(1+\alpha)}$ and $|x|^{-(1+\beta)}$ [5]. As a result, bulk non-Gaussian motions emerge, and the mean squared displacement (MSD) no longer increases with a linear dependence on time. In the most basic form, the one-dimensional MSD is expressed by a composite power law as:

$$\langle x^2(t)\rangle \sim t^{2\alpha/\beta}. \tag{2}$$

When $\alpha = 1$ and $\beta = 2$, Equation (2) grows linearly with the diffusion time, and Equation (1) collapses to the integer-order partial differential equation to describe Gaussian diffusion. When the ratio $2\alpha/\beta < 1$, the MSD dynamics are, overall, subdiffusive, and when $2\alpha/\beta > 1$, the dynamics diverge to be described to have superdiffusive behavior (e.g., pseudo-transport) [3,5,7]. The MSD trajectories of Gaussian, sub- and super-diffusion are illustrated in Figure 1, where in the case of subdiffusion, the variance grows slower than the Gaussian case, where in the case of superdiffusion, the variance increases faster than the linear growth of the Gaussian. In the context of the MSD, when the power law properties of the random walk distributions are connected with the order of the fractional derivatives, α and β encode either a fading memory in time or extended neighbor interactions in space as a consequence of the mathematical operators in the fractional partial differential equation. As a new perspective to view the utility of fractional order calculus, we consider power law diffusion dynamics from the perspective of information theory and make use of entropy as a measurement tool for the characteristic function (CF) of the diffusion propagator. Since the fractional derivative is a heuristic tool that includes, in its very definition, a distribution of time and space constants in order to represent multi-scale systems, we find that fractional order models convey more information, via entropy, about the underlying structure and dynamics of complex systems. Furthermore, we compute the kurtosis of the time-fractional solution to the diffusion equation as an alternative measure to demonstrate how the fractional operator encodes

heterogeneous, non-Gaussian dynamics. Finally, we demonstrate the utility of the fractional order derivative, entropy and kurtosis to provide new contrast to characterize biological tissue microstructure beyond the classical model for diffusion MRI measurements in the brain of a chronic ischemic stroke patient.

Figure 1. Sketches of the mean squared displacement for the cases of Gaussian diffusion ($2\alpha/\beta = 1$), subdiffusion ($2\alpha/\beta < 1$) and superdiffusion ($2\alpha/\beta > 1$).
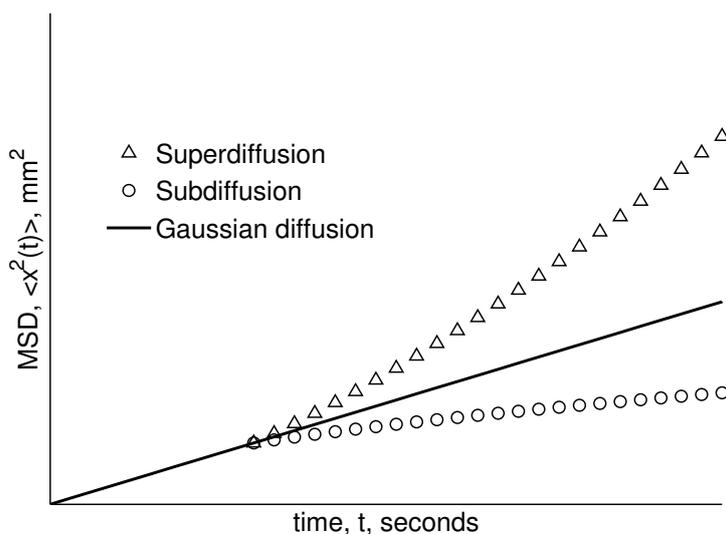
2. Experimental Section

2.1. Theory

In the general case, the order of the fractional partial derivatives, α and β, can span a range of values that encompass processes characterized as both subdiffusion and as superdiffusion in the diffusion phase diagram shown in Figure 2. This distinction is made by comparison with Brownian motion and the second moment of the Gaussian distribution, which increases as a linear function of time ($\alpha = 1$ and $\beta = 2$). This diagram is formally derived in CTRW theory, where $P(x,t)$ expresses the probability density of finding the molecule at position x at time t, assuming that at $t = 0$, all of the material is concentrated at the origin, where $P(x,0) = \delta(x)$ as a Dirac delta function[1,5].

Conveniently, utilizing Fourier and Laplace transforms, Equation (1) can be succinctly expressed in the form, $P(x,t) \rightarrow p(k,s)$:

$$p(k,s) = \frac{1}{s + D_{\alpha,\beta} s^{1-\alpha} |k|^{\beta}}. \tag{3}$$

Then, by applying the inverse Laplace transformation to Equation (3), we obtain a simple expression for the relaxation for each wavenumber, k:

$$p(k,t) = E_{\alpha}\big(- D_{\alpha,\beta} |k|^{\beta} t^{\alpha} \big). \tag{4}$$

Figure 2. Anomalous diffusion phase diagram with respect to the order of the fractional derivative in space, β, and the order of the fractional derivative in time, α.
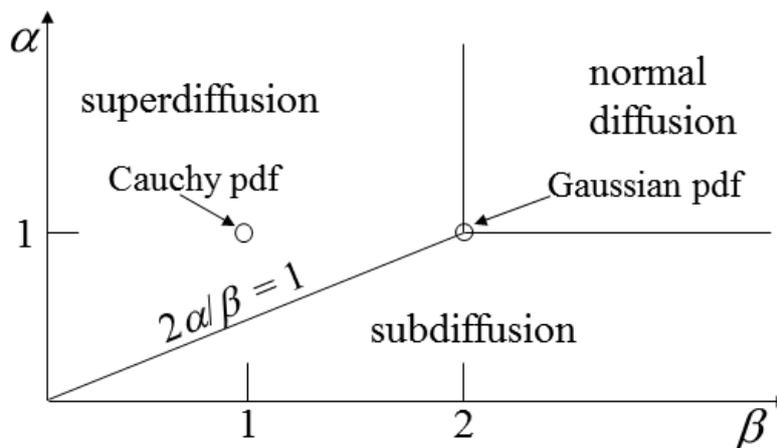
The symbol E_α represents the single-parameter Mittag–Leffler function (MLF), which describes a generalized power law decay for the range, $0 < \alpha < 1$, and a damped, but oscillating function for $1 < \alpha < 2$ [9,10]. Specific functions encapsulated by the MLF are a monoexponential decay when $\alpha = 1$ and the cosine function when $\alpha = 2$. Thus, for a fixed wavenumber and diffusion coefficient, Equation (4) corresponds to the time decay (with or without oscillations) of one spectral component. As the order of the fractional time derivative α decreases from one, there is a shift away from the single exponential decay toward an apparent multi-exponential relaxation in time. In the case of $\alpha = 1$ and $\beta = 2$, we have the characteristic form for a Gaussian distribution,

$$p(k,t) = exp\big(-D_{1,2}|k|^2 t\big). \tag{5}$$

Clearly, as either the time or the diffusion coefficient increase, the spread of wavenumbers decreases, which, through the scaling property of the Fourier transform, gives a correspondingly wider distribution in space. Conversely, as the time or the diffusion coefficient decrease, the spread of the wavenumbers increases and the distribution of the diffusion process in space is more localized. This phenomena is described in the space-time domain for normal diffusion by the probability diffusion function (pdf) of the Gaussian:

$$P(x,t) = \frac{1}{\sqrt{4\pi D_{1,2}t}} exp\Big(-\frac{x^2}{4D_{1,2}t}\Big), \tag{6}$$

where the mean squared displacement in space is simply the second moment (or variance),

$$\langle x^2(t) \rangle = 2D_{1,2}t. \tag{7}$$

In comparison to the Gaussian case, there is the condition of time-fractional subdiffusion, $0 < \alpha < 1$ and $\beta = 2$. Equation (3) becomes,

$$p(k,s) = \frac{1}{s + D_{\alpha,2}s^{1-\alpha}|k|^2}, \tag{8}$$

and Equation (4) becomes,

$$p(k,t) = E_\alpha\big(-D_{\alpha,2}|k|^2 t^\alpha\big). \tag{9}$$

By first utilizing the simple form of the Laplace–Fourier solution to time-fractional subdiffusion in Equation (8), the MSD can be computed by taking the second derivative of Equation (8) with respect to k as the limit of $k \to 0$ and then performing a Laplace inversion,

$$\langle x^2(t) \rangle = \mathcal{L}^{-1} \lim_{k \to 0} \{ -\frac{d^2 p(k,s)}{dk^2} \}, \tag{10}$$

which gives,

$$\langle x^2(t) \rangle = \mathcal{L}^{-1} \{ 2D_{\alpha,2} s^{-(\alpha+1)} \}, \tag{11}$$

and utilizing the common Laplace and time domain transform pair,

$$\frac{t^\alpha}{\Gamma(\alpha+1)} = \mathcal{L}^{-1} \{ s^{-(\alpha+1)} \}, \tag{12}$$

yields the form of the MSD,

$$\langle x^2(t) \rangle = \frac{2D_{\alpha,2}}{\Gamma(\alpha+1)} t^\alpha, \tag{13}$$

as reported in [5].

For the antithetical case of superdiffusion, $\alpha = 1$ and $\beta < 2$, Equation (3) becomes,

$$p(k,s) = \frac{1}{s + D_{1,\beta}|k|^\beta}, \tag{14}$$

and applying the inverse Laplace transform yields a stretched exponential about k,

$$p(k,t) = exp\big[- D_{1,\beta}|k|^\beta t \big], \tag{15}$$

which is the CF for a centered, symmetric Levy distribution [8]. Unfortunately, Equation (15) cannot be analytically expanded at $k = 0$ when $\beta < 2$, and so, the second moment is undefined for the diffusion propagator, $P(x,t)$. In contrast, the CF for subdiffusion in Equation (8) has a closed form representation of the MSD in Equation (13), so it is possible to compute the kurtosis for the diffusion propagator as a function of the order of the fractional time derivative as follows.

2.2. Kurtosis

Qualitatively, kurtosis is a measure to nonspecifically describe the peakedness and/or heavy tail shape in a probability distribution via the standardized fourth moment [11]. For example, a Gaussian probability distribution function (pdf) has a kurtosis value of three, whereas the hyperbolic secant pdf has a kurtosis value of five, as its shape is both more peaked and heavier-tailed than the Gaussian shape. A convenient index called excess kurtosis is defined as the difference between the estimated kurtosis value for a given distribution with the value of the Gaussian pdf (*i.e.*, three), such that the excess kurtosis of a Gaussian pdf is zero. In the context of the mathematical definition of excess kurtosis, diffusion processes with positive excess kurtosis imply that the diffusion propagator has a standardized fourth moment, which is broader than the Gaussian case,

$$K \equiv \frac{\langle x^4 \rangle}{\langle x^2 \rangle^2} - 3. \tag{16}$$

Returning back to the formalism of operating in the Laplace domain, the fourth moment of Equation (8) can be expressed as,

$$< x^4(t) >= \mathcal{L}^{-1}\{24(D_{\alpha,2})^2 s^{-(2\alpha+1)}\}, \tag{17}$$

and inverting into the time domain using the transform in Equation (12) gives,

$$< x^4(t) >= \frac{24(D_{\alpha,2})^2}{\Gamma(2\alpha+1)}t^{2\alpha}. \tag{18}$$

Inserting Equation (13) and Equation (18) into Equation (16), the excess kurtosis of the MLF for time-fractional subdiffusion is,

$$K_{MLF} = 6\frac{\Gamma^2(\alpha+1)}{\Gamma(2\alpha+1)} - 3. \tag{19}$$

When $\alpha = 1$, the MLF becomes the monoexponential CF and $K_{MLF} = 0$. For $0 < \alpha < 1$, $K_{MLF} > 0$ with the maximum excess kurtosis value limited to $\max(K_{MLF}) = 3$ when $\alpha \to 0$. Equation (19) is plotted in Figure 3, which shows a nearly inverse linear relationship between K_{MLF} and α.

As mentioned above, the CF for Equation (15) cannot be expanded as a Maclaurin series ($k = 0$). However, as Equation (15) exists in closed form as a stretched exponential function, it is possible to consider entropy in the Fourier domain. Furthermore, in general, since $p(k, t)$ is available in closed form as the MLF in Equation (4) for all diffusion regimes, we are able to compute the entropy for the cases of Gaussian, sub- and super-diffusion from the perspective of information theory.

Figure 3. Plot of Equation (19) for the kurtosis, K_{MLF}, computed in the Mittag–Leffler representation of subdiffusion *versus* the time-fractional derivative, α.
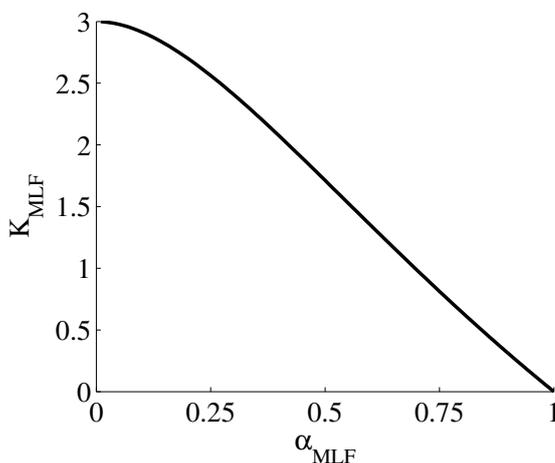
2.3. Entropy

The most fundamental formula to compute entropy, or the amount of uncertainty in a discrete pdf, $P(x)$ can be measured with,

$$H_x \equiv -\sum_{i=1}^{N} P(x_i)ln\big(P(x_i)\big), \tag{20}$$

where $P(x_i)$ are the individual samples in the the discrete pdf and H_x is the Shannon information entropy [12]. With the consideration of information formulated in the context of statistical uncertainty,

we have a tool to compare systems governed by differing stochastic processes. For example, when comparing two distributions, both normalized with the same full-width, half maximum values, the heavier-tailed distribution can be shown to have greater information entropy, pertaining to greater statistical uncertainty. Another approach to measure the uncertainty in a system is to analyze the characteristic function in terms of the Fourier transform in space, $P(x) \rightarrow p(k)$, with spectral entropy,

$$H_k \equiv -\sum_{i=1}^{N} \frac{\hat{p}(k_i) ln(\hat{p}(k_i))}{ln(N)}, \tag{21}$$

where $\hat{p}(k_i) = p(k_i)p^*(k_i)$ reflects the individual wavenumber's contribution to a normalized power spectrum of the Fourier transform, p_k, and the term, $ln(N)$ (*i.e.*, discrete uniform distribution of N samples), is a normalization factor applied so that the spectral entropy, H_k, is between zero and one [13,14]. Furthermore, as Equation (21) is generally defined to measure the uncertainty of a characteristic function, we can adapt this formalism to compare the CFs for anomalous diffusion,

$$H[\hat{p}(k,t)] \equiv -\sum_{i=1}^{N} \frac{\hat{p}(k,t)_i ln(\hat{p}(k,t)_i)}{ln(N)}. \tag{22}$$

By inserting the characteristic function in Equation (4) into Equation (22), the entropy in arbitrary diffusion processes can be measured. Furthermore, in the context of MRI, in which the reconstructed signal is considered as the Fourier transform in space, Equation (22) can be rewritten as a model for the diffusion MRI weighting factor $b \equiv q^2 \bar{\Delta}$, where by convention, q is diffusion gradient strength sensitization (e.g., units of mm^{-1}) and $\bar{\Delta}$ is the effective diffusion time (e.g., units of ms) to be rewritten as,

$$H[\hat{p}(q,\bar{\Delta})] \equiv -\sum_{i=1}^{N} \frac{\hat{p}(q,\bar{\Delta})_i ln(\hat{p}(q,\bar{\Delta})_i)}{ln(N)}. \tag{23}$$

to demonstrate the utility of Equations (9), (19) and (23) as models to characterize biological tissue in non-Gaussian diffusion MRI measurements. Furthermore, we examine, via spectral entropy (Equation (22)) the properties of the MLF to encapsulate the special cases of time- and space-fractional diffusion processes, which, in general, are shown to have more information content than the case of a Gaussian diffusion process.

2.4. Methods for Diffusion MRI Experiments

To demonstrate the utility of the fractional order of the time derivative, entropy and kurtosis in the diffusion of water within biological tissue, one patient with chronic ischemic stroke was scanned on a 3 Tesla Siemens Trio MRI scanner (Siemens Medical Solutions, Erlangen, Germany). Diffusion-weighted spin echo-echo planar imaging (SE-EPI) experiments were performed with the following pulse sequence parameters: echo time $TE = 102$ ms, repetition time $TR = 6$ s, diffusion time of $\Delta = 41.2$ ms, pulse duration of $\delta = 40.6$ ms $\bar{\Delta} = \Delta - \delta/3 = 27.7$ ms, diffusion weightings of $b = 0, 500, 1000, 3000, 4000$ s/mm^2, 3 orthogonal diffusion weighted directions, number of averages $NA = 6$, in-plane voxel resolution $= 2 \times 2$ mm, voxel thickness $= 4$ mm, 20 axial slices, scan time ~ 6 minutes. The raw diffusion weighted data were Rician noise corrected by estimating the variance (σ^2) in the signal

intensity of the ventricle at each b-value, such that $S_{rn} = \sqrt{S^2 - \sigma^2}$. The Rician noise-corrected diffusion weighted images were skull-stripped utilizing the Brain Extraction Tool [15]. All skull-stripped and Rician noise-corrected diffusion weighted images were co-registered to the $b = 0$ image space using statistical parametric mapping software (SPM8). Using the Levenberg–Marquardt minimization algorithm in MATLAB (Natick, MA, USA), the average of the three diffusion weighted direction data were fit on a voxel-wise basis to Equation (9) with the MLF algorithm in [16,17]. Following estimations of D and α, the excess kurtosis, K_{MLF}, was computed using the conversion provided in Equation (19). Following estimations of D and α, the CF in Equation (9) for $p(k, t)$ was constructed using $N = 100$ increments arrayed over variable b-values between 0 and 10 000 s/mm². Then, the entropy (defined in Equation (22)) in the diffusion process, as modeled by the MLF, was computed as H_{MLF}. The isotropic parameter maps of D, α, K_{MLF} and H_{MLF} for the same axial slice through the stroke patient's brain are shown in Figure 4.

2.5. Methods to Evaluate Entropy in the Mittag–Leffler Function

We computed the entropy for the CTRW model of diffusion by insertion of Equation (4) into Equation (21) by considering cases of $0 < \alpha \leq 2$ and $0 < \beta \leq 4$. Using MATLAB, the MLF was computed using the algorithm in [16,17]. For simplicity, the units are arbitrary with a fixed diffusion coefficient of $D = 1$; the diffusion times are evaluated at four cases $t = 0.5 - 2$, and the wavenumbers are incremented in an array $k_i = 0 - 5$ for $N = 500$ points. The overall results for entropy in the MLF are presented in Figure 5 as a three-dimensional entropy surface drawn above a plane defined by the positive values of α and β. The floor of the plot is essentially the phase diagram shown in Figure 2. Figures 6 and 7 plot the entropy in the diffusion process as a function of the order of the fractional derivative in space and time, respectively. Figures 8 and 9 plot the individual contributions to the total entropy of the diffusion process as a function of the wavenumber, k_i, for selected cases of α and β.

3. Results and Discussion

3.1. Diffusion MRI Experiments

The ischemic tissue (IT), in the right hemisphere of the patient's brain (left side of the image), has a diffusion coefficient, D, value ($\sim 3 \times 10^{-3}$ mm²/s), which is similar to the typical value found for the cerebral spinal fluid (CSF) of the ventricles. As can be seen in the contralateral hemisphere, prior to the onset of the stroke, the brain slice would have appeared symmetrical with white matter (WM) and gray matter (GM) voxels. However, as these data were acquired ~ 2 years following onset, the IT microstructure has degenerated (necrosis), such that the bulk diffusion coefficient has increased to an unhindered value. Furthermore, the diffusion in the IT is close to Gaussian as $\alpha \sim 1$, indicating a monoexponential behavior, which is also the case for the CSF. The trace values for D in the healthy WM and GM are $\sim 1/3$ of the values in the IT and CSF, with the WM possessing an overall slower diffusion than measured in the GM.

Figure 4. Trace parameter maps of α, D, K_{MLF} and H_{MLF} for an axial slice through a brain of a chronic stroke patient.
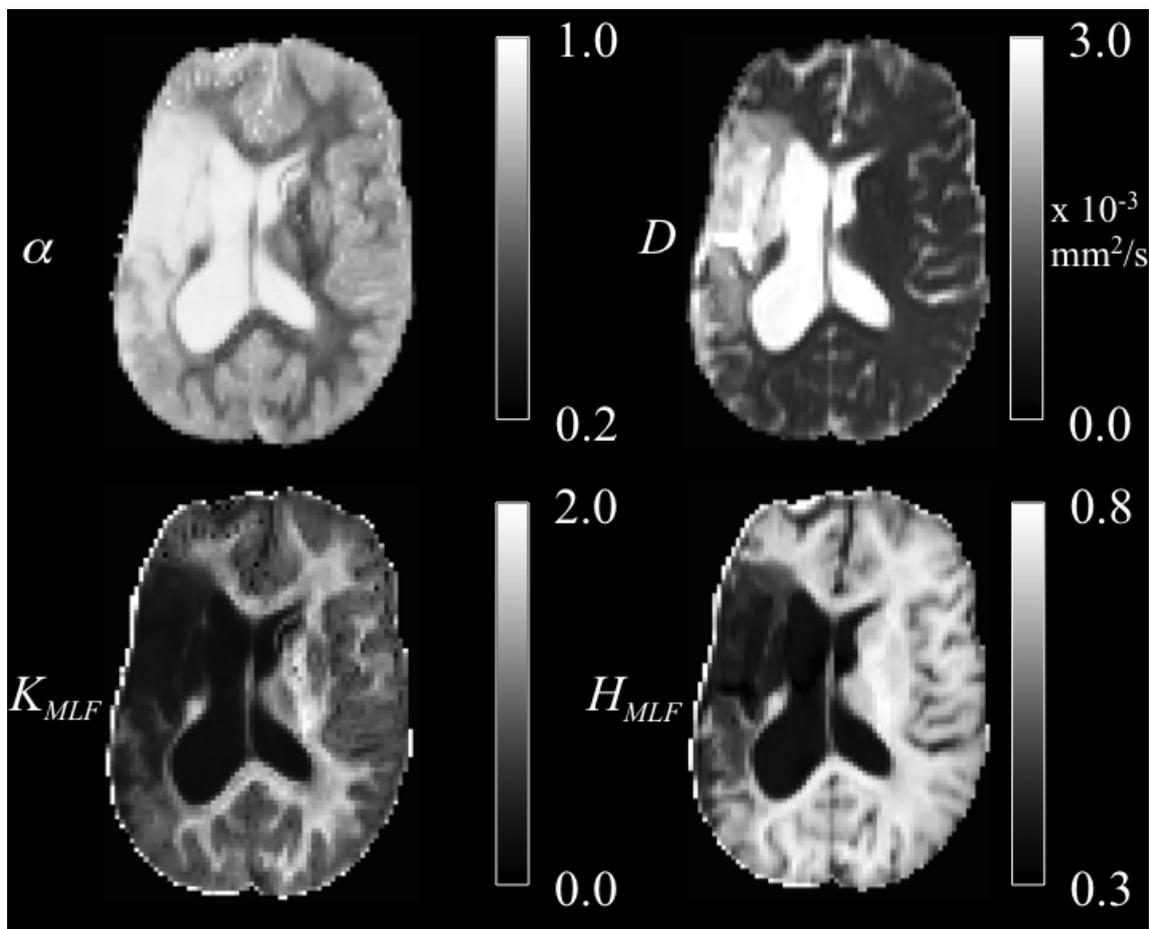
As the scale in the D map in Figure 4 spans 3×10^{-3} mm^2/s, the contrast between WM and GM is difficult to discern; however, in the α map, the WM/GM contrast is clearly visible with the WM demonstrating more subdiffusive behavior compared to the GM. The K_{MLF} map also has clearly visible GM/WM contrast and appears as a negative image in the α map, due to the nearly inverse relationship between K_{MLF} and α in Equation (19). The entropy, H_{MLF}, map provides a stable image in which there is visible and smooth GM/WM contrast, with the IT and CSF exhibiting low values of entropy due to the unhindered Gaussian diffusion dynamics. Interestingly, voxels exhibiting low values of α, representing highly subdiffusive dynamics, also have relatively high entropy estimations (particularly in the WM) to indicate a diffusion propagator with a heavy-tailed pdf. In correspondence with entropy, there is high kurtosis estimated for the diffusion propagator pdf in regions of low values of α. However, kurtosis and entropy are not interchangeable measures of the diffusion propagator pdf, evident not only in the images shown in Figure 4, but also mathematically distinct in the forms of Equations (19) and (22). Specifically, in consideration of the CF, α and K_{MLF} are measures of the deviation from a monoexponential form with a rate of the diffusion coefficient, D, whereas entropy considers the entire CF, which includes both the monoexponential component, D, and the non-Gaussian component, α, of the diffusion profile. As kurtosis is defined as the normalized fourth moment in Equation (16), it can be readily seen that the variance, or D, is canceled out when dividing Equation (18) by Equation (13).

Utilizing the moment expansion of the time-fractional form of the MLF in Equations (13) and (18) provides a direct link for kurtosis to subdiffusion through the Γ function and α. As Equation (9) is an analytic and monotonically decreasing function, the MLF provides the opportunity to more completely sample wavenumber and diffusion time distributions, to more accurately estimate the true kurtosis of the diffusion propagator, which is an advantage compared to other approaches that estimate kurtosis through Taylor expansion of the exponential function [18].

3.2. Measuring the Mittag–Leffler Function with Spectral Entropy

In Figure 5, the overall shape of the surface resembles plateaus of high entropy at non-integer values of α and β with a canyon of low entropy (near $\alpha = 1$) that flows in the direction of increasing β. The Gaussian case ($\alpha = 1$ and $\beta = 2$) is at a minimum for permutations of $\alpha \leq 1$ and $\beta \leq 2$. Of course, the specifics of the cross-sectional shape and entropy values of the surface is subject to change, with values chosen for the values of D and t as detailed in the analyses provided by Figures 6–9.

Figure 5. Spectral entropy surface plot for the Mittag–Leffler function (MLF) in Equation (4) with respect to the order of the fractional space derivative, β, and the order of the fractional time derivative, α ($D_{\alpha,\beta} = 1, t = 1$). The floor of the plot corresponds to the anomalous diffusion phase diagram.
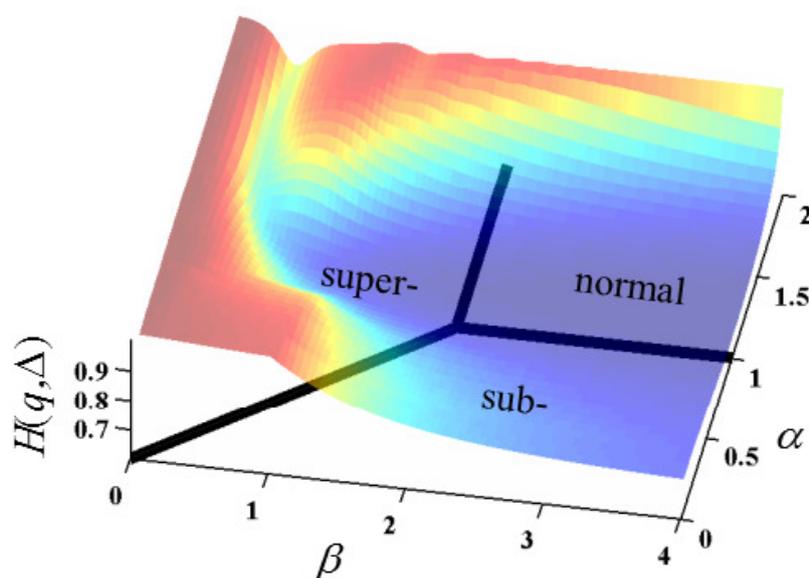
Figure 6 is a slice of the spectral entropy surface (for $\alpha = 1$ and four values of time) from the $\beta = 0$ rim out to the distance of $\beta = 4$. Selecting one case of the argument, say $D_{1,\beta} = 1, t = 1$, and starting at $\beta = 2$, we observe that the entropy increases as β gets smaller, with an approximately 20% increase in the normalized spectral entropy when $\beta = 1$ (the Cauchy distribution), whereas travel in the direction of increasing β is mostly flat by this measure of entropy. From the Gaussian location, $\beta = 2$, the entropy appears to converge to a value near 0.5 for increasing β, while for decreasing β, the entropy increases in a monotonic manner at short times. The effect of increasing diffusion time (or larger values of the diffusion coefficient) results in a decrease in the magnitude for the normalized entropy values, as demonstrated going from $t = 0.5$ to $t = 2$. In the phase diagram (Figure 2: $\alpha = 1$ and $\beta < 2$),

Figure 6 evaluates the case of super-diffusion, and it is encouraging that this perspective portrays this regime, which includes the CF of the Cauchy distribution, as one of higher entropy (in comparison with the Gaussian diffusion case).

Figure 6. Spectral entropy for Equation (15) with respect to the order of the fractional space derivative, β, with diffusion time cases where $t = 0.5, 1, 1.5, 2$ for $\alpha = 1$ and $D_{1,\beta} = 1$.
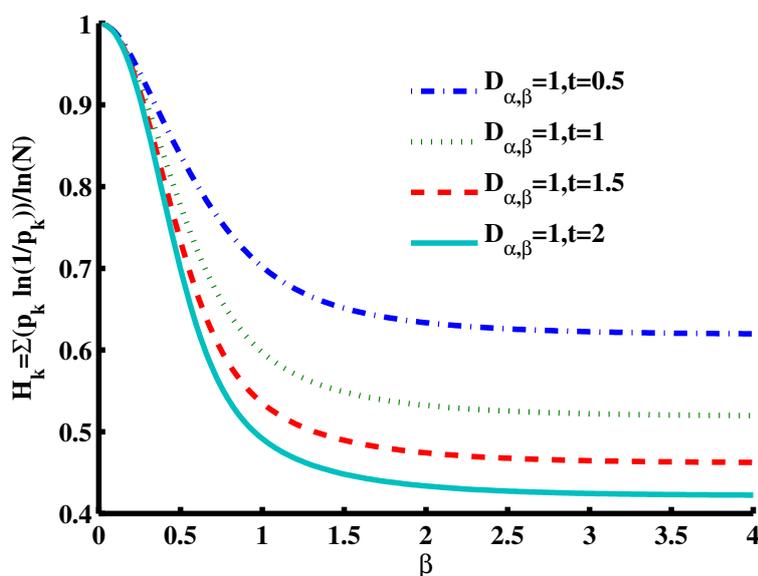
Figure 7. Spectral entropy for Equation (9) with respect to the order of the fractional time derivative, α, for four diffusion time cases where $t = 0.5, 1, 1.5, 2$ for $\beta = 2$ and $D_{\alpha,2} = 1$.
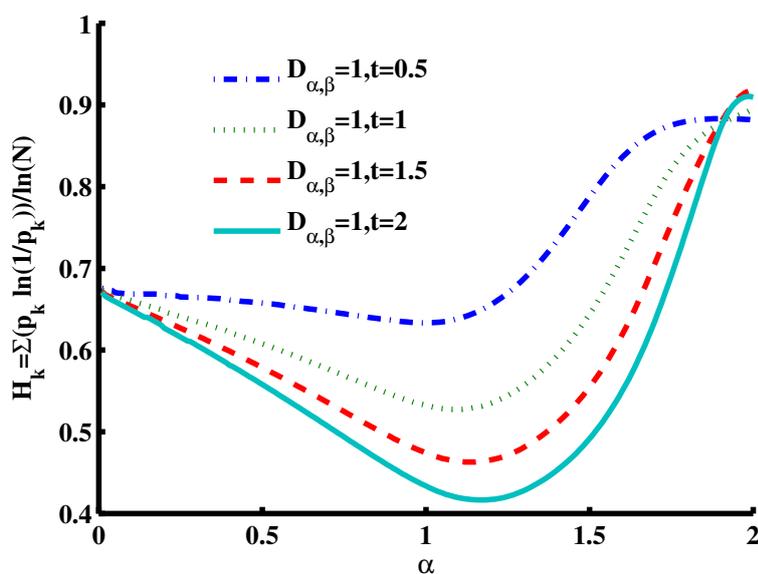
Figure 7 is a slice of the spectral entropy surface (for $\beta = 2$ and four values of time) from the $\alpha = 0$ rim out to the distance of $\alpha = 2$. Selecting one case of the argument, say $D_{\alpha,2} = 1, t = 1$, and starting at $\alpha = 1$, we observe the entropy increasing in both directions, overall. Again, the depth of the minimum grows for longer times, but in this cross-sectional view, the location is in the direction of higher values of α. As is shown in the phase diagram (Figure 2), when $\beta = 2$, values of $\alpha > 1$

are in a region of super-diffusion and values of $\alpha < 1$ are in a region of subdiffusion. Furthermore, in Figure 7, we observe that for a specific value of time (and diffusion coefficient constant), the entropy generally increases (from the Gaussian diffusion case of $\alpha = 1$) as the value of α increases, as well as as it decreases. Thus, both higher and lower values of the order of the fractional derivative α (relative to $\alpha = 1$) give higher entropy values.

In both Figure 6 and Figure 7, it is interesting to note that as the product of the diffusion coefficient and the time increases, the spectral entropy decreases. Mathematically, this behavior is consistent with the Fourier- transform duality between the space and the wavenumber (spatial frequency) domains, in which the diffusion coefficient and time, Dt, change position from the denominator to the numerator of the argument (see Equation (5) and Equation (6) for the case of a Gaussian pdf). Thus, as diffusion time increases, in the framework of the space domain, we expect the distribution to widen and the entropy to increase (increasing variance for the Gaussian). Conversely, as the diffusion time increases, in the framework of the spatial frequency domain, we expect the distribution to narrow and the entropy to decrease. From a CTRW physical model perspective, as the diffusion time increases in the spatial domain, we argue that the distribution widens and the entropy increases as a dynamic measure by which the uncertainty in predicting the location of the diffusing particle increases. As such, more information is required to specify the spatial location of the particle as the diffusion time increases. Conversely, as the diffusion time increases in the spatial frequency domain, we argue that the distribution narrows and the entropy decreases as a dynamic measure by which the amount of information to be gained about the diffusion environment decreases. Therefore, as the CTRW process progresses in time, the environment becomes completely explored, and no new information can be captured about the system, albeit at the cost of maximum uncertainty about the particle's location in space.

Figure 8. Plot of the individual wavenumber contributions to the spectral entropy of Equation (15) when the order of the fractional space derivative $\beta = 0.5, 0.75, 1, 2, 4$ for $\alpha = 1$, $D_{1,\beta} = 1$ and $t = 1$.
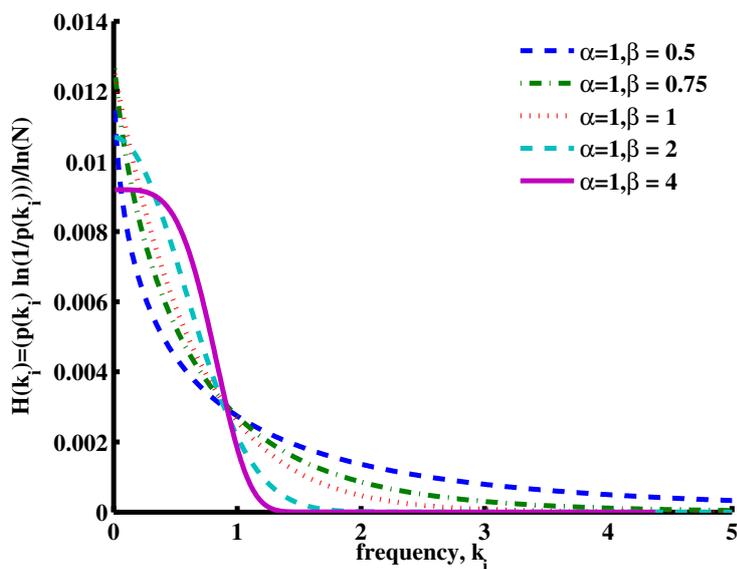
In order to examine further the factors that are summed in Equation (22), we have plotted (for a fixed diffusion coefficient and time) a single spectral entropy term as a function of the wavenumber for a series of β values in Figure 8 (stretched exponential function) and a series of α values in Figure 9 (stretched MLF). Here, for $\beta = 2$, the characteristic Gaussian shape is apparent, and as β decreases into the domain of super-diffusion, the spectrum appears to narrow, but in fact, due to the long power law tail, it actually spreads out, expanding the number and the range of higher wavenumber components. The sum of many of these terms can be interpreted as adding information to the corresponding spatial distribution, increasing its variance and its entropy. Furthermore, in this figure, we note that the Cauchy distribution ($\beta = 1$) has, in comparison with the Gaussian distribution, a wider spectral distribution, with a corresponding increase in spatial complexity and entropy.

Figure 9. Plot of the individual wavenumber contributions to the spectral entropy of Equation (9) when the order of the fractional time derivative $\alpha = 0.5, 1, 1.5, 2$ for $\beta = 2$, $D_{\alpha,2} = 1$ and $t = 1$.
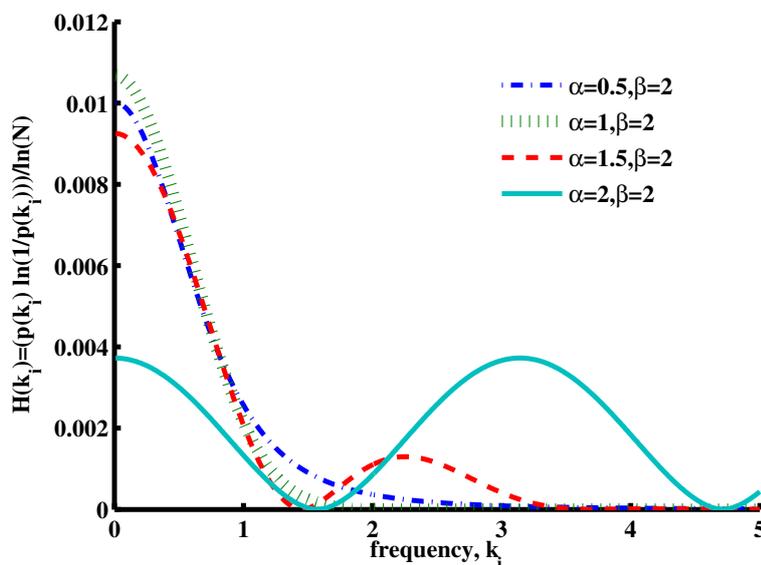
The spectral entropy plotted in Figure 9 has similar features. For $\alpha = 1$, the expected Gaussian distribution of spectral entropy is apparent. When α is reduced to 0.5, the spectra expand (higher uncertainty, higher entropy), and when α is increase to 1.5 and to 2, an oscillation appears in the spectra due to the behavior of the MLF, which again pushes more wavenumber components into the higher range. Such components would be expected to add uncertainty and entropy to the spatial distribution. In the case of $\alpha = 2$, we have a cosine function in the wavenumber, which corresponds to a single, very small spatial feature (a Dirac Delta function) in space.

4. Conclusions

In this study, we have shown that entropy can be used as a measure of the information content, or uncertainty, in the diffusion propagator pdf for cases of Gaussian, sub- and super-diffusion. Both the space (β) and the time (α) fractional order dependence are expressed separately as special cases governed by the Mittag–Leffler function. The classical, Gaussian case of normal diffusion ($\alpha = 1$; $\beta = 2$) is at a

minimum of a surface plot of the spectral entropy for the selected range of α and β in the Mittag–Leffler function. In all directions from the minimum on this surface, the entropy increases, both for increasing and for decreasing values of the orders of fractional differentiation. When either α or β diverge from the Gaussian case, ($\alpha = 1$; $\beta = 2$), the spectra for each component of the total entropy expand or contract in a manner that captures greater overall information about the system. There is an overall reduction in the total spectral entropy as time (or the diffusion coefficient) increases, corresponding to a wider spatial distribution of the individual diffusing components, which is consistent with the noted contraction of the wavenumbers in the Fourier spectral domain. Additionally, for the case of time-fractional subdiffusion in which the second and fourth moments are finite, we have presented a new formulation for excess kurtosis, K, which has a nearly linear increase when the time derivative, α, decreases away from the integer order. Finally, we performed diffusion weighted MRI measurements in the brain of a patient with chronic ischemic stroke in order to demonstrate the utility of α, K_{MLF} and entropy H_{MLF} measures to provide additional information about biological tissue microstructure beyond the classical diffusion coefficient, D.

Acknowledgments

Author Contributions

Carson Ingo designed the research project and performed the data analyses. Carson Ingo and Richard L. Magin performed the theoretical analyses. Todd B. Parrish designed and performed the MRI experiments. Carson Ingo, Richard L. Magin, and Todd B. Parrish wrote the article. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Gorenflo, R.; Vivoli, A.; Mainardi, F. Discrete and continuous random walk models for Space-Time fractional diffusion. *Nonlinear Dyn.* **2004**, *38*, 101–116.
2. Brockmann, D.; Hufnagel, L.; Geisel, T. The scaling laws of human travel. *Nature* **2006**, *439*, 462–465.
3. Ingo, C.; Magin, R.L.; Colon-Perez, L.; Triplett, W.; Mareci, T.H. On random walks and entropy in diffusion-weighted magnetic resonance imaging studies of neural tissue. *Magn. Reson. Med.* **2014**, *71*, 617–627.
4. Metzler, R.; Glackle, W.G.; Nonnenmacher, T.F. Fractional model equation for anomalous diffusion. *Physica A* **1994**, *211*, 13 – 24.

5. Metzler, R.; Klafter, J. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* **2000**, *339*, 1–77.

6. Magin, R.L. *Fractional Calculus in Bioengineering*; Begell House: Redding, CT, USA, 2006.

7. Zaslavsky, G.M. *Hamiltonian Chaos and Fractional Dynamics*; Oxford University Press: New York, NY, USA, 2005.

8. Meerschaert, M.M.; Sikorskii, A. *Stochastic models for fractional calculus*; De Gruyter: Berlin, Germany, 2012; Volume 43.

9. Mittag-Leffler, G. Sur la nouvelle fonction Eα (x). *C. R. Acad. Sci.* **1903**, *137*, 554–558.

10. Mittag-Leffler, G. Sur la representation analytique d'une branche uniforme d'une fonction monogene. *Acta Math.* **1905**, *29*, 101–181.

11. Balanda, K.P.; MacGillivray, H. Kurtosis: A critical review. *Am. Stat.* **1988**, *42*, 111–119.

12. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.

13. Viertiö-Oja, H.; Maja, V.; Särkelä, M.; Talja, P.; Tenkanen, N.; Tolvanen-Laakso, H.; Paloheimo, M.; Vakkuri, A.; Yli-Hankala, A.; Meriläinen, P. Description of the entropy algorithm as applied in the Datex-Ohmeda S/5 entropy module. *Acta Anaesth. Scand.* **2004**, *48*, 154–161.

14. Kirsch, M.R.; Monahan, K.; Weng, J.; Redline, S.; Loparo, K.A. Entropy-based measures for quantifying sleep-stage transition dynamics: Relationship to sleep fragmentation and daytime sleepiness. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 787–96.

15. Smith, S.M. Fast robust automated brain extraction. *Hum. Brain Mapp.* **2002**, *17*, 143–155.

16. Podlubny, I. The Mittag–Leffler function. Available online: www.mathworks.com/ matlabcentral/ fileexchange/8738-mittag-leffler-function (accessed on 5 November 2014).

17. Gorenflo, R.; Loutchko, J.; Luchko, Y. Eα, β (z) and its derivative. *Fract. Calc. Appl. Anal.* **2002**, *5*, 491–518.

18. Jensen, J.H.; Helpern, J.A.; Ramani, A.; Lu, H.; Kaczynski, K. Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **2005**, *53*, 1432–1440.