Mesoscopic structure of neuronal tracts from time-dependent diffusion

Lauren M. Burcaw, Els Fieremans, Dmitry S. Novikov

Bernard and Irene Schwartz Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, New York, NY 10016, USA

A R T I C L E   I N F O

Article history:
Received 22 February 2014
Accepted 23 March 2015
Available online 30 March 2015

Keywords:
Diffusion
MRI
Time dependence
White matter
Microstructure
Axonal diameters

A B S T R A C T

Interpreting brain diffusion MRI measurements in terms of neuronal structure at a micrometer level is an exciting unresolved problem. Here we consider diffusion transverse to a bundle of fibers, and show theoretically, as well as using Monte Carlo simulations and measurements in a phantom made of parallel fibers mimicking axons, that the time dependent diffusion coefficient approaches its macroscopic limit slowly, in a \( \langle \ln t \rangle/t \) fashion. The logarithmic singularity arises due to short range disorder in the fiber packing. We identify short range disorder in axonal fibers based on histological data from the splenium, and argue that the time dependent contribution to the overall diffusion coefficient from the extra-axonal water dominates that of the intra-axonal water. This dominance may explain the bias in measuring axon diameters in clinical settings. The short range disorder is also reflected in the asymptotically linear frequency dependence of the diffusion coefficient measured with oscillating gradients, in agreement with recent experiments. Our results relate the measured diffusion to the mesoscopic structure of neuronal tissue, uncovering the sensitivity of diffusion metrics to axonal arrangement within a fiber tract, and providing an alternative interpretation of axonal diameter mapping techniques.

Introduction and overview of results

Diffusion-weighted magnetic resonance imaging (dMRI) has become a powerful tool for imaging the central nervous system, particularly with the advent of diffusion tensor imaging (DTI) (Basser et al., 1994). Due to the directional and microstructural sensitivity of the diffusion-weighted signal (which has a meaning of the diffusion propagator), dMRI has become an indispensable tool for studying white matter (Beaulieu, 2002).

The diffusion propagator depends separately on the molecular displacement \( \Delta x \) via the spatial Fourier wave vector \( q \), and on the diffusion time \( t \). Fig. 1. So far, the brain dMRI development has mainly focused on exploring as much of the \( q \)-space as is possible given clinical limitations, under the overarching theme of \( q \)-space imaging (Callaghan et al., 1988). This corresponds to proceeding vertically in Fig. 1, with diffusion time fixed at \( t \sim 50–100 \text{ ms} \) in the clinical settings.

Here we explore the horizontal direction in Fig. 1, along the \( t \)-axis. We focus on the origin of the time dependence of mean squared displacement \( \langle \Delta x(t)^2 \rangle \) of water molecules, characterized by the diffusion coefficient \( D(t) = \langle \Delta x(t)^2 \rangle/2t \).

As axonal walls are believed to be the major restrictions to diffusion in the white matter (Beaulieu, 2002), the biophysical challenge is to identify which \( \mu \)-level features of axonal packing within a fiber tract are most pronounced in a dMRI measurement with a macroscopic, mm-level resolution. It turns out that the relative importance of different features of axonal geometry depends on the measurement time scales.

Early understanding of \( D(t) \) relied on the result of Mitra et al. (1992), who found that the initial decrease of \( D(t) \) is determined solely by the net surface area of all barriers, such as cell walls, irrespective of their positions. With transverse dimensions of axons and dendrites of \( \sim 1 \mu m \) (Lamantia and Rakic, 1990; Aboitiz et al., 1992), and water diffusion coefficient \( \sim 1 \mu m^2/\text{ms} \), this limit demands time scales below 1 ms. These are unrealistically short for the human brain dMRI.

Fortunately, the time dependence observed in neuronal tissue (Horsfield et al., 1994; Does et al., 2003; Assaf et al., 2008; Alexander et al., 2010; Aggarwal et al., 2012; Portnoy et al., 2013; Kunz et al., 2013; Pyatigorskaya et al., 2014; Kershaw et al., 2013; Baron and Beaulieu, 2014; Van et al., 2014; Burcaw et al., 2014) extends onto feasible times \( t \sim 1–10 \text{ ms} \) and even beyond, which prompts us to explore the long time limit of \( D(t) \), and its connection to axonal tract geometry. This is the main objective of this work.

For \( t \) long enough so that water molecules travel past one or more axons, it is not just the net amount of the restrictions, such as axonal membranes and myelin, but the correlations in their positions and orientations that determine the diffusion propagator (Novikov and Kiselev, 2010) and the derived diffusion metrics (Szafer et al., 1995; Fieremans et al., 2008; Novikov and Fieremans, 2012; Fieremans et al., 2012). In particular, when \( D(t) \) is approaching its macroscopic \( t \to \infty \) (tortuosity) asymptote \( D_m \), the short-distance details become irrelevant, while the behavior of \( D(t) \) is determined by how spatially correlated are the restrictions at large distances (Novikov et al., 2014). These correlations are characterized by the structural exponent \( p \), cf. Theory section.
In this work, we apply the framework (Novikov et al., 2014) to diffusion in the extra-axonal space at experimentally feasible diffusion times, and obtain the following four results:

(i) By analyzing histology data, we identify the short range disorder in the two-dimensional fiber tract cross-section (exponent \( p = 0 \)), and determine how this randomness in fiber packing manifests itself in the long-time behavior of the common diffusion metrics. These results, summarized in Table 1, are independent of whether axonal walls are permeable or not.

(ii) By performing time-dependent pulsed-gradient (PG) DTI on a fiber phantom designed to mimic the diffusion in the extra-axonal space, as well as using Monte Carlo simulations, we confirm the logarithmic singularity in \( D(t) \), cf. Table 1:

\[
D(t) = D_0 + A \ln \left( \frac{t}{\delta_c} \right), \quad t \gg \delta_c. \tag{1}
\]

Here the coefficient \( A \) is proportional to the value \( \Gamma(k) |_{k \to 0} \) of the power spectrum of fibers within the two-dimensional (2d) cross-section [cf. Appendix A, Eq. (A.4)], and \( \delta_c \) represents the time to diffuse across the correlation length \( l_c \) of the 2d random packing geometry. For realistic axonal packings, \( l_c \) closely follows mean external axonal radius, with \( t_c \leq 1 \) ms for the brain, while \( A \) roughly scales with \( l_c^2 \), Table 2. The \( \ln(t)/t \) singularity occurs due to the short range 2d disorder in fiber packing (exponent \( p = 0 \)), and is absent in an ordered (lattice) arrangement (\( p = \infty \)), as our simulations demonstrate. For PG with pulse width \( \delta > t_c \), \( \ln(t)/t \) in Eq. (1) crosses over to \( \ln(t)/t \), \( t \gg \delta \).

(iii) Equivalently, the 2d short-range disorder leads to the characteristic \( -|\omega| \) frequency dependence, cf. Table 1,

\[
\text{Re}D(\omega) = D_0 + A \cdot \frac{\pi}{2} |\omega|, \quad |\omega| t_c \ll 1 \tag{2}
\]

of the oscillating gradient (OG) measured diffusion coefficient \( D(\omega) \) transverse to axons. By virtue of the cumulant expansion (Kiselev, 2010), Eq. (1) or (2) allows one to calculate the effect of the disordered fiber geometry on any gradient wave form. We show that recent OG ex vivo data (Portnoy et al., 2013) exhibit the behavior (2). This confirms our conclusion about the short range disorder, \( p = 0 \), in fiber packing. Hence, the structural disorder renders OG measurements insensitive to intra-axonal water at low \( \omega \), as the latter contributes a less pronounced, quadratic frequency dependence \( -\omega^2 \ll |\omega| \) for \( \omega \to 0 \).

(iv) The singularity (1) provides a different interpretation of axonal diameter mapping results. Under a common assumption of no exchange between compartments, we argue that the contribution (1) is more relevant (decays slower) than the \( 1/t \) contribution from water confined inside axons. Hence, structural disorder in axonal packing amplifies the role of extra-axonal water. This may explain (Fig. 2 and Table 3) the well-known overestimation of axon diameters, by factors of \(-3\)–\(-5\) or even more, in clinical dMRI (Alexander et al., 2010; Zhang et al., 2011). This bias has been previously attributed to the insensitivity of the AxCaliber (Assaf et al., 2008) scheme to small axons (Dyrby et al., 2013) or to the noise. We believe the reason may be more fundamental: when quantifying cell sizes, the residual time-dependent (and hence non-Gaussian) nature of diffusion in the extracellular space should not be neglected. This is especially relevant for clinical dMRI, where the extra-axonal signal is dominated by its first few cumulants.

**Table 1**
A summary of the long-time dependencies of the instantaneous diffusion coefficient, \( D_{\text{inst}}(t) \), Eq. (4), PG-measured cumulative diffusion coefficient, \( D(t) \), Eq. (3), and the OG-measured frequency-dependent diffusion coefficient, \( D(\omega) \), Eq. (15), in 2d. Here \( \delta_c \) is defined in Eq. (1).

<table>
<thead>
<tr>
<th>Structure</th>
<th>( D_{\text{inst}}(t) )</th>
<th>PG: ( D(t) )</th>
<th>OG: Re(D(\omega))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordered (( p = \infty )), or confining</td>
<td>( e^{-t/\delta_c} )</td>
<td>( \frac{t}{\delta_c} )</td>
<td>( \frac{\omega^2}{</td>
</tr>
<tr>
<td>Short range disordered (( p = 0 ))</td>
<td>( \frac{1}{\delta} )</td>
<td>( \ln(\delta/c_1) )</td>
<td>(</td>
</tr>
</tbody>
</table>

**Table 2**
Results from experiment, and from MC simulations in random disk packings for the geometries of the Dynenea phantom (Fig. 6b) and of sectors \( 2, 4, 6 \) of monkey CC (Fig. 11). Parameters \( D, A \) and \( \delta_c \) are obtained by fitting Eqs. (1) and (8) to experiment and MC data. The correlation length estimated as \( l_c \) is \( 4D t_c \), closely follows the mean external radius \( \langle r_{\text{ex}} \rangle \) of the packings, while the amplitude \( A \) scales roughly with \( l_c^2 \), as \( A \approx 0.2 \cdot l_c^2 \), cf. Supplementary fig. 54.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Experiment</th>
<th>MC phantom</th>
<th>Sector 2</th>
<th>Sector 4</th>
<th>Sector 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_0 ), ( \mu )m(^2)/ms</td>
<td>1.8</td>
<td>1.8</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>( D_0 ), ( \mu )m(^2)</td>
<td>0.65</td>
<td>0.66</td>
<td>0.81</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>( \delta_c ), or ( t_c ), ms</td>
<td>13.90</td>
<td>13.76</td>
<td>0.132</td>
<td>0.319</td>
<td>0.276</td>
</tr>
<tr>
<td>( l_c ), ( \mu )m</td>
<td>6.0</td>
<td>6.0</td>
<td>0.65</td>
<td>1.01</td>
<td>0.94</td>
</tr>
<tr>
<td>( \langle r_{\text{ex}} \rangle ), ( \mu )m</td>
<td>8.5</td>
<td>8.5</td>
<td>0.53</td>
<td>0.73</td>
<td>0.88</td>
</tr>
</tbody>
</table>

**Fig. 2.** Apparent internal axonal radius \( r_{\text{app}} \) for rhesus monkey corpus callosum sectors \( 2, 4, 6 \) (red, black and green), obtained from MC-simulated overall diffusion coefficient \( D(t) \) in the narrow-pulse limit (dashed) and with pulse width \( \delta = 10 \) ms (derived from MC data, with diffusion time \( t_c \Delta > \delta \)), by attributing all time dependence to the intra-axonal water. [Results apply for regions I and II of the phase diagram, cf. Discussion, Fig. 12; see Eqs. (34) and (39), and Table 3.] Note that \( r_{\text{app}}(t) \) greatly exceeds the corresponding mean radius \( \langle r \rangle \) (thin lines); for \( \delta = 10 \) ms the extra-axonal contribution is completely dominant. Singularity (1) is reflected in \( r_{\text{app}} \sim \ln t, \delta = 0 \), and \( r_{\text{app}}^p \sim \ln t, \delta > t_c \).
Table 3
The relative contribution of the time-dependent part of the extra-axonal diffusion coefficient $D_{\text{ext}}(t)$ in the time dependence of the total $D(t)$, cf. Fig. 11, and the corresponding overestimation of $r_{\text{app}}$ compared to mean internal axon radius $r\text{'}$, for $t = 10$ ms and $100$ ms, cf. Fig. 2.

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Sector 2</th>
<th>Sector 4</th>
<th>Sector 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{\text{ms}}$</td>
<td>$D_{\text{ext}}(t) - D_{\text{inst}}$</td>
<td>$D(t) - D_{\text{inst}}$</td>
<td>$r_{\text{app}}: r\text{'} = 0$</td>
</tr>
<tr>
<td>10 ms</td>
<td>0.84</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>100 ms</td>
<td>3.33</td>
<td>3.99</td>
<td>3.47</td>
</tr>
</tbody>
</table>

The outline of this paper is as follows. Theoretical arguments in favor of results (i)–(iii) and Eqs. (1)–(2), as well as establishing $p = 0$ for a fiber tract, are given in Theory section. After we present our methods, the experimental and numerical validation of (ii) and (iii) and their implications for OG and PG measurements are given in Results section. Our new interpretation (iv) of axonal diameter mapping, the phase diagram for PG measurements transverse to axons (Fig. 12), and the explanation of how Fig. 2 has been obtained, are outlined in Discussion section.

**Theory**

One of the fundamental questions of brain dMRI is to identify relevant μm-level structural features that affect the observed signal the most. Addressing it would help develop parsimonious models of diffusion in neuronal tissue in order to become specific to pathological changes at the earliest stages of disease.

Here we identify and quantify the effect of the randomness (disorder) in packing of neuronal fibers in a bundle (Fig. 3) as the relevant structural characteristic observable in brain dMRI, especially at the clinically feasible diffusion times. Our main object of investigation will be the power spectrum $\Gamma(k)$ of the restrictions, Fig. 4, which is the Fourier transform of their density correlation function $\Gamma(r)$ (cf. Appendix B). We will demonstrate that the way nature packs axons within a bundle, Figs. 3b–f, is qualitatively similar to the random packings produced in Figs. 3b–d in all cases, the packings are of the so-called short-range disorder type. The similarity in the structure will lead to similarly behaving diffusion metrics and to the qualitative differences from those in the ordered packing of Fig. 3a, see Table 1.

**Relation of time dependent diffusion to structural correlations**

The PG diffusion weighting in the narrow-pulse limit yields the time-dependent diffusion coefficient, commonly defined as (Callaghan, 1991)

$$D(t) = \frac{\langle |\delta x(t)|^2 \rangle}{\langle \delta x(t)^2 \rangle} = \frac{\langle |\delta x(t)|^2 \rangle}{2t}, \quad t \geq 0.$$  

(3)

where $\delta x = x(t) - x(0)$ is the displacement in a chosen direction $\hat{x}$ along the Brownian path $\mathbf{r}(t)$. For the purpose of investigating the temporal evolution, we find it useful to define the instantaneous diffusion coefficient via a time derivative,

$$D_{\text{inst}}(t) = \frac{\partial}{\partial t} \langle |\delta x(t)|^2 \rangle = \frac{\partial}{\partial t} \langle D(t) \rangle, \quad t \geq 0.$$  

(4)

While the PG diffusion coefficient (Eq. (3)) characterizes the average rate of the evolution of the mean squared displacement over the whole measurement interval $t$ and is a cumulative quantity, the instantaneous one describes its rate of change at the moment $t$. In uniform fluids, $D = D_{\text{inst}} = \text{const}$. In heterogeneous media such as tissues, it is $D_{\text{inst}}(t)$ that is most directly related to structure probed at the current diffusion length scale $L(t) = \sqrt{\langle |\delta x(t)|^2 \rangle}$.

**Fig. 3.** Examples of 2d packings, relevant for the extra-axonal diffusion transverse to fibers, with their correlation functions shown in Fig. 4. a, Order: a square lattice of disks. b, Random non-overlapping packing of disks with a volume fraction $\delta = 0.75$. c, The same as for b, but with a volume fraction $\delta = 0.5$. Note the disks have the same distribution of radii as in b, but we show a larger subset of the full packing. d, Random disk pack where disks are allowed to freely overlap (Poissonian disorder), $\delta = 0.55$. e, Electron microscope image of the splenium of a 20-week-old female C57BL16 mouse (Zurek et al., 2014). f, Axons from e, including myelin, outlined by hand, with calculated volume fraction $\delta = 0.7$, exemplifying restrictions to diffusion in extra-axonal space.
Fig. 4. Angular-averaged Fourier transform \( \Gamma(k) \) of the normalized density correlation function (B.1) (equivalently, the power spectrum, Eq. (B.2)) for the five structures shown in Fig. 3, as function of radial \( k \), which is made dimensionless by multiplying by the average object radius \( \text{rad}_{\text{ext}} \) for each packing (cf. Appendix B). Square lattice of disks (dashed red line), Fig. 3a, exhibits Bragg peaks, such that \( \Gamma(k) \equiv 0 \) for \( k \) below the first peak, corresponding to exponent \( p = \infty \). Randomly placed disks without overlap with a volume fraction \( \phi = 0.75 \) (blue), Fig. 3b; \( \phi = 0.5 \) (magenta), Fig. 3c; and with overlap (black), Fig. 3d, exhibit short-range correlations, with the finite plateau \( \Gamma(k) \mid_{k \rightarrow \infty} = \text{const} > 0 \), corresponding to exponent \( p = 0 \). As expected by how they are packed, these arrangements are progressively more disordered, consistent with an increase in the plateau value \( \Gamma(k) \mid_{k \rightarrow 0} \). The density correlator for axes, Fig. 3f, with \( \phi = 0.7 \), seems to approach a finite plateau \( \Gamma(k) \mid_{k \rightarrow 0} \), corresponding to the exponent \( \phi \). The plateau value is roughly that of the randomly packed disks with \( \phi = 0.5 \), indicating a fairly disordered arrangement. For all packings, the disorder correlation length \( \delta \) decreases at large \( k \) for all curves is due to the 2d curved boundaries (see Appendix B), a short-scale signature of the individual objects that is not relevant for the long-time behavior of diffusion.

Definitions (3) and (4) assume the narrow-pulse limit, where the pulse width \( \delta \) is smaller than any characteristic time scale. For understanding the physics of diffusion in restricted geometry, these fundamental metrics are the most transparent ones to consider; the effects of finite \( \delta \) will be derived from them below.

At long \( t \), the way of approaching the macroscopic limit \( D_{\text{in}} \)

\[
D_{\text{in}}(t) \approx D_{\text{in}} + \text{const} \cdot t^{-\theta}, \quad t \to \infty
\]  

is characterized by a dynamical exponent \( \theta \) whose value is directly connected to the structure via (Novikov et al., 2014)

\[
\theta = (p + d)/2
\]  
in \( d \) spatial dimensions, generalizing earlier approaches (Ernst et al., 1984; Machta et al., 1984; Vischer, 1984). The structural exponent \( p \) determines the character of the long-range spatial correlations of the restrictions, via the low-\( k \) behavior of their power spectrum (Appendix B)

\[
\Gamma(k) \sim k^{\theta}, \quad k \to 0.
\]

The relations (5) and (6) enable grouping differently looking media (tissues) into a relatively few distinct structural universality classes defined in terms of their structural exponent \( p \). In other words, the dynamics of the observable diffusion coefficient is qualitatively similar for all \( d \)-dimensional media with the same structural exponent \( p \) describing the long-range \((k \to 0)\) behavior (Eq. (7)) of the power spectrum of the restrictions (Novikov et al., 2014).

Here we consider the two most common classes, short range disorder \((p = 0)\) and order \((p = \infty)\). We will show that the former corresponds to neuronal tissue, and use the comparison with the latter to exemplify the role of the structural disorder. Other disorder universality classes (Novikov et al., 2014) are characterized by distinct values of the exponent \( p \) in Eq. (7).

Fig. 4 shows the 2d angular-averaged density power spectrum, or equivalently, the spatial Fourier transform \( \Gamma(k) \), Eq. (B.2), of the density correlators (Eq. (B.1)), for our examples from Fig. 3 (see Appendix B for details). We observe that the low-\( k \) behavior, which reflects the long range fluctuations, is qualitatively similar for the disordered arrangements of Fig. 3, Panels b–d, and distinctly different in the periodic (ordered) case of Fig. 3a.

Short range disorder is characterized by a finite plateau \( \Gamma(k) \mid_{k \to 0} > 0 \). It corresponds to exponent \( p = 0 \) in Eq. (7), and reflects the absence of long range correlations at distances beyond the disorder correlation length \( l_{c} \sim 1/k \), where \( k_{c} \) is of the order of the extent of the low-\( k \) plateau (i.e. \( \Gamma(r) \) is a peak of the width \( \sim l_{c} \)). This is the most prevalent disorder class, qualitatively similar to Poissonian disorder (for which there are no spatial correlations, e.g. for the centers of the disks in Fig. 3d).

For the opposite case of ordered arrangement, Fig. 3a, \( \Gamma(k) \) is a set of \( \delta \)-function peaks at the reciprocal lattice vectors \( k_{0} \), so that \( \Gamma(k) \equiv 0 \) when \( k \) is within a circle of radius less than the distance to the smallest reciprocal lattice vector \( k_{0} \). While this \( \Gamma(k) \) does not possess axial symmetry, we can still consider the angular-averaged \( \Gamma(k) \) so as to compare with the disordered cases above. This \( \Gamma(k) \) vanishes identically, \( \Gamma(k) = 0 \), for \( k < k_{c} \). In the case of Fig. 3, \( k_{0} = \pi/k_{0} \), where \( k_{0} \) is the lattice constant (shortest distance between disk centers). From the point of the general classification (Novikov et al., 2014), this means that \( \Gamma(k_{0}) \) is smaller than any power law \( k^{\theta} \) with arbitrary large \( p \) as \( k \to 0 \), hence formally \( p = \infty \) for any perfect lattice.

The relations (5) and (6) now predict qualitatively different behavior for \( D_{\text{in}}(t) \). The short range disorder in Figs. 3b–d yields the exponent \( \theta = 1 \) in \( d = 2 \) dimensions, resulting in

\[
p = 0: \quad D_{\text{in}}(t) \approx D_{\text{in}} + A \cdot t^{\theta}, \quad t \gg t_{c}.
\]  

where \( t_{c} \) is the time to diffuse across the disorder correlation length \( l_{c} \) (in our examples, \( l_{c} \) is of the order of the disk size, see Table 2 and Fig. 4). In Appendix A we show that the coefficient \( A \) is proportional to the plateau value \( \Gamma(k) \mid_{k \to 0} \) in Fig. 4. Roughly speaking, the value of \( A \) reflects the “strength” of the short range disorder component, and determines how pronounced is the power law tail (Eq. (8)).

On the other hand, \( D_{\text{in}}(t) \) in an ordered medium approaches its macroscopic limit exponentially fast, as there is no plateau, \( \Gamma(k) \mid_{k \to 0} \equiv 0 \), and thus no power law, \( A \equiv 0 \):

\[
p = \infty: \quad D_{\text{in}}(t) \approx D_{\text{in}} + \text{const} \cdot e^{-t/t_{0}}, \quad t \gg t_{0}.
\]  

The exponential approach of \( D_{\text{in}} \) is faster than any inverse power law \( t^{-\theta} \). This allows us, using the nomenclature of Novikov et al. (2014), to formally assign the dynamical exponent \( \theta = \infty \) to any lattice; this value of \( \theta \) is then in agreement with Eq. (6) for \( p = \infty \). Qualitatively, one can view any lattice as a limiting case of a so-called hyperuniform arrangement (Torquato and Stillinger, 2003) (with infinitely small long-range fluctuations).

The fast decrease (Eq. (9)) in \( D_{\text{in}}(t) \) occurs for times exceeding \( t_{0} \), which is of the order of the diffusion time across the period \( t_{0} \). Note that a similar, exponentially fast approach of the value \( D_{\text{in}} = 0 \) also occurs when random walkers are confined in an impermeable pore, cf. Appendix C, Eq. (C.5) and Fig. C.1, for an impermeable cylinder example. As we can see, \( D_{\text{in}}(t) \) can be used to probe long-range correlations in tissue architecture. Measurement of the exponent \( \theta \), with the knowledge of sample’s dimensionality \( d \), can allow one to identify whether the sample is ordered or disordered, and determine the disorder universality class. The latter can be also identified by determining the exponent \( p \) from histology. Below we determine \( p \), as well as find out how \( \theta \) can be identified in practically measurable PG and OG diffusivities.
Structural and dynamical exponents transverse to neuronal fibers

What is the structural exponent \( p \) for white matter? To answer this question, we obtain an electron microscope histology image of a mouse splenium (Zurek et al., 2014), shown in Fig. 3e, and identify myelinated axons as the restrictions for water molecules in the extra-axonal space. We then manually outline all these axons, including both intra-axonal space and myelin, assign to them a unit density \( \rho(x) = 1 \), and generate an image shown in Fig. 3f, similar to those for the disk packings. The correlation analysis as described above now yields the finite plateau \( \Gamma(k) \propto \text{const} > 0 \) in Fig. 4, corresponding to exponent

\[
p = 0 \quad \text{transverse to white matter fibers} \quad (10)
\]

and a very short disorder correlation length \( l_c - 1/k_c - (\rho_{\text{ex}}) \leq 1 \mu m. \) While the plateau in \( \Gamma(k) \) is not fully pronounced due to a relatively small number of axons within the available histology slide, we believe that there is neither a tendency of ordering (which would result in the decrease of \( \Gamma(k) \) as \( k \to 0 \)), nor strong fluctuations (which would cause the divergence in \( \Gamma(k) \) as \( k \to 0 \)).

We note that \( p \) may seem slightly negative, which would indicate density fluctuations somewhat stronger than those of short-range disorder. We currently believe that this is most likely an artifact of a finite-size histology slide: indeed, the exponents \( p \) take well-defined discrete values, and we are not aware of a universality class with a very short disorder correlation length.

To conclude, the analysis of the available histology is compatible with the short range disorder in the packing of neuronal fibers within a bundle.

The observation (Eq. (10)) has a number of consequences.

1. Short range disorder is the most common disorder type — nature did not surprise us here. We could have expected reduced spatial fluctuations, if evolution’s goal were to pack axons most efficiently. That would have led to \( p > 0 \), similar to the exponent \( p \approx 1 \) for the “maximally random jammed” packing of spheres in \( d = 3 \) dimensions (Donev et al., 2005a). Conversely, axons do not look particularly “ jammed”.

2. The absence of very strong long-range fluctuations (which would be characterized by a strongly diverging \( \Gamma(k) \) with \( p < -d \)) is consistent with having a finite nonzero macroscopic diffusion constant \( D_{\text{w}} \) transverse to the axons, and is incompatible with the so-called anomalous diffusion (Bouchaud and Georges, 1990) for which \( D_{\text{w}} \) does not exist. In other words, under the common assumption of cell walls providing the most important restrictions to diffusion, at distances beyond \( l_c \) the hindered extra-axonal diffusion should asymptotically become normal. Our histology-based evidence contradicts the models of the “anomalous”, or “fractal” diffusion in the brain (continuous time random walks, stretched exponentials). Here we showed that the Gaussian limit is approached slowly, so that the residual non-Gaussianity can matter at long but finite \( t \). This slowness is a likely reason for relatively good fits provided by unphysical stretched-exponential models in a finite range of \( q \) and \( t \).

3. Based on Eq. (6), we predict the dynamical exponent

\[
\vartheta = 1 \quad \text{transverse to white matter fibers} \quad (11)
\]

which will result in the behavior (1) and (2) as we discuss below. (In the unlikely case of \( p \) turning out to be small and negative as mentioned after Eq. (10), the corresponding \( \vartheta < 1 \) would make the time-dependence of diffusion in the extra-axonal space even more dominant than we anticipate.)

4. The exponent (Eq. (10)) gives us a practical recipe to represent and model the packing disorder in white matter. For instance, Monte Carlo simulations should be performed for the random packings rather than for the periodic ones, even at the cost of more challenging implementation. Below, we will identify the dynamical exponent (Eq. (11)) in the Monte Carlo data for diffusion transverse to simulated phantoms such as shown in Figs. 3b, and c.

5. The short range disorder identified in the white matter justifies building artificial diffusion phantoms in which the same disorder class occurs, e.g. due to random packing of hydrophobic fibers in a bundle. In what follows, we will describe how we identify the dynamical exponent \( \vartheta = 1 \), and thereby the structural exponent \( p = 0 \), in such a phantom.

Origin of the logarithmic singularity, Eq. (1)

While the power law dependence of \( D_{\text{inst}}(t) \) reveals profound differences in the structure, a narrow-pulse PG MR measurement yields the cumulative diffusion coefficient (Eq. (3)), in which these differences may be masked. From Eq. (4),

\[
D(t) = \frac{1}{T} \int D_{\text{inst}}(t') \, dt'.
\]

One can see that, whenever we are dealing with structures for which \( D_{\text{inst}}(t) < D_{a} \) decreases faster than \( 1/t \), the corresponding difference \( D(t) - D_{a} \) always behaves as \( 1/t \) at long \( t \). This is because the integral in Eq. (12) converges as \( t \to \infty \), and hence, to the leading approximation, it can be substituted by its \( t = \infty \) limit,

\[
\vartheta \equiv \lim_{t \to \infty} \int_{0}^{t} \cdots \int_{0}^{t} = \text{const} < \infty \.
\]

Therefore, e.g. both for a periodic (ordered) medium, and for fully confining pores (cells), the cumulative diffusion coefficient

\[
\vartheta \approx 1 : \quad D(t) \approx D_{a} + \frac{\text{const}}{t}, \quad t \gg t_{c}.
\]

Overall, in all cases with the dynamical exponent \( \vartheta > 1 \), i.e. when \( p > -2 - d \), \( D(t) \) will behave in the same way, Eq. (13), which makes it practically harder to distinguish between media with different types of structural correlations using PG DTI.

The exponent \( \vartheta = 1 \), cf. Eq. (8), which is the subject of our focus, causes a logarithmic divergence \( \int_{0}^{t} \int' \cdots = \ln (t/t_{c}) \) in Eq. (12). Hence, its presence survives, albeit marginally, the temporal averaging inherent to the cumulative \( D(t) \), and yields a logarithmically enhanced time dependence (Eq. (1)) as a function of the upper limit, \( t \), relative to that in an ordered or confining geometry, Eq. (13). This is one of our main results (cf. Table 1) to be checked in simulations and experiment. The lower limit \( t_{c} \) is the time scale beyond which the tail (Eq. (8)) enters the cumulative \( D(t) \) as a function of the upper limit, \( t \), relative to that in an ordered or confining geometry, Eq. (13).

To summarize, the logarithmic singularity (Eq. (1)) in \( D(t) \) arises from averaging over time the slow \( 1/t \) tail in \( D_{\text{inst}}(t) \) coming from the short range disorder in two spatial dimensions.

Implications for OG-measured \( D(\omega) \), Eq. (2)

Practically, the long-time limit for neuronal tissue occurs already for relatively short times from the point of a conventional PG measurement. Hence, to increase the diffusion weighting and to access shorter time scales, it is convenient to utilize oscillating gradient methods (Does et al., 2003).
We begin with the relation [Eq. (8) of Novikov et al. (2014)]

\[
D(\omega) = -io \int_0^\infty dt e^{i\omega t} D_{\text{inst}}(t)
\]  

(14)

between \(D_{\text{inst}}(t)\) and the dispersive diffusivity defined via velocity autocorrelation function (Novikov and Kiselev, 2010)

\[
D(\omega) = \int_0^\infty \langle v(t)v(0)\rangle e^{i\omega t} dt.
\]  

(15)

Eq. (14) is a nonlocal in time relation between \(D_{\text{inst}}(t)\) and \(D(\omega)\), with these two quantities having identical information content. We also note (Novikov and Kiselev, 2010, 2011) that the OG-measured frequency-dependent diffusivity in the limit of large number of oscillations

\[
\frac{1}{T} \int_0^T dt e^{i\omega t} \langle v(t)v(0)\rangle = \text{Re} \, D(\omega)
\]  

(16)

equals the real part of \(D(\omega)\) defined in Eq. (15). Hence, the general relation between the OG measurement and \(D_{\text{inst}}(t)\) is

\[
\text{OG} : \quad \text{Re} \, D(\omega) = \frac{A}{\omega} \int_0^\infty dt \sin \omega t D_{\text{inst}}(t).
\]  

(17)

Let us now focus on the low-frequency behavior of \(D(\omega)\), corresponding to the 1/t tail (Eq. (8)). Using \(\int_0^\infty dt (\sin t)/t = (\pi/2) \text{sgn} \omega\), Eq. (17) yields the result (Eq. (2)) advertised in the Introduction section, which is the OG counterpart of Eqs. (1) and (8). Hence, the dynamical exponent \(d = 1\) is reflected in the 1/t \(\sim (\pi/2)|\omega|\) correspondence between the tails in \(D_{\text{inst}}(t)\) and in \(\text{Re} \, D(\omega)\). We also note that the 1/t tail in \(D_{\text{inst}}\) is equivalent to the tail

\[
p = 0 : \quad \langle v(t)v(0)\rangle \sim -\frac{A}{\omega^2}, \quad t \gg t_c.
\]  

(18)

(cf. Ernst et al., 1984) in the velocity autocorrelation function, in contrast to its exponentially fast decay \(\sim -e^{-t/t_0}\) for ordered \((p = \infty)\) or confined geometries.

It is instructive to find the full \(D(\omega)\), Eq. (14). Estimating

\[
\int_{t_c}^\infty \frac{dt}{t} e^{i\omega t} \approx \int_{t_c}^{1/|\omega| t_c} \frac{dt}{t} = \ln \frac{1}{|\omega| t_c}, \quad |\omega| t_c \ll 1
\]

with logarithmic accuracy (neglecting terms \(-1\) as compared to the large logarithm above), we choose \(t_c\) as the lowest limit of integration (time scale from which the tail begins), and \(-1/\omega\) as a long-time cutoff provided by the oscillating \(e^{i\omega t}\). Thus

\[
p = 0 : \quad D(\omega) \sim D_0 - \frac{i\omega t_0}{1 - i\omega t_0} = -i\omega t_0 + \omega^2 t_0^2 + \ldots
\]  

(20)

cf. Eq. (D.1). Note the crucial difference from Eq. (19): Eq. (20) is analytic in \(\omega\) at \(\omega = 0\), i.e. it has a regular Taylor expansion with a finite convergence radius \(1/t_0\) in the complex plane of \(\omega\).

---

**Fig. 5 a.** The real (solid blue line) and imaginary (dashed red line) parts of \(D(\omega)\). Note that \(\text{Re} \, D(\omega)\) is an even function, which requires even powers of \(\omega\) in its Taylor expansion, in case it is analytic around \(\omega = 0\). By this token, an OG-measured linear-in-frequency behavior of \(\text{Re} \, D(\omega)\) signifies the non-analytic behavior (Eq. (19)) for the full \(D(\omega)\), involving the modulus \(|\omega|\) rather than \(\omega\) for \(\text{Re} \, D(\omega)\) to remain even. b. A graphical representation of Eq. (31). At low \(\omega\), the singular (non-analytic) \(|\omega|\) term, which is a signature of short-range disorder, will dominate the regular (analytic) \(\omega^2\) contribution to \(D(\omega)\).
Since \( \langle v(t) v(0) \rangle \) is real-valued, the real part of its Fourier transform (Eq. (15)) must be even in \( \omega \). If \( D(\omega) \) is also analytic around \( \omega = 0 \), its Taylor expansion must involve only even powers of \( \omega \), starting from a quadratic one,

\[
p = \infty : \quad \text{Re} \, D(\omega) = D_0 + \text{const} \cdot \omega^2,
\]

in agreement with Eq. (20) and with Appendix D.

In contrast, for the random geometry, the convergence radius of \( D(\omega) \) at \( \omega = 0 \) is zero, due to the non-analytic term \(-i \omega^6 \) originating from the long tail (Eq. (5)). If \( \omega > 2 \), the lowest order behavior (Eq. (21)) will mask the existence of the term \(-i \omega^6 \), much like the \( 1/t \) term masks the long-time tail in PG \( D(t) \), Eq. (13). In our case of \( \omega = 1 \), \((-i \omega)^6 \rightarrow -i \omega \ln(-i \omega) \), so already the lowest order in \( \omega \) in Eqs. (2) and (19) is explicitly non-analytic at \( \omega = 0 \).

The absolute value (modulus) \( |\omega| \) in the real part (Eq. (2)) is essential: \( \text{Re} \, D(\omega) \) cannot have odd powers of \( \omega \) (without the modulus) in its low-frequency expansion (see Fig. 5a) as it must stay even in \( \omega \). The only way to have a term linear in frequency in \( \text{Re} \, D(\omega) \) is via a non-analytic kink \(-|\omega| \) (no Taylor expansion at \( \omega = 0 \)), Fig. 5b.

We emphasize that the absence of linear-in-frequency term in the OG-measured diffusivity (Eq. (21)) for any regular or confined geometry is a universal consequence of the analytic behavior of \( D(\omega) \) in these cases. For instance, a naive dimensional estimate would give \(-\omega^2 / D \), with \( D \) in an impermeable pore of size \( \sim a \), similar to the PG estimate \(-\omega^2 / t \); however, this term enters \( D(\omega) \) with an imaginary \( i \) and therefore is not captured by the OG (in the limit of large number of oscillations), which is only sensitive to the second-order real term \(-D_0 : (\omega \delta_0)^2 - \omega^2 \delta_0^2 / D_0 \) where \( D_0 \) is the free diffusion coefficient and \( \delta_0 \) to \( \omega^2 \delta_0^2 / t \).

\[\text{Effect of finite PG pulse duration } \delta \gg t_e\]

\( D(\omega) \) is the fundamental characteristic using which one can calculate the effect of any sequence for the lowest (second) order cumulant of the dMRI signal (Novikov and Kiselev, 2011)

\[\text{−Ln}S = \frac{d}{dt} \int_{-\infty}^{\infty} D(\omega) \, q_\omega \, + \, C^i(g^2)\]

where \( q_\omega \) is the Fourier transform of the integral \( q(t) = \int dt \, g(t) \) of the applied Larmor frequency gradient \( g(t) \). The finite-pulse-width PG corresponds to \( \text{Callaghan, 1991} \)

\[q_\omega = \frac{g}{(i\omega)^2} (e^{i\omega \delta} - 1)(e^{i\omega \delta} - 1),\]

where \( g \) is the maximum gradient value, and \( \Delta \) is the diffusion time, defined as the interval between the front edges of the pulses (whenever it does not cause confusion, we denote this diffusion time \( \Delta \) or \( \Delta \) similar to the narrow-pulse case). The PG measurement becomes a low-pass filter \( |q_\omega|^2 \) which resolves the time-dependent tail (Eq.(18)) only for \( t = \Delta - 1 / \omega \) or \( \delta \), in which case \( q_\omega \approx (g\delta)(e^{i\omega \delta} - 1)(i\omega) \) coincides with that for the narrow-pulse limit. While the details of the calculation for the low-pass filter effect will be reported elsewhere (Lee et al., 2015), here we will announce the result for \( \delta = 1 \). The frequency integral in the first term of Eq. (22) is found using Eqs. (19) and (23) by deforming the integration contour in the complex plane of \( \omega \) from the real axis to the one passing along both edges of the branch cut of \( \ln \omega \), connecting \( \omega = 0 \) and \( \omega = \infty \) via the lower-half-plane, yielding

\[\text{−Ln}S = bD_a + (g\delta)^2 \cdot AF(\Delta / \delta) + C^i(g^2),\]

\[F(x) = \frac{x^2}{2} \ln \left( 1 + \frac{1}{x^2} \right) + \frac{1}{2} \ln \left( x^2 - 1 \right) + x \ln \frac{x + 1}{x - 1}\]

where the diffusion weighting parameter \( b = (g\delta)^2 (\Delta - \delta / 3) \). Note that \( F(1) = 2 \ln 2 \approx 1.38 \) and \( F(\infty) \approx \ln \pi + \frac{\delta}{2} \). The latter asymptotic behavior practically works already for \( x \ll 1 \). Eqs. (24) and (25) establish the behavior (Eq. (1)) for \( \delta \gg t_e \).

Physically, finite PG pulse width \( \delta \) allows the tail (Eq. (8)) to be manifest only at the time scales \( t \approx t_e \), otherwise \( \delta \). Therefore, finite \( \delta \) may mask the correlation time \( t_e \) and hence, the correlation length \( l_e \). However, since \( F(x)\rangle \approx x^2 \) and grows monotonically for \( x \geq 1 \), the time-dependent part \( -AF(\delta / x) / (x - \delta / 3) \) of (D) defined via \( S(t) = -bD(t) + C^i(g^2) \) t \( \Delta \), decreases slower than \( 1/t \), asymptotically approaching the \( \left[ \ln \left( \frac{t}{t_c} \right) \right] / t \) behavior (Eq. (1)). Hence, finite \( \delta \) does not change the overall (\ln)/t dependence (Eq. (1)), only affecting the higher-order corrections \( 1/t \) to the “ideal” narrow-pulse \( D(t) \).

Table 1 summarizes the results of the Theory section.

**Methods**

To demonstrate the effect of disorder on the two-dimensional diffusion in the long time limit, we employ an artificial model system mimicking extra-axonal space. By utilizing fibers with a mean diameter larger than axonal diameters by about tenfold, the required diffusion times are increased by about a factor of a hundred, from a few milliseconds to hundreds of ms, which are much easier to observe on a standard clinical MR system. To access such long diffusion times (of up to a second in our case), we use a stimulated echo sequence to reduce the echo attenuation due to \( T_2 \) decay, as we describe in more detail below.

**Phantom construction**

The diffusion phantom for this study was constructed using Dyneema® fibers held together with a flexible, polyolefin low temperature-shrinking tube (Fieremans, 2008; Fieremans et al., 2008). The fibers are 17 ± 2.6 μm in diameter and are comprised of an ultra-high-molecular-weight polyethylene (UHMWPE) which is ultrahydrophobic and impermeable to water. Approximately 195,000 fibers were placed in a shrinking tube measuring 8 cm long, and the fiber and tube bundles were placed in a hot water bath with temperature 90 °C for 10 min to shrink the tube. To facilitate the removal of air bubbles, which could contribute to susceptibility artifacts, the tubes were squeezed and shaken as they were shrunk in the hot water bath. The fiber bundle was suspended in the center of a 1.5 L plastic bottle using zip ties. The bottle was then filled with a distilled water solution of 0.09% w/v NaCl to reduce \( B_0^\perp \) field inhomogeneities.

**MRI measurements**

Imaging was performed at 16 °C on a 7 T Siemens clinical whole-body MRI scanner using a 28 channel knee coil. A high resolution (1 mm × 1 mm × 1 mm) MPRAGE image was first acquired to locate the presence of any air bubbles within the fiber bundle to determine where best to place the diffusion tensor imaging slices. DTI was carried out using the stimulated echo acquisition mode (STEAM) sequence which allows for long diffusion times while minimizing echo attenuation caused by \( T_2 \) relaxation. Twenty five measurements were performed at b values of 0 and 1 ms/μm² in 20 directions, each with a constant TE = 68 ms, δ = 17.5 ms, and a TM ranging from 10 ms to 980 ms, corresponding to diffusion times t = TE/2 + TM of 44 ms to 1014 ms. Two slices of resolution 2.2 mm × 2.2 mm × 10 mm were used. The fiber bundle was placed parallel to the \( B_0^\perp \) field to eliminate the possibility of susceptibility-induced internal field inhomogeneities produced by the fibers.
Monte Carlo simulations in the intra- and extra-axonal spaces

We generate three types of impermeable two-dimensional disk packings – one to simulate our fiber phantom, one to simulate a periodic packing, and another to simulate white matter. The disk packing mimicking the fiber phantom is comprised of \( N \approx 10,000 \) disks with a radius of \( 8.5 \pm 1.3 \mu m \) packed to a density of 75%. This density is chosen as it best matches the FA value of 0.54 found for the phantom for \( t > 700 \) ms. For the periodic packing we generate a lattice of disks with a radius 8.5 \( \mu m \) packed to a density of 72%; a higher density lattice is unachievable for our current simulation.

For the white matter, we generated three different packings based on the axon inner diameter distributions in sectors 2, 4, and 6 of the rhesus monkey corpus callosum determined by Lamantia and Rakic (1990), and packed to a density of 70% for all sectors, i.e. extra-axonal space occupies fraction \( \phi_{ext} = 0.3 \) of the volume. To determine these distributions, we first digitized the histograms provided in Lamantia and Rakic (1990) for the internal axonal diameters 2r. The corresponding external diameters \( 2r_{ext} \) were found by dividing \( 2r \) by the myelin g-ratio \( g_{m} = r/r_{ext} = 0.6 \) (Rushton, 1951) which we assumed to be the same for all simulated axons. We do not account for the possible shrinkage of the axons due to a conflicting evidence of the extent of this effect; our results can be readily rescaled to account for the shrinkage. These external diameters were then used for randomly packing \( N \) impermeable disks for the simulations, see Supplementary figs. S1–S3.

All packings (except for the periodic one) were generated using a collision-driven molecular dynamics simulation script written by Donev et al. (Donev et al., 2005b,c).

To simulate the narrow-pulse cumulative diffusion coefficient \( D(t) \) in the extra-axonal space, \( 1 \times 10^{6} \) tracers are placed randomly outside the disks, and allowed to perform 2D Brownian motion restricted by the impermeable disk walls, following the Monte Carlo dynamics previously described in detail (Fieremans et al., 2010). Briefly, during each time step \( dt \), a tracer steps a fixed distance \( dr \) given by \( \sqrt{2 \delta D_{0} dt} \) with \( \delta = 2 \) in a randomly chosen direction. For the phantom simulation, \( \delta = 7.5 \times 10^{-4} \) ms, and for the white matter simulation, \( \delta = 7.5 \times 10^{-5} \) ms, corresponding to step lengths of 0.0775 \( \mu m \) and 0.0245 \( \mu m \), respectively. If a tracer encounters a barrier, it will elastically reflect off the barrier. We verified our MC algorithm by simulating \( D(t) \) inside one impermeable cylinder, and obtained a perfect agreement with the exact result, see Fig. C.1 in Appendix C. Expanding upon this, we also simulate the intra-axonal \( D(t) \) within the impermeable cylinders of all 3 sectors assuming inner radii based on the reported values in Lamantia and Rakic (1990), which agrees with the asymptotic behavior (Eq. (27)) below. For all simulations, \( D_{0} = 2 \mu m^{2}/ms \).

The time-dependent diffusion coefficient in direction \( n \) is calculated via Eq. (3), with \( 6x(t) \in \mathbb{R} \cdot n \), where \( \mathbb{R}(t) \) is the cumulative displacement for a given tracer over the Monte Carlo time \( t \). In this paper, \( n \) is taken in the two orthogonal directions separately, and the resulting mean square displacements are further averaged between the two directions to yield the isotropic diffusion coefficient transverse to cylinders (disks).

We do not simulate the diffusion-weighted signal, as the effect of the finite gradient pulses can be derived from the narrow-pulse MC using Eqs. (24) and (25). This makes results more reliable, since simulating the signal with finite pulses involves adding oscillating phases, which notably reduces the precision.

Model for diffusion inside axons and for overall \( D(t) \)

For comparing the roles of extra- and intra-axonal spaces, we use a common assumption of effectively impermeable axonal walls (Assaf et al., 2008; Alexander et al., 2010; Zhang et al., 2011) which is reasonable given thick myelin layers covering axons within major fiber tracts. Hence, the total time-dependent diffusion coefficient (Eq. (3)), defined via the second cumulant of the total signal (Appendix E), is given by the weighted average of the diffusion coefficients of the intra- and extra-axonal compartments (Fieremans et al., 2010):

\[
D(t) = \frac{f_{int}}{D_{int}(t)} + \frac{f_{ext}}{D_{ext}(t)},
\]

where \( f_{int} + f_{ext} = 1 \) are the fractions of intra- and extra-axonal water, and \( D_{ext}(t) \) is simulated as described above. We assume that water within myelin is effectively invisible in diffusion due to its short \( T_{2} \).

We model diffusion in the intra-axonal space in the narrow-pulse limit using the impermeable cylinder model (Stepišnik, 1993; Callaghan, 1995). As it follows from our analysis of the exact solution in Appendix C, for the relevant diffusion times, within each cylinder with internal radius \( a \) the diffusion coefficient is very well approximated by its asymptotic limit \( D(t) = a^{2}/4t \), Eq. (C.4) and Fig. C.1. We obtain \( D_{int}(t) \) from impermeable cylinders according to the distribution of axonal sizes from Lamantia and Rakic (1990). As a result, the diffusion coefficient from inside the axons entering Eq. (26) is found as

\[
D_{int}(t) = \frac{\langle r^2 \rangle_{i} }{4t} \phi_{int} = \frac{N}{\phi_{int}} \phi_{int} = \frac{N r_{int}^2}{V}.
\]

In terms of the moments \( \langle r^2 \rangle \) of the axonal size distribution,

\[
D_{int}(t) = \frac{\langle r^2 \rangle _{i} }{4t}.
\]

With a fixed value of the g-ratio, the fractions \( \phi_{ext}, \phi_{int} \) and \( (1/g^2 - 1)\phi_{int} \) are the volume fractions of extra-axonal water, intra-axonal water and myelin, correspondingly, such that \( \phi_{ext} + \phi_{int}/g^2 = 1 \). The water fractions entering Eq. (26) are

\[
f_{ext} = \frac{\phi_{ext}}{\phi_{int} + \phi_{ext}}, \quad f_{int} = \frac{\phi_{int}}{\phi_{int} + \phi_{ext}}.
\]

Hence, for sectors 2, 4, 6, \( \phi_{ext} = 0.3, \phi_{int} = 0.7 \cdot 0.5^2 = 0.252 \), so that \( f_{ext} = 0.54, f_{int} = 0.46 \) in Fig. 11.

Results

Identifying the logarithmic singularity, Eq. (1)

The eigenvalues from the experimentally measured diffusion tensor in our fiber phantom are plotted in Fig. 6a, with the principle eigenvalue, \( \lambda_{1} \), shown by the circles, and the transverse eigenvalues, \( \lambda_{2} \) and \( \lambda_{3} \), shown by the triangles and squares, respectively. Since the direction of \( \lambda_{1} \) is aligned parallel to the fibers for all \( t \), it corresponds to practically unrestricted diffusion, as seen in the figure. However, as \( \lambda_{2} \) and \( \lambda_{3} \) are measured transverse to the bundle, they indeed decrease with \( t \) due to geometric restrictions by impermeable fibers. Note that differences between \( \lambda_{2} \) and \( \lambda_{3} \) are probably the result of the eigenvalue repulsion due to noise. The larger errorbars associated with these two eigenvalues likely arise due to the Gibbs ringing present in transverse direction.

We see similar behavior in our simulation data as indicated by the solid black line. To better represent the experimental conditions, which were at 16 °C rather than 20 °C, we applied a rescaling factor of 1.8/2.0 = 0.9 to \( D(t) \) and 1/0.9 to \( t \) based on the free \( D_{0} = 1.8 \mu m^{2}/ms \) for water at 16°C (Tofts et al., 2000).

To study the nature of the time dependence, the transverse eigenvalues are averaged together to obtain the isotropic two-dimensional component of the transverse diffusivity

\[
D(t) = \frac{\lambda_{2} + \lambda_{3}}{2}.
\]

Taking the average (30) significantly reduces bias due to noise. Indeed, the effect of eigenvalue repulsion is exactly compensated by
taking the mean of eigenvalues in a 2 × 2 tensor; the residual repulsion in our 3 × 3 case coming from the presence of \( \lambda_3 \) is small, since \( \lambda_2 \ll \lambda_3 \ll \lambda_2,3 \). We use the narrow-pulse PG notation (Eq. (3)) in Eq. (30) since we are studying \( t \sim 100 - 1000 \) ms, orders of magnitude longer than \( \delta = 17.5 \) ms.

The approach of \( D_\perp(t) \) towards \( D_\perp \) is shown in Fig. 6b, where \( D_\perp \) is determined via a fit to Eq. (1). The fit parameters \( D_\perp, A, \) and \( t_\perp \) are shown in the “Experiment” column of Table 2. We also observe that the value \( t_\perp \) is close to \( \delta \), so while the agreement with \( t_\perp \) from MC simulations is very good, this fit parameter value might be rather indicating the “filter effect” of a finite-\( \delta \) PG, besides, it is generally hard to obtain reliable fit results for quantities under the logarithm.

Even on a qualitative level, it is clear that the slope of the experimental curve is not constant. Hence, the time-dependence clearly deviates from \( 1/t \), as the fit to \( D_\perp + c/t \) shows. This fit also looks particularly poor due to the double-logarithmic axes exaggerating the small deviations the fit has from the data. The fit to Eq. (1) shown by the black dashed line is considerably better (the \( \hat{R}^2 \) value increases from 0.82 to 0.95). This is the first indication of the anticipated logarithmic singularity (Eq. (1)). The simulation data (blue dotted line) shows a trend similar to the experimental data, with the parameters summarized in the “MC phantom” column of Table 2.

Another way to see the effect of the \( \ln(t) \) dependence is presented in Fig. 7. Here, the transverse component (Eq. (30)) is plotted with respect to \( 1/t \) (Figs. 7a, c) and with respect to \( \ln(t/t_c) \) (Figs. 7b, d). The top row shows Monte Carlo simulated fiber phantom diffusion data, and the bottom row is the experimental data. As is evident from Fig. 7a, the slope of the data keeps changing with respect to \( 1/t \), indicating that the behavior is slower than \( 1/t \). However, when the data is plotted with respect to \( \ln(t/t_c) \) as in Fig. 7b, this curve straightens out, indicating the logarithmic singularity. We also see this trend in our experimental data as shown in Figs. 7c and d.

In contrast, Fig. 8a shows the \( D(t) \) from the exterior of the simulated square lattice of disks, exhibiting a \( 1/t \) dependence as expected from Eq. (13).

**Identifying the dynamical exponent \( \vartheta \)**

Alternatively, we can determine the instantaneous diffusion coefficient \( D_{\text{inst}}(t) \) using Eq. (4) from both experimental and simulated data, and confirm the value \( \vartheta = 1 \) of the dynamical exponent, which is equivalent to the singular behavior (Eq. (1)). To do this, we use a Savitzky–Golay (Savitzky and Golay, 1964) second order polynomial smoothing procedure written in house using Matlab (Mathworks, Natick, MA) to smooth and differentiate our data.

The results of this procedure are shown in Figs. 8 and 9.

In Figs. 8a and b, \( D(t) \) and the derived \( D_{\text{inst}}(t) \) is shown for the Monte Carlo simulated data for the square lattice disk pack. The semilog scale in Fig. 8b indicates that, in contrast to the dispersed samples, \( D_{\text{inst}}(t) \) approaches its \( D_\perp \) value in the –\( \exp(-t/\xi) \) fashion, as expected from Eq. (9), corresponding to the dynamical exponent \( \vartheta = \infty \). A fit of this data (shown by the black line) reveals \( \xi_0 \approx 3.1 \) ms, which is somewhat small compared to the diffusion time of \( \approx 20 \) ms across the lattice constant. This smallness is similar to what is observed in the \( d = 1 \) lattice (Novikov et al., 2014).

Fig. 8c contrasts \( D_{\text{inst}}(t) \) for both the periodic and random geometries, plotted with respect to \( 1/t \). Note that \( D_{\text{inst}}(t) \) for both geometries have markedly different approaches to their corresponding \( D_\perp \) values, indicating the strong sensitivity of \( D_{\text{inst}}(t) \) to order/disorder at long times.

In Fig. 9, to emphasize the \( r^{-1} \) behavior, we plot \( D_{\text{inst}}(t) \) as function of \( 1/t \) for our experimental and simulated data, respectively. In Panel (a), different symbol shapes indicate different Savitzky–Golay smoothing window sizes (see legend). We note that at long times (i.e. small \( 1/t \)), no matter how much the data is smoothed, there is an asymptotically linear dependence on \( 1/t \) (indicated by the solid black line), showing that \( D_{\text{inst}}(t) \) approaches its \( D_\perp \), limit as \( t \to \infty \), in agreement with the exponent \( \vartheta = 1 \) in Eq. (5), expected in a 2d random packing geometry.

In Fig. 9b, the same \( D_{\text{inst}}(t) \) as in Fig. 8c is plotted for the long \( t \) (small \( r^{-1} \)). For MC data, we use adaptive Savitzky–Golay window growing with \( t \) from 21 to 2001 MC time points.

While it may seem that the time dependence of the blue curve in Fig. 6c is far from \( r^{-1} \), this is only because the shortest times are overemphasized due to the \( r^{-1} \) inversion of the time axis. The asymptotically linear dependence on \( r^{-1} \) becomes evident in Fig. 9b, which displays about 90% of all our MC time points, corresponding to 100 ms \( \leq t \leq 1000 \) ms.

What does this mean for diffusion transverse to WM fibers? Since diffusion time scales as length squared, and given that axonal diameters are about 10 times smaller than those of the Dyneema fibers, Fig. 9...
suggests that the long-time $1/t$ tail in $D_{\text{inst}}(t)$ and, thereby, the logarithmically enhanced behavior (Eq. (1)) would manifest themselves in WM as long as the diffusion time is about 100 times smaller than its onset of $\approx 100 \text{ ms}$ in Fig. 9, i.e. for $t \gtrsim 1 \text{ ms}$ in vivo, cf. Table 2. This, for all practical purposes, covers all MRI-observable diffusion times. From Fig. 9, we also note that the dynamical exponent $\nu = 1$ becomes manifest when $D_{\text{inst}}(t)$ is within roughly 10% of $D_\infty$ for both experiment (a) and simulation (b). Our MC simulations for the realistic axonal geometries discussed below and shown in Fig. 11 are in agreement with the above estimates.

---

Fig. 7. Identifying the logarithmic singularity (1) in the cumulative diffusion coefficient $D(t)$, cf. Eq. (3), for the randomly packed fibers. a, MC-simulated $D(t)$ plotted as function of $1/t$ for diffusion in the exterior of a random disk packing (fragment shown in Fig. 3b). The logarithmic singularity is manifest in the “bend”. b, This “bend” straightens up when $D(t)$ is replotted with respect to $\ln(t/t_c)/t$. c, d: The same as for a and b, but with experimental $D_{\perp}(t)$, Eq. (30) and Fig. 6, transverse to the fiber phantom.

Fig. 8. Effects of the periodic geometry. a, $D(t)$ for MC simulated diffusion in the exterior of periodically ordered disks. Note the absence of the “bend” (cf. Fig. 7), illustrating that the logarithmic singularity is absent in an ordered medium. b, $D_{\text{inst}}(t)$ derived from the same simulated data, and plotted on a semilog plot, indicating an exponentially fast approach of $D_\infty$. A fit to $c t^{-1} e^{-t/t_c}$ is shown by the black line. c, The qualitatively different approach to $D_\infty$ of MC-derived $D_{\text{inst}}(t)$ for the periodic (green) and randomly packed disks (blue). While $D_{\text{inst}}(t)$ for the regular lattice of disks reaches its $D_\infty$ very rapidly (within about 10 ms), the blue curve (mimicking a randomly packed fiber phantom) slowly straightens out, as $t \to 0$, cf. Fig. 5b for the magnified time domain 100 ms $\leq t \leq 1000$ ms.
Dominant role of extracellular water in OG diffusion transverse to neuronal fibers

Let us now turn to the recent OG experiment (Portnoy et al., 2013), whose data we replot in Fig. 10. In this figure, the frequency dependent diffusion coefficient was measured inside an ex vivo rat hippocampus with 10 cosine modulated diffusion gradients ranging from 67 Hz to 1 kHz on a 600 MHz Bruker spectrometer. The data shown is from the stratum radiatum section of the hippocampus transverse to the fibers. We can clearly see the linear dependence (Eq. (2)) at low ω (as long as Re D(ω) is within about 20% of D∞), corresponding to the dynamical exponent p = 1. From that, it follows that the structural exponent θ = 1. Therefore, at low ω, due to the structural disorder, the dispersive contribution to D(ω) of the hindered component wins over that of the fully restricted component in a 2d geometry. In other words, the contribution of the confined water to the overall D(ω), and with that, any information about the inner fiber diameters, is practically unobservable with diffusion measured using OG transverse to the fibers in the small ω limit.

We note that extra-axonal D(ω) also has regular contributions, such as ω2, Eq. (21), much like the intra-axonal one, Eq. (D.4). Hence, if one further increases ω to “catch” the ω2 term on the background of |ω|, it does not help to quantify internal axonal diameters without knowing which part of this term comes from intra-axonal water.

In Appendix F we argue that the |ω| behavior in D(ω) is determined by the random packing geometry irrespective of the exchange rate τex between intra- and extra-axonal compartments, and that OG is thus asymptotically insensitive to internal axonal diameters at low ω even when exchange is present.

Dominate role of extra-axonal water in narrow-pulse PG diffusion transverse to axons

Consider first the common assumption of negligible exchange, which is likely to hold for major fiber tracts. In this limit, the overall diffusion coefficient is given by Eq. (26). We now show, similar to the frequency case (Eq. (31)), that the extra-axonal contribution is dominant at long times. Here we focus on the narrow-pulse PG limit in order to understand the ideal situation. In Discussion section, we will use these results to argue that for finite p, the extra-axonal contribution becomes relatively more significant and biases axonal diameter results, Fig. 2.

The extra-axonal (ln t)/t contribution (Eq. (1)) to the overall narrow-pulse diffusion coefficient (Eq. (26)) should become gradually more...
pronounced, due to its slower decrease with time, than the $1/t$ contribution of the confined intra-axonal water, Eq. (28), since

$$\tilde{c}_1 \frac{1}{T} + \tilde{c}_2 \frac{\ln t}{T} \approx \tilde{c}_2 \frac{\ln t}{T}, \quad t \gg t_c. \tag{32}$$

The extra-axonal dominance should be less dramatic for PG than for OG, Eq. (31), as $\ln t$ is a relatively slowly growing function.

How significant is this effect for realistic axonal packings? To address this question quantitatively, we calculate the time dependence of the diffusion coefficient (Eq. (26))

$$D(t) - D_\infty = f_{\text{int}} D_{\text{int}}(t) + f_{\text{ext}} \left( D_{\text{ext}}(t) - D_\infty^{\text{ext}} \right). \tag{33}$$

where the overall $D_\infty = f_{\text{ext}} D_\infty^{\text{ext}}$. In Fig. 11, we show the results of Monte Carlo simulations for $D_{\text{ext}}(t)$ for the three cylindrical (disk) packings based on sectors 2, 4, and 6 of the rhesus monkey corpus callosum, as well as analytical and numerical results for $D_{\text{int}}(t)$ (see Methods section). The left column of Fig. 11 shows small fragments of the cylinder packs used in the simulation with the corresponding histograms of the radii (Lamantia and Rakic, 1990) in the center column (for full packings, see Supplementary fgs. S1–S3). The main results are shown in the right column, where we plot the first and the second terms of Eq. (33) along with the total $D(t) - D_\infty$.

In all three cases, the extra-axonal contribution dominates the overall time dependence at practically relevant times ($t \geq 1$ ms). The long time limit begins at approximately 1 ms, meaning that this dominance is always present, as already anticipated from the results of Fig. 9.

At the experimentally feasible time scales of $t = 10$ and 100 ms, the $t$-dependent part of $D_{\text{ext}}(t)$ comprises over 80% of the total $D(t)$ for all sectors as seen in Table 3, and logarithmically grows with $t$.

In other words, similar to the OG case, the intra-axonal compartment becomes asymptotically invisible on the background of the extra-axonal contribution.

![Fig. 11. Simulated data for sectors 2, 4 and 6 (rows a, b and c) of the rhesus monkey corpus callosum based on data from Lamantia and Rakic (1990). Left column: Sample sections of the disk packs used in the MC simulation. For full packs, see Supplementary fgs. S1–S3. Intra-axonal areas are indicated by the red pixels, with myelin and extra-axonal areas shown in green and blue, respectively. Center column: histograms of the internal radii used in each packing. Right column: the intra- and extra-axonal components of the long time diffusion coefficient. The total $D(t) - D_\infty$, Eq. (33), is given by the black solid line, the weighted intra-axonal component of the simulation, $f_{\text{int}} D_{\text{int}}(t)$, is given by the red dashed line, the weighted intra-axonal component calculated from Eq. (28) is given by the cyan dotted line, and the weighted extra-axonal component, $f_{\text{ext}} D_{\text{ext}}(t)$, from the Monte Carlo simulation (see Methods) is given by the blue dot dashed line.](image-url)
The results of fitting of $D_{\text{ext}}(t)$ to Eq. (1) are presented in the last three columns of Table 2. As the three-parameter fitting can become unstable, especially for $t_c$ entering under the weakly-varying logarithm, we first derive the corresponding instantaneous diffusion coefficient (Eq. (8)) using the Savitzky–Golay procedure outlined above, which allows us to eliminate $t_c$ (Supplementary fig. S5); after $D_a$ and $A$ are fit, they are fixed in the fitting of $D_{\text{ext}}(t)$ to Eq. (1) in order to obtain $t_c$. We observe that the correlation length $l_\text{c} \approx (r_{\text{app}})$ does not parametrically exceed the mean external axonal radius. Its values are also consistent with the estimates made via the extent of the plateau of $T$ in Fig. 4. This indicates that local order due to a mutual “repulsion” of disks in all our packings disappears already at the level of nearest neighbors. (Presumably, packing at a higher density may lead to emergence of short range order, $l_\text{c} \gg (r_{\text{app}})$.) We also note an approximate scaling $A \propto b^2$, Supplementary fig. S4, expected from the dimensional considerations: At $t \gg t_c$, the $t$-independent length scale that the system still “remembers” after the coarse-graining is the correlation length $l_\text{c} = \sqrt{2ddt_c}$, and $A$ has dimensions of $(\text{length})^2$. This scaling is only approximate since the plateau $T(k)$ is governed not only by $b^2$ but by the packing details, see Fig. 4 and an estimate at the end of Appendix A.

If exchange is present, our argument in Appendix F translated to the time domain suggests that the $(\ln t)/t$ term due to extra-axonal space geometry will similarly prevail for the slow exchange $\tau_{\text{ex}} > t_c$ and will apply to all water molecules in the fast exchange limit $\tau_{\text{ex}} < t_c$.

Discussion: what do we measure with dMRI transverse to axonal fibers?

Quantifying $\mu$m-level structure of neuronal tracts in vivo has been brought to the forefront of neuroscience research primarily due to the axonal diameter mapping (ADM) concept, developed within AxCaliber framework (Assaf et al., 2008; Barazany et al., 2009) and their extensions (Alexander et al., 2010; Zhang et al., 2011). Their common theme is the focus on the intra-axonal compartment (assuming no exchange), which is typically done by approximating axons as impermeable cylinders, and building on exact solutions (Neuman, 1974; Stepišnik, 1993; Callaghan, 1995). Water diffusion in the extra-axonal space in all of the above approaches is assumed to be Gaussian (time-independent).

Our findings of the unexpectedly dominant role of the time-dependence of dMRI signal from extra-axonal water prompt us to reconsider the applicability and interpretation of ADM.

ADM in the narrow-pulse limit $\delta < t_c$

In Appendix E, by attributing the time-dependencies of both $D_{\text{int}}(t)$ and $D_{\text{ext}}(t)$ in Eq. (33) to the intra-axonal water, i.e. assuming Gaussian extra-axonal diffusion, up to order $O(q^2)$ in the cumulant expansion we obtain an apparent internal axonal radius $r_{\text{app}}(t)$ growing with diffusion time $t$:

$$r_{\text{app}}^2(t)|_{\delta < t_c} \approx \frac{\alpha^2}{\alpha^2} + \frac{4A}{f_{\text{int}}} \int_0^t \ln \frac{t}{t_c}. \tag{34}$$

While the first term in Eq. (34) reflects the volume-weighted contribution of water inside all axons, cf. Eq. (28), the second term comes from the singular time-dependent contribution (1) of the extra-axonal water. It results in the overestimation of the value of the axonal radius. This bias has been illustrated in Fig. 2 of Introduction, where $r_{\text{app}}(t)$ (dashed lines) is calculated from Eq. (34), using $A$ and $t_c$ determined from Fig. 11, Table 2. Based on these parameters, the second term in Eq. (34) exceeds not only $r^2$ but even $r^4/r^2$, so that the apparent internal axonal radii could be significantly biased by the extra-axonal contribution. Indeed, as seen from the second row of Table 3, the calculated value of $r_{\text{app}}$ is roughly 3–4 times greater than $r$ for all three corpus callosum sectors.

Furthermore, our analysis in Appendix E shows that including the $O(q^2)$ terms leads to the further increase of the apparent $r_{\text{app}}(t)$, making it grow even faster with time, as $(\ln t)^{1/4}$, Eq. (E9), which occurs due to the corresponding $\ln(t)/t$ behavior of the Kurtosis $K_{\text{ext}}(t)$ in the extra-axonal space, Eq. (E4). While being numerically too small to affect fits of quantities such as axonal (or neurite) water fraction especially at the clinical times $t \sim 10 \text{–} 100 \text{~ms}$ (Assaf et al., 2008; Jespersen et al., 2007, 2010; Fieremans et al., 2010), $K_{\text{ext}}(t)$ generally cannot be neglected for more subtle techniques such as axonal diameter mapping, since all the (time-dependent) corrections originating from finite axonal diameters are comparably small, cf. Fig. 11.

To get a sense when Eq. (34) applies, let us employ dimensionless time and Larmor frequency gradient variables

$$\bar{t} = \frac{t}{t_c} = \frac{\Delta}{t_c} \quad \bar{\delta} = \frac{\delta}{t_c} \quad \bar{g} = \frac{g}{g_r} \quad \bar{g}_r = \frac{\gamma_0}{g_r} \frac{D_0}{\tau_c} \frac{1}{\Gamma_c} \tag{35}$$

where the disorder correlation time scale $t_c \sim r^2/D_a - \tau_D = r^2/D_a - 0.1 \sim 1$ ms is practically similar to the diffusion time $\tau_D$ within axon with typical radius $r$ (Table 2), and $g_r$ is the characteristic scale for the Larmor frequency gradient needed to efficiently probe distances $r \sim 1$ ms. Using the variables (Eq. (35)), we make order-of-magnitude estimates by dropping all coefficients $O(1)$ in Appendix E:

$$-\ln \bar{S}_{\text{int}} \sim \alpha^2 + O\left(\alpha^4\right) \quad -\ln \bar{S}_{\text{ext}} \sim \alpha^2 \left(\bar{t} + A \ln \bar{t}\right) + O\left(\alpha^4\right) \tag{36}$$

where $\bar{A} = A/r^2 - 0.1 \sim 1$, and the dephasing strength $\alpha = qr = gr_0 \equiv g_0$ is usually small, as diffusion gradients are generally weak relative to the scale $g_0$. Indeed, the scale $g_r \sim 20(\mu m \cdot ms)^{-1}$ for $r = 1 \mu m$, assuming free axoplasm diffusivity $D_0 = 2.4 \mu m^2/\text{ms}$ (Beaulieu, 2002), whereas $g = \gamma - 40 \text{mT/cm} = 0.0107(\mu m \cdot ms)^{-1}$ for a typical clinical scanner, and $g \sim 0.1(\mu m \cdot ms)^{-1}$ for animal systems. This makes $g \sim 10^{-2} \sim 10^{-3}$; assuming $\bar{\delta} < t_c$, makes $\alpha \ll 1$.

We can see that in the $\bar{\delta} = 1$ limit considered in Eq. (36), the three dimensionless quantities $I, \bar{g}$, and $\bar{g}_r$ enter only in the two combinations: $\alpha$ and $I$. Hence, the phase diagram in this limit is two-dimensional, and is schematically outlined in the left panel of Fig. 12. The extra-axonal signal is effectively suppressed when $\alpha^2 \bar{I} > 1$, i.e. above the green line \(\alpha \sim \bar{I}^{-1/2}\). As $\alpha \ll 1$, the intra-axonal signal is practically never small (we are always below the red line). The ADM applies in the green shaded region where the extra-axonal contribution can be neglected, while the intra-axonal signal is not too small. However, below the green line (regime I), when both $S_{\text{int}}$ and $S_{\text{ext}}$ are not exponentially suppressed, the above results of Appendix E apply, and Eq. (34) becomes asymptotically exact for $\alpha \ll \bar{I}^{-1/2}$. While it is advantageous for the ADM to increase the diffusion time $\bar{t}$, the right edge of the phase diagram is limited by $T_2$, so that $t \leq 100 \text{~ms}$, $\bar{t} \leq 100^2$, and so for $\bar{\delta} = 1$, we are practically always in the regime I, $\alpha \ll \bar{I}^{-1/2}$, where Eq. (34) asymptotically applies.

ADM on clinical scanners, $\delta \gg t_c$

Clinical acquisition employs wide pulses, with duration $\delta \geq 10$ ms exceeding both the diffusion time $\tau_D = r^2/D_a$ across the inner radius of a typical axon, and the correlation time $t_c$ in the extra-axonal space. In this limit, the above analysis gets modified. It turns out that diffusion attenuation (at a given $b$-value) becomes generally weaker in both intra- and extra-axonal space. However, while extra-axonal contribution changes only logarithmically, $\ln(t/t_c) \sim \ln(\Delta/\delta)$, cf. Eq. (24), the intra-axonal attenuation becomes smaller by the factor $t_{\text{app}}/\delta \ll 1$. Physically, this ineffectiveness of diffusion gradients stems from the
To get an idea of just how insensitive clinical systems are to typical axonal sizes, we find the attenuation \(\text{Neuman, 1974}\)

\[
\ln S_{\text{ext}} \mid_{g \to g_0} \approx \frac{g}{96 D_0} \cdot 2 \delta \approx 2.17 \cdot 10^{-5}
\]

where we used \(g = 0.0107 \, (\mu m \cdot ms)^{-1}\) corresponding to 40 mT/m gradient, \(D_0 = 2.4 \, \mu m^2/\mu s\), \(\delta = 50 \, ms\), and a typical internal axonal diameter \(2r = 1 \, \mu m\). Even with the \(r^4\) scaling, the very large axons with \(2r = 4 \, \mu m\) give \(\ln S_{\text{ext}} \approx 0.0056\), i.e. less than 1% attenuation. In human corpus callosum, fibers larger than 3 \(\mu m\) made no more than 1% of the total fibers (Aboitiz et al, 1992; Innocenti et al., in press).

For dimensionless estimates, the variables (35) are \(\tilde{g} \sim 10^{-3}\); \(\tilde{T} \equiv T/t_c \sim 10^{-2} - 10^{2}; \tilde{\delta} \sim 10 - 10^2\). The contributions to signal \(S = S_{\text{int}} + S_{\text{ext}}\) are Eqs. (37) and (24)–(25) respectively,

\[
\ln S_{\text{int}} \sim \tilde{g}^2, \quad -\ln S_{\text{ext}} \sim \tilde{g}^{2-2} \tilde{T} + \tilde{g}^{2-2} A F(\tilde{T}/\tilde{\delta}).
\]

The full phase diagram is now 3-dimensional. To simplify it and to gain further intuition, we assume \(T \sim \tilde{t}\), i.e. pulses are as wide as possible, \(\delta - \Delta\), which is often the case. This leaves \(\tilde{g}\) and \(\tilde{T}\) as variables, cf. right panel of Fig. 12. The extra-axonal signal is suppressed and ADM applies above the green line, scaling as \(\tilde{g} \sim T^{-3/2}\). Below this line (regime II), both \(S_{\text{int}}\) and \(S_{\text{ext}}\) are not suppressed, and we can again use their cumulant expansions to estimate the role of \(t\)-dependence of \(S_{\text{ext}}\) in the ADM.

Restoring numerical coefficients in Eq. (38), expanding up to \(O(\tilde{g}^3)\) similar to Appendix E, and assigning all \(t\)-dependence to \(-\ln S_{\text{int}} \approx \tilde{t}^4_{\text{app}}(\Delta, \delta)\), cf. Eq. (37), as it would effectively occur in a fit to ADM models, we obtain in regime II

\[
\tilde{t}^4_{\text{app}}(\Delta, \delta) \mid_{\tilde{t} \gg \tilde{t}_c} \approx \frac{\langle \tilde{t}^4 \rangle}{\langle \tilde{t}^2 \rangle} + \frac{48 f_{\text{ext}}}{\tilde{T} f_{\text{int}}} \cdot D_0 \delta \cdot AF(\Delta/\delta).
\]

The apparent axonal radius according to Eq. (39) is shown in Fig. 2 (solid lines). It exceeds the one from Eq. (34) due to the relative smallness of the intra-axonal contribution (Eq. (37)) in the \(\delta \gg t_c\) limit. For the corpus callosum parameters determined from MC, Table 2, the intra-axonal term \(\langle\tilde{t}^6\rangle/\langle\tilde{t}^4\rangle\) of Eq. (39) gives a less than 1% contribution to the right-hand side. Therefore, the apparent radii \(\tilde{r}_{\text{app}}\) found in regime II of Fig. 12, third line of Table 3, exceeding the ones from histology (Lamantia and Rakic, 1990) by almost an order of magnitude, are completely determined by the extra-axonal geometry, embodied in the parameter \(A\), Eq. (A.4), and have little to do with intra-axonal diameters. This observation makes studying the effect of spatial correlations of axonal packing on the dMRI signal highly relevant. It is also likely to explain the noted overestimation of axonal diameters (Alexander et al., 2010; Zhang et al., 2011), in ADM at clinical gradient strengths, when the overall signal is mainly determined by its few lowest-order cumulants (regime II). The cumulant expansion performs within its convergence radius, \(\delta < b_t\) (Kiselev, 2010), which can be roughly estimated as \(b_t \approx 3 \, ms/\mu m^2\) for the white matter.

In this picture, the main contribution to Eq. (39) coming from the term \(A\) growing with the external radii (Table 2), can rationalize correlation between \(r_{\text{app}}\) and neuronal signal propagation speed (amount of myelin) (Horowitz et al., 2014). The observed decrease of \(r_{\text{app}}\) with increasing gradients (Huang et al., 2015) is consistent with moving upwards from regime II towards the green domain in Fig. 12. Dedicated suppression of extra-axonal contribution (Dhital et al., 2014) requires strong gradients, or the stimulated echo weighting. OG measurements, much like PG, strongly benefit from larger axons, such as in spinal cord (Xu et al., 2014). For OG, it is essential to reach high enough \(\omega\) such that \(\mathcal{D}(\omega)\) saturates, and ideally to combine this with axonal water fraction estimation (at high \(b\) in the \(t \to \infty\) limit, in order to be able to resolve intra- and extra-axonal contributions in the view of the remark after Eq. (31).

Mesoscopic effects: a novel kind of dMRI contrast

The unexpectedly strong contribution of the extra-axonal water to the overall time-dependent diffusion cumulants such as diffusion coefficient and kurtosis, and their sensitivity to the fiber packing embodied in the structure correlation function \(\Gamma(k)\), opens up a possibility to sense pathologies affecting the extra-axonal space geometry, such as demyelination and axonal loss, as well as possibly signal propagation speed, with diffusion.

Overall, our work brings to the forefront the mesoscopic effects — i.e. effects of spatial arrangement, correlations and structural disorder — in dMRI. Embracing the complexity of cellular architecture at the mesoscopic scale, intermediate between the molecular level (where local NMR parameters originate) and the macroscopic MRI resolution, offers a variety of physical effects relevant to quantifying tissue architecture. Mesoscopic transport, an area of condensed matter physics investigating the role of disorder in transport properties of complex systems, has enjoyed over a half-century of development motivated primarily by the electron conductivity (Anderson, 1958; Abrahams et al., 1979; Bouchaud and Georges, 1990; Kamenev, 2011; Mott and Davies, 1971; Shklovskii and Efros, 1984). This spurred the development of
sophisticated methods and physical intuition which extends far beyond semiconductor physics. The rapidly improving quality of dMRI acquisition is now enabling us to investigate the mesoscopic effects in vivo (Novikov et al., 2014). This novel kind of sensitivity to structure requires going beyond the conventional modeling paradigm of Gaussian compartments (with or without exchange), such as the Kärger model (Fieremans et al., 2010) and its multieponential anisotropic extensions (Panagiotaki et al., 2012). The focus of the mesoscopic approach is on the averaging of the diffusion propagator over the disorder in positions of restrictions at the mesoscopic scale. Remarkably, methods developed previously in a different physical context are now becoming relevant to neuroscience and to noninvasive imaging, with an exciting prospect of identifying highly specific markers of pathology in neuronal tissue.

The mesoscopic effects may also affect the intra-axonal contribution to the overall $D(t)$ and to $f_{ans}$, cf. Eqs. (34) and (39), practically relevant for ADM on systems with high gradients (e.g. animal dMRI). The tail of the axonal diameter distribution strongly affects the moments $\langle r^4 \rangle$ and $\langle r^6 \rangle$, giving a large weight to a relatively small number of axons with largest diameters. Hence, these contributions are susceptible to the mesoscopic fluctuations governed by the tail of the distribution. This general phenomenon of rare disorder configurations determining the measurement outcome has parallels with similar effects found in hopping conduction in disordered semiconductors, kinetics of reaction-diffusion systems, and other phenomena in disordered media (Lifshitz, 1964; Mott and Davies, 1971; Shklovskii and Efros, 1984). In our case, an accidently large number of thicker axons, or even a few very thick axons, may significantly skew the overall $D_{an}(t)$ for a particular voxel. This could lead to strongly enhanced variations (relative to those expected based on the measurement noise alone) in the corresponding parametric maps. Addressing how the mesoscopic fluctuations affect the measurement, how representative is a $\langle r^4 \rangle$ voxel in a region of interest, how reliably we know the axonal diameter distribution strongly affects the moments $\langle r^4 \rangle$ and $\langle r^6 \rangle$, where the structural exponent $\theta = 1$ in the extra-neurite space will persist, and $\langle r^6 \rangle$ in the extra-neurite water characterized by the dynamical exponent $\omega = 1/2$ dominating $|\ln(t)/t|^{1/2}$ terms as a result of disordered arrangement of axons in a bundle, again amplifying the role of extra-axonal water.

Conclusions

We showed that the long time behavior of diffusion coefficient transverse to a fiber bundle is determined by the disordered packing of the fibers. We identified short range disorder in the fibrous geometry in brain white matter, which may be a common feature of packings of biological building blocks.

As a consequence of the structural disorder, we predicted a characteristically singular slow $|\ln(t)/t|^{1/2}$ decrease of the diffusion coefficient transverse to the axons, Eq. (1). The logarithmic singularity (Eq. (1)) coming from extra-axonal space is shown to be more relevant than the $1/t$ decrease from water confined within axons, as $|\ln(t)/t|^{1/2}$ eventually takes over the $1/t$ decrease. This may explain the strong bias in the values of apparent axonal diameters, especially on human scanners.

The unanticipated effect the random spatial arrangement of axons has on dMRI measurements, leads us to conclude that the time-dependent dMRI signal acquired transverse to a fiber tract at clinical diffusion weightings may be more sensitive to the extra- than to intra-axonal water, as well as to how randomly fibers are arranged within a bundle. This observation points at a crucial role of the mesoscopic effects which can be quantified using methods of condensed matter physics outlined here. Understanding this rich physics may open up exciting prospects of designing new imaging biomarkers of neuronal integrity based on changes of the geometry of extra-axonal space, such as axonal loss and demyelination, as a result of disease, aging, or normal development.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.03.061.

Acknowledgments

It is a pleasure to thank Sharon Portnoy and Greg Stanisz for sharing with us their experimental data published in Portnoy et al. (2013) and replotted here in Fig. 10. We also thank Fengxia Liang from the NVU histology core for providing us with the mouse histology image in Fig. 3e, and Valerij Kiselev for his helpful comments on the manuscript. This work was supported by the Fellowship from Raymond and Beverly Sackler Laboratories for Convergence of Physical, Engineering and Biomedical Sciences, by the Litwin Foundation for Alzheimer’s Research, and by the National Institute of Neurological Disorders and Stroke of the NIH under award number R01NS088040.

Appendix A. Relation of the diffusion coefficient to the structure correlation function

Knowledge of the structure correlator allows us to tie in long-time diffusion behavior with the disorder in a system via the homogenization argument, as introduced in Ref. Novikov et al. (2014). Briefly, this argument proceeds in two steps.

At the first step, one realizes that, no matter how strongly heterogeneous the system is microscopically, at long times these inhomogeneities become effectively coarse grained (homogenized) over the increasing diffusion length scale $L(t) = (\langle |\delta x(t)|^2 \rangle)^{1/2} \approx \sqrt{2Dt}$, where $D_\alpha$ is the finite
bulk diffusion coefficient and $\dot{x} = x(t) - x(0)$. In other words, while at short times the diffusing molecules “see” individual axonal walls and other sharp pm-level features, at longer times these features are progressively blurred, becoming indistinguishable, from the diffusion standpoint, from some effective, almost uniform medium. This looks as if a Gaussian filter with a window $\sim L(t) \gg \langle \tau_{\text{ext}} \rangle$ were applied to a medium. At this point, it is justifiable to introduce a coarse-grained local diffusion coefficient $D(r)$ which smoothly varies on the scale $\geq L(t)$, and whose variation $\partial D(r) = D(r) - D_\infty$ becomes small, asymptotically approaching zero as $t \to \infty$. This smallness in turn justifies applying the lowest-order perturbation theory for the effective diffusivity in a medium with a weakly-varying local $D(r)$. Technically, this theory is most directly built in the frequency domain, for the dispersive diffusivity $D(\omega)$ (Ernst et al., 1984; Novikov and Kiselev, 2010; Novikov et al., 2014):

$$D(\omega) = \frac{-i\omega}{2} \int \frac{d^2k}{(2\pi)^2} \Gamma_0(k) \left( \frac{\partial D}{\partial \omega} \right) \omega,$$

where $\Gamma_0(k) = \int \Gamma_0(r)e^{-\omega r}d^3r$ is the Fourier transform of the two-point correlation function $\Gamma_0(r) = \langle \partial D(t_0 + r)\partial D(t_0) \rangle$ of the coarse-grained weakly-varying diffusion coefficient in $d$ spatial dimensions. Here we consider a statistically isotropic case, so that $\Gamma_0(k)$ depends only on $k = |k|$. In the time domain, Eq. (A.1) yields

$$D_{\text{int}}(t) \equiv D_\infty + \frac{1}{2} \int \frac{d^2k}{(2\pi)^2} \Gamma_0(k)e^{-\omega_0k^2} t.$$

At the second step, one needs to relate a somewhat artificial object, $\Gamma_0$, entering Eqs. (A.1) and (A.2), to the correlation function $\Gamma(r) \equiv \langle \rho(\mathbf{r} + \mathbf{r}_0)\rho(\mathbf{r}_0) \rangle$ of the actual restrictions, Eq. (B.1) below. As it has been realized in Ref. Novikov et al. (2014), in the long-distance (or low $k$) limit, these correlation functions become asymptotically proportional to each other:

$$\Gamma_0(k) \propto C \cdot \Gamma(k), \quad C = \left( \frac{\partial D}{\partial \rho} \right)^2.$$

The dependence $D_{\text{int}}(\rho)$ of the macroscopic diffusion coefficient on the macroscopic density $\rho = \langle \rho(\mathbf{r}) \rangle$ of the restrictions (e.g., axons) is non-singular, such that $0 < C < \infty$, as long as the system is not at its percolation transition, so that $D_\infty > 0$ in the extracellular space, which is always the case for neural tracts. Therefore, in a general situation, i.e. away from the single packing density (for which $D_\infty = 0$), the $\omega \to 0$ and $t \to \infty$ asymptotic forms of Eqs. (A.1) and (A.2) can be (up to a coefficient $C$) deduced by plugging in the corresponding $k \to 0$ behavior of the structure correlator $\Gamma(k) \sim k^p$.

For the short-range disorder, $p = 0$, the correlation function $\Gamma(k)$ has a plateau $\Gamma(k)_{\text{lc}} \to \rho^2 > 0$ as $k \to 0$, Fig. A.4; hence, a similar plateau $\Gamma(k)_{\text{lc}} \to \rho^2$ enters Eq. (A.1). Substituting this value into Eqs. (A.1) and (A.2) for $d = 2$, we obtain Eqs. (1), (2), (8), (18), and (19) with the coefficient

$$A = \frac{C \cdot \Gamma(0)}{8\pi D_{\text{int}}^2}.$$

An accurate theory of the parameter $A$ is nontrivial, as it involves finding analytically the dependence $D_{\text{int}}(\rho)$, and is beyond the scope of this work. Here, we roughly estimate

$$A \sim \frac{\Gamma(0)}{D_{\text{int}}} \sim \frac{\text{var}[D(r)]}{D_{\text{int}}} \frac{L_{\text{int}}^2}{D_{\text{int}}} \frac{D_{\text{int}}^2}{D_{\text{int}}^2} = \lambda^2,$$

based on the fact that, similar to the correlation function for the Poissonian disorder, $\Gamma(k)$ is dominated by a 6-function peak around $r = 0$ of amplitude $-\text{var}[D(r)] \sim (\Delta^2)^2$ and width $L_{\text{int}} - L_i$, corresponding to the minimal diffusion length scale at which we coarse-grained the medium and applied Eq. (A.1). This peak sharply decreases to 0 for $r \gg L_i$. Coarse-graining past $L_i$ means that the fluctuations of $D(r)$ are already small, $\text{var}[D(r)] \leq D_{\text{int}}^2$. This estimate for $A$ is in agreement with our data summarized in Table 2. The scaling $A \sim \frac{L_i}{\lambda}$ (Supplementary fig. S4) is only approximate, since the plateau $\Gamma(k)_{\text{lc}} \to \rho^2 > 0$ is determined not only by $L_i$ but also by how strongly the objects “repel” each other, cf. Fig. 4 showing fairly strong sensitivity of the plateau value $\Gamma(k)_{\text{lc}} \to \rho^2$ to the arrangement of the objects. Now it is also evident that the units for $\rho(r)$ in the examples of Fig. 3 (cf. Appendix B) are inessential, compensated by the respective units in the derivative in C, Eq. (A.3).

Appendix B. The structure correlation function $\Gamma(k)$

In order to formalize the qualitative difference between the computer-generated ordered packing in Panel a and the disordered ones in Panels b–d of Fig. 3, consider the density correlation function defined in a standard way,

$$\Gamma(k) = \langle \rho(\mathbf{r} + \mathbf{r}_0)\rho(\mathbf{r}_0) \rangle \equiv \int \rho(\mathbf{r} + \mathbf{r}_0)\rho(\mathbf{r}_0) \frac{d\mathbf{r}_0}{V}.$$

where $\rho(r)$ is the density of restrictions (black pixels in our examples in Fig. 3) at the point $r$, and $V$ is the sample volume (area). The units of $\rho(r)$ and the overall normalization are not important for the qualitative distinction we are going to illustrate (cf. Appendix A); here we assumed $\rho = 1$ for the black pixels and $\rho = 0$ for the white pixels in Fig. 3.

Practically, the correlation functions of the restrictions in Fig. 4 are calculated in the Fourier domain, using the equivalence with the power spectral density

$$\Gamma(k) = \int e^{-i\mathbf{kr}} \rho(\mathbf{r} + \mathbf{r}_0)\rho(\mathbf{r}_0) \frac{d\mathbf{r}_0}{V} \equiv \frac{1}{V} \rho(-\mathbf{k})\rho(\mathbf{k}),$$

where $\rho(k) = \int d\mathbf{r} e^{-i\mathbf{kr}} \rho(\mathbf{r}) = \rho(\mathbf{r})$. In particular, given the shape functions, $\psi(\mathbf{r}) = 1$ inside an object and 0 outside (cf. Fig. 3), we find $\rho(k) = \sum_1^N \psi(\mathbf{r}_j \sim -\mathbf{r}_j)$ for any array of shapes, with the Fourier transform given by

$$\rho(k) = \sum_{j=1}^N \psi(\mathbf{r}) e^{-i\mathbf{kr}}.$$

We note a few important properties of the function $\rho(k)$.

1. Large-k behavior For a disk of radius $a$, $\psi(r) \sim 1$, $\psi(k) = 2\pi a^2 j_0(ka)/k$, where $j_1$ is a Bessel function of the first kind. We note that $\psi_{\text{lim}} \sim 1 - k^{-1/2}$. This scaling at large $k$ is a general signature of the sharp boundaries with finite curvature in two dimensions, and is indeed present in Fig. 4 in the form $\Gamma(k) \sim |\psi(k)|^2 \sim k^{-3}$ as $k \to \infty$ for both the disks and for the randomly-shaped axons from Fig. 3f.

2. Normalization Consider the case of totally uncorrelated positions $\mathbf{r}_j$ of the objects (the Poissonian disorder), and let us also for simplicity assume them identical, $\psi \equiv 1$ in Eq. (B.3). Substituting Eq. (B.3) in Eq. (B.2), we obtain

$$\Gamma_{\text{Pois}}(k) = \frac{|\psi(k)|^2}{V} \sum_{j=1}^N \delta^2(\mathbf{r}_j - \mathbf{r}_j) = n|\psi(k)|^2, \quad n = N/V.$$

since only the terms with $j = j'$ survive and the rest are averaged to zero due to the random phases. Hence, the plateau value for the Poissonian disorder is $\Gamma_{\text{Pois}}(k)_{\text{lc}} \to \rho^2 = n\psi^2(0)$, where $\psi(0) = \psi(\mathbf{r})_{\text{lc}} = 0$ is the volume (area) of a single object, and $\phi = n\psi^2(0) = \psi^2(0)$ the volume (area) fraction of the objects. The Poissonian disorder is a good reference point, hence we have chosen to normalize all correlation functions in Fig. 4 onto the value $\psi_{\text{lim}}$. For the axons in Fig. 3f, we defined $\psi_{\text{lim}} = \phi/n$, and the (external) object radius $\langle \tau_{\text{ext}} \rangle = \sqrt{V/\pi}$ for the $k$-axis normalization.
We can see that the top curve, corresponding to Fig. 3d, plateaus at almost unity. (It does not reach unity because we assign \( \rho = 1 \) instead of \( \rho = m \) to the pixels where \( m \) disks overlap, so there is a bit of anti-correlation baked in this \( \rho(r) \).) Interestingly, all other correlation functions have plateaus notably below \( \rho_0 \), which is expected since the objects are anti-correlated due to their “repetition” (non-overlap condition).

### 3. Discontinuity at \( k = 0 \)

Since \( \rho(k)|_{k=0} = V_0 \), which is the volume (area) occupied by the objects, the value \( \Gamma(k)|_{k=0} = \pi V_0 \). This value is different from the plateau \( \Gamma(k)|_{k \to +0} \) described above. This reflects the fact that \( \Gamma(k) \) is discontinuous at \( k = 0 \). physically, we are interested in the plateau \( \Gamma(k)|_{k \to +0} \), as it reflects the statistical properties of the medium at distances \( \sim 1/k \) which are large, but still smaller than the system size. Conversely, the value \( \Gamma(k)|_{k \to 0} \) is irrelevant since it is only approached for \( k \) of the order of inverse system size (i.e. it is never approached in the thermodynamic limit \( N, V \to \infty \)), and thereby it cannot affect any observable physical quantities.

### Appendix C. Time-dependent diffusion inside a perfectly reflecting cylinder in the narrow-pulse limit

The diffusion propagator for a perfectly reflecting cylinder of radius \( a \) is given by Stepnišnik (1993) and Callaghan (1995)

\[
S(t, q) = \sum_{k=1}^{n} 4 \pi e^{-q \beta_{ik} a / \sqrt{D}} \left[ (q^2 f_0(q^2 a^2)) \right] + \sum_{n, k=1}^{\infty} 8e^{-q^2 \beta_{ik} a / \sqrt{D}} \left[ \beta_{ik} \left( q^2 f_0(q^2 a^2) \right) \right] \]

where \( J_n \) are Bessel functions of the first kind and \( \beta_{ik} \) is the kth root of their derivative \( J_n'(x) \).

Taylor expanding up to \( q^2 \) reveals that only the Bessel functions \( J_n \) with \( n = 0 \) and 1 contribute at this level, as higher \( n \) result in the powers \( q^8 \). Similarly, the only \( k \) value needed for \( \beta_{ik} \) in the first sum is \( k = 1 \) with \( \beta_{i1} = 0 \). Using \( J_0(x) \sim 1 - x^2 / 4 + x^4 / 64 - x^6 / (36 \cdot 64) \) and \( J_1(x) = x / 2 \) for \( x \ll 1 \), Eq. (C.1) simplifies to

\[
S(t, q) = 1 - \frac{q^2}{\sqrt{D}} t + O(q^4) \]

From the above equation, we find

\[
D(t) = \frac{q^2}{4t} - 2\frac{q^2}{t} \sum_{k=1}^{n} e^{-q \beta_{ik} a / \sqrt{D}} \frac{1}{\beta_{ik}^2 (\beta_{ik}^2 - 1)}.
\]

This expression is plotted in Fig. C.1. Using (Neuman, 1974)

\[
\frac{1}{\beta_{ik}^2 - 1} = 1 - \frac{1}{2}, \quad \frac{1}{\beta_{ik}^2 (\beta_{ik}^2 - 1)} = \frac{1}{8}.
\]

the short-time limit of Eq. (C.2) is indeed \( D(t)|_{t \to 0} = D_0 \). The next-order correction, to which all the eigenmodes in Eq. (C.2) contribute, is the universal \( -\sqrt{t} \) term (Mitra et al., 1992), shown in Fig. C.1. It works for \( t < \tau_0 \), where \( \tau_0 = a^2 / D_0 \) is the diffusion time for the cylinder. To see how all the eigenmodes contribute to the Mitra limit, one can substitute the summation by integration via approximating \( \beta_{ik} = n k \) for \( k \gg 1 \), and regularize the resulting integral

\[
\int_{-\infty}^{\infty} \frac{dk}{\pi k^2} e^{-k^2 / \tau_0} \left( \frac{\tau_0}{2 \pi} \right)^{1/2} \Gamma\left( -\frac{3}{2} \right)
\]

\[\text{Fig. C.1. Cumulative diffusion coefficient (Eq. (C.2)) for an impermeable cylinder for typical axonal parameters,} \quad D_0 = 2 \mu^2 / \text{ms} \text{ and radius} \quad a = 0.5 \mu \text{m. The characteristic diffusion time} \quad \tau_0 \text{ is shown by the vertical black dashed line. The diffusion times commonly employed in dMRI fall into the long time limit for typical axons. Note a much faster, exponential decrease of} \quad D_{\text{mut}}(t) \text{ given by the first term of the sum (Eq. (C.3)) (dashed magenta), as compared to} \sim 1/t \text{ asymptote (Eq. (C.4)) of} \quad D(t) \text{ (dot-dashed red). The MC simulated} \quad D(t) \text{ (open circles) perfectly agrees with the theoretical value (C.2).}
\]

using Euler’s Gamma function \( \Gamma(-3/2) = 4 \pi / 3 \). (In the above equation, we extended the lower limit of integration to \( k = 0 \). with the understanding that there is no actual divergency at \( k \to 0 \) in the original discrete sum; the formally divergent integral \( \int_0^\infty dx x^{1/2} e^{-x} \) with \( x = \pi^2 k^2 / \tau_0 \) is then understood in terms of the Gamma function \( \Gamma(z) \) after analytic continuation in the complex plane of \( z \) from \( z > 0 \) to \( z = -3/2 \).) Collecting all the coefficients yields the exact Mitra \( \sqrt{t} \) correction, with \( S/V = 2 \mu a^2 / 2a \).

In our example in Fig. C.1, we use \( D_0 = 2 \mu^2 / \text{ms} \) and a cylinder radius \( a = 0.5 \mu \text{m} \) for a typical axon with an internal diameter of 1 \( \mu \text{m} \). For this case, \( \tau_0 = 0.13 \text{ ms} \), therefore, for typical axonal length scales, all dMRI measurements will be in the long time limit. For the long times, the sum over the eigenmodes in Eq. (C.2) becomes negligible, and

\[
D(t) = \frac{q^2}{4t}, \quad t \gg \tau_0.
\]

This limit is utilized in Eq. (27) of the main text.

From Eq. (C.2), using Eq. (4), we obtain the instantaneous

\[
D_{\text{mut}}(t) = \sum_{k=1}^{n} 2D_0 \frac{e^{-q \beta_{ik} a / \sqrt{D}}}{\beta_{ik}^2 (\beta_{ik}^2 - 1)}
\]

\[\text{Fig. C.1 shows how distinctly faster is the exponential decrease of} \quad D_{\text{mut}}(t) \text{ (solid black line) as compared to the power law decrease (Eq. (C.4)). Indeed, practically,} \quad D_{\text{mut}}(t) \text{ is already captured by the lowest} \quad k = 1 \text{ eigenmode in the sum (Eq. (C.5)), as shown by the dashed magenta line (cf. Eq. (9) of the main text). The higher modes are only really needed to get the correct} \quad t \ll \tau_0 \text{ behavior of Eqs. (C.2) and (C.5).}
\]

### Appendix D. Frequency-dependent diffusion coefficient inside a perfectly reflecting cylinder

Taking the Fourier transform (Eq. (14)) of Eq. (C.5), we obtain

\[
D(\omega) = \sum_{k=1}^{n} 2D_0 \frac{-i\omega}{\beta_{ik}^2 D_0 / a^2 - i\omega}.
\]

The lowest-order term in \( \omega \) of Eq. (D.1) is purely imaginary, \( D(\omega) = -i\omega^2 / 4 + O(\omega^3) \), and corresponds to the long-time limit (C.4). The
The result (Eq. (D.2)), together with its low frequency limit (Eq. (D.4)), and its high frequency limit based on Ref. Novikov and Kiselev (2011) with \( S/V = 2/a \), are plotted in Fig. D.1. We observe that for axons (generally, for impermeable cells) of \( -1 \) mm size, the high-frequency limit is practically unattainable, as it requires frequencies beyond \( \sim 10 \) kHz, and the OG-measured diffusivity has a clear low-frequency parabolic dependence (Eq. (D.4)) within a few kHz range. As with Eq. (C.5), we again see that the lowest \( k = 1 \) eigenmode very well captures the low-frequency behavior of \( D(\omega) \). Indeed, the sums in Eqs. (C.3) and (D.3) converge very rapidly and are well approximated by their respective first terms involving only \( \beta_1 \) \( \approx \pm 1.8412 \). For instance, keeping only the \( k = 1 \) term in Eqs. (D.1)–(D.3), we get the \( \omega^2 \) behavior with a coefficient \( 2 \left( \beta_1 \right)^2 \left( \beta_1^2 - 1 \right)^{-1} \approx 0.0728 \), approximating the exact \( \beta_1^2 + 0.5 \) to within a 0.14% accuracy.

We note that a qualitatively similar dominance of the lowest eigenmode also occurs for periodic lattices (Dunn and Bergman, 1995; Sen et al., 1994; Sukstanski et al., 2004; Novikov et al., 2014). The physical reason is the same: the spectrum of discrete eigenmodes, which, for sufficiently long \( t \gg \Delta D_0 \), where \( \Delta \) is the lattice constant, allows one to keep the lowest one, cf. Eq. (9), that is separated by a gap \( \sim 2 \Delta D_0 / \Delta L \) from the rest, whose contributions exponentially die out even faster. This is the rationale behind the simple Lorentzian representation (Eq. (20)) used in the main text.

In disordered media, when there is no spectral gap, the sum (Eq. (A.2)) over the finite density of exponentially decaying eigenmodes yields a power law (Ernst et al., 1984; Macht et al., 1984; Vischer, 1984; Novikov et al., 2014). More generally, this is why transport properties of ordered and disordered systems are so qualitatively distinct, even at the level of purely classical diffusion, with this distinction becoming even more pronounced in the case of quantum transport, where disorder leads to localization of single-particle wave functions (Anderson, 1958; Abrahams et al., 1979).

**Appendix E. Axonal diameter mapping up to \( \mathcal{O}(q^4) \) in the narrow-pulse limit**

Consider first for simplicity the case of all axons to be of the same internal radius \( r = a \), and the narrow pulse limit \( \delta \ll t_c \) of diffusion weighting. The total signal is additive,

\[
S(t, q) = f_{\text{int}} S_{\text{int}}(t, q) + f_{\text{ext}} S_{\text{ext}}(t, q), \quad q = \delta \theta, \tag{E.1}
\]

where \( \delta \) is the diffusion-sensitizing gradient value (in the units of Larmor frequency per unit length) during the short pulse. We will now expand both terms of Eq. (E.1) up to \( \mathcal{O}(q^4) \).

The intra-axonal signal in the limit \( t \gg q^2 / D_0 \) is given by the \( k = 1 \) term in the first sum in Eq. (C.1). Its expansion

\[
S_{\text{int}}(t, q) = 1 - \frac{4 a q^2}{D_0} - \frac{5}{6} \frac{a q^4}{D_0^2} + \mathcal{O}(q^6), \tag{E.2}
\]

is equivalent to the kurtosis value \( K_{\text{int}}(\omega) = -1/2 \) transverse to an impermeable cylinder. The cumulative expansion of the extra-axonal signal for \( t \ll t_c \)

\[
\ln S_{\text{ext}} = -\left( D_0 + A \ln \frac{L_c}{l} \right) q^2 t + \frac{K_{\text{ext}}(t)}{6} \left( D_0 q^2 t \right)^2 + \cdots, \tag{E.3}
\]

Using the approach (Novikov and Kiselev, 2010), we found (details to be published elsewhere) the kurtosis due to the disordered packing, as

\[
K_{\text{ext}}(t) = \frac{A}{D_0} \ln \frac{t}{t_c}, \quad t \gg t_c. \tag{E.4}
\]

One can already expect the behavior \( K_{\text{ext}} \sim A / (D_0 t) \) on purely dimensionless grounds (kurtosis is dimensionless and must scale with the disorder strength \( A \)). The \( \ln t \) singularity is not trivial and has the same origin as in Eq. (1). In Eq. (E.3), within the accuracy of keeping the terms \( \sim A \ln t \), we approximated \( D(t) \) by \( D_0 \) in the term \( D_0 q^2 t^2 \).

Substituting Eqs. (E.2) and (E.3) into Eq. (E.1) and regrouping the terms into the moments \( \langle \chi^4 \rangle \) of the total signal, we obtain

\[
S = 1 - \frac{q^2}{2!} \langle \chi^2 \rangle / 2! + q^4 \langle \chi^4 \rangle / 4! + \mathcal{O}(q^6), \tag{E.5}
\]

\[
\frac{q^2}{2!} = f_{\text{int}} A^2 + f_{\text{ext}} \left( D_0 t + A \ln \frac{L_c}{l} \right) \tag{E.6}
\]

\[
\frac{q^4}{4!} = f_{\text{int}} \frac{5}{3} \frac{a q^4}{D_0} + f_{\text{ext}} \left[ \frac{K_{\text{ext}}}{6} \left( D_0 q^2 t \right)^2 + \frac{1}{2} \left( D_0 t + A \ln \frac{L_c}{l} \right)^2 \right]. \tag{E.7}
\]

Consider now the result of matching the \( q^2 \) and \( q^4 \) terms of Eq. (E.5) to those of the conventional model

\[
S_0(t, q) = f_{\text{int}} S_{\text{int}}(t, q)\big|_{q \to q_{\text{app}}} + f_{\text{ext}} e^{-D_0 q_{\text{app}}^2 t}, \tag{E.8}
\]

which assumes time-independent Gaussian diffusion in the extra-axonal space, with some apparent axonal radius \( r_{\text{app}} \) instead of \( a \) in the intra-axonal signal expansion (Eq. (E.2)).
Matching of \((x^2)\) terms of models in Eqs. (E.5) and (E.8) yields

\[
f_{\text{int}} \cdot \frac{r_{\text{app}}^2}{4} = f_{\text{int}} \frac{a^2}{4} + f_{\text{ext}} \cdot A \ln \frac{t}{\tau_c},
\]

which gives Eq. (34) after generalizing \(a^2 \to (\bar{r}^2)/\rho^2\) onto the distribution of the internal radii \(\bar{r}\) by taking the volume average over the distribution of axons. Hence, fitting of the actual signal (Eq. (E.5)) to the conventional model (Eq. (E.8)) at the lowest diffusion weighting would give \(r_{\text{app}}(t) > a\).

The increase in the diffusion weighting would further bias the result for \(r_{\text{app}}\). Matching of the \((x^2)\) terms, we neglect the \(-A^2\) term as a small correction, yields a different expression for \(r_{\text{app}}(t)\). Using Eq. (E.4), we find

\[
r_{\text{app}}^4(t) \approx a^4 + \frac{2}{5} \cdot \frac{3 \cdot 64}{f_{\text{ext}}} AD_{\text{ext}} \ln \frac{t}{\tau_c}, \quad t \gg \tau_c,
\]

where we neglected the \(-A^2\) term as a small correction, yields a different expression for \(r_{\text{app}}(t)\). Using Eq. (E.4), we find

\[
r_{\text{app}}^4(t) \equiv \frac{5}{3} \cdot \frac{f_{\text{app}}^2}{6} A \cdot D_{\text{ext}} \cdot A \ln \frac{t}{\tau_c},
\]

where we neglected the \(-A^2\) term as a small correction, yields a different expression for \(r_{\text{app}}(t)\). Using Eq. (E.4), we find

\[
r_{\text{app}}^4(t) \approx \frac{5}{3} \cdot \frac{f_{\text{app}}^2}{6} A \cdot D_{\text{ext}} \cdot A \ln \frac{t}{\tau_c}, \quad t \gg \tau_c.
\]

References


