Abstract. Diffusion magnetic resonance imaging (dMRI) is a non-invasive tool to obtain information on the neural architecture in vivo. Among MR diffusion assessment techniques, diffusion tensor imaging (DTI) has become the most popular. However, DTI suffers from intrinsic methodological limitations. The main drawback is the inability of the model to cope with non-Gaussian diffusion. DTI assumes that the diffusion in neural tissue is Gaussian (i.e. free, fast exchange across cellular compartments and with no restriction). DTI is unable to separate multiple compartments (partial volume effects, multiple fiber orientations etc.). The objectives of the present paper are to learn about the limits of the Gaussian diffusion model when applied to biological tissues and about model-free MR diffusion techniques.

Key words: diffusion, magnetic resonance imaging, Gaussian, non-Gaussian.

1. INTRODUCTION

Diffusion is one of several “transport processes” that occur in nature. An interesting feature of diffusion is that it occurs even in thermodynamic equilibrium. This is quite remarkable because the picture of diffusion expressed in Fick’s First Law was that particle, like heat, flow from regions of high concentration to low concentration. When these gradients vanished, there was no net flux. While the net flux vanished, there are still diffusive fluxes nonetheless, however they cancel each other [1].

Diffusion process describe liquids and gases molecules at a temperature greater than absolute zero (K) which are in constant motion and collide with one another. The greater the kinetic energy, the faster the movement. With each collision, a water molecule experiences a random displacement. As a result, a group of molecules that start at the same location will be spread out over time. This random motion also leads to microscopic motion of suspended particles and was
observed by Brown in 1828 for grains of pollen in water and is referred to as “Brownian motion”. Figure 1 displays the probability of molecules displacement ($P_{1D}$) for 1-D diffusion having a Gaussian shape for the unrestricted diffusion.

![Figure 1](image)

Fig. 1 – The top of the histogram is centered on zero, indicating that most molecules had the same position at $t = 0$ and $t = \Delta$; $\Delta$ is the experiment diffusion time, $P(x/\Delta)$ is the molecule displacement probability.

The most popular is diffusion model is the Gaussian one (Fig. 1). However, water motion in biological tissue can hardly be described as an unrestricted (Gaussian) diffusion. Impermeable membrane, dense geometrical arrangement, different cellular environment may cause a deviation of diffusion from Gaussian behavior. A typical voxel in a diffusion MRI experiment is of the order of 10 mm$^3$ and thus contains thousands of cells and tissue components. The diffusion of water molecules in each compartment (e.g. extra/intracellular space, axons, dendrites) is affected by the local viscosity, composition, geometry, and membrane permeability.

During typical diffusion times of about 50 msec used in diffusion MRI experiments, water molecules move in the brain on average over distances around 10 µm, bouncing, crossing or interacting with many tissue components such cell membranes, fibers or macromolecules [2]. Just because diffusional processes are influenced by the geometrical structure of the environment, magnetic resonance (MR) technique can be used to probe the structural environment non-invasively.

In 1965, Stejskal and Tanner introduced innovations that made modern diffusion experiments by MR imaging possible. They introduced the pulsed gradient spin-echo (PSGE) where the diffusion-encoding parameters are the size and direction of the diffusion-encoding gradient vector, $G$, the time difference between the onset of the two diffusion-encoding gradients, $\Delta$, and the duration of each of the diffusion encoding gradients, $\delta$.

The sequence allowed a clear distinction between the encoding time (gradient pulse duration, $\delta$) and the diffusion time (separation of the two pulses, $\Delta$). In this sequence, the diffusion gradients can be so short that diffusion taking place during the application of these pulses can be neglected.
Protons precession is proportional to the magnet strength so, in this “narrow pulse” regime, the net phase change induced by the first gradient pulse is given simply by:

\[ \varphi_1 = -q \cdot x_1, \]

where is \( x_1 \) is the position of the particle during the application of the first pulse and \( q = \gamma \delta G \), where \( \delta \) and \( G \) are the duration and the magnitude of the gradient pulses, respectively. Similarly, if the particle is situated at position \( x_2 \) during the application of the second pulse, the net phase change due to the second pulse is given by:

\[ \varphi_2 = -q \cdot x_2. \]

The 180° rf pulse applied in between the two gradient pulses reverses the phase change by the first gradient pulse. Therefore, the aggregate phase change that the particle suffers is given by:

\[ \varphi_2 - \varphi_1 = -q \cdot (x_2 - x_1). \]

Clearly, if particles remained stationary, i.e., \( x_1 = x_2 \) the net phase shift would vanish. In this case, and in the case in which all spins are displaced by the same constant amount, the magnitude of the echo will be unchanged (except for the \( T_1 \) and \( T_2 \) decay that is occurring throughout the sequence). However, if particles diffuse randomly throughout the excited volume, the phase increment they gain during the first period does not generally cancel the phase decrement they accrue during the second period. This incomplete cancellation results in phase dispersion or a spreading of phases among the randomly moving population of spins. Therefore, the overall signal, given by the sum of the magnetic moments of all spins, is attenuated due to the incoherence in the orientations of individual magnetic moments.

It is more convenient to introduce a new quantity called MR signal attenuation than to deal with the MR signal itself. Signal attenuation is obtained by dividing the diffusion-attenuated signal by the signal in the absence of any gradients \( \left( \frac{S_{(G, \delta)}}{S_0} \right) \).

The signal attenuation in the MRI voxel induced by the diffusion encoding gradients will be given by the average dephasing of the spins of the molecules in the voxel. The general form of this dependence is described by:

\[ \frac{S_{(G, \delta)}}{S_0} = \langle e^{in} \rangle. \]
Fig. 2 – Stejskal and Tanner MR diffusion experiment. They introduced the pulsed gradient spin-echo (PSGE) where the diffusion-encoding parameters are the size and direction of the diffusion-encoding gradient vector, $G$, the time difference between the onset of the two diffusion-encoding gradients, $\Delta$, and the duration of each of the diffusion encoding gradients, $\delta$.

More specifically, in a diffusion experiment, if one ignores the relaxation-related signal attenuation, the MR signal attenuation is given by:

$$\frac{S_{(G,\Delta,\delta)}}{S_0} = \int \int P(x_1, x_2, \Delta) \cdot e^{-i\mathbf{x}_2 \cdot \mathbf{x}_1} \, dx,$$

where $\rho$ is the spin density and $P$ is the probability that a particle initially located at position $x_1$ will have ended up at $x_2$ after a time delta.

The parameter of interest (to be obtained) is the displacement probability function ($P$). The displacement probability function provides the probability that a randomly chosen water molecule within the volume of interest (i.e., a single voxel) will have a particular displacement over the diffusion time. This parameter contains information on local tissue structure.

If diffusion is free (isotropic), the MR signal attenuation is given by:

$$\frac{S_{(G,\Delta,\delta)}}{S_0} = e^{-bD},$$

where $b = \left(\gamma \delta G\right)^2 \cdot (\Delta - \delta / 3)$ and $D$ is the classical diffusion coefficient (appearing also in Fick’s law). This equation is known as Stejskal-Tanner equation.

In other words, the unrestricted diffusion is modeled by a mono-exponential decay of the MR signal.

In tissues, such as brain gray matter, where the measured apparent diffusivity is largely independent of the orientation of the tissue (i.e. isotropic) at the voxel length scale, it is usually sufficient to describe the diffusion characteristics with a single (scalar) diffusion coefficient ($D$ known as $ADC$ (apparent diffusion...
coefficient) in MRI experiments) (Fig. 3, upper image). However, in anisotropic media, such as white matter, cardiac and skeletal muscle where the measured diffusivity is known to depend upon the orientation of the tissue, a single ADC cannot characterize the orientation-dependent water mobility. The causes of diffusion anisotropy have not been fully elucidated in brain parenchyma, although most investigators ascribe it to ordered, heterogeneous structures, such as large oriented extracellular and intracellular macromolecules, supermacromolecular structures, organelles, and membranes. In the central nervous system (CNS), diffusion anisotropy it is mostly regarded as a parallel orientation of myelinated axons consequence (Fig. 3, lower image) but it is not simply caused by myelin since several studies have shown that even before myelin is deposited, diffusion anisotropy can be measured using MRI [3].

2. DIFFUSION TENSOR IMAGING

When the diffusion process is different for different directions, the next model that can describe anisotropic diffusion replaces the scalar diffusion coefficient with a symmetric diffusion tensor of water; the diffusion is described by a symmetric tensor with 9 components. The method is called Diffusion tensor imaging (DTI) [3].

![Diffusion Tensor Imaging Diagram](image)

Fig. 3 – The parallel orientation of axons in white matter is, in good part, responsible for the preferential (anisotropic) diffusion of water (top figure). In gray matter, the randomly distributed cells produce the isotropic diffusion of water (bottom figure – copyright license number: 3230730540453, Elsevier).

In vivo, the orientation of the main axis of diffusion (the orientation of fibers) might not be the same to the orientation of the laboratory frame (field orientation).
In this case, the orientation of the ellipsoid in the laboratory frame is determined by the rotation matrix containing the 3 eigenvectors (Fig. 4).

In order to assess such diffusion anisotropy in an orientation-independent manner, the diffusion tensor, measured in the laboratory frame needs to be diagonalized to determine the magnitude of the apparent diffusion constants in each voxel. The magnitude of the three resulting ADCs, the eigenvalues ($\lambda_1$, $\lambda_2$, $\lambda_3$) determine the shape of the so-called ellipsoid.

$$\mathbf{D} = \mathbf{D}_{\text{II}} = \mathbf{D} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xz} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

$$\mathbf{D} = \left( \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \right) \cdot \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \cdot \left( \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \right)^T$$

After diagonalizing the tensor, the eigenvalues can be used to calculate a multitude of diffusion anisotropy parameters.

Fractional anisotropy is the most used diffusion parameter. It's computed by comparing each eigenvalue with the mean of all the eigenvalues ($\lambda$) as in the following equation:

$$FA = \sqrt{\frac{3}{2}} \cdot \sqrt{\frac{\left(\lambda_1^2 - \lambda\right) + \left(\lambda_2^2 - \lambda\right) + \left(\lambda_3^2 - \lambda\right)}{\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2\right)}}.$$

FA describes the degree of anisotropy of a diffusion process. A value of zero means that diffusion is isotropic, i.e. it is unrestricted (or equally restricted) in all directions. A value of one means that diffusion occurs only along one axis and is fully restricted along all other directions.

Diffusion tensor imaging is the most popular model for diffusion. Despite that, one should bare in mind that DTI has intrinsic methodological limitations, to mention some of them: (1) the assumption that diffusion in neuronal tissue is Gaussian (i.e. free, fast exchange across cellular compartments and with no restriction), (2) inability to separate multiple compartments (including partial volume effect and multiple fiber orientations), (3) high noise sensitivity.

These limitations cause post-processing artifacts that reduce the reliability of the extracted indices and alter the accuracy of reconstructed fiber systems.

3. BEYOND DTI

When diffusion is no-longer free or happens in multiple compartments (that are not in fast exchange) a non-Gaussian behavior will be observed. Cells, neuronal
fibers and extra-cellular matrix are the compartments of the neuronal tissue. The exact diffusion properties (diffusion coefficient, anisotropy, restricted diffusion and exchange) of each of these compartments are unknown; however those factors do contribute significantly to the diffusion signal decay. Practically, when a diffusion experiment is done and a non-mono-exponential signal decay is observed, neither the Stejskal-Tanner relation nor the most basic DTI analysis can be used.

The deviation from mono-exponential diffusion is obvious when experimental condition like strong gradient pulses and long diffusion times are undertaken - high b value experiments (i.e. b value higher than 2,500 s/mm²). Facing a non-mono-exponential decay there are two ways to analyze the data. The one is modeling the signal or the cellular system; for it, a basic model - a bi-exponential decay function has been suggested. Another option is to use a model-free approach – the q-space approach [4, 5].

3.1. BI-EXPONENTIAL MODEL

The most basic model used to analyze high b-value data (when deviation from Gaussian diffusion is significant) is the bi-exponential model. This model is based on the assumption of two non-exchanging compartments: one exhibiting fast diffusion (extracellular space) and the other slow diffusion (intracellular space).

For the WM, assigning the highly restricted water diffusion inside the axons to the slow compartment and the less hindered diffusion in the extra-axonal space to the fast compartment, has been justified experimentally.

Parametric images of the biexponential parameters available from pixel-by-pixel fits of diffusion decay datasets in adult brain and newborn brain. The amplitude of the fast diffusion component ($A$) is significantly higher in newborn brains which have only a small slow diffusion component amplitude ($B$) than in adult brains (Fig. 4).

However, other anatomical factors can influence the diffusion anisotropy measurement. For example, in a bi-exponential model, the intra-extra cellular exchange is neglected.

Therefore, more dedicated, model-free (non-Gaussian) methods for diffusion have been introduced. Parameters like Mean displacement using q-space imaging (QSI) and Diffusional kurtosis using diffusional kurtosis imaging (DKI) have been increasingly used in reports dealing with clinical images [6,7]. Mean displacement and diffusional kurtosis proved to have advantages over parameters obtained by diffusion tensor imaging parameters (fractional anisotropy) in assessing e.g. cerebral ischemia or in observing neuronal remodeling after brain injury or stroke.
Fig. 4 – Parametric images of the biexponential parameters available from pixel-by-pixel fits of diffusion decay datasets in adult brain (left column) and newborn brain (right column). The amplitude of the fast diffusion component ($A$) is significantly higher in newborn brains which have only a small slow diffusion component amplitude ($B$) than in adult brains (copyright license number: 3232460161188, Elsevier).

3.2. Q-SPACE DIFFUSION

According to basic MR imaging principles, the measured signal at conventional MR imaging is phase and frequency encoded. It is the result of the application of gradients in different directions and with different intensities at specific moments of the acquisition sequence. The values of the measured signal are organized in a coordinate system known as $k$-space. Performing the acquisition enables the filling of $k$-space (a spatial frequency space). To transform the raw MR imaging data from $k$-space into a position-encoded visual image, a mathematical operation known as a Fourier transform is applied.

The process in diffusion MR imaging is analogous; digitized data from diffusion are stored in a similar space called $q$-space. The $q$-space approach assumes no specific model of water diffusion. Callaghan introduced the $q$-vector as:

$$q = \frac{\gamma \delta}{2\pi} \cdot G,$$

where $G$ is the vector corresponding to the direction and magnitude of the diffusion gradient.

This is analogous to the vector

$$k = \frac{\gamma}{2\pi} \cdot \int G(t) dt,$$

which is at the basis of the MRI theory. If the diffusion gradients in a Stejskal-Tanner sequence are infinitely narrow ($\delta \rightarrow 0$), the signal attenuation ratio of the experiment is given by equation. e.g.:

$$\frac{S_{(G,\Delta,\delta)}}{S_0} = \int \rho(x_1) \int P(x_1, x_2, \Delta) \cdot e^{-i \omega (x_2 - x_1)} dx_2$$

where $P(x_1, x_2, \Delta)$ is the conditional probability that a molecule (initially at the origin) is displaced over in a time $t$. By taking the inverse Fourier transform, the probability density function can be recovered.
By sampling the signal attenuation \( \frac{S_{(G,A,\delta)}}{S_0} \) for a series of locations in \( q \)-space, the measured profile can be directly related to molecular displacements due to incoherent random motion without assuming Gaussian behavior.

The actual steps for obtaining QSI metrics by MR diffusion technique are:

- Acquisition of diffusion weighted images with multiple b-values (instead of one b-value of usually 1000 s/mm\(^2\) in DTI); the b-values in QSI should be, for example, 0, 1000, 2000, 3000, ... 11,000 s/mm\(^2\). A q-value signal intensity curve is obtained at each pixel. After the Fourier transformation of the signal decay with respect to q produced a non-Gaussian displacement distribution profile for each pixel in the image, probability density function curves is calculated. The mean displacement (calculated from the full width at half height) and the probability for zero displacement (given by the height of the profile at zero displacement) are important indices. In Fig. 5, the yellow line indicates white matter (WM), the gray line indicates gray matter (GM), and the blue line indicates cerebrospinal fluid (CSF) [6].
This method is theoretically superior to conventional Gaussian diffusion analysis, but few reports apply it to human studies because QSI is time-consuming for daily clinical use, generally taking a minimum of 10 min.

3.3. DIFFUSIONAL KURTOSIS

DKI has a close relationship to q-space imaging and QSI methods have recently been employed to estimate diffusional kurtosis (DK). As it's less demanding from technological point of vue and as acquisition time than other QSI methods, DKI is more used in clinical practice.

The kurtosis is a dimensionless statistical metric for quantifying the non-Gaussianity of an arbitrary probability distribution; kurtosis is actually a measure of a function sharpness (excess). If one uses mathematical moments to characterize the probability distribution shape, kurtosis may be defined as $k = \frac{M_4}{M_2^2} - 3$, where $M_4$ is the 4$^{th}$ moment about its mean and $M_2$ is the second moment about its mean (variance).

For any Gaussian distribution, $K = 0$. If a distribution has less weight on its center and tails compared to a Gaussian with the same variance, then $K < 0$, and if the distribution has more weight on its center and tails, then $K > 0$.

The metrics of kurtosis [8] do not need the full diffusion displacement probability distribution, so the technique is less demanding than QSI in terms of imaging time and gradient strengths. The important point of Jensen's kurtosis [8] is that the excess DK may be approximately determined from just the first 3 terms of an expansion of the logarithm of the MR signal intensity in powers of $b$. This is why DK measurement requires only modest increases in $b$-values beyond those typically employed for DWI making it more accessible for clinical practice.

![Fig. 6 – The kurtosis is a measure of the sharpness of a function and is demonstrated here for functions with varying kurtosis values. The plain curve represents a Gaussian curve, the others are deviations from Gaussian curve.](image-url)
Recently, it has been shown how to estimate, in brain, the kurtosis of the water diffusion displacement PDF with relatively simple diffusion-weighted imaging protocols that are suitable for clinical MRI systems. This method has been referred to as diffusional kurtosis imaging (DKI) and is a natural extension of diffusion tensor imaging (DTI). With DKI, one obtains estimates for all the standard DTI diffusion metrics, such as the mean diffusivity (MD) and the fractional anisotropy (FA), and also for several additional metrics related to the diffusional kurtosis. In this way, DKI provides for a more complete characterization of water diffusion in brain [8].

In general, the measured diffusional kurtosis depends on the direction of the diffusion sensitizing gradients [6]. This dependence on direction can be described by a tensor with 15 independent components; therefore to determine the full diffusional kurtosis tensor, the diffusional kurtosis must be measured in at least 15 different directions. The metric that is usually used is the average value of the directional-dependent kurtosis, i.e., the mean kurtosis (MK). Some report that a minimum of five b values should be used, and, moreover, since DKI is based on a series expansion, it is unsuitable for b values larger than 3000 s/mm² (Fig. 7).

Fig. 7 – Technical requirements for DKI compared to DTI and advantage of DKI in modelling the actual data (copyright license number: 3230741260168, Elsevier).

There are many freely downloadable diffusion MRI software packages (Hassan). Most DTI quantitative packages produce diffusion tensor, FA, eigenvalue and principal eigenvector volumes from the DWI data acquired on each subject. There are also commercial packages with DTI capabilities such as BrainVoyager (http://www.brainvoyager.com/index.html). The free diffusion MRI software packages that have been frequently reported in the literature are: FSL-FDT (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT) (running under Linux) which includes tools for data preprocessing, local diffusion modeling and tractography,
and FSL-TBSS (Tract-Based Statistics) which aims to improve the analysis of multi-subject diffusion imaging studies (group analysis), MedInria (http://med.inria.fr/) (running under both Windows and Linux) with features that offer tensor estimation and tractography, SPM (a platform with Matlab is required) (http://www.fil.ion.ucl.ac.uk/spm/ext/), DTIStudio (https://www.dtistudio.org/) (running under Windows).

4. TRACTOGRAPHY

Brain fiber tractography is a rendering method for improving the depiction of data from diffusion imaging of the brain. The primary purpose of tractography is to clarify the orientational architecture of tissues by integrating pathways of maximum diffusion coherence. Fibers are grown across the brain by following from voxel to voxel the direction of the diffusion maximum. The fibers depicted with tractography are often considered to represent individual axons or nerve fibers, but they are more correctly viewed in physical terms as lines of fast diffusion that follow the local diffusion maxima and that only generally reflect the axonal architecture. Tractography adds information and interest to the MR imaging depiction of the human neuronal anatomy.

The connectivity maps obtained with tractography vary according to the diffusion imaging modality used to obtain the diffusion data.

**Deterministic tractography** or, deterministic streamline tractography. The major eigenvector of the diffusion tensor (also referred to as the principal diffusion direction) is typically assumed to provide a suitable estimate of the fiber orientation within each imaging voxel [7]. The simplest method to obtain an estimate of this orientation at any location is then to use nearest-neighbor interpolation. Deterministic tractography rely on: the identification of a suitable position from which to initiate the algorithm (seed point); the propagation of the track; the termination of the track when appropriate termination criteria are met (the most common such criterion is to impose a threshold based on a measure of diffusion anisotropy (typically FA)). An alternative approach to manual seed point or region selection is to use the so-called “brute-force” approach whereby tracking is initiated from all voxels in the brain.

Diffusion tensor imaging provides a Gaussian approximation of the actual displacement distribution, and since the representation of that distribution is restricted to variations of an ellipsoid, this method creates various biases in the tractography result. In contrast, q-space diffusion imaging with tractography overcomes many of those biases and allows more realistic mapping of connectivity. The tractography result also depends on the tracking algorithm used. Deterministic fiber tracking from diffusion tensor imaging uses the principal direction of diffusion to integrate trajectories over the image but ignores the fact that fiber...
orientation is often undetermined in the diffusion tensor imaging data [9]. To overcome this limitation of the data, Hagmann and colleagues, as well as other investigators, investigated statistical fiber tracking methods based on consideration of the tensor as a probability distribution of fiber orientation [9].

Fig. 8 – Figures showing the fibers within the centrum semiovale of in vivo human brain using q-space imaging (A, C) and DTI (B, D) reconstruction at low and high power magnification. The callosal fibers (red, orange), association fibers (green) and corona radiata fibers (blue) are seen with QSI. With DTI, commissural fibers are greatly reduced and do not reliably penetrate the corona radiata, and such association fibers as are seen have spurious trajectories. Assn, long association fibers; CC, corpus callosum; Cd, caudate nucleus; Comm, commissural fibers; ICp, posterior limb of internal capsule; SB, subcortical bundle projection fibers; Th, thalamus (Copyright license number: 3230730854809, Elsevier).

Probabilistic tractography. Unfortunately, deterministic tractography algorithms only provide a single estimate of the path of white matter fibers from each supplied seed point. Probabilistic approach attempts to address this limitation by providing their results in the form of a probabilistic distribution, rather than a single “best fit” estimate [3]. Probabilistic algorithms attempt to account for crossing fibers, by allowing some spread around the estimated orientations [3]. Consequently, the description of crossing fibers as well as of highly angulated fibers is better in probabilistic tractography obtained from q-space imaging data (figure 9).

5. CONCLUSIONS

The reduction of the anatomical information to a tensor and then to a scalar value implies that when changes or differences are found in one of the scalar metrics, it is often difficult to draw any conclusions about the exact cause at the cellular level. While this can be considered as a drawback of DTI, the systematic information reduction can also be viewed as an advantage: with more than 100
billion neurons, an equal number of axons, 100 trillion synapses, and hundreds of billions of astrocytes, the human brain is a hugely complex system, the complete characterization of which is currently beyond our ability.

When trying to characterize anatomical status and compare it between different populations [10], a quantitative method that can systematically reduce the anatomical information into a manageable size is required, which is exactly what DTI can offer noninvasively and within a short time-frame.

Acknowledgements. This work was supported by UEFISCDI – Unit for Higher Education Research Development and Innovation Funding (grant number 84/2012).

REFERENCES