Probing Tissue Microstructure Using Oscillating Spin Echo Gradients

by

Morgan Elias Mercredi

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Physics and Astronomy
University of Manitoba
Winnipeg

Abstract

The central nervous system (CNS) is made up of neurons and glial cells. Information is transmitted along axons to other neurons, muscles, or glands. Recent studies indicate possible changes in axon diameter distributions associated with diseases such as Alzheimer's disease, autism, dyslexia, and schizophrenia. Magnetic resonance imaging (MRI) techniques such as diffusion MRI can be used to probe the tissue microstructure of the brain noninvasively. Current MRI axon diameter measurements rely on the pulsed gradient spin echo sequence which cannot provide short enough diffusion times to measure small axon diameters. Recent advances have allowed oscillating gradient (OG) diffusion MRI to infer the sizes of micron-scale axon diameters. Monte Carlo simulations of cosine OG sequences were conducted on a parallel cylinder (diameters 1 to 10 μm) geometry. For feasible experiments on a Bruker BG6 gradient set, the simulations inferred diameters as small as 1 μm on square packed and randomly packed cylinders. The accuracy of the inferred diameters was found to be dependent on the signal-to-noise ratio (SNR) with smaller diameters more affected by noise although all diameter distributions were distinguishable from one another for all SNRs tested. Five frequencies were adequate for $d = 3 - 5$ μm with single-sized cylinders and for effective mean axon diameters (*AxD*) greater than 2 μm for cylinders with a distributions of diameters. There was some improvement in precision for $d = 1 - 2$ μm with 10 frequencies. It was better to repeat measurements at higher gradient strengths than to use a range of gradient strengths. Data were collected from a portion of normal-appearing corpus callosum from an autopsy human brain, which did not demonstrate any pathological changes. The average fitted *AxD* was 2.0 ± 0.2 μm, while *AxD* obtained from electron microscopy was 1.4 ± 0.2 μm. Fitted *AxD* showed more variability below 7 OG frequencies and little change when using two or three gradient strengths, agreeing

with the simulations.

Acknowledgements

I would like to thank my advisor, Dr. Melanie Martin for taking me in as a graduate student and for her help and support over the past five years. I would also like to thank my committee members for their suggestions: Dr. Boyd McCurdy, Dr. Chase Figley, Dr. Scott King, and Dr. Jeffrey Dunn.

I would specifically like to thank the following members and former members of Melanie's lab: Sheryl Herrera, Trevor Vincent, and Dr. Richard Buist, the MRI system manager. Sheryl Herrera performed the MRI experiments on the brain sample and prepared the images for analysis. She also measured each axon diameter in the electron microscopy images. Trevor Vincent helped develop the Monte Carlo simulation code and taught me how to use it.

I would like to thank Dr. Kant Matsuda for preparing the brain sample for MRI and electron microscopy. I would also like to thank Dr. Chris Bidinosti for allowing us to use his GPU workstation for our simulations, saving us a lot of time.

Thanks to the staff and faculty in the Department of Physics & Astronomy at the University of Manitoba. I would also like to thank the staff and faculty in the Department of Physics at the University of Winnipeg. They were always willing to help the students in whatever ways they could. There were always research opportunities available each summer. I worked for Dr. Murray Alexander for three summers. My classmates were also very helpful whenever there were problems in school.

I should also thank my parents for putting up with me and supporting me over the years.

Contents

List of Acronyms

MRI	magnetic resonance imaging
CNS	central nervous system
OGSE	oscillating gradient spin echo
PGSE	pulse gradient spin echo
RF	radiofrequency
$\boldsymbol{\mu}$	nuclear magnetic dipole moment
$\boldsymbol{B}$	magnetic field vector
$\boldsymbol{M}$	magnetization vector
M_0	magnetization at equilibrium
ω_0	Larmor frequency
γ	gyromagnetic ratio
T_1	T_1 relaxation time
T_2	T_2 relaxation time
T_E	echo time
T_I	inversion time
S	MRI signal
$P(\boldsymbol{r}',t'\vert\boldsymbol{r},t)$	conditional probability (probability for particle to move from $\boldsymbol{r}$ to $\boldsymbol{r}'$)
$\boldsymbol{R}$	particle displacement vector
n_d	dimensionality (1, 2, or 3)
G	gradient strength
$\boldsymbol{q}$	q-value ($= \gamma\boldsymbol{G}\delta/2\pi$)
G^*	effective gradient strength
Δ	diffusion time (PGSE)
δ	diffusion gradient duration (PGSE)
T	diffusion gradient duration (OGSE)
N_{OG}	number of OGSE oscillations
ω_m	OGSE angular frequency
f_m	OGSE temporal frequency
τ	half-echo time
b	b-value
b_{sin}	b-value (sine-OGSE)
b_{cos}	b-value (cosine-OGSE)
D	diffusion coefficient
D_0	diffusion coefficient (a specific value)
$E(t)$	diffusion echo attenuation
$D(\omega)$	diffusion spectrum
$F(\omega)$	gradient frequency spectrum
ADC	apparent diffusion coefficient

K_{app}	apparent kurtosis
$\mathbf{D}$	diffusion tensor
ADD	axon diameter distribution
E_r	restricted compartment diffusion signal
E_h	hindered compartment diffusion signal
D_h	hindered diffusion coefficient
f_h	hindered compartment fraction
Intra-axonal simulation	a simulation with particles inside the axons
Exra-axonal simulation	a simulation with particles outside the axons
Full simulation	a simulation with particles inside and outside the axons
$w(a, \boldsymbol{p})$	diameter distribution function parameterized by $\boldsymbol{p}$
$(\alpha_\Gamma, \beta_\Gamma)$	gamma distribution parameters (AxCaliber model)
(μ_N, σ_N)	Gaussian distribution parameters (AxCaliber model)
a	axon radius
d	axon diameter
ϕ_{bead}	porosity of bead suspension
λ_τ^2	tortuosity constant
AxD	effective mean axon diameter
D_{in}	intra-axonal diffusion coefficient
D_{ex}	extra-axonal diffusion coefficient
f_{axon}	axon packing fraction
μ_D	mean fitted intra-axonal diffusion coefficient
μ_f	mean fitted axon packing fraction
μ_d	mean fitted axon diameter
μ_{AxD}	mean fitted AxD
σ_D	standard deviation of fitted intra-axonal diffusion coefficients
σ_f	standard deviation of fitted axon packing fractions
σ_d	standard deviation of fitted axon diameters
σ_{AxD}	standard deviation of fitted AxD
G_{max}	maximum gradient strength
G_{high}	highest gradient strength in a particular subset
n_f	number of OGSE frequencies in a data set
S/V	surface to volume ratio
$\mu_{S/V}$	mean fitted surface to volume ratio
$\sigma_{S/V}$	standard deviation of fitted surface to volume ratios

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Purpose(s)

Magnetic Resonance Imaging (MRI) is a powerful tool for neuroscience. It can provide information about the structure of central nervous system (CNS) tissue non-invasively. The resolution of human MRI is typically on the order of 1 mm^3 and for mouse MRI around 100 μm^3 [85]. This resolution is not high enough to see very small structures, such as axons, in the images. Axon diameters are on the order of 1 μm, so a single voxel can contain hundreds of thousands of axons [74]. Thus MRI provides macroscopic tissue information.

Recent advances in the development of the capabilities of MRI have allowed MRI to infer the sizes of compartments in samples that are on the order of microns in diameter. Recently there has been a push in MRI to determine axon diameter distributions and density of axons in fibres within the CNS [7, 11]. The methods are currently used on humans and can infer axon diameters as small as 5 μm [45]. These new developments rely on diffusion-weighted MRI, which is described below.

Mouse axons are much smaller than 1 μm [175] which makes the current method irrelevant for rodent studies. Rodent models of CNS disorders provide researchers with an important tool to study CNS disorders. Because existing methods can infer diameters above 5 μm, this thesis mainly focuses on diameters between 1 to 5 μm. With the method developed it can be applied to rodent models to study the changes in fibre composition due to CNS disorders such

1

as Alzheimer's, schizophrenia, autism, or cancer.

1.2 Thesis Organization

This section gives brief descriptions of the upcoming chapters of this thesis. It provides an overview of the projects and purpose of the work done as part of this thesis. MRI is explained, diffusion is explained, and the projects completed as part of the thesis are explained.

1.2.1 An Introduction to MRI

MRI uses magnetization of hydrogen nuclei when placed in a magnetic field to generated signals and images. Magnetic resonance signals and images in the body depend on microscopic properties of the tissues being imaged. Because of the abundance of water in the human body, soft tissue signal from the body or brain typically comes from the water in the tissue. Differences in water content and tissue types are a source of image contrast (for example cerebrospinal fluid (CSF), grey matter and white matter in the brain). Pulse sequences manipulate hydrogen nuclei signals, making them depend on quantities such as proton density (see Section 2.3) or T_1 and T_2 (see Section 2.1) relaxation times. Basic principles behind MRI are described in Chapter 2.

Quantitative MRI methods combine many types of images to calculate quantities like diffusivity of water, T_1, and T_2 relaxation times, or other tissue properties. We use diffusion MRI because of its ability to probe microscopic length scales. A brief description follows, but the full description of the methods and capabilities of diffusion MRI are given in Chapter 3.

1.2.2 An Introduction to Diffusion

Molecules in a liquid undergo diffusion. Their mean square displacement depends on the diffusion time, Δ, as described by Einstein's relation $\langle x^2 \rangle = 2D\Delta$ (in one dimension), where D is the diffusion coefficient [77]. Molecules diffusing in a uniform medium with no barriers experience unrestricted diffusion. In many media (e.g. porous samples and cellular tissues) barriers, such as cell walls, restrict molecular displacements so that the diffusion depends on

the time scale of the study and the permeability of the barriers. Einstein's relation can be used to find an apparent diffusion coefficient (*ADC*) that is less than *D* because of restrictions. Understanding the influence of the restrictions on the *ADC* gives information about the geometry of the boundaries of the surrounding medium.

Assuming water molecules are diffusing within an impermeable cylinder of radius *a*, one can find the theoretical formula for *ADC* as a function of Δ [10, 15, 127, 128, 184]. If the sample contains many parallel cylinders of different radii, a_i, one can sum together the expected formulae with proper weighted factors to find the expected *ADC* for the sample as a function of Δ. With MRI data, one can use the measured *ADC* as a function of time with the formula to find the radii and weighting factors so that the radii and density of cylinders in the sample can be inferred.

1.2.3 Diffusion MRI

Traditional MR measurements of the *ADC* in different samples use the Pulsed Gradient Spin Echo (PGSE) sequence. After excitation, a magnetic field gradient pulse is applied to the sample for a short time. This causes the magnetic moments of the hydrogen nuclei (or "spins") to obtain a phase based on their position at the time of the pulse. A 180° radiofrequency (RF) pulse is then applied to the system which reverses the phase of the spins. Another identical gradient pulse is applied to the sample which changes the phase of the spins based on their position at the time of the second pulse (Δ). If no diffusion has occurred, the phase acquired from the second pulse will be equal and opposite to the phase of the spin just before the pulse resulting in a net phase of zero. If diffusion occurs, the mean squared phase of all spins will be nonzero and cause a loss in MR signal which can be used to calculate the *ADC*. Figure 1.1 shows a diagram of the spins in a PGSE sequence.

The key to providing a good inference of radii is the choice of the times, Δ, of the measurements. For very short Δ, few water molecules have interacted with the boundaries so the *ADC* is large. As Δ increases slowly, more and more water molecules interact with the boundaries causing a rapid decrease in the *ADC*. Around the time when $\langle x^2 \rangle = a^2$ for a cylinder of radius *a*, most water molecules will have interacted with the barriers at least once [145]. At much
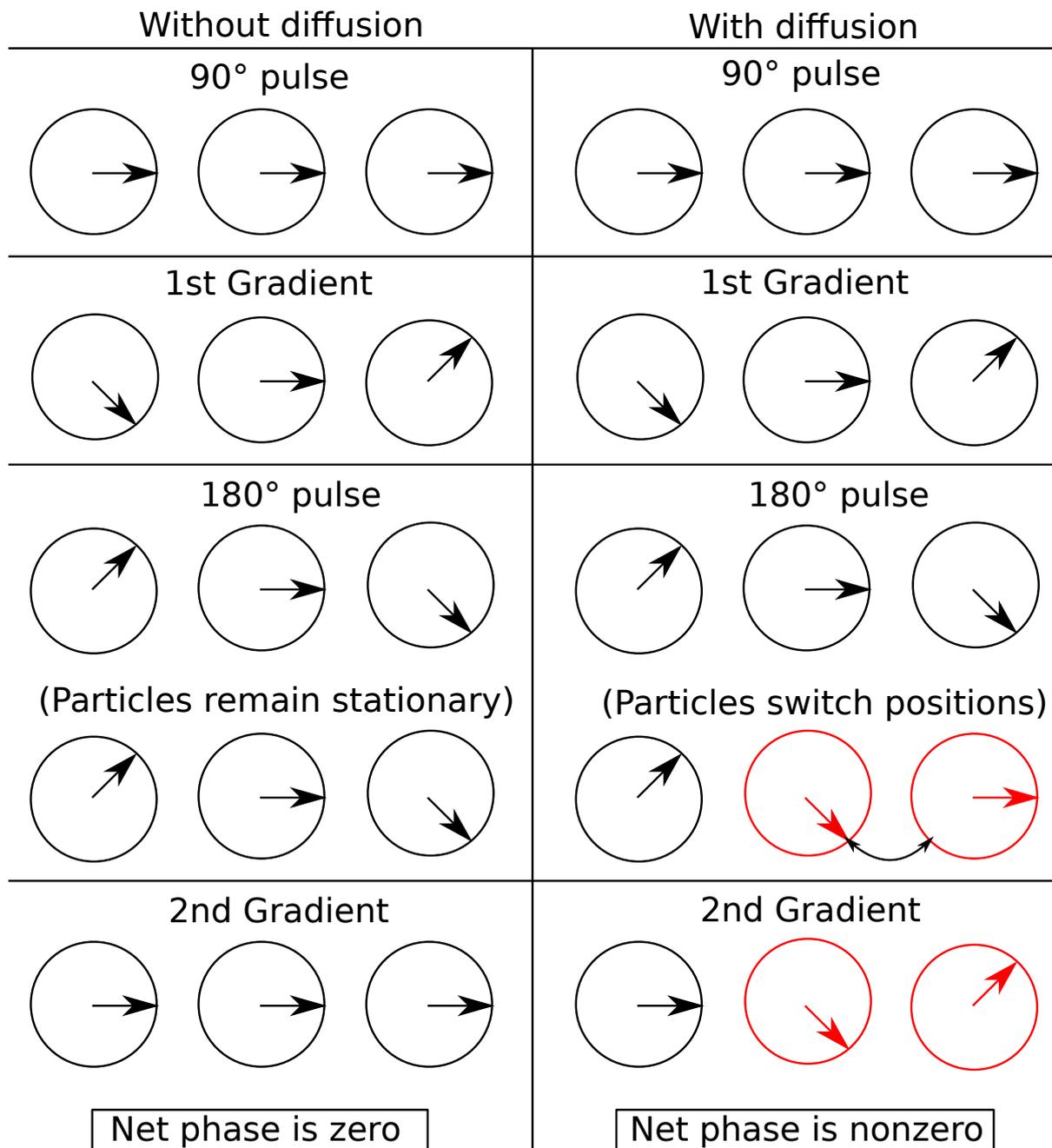
Figure 1.1: After the 90° RF pulse, a rectangular gradient is applied to the sample. This causes the spins to obtain a phase dependent upon their position at the time of the pulse. The 180° pulse reverses the phases of the spins. The second gradient pulse again causes the spins to obtain a phase dependent upon their position. In the case of no diffusion, the net phase of the spins will be zero. If diffusion has occurred, the net phase of the spins is nonzero, leading to a smaller signal.

longer times the *ADC* becomes essentially constant making the inference of the restriction size impossible with this method.

The PGSE sequence requires that Δ be large with respect to the restriction sizes in biological tissues, thus limiting the information that could be obtained from the measurements. For instance, water inside an axon can be approximated as water inside a cylinder in the above model. To infer the diameter and surface to volume ratio of the axon, ideally Δs around $a^2/2D$ need to be used [145]. Unfortunately, the PGSE sequence cannot be used with these short diffusion times because of hardware restrictions such as gradient amplitudes, and timing restrictions, such as RF pulse times.

Using the oscillating gradient spin echo (OGSE) sequences, square gradient pulses are changed to sinusoidal pulses. Each sine wave acts as diffusion-weighting, so with high frequency sine waves, the effective diffusion-time can be made small [35, 43, 58, 81, 106, 139, 140, 143, 144, 154, 190]. Additional sine waves can be used to increase the diffusion-weighting so that the *ADC* can be measured. However, only in the case of unrestricted diffusion can these multiple diffusion-weighting periods be considered independent. Thus, the signal attenuation should be described in terms of a frequency spectrum rather than the diffusion-time. Choosing the appropriate OGSE sequence allows for sampling a narrow frequency domain of the diffusion spectrum and thus provides a straightforward means of characterizing this spectrum. Using temporal diffusion spectroscopy [140], the effective diffusion times are changed by changing gradient frequencies, and a spectrum of diffusion rates can be measured which describe the biological tissue microenvironment [43, 56, 98, 139, 140, 192, 193, 194, 196]. The principles of diffusion MRI are described more fully in Chapter 3.

1.3 Inferring Restriction Sizes Using OGSE MRI

We have developed a Monte Carlo computer simulation of OGSE MRI sequences (described in Section 3.5). The simulations allow us to generate diffusion-weighted signals for different tissue geometries. We found that OGSE frequencies corresponding to scales in the range between completely restricted and free diffusion best distinguish restriction sizes (e.g. axon diameters)

under 10 μm in size. Combining these simulated measurements (or signals) with tissue models such as AxCaliber [11] (described in more detail in Chapter 4) and ActiveAx [6, 7] allows for the inference of restriction sizes in the samples. The AxCaliber model estimates axon diameter distributions and axon density, while the ActiveAx estimates a single axon diameter index, also called the "effective mean axon diameter."

In the AxCaliber model, axons are assumed to consist of non-overlapping parallel cylinders. Experimental (or simulated) MRI signals can then be fit to a theoretical model so that axon distributions can be calculated. In Chapter 4, we perform simulations with a variety of axon diameter distributions and axon densities. Our simulations show that it is possible to infer diameters as small as 1 μm using gradient strengths available on our Bruker BG6 gradient set (maximum gradient strength 1.01 T/m). For now, the method requires the axon orientation to be known, but this requirement may be relaxed in the future. We also tested the model with different levels of signal noise. Much of the material in Chapter 4 has been published as a journal article entitled "Assessing the accuracy of using oscillating gradient spin echo sequences with AxCaliber to infer micron-sized axon diameters" in Magnetic Resonance Materials in Physics, Biology, and Medicine [118].

In Chapter 5, we use a two compartment model on our simulated data (including signal noise). Instead of extracting the full axon distribution, the model extracts an index of axon diameter *AxD* [7]. Because image acquisition can be very time-consuming, we are interested in how data with fewer OGSE frequencies and fewer gradient strengths affects parameter accuracy and precision. We also look at how the model performs with different maximum gradient strengths. Much of the material in Chapter 5 has been published as a journal article entitled "Toward faster inference of micron-scale axon diameters using Monte Carlo simulations" in Magnetic Resonance Materials in Physics, Biology, and Medicine [117].

In Chapter 6, we use an *ADC* model valid at short-times on our simulated data [90, 94, 122, 124, 140, 155, 157, 169, 200]. This model is used to extract surface to volume ratios and diffusion coefficients. We are interested in the accuracy of the parameters, with and without signal noise. We first explore the conditions for acceptable accuracy, then we experiment with using fewer OGSE frequencies and fewer gradient strengths, with an interest in knowing how

removal affects accuracy and precision of parameters.

In Chapter 7, we use a two compartment model on MR data acquired from a portion of human brain. We are interested in how data with fewer OGSE frequencies and fewer gradient strengths affects parameter accuracy and precision.

1.4 Conclusions

Chapter 8 summarizes the thesis and discusses the next steps for the project.

Chapter 2

Nuclear Magnetic Resonance

Nuclear magnetic resonance uses a strong magnetic field and radiofrequency pulses to generate images. A basic outline of its principles is described in this chapter.

2.1 Nuclear Magnetic Resonance

2.1.1 Magnetic Moments

Nuclei possess an intrinsic angular momentum called spin. The existence of spin gives rise to nuclear magnetism. The magnetic field from the spin would be similar to that of a bar magnet, where field lines leave one pole of the proton and return to the other pole. Let $\boldsymbol{\mu}$ be the nuclear magnetic dipole moment and $\boldsymbol{J}$ be the spin angular momentum. Then

$$\boldsymbol{\mu} = \gamma \boldsymbol{J} \tag{2.1}$$

where γ is called the gyromagnetic ratio and depends on the type of nucleus. Sometimes, a quantity $\gamma/2\pi$, having units of MHz/T, is given instead. Typical values of $\gamma/2\pi$ for magnetic resonance are 42.58 MHz/T for hydrogen nuclei (^{1}H) and 10.71 MHz/T for ^{13}C [100].

From quantum mechanics, we know that all the components of $\boldsymbol{\mu}$ cannot be measured simultaneously. The magnitude of the magnetic moment is

$$|\boldsymbol{\mu}| = \gamma \hbar \sqrt{I(I+1)} \tag{2.2}$$

9

where I is the spin quantum number, which depends upon the specific nuclei and $\hbar$ is the reduced Planck's constant ($\hbar = 1.05 \times 10^{-34}$ J·s). The spin quantum number can take on zero, half-integer, or integer values ($I = 0, 1/2, 1, 3/2, \ldots$). There are a few general rules for determining I. Nuclei with odd mass numbers have half-integer spins. Those with even mass number and even charge have $I = 0$. Those with even mass number and odd charge have integer spins [86]. Nuclei such as ^{1}H, ^{13}C, ^{19}F, and ^{31}P are all spin-1/2 systems where $I = 1/2$ [100]. Only nuclei with nonzero spin can be excited during an MR experiment.

Without an external magnetic field, each individual $\boldsymbol{\mu}$ is randomly oriented due to thermal motion. This situation is analogous to compass needles on a vibrating table. Therefore, there is no net macroscopic magnetization. To generate a net magnetization, we need to align the $\boldsymbol{\mu}$ vectors. We use a strong external magnetic field B_0. Assume a static field pointing along the z-axis $\boldsymbol{B} = B_0 \hat{\boldsymbol{k}}$ (the z-axis is chosen by convention). We know from quantum mechanics that only one component of $\boldsymbol{\mu}$ can be measured at a time and that this component is discretized. Assuming that the z-component is measured, we have

$$\mu_z = \gamma m_I \hbar \tag{2.3}$$

where m_I is the magnetic quantum number. For a given I, m_I ranges from $-I$ to I in integer steps ($m_I = -I, -I+1, \ldots, I-1, I$). This means that μ_z can take on $2I+1$ possible values. The angle between $\boldsymbol{B}$ and $\boldsymbol{\mu}$ can be expressed as

$$\cos\theta = \frac{\mu_z}{|\boldsymbol{\mu}|} = \frac{m_I}{\sqrt{I(I+1)}} \tag{2.4}$$

The magnetization is quantized along the direction of the main field. The transverse component of $\boldsymbol{\mu}$ is randomly oriented in the transverse plane. Therefore, $\boldsymbol{\mu}_{xy}$ can be written as $\boldsymbol{\mu}_{xy} = \mu_x \hat{\boldsymbol{i}} + \mu_y \hat{\boldsymbol{j}}$. Now [100],

$$\mu_x = |\mu_{xy}| \cos\xi$$
$$\mu_y = |\mu_{xy}| \sin\xi \tag{2.5}$$

where ξ is a uniform random variable in the interval $[0, 2\pi)$. The magnitude of the transverse

component is

$$|\boldsymbol{\mu}_{xy}| = \sqrt{\mu^2 - \mu_z^2} = \gamma\hbar\sqrt{I(I+1) - m_I^2} \tag{2.6}$$

As an example, we can calculate θ and $|\boldsymbol{\mu}_{xy}|$ for a spin-1/2 system ($I = 1/2$, $m_I = \pm 1/2$). In a spin-1/2 system spin vectors can take on two possible orientations, either parallel or anti-parallel. The angle between $\boldsymbol{\mu}$ and $\boldsymbol{B}$ is $\theta = \pm 54°$, while the magnitude of the transverse component is $|\boldsymbol{\mu}_{xy}| = \gamma\hbar/\sqrt{2}$.

Next, we derive, with a classical treatment, the behaviour of $\boldsymbol{\mu}$ in a static magnetic field $B_0\hat{\boldsymbol{k}}$. The time derivative of the spin angular momentum J is

$$\frac{d\boldsymbol{J}}{dt} = \boldsymbol{\mu} \times B_0\hat{\boldsymbol{k}} \tag{2.7}$$

which can be rewritten as an equation of motion for isolated spins,

$$\frac{d\boldsymbol{\mu}}{dt} = \gamma\boldsymbol{\mu} \times B_0\hat{\boldsymbol{k}} \tag{2.8}$$

using Eq. 2.1. The solution to Eq. 2.8 is

$$\mu_{xy}(t) = \mu_{xy}(0)e^{-i\gamma B_0 t}$$
$$\mu_z(t) = \mu_z(0) \tag{2.9}$$

where the transverse component $\mu_{xy} = \mu_x + i\mu_y = \mu_x\hat{\boldsymbol{i}} + \mu_y\hat{\boldsymbol{j}}$ has been rewritten as a complex exponential. Eq. 2.9 describes a precession of $\boldsymbol{\mu}$ about $\boldsymbol{B}$, a phenomenon called nuclear precession. The precessing magnetization has angular frequency γB_0 and rotates in the clockwise direction. The precession frequency $\omega_0 = \gamma B_0$ is known as the Larmor frequency. Precession could also be described using rotation matrices. If $\mathbf{R} = \mathbf{R}(\theta)$ describes a rotation of θ about the z-axis, then we can write the time dependence of $\boldsymbol{\mu}$ as $\boldsymbol{\mu}(t) = \mathbf{R}(\omega_0 t)\boldsymbol{\mu}(0)$ [100]. A representation of the precession is shown in Figure 2.1.
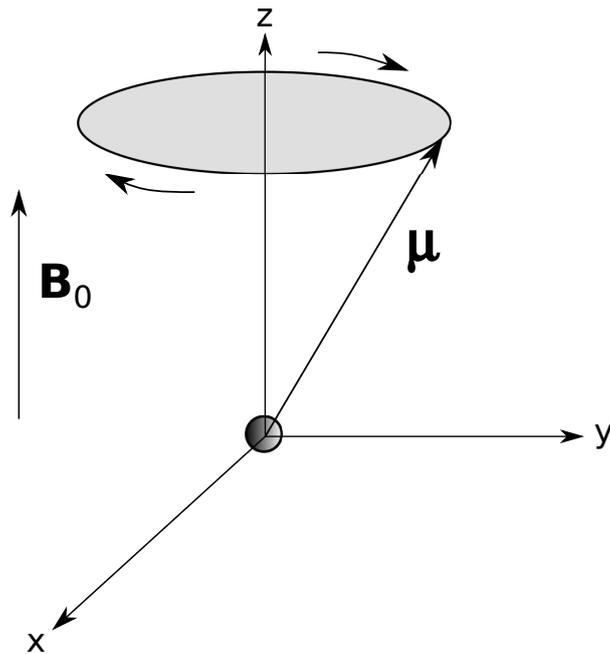
Figure 2.1: Precession of nuclear spin ($\boldsymbol{\mu}$) about an external magnetic field $\boldsymbol{B}_0$.

2.1.2 Bulk Magnetization

We denote the sum of all microscopic magnetic moments in a macroscopic object as the bulk magnetization M. If we have a collection of N_s magnetic moments μ_n, then M is

$$\boldsymbol{M} = \sum_{n=1}^{N_s} \boldsymbol{\mu}_n \tag{2.10}$$

We mentioned in Section 2.1.1 that each $\boldsymbol{\mu}_n$ takes on a different orientation in the B_0-field. The energy of a magnetic dipole in a magnetic field is

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\gamma\hbar m_I B_0 \tag{2.11}$$

There are two energy states, $E_\uparrow$ and $E_\downarrow$, corresponding to the two possible values of m_I ($\pm 1/2$). One is a low energy state and the other is a high energy state, separated by an energy difference

$$\Delta E = E_\downarrow - E_\uparrow = \gamma\hbar B_0 \tag{2.12}$$

The energy separation in a magnetic field is called Zeeman splitting, shown in Figure 2.2.

The spin population difference is related to the energy difference between states and to the
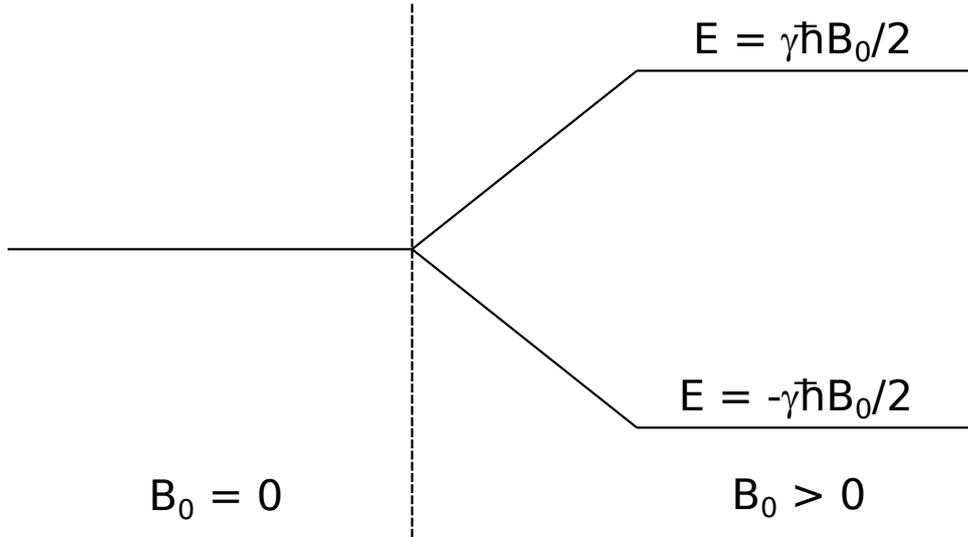
Figure 2.2: Zeeman splitting in a spin-$\frac{1}{2}$ system. When a nonzero magnetic field, $B_0\hat{\boldsymbol{k}}$, is added to a spin-1/2 system two energy levels exist with energies $-\gamma\hbar B_0/2$ and $+\gamma\hbar B_0/2$. The energy difference is $\gamma\hbar B_0$.

temperature of the system (T_s). We can make use of the Boltzmann relation [31],

$$\frac{N_\uparrow}{N_\downarrow} = \exp\left(\frac{\Delta E}{k_B T_s}\right) \tag{2.13}$$

where k_B is the Boltzmann coefficient ($k_B = 1.38 \times 10^{-23}$ J/K). Usually $\Delta E \ll k_B T$, so that we can use small exponent approximation,

$$\frac{N_\uparrow}{N_\downarrow} \approx 1 + \frac{\gamma\hbar B_0}{k_B T_s} \tag{2.14}$$

From here, we can find the population difference

$$N_\uparrow - N_\downarrow \approx N_s \frac{\gamma\hbar B_0}{2k_B T_s} \tag{2.15}$$

There is a small excess of spins in the lower energy state. The lower energy state is more stable, so more spins choose it. Even though ΔN is small, it still generates an observable macroscopic magnetization in the sample. Figure 2.3 shows the basic idea behind spin alignment.

If we expand $\boldsymbol{M}$ into its components, we have

$$\boldsymbol{M} = \sum_{n=1}^{N_s} \boldsymbol{\mu}_{x,n} + \sum_{n=1}^{N_s} \boldsymbol{\mu}_{y,n} + \sum_{n=1}^{N_s} \boldsymbol{\mu}_{z,n} \tag{2.16}$$

13

a

B = 0

b

B > 0

M > 0
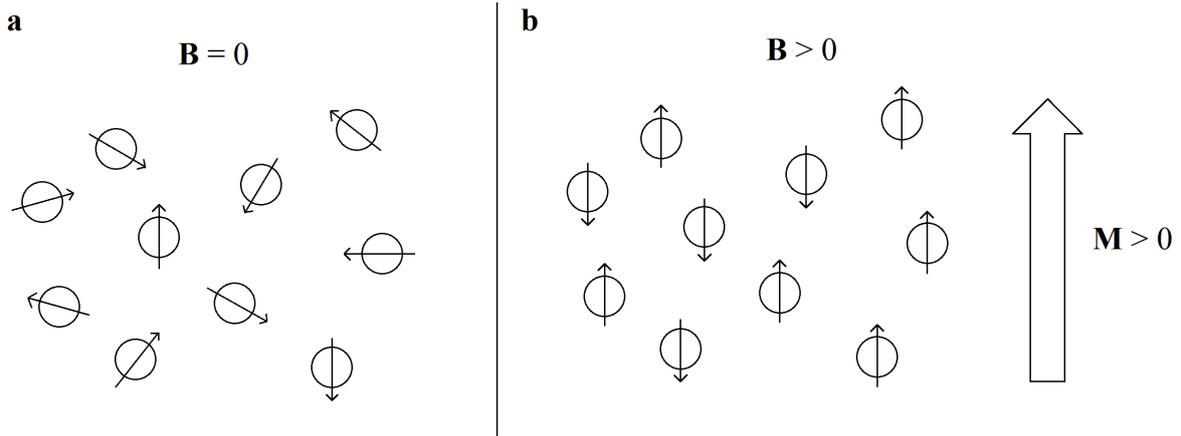
Figure 2.3: Nuclear magnetic moment vectors (a) randomly oriented (with no field) (b) aligned parallel or anti-parallel to an external magnetic field. Without a magnetic field the net magnetization is zero. With a magnetic field, the net magnetization is nonzero and points in the same direction as the field.

where $\boldsymbol{\mu}_{x,n}$, $\boldsymbol{\mu}_{y,n}$, and $\boldsymbol{\mu}_{z,n}$ are the projections of $\boldsymbol{\mu}_n$ along each axis. we know that $\boldsymbol{\mu}_{z,n}$ is $\pm\gamma\hbar/2$, depending on the whether $\boldsymbol{\mu}$ is parallel or anti-parallel. The sum of the x- and y-components of $\boldsymbol{\mu}_n$ are zero because the individual $\boldsymbol{\mu}_{xy}$ (transverse components) have random phases.

In terms of the population difference ΔN, the magnetization is

$$\boldsymbol{M} = \left(\sum_{n=1}^{N_\uparrow} \tfrac{1}{2}\gamma\hbar - \sum_{n=1}^{N_\downarrow} \tfrac{1}{2}\gamma\hbar \right) = \frac{1}{2}(N_\uparrow - N_\downarrow) \tag{2.17}$$

At equilibrium it points in the z-direction. Its magnitude is

$$M_z^0 = |\boldsymbol{M}| = \frac{\gamma^2\hbar^2 B_0 N_s}{4k_B T_s} \tag{2.18}$$

It is linearly proportional to the field strength and to the total number of spins in the system. It is inversely proportional to the temperature. In an object, the number of spins available is fixed, so one can vary B_0 or T_s to change the magnetization. MRI experiments are typically going to be performed at room temperature (or at body temperature), so that T_s is fixed. The only way to increase the magnetization is to increase the field strength.

On clinical scanners, B_0 typically ranges from 0.5 T to about 3 T [14]. As an example of the magnitude of the quantities, assume a magnetic field of 1 T. The population difference as a fraction of the total number of spins is $\Delta N/N_s \approx 3$ per million, so approximately 3 protons per

million are available for generating a MR signal [82]. At 3 T, the proportion would be three times greater (i.e. approximately 9 protons per million).

The Larmor frequency is the natural frequency of the spin system. The angular frequency depends on B_0 and gyromagnetic ratio of the material, so each species of nuclei will have its own Larmor frequency. For example, ^{1}H has $\omega_0/2\pi$ = 42.58 MHz in a 1 T field [100]. This makes it possible to choose one species of nuclei to image without disturbing the others.

2.1.3 Magnetic Field Inhomogeneities

Many spin systems contain a range of resonance frequencies. This might happen if there are inhomogeneities in the applied magnetic field. When the B_0-field is inhomogeneous, then the Larmor frequency is spatially dependent. We call a group of nuclear spins with the same resonance frequency an isochromat. Another possible reason for a range of Larmor frequencies is due to the chemical shift effect which occurs because nuclei in different chemical environments resonate at slightly different frequencies [96]. The surrounding electrons shield the nucleus, so that it experiences an effective field

$$\hat{B}_0 = B_0(1 - \delta) \tag{2.19}$$

where δ is called the shielding constant. Its resonance frequency is shifted by $\Delta\omega = \delta\omega_0$. Chemical shift means that even if B_0 is homogeneous, there may be a range of resonance frequencies present in the sample. The exact value of δ depends on the chemical environment near the nucleus. For example, some hydrogen nuclei in fat (CH_2), have a chemical shift of around 3.35 ppm from water (H_2O) [100]. In biological samples, we could have a range of δ values. Suppose that the maximum chemical shift in a sample is $\omega_c/2$, where ω_c is called the chemical shift frequency bandwidth. We could express the range of resonance frequencies as $|\omega - \omega_0| < \omega_c/2$. Chemical shifts are used to identify molecular structure using nuclear magnetic spectroscopy.

2.1.4 RF Excitation

An applied magnetic field $\boldsymbol{B}$ generates a net magnetization $\boldsymbol{M}$ in an object that points along the direction of the field. Each individual microscopic moment has a transverse component, but because they have random phases, there is no macroscopic transverse component. Resonance refers to the establishment of phase coherence among the spins. One way to think about resonance is to imagine children on a set of identical swings at a playground. Usually, the children are out of phase with each other. If an external force is applied at the natural (or resonant) frequency of the swings, then eventually the children will be in phase with each other. The situation in MRI is similar. We need to apply some kind of external force to the spins to make them go in phase. For a magnetized system, we use a small oscillating magnetic field $B_1(t)$ to generate phase coherence. The B_1-field has the same frequency as the spin-system.

To describe the resonance phenomenon fully, we need a quantum mechanical model. According to Planck's law, the energy of electromagnetic radiation with frequency ω_{rf} is $E = \hbar\omega_{\text{rf}}$. To generate a coherent transition of spins from one energy state to another, the radiation energy needs to be the same as the energy difference between states:

$$\hbar\omega_{\text{rf}} = \Delta E = \gamma\hbar B_0 \tag{2.20}$$

Therefore, the resonance condition requires that $\omega_{\text{rf}} = \omega_0$.

2.1.5 RF Pulses

The B_1-field is of short duration (ms or μs) and in the radio-frequency range. It is also much weaker than the main field (e.g. $B_1 = 50$ mT vs $B_0 = 1$ T [100]). A typical B_1-field could be described by

$$\boldsymbol{B}_1(t) = 2B_1^e(t)\cos\left(\omega_{rf}t + \phi\right)\hat{\boldsymbol{i}} \tag{2.21}$$

where $B_1^e(t)$ is called the envelope function and ϕ is a phase angle [100]. This choice of $B_1(t)$ is said to be linearly polarized along the x-axis because it oscillates in the x-direction. Eq. 2.21 can be decomposed into two circularly polarized fields, one traveling clockwise and the other

counterclockwise:

$$\boldsymbol{B}_1(t) = B_1^e(t)[\cos(\omega_{rf}t + \phi)\hat{\boldsymbol{i}} - \sin(\omega_{rf}t + \phi)\hat{\boldsymbol{j}}]$$
$$+ B_1^e(t)[\cos(\omega_{rf}t + \phi)\hat{\boldsymbol{i}} + \sin(\omega_{rf}t + \phi)\hat{\boldsymbol{j}}] \tag{2.22}$$

The counterclockwise component rotates in the opposite direction as the spins and has little effect on the system, so we can just consider the clockwise component:

$$B_1(t) = B_1^e(t)[\cos(\omega_{rf}t + \phi)\hat{\boldsymbol{i}} - \sin(\omega_{rf}t + \phi)\hat{\boldsymbol{j}}] \tag{2.23}$$

Note that this can be written in complex form:

$$B_1(t) = B_1^e(t)e^{-i(\omega_{rf}t + \phi)} \tag{2.24}$$

Since ϕ has no real effect on the excitation result, we take it to be zero from now on. The RF frequency ω_{rf} is usually constant over time and is chosen based on the resonance condition. The envelope function is important because it determines the shape and duration of the pulse. RF pulses are usually named based on the shape of $B_1^e(t)$ (e.g. whether it be a sinc function, a rectangular function, etc). More discussion as to why the shape of the RF pulse is important can be found in Section 2.3.

2.1.6 Rotating Frame

It is easier to describe the excitation process using a rotating frame of reference that rotates at angular frequency ω about the z-axis and has orthogonal coordinate axes x', y', z'. The unit vectors $\hat{\boldsymbol{i}}'$, $\hat{\boldsymbol{j}}'$, $\hat{\boldsymbol{k}}'$ in the rotating frame can be related to the stationary unit vectors [100]:

$$\hat{\boldsymbol{i}}' = \cos(\omega t)\hat{\boldsymbol{i}} - \sin(\omega t)\hat{\boldsymbol{j}}$$
$$\hat{\boldsymbol{j}}' = \sin(\omega t)\hat{\boldsymbol{i}} + \cos(\omega t)\hat{\boldsymbol{j}} \tag{2.25}$$
$$\hat{\boldsymbol{k}}' = \hat{\boldsymbol{k}}$$

There are two important rotating frames in MR. One is the frame where $\omega = \omega_0$, called the Larmor rotating frame. The other is the frame where $\omega = \omega_{rf}$, called the RF rotating frame. If $\omega_{rf} = \omega_0 = \gamma B_0$, then both frames coincide. The magnetization in the rotating frame ($\boldsymbol{M}_{rot}$) is related to the magnetization of the laboratory frame through the following equation:

$$
\begin{bmatrix} M_{x'} \\ M_{y'} \\ M_{z'} \end{bmatrix} = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) & 0 \\ \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} \tag{2.26}
$$

where $M_{x'}$, $M_{y'}$, and $M_{z'}$ are components of $\boldsymbol{M}_{rot}$. Similarly, the RF field in the rotating frame ($\boldsymbol{B}_{1,rot}$) is related to the RF field in the laboratory frame through

$$
\begin{bmatrix} B_{1,x'} \\ B_{1,y'} \end{bmatrix} = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix} \begin{bmatrix} B_{1,x} \\ B_{1,y} \end{bmatrix} \tag{2.27}
$$

where $B_{1,x'}$ and $B_{1,y'}$ are components of $\boldsymbol{B}_{1,rot}$. We could also rewrite Eqs. 2.26 and 2.27 using complex notation. For the B_1-field, we have

$$
B_{1,rot} = B_1(t)e^{i\omega t} \tag{2.28}
$$

where $B_{1,rot}(t) = B_{1,x'} + iB'_{1,y'}$ and $B_1(t) = B_{1,x}(t) + iB_{1,y}(t)$. Similarly,

$$
M_{x'y'} = M_{xy}(t)e^{i\omega t} \tag{2.29}
$$

where $M_{x'y'} = M_{x'} + iM_{y'}$ and $M_{x'y'} = M_x + iM_y$. For example, suppose that

$$
\boldsymbol{B}_1(t) = B_1(t)\cos(\omega_{rf}t)\hat{\boldsymbol{i}} - B_1(t)\sin(\omega_{rf}t)\hat{\boldsymbol{j}} \tag{2.30}
$$

Using Eqs. 2.26 and 2.27, we can show that the B_1-field in the RF rotating frame is $B_1(t)\hat{\boldsymbol{i}'}$ [23]. Here, the B_1-field points along the x'-axis in the RF rotating frame.

We can also find the rate of change of $\boldsymbol{M}$ observed in the rotating frame, given that $d\boldsymbol{M}/dt$ is the rate of change observed in the laboratory frame. It can be shown that the equation of

motion becomes

$$\frac{d\boldsymbol{M}}{dt} = \frac{\partial \boldsymbol{M}_{\text{rot}}}{\partial t} + \boldsymbol{\omega} \times \boldsymbol{M}_{\text{rot}} \tag{2.31}$$

where $\partial \boldsymbol{M}_{\text{rot}}/\partial t$ is the rate of change of $\boldsymbol{M}$ in the rotating frame [100]. Switching to the laboratory frame introduces a cross term involving $\boldsymbol{\omega}$ and $\boldsymbol{M}_{\text{rot}}$.

2.1.7 Bloch Equation

The time-dependent behaviour of $\boldsymbol{M}$ in an applied magnetic field is described by the Bloch equation [25],

$$\frac{d\boldsymbol{M}}{dt} = \gamma \boldsymbol{M} \times \boldsymbol{B} - \frac{M_x\hat{\boldsymbol{i}} + M_y\hat{\boldsymbol{j}}}{T_2} - \frac{(M_z - M_z^0)\hat{\boldsymbol{k}}}{T_1} \tag{2.32}$$

where $M_z^0\hat{\boldsymbol{k}}$ is the equilibrium magnitude of $\boldsymbol{M}$ and T_1 and T_2 are time constants describing relaxation of the system after excitation (described in Section 2.1.10). If the RF pulse duration is much less than T_1 and T_2, then we can neglect the final two terms when modeling RF excitation [100]. The simplified Bloch equation becomes

$$\frac{d\boldsymbol{M}}{dt} = \gamma \boldsymbol{M} \times \boldsymbol{B} \tag{2.33}$$

In the rotating frame, the Bloch equation is

$$\begin{aligned}
\frac{\partial \boldsymbol{M}_{\text{rot}}}{\partial t} &= \gamma \boldsymbol{M}_{\text{rot}} \times \boldsymbol{B}_{rot} - \boldsymbol{\omega} \times \boldsymbol{M}_{\text{rot}} \\
&= \gamma \boldsymbol{M}_{\text{rot}} \times \left(\boldsymbol{B}_{rot} + \frac{\boldsymbol{\omega}}{\gamma} \right) \\
&= \gamma \boldsymbol{M}_{\text{rot}} \times \boldsymbol{B}_{eff}
\end{aligned} \tag{2.34}$$

where

$$\boldsymbol{B}_{eff} = \boldsymbol{B}_{rot} + \frac{\boldsymbol{\omega}}{\gamma} \tag{2.35}$$

is the effective magnetic field experienced by the magnetization in the rotating frame. Suppose that the rotating frame rotates at the Larmor frequency, $\boldsymbol{\omega} = -\gamma B_0\hat{\boldsymbol{k}}$, in a static magnetic field $\boldsymbol{B} = B_0\hat{\boldsymbol{k}}$. According to Eq. 2.35, $\boldsymbol{B}_{\text{eff}} = 0$. Since there is no effective longitudinal field in the rotating frame, $\boldsymbol{M}_{\text{rot}}$ does not precess.

2.1.8   On-resonance excitation

We want to know the behaviour of $\boldsymbol{M}$ during excitation. For simplicity, we assume a spin system with one isochromat of resonance frequency $\omega_0 = \gamma B_0$. The RF pulse in the rotating frame is

$$\boldsymbol{B}_{1,rot} = B_1^e(t)\hat{\boldsymbol{i}}'$$

(2.36)

The effective magnetic field in the rotating frame is

$$\boldsymbol{B}_{\text{eff}} = \left(B_0 - \frac{\omega_{rf}}{\gamma} \right)\hat{\boldsymbol{k}}' + B_1^e(t)\hat{\boldsymbol{i}}'$$

(2.37)

When the RF frequency is at resonance, $\omega_{\text{rf}} = \omega_0$, the z'-component of the B-field vanishes and all that remains is the RF field component $B_1^e(t)\hat{\boldsymbol{i}}'$. To find $\boldsymbol{M}_{\text{rot}}$, we use the Bloch equation in Eq. 2.34:

$$\frac{\partial \boldsymbol{M}_{\text{rot}}}{\partial t} = \gamma \boldsymbol{M}_{\text{rot}} \times B_1^e(t)\hat{\boldsymbol{i}}'$$

(2.38)

The general solution to Eq. 2.38 is [100]

$$\begin{aligned}
M_{x'}(t) &= 0 \\
M_{y'}(t) &= M_z^0 \sin\left(\int_0^t \gamma B_1^e(t')dt' \right) \\
M_{z'}(t) &= M_z^0 \cos\left(\int_0^t \gamma B_1^e(t')dt' \right)
\end{aligned}$$

(2.39)

showing that $\boldsymbol{M}_{\text{rot}}$ precesses about the x'-axis, which makes sense, because the RF field points along the x'-axis. Figure 2.4 shows how a $\boldsymbol{B}_1$-field causes excitation. As an example, suppose the RF envelope is a rectangular function with amplitude B_1. The equation for the magnetization (Eq. 2.39) during excitation becomes

$$\begin{aligned}
M_{x'}(t) &= 0 \\
M_{y'}(t) &= M_z^0 \sin(\omega_1 t) \\
M_{z'}(t) &= M_z^0 \cos(\omega_1 t)
\end{aligned}$$

(2.40)

where $\omega_1 = \gamma B_1$. The magnetization precesses at frequency ω_1 about the x'-axis. In the laboratory frame, $\boldsymbol{M}$ spirals down to the transverse plane [29].
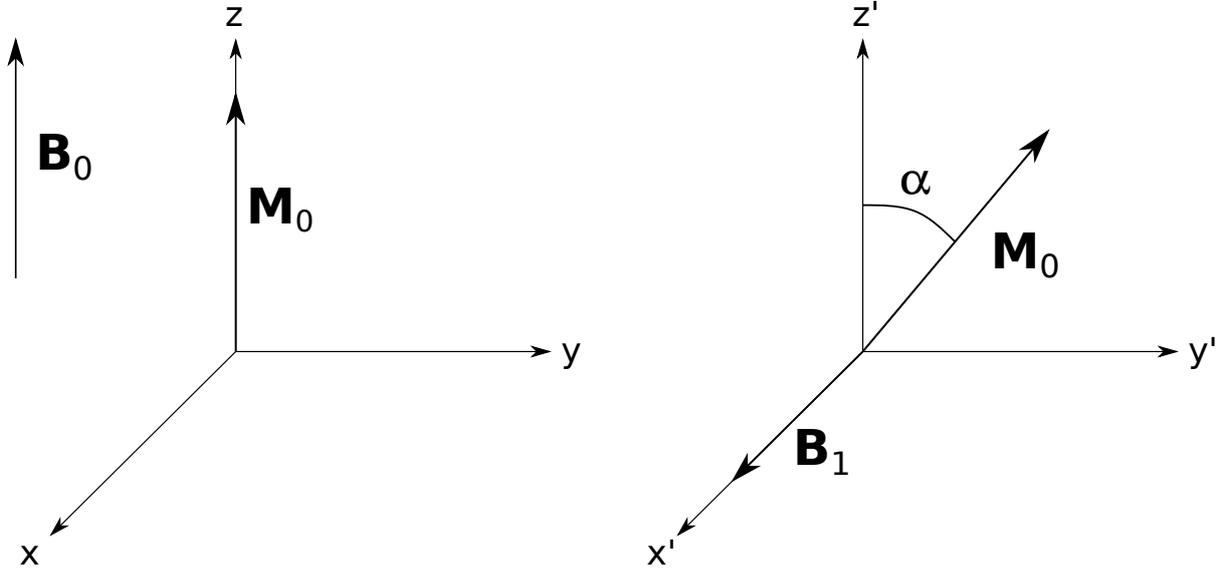
Figure 2.4: (a) Bulk magnetization $\boldsymbol{M}_0$ in the laboratory frame with a static magnetic field $\boldsymbol{B}_0$. (b) Bulk magnetization $\boldsymbol{M}_0$ in the RF rotating frame during an RF pulse $\boldsymbol{B}_1(t)$. The RF pulse, applied in the x'-direction, causes $\boldsymbol{M}_0$ to rotate through an angle α towards the y'-axis.

2.1.9 Frequency Spectrum of RF Pulse

We want to know how a pulse of the form $B_1^e(t)e^{-i\omega_{\mathrm{rf}}t}$ affects a system with more than one isochromat. Let $\{\mathcal{F}B_1^e\}(\omega)$ denote the Fourier transform of $B_1^e(t)$. We can write the RF pulse as [100]

$$B_1(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\{\mathcal{F}B_1^e\}(\omega)e^{-i(\omega+\omega_{\mathrm{rf}})t}d\omega \tag{2.41}$$

We can think of Eq. 2.41 as decomposing $B_1(t)$ into vectors of amplitude $\{\mathcal{F}B_1^e\}(\omega)d\omega$ and frequency $\omega + \omega_{\mathrm{rf}}$. In the linear system assumption, we assume that the total excitation effect of $B_1(t)$ is equal to the sum of excitation effects of its components $\{\mathcal{F}B_1^e\}(\omega)d\omega e^{-i(\omega+\omega_{\mathrm{rf}})t}$. Ideally, each component excites an isochromat of resonant frequency $\omega + \omega_{\mathrm{rf}}$. In reality, there will be a phase dispersal among isochromats along the transverse plane. Each isochromat also goes through a different flip angle. Both properties are controlled by the specific form of $\{\mathcal{F}B_1^e\}(\omega)$ [100].

There are two types of RF pulses. Hard pulses are short (~ 1 μs) pulses meant to excite

as many spins as possible. Soft pulses are longer pulses meant to excite a small range of frequencies in the sample. Soft pulses are important for slice selection in imaging.

2.1.10 Relaxation

After being disturbed, the spins eventually return to equilibrium. During this process, the spins undergo free precession, where $\boldsymbol{M}$ precesses about the B_0-field. Longitudinal relaxation describes the return of M_z to its equilibrium value M_0, while transverse relaxation describes the destruction of M_{xy}. Both relaxation processes are due to the existence of microscopic time-dependent magnetic fields surrounding the nuclei, though the exact mechanism is complicated [2, 26, 62]. The Bloch equation describing the process is

$$\begin{aligned}
\frac{dM_{z'}}{dt} &= -\frac{M_{z'} - M_z^0}{T_1} \\
\frac{dM_{x'y'}}{dt} &= -\frac{M_{x'y'}}{T_2}
\end{aligned} \tag{2.42}$$

The solutions to Eq. 2.42 are

$$\begin{aligned}
M_{x'y'}(t) &= M_{x'y'}(0) \exp\left(-t/T_2\right) \\
M_{z'}(t) &= M_z^0(1 - \exp\left(-t/T_1\right)) + M_{z'}(0) \exp\left(-t/T_1\right)
\end{aligned} \tag{2.43}$$

where $M_{x'y'}(0)$ and $M_{z'}(0)$ are transverse and longitudinal magnetizations immediately after an RF pulse (at $t = 0$). Figure 2.5 shows relaxation curves after a $90°$ pulse. This description applies only to weak-spin interactions like those found in liquids, which include most biological applications of MRI [31]. At $t = T_1$, $M_z(t)$ has grown to 63% of its equilibrium value. At $t = T_2$, M_{xy} has decayed to 37% its initial value. Usually, T_1 ranges from 300 to 2000 ms, while T_2 typically ranges from 30 to 300 ms [100]. If we transformed Eq. 2.43 to the laboratory frame, we would find that the transverse component precesses about the B_0-field while undergoing exponential decay. The total bulk magnetization spirals back up to the z-axis (but its magnitude is not preserved). Since T_2 is on the order of tens of milliseconds in biological tissues, the free precession period lasts long enough for MR signals to be detected.
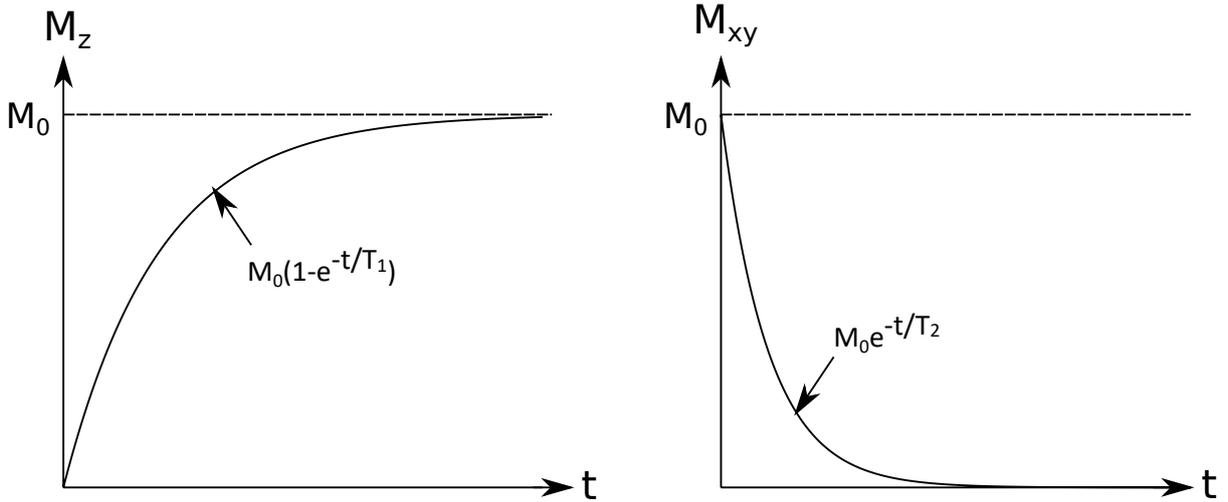
Figure 2.5: Relaxation curves showing (a) recovery of longitudinal magnetization, and (b) decay of transverse magnetization after a 90° pulse. The longitudinal magnetization exponentially grows to its equilibrium value M_0 while the transverse magnetization exponentially decays to zero.

2.2 Magnetic Resonance Signals

In this section, the detection of MR signals (2.2.1) will be discussed. Section 2.2.2 discusses how the signal relates to the density of spins in the sample. The section concludes with discussion of different ways MR signals can be created, such as free induction decay (2.2.3), spin echo (2.2.4), and gradient echo (2.2.5).

2.2.1 Signal Detection

MR signal detection depends on Faraday's law of induction and the principle of reciprocity. Faraday's law states that a time varying magnetic flux through a receiver coil induces an electromotive force (emf) proportional to the rate of change of that flux. Let $\boldsymbol{B}_r(\boldsymbol{r})$ be the laboratory frame magnetic field generated by a unit current in the coil. The flux through the coil due to a bulk magnetization $\boldsymbol{M}(\boldsymbol{r},t)$ is

$$\Phi(t) = \int_{object} \boldsymbol{B}_r(\boldsymbol{r}) \cdot \boldsymbol{M}(\boldsymbol{r},t) d\boldsymbol{r} \tag{2.44}$$

23

By Faraday's law, the voltage induced in the receiver coil is

$$V(t) = -\frac{\partial \Phi(t)}{\partial t} = -\frac{\partial}{\partial t} \int_{object} \boldsymbol{B}_r(\boldsymbol{r}) \cdot \boldsymbol{M}(\boldsymbol{r},t) d\boldsymbol{r} \tag{2.45}$$

We can make a few assumptions to simplify Eq. 2.45 [100]. We assume that $\boldsymbol{M}(\boldsymbol{r},t)$ is undergoing free precession with frequency $\omega(\boldsymbol{r})$. The frequency $\omega(\boldsymbol{r})$ is made up of a Larmor component ω_0 and a spatially dependent offset $\Delta\omega(\boldsymbol{r})$. The offset could be caused by an inhomogeneous field. Since $M_z(\boldsymbol{r},t)$ is slow varying compared to $M_x(\boldsymbol{r},t)$ and $M_y(\boldsymbol{r},t)$, we can neglect it. We assume free precession occurs at a much faster rate than T_2 relaxation. Normally, $V(t)$ is a high frequency signal because the transverse magnetization precesses near the Larmor frequency. Signal demodulation is used to remove the Larmor component, leaving only the low frequency component $\Delta\omega(\boldsymbol{r})$. Signal demodulation involves multiplying $V(t)$ by a reference signal with frequency ω_0 and then applying a low-pass filter [31].

Usually, the signal is acquired using quadrature detection, which uses two orthogonal detectors to measure the in-phase and out-of-phase components of $V(t)$ [100]. The signal can be written in complex notation, $S(t) = S_r(t) + iS_i(t)$, where S_r and S_i are the two orthogonal components that are measured. If one assumes a homogeneous B_r-field over the region of interest, then $S(t)$ is

$$S(t) \propto \int_{object} M_{xy}(\boldsymbol{r},0) e^{-i\Delta\omega(\boldsymbol{r})t} d\boldsymbol{r} \tag{2.46}$$

where $M_{xy}(\boldsymbol{r},0)$ is the transverse component of the magnetization ($M_{xy} = M_x + iM_y$). If the offset $\Delta\omega(\boldsymbol{r})$ was caused by an inhomogeneous applied field, we could let $\Delta\omega = \gamma\Delta B(\boldsymbol{r})$, where $\Delta B(\boldsymbol{r})$ describes the spatially varying field inhomogeneity.

2.2.2 Spin Spectral Density Function

The spin density spectral function $\rho(\omega)$ describes the frequency distribution of a spin system. It is related to the magnetization density as $dM(\omega) = \rho(\omega)d\omega$. Integration of $dM(\omega)$ over all frequencies gives the total bulk magnetization of the sample, $\int \rho(\omega)d\omega = M$. In terms of $\rho(\omega)$, the MR signal is

$$S(t) = \int_{-\infty}^{\infty} \rho(\omega) \exp(-t/T_2(\omega)) \exp(-i\omega t) d\omega \tag{2.47}$$

If one neglects T_2 relaxation (or takes T_2 to be infinite), then $\rho(\omega)$ is the frequency spectrum of the signal $S(t)$. Suppose we have a spin system with N isochromats, each with distinct frequency ω_n. The spectral density function is

$$\rho(\omega) = \sum_{n=1}^{N} M_{z,n}^0 \delta(\omega - \omega_n) \tag{2.48}$$

where $M_{z,n}$ is the equilibrium magnetization for an isochromat with frequency ω_n. Another possible spectral density is a Lorentzian distribution, which occurs with inhomogeneous fields [100].

2.2.3 Free Induction Decays

The free induction decay (FID) is the signal generated by free precession of a bulk magnetization $\boldsymbol{M}$ about the B_0-field. It is the most basic form of transient signal from a spin system after excitation. For a flip angle α, the FID signal is

$$S(t) = \sin(\alpha) \int_{-\infty}^{\infty} \rho(\omega) \exp\left(-t/T_2(\omega)\right) \exp\left(-i\omega t\right) d\omega \tag{2.49}$$

The maximum amplitude occurs at $t = 0$,

$$A_f = \sin(\alpha) \int_{-\infty}^{\infty} \rho(\omega) d\omega = M_z^0 \sin(\alpha) \tag{2.50}$$

and depends on the flip angle and the equilibrium magnetization value. The spectral distribution affects the exact form of the signal. For a system with only one spectral component at ω_0, the spectral density is $\rho(\omega) = M_z^0 \delta(\omega - \omega_0)$ and the FID is

$$S(t) = M_z^0 \sin(\alpha) \exp\left(-t/T_2\right) \exp\left(-i\omega_0 t\right) \tag{2.51}$$

For one spectral component and a uniform field, the FID decays exponentially with T_2. Inhomogeneous fields lead to faster decay because different isochromats quickly go out of phase, shrinking the bulk magnetization. Signal decay due to field inhomogeneity is characterized

by a time constant T_2^*. Assuming that the field inhomogeneity causes a Lorentzian spectral distribution, then signal decay is approximately

$$S(t) \propto \exp\left(-t/T_2^*\right) \tag{2.52}$$

where

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \Delta B \tag{2.53}$$

and ΔB measures the spread in magnetic field values over the sample. If the spectral density is not Lorentzian, then Eq. 2.53, which describe simple T_2^* decay, does not strictly hold [23]. Figure 2.6 shows a FID with time constant T_2^*.
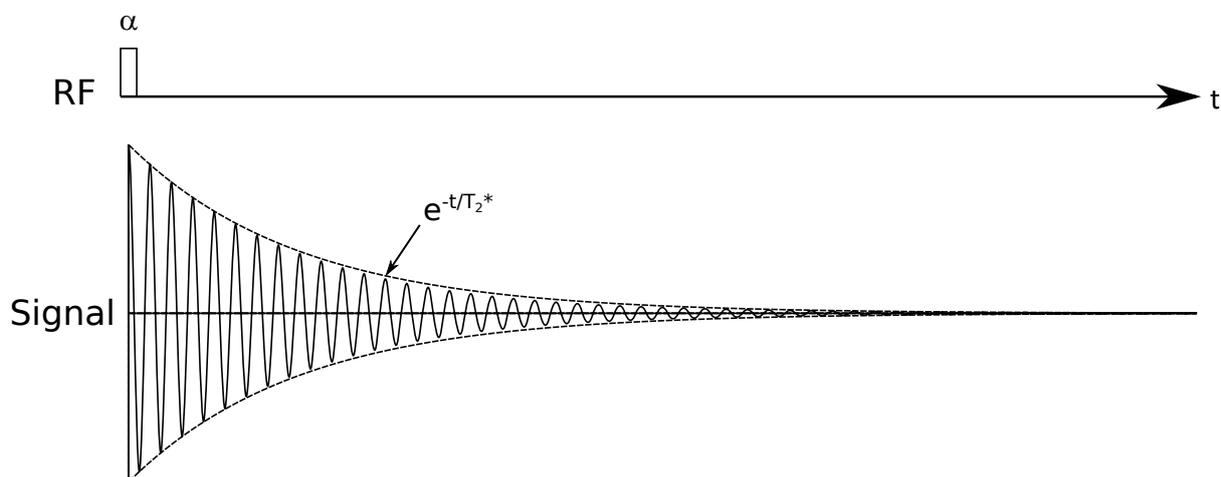
Figure 2.6: Free induction decay after an α pulse. The FID decays with approximate time constant T_2^*. As it decays, the FID oscillates with frequency ω_0.

2.2.4 Spin Echoes

When spins are dephased and then subsequently rephased, an RF echo is generated. It is a two-sided signal. It was discovered by E. Hahn in 1950 and has been widely used in MRI experiments ever since [59, 60].

To generate an RF echo, we need two or more RF pulses. The simplest sequence uses a $90°$ excitation pulse and a $180°$ inversion pulse separated by a time delay τ ($90°$–τ–$180°$). This sequence generates a spin echo. To illustrate how it works, suppose we have two isochromats, like in Figure 2.7. One is "slow," having frequency ω_s. The other is "fast," having frequency

ω_f. We apply a 90° RF pulse along the x'-axis. After excitation, both isochromats precess about the z-axis, with one (the "fast" one) precessing slightly faster than the other. Since they precess at different frequencies, they gradually lose phase coherence. At time τ, the phase difference is $(\omega_f - \omega_s)\tau$. We then apply a 180° RF pulse along the y'-axis. Both vectors flip over the transverse plane, but now the faster vector (ω_f) is behind slower vector (ω_s). During the next time period τ, the faster vector catches up to the slower vector. At $t = 2\tau$, called the echo time, the spins are rephased.
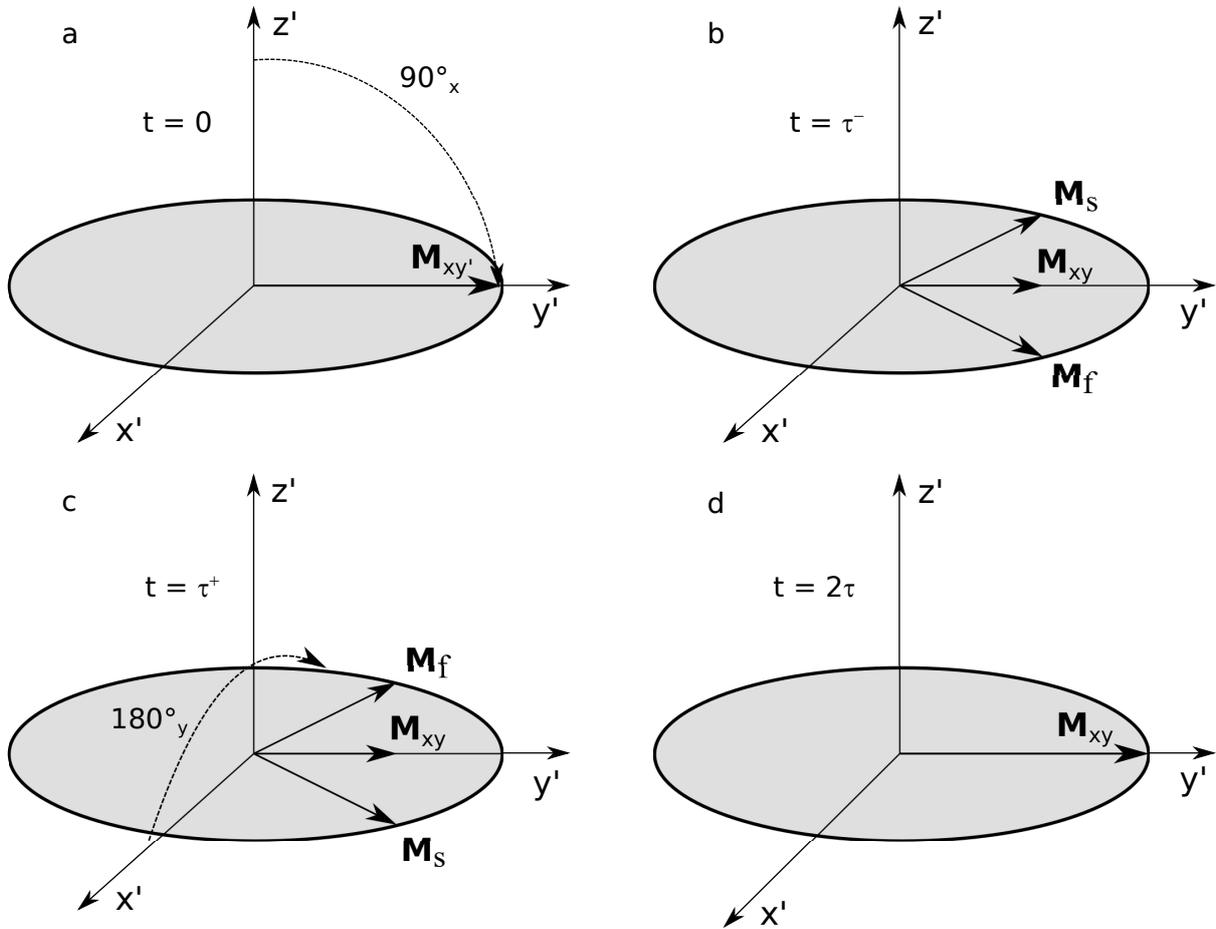
Figure 2.7: Isochromats ($\boldsymbol{M}_f$ and $\boldsymbol{M}_s$) in the Larmor rotating frame during a spin-echo experiment. In (a), all isochromats are aligned along the y'-axis after the 90°_x pulse. In (b), isochromats are dephased because of magnetic field inhomogeneities. In (c), a 180°_y pulse has inverted the isochromats ($t = \tau$). In (d), isochromats are completely rephased ($t = 2\tau$).

The same reasoning applies to the more general case of an arbitrary number of isochromats. Usually, phase coherence is totally lost by $t = \tau$, so that $M_{xy}(t)$ vanishes and there is no measurable signal. After the 180° pulse, the echo forms when all isochromats are rephased. The same mechanism that causes spins to dephase also causes them to rephase, so that if we ignore

T_2, the echo is symmetric about $t = \tau$:

$$|M_{x'y'}(\tau - t)| = |M_{x'y'}(\tau + t)| \qquad 0 \leqslant t \leqslant \tau \qquad (2.54)$$

Sometimes, $M_{x'y'}(t)$ after $t = \tau$ is referred to as the recalled transverse magnetization. Dephasing due to T_2 processes are irreversible, so that the amplitude of a spin echo is T_2-weighted with the whole signal having a e^{-t/T_2} envelope (see Figure 2.8).
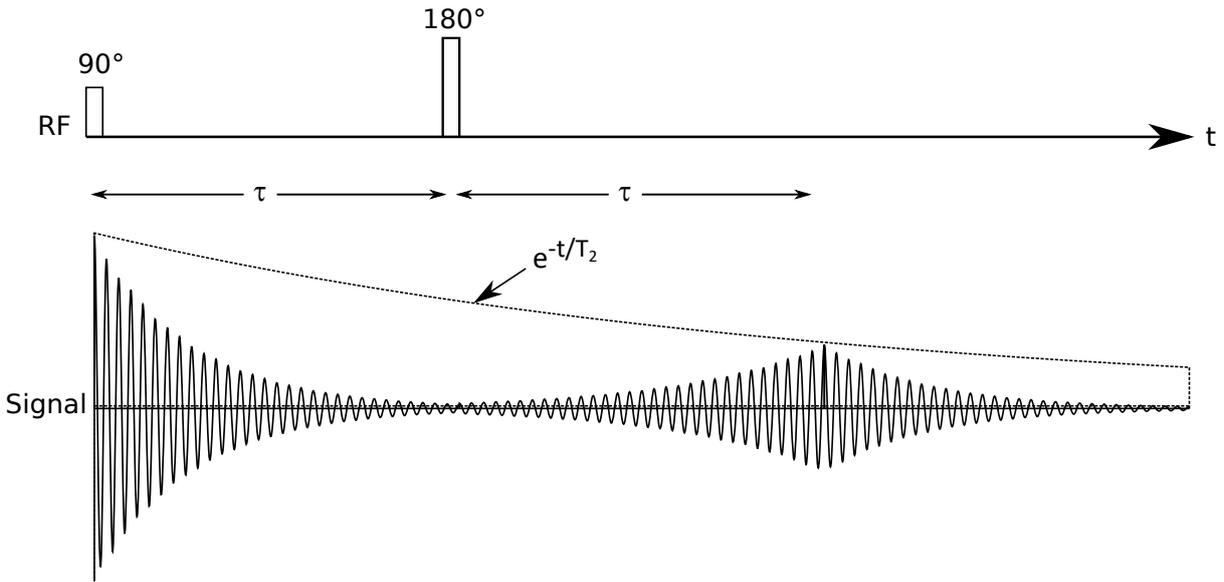
Figure 2.8: Spin echo pulse sequence. After the initial $90°$ pulse, the signal quickly disappears. Following the $180°$ pulse, an echo appears at $t = 2\tau$. The echo amplitude is proportional to $e^{-2\tau/T_2}$.

2.2.5 Gradient Echoes

Gradient echoes are generated using time-varying magnetic field gradients to dephase and rephase the spins. We define an inhomogeneous gradient field $\boldsymbol{B}_G$ whose z-component varies linearly along a given direction. For example, we could have

$$B_{G,z} = G_x x \qquad (2.55)$$

where G_x is an x-gradient. There are similar expressions for the y- and z-gradients. Technically, there are other components of $\boldsymbol{B}_G$ other than $B_{G,z}$, but these are usually much smaller than the

B_0-field and can usually be ignored [23]. Then $B_{G,z} = B_G$. The total magnetic field is

$$\boldsymbol{B} = (B_0 + B_{G,z})\hat{\boldsymbol{k}} \tag{2.56}$$

Figure 2.9 illustrates what happens to the magnetic field when we turn on a gradient. We can define a gradient vector $\boldsymbol{G} = (G_x, G_y, G_z)$ made up of the x-, y-, and z-gradient components. With this notation, the gradient field can be written as

$$B_{G,z} = \boldsymbol{G} \cdot \boldsymbol{r} \tag{2.57}$$

To show how we can generate a gradient echo, suppose we have an x-gradient. After the $90°$ excitation pulse, we turn on an x-gradient $(-G_x)$ for a time τ. Because the field varies spatially, spin frequencies will vary spatially. At $t = \tau$, spins at different locations will have acquired different phases. For $t < \tau$, the phase at location x and time t can be written as

$$\phi(x,t) = \gamma \int_0^t -G_x x\, dt' = -\gamma G_x x t \qquad 0 \leqslant t \leqslant \tau \tag{2.58}$$

Phase coherence quickly disappears and the signal vanishes. Afterwards, an opposite gradient $(+G_x)$ is applied to rephase the spins and recall the signal. The phase during this period of time is

$$\begin{aligned} \phi(x,t) &= -\gamma G_x x \tau + \gamma \int_\tau^t G_x x\, dt' \\ &= -\gamma G_x x \tau + \gamma G_x x (t - \tau) \qquad \tau < t \leqslant 2\tau \end{aligned} \tag{2.59}$$

At $t = 2\tau$, $\phi(x,t)$ is zero for all x, showing that all spins have rephased. In the presence of magnetic field inhomogeneities, the amplitude of a gradient echo is T_2^*-weighted (see Figure 2.10).

2.3 Magnetic Resonance Imaging

The MR signal generated by an object is a sum of all the local signals from different parts of the object. If the object is homogeneous, then all the local signals will be the same. But
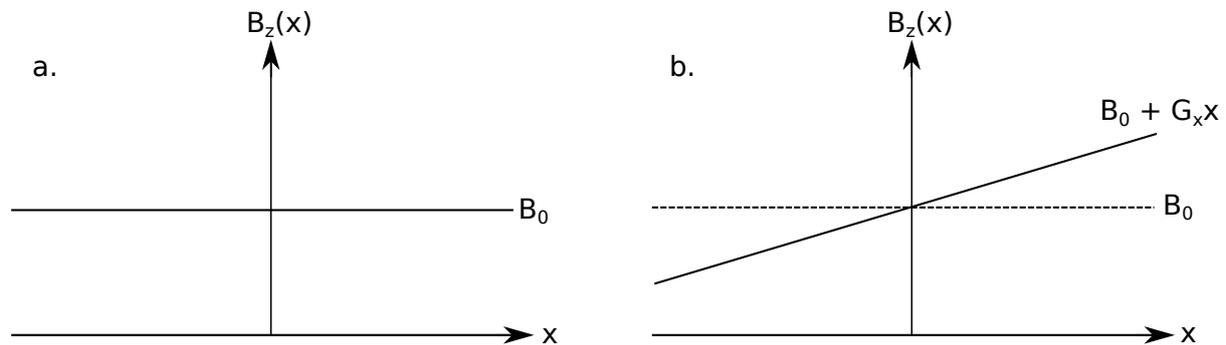
Figure 2.9: (a) Spatially independent magnetic field $\boldsymbol{B} = B_0\hat{\boldsymbol{k}}$. (b) Spatial dependence of the magnetic field in the presence of an x-gradient G_x. The magnetic field becomes $\boldsymbol{B} = (B_0 + G_x x)\hat{\boldsymbol{k}}$. Assuming G_x is positive, the magnitude of $\boldsymbol{B}$ increases linearly in the x-direction.
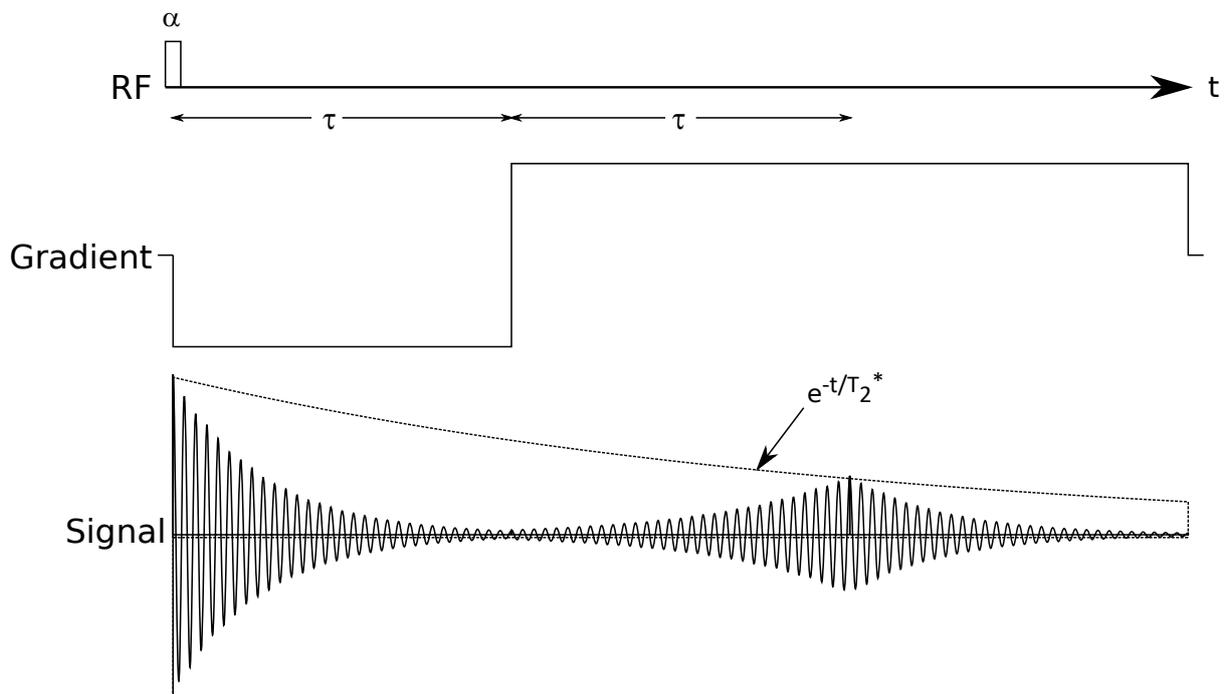
Figure 2.10: Gradient echo pulse sequence. After the application of an α pulse, a negative gradient is is applied for a time τ to dephase the spins. The spins are then rephased by an equally strong positive gradient applied for an equal amount of time τ, so that the signal is rephased at 2τ. The echo amplitude is proportional to $e^{-2\tau/T_2^*}$.

in biological samples, which are heterogeneous, we need a way to distinguish signals from different spatial locations. There are two types of spatial localization: selective excitation (see Section 2.3.1) and spatial encoding (see Section 2.3.2) [91, 109, 110, 111].

2.3.1 Selective Excitation

Slice selection has two requirements: a shaped RF pulse and an applied gradient field. A slice of thickness Δs can be described by

$$|\boldsymbol{\mu}_s \cdot \boldsymbol{r} - s_0| < \Delta s/2 \tag{2.60}$$

where $\boldsymbol{\mu}_s$ is a vector that points away from the surface of the slice and s_0 is the distance to the center of the slice from the origin. MRI is unique and powerful because the slice orientation can be in any arbitrary direction. Three special slice orientations are those along the x-, y-, and z-directions. For example, a slice at $z = z_0$ along the z-axis is given by $|z - z_0| < \Delta z/2$. We apply a gradient to make the resonant frequency vary linearly along the sample in the direction of the slice orientation. This gradient is called the slice selection gradient, $\boldsymbol{G}_{ss}$. Appropriate choice of $B_1(t)$ lets us excite spins at a specific slice. The magnitude of $\boldsymbol{G}_{ss}$ depends on the desired thickness of the slice. If we want a thin slice, then we need a larger slice selection gradient.

After choosing a slice orientation, we have to choose the temporal waveform of the RF pulse $B_1(t)$. Both the envelope function $B_1^e(t)$ and frequency ω_{rf} need to be chosen. In this section, we use a Fourier transform approach to find these quantities [100].

We describe the slice profile using a spatial selection function $p_s(z)$ (for simplicity, a z-direction slice orientation is used in this section). Since we want a rectangular slice centered at z_0, $p_s(z)$ will be a boxcar function:

$$p_s(z) = \Pi\left(\frac{z - z_0}{\Delta z}\right) \tag{2.61}$$

Suppose we turn on a z-gradient, $\boldsymbol{G}_{ss} = G_z \hat{\boldsymbol{k}}$. The Larmor frequency at position z is

$$\omega(z) = \omega_0 + \gamma G_z z \tag{2.62}$$

or

$$f(z) = f_0 + \gamma G_z z / 2\pi \tag{2.63}$$

Therefore, the frequency selection function is

$$p_s(f) = p_s\left(\frac{2\pi f}{\gamma G_z}\right) = \Pi\left(\frac{f - f_c}{\Delta f}\right) \tag{2.64}$$

where $f_c = f_0 + \gamma G_z z_0 / 2\pi$ and $\Delta f = \gamma G_z \Delta z / 2\pi$. Assuming the spins behave like a linear system, the Fourier transform of $B_1(t)$ is $p(f)$ [100]:

$$B_1(t) \propto \int p_s(f) e^{-i2\pi f t} df \tag{2.65}$$

Eq. 2.65 allows us to find both ω_{rf} and $B_1^e(t)$. If we substitute Eq. 2.64 into Eq. 2.65, we find that

$$B_1(t) \propto \Delta f \operatorname{sinc}(\pi \Delta f t) e^{-i2\pi f_c t} \tag{2.66}$$

If we want a rectangular slice, the envelope function will be a sinc function. The proportionality constant in Eq. 2.66 is determined by the desired flip angle. The desired RF frequency is

$$\omega_{\mathrm{rf}} = 2\pi f_c = \omega_0 + \gamma G_z z_0 \tag{2.67}$$

Figure 2.11a summarizes the idea behind spatial selection. A more accurate approach to finding $B_1^e(t)$ involves solving the Bloch equation. Although the Bloch equation approach is more accurate, it is also much more complicated [29, 31, 100, 105].

There are a few other considerations that need to be taken into account [100]. Suppose the slice selection gradient is on for time τ_p. When the slice selection gradient is turned off, there will be a position dependent phase shift $\gamma G_z(z - z_0)\tau_p/2$ across the slice. If ignored, it can cause unwanted signal loss. To correct for it, a refocusing gradient $G_{r,z}$ is turned on after

excitation to rephase the slice. If the refocusing gradient is on for a time τ_r, the phase shift at the end will be

$$\phi(z, t = \tau_p + \tau_r) = \gamma G_z(z - z_0)\frac{\tau_p}{2} + \gamma G_{r,z}(z - z_0)\tau_r \tag{2.68}$$

If we choose G_z and τ_r correctly, we can make the phase shift disappear. The condition is

$$G_{r,z}\tau_r = -\frac{1}{2}G_z\tau_p \tag{2.69}$$

Usually, we set $G_{r,z} = -G_z$ and $\tau_r = \frac{1}{2}\tau_p$, so that the refocusing gradient has opposite polarity compared to the slice select gradient and is half the duration. Figure 2.11b shows a slice selection gradient and a refocusing gradient.

The sinc function also has to be truncated. Truncation is equivalent to multiplication by a rectangular window function. Its Fourier transform $p'_s(f)$ is a convolution of an ideal rectangular function and a sinc function. The actual slice excitation profile will be non-uniform and there will be some excitation in nearby slices (known as a "cross-talk" artifact). Other envelopes for the RF pulses can be used which could provide smoother slice profiles. For instance, a Gaussian RF pulse provides excitation of smaller side lobes compared to a truncated sinc pulse [23].

2.3.2 Spatial Encoding

With spatial encoding, spatial information is encoded during the free precession period. There are two types: frequency encoding and phase encoding.

2.3.2.1 Frequency encoding

In frequency encoding, the oscillation frequency of the MR signal is designed to be linearly dependent on spatial position. Suppose we have a one dimensional spin distribution $\rho(x)$. If we have a homogenous magnetic field $B_0\hat{k}$ and turn on an x-gradient (G_x), the Larmor frequency at position x is

$$\omega(x) = \omega_0 + \gamma G_x x \tag{2.70}$$
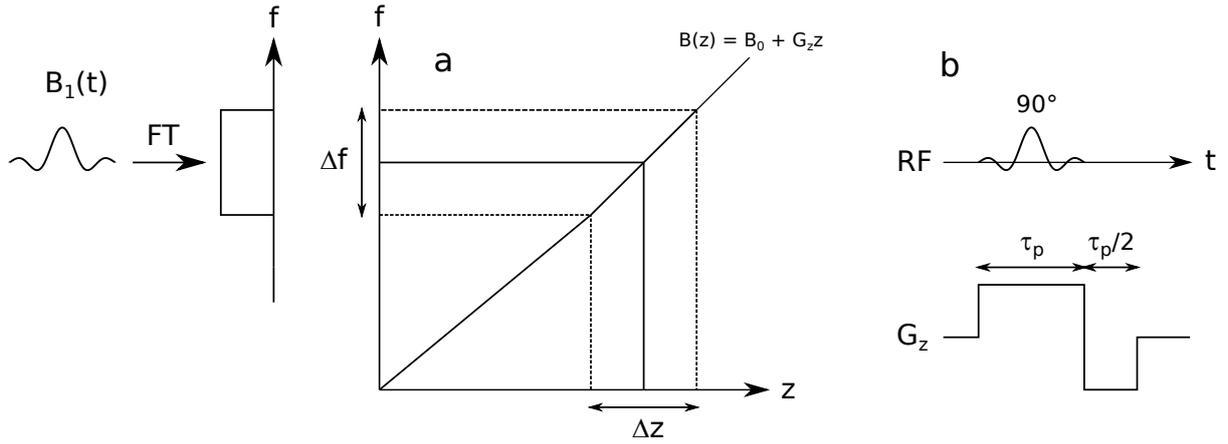
Figure 2.11: (a) Mapping slice width to frequency bandwidth. An RF pulse $B_1(t)$ with a sinc envelope has a rectangular frequency profile. The rectangular frequency profile excites spins in a frequency range Δf. Because of the linear G_z-gradient, spins in frequency range Δf map to slice width Δz. (b) Slice-selective 90° RF pulse and slice selection gradient. The slice selection gradient is applied for time τ_p. Afterwards, a refocusing gradient is applied for time $\tau_p/2$.

The signal due to the water at x in the interval dx is

$$dS(x,t) \propto \rho(x)dx e^{-i\gamma(B_0 + G_x x)t} \tag{2.71}$$

where the flip angle, magnetization M_z^0, and other constants are included in the proportionality constant. The signal is frequency encoded because $\omega(x)$ depends on x, provided the gradient is turned on. The signal from the entire object is

$$S(t) = \int_{object} \rho(x)e^{-i\gamma G_x x t}dx \tag{2.72}$$

where demodulation removes the carrier signal $e^{(-i\omega_0 t)}$. Because resonant frequencies are spatially dependent when the frequency encoding gradient is on, loss of phase coherence causes the signal to decay at a faster rate. More generally, the frequency encoding gradient can point in any direction, $\boldsymbol{G}_{FE} = (G_x, G_y, G_z)$. The general form of the signal becomes

$$S(t) = \int_{object} \rho(\boldsymbol{r})e^{-i\gamma \boldsymbol{G}_{FE} \cdot \boldsymbol{r} t}d\boldsymbol{r} \tag{2.73}$$

2.3.2.2 Phase encoding

Phase encoding involves turning on a gradient for a time T_{PE} after the RF pulse. While the gradient is on, spins are encoded along the gradient direction. After the gradient is turned off, spins have acquired different phases depending on their position. After the preparatory period, the phase at location $\boldsymbol{r}$ is

$$\phi(\boldsymbol{r}) = -\gamma \boldsymbol{G}_{PE} \cdot \boldsymbol{r} T_{PE} \tag{2.74}$$

where $\boldsymbol{G}_{PE}$ is the phase encoding gradient and T_{PE} is the phase encode interval. Because the phase is linearly dependent on position, the resulting signal is said to be phase encoded. The signal is the sum of all local signals,

$$S(t) = \int_{object} \rho(\boldsymbol{r}) e^{-i\gamma \boldsymbol{G}_{PE} \cdot \boldsymbol{r} T_{PE}} d\boldsymbol{r} \tag{2.75}$$

where again, the carrier signal $e^{(-i\omega_0 t)}$ has been removed with demodulation. The phase $\phi(\boldsymbol{r})$ can be adjusted by changing either the phase encode interval T_{PE} or the gradient magnitude $|\boldsymbol{G}_{PE}|$.

2.3.3 k-space

We can make the connection between spatial encoding (frequency and/or phase) and Fourier transforms more explicit by introducing the concept of k-space [103, 119, 182]. Suppose we have a frequency encoding gradient $\boldsymbol{G}_{FE}$. If we let

$$\boldsymbol{k} = \begin{cases} \gamma \boldsymbol{G}_{FE} t / 2\pi & \text{FID signals} \\ \gamma \boldsymbol{G}_{FE}(t - T_E) / 2\pi & \text{echo signals} \end{cases} \tag{2.76}$$

then we can write the signal as

$$S(\boldsymbol{k}) = \int_{object} \rho(\boldsymbol{r}) e^{-i2\pi \boldsymbol{k} \cdot \boldsymbol{r}} d\boldsymbol{r} \tag{2.77}$$

In this formulation, spin density and signal are Fourier transforms of each other. The gradient $\boldsymbol{G}_{FE}$ maps a time signal to a k-space signal. If $\boldsymbol{G}_{FE} = 0$, all time data points get mapped to the k-space origin. If $\boldsymbol{G}_{FE}$ is nonzero, then the time signal is mapped to a line in k-space. This set of points defines the k-space trajectory. In practice, $S(\boldsymbol{k})$ is available for a limited set of points.

For example, suppose we turn on x- and y-gradients G_x and G_y. The x and y components of $\boldsymbol{k}$ are

$$\begin{cases} k_x = \gamma G_x t / 2\pi \\ k_y = \gamma G_y t / 2\pi \end{cases} \tag{2.78}$$

As time moves on, we trace out a line in k-space that starts at the origin and makes an angle

$$\phi = \tan^{-1}\left(\frac{G_y}{G_x}\right) \tag{2.79}$$

with the k_x axis. If we chose to include an echo, the k-space components would be

$$\begin{cases} k_x = \gamma G_x(t - T_E)/2\pi \\ k_y = \gamma G_y(t - T_E)/2\pi \end{cases} \tag{2.80}$$

Instead of starting at the origin, the k-space trajectory starts at $\boldsymbol{k} = (-\gamma G_x T_E/2\pi, -\gamma G_y T_E/2\pi)$. We can also incorporate time varying gradients,

$$\boldsymbol{k}(t) = \frac{\gamma}{2\pi} \int_0^t \boldsymbol{G}_{FE}(\tau)d\tau \tag{2.81}$$

where the lower limit represents the time just after RF excitation. Eq. 2.81 allows nonlinear trajectories in k-space [100].

The phase encoding signal in k-space notation is

$$S(\boldsymbol{k}) = \int_{object} \rho(\boldsymbol{r})e^{-i2\pi\boldsymbol{k}\cdot\boldsymbol{r}}d\boldsymbol{r} \tag{2.82}$$

where

$$\boldsymbol{k} = \gamma \boldsymbol{G}_{PE} T_{PE}/2\pi \tag{2.83}$$

36

In frequency encoding, $\boldsymbol{k}$ was a function of t. In phase encoding, $\boldsymbol{k}$ is a single point, provided $\boldsymbol{G}_{PE}$ and $\boldsymbol{T}_{PE}$ are given. This means that phase encoding affects the starting point and not the trajectory in k-space. For a time varying gradient,

$$\boldsymbol{k} = \frac{\gamma}{2\pi} \int_0^{T_{PE}} \boldsymbol{G}_{PE}(\tau)d\tau \tag{2.84}$$

The final point in k-space depends only on the area of the integral in Eq. 2.84. Provided their integrals are the same, differently shaped gradient waveforms end up at the same point (though they approach it at different rates) [100].

2.3.4 Imaging Methods

Suppose we want a two-dimensional image. We can have either a 2D projection or a 2D slice. To get a projection, we can use a non-selective hard pulse to excite all the spins. If we want only a slice, we need to use a selective soft pulse to excite the spins in a specific slice. Given a spin density $\rho(\boldsymbol{r})$, we can define the image function to be

$$I(x,y) = \int_{-\infty}^{\infty} \rho(x,y,z)dz \tag{2.85}$$

for a two dimensional projection and

$$I(x,y) = \int_{z_0-\Delta z/2}^{z_0+\Delta z/2} \rho(x,y,z)dz \tag{2.86}$$

for a slice of thickness Δz.

In k-space notation, the signal $S(k_x,k_y)$ is related to the image function in the following way:

$$S(k_x,k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y)e^{-i2\pi(k_x x+k_y y)}dxdy \tag{2.87}$$

We need to measure $S(\boldsymbol{k})$ at enough points in k-space so that we can perform the inversion in Eq. 2.87. The conventional method involves collecting n sets of signals $S_n(t)$, each encoded in a slightly different way. All $S_n(t)$ taken together cover k-space. If we use this method, we need to use multiple excitations so that we can cover multiple lines in k-space.

2.3.5 k-space Acquisition

We need to decide how to encode each $S_n(t)$ so that we cover k-space. One relatively simple choice is to use a Cartesian grid [49]. After initial excitation, we first phase encode the signal along the y-direction, then acquire the signal in the presence of a frequency encoding gradient.

Consider the nth excitation. During the phase encode step and preparatory readout, we have

$$\begin{cases} k_x = \gamma G_x(t - t_0)/2\pi \\ k_y = \gamma n \Delta G_y(t - t_0)/2\pi \end{cases} \quad t_0 < t < T_{acq}/2 + t_0 \tag{2.88}$$

where we assumed that $G_y = n\Delta G_y$. When both gradients are turned off, $\boldsymbol{k}$ is at location

$$\boldsymbol{k}_A = (\gamma G_x T_{PE}/2\pi, \ \gamma n \Delta G_y T_{PE}/2\pi) \tag{2.89}$$

The $180°$ refocusing pulse inverts $\boldsymbol{k}$, changing it to $-\boldsymbol{k}$. During data acquisition, we have

$$\begin{cases} k_x = \gamma G(t - T_E)/2\pi \\ k_y = \gamma n \Delta G_y T_{PE}/2\pi \end{cases} \quad |t - T_E| < T_{acq}/2 \tag{2.90}$$

While frequency encoding gradient is on, $\boldsymbol{k}$ moves in a horizontal line, intercepting the k_y axis at $\gamma n \Delta G_y T_{PE}/2\pi$. By varying G_y, we can control the intercept. After each excitation, we trace out a new horizontal line in k-space. This scheme is called the phase encoding method.

An alternative acquisition scheme involves sampling radial lines in k-space. In this scheme, readout occurs in the presence of both the G_x and G_y gradients. Consider the nth excitation. Between the $90°$ and $180°$ pulses, we have

$$\begin{cases} k_x = \gamma G(t - t_0)\cos(\phi_n)/2\pi \\ k_y = \gamma G(t - t_0)\sin(\phi_n)/2\pi \end{cases} \quad t_0 < t < T_{acq}/2 + t_0 \tag{2.91}$$

where ϕ_n is

$$\phi_n = \tan^{-1}\left(\frac{G_{n,x}}{G_{n,y}}\right) \tag{2.92}$$

When the gradients are turned off, $\boldsymbol{k}$ is

$$\boldsymbol{k}_A = \gamma G \frac{T_{acq}}{2 \cdot 2\pi}(\cos(\phi_n), \; \sin(\phi_n)) \tag{2.93}$$

The 180° pulse sends $\boldsymbol{k}_A \to -\boldsymbol{k}_A$. During readout, we trace a k-space trajectory defined by

$$\begin{cases} k_x = \gamma G \cos(\phi_n)(t - T_E)/2\pi \\ k_y = \gamma G \cos(\phi_n)(t - T_E)/2\pi \end{cases} \quad |t - T_E| < T_{acq}/2 \tag{2.94}$$

which is a line passing through the origin and ending at $\boldsymbol{k}_A$. We can trace a different line by adjusting the relative strengths of $G_{n,x}$ and $G_{n,y}$ (or ϕ_n) in the next excitation. This scheme is called radial projection imaging. Both acquisition schemes are shown in Figure 2.12.

2.3.6 Sampling Criteria

The Nyquist sampling theorem states that if $G(t)$ is a bandlimited function with maximum frequency f_{max}, then $f_s = 2f_{max}$ is the minimum sampling rate needed to recover $G(t)$. Here, the reciprocal variables were time and frequency. In MRI, the reciprocal variables are k (analogous to time) and position (analogous to frequency). Assume we have an object bounded in a rectangular region of widths W_x and W_y. According to the Nyquist sampling theorem, k-space should be sampled so that

$$\begin{cases} \Delta k_x \leqslant 1/W_x \\ \Delta k_y \leqslant 1/W_y \end{cases} \tag{2.95}$$

If we use the phase encoding method described in Section 2.3.5, then Δk_x and Δk_y are

$$\Delta k_x = \gamma |G_x| \Delta t / 2\pi$$
$$\Delta k_y = \gamma \Delta G_y T_{PE} / 2\pi \tag{2.96}$$
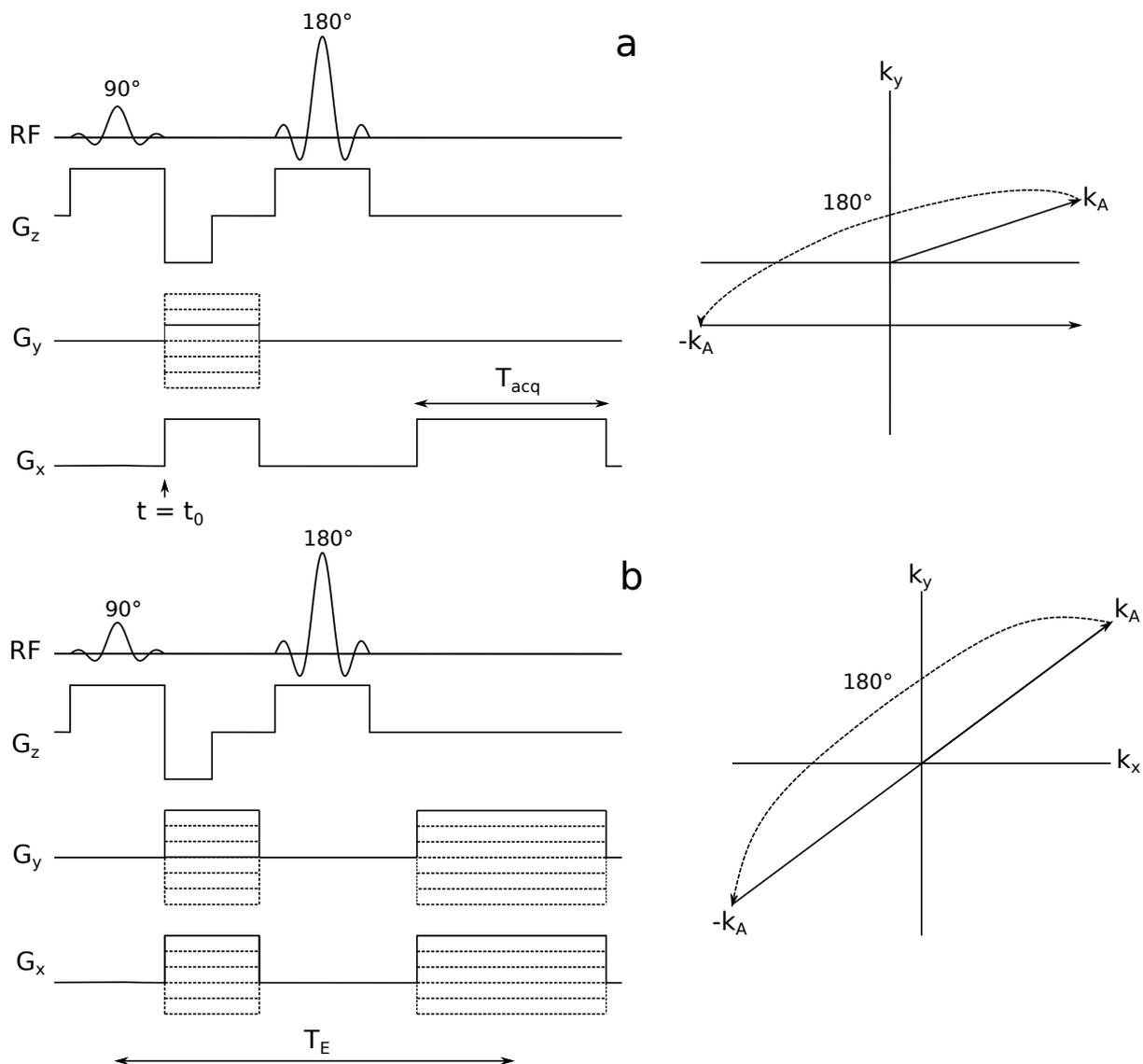
Figure 2.12: Spin echo imaging sequences and k-space trajectories. (a) 'Spin-warp' trajectory. A $90°$ slice selective pulse with a refocusing gradient, here in the z-direction, is applied to excite the spins. A phase encoding gradient, shown here in the y-direction, is applied at the same time as the frequency dephasing gradient to move to the starting point in k-space ($\boldsymbol{k}_A$). A $180°$ pulse is applied to move $\boldsymbol{k}_A \rightarrow -\boldsymbol{k}_A$. Here the pulse is shown with a slice selective gradient in the z-direction. A frequency encoding gradient of duration T_{acq}, shown here in the x-direction, is used to collect a horizontal line in k-space. (b) Radial projection trajectory. A $90°$ slice selective pulse with a refocusing gradient, here in the z-direction, is applied to excite the spins. Two gradients are applied simultaneously in the x- and y-direction to move the starting point in k-space ($\boldsymbol{k}_A$). A $180°$ to move $\boldsymbol{k}_A \rightarrow -\boldsymbol{k}_A$. Two gradients are then applied simultaneously in the x- and y-direction to collect a radial line in k-space.

Solving for Δt and ΔG_y and making use of Eq. 2.95, we have the following conditions:

$$\Delta t \leqslant \frac{2\pi}{\gamma |G_x| W_x}$$
$$\Delta G_y \leqslant \frac{2\pi}{\gamma T_{PE} W_y} \tag{2.97}$$

Eq. 2.97 gives the sampling requirements needed to make an image. Similar ideas can be used to find minimum sampling requirements for polar sampling [100].

2.3.7 Image Reconstruction

Until now, we assumed that $S(k)$ and $I(x)$ were continuous. In reality, $S(k)$ is made up of discrete points $S(k_n) = S(n\Delta k)$, for integer n. The imaging equation then becomes

$$S[n] = S(n\Delta k) = \int_{-\infty}^{\infty} I(x) e^{-i2\pi n\Delta k x} \tag{2.98}$$

Provided that the Nyquist condition is satisfied (or that $I(x) = 0$ past a certain range), we can write

$$I(x) = \Delta k \sum_{n=-N/2}^{N/2-1} S[n] e^{i2\pi n\Delta k x} \qquad\qquad |x| < \frac{1}{\Delta k} \tag{2.99}$$

The region $|x| < W_x/2$ is called the field of view (FOV). Eq. 2.99 is called the Fourier reconstruction formula. Truncation of the series leads to a Gibbs ringing artifact. It can be reduced by multiplying $S[n]$ by a window function w_n (such as a Hamming function), but at the cost of reduced spatial resolution [100].

We can turn $I(x)$ into $I(m\Delta x) = I[m]$ if the Nyquist condition is fulfilled. Since $I(x)$ is frequency bandlimited (finite sampling, $|k| \leqslant (N/2)\Delta k$), we can recover $I(x)$ from $I[m]$ as long as

$$\Delta x \leqslant \frac{1}{\Delta k} \tag{2.100}$$

Eq. 2.100 sets a limit on the largest pixel size. When the equality in Eq. 2.100 holds, Δx is

known as the Fourier pixel size. So

$$I[m] \propto \sum_{n=-N/2}^{N/2-1} S[n] e^{i2\pi nm/N} \qquad\qquad -N/2 \leqslant m < N/2 \qquad (2.101)$$

The reconstruction in Eq. 2.101 requires N^2 multiplications, so it can be slow. Nowadays, we would use the Fast Fourier Transform (FFT) developed in the 1960s by Tukey and Cooley [38]. The simple version of the Fast Fourier transform requires the number of data points to be a power of two, so one might use zero-padding to make the data a power of two.

2.3.8   Image Contrast

Contrast is important because it allows humans to distinguish different parts of an image. Let I_A be the image intensity of tissue A and I_B be the image intensity of tissue B. Image contrast is defined by

$$C_{AB} = \frac{I_A - I_B}{I_{ref}} \qquad\qquad (2.102)$$

where I_{ref} is reference intensity. Image intensity, as well as contrast, depends on many factors. Some of these factors include spin density ρ, relaxation times T_1, T_2, T_2^*, and diffusion coefficient D. Generally, the contrast can be written as

$$C_{AB} = f(\rho, T_1, T_2, T_2^*, D, \dots) \qquad\qquad (2.103)$$

where the function f depends on the data acquisition procedure. Each acquisition procedure, called a pulse sequence, is usually designed to isolate one or more causes of contrast in the image.

2.3.9   Example Pulse Sequences

This section describes several common pulse sequences designed to isolate contrast depending on relaxation times.

2.3.9.1 Saturation Recovery

The first such sequence is a saturation-recovery. It consists of a series of N $90°$ pulses separated by a repetition time T_R. If we want to generate a spin echo signal, we also need to include a $180°$ pulse after the $90°$ pulse. The saturation condition ($T_R \gg T_2$) ensures that no transverse magnetization remains at the time the next pulse is applied. If necessary, we can use spoiler gradients to destroy any remaining transverse magnetization [28]. Provided that $T_E \ll T_R$, then the image intensity is [65]

$$I(\boldsymbol{r}) \propto \rho(\boldsymbol{r}) \left(1 - e^{-T_R/T_1} \right) e^{-T_E/T_2} \tag{2.104}$$

All three weightings (T_1, T_2, ρ) are present. If T_E is very short, T_2-weighting is nearly eliminated. If T_R is also chosen to be long (> 1500 ms), then $I(\boldsymbol{r}) \approx \rho(\boldsymbol{r})$ and the image is spin density weighted. But if T_R is short (< 500 ms) instead, then the image is T_1-weighted. A saturation recovery sequence is shown in Figure 2.13.

2.3.9.2 Inversion Recovery

Another useful sequence is inversion-recovery. It begins with a $180°$ inversion pulse followed by an inversion time T_I, then a $90°$ pulse is applied, followed by a recovery (or delay) time T_D. If we want to generate a spin echo signal, we also need to include another $180°$ pulse after the $90°$ pulse. The image intensity goes as [100]

$$I(\boldsymbol{r}) \propto \rho(\boldsymbol{r}) \left[1 - 2e^{-T_I/T_1(\boldsymbol{r})} + e^{-T_D/T_1(\boldsymbol{r})} \right] e^{-T_E/T_2} \tag{2.105}$$

where T_E is assumed to be much less than $T_R = T_I + T_D$. Two parameters, T_I and T_D, can be adjusted for optimal T_1 contrast. Appropriate choice of T_I can be even used to suppress signals from certain tissues (e.g. fat, CSF, etc). If

$$T_I = [\ln 2 - \ln (1 + e^{-T_R/T_1^0})] T_1^0 \tag{2.106}$$

then signals from tissues with $T_1 = T_1^0$ will be suppressed and not contribute to the image [30]. This is known as signal-nulling. Inversion-recovery can generate more T_1 contrast than saturation recovery, so it is useful for differentiating tissues with similar (but different) T_1 values. Figure 2.14 shows an inversion recovery sequence.

2.3.9.3 Gradient Echoes

Gradient echo imaging replaces the 180° RF pulses with rephasing gradients. Gradient echo sequences allow for much richer contrast mechanisms. Assume that $T_R \gg T_2$, so the transverse magnetization is totally dephased before the next excitation. For an arbitrary flip angle α, the echo amplitude is [100]

$$A_E = \frac{M_z^0(1 - e^{-T_R/T_1})}{1 - \cos\alpha e^{-T_R/T_1}} e^{-T_E/T_2^*} \tag{2.107}$$

The echo amplitude carries T_1 and T_2^* weighting. The T_1-weighting depends mainly on flip angle, not T_R. If the flip angle is very small, T_1-weighting is eliminated. As flip angle increases, T_1-weighting also increases. The amount of T_2^*-weighting is controlled by T_E. Figure 2.15 shows an example of a gradient echo sequence.

2.3.9.4 Discussion

There are three main types of contrast in MRI imaging. Spin density contrast is proportional to spin density differences between tissues. The T_1 and T_2 contrasts are exponentially dependent on T_1 and T_2 values of the tissue. In soft tissues, $\rho(\boldsymbol{r})$ is mostly uniform. For clinical fields strengths, since T_1 values are different, T_1-weighted images are useful for viewing anatomical definition. Sometimes, T_2 changes occur in diseased states (such as tumours), so T_2-weighted images are useful for disease detection [100]. For high field strengths, such as the 7 T used in the experiments and simulations described in this thesis, T_1 values of tissues are similar so T_2-weighted images are used for viewing anatomical definition. Usually, we want to emphasize one type of contrast (T_1, T_2, or spin density). That contrast is determined by the relative values of T_E and T_R. A long T_R and short T_E gives spin density weighted images, a long T_R and intermediate T_E gives T_2-weighted images, and intermediate T_R and short T_E gives T_1-weighted images. The smallest achievable T_E is limited by hardware. The maximum achievable T_R is

constrained by imaging times. Table 2.1 summarizes the contrast mechanisms.

	Short T_R	Intermediate T_R	Long T_R
Short T_E		T_1-weighted	Proton density
Intermediate T_E			T_2-weighted

Table 2.1: Image contrasts of a saturation-recovery spin echo sequence.
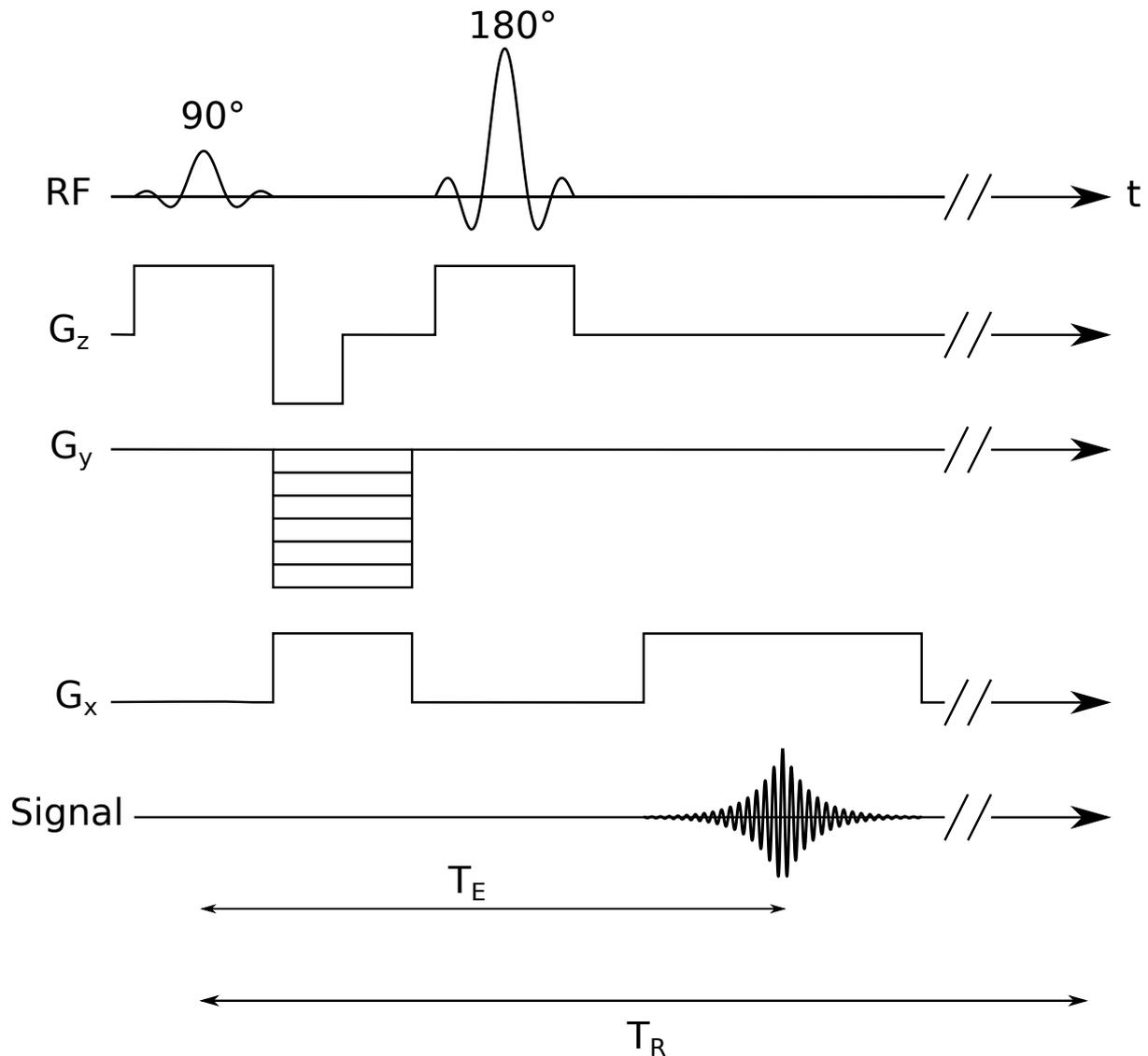
Figure 2.13: Spin echo imaging sequence. A 90° slice selective pulse with a refocusing gradient, here in the *z*-direction, is applied to excite the spins. A phase encoding gradient, shown here in the *y*-direction, is applied at the same time as the frequency dephasing gradient to move to the starting point in *k*-space. A 180° pulse is applied to create a spin echo, here the pulse is shown with a slice selective gradient in the *z*-direction. A frequency encoding gradient, shown here in the *x*-direction is used to collect an echo that represents a line in *k*-space. The pulses are then repeated after a time T_R to collect another line of *k*-space.
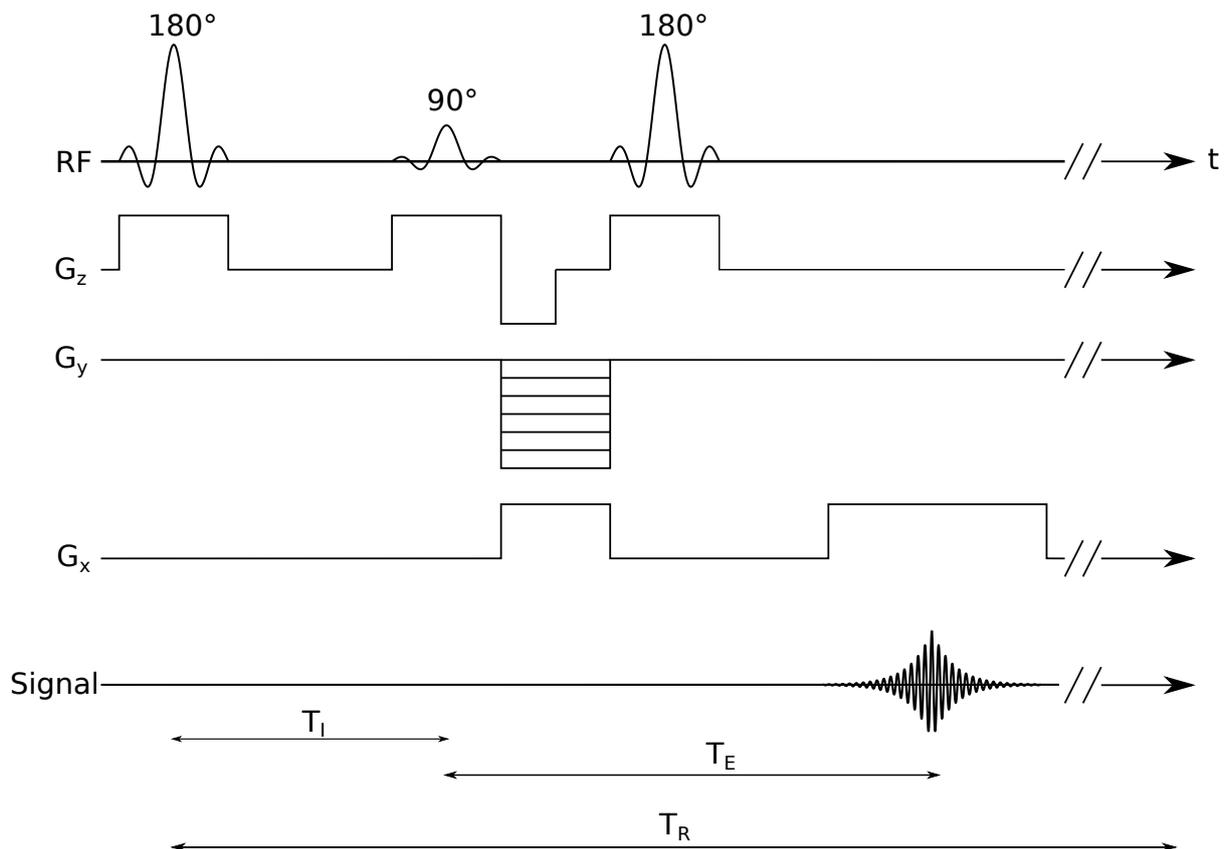
Figure 2.14: Inversion recovery imaging sequence. A 180° slice selective pulse, here in the z-direction, is applied to invert the spins so that they point in the $-z$-direction. After a time T_I, a 90° slice selective pulse with a refocusing gradient, here in the z-direction, is applied to flip the spins into the transverse plane. A phase encoding gradient, shown here in the y-direction, is applied at the same time as the frequency dephasing gradient to move to the starting point in k-space. A 180° pulse is applied to create a spin echo, here the pulse is shown with a slice selective gradient in the z-direction. A frequency encoding gradient, shown here in the x-direction is used to collect an echo that represents a line in k-space. The pulses are then repeated after a time T_R to collect another line of k-space.
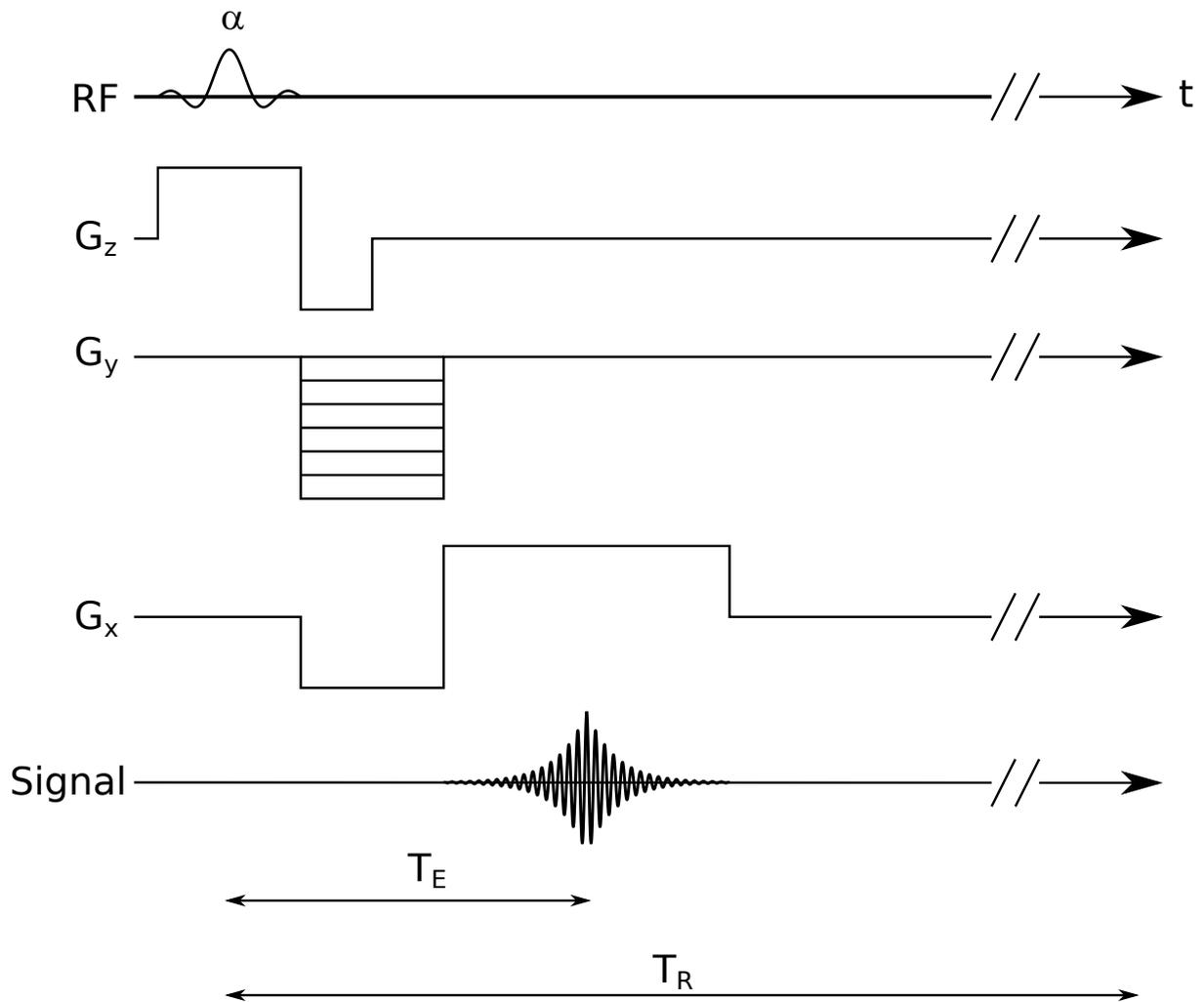
Figure 2.15: Gradient echo imaging sequence. An α slice selective pulse with a refocusing gradient, here in the z-direction, is applied to excite the spins. A phase encoding gradient, shown here in the y-direction, is applied at the same time as the frequency dephasing gradient to move to the starting point in k-space. A frequency encoding gradient, shown here in the x-direction is used to collect a gradient echo that represents a line in k-space. The pulses are then repeated after a time T_R to collect another line of k-space.

Chapter 3

Diffusion Magnetic Resonance Imaging

In Chapter 2, we saw how to generate echoes using either RF pulses or magnetic field gradients. In the presence of magnetic field gradients, full echo formation depends on the spins remaining in their original positions for the duration of the experiment [77]. Spin displacements lead to echo signal attenuation. This means that we can use magnetic field gradients to measure translational motion. When a gradient is turned on, the spins' Larmor frequencies become position dependent, and so their accumulated phases are also position dependent. If the particles diffuse during the experiment, the 180° refocusing pulse and subsequent gradient do not rephase all the spins, causing signal attenuation.

3.1   Diffusion and Brownian motion

Diffusion refers to the random motion of particles over time. At the molecular level, this is caused by collisions between molecules. In 1827, Robert Brown observed suspended pollen grains undergoing jittery motion in water, a phenomenon now called Brownian motion [77]. Einstein was able to connect Brownian motion with diffusion. He explained that water molecules (too small to be seen) were colliding with the pollen grain, causing it to move. He also showed that the root-mean square displacement was proportional to time,

$$\langle \boldsymbol{r}^2 \rangle = 2n_d Dt \tag{3.1}$$

where D is the diffusion coefficient and n_d is the dimensionality [50]. Figure 3.1 shows a collection of particles undergoing diffusion. Figure 3.2 shows an example of a single particle undergoing diffusion.
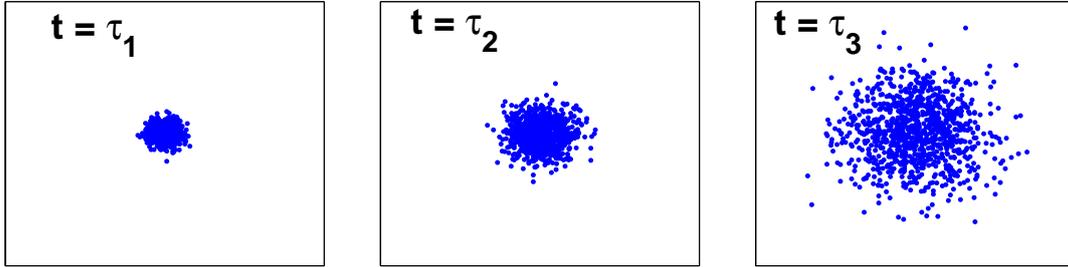
Figure 3.1: Displacement of particles as a function of time ($\tau_1 < \tau_2 < \tau_3$). At short times (τ_1) the spread is small, but at long times (τ_2 and τ_3), the particles have a bigger spread.
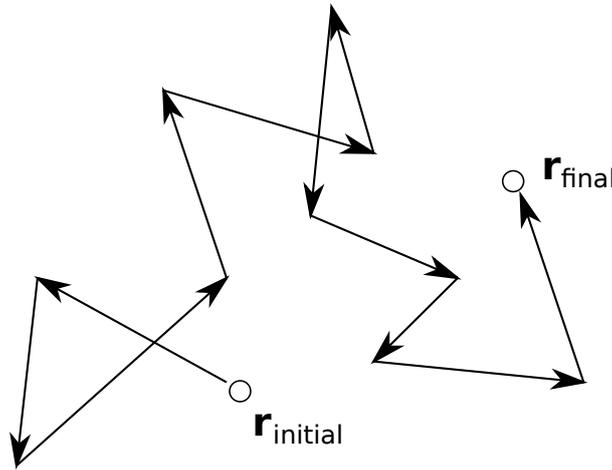
Figure 3.2: Single particle undergoing a random walk. We observe the particle starting at $\boldsymbol{r}_{initial}$ and ending at $\boldsymbol{r}_{final}$. The particle changes direction when it collides with another particle causing the motion to appear as a random walk.

3.1.1 The Diffusion Propagator

Since MR signals come from the superposition of signals from many spins, and since the motion of each individual particle will be different, an ensemble-based approach is helpful [31, 77, 129, 168].

Let the position of particle i at time t be $\boldsymbol{r}_i(t)$. We would like to describe the probability of a particle starting at $\boldsymbol{r}$ and moving to $\boldsymbol{r}'$ over the time t (see Figure 3.3). This probability may depend on the starting position itself. This probability is given by the self-correlation function $P_s(\boldsymbol{r}|\boldsymbol{r}',t)$, also called the conditional probability function. The self-correlation function

$P_s(\boldsymbol{r}|\boldsymbol{r}',t)$ has initial condition

$$P_s(\boldsymbol{r}|\boldsymbol{r}',0) = \delta(\boldsymbol{r}' - \boldsymbol{r}) \tag{3.2}$$

The probability of finding a particle at $\boldsymbol{r}'$ at time t is given by

$$P(\boldsymbol{r}',t) = \int P(\boldsymbol{r},0)P_s(\boldsymbol{r}|\boldsymbol{r}',t)d\boldsymbol{r} \tag{3.3}$$

where $P(\boldsymbol{r},0)$ is the initial probability of finding a particle at $\boldsymbol{r}$ at $t = 0$. We can take it to be the particle density $\rho(\boldsymbol{r})$.
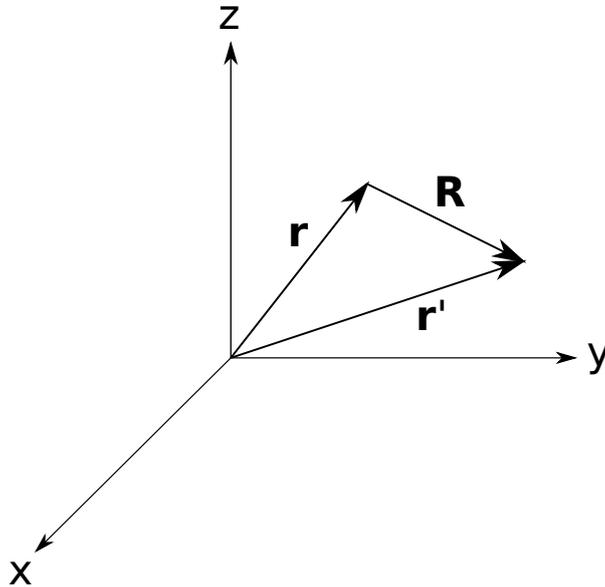
Figure 3.3: Displacement of a particle from $\boldsymbol{r}$ to $\boldsymbol{r}'$. The displacement vector is $\boldsymbol{R}$. The probability for a particle to move from $\boldsymbol{r}$ to $\boldsymbol{r}'$ over a time t is $P_s(\boldsymbol{r}|\boldsymbol{r}',t)$.

Classically, diffusion is described using Fick's law. Fick's law states that the particle flux (number of particles per unit area per unit time) is proportional to the concentration gradient. For self-diffusion there is no net concentration gradient, but the probability function $P(\boldsymbol{r}',t)$ obeys Fick's law. The concentration at $\boldsymbol{r}$ is analagous to the probability of finding a particle at $\boldsymbol{r}$. Since $P(\boldsymbol{r}',t)$ obeys Fick's law, so does the conditional probability $P_s(\boldsymbol{r}|r',t)$. This is Fick's first law,

$$\boldsymbol{J} = -D\nabla' P_s \tag{3.4}$$

where $\boldsymbol{J}$ is a conditional probability flux, D is the diffusion coefficient, and ∇ is the gradient

operator. Using the continuity equation,

$$\nabla \cdot \boldsymbol{J} = -\frac{\partial P_s}{\partial t} \tag{3.5}$$

which says that probability is conserved, we can write Fick's second law,

$$\frac{\partial P_s}{\partial t} = D\nabla'^2 P_s \tag{3.6}$$

For free, unrestricted diffusion, $P_s \rightarrow 0$ as $\boldsymbol{r} \rightarrow \infty$. The solution to Eq. 3.6 is a Gaussian function

$$P_s(\boldsymbol{r}|\boldsymbol{r}',t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{(\boldsymbol{r}'-\boldsymbol{r})^2}{4Dt}\right) \tag{3.7}$$

The Markov nature of Brownian motion is reflected in the fact that P_s depends only on the particle displacement $\boldsymbol{r}' - \boldsymbol{r}$. In PGSE MRI, particle displacements are indirectly measured, so let $\boldsymbol{R} = \boldsymbol{r}' - \boldsymbol{r}$ denote the net particle displacement over time t. Then Eq. 3.7 becomes

$$P_s(\boldsymbol{r}|\boldsymbol{r}+\boldsymbol{R},t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{\boldsymbol{R}^2}{4Dt}\right) \tag{3.8}$$

A related quantity is the averaged propagator $\overline{P_s(\boldsymbol{R},t)}$, defined by

$$\overline{P_s(\boldsymbol{R},t)} = \int P_s(\boldsymbol{r}|\boldsymbol{r}+\boldsymbol{R},t)\rho(\boldsymbol{r})d\boldsymbol{r} \tag{3.9}$$

The average propagator gives the average probability for a particle to have displacement $\boldsymbol{R}$ over time t. In cases such as free diffusion, $P(\boldsymbol{r}|\boldsymbol{r}+\boldsymbol{R},t)$ is independent of starting position. The average propagator is the same for all particles so that $\overline{P_s(\boldsymbol{R},t)} = P_s(\boldsymbol{R}|\boldsymbol{r}+\boldsymbol{R},t)$. Fick's law (Eq. 3.6) can be rewritten to account for flow of velocity $\boldsymbol{v}$ [126, 162]. In that case, the propagator becomes

$$P_s(\boldsymbol{R},t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{(\boldsymbol{R}-\boldsymbol{v}t)^2}{4Dt}\right) \tag{3.10}$$

which is a Gaussian centered at $\boldsymbol{R} = \boldsymbol{v}t$ that widens over time [31].

If we are interested in motion along one dimension, then we can integrate $P_s(\boldsymbol{R},t)$ over the

other two dimensions. For example, if we were interested in motion along the x direction,

$$
\begin{aligned}
P_s(X,t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{\boldsymbol{R}^2}{4Dt}\right) dY\,dZ \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{X^2}{4Dt}\right) \exp\left(-\frac{Y^2}{4Dt}\right) \exp\left(-\frac{Z^2}{4Dt}\right) dY\,dZ \quad (3.11) \\
&= \frac{1}{(4\pi Dt)^{1/2}} \exp\left(-\frac{X^2}{4Dt}\right)
\end{aligned}
$$

The time evolution of $P_s(X,t)$ is shown in Figure 3.4. Using $P_s(R,t)$ (Eq. 3.11), we can show that the average displacement is zero,

$$
\begin{aligned}
\langle X \rangle &= \int_{-\infty}^{\infty} X P_s(X,t) dX \\
&= 0
\end{aligned}
\tag{3.12}
$$

We can also find the mean square displacement $\langle X^2 \rangle$,

$$
\begin{aligned}
\langle X^2 \rangle &= \int_{-\infty}^{\infty} X^2 P_s(X,t) dX \\
&= 2Dt
\end{aligned}
\tag{3.13}
$$

which is the same result as 3.1. Eqs. 3.7 and 3.11 are valid for the specific case of free and unrestricted diffusion.
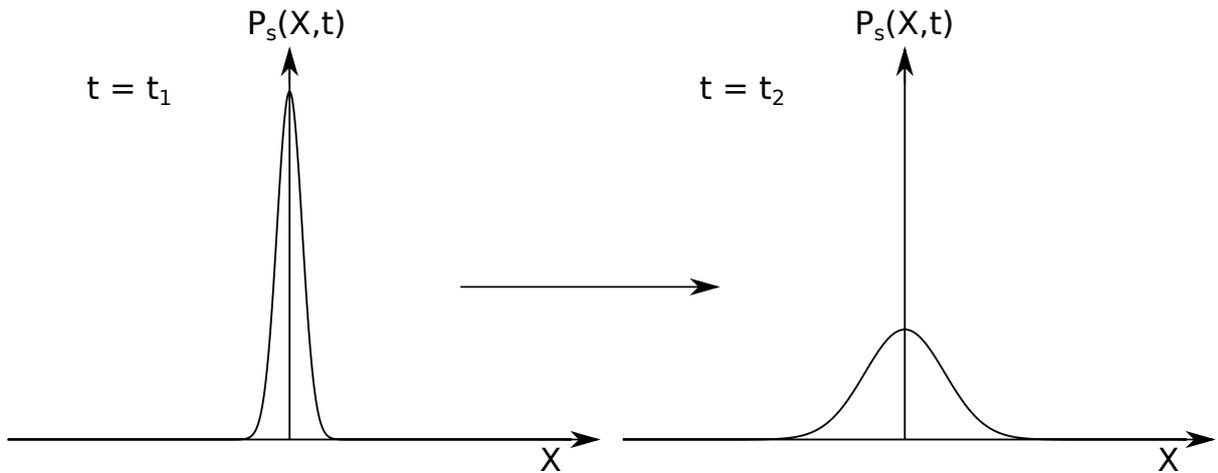
Figure 3.4: Time evolution of the conditional probability function $P_s(X,t)$ for free diffusion from time t_1 to $t_2 > t_1$. As time increases the probability that a particle will be farther away from its original point increases.

3.1.2 Restricted Diffusion

In Section 3.1.1, we saw that the propagator was a Gaussian when diffusion was unrestricted. We now look at the case of restricted diffusion, or diffusion in the presence of boundaries. The propagator can expanded in terms of its eigenfunctions [165],

$$P(\boldsymbol{r},t|\boldsymbol{r}',t') = \sum_n e^{-\lambda_n D|t'-t|} u_n(\boldsymbol{r}) u_n(\boldsymbol{r}') \tag{3.14}$$

where λ_n are the corresponding eigenvalues. The time dependence in the propagator is described by exponential decay. The eigenfunctions are solutions to the equation

$$-\lambda_n u_n(\boldsymbol{r}) = \nabla^2 u_n(\boldsymbol{r}) \tag{3.15}$$

They have the property that

$$\sum_n u_n(\boldsymbol{r}) u_n(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{3.16}$$

The eigenfunctions depend on the geometry of the pore through the boundary conditions. If we have impermeable boundaries, the boundary condition takes the form

$$\nabla P \cdot \hat{\boldsymbol{n}} = 0 \tag{3.17}$$

where $\hat{\boldsymbol{n}}$ is a vector normal to the surface [32]. Here no flux passes through the barrier. Eq. 3.15 has been solved analytically for a limited number of geometries, including plane separation, cylinders, and spheres.

3.1.2.1 Planes

In one dimension, the eigenfunction equation is

$$-\lambda_n u_n(x) = \frac{d^2 u_n}{dx^2} \tag{3.18}$$

The general solution to Eq. 3.18 has the form $A\cos(\lambda_n x) + B\sin(\lambda_n x)$. The boundary conditions for impermeable barriers at $x = 0$ and $x = L$ are

$$\begin{cases} u'_n(0) = 0 \\ u'_n(L) = 0 \end{cases} \tag{3.19}$$

The normalization condition is

$$\left(\int_0^L u_n(x)dx \right)^2 = 1 \tag{3.20}$$

The eigenvalues are

$$\lambda_n = \frac{n^2 \pi^2}{L^2} \tag{3.21}$$

The eigenfunctions are

$$\begin{cases} u_0(x) = \sqrt{\frac{1}{L}} \\ u_n(x) = \sqrt{\frac{2}{L}} \cos\left(\frac{n\pi x}{L}\right) & \text{if } n \neq 0 \end{cases} \tag{3.22}$$

Therefore, the propagator is [32]

$$P(x,t_1|x',t_2) = \frac{2}{L}\left[\frac{1}{2} + \sum_{n=1}^{\infty} \cos\left(\frac{n\pi x}{L}\right)\cos\left(\frac{n\pi x'}{L}\right)e^{-(n\pi/L)^2 D|t_1 - t_2|}\right] \tag{3.23}$$

3.1.2.2 Cylinders

The propagator for a cylindrical pore with an impermeable barrier at $r = a$ is [32]

$$P(\mathbf{r},t|\mathbf{r}',t') = \sum_k \sum_n \exp\left[-v_{nk}^2 D\Delta/a^2\right]\left[\left(\frac{2}{\pi a^2}\right)\left(\frac{v_{nk}^2}{J_k^2(v_{nk})}\right)\left(\frac{1}{v_{nk}^2 - k^2}\right)\right]$$
$$\times J_k(v_{nk}r/a)J_k(v_{nk}r'/a)\cos(k\theta)\cos(k\theta') \tag{3.24}$$

where v_{nk} is the nth root of $J_k'(v) = 0$ and J_k is a Bessel function of the first kind.

3.1.2.3 Spheres

The propagator for a spherical pore with an impermeable barrier at $r = a$ is [32]

$$P(\mathbf{r},t|\mathbf{r}',t') = \sum_k \sum_n \exp\left[-\alpha_{nk}^2 D\Delta/a^2\right]\left(\frac{2k+1}{2\pi}\right)\left(\frac{1}{a^3(j_k^2(\nu_{nk}) - j_{k-1}(\nu_{nk})j_{k+1}(\nu_{nk}))}\right)$$
$$\times j_k(\alpha_{nk}r/a)j_k(\nu_{nk}r'/a)P_k(\cos(\theta))P_k(\cos(\theta')) \tag{3.25}$$

where j_k are spherical Bessel functions, ν_{nk} is the nth root of $j_k'(\nu)$, and P_k are Legendre polynomials.

3.1.3 Velocity Autocorrelation and Self-diffusion Tensor

The autocorrelation function of a time-dependent quantity $A(t)$ is [31]

$$ACF(t) = \int_0^\infty A(t')A(t+t')dt' \tag{3.26}$$

Usually we assume we have a stationary ensemble, where only time differences are important. In practice, this means that the results of an MR experiment are not dependent on the starting time. Stationary ensembles also have the property that a time average of a quantity, like in Eq. 3.26, can be written as an average over all the particles in the ensemble. So we could also write Eq. 3.26 as $ACF(t) = \langle A(0)A(t)\rangle$. If we know $A(t)$ as a function of position, we can also write $ACF(t)$ as

$$\langle A(0)A(t)\rangle = \int\int \rho(\mathbf{r})P_s(\mathbf{r}|\mathbf{r}',t)A(\mathbf{r}')A(\mathbf{r})d\mathbf{r}'d\mathbf{r} \tag{3.27}$$

The autocorrelation function measures how quickly $A(t)$ loses memory of its previous values. The correlation time τ_c describes the time-scale of the memory loss and is defined by

$$\tau_c = \frac{\int_0^\infty A(0)A(t)dt}{A(0)^2} \tag{3.28}$$

When describing translational motion, we are particularly interested in the spectrum of the

velocity correlation function,

$$D_{\alpha\beta}(\omega) = \frac{1}{2} \int_{-\infty}^{\infty} \langle v_\alpha(0) v_\beta(t) \rangle \exp(i\omega t) dt \qquad (3.29)$$

where α and β denote Cartesian directions. The quantity defined in Eq. 3.29 is called the self-diffusion tensor. Unless the applied gradients fluctuate in direction during a measurement, cross-terms can be neglected. Because the gradients we use are applied in a single direction, we will drop the indices and assume that $\alpha = \beta$.

If a particle is undergoing a random walk with uncorrelated, infinitesimally spaced jumps, then the velocity correlation function is

$$\langle v(0) v(t) \rangle = 2D\delta(t) \qquad (3.30)$$

The corresponding diffusion spectrum is constant [31]. The velocity correlation for restricted diffusion has a peak at $t = 0$ and a residual negative correlation at small times [140]. The corresponding diffusion spectrum is equal to the intrinsic diffusion coefficient at high frequencies, but decreases as $\omega \to 0$. If the pore geometry is closed, then $D(\omega)$ goes to zero. In open pores, $D(\omega)$ is still reduced at low frequencies, but not all the way to zero. The reduction in $D(\omega)$ occurs at frequencies accessible in MR experiments. Figure 3.5 shows the velocity correlation functions and diffusion spectra for free and restricted diffusion.

3.2 Diffusion-weighting Gradients

At the beginning of this chapter, we said that diffusion leads to a reduced echo signal. In diffusion imaging, large diffusion-weighting gradients are used to sensitize the MRI signals to molecular diffusion. In a spin echo sequence, the area of the diffusion gradients on either side of the refocusing pulse should be equal. If there is no refocusing pulse (a gradient echo), then the net area of the diffusion gradients should be zero. These conditions ensure that the net phase of stationary molecules at the echo is zero. In this section, we introduce some diffusion pulse sequences used in this thesis.
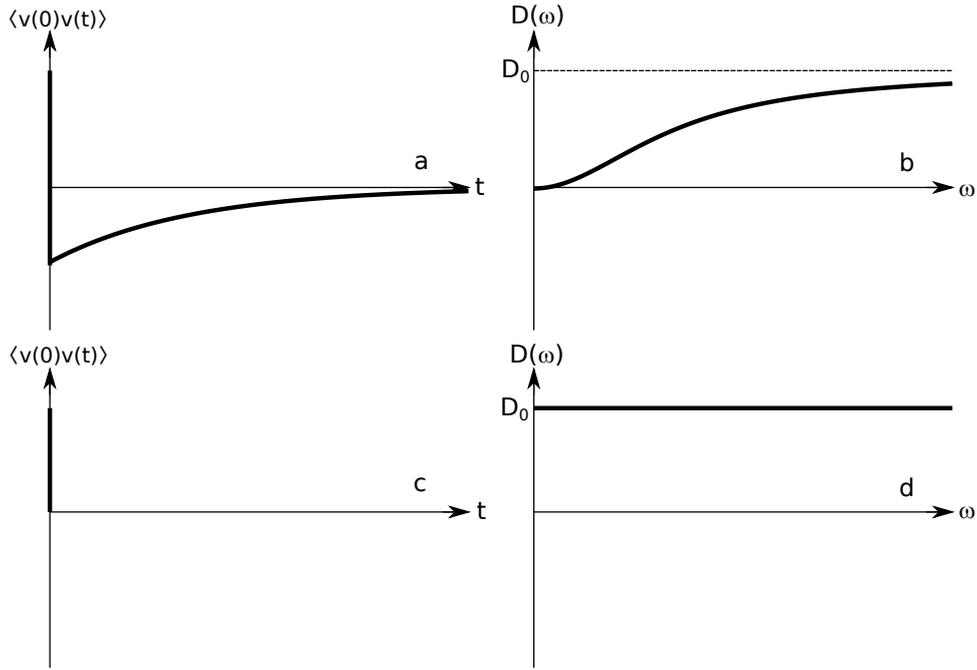
Figure 3.5: Velocity correlation functions $\langle v(0)v(t)\rangle$ and associated diffusion spectra $D(\omega)$. In (c) and (d), diffusion is free, with diffusion coefficient D_0. The velocity correlation is a δ-function and the diffusion spectrum is a constant (D_0) over frequency. In (a) and (b), diffusion is restricted. The velocity correlation is negative at short times and the diffusion spectrum is reduced at low frequencies, but rises to D_0 at high frequencies.

3.2.1 Pulsed Gradients

The commonly-used pulsed gradient spin echo sequence (PGSE) was developed by Steskjal and Tanner in 1965 [163]. It consists of two rectangular gradients with amplitude G and duration δ separated by a time Δ (see Figure 3.6). These three parameters (Δ, δ, and G) can be adjusted to give a different weighting to the signal. The 180° refocusing pulse is applied at time τ. A pulse sequence measures molecular displacements over a time scale called the diffusion time. This time scale is well-defined only when the gradient duration is very small (see Section 3.3.2). In a realistic situation, the time scale is described by an 'effective diffusion time.' For PGSE, the effective diffusion time is determined by the spacing and duration of the pulses. Bringing the two pulses closer together shortens the diffusion time. The effective diffusion time for a pulsed gradient sequence is

$$\Delta_{eff} = \Delta - \delta/3 \tag{3.31}$$

If the pulses were very narrow, the diffusion time would simply be Δ.
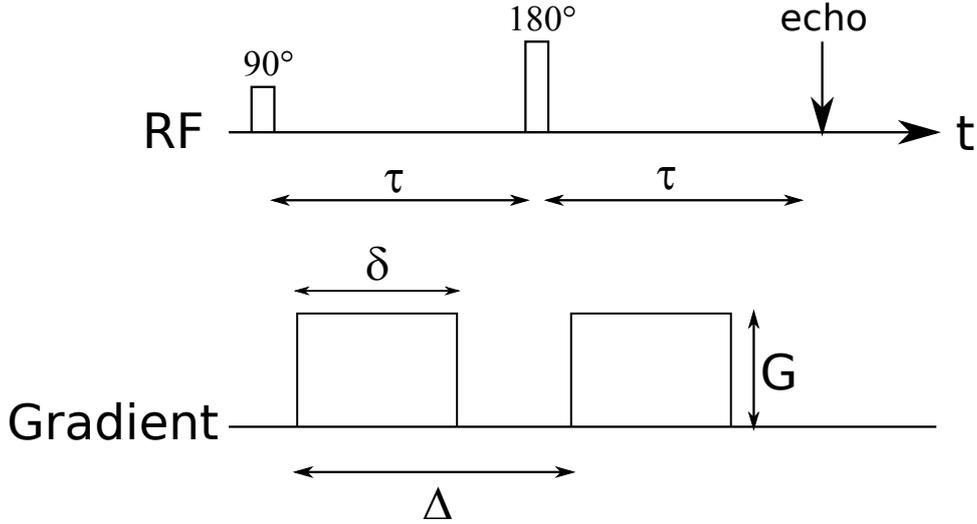
Figure 3.6: The PGSE sequence uses two rectangular gradients of duration δ and strength G. The first occurs just after the initial 90° RF pulse, with the second occurring a time Δ afterwards. The effective diffusion time is $\Delta - \delta/3$.

3.2.2 Oscillating Gradients

The oscillating gradient spin echo (OGSE) sequence replaces the pulsed gradients in PGSE with sinusoidally oscillating gradients applied for a duration T [58,154]. There are two versions used through this thesis: a sine gradient and a cosine gradient. Each gradient has N_{OG} lobes and amplitude G (see Figure 3.7). They can also be described by their angular frequency ω_m. The 180° refocusing pulse is applied at time τ. Mathematically, we have

$$\boldsymbol{G}(t) = \begin{cases} G\sin(\omega_m t) & 0 \leqslant t < T \\ G\sin(\omega_m(t-\tau)) & \tau \leqslant t < \tau + T \\ 0 & \text{otherwise} \end{cases} \tag{3.32}$$

for a sine gradient and

$$\boldsymbol{G}(t) = \begin{cases} G\cos(\omega_m t) & 0 \leqslant t < T \\ G\cos(\omega_m(t-\tau)) & \tau \leqslant t < \tau + T \\ 0 & \text{otherwise} \end{cases} \tag{3.33}$$

for the cosine gradient. More commonly, an OGSE sequence is described by its temporal frequency $f_m = \omega_m/2\pi$. The effective diffusion time for a sine gradient is [43]

$$\Delta_{eff} = \frac{3T}{8N_{OG}} \tag{3.34}$$

For a cosine gradient, the effective diffusion time is [43]

$$\Delta_{eff} = \frac{T}{4N_{OG}} \tag{3.35}$$

Diffusion times with oscillating gradients are proportional to the period T/N_{OG}, so higher gradient frequencies in the OGSE sequence correspond to a probe of shorter diffusion times. Theoretically oscillating gradients have the ability to probe smaller restriction sizes than pulsed gradients [154].

3.2.3 Diffusion Imaging

In diffusion-weighted imaging, we want diffusion measurements on a voxel-by-voxel basis. There are several ways to do this [79, 93].

One possibility is to add diffusion gradients to a spin echo imaging sequence. The diffusion gradients can be placed at the beginning of the sequence, around the 180° pulse, before the read gradient and signal acquisition (Figure 3.8). This method is what we assumed in Sections 3.2.1 and 3.2.2. We could also use a gradient echo [54]. Diffusion has the same effect on gradient echoes as on spin echoes, but because gradient echoes are affected by T_2^* and not T_2, echo times need to be shorter to maintain the same SNR.

A stimulated echo sequence uses three RF pulses to generate a stimulated echo [53, 116]. Diffusion gradients can be placed after the first and third pulses. One advantage is that it allows the effective diffusion time to be increased without additional T_2 decay.

Another possibility is to use an echo planar imaging (EPI) sequence [108,136]. This method is useful because it acquires images fast, often in a single shot (i.e. with a single RF excitation). Because of the increased speed, many images can be acquired and averaged in a short period of time. Its speed also helps to reduce motion artifacts. Possible downsides include lower
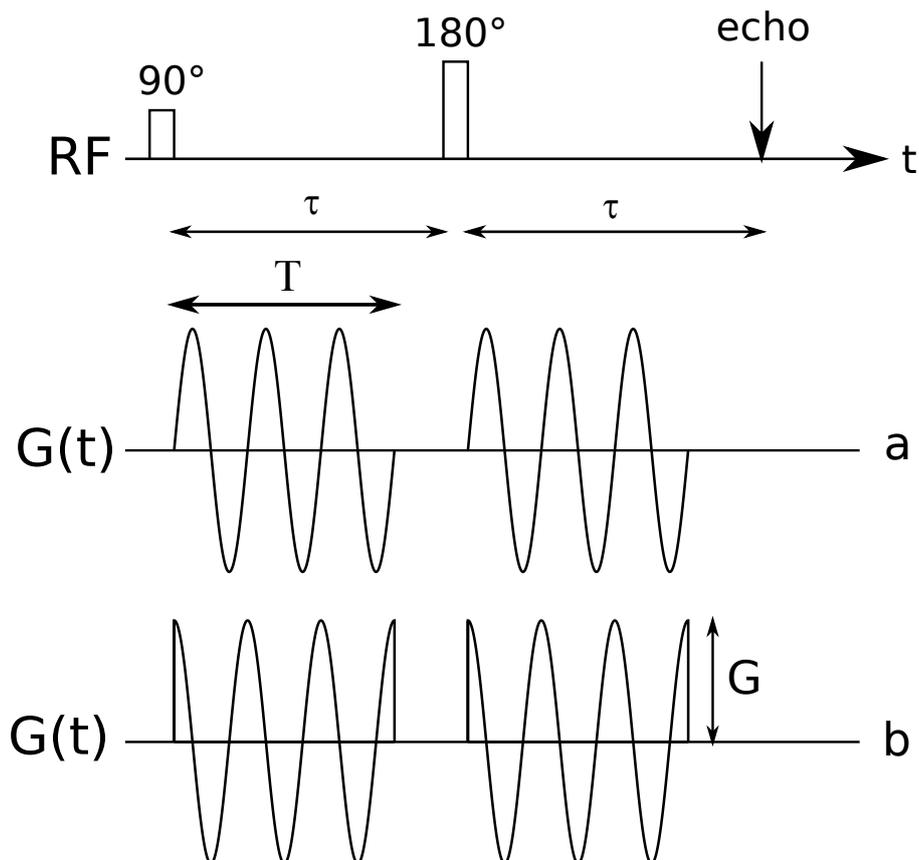
Figure 3.7: The OGSE sequence uses two sinusoidal gradients. The first occurs just after the initial 90° RF pulse, with the second occurring a time Δ afterwards, after the 180° pulse. (a) The sine-OGSE sequence uses two sinusoidal gradients of duration T and strength G. Each pulse has N_{OG} lobes and angular frequency $\omega_m = 2\pi N_{OG}/T$. Here $N_{OG} = 3$. (b) The cos-OGSE sequence uses two sinusoidal gradients of duration T and strength G.

resolution and vulnerability to susceptibility artifacts.

Echo planar imaging reads all of k-space in one excitation. It begins with a 90° excitation pulse, which is followed by two diffusion weighting gradients straddled around the 180° refocusing pulse (see Figure 3.9). During readout, it generates a series of gradient echoes by alternating the polarity of the readout gradient. Before each echo, a small phase encode gradient is applied. The effect is that k-space is traversed in a zig-zag pattern where the readout gradients horizontally sweep out k-space and the phase encoding gradients increment k_y

3.3 Diffusion Signals

This section explains basic principles of creating a magnetic resonance signal with diffusion weighting. It explains how gradients are used to weight the signal to diffusion and the concept
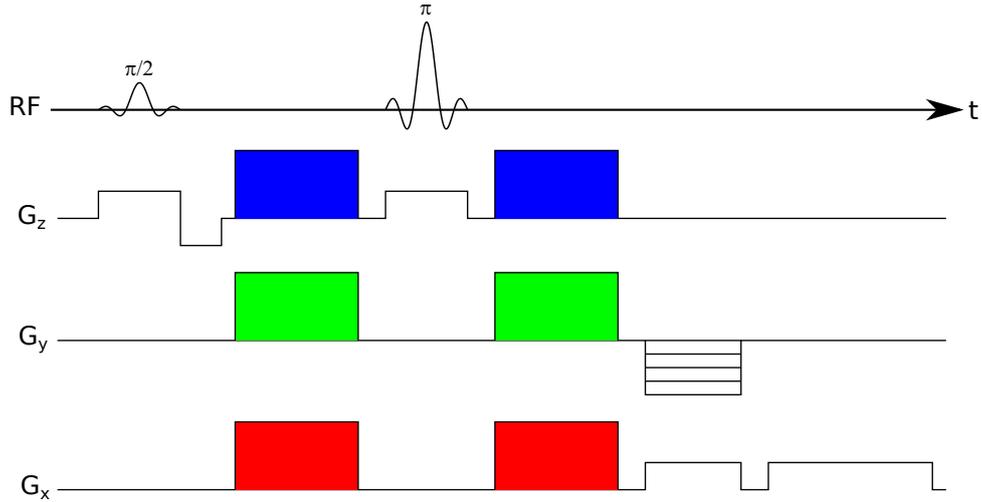
Figure 3.8: Diffusion-weighted spin echo MRI sequence. The diffusion weighting can be applied in any arbitrary direction using different combinations of G_x, G_y, and G_z. This figure assumes the diffusion gradients are rectangular in shape. The difference between this sequence and the spin echo sequence shown in Figure 2.13 is the addition of the diffusion gradients.

of an apparent diffusion coefficient as well as a diffusion tensor and how to measure them with MRI.

3.3.1 Bloch-Torrey Equation

When analyzing diffusion weighted spin echo sequences, it is useful to define an 'effective gradient' $\boldsymbol{G}^*(t)$. It lets us ignore the 180° refocusing pulse in the analysis. A positive gradient followed by a 180° RF pulse has the same effect on the phases of the spins as if a negative gradient had been applied without the 180° RF pulse. In other words, a spin echo and a gradient echo have the same diffusion weighting. We define $\boldsymbol{G}^*(t)$ as [31]

$$\boldsymbol{G}^*(t) = \begin{cases} -\boldsymbol{G}(t) & \text{before 180° pulse} \\ +\boldsymbol{G}(t) & \text{after 180° pulse} \end{cases} \tag{3.36}$$

All gradients occurring before the 180° pulse have been reversed. Figure 3.10 shows the effective gradient for a standard PGSE sequence.

We want to know how a particular diffusion-weighting gradient attenuates the MR signal. One way is to use the Bloch-Torrey equation. Torrey introduced a diffusion term into the Bloch equation (Eq. 2.32). Ignoring any RF pulses and working in the rotating frame, then the time
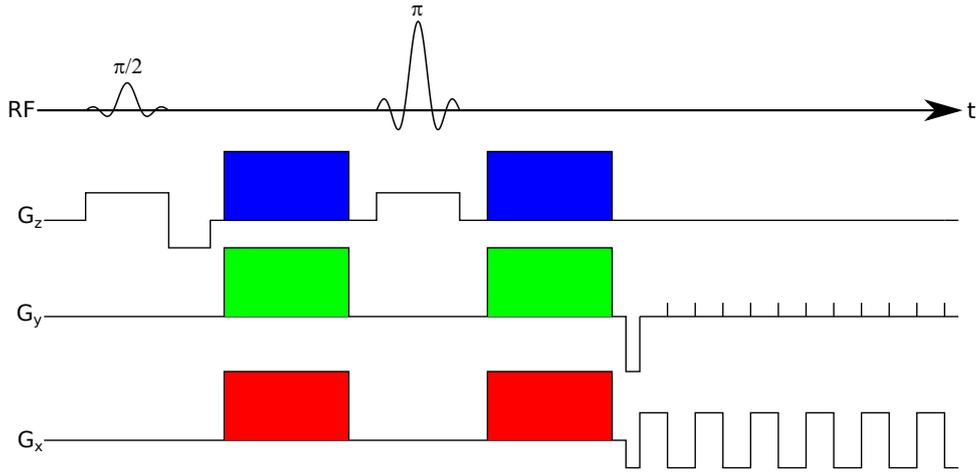
Figure 3.9: Diffusion-weighted EPI sequence. The diffusion weighting can be applied in any arbitrary direction using different combinations of G_x, G_y, and G_z. This figure assumes the diffusion gradients are rectangular in shape. The EPI sequence uses alternating gradients in the read direction with phase gradients applied between each change in read gradient to cover many, if not all, lines of k-space after just one excitation. This allows images to be collected faster.

evolution of the complex magnetization is [31, 176]

$$\frac{\partial M_{x'y'}}{\partial t} = i\gamma \boldsymbol{r} \cdot \boldsymbol{G} M_{x'y'} - \frac{M_{x'y'}}{T_2} + D\nabla^2 M_{x'y'} \tag{3.37}$$

where D is the diffusion coefficient. If we assume isotropic diffusion, then the solution to Eq. 3.37 can be found by assuming an $M_{x'y'}(r,t)$ of the form [31]

$$M_{x'y'}(\boldsymbol{r},t) = E(t)\exp\left(-i\gamma\boldsymbol{r}\cdot\int_0^t \boldsymbol{G}(t')dt'\right)\exp\left(-t/T_2\right) \tag{3.38}$$
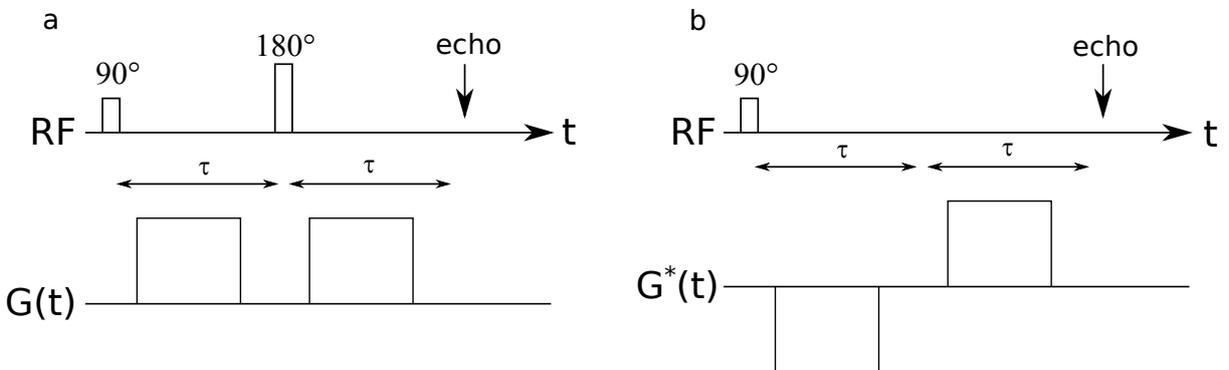
Figure 3.10: Applied gradient $G(t)$ and effective gradient $G^*(t)$ for a pulsed gradient sequence. In (a), two positive gradients are separated by a $180°$. In (b), instead of a $180°$ pulse, the first gradient is negative. The echo occurs at $t = 2\tau$. The effective gradient can be used to simplify calculations because the diffusion weighting it gives is the same as the applied gradient.

63

where $E(t)$ is a diffusion attenuation factor and the final term describes T_2 relaxation. At the echo centre ($t = 2\tau$), the exponent in the middle factor is zero because $\int_0^{2\tau} \mathbf{G}^*(t')dt' = 0$. Substituting Eq. 3.38 into the Bloch-Torrey equation leads to a differential equation for $E(t)$,

$$\frac{dE(t)}{dt} = -D\gamma^2 \left(\int_0^t \mathbf{G}^*(t')dt' \right)^2 E(t) \tag{3.39}$$

which has the solution [31, 79]

$$E(t) = \exp \left[-D\gamma^2 \int_0^t \left(\int_0^{t'} \mathbf{G}(t'')dt'' \right)^2 \right] \tag{3.40}$$

We can write Eq. 3.40 as

$$E = e^{-bD} \tag{3.41}$$

where

$$b = \gamma^2 \int_0^t \left(\int_0^{t'} \mathbf{G}^*(t'')dt'' \right)^2 dt' \tag{3.42}$$

Eq. 3.42 defines the b-value and gives the amount of signal weighting imparted by the diffusion gradients. The b-value depends on the gradient waveform. Eq. 3.41 can be used to calculate the diffusion coefficient of a liquid. We measure the signal with different b-values and then solve for D. This requires as few as two measurements:

$$D = \frac{\ln(E_2) - \ln(E_1)}{b_2 - b_1} \tag{3.43}$$

where E_1 and E_2 are diffusion-weighted signals measured with weightings of b_1 and b_2 [93].

For a given pulse sequence $G(t)$, one can calculate $E(t)$ to find the signal attenuation due to diffusion. For example, in the pulsed gradient spin echo sequence we have

$$G^*(t) = \begin{cases} -G & 0 < t < \delta \\ +G & \Delta < t < \Delta + \delta \\ 0 & \text{otherwise} \end{cases} \tag{3.44}$$

The relevant integrals are

$$\int_0^t G^*(t')dt' = \begin{cases} -Gt & 0 < t < \delta \\ -G\delta & \delta < t < \Delta \\ -G\delta + Gt & \Delta < t < \Delta + \delta \end{cases} \tag{3.45}$$

$$\left(\int_0^t G^*(t')dt'\right)^2 = \begin{cases} G^2t^2 & 0 < t < \delta \\ G^2\delta^2 & \delta < t < \Delta \\ G^2(\delta + t)^2 & \Delta < t < \Delta + \delta \end{cases} \tag{3.46}$$

and

$$\int_0^\delta G^2t^2 dt + \int_\delta^\Delta G^2\delta^2 dt + \int_\Delta^{\Delta+\delta} G^2(\delta+t)^2 dt$$
$$= G^2(\delta^3/3 + \delta^2(\Delta - \delta) + [(2\delta + \Delta)^3 - (\delta + \Delta)^3)]/3) \tag{3.47}$$
$$= G^2\delta^2(\Delta - \delta/3)$$

The echo attenuation is

$$E(2\tau) = \exp\left(-\gamma^2 G^2\delta^2 D(\Delta - \delta/3)\right) \tag{3.48}$$

From Eq. 3.48, the *b*-value for a PGSE sequence is

$$b_{PG} = -\gamma^2 G^2\delta^2(\Delta - \delta/3) \tag{3.49}$$

We can also calculate the echo attenuation from a single sine gradient with angular frequency ω_0 and duration T, where $G(t) = G\sin(\omega_m t)$. The relevant integrals are

$$\left(\int_0^t G(t)\right)^2 dt = \left(\frac{G}{\omega_m}\right)^2 (1 - \cos(\omega_m t))^2 \tag{3.50}$$

and

$$\left(\frac{G}{\omega_m}\right)^2 \int_0^T (1-\cos(\omega_m t))^2 dt$$

$$= \frac{G^2}{\omega_m^2}(6\omega_m t - 8\sin(\omega_m t) + \sin(2\omega_m t))/(4\omega_m)\Big|_0^T \qquad (3.51)$$

$$= \frac{3G^2 T}{2\omega_m^2}$$

The echo attenuation is

$$E(2\tau) = \exp\left(-\frac{3\gamma^2 G^2 T}{2\omega_m^2}D\right) \qquad (3.52)$$

The b-value is

$$b_{sin} = \frac{3\gamma^2 G^2 T}{2\omega_m^2} \qquad (3.53)$$

In an actual experiment, there will be two sine gradients, so the total b-value will be twice that of Eq. 3.53. Similarly, it can be shown that the b-value for a cosine gradient with duration T and angular frequency ω_m is

$$b_{cos} = \frac{\gamma^2 G^2 T}{2\omega_m^2} \qquad (3.54)$$

Signal attenuation in both sine and cosine sequences depend on G^2/ω_m^2. If we increase the gradient frequency and want to keep the same signal weighting, we have to increase G.

3.3.2 Narrow Pulse Approximation

Another method for calculating a diffusion signal involves the narrow pulse approximation, which assumes an idealized pulse sequence with infinitesimally short gradients [31, 129, 172]. After excitation in the PGSE sequence, a magnetic field gradient pulse is applied to the sample for a short time δ. The narrow-pulse condition assumes that the applied gradient pulses are short enough that motion during them can be neglected ($\delta \ll \Delta$). This causes the spins to obtain a phase based on their position at the time of the pulse. After the pulse, a particle at $\boldsymbol{r}$ will have a phase $\gamma \boldsymbol{G} \delta \cdot \boldsymbol{r}$. A 180° RF pulse is then applied to the system which reverses the phase of the spins. Another identical gradient pulse is then applied to the sample a time Δ after the first gradient pulse. Assuming the particle is now at $\boldsymbol{r}'$, then the phase change after the pulse

will be $\gamma \boldsymbol{G}\delta \cdot \boldsymbol{r}'$. The net phase shift is

$$\Delta \phi = \gamma \delta \boldsymbol{G} \cdot (\boldsymbol{r}' - \boldsymbol{r}). \tag{3.55}$$

If no diffusion occurs ($\boldsymbol{r} = \boldsymbol{r}'$), the phase acquired from the second pulse will be equal and opposite to the phase of the spin just before the pulse resulting in a net phase of zero. If diffusion occurs ($\boldsymbol{r} \neq \boldsymbol{r}'$), the mean squared phase of all spins will be nonzero and will cause a loss in MR signal which can be used to calculate the diffusion coefficient. The echo signal will then be the sum of transverse magnetizations, weighted by the probability of moving from $\boldsymbol{r}$ to $\boldsymbol{r}'$, which is $\rho(\boldsymbol{r})P_s(\boldsymbol{r}|\boldsymbol{r}',\Delta)$. The attenuation will be

$$E(\boldsymbol{G},\Delta) = \int \int \rho(\boldsymbol{r})P_s(\boldsymbol{r}|\boldsymbol{r}',\Delta)\exp\left(i\gamma\delta\boldsymbol{G}\cdot(\boldsymbol{r}'-\boldsymbol{r})\right)d\boldsymbol{r}d\boldsymbol{r}' \tag{3.56}$$

In terms of the particle's net displacement, $\boldsymbol{R} = \boldsymbol{r}' - \boldsymbol{r}$, the signal attenuation is

$$E(\boldsymbol{G},\Delta) = \int \int \rho(\boldsymbol{r})P_s(\boldsymbol{r}|\boldsymbol{r}+\boldsymbol{R},\Delta)\exp\left(i\gamma\delta\boldsymbol{G}\cdot\boldsymbol{R}\right)d\boldsymbol{r}d\boldsymbol{R} \tag{3.57}$$

The integral $\int \rho(\boldsymbol{r})P_s(\boldsymbol{r}|\boldsymbol{r}+\boldsymbol{R},\Delta)d\boldsymbol{r}$ is the definition of the average propagator $\overline{P(\boldsymbol{R},\Delta)}$. Therefore, we can write $E(\boldsymbol{G},\Delta)$ as

$$E(\boldsymbol{G},\Delta) = \int \overline{P(\boldsymbol{R},\Delta)}\exp\left(i\gamma\delta\boldsymbol{G}\cdot\boldsymbol{R}\right)d\boldsymbol{R} \tag{3.58}$$

Eq. 3.58 plays an important role in q-space imaging. In q-space imaging, we define $\boldsymbol{q} = \gamma \boldsymbol{G}\delta/2\pi$, so that $\overline{P(\boldsymbol{R},\Delta)}$ and $E(q,\Delta)$ become a Fourier transform pair. Measuring $E(\boldsymbol{q})$ at different $\boldsymbol{q}$ values (with fixed Δ) and then taking the Fourier transform can be used to map the averaged propagator [13, 84, 104, 188, 191]. On the other hand, if we know $\overline{P(\boldsymbol{R},\Delta)}$, we can obtain an analytical expression for $E(\boldsymbol{G},\Delta)$.

If the propagator is a Gaussian (Eq. 3.11) with gradients applied along the x-direction, we can write $E(G,\Delta)$ as

$$E(G,\Delta) = \exp\left(-\gamma^2 G^2 \delta^2 D\Delta\right) \tag{3.59}$$

The exponent can also be written as $-4\pi q^2 D\Delta$. A comparison of Eq. 3.41 with Eq. 3.59 shows that the b-value for a narrow pulse experiment is

$$b_{NP} = \gamma^2 G^2 \delta^2 \Delta \tag{3.60}$$

Eqs. 3.59 and 3.60 are valid if the narrow-pulse approximation holds.

3.3.3 Arbitrary Gradients

The narrow pulse approximation has limited applicability, so we want a more general method of analyzing diffusion signals. The phase shift of the ith particle at time t can be written as

$$\phi_i(t) = \gamma \int_0^t \boldsymbol{G}^*(t') \cdot \boldsymbol{r}_i(t')dt' \tag{3.61}$$

In general, the echo attenuation requires calculation of an ensemble average over all the spins, so the echo signal is [31, 164, 165, 166]

$$E(2\tau) = \langle \exp(i\phi_i(2\tau)) \rangle = \left\langle \exp\left(i\gamma \int_0^{2\tau} \boldsymbol{G}^*(t') \cdot \boldsymbol{r}_i(t')dt'\right) \right\rangle \tag{3.62}$$

The phase shift can also be written in terms of particle velocity $\boldsymbol{v}_i(t)$. Integrating Eq. 3.61 by parts and making use of the echo condition ($\int_0^{2\tau} \boldsymbol{G}^*(t')dt' = 0$), we find [165, 166]

$$\phi_i(2\tau) = -\gamma \int_0^t \boldsymbol{F}(t') \cdot \boldsymbol{v}_i(t')dt' \tag{3.63}$$

where $\boldsymbol{F}(t')$ is the time integral of the applied gradient,

$$\boldsymbol{F}(t) = \int_0^t \boldsymbol{G}^*(t')dt' \tag{3.64}$$

The steps for finding Eq. 3.63 are shown below [165, 166]:

$$
\begin{aligned}
\phi_i(2\tau) &= \gamma \int_0^{2\tau} \boldsymbol{G}^*(t') \cdot \boldsymbol{r}_i(t') dt' \\
&= \gamma \left(\int_0^{t'} \boldsymbol{G}^*(t'') dt'' \right) \cdot \boldsymbol{r}_i(t') dt' \Big|_0^{2\tau} - \gamma \int_0^{2\tau} \left(\int_0^{t'} \boldsymbol{G}^*(t'') dt'' \right) \cdot \frac{d\boldsymbol{r}_j}{dt'} dt' \\
&= 0 - \gamma \int_0^{2\tau} \left(\int_0^{t'} \boldsymbol{G}^*(t'') dt'' \right) \cdot \boldsymbol{v}_j(t') dt' \\
&= -\gamma \int_0^{2\tau} \boldsymbol{F}(t') \cdot \boldsymbol{v}_j(t') dt'
\end{aligned}
\tag{3.65}
$$

where we used the echo condition in the third line ($\int_0^{2\tau} \boldsymbol{G}^*(t) dt = 0$). Eq. 3.63 allows us to write the echo signal as

$$
E(2\tau) = \left\langle \exp\left(-i \int_0^{2\tau} \boldsymbol{F}(t') \cdot \boldsymbol{v}_j(t') dt' \right) \right\rangle
\tag{3.66}
$$

3.3.3.1 Gaussian approximation

Suppose we have a continuous random variable A, whose distribution is described by the probability density $p(A)$. If we want to find the ensemble average of $\langle e^{iA} \rangle$, we can perform a Taylor expansion, so that [79]

$$
\begin{aligned}
\langle e^{iA} \rangle &= \sum_{n=0}^{\infty} \frac{i^n}{n!} \langle A^n \rangle \\
&= 1 + i\langle A \rangle - \tfrac{1}{2}\langle A^2 \rangle +
\end{aligned}
\tag{3.67}
$$

Here $\langle A^n \rangle$ is called the nth raw moment. The Taylor expansion transfers the ensemble average from the exponential to powers of A.

Similarly, it can be shown that the logarithm of $\langle e^{iA} \rangle$ is [40, 79]

$$
\begin{aligned}
\ln(\langle e^{iA} \rangle) &= \sum_{n=1}^{\infty} \frac{i^n}{n!} \langle A^n \rangle_c \\
&= i\langle A \rangle_c - \tfrac{1}{2}\langle A^2 \rangle_c + \dots
\end{aligned}
\tag{3.68}
$$

where $\langle A^n \rangle_c$ is the nth order cumulant of A. Each cumulant is formed from combinations of

$\langle A^n \rangle$. The first three cumulants are

$$\langle A \rangle_c = \langle A \rangle$$
$$\langle A^2 \rangle_c = \langle A^2 \rangle - \langle A \rangle^2 \tag{3.69}$$
$$\langle A^3 \rangle_c = \langle A^3 \rangle - \langle A^3 \rangle_c - 3\langle A \rangle_c \langle A^2 \rangle_c$$

The first order cumulant is just the mean of A and the second order cumulant is the variance of A.

The general form of an echo described by Eq. 3.62 is an ensemble average of an exponential, where the random variable is the net phase shift $\phi_j(t)$. More specifically, the random variable is either the position $\boldsymbol{r}_j(t)$ (Eq. 3.61) or velocity $\boldsymbol{v}_j(t)$ (Eq. 3.63). Therefore, we can expand the echo signal in a series of moments (or cumulants).

Suppose we expand the signal in terms of position. If gradient magnitudes are small, then we can ignore higher order terms in the expansion, leading to the Gaussian phase approximation [79, 133, 165, 192],

$$E(2\tau) = \exp\left(-\frac{\gamma^2}{2} \int_0^{2\tau} \int_0^{2\tau} \boldsymbol{G}(t_1) \cdot \langle \boldsymbol{r}(t_1)\boldsymbol{r}(t_2) \rangle_c \cdot \boldsymbol{G}(t_2) dt_1 dt_2 \right) \tag{3.70}$$

It turns out that if there is no net flow, then all odd-order terms are zero [79]. The next highest term in the expansion would be a 4th-order term. If diffusion is unrestricted, higher order terms are zero. If we write the signal in terms of velocity cumulants, we have [55, 135, 165, 166]

$$E(2\tau) = \exp\left(-\frac{\gamma^2}{2} \int_0^{2\tau} \int_0^{2\tau} \boldsymbol{F}(t_1) \cdot \langle \boldsymbol{v}(t_1)\boldsymbol{v}(t_2) \rangle_c \cdot \boldsymbol{F}(t_2) dt_1 dt_2 \right) \tag{3.71}$$

Again, if gradient magnitudes are small, then higher order terms can be ignored. Note that if we substitute $\langle v(t_1)v(t_2) \rangle = 2D\delta(t_1 - t_2)$ (Section 3.1.3) into Eq. 3.71 we get

$$E(2\tau) = \exp\left(-\gamma^2 D \int_0^{2\tau} [F(t_1)]^2 dt_1 \right) \tag{3.72}$$

which is the same echo attenuation found using the Bloch-Torrey equation (Eq. 3.40).

3.3.4 Diffusion Spectrum

The diffusion signal can also be expressed in terms of the diffusion spectrum (Section 3.1.3). The diffusion spectrum is the Fourier transform of the velocity correlation function (Eq. 3.29), so we can also write the inverse transform [165]:

$$\langle v(t_1)v(t_2) \rangle = \frac{1}{\pi} \int_{-\infty}^{\infty} D(\omega) \exp(i\omega(t_1 - t_2)) d\omega \tag{3.73}$$

Substituting Eq. 3.73 into Eq. 3.71 gives another form for the signal [140, 167]

$$E(2\tau) = \exp\left(-\frac{\gamma^2}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 D(\omega) d\omega\right) \tag{3.74}$$

where

$$F(\omega) = \int_0^{2\tau} F(t) e^{i\omega t} dt \tag{3.75}$$

is the Fourier transform of $F(t)$. The signal attenuation depends on the shape of the gradient waveform, through $F(\omega)$, and on the diffusion spectrum $D(\omega)$. Eq. 3.74 also shows that different gradient waveforms probe the diffusion spectrum differently. The next section looks at a few waveforms.

3.3.4.1 Frequency spectrum

In this section we look at $|F(\omega)|^2$ for some typical gradient waveforms.

3.3.4.2 Pulsed gradients

The frequency spectrum for pulsed gradients of width δ and gradient amplitude G is [31]

$$|F(\omega)|^2 = \left[\frac{4G}{\omega^2} \sin\left(\frac{1}{2}\omega\delta\right) \sin\left(\frac{1}{2}\omega\Delta\right)\right]^2 \tag{3.76}$$

At zero frequency, the spectrum has a large lobe with a width on the order of $1/\Delta$. Shorter diffusion times broaden the peak. This peak makes the sequence unsuitable for extracting high frequency information from $D(\omega)$. If we let $D(\omega) = D_0$ (free diffusion) and integrate

$|F(\omega)|^2 D_0$, we recover the usual PGSE attenuation coefficient (Eq. 3.49) [31].

3.3.4.3 Oscillating gradients

Oscillating gradients provide a way to probe the high frequency regions of $D(\omega)$. It can be shown that $|F(\omega)|^2$ for a single sine gradient with N_{OG} lobes and angular frequency ω_m is [31]

$$|F(\omega)|^2 = \left[\frac{2G\omega_m}{\omega(\omega+\omega_m)(\omega-\omega_m)} \sin\left(N_{OG}\pi\omega/\omega_m\right) \right]^2 \tag{3.77}$$

Two main features in the spectrum are a large lobe at $\omega = 0$ and a smaller lobe at $\omega = \omega_m$. The spectrum for a single cosine gradient is

$$|F(\omega)|^2 = \left[\frac{2G}{(\omega-\omega_m)(\omega+\omega_m)} \sin\left(N_{OG}\pi\omega/\omega_m\right) \right]^2 \tag{3.78}$$

The frequency spectrum of a cosine has a lobe at $\omega = \omega_m$. But unlike the sine gradient, there is no lobe at $\omega = 0$. In the limit of large N_{OG} (infinite pulses), both Eq. 3.77 and Eq. 3.78 simplify. For a sine gradient, the spectrum becomes [31, 43]

$$|F(\omega)|^2 = \frac{\pi^2 G^2}{\omega_m^2} \left[4\delta(\omega) + \delta(\omega+\omega_m) + \delta(\omega-\omega_m) \right] \tag{3.79}$$

The signal attenuation is

$$E = \exp\left(-b_{sin} \cdot \left[\tfrac{2}{3}D(0) + \tfrac{1}{3}D(\omega_m) \right] \right) \tag{3.80}$$

At large n, the sequence picks out $D(\omega = 0)$ and $D(\omega = \omega_m)$. For free diffusion, $D(\omega)$ is independent of frequency ($D(\omega) = D_0$), so the sine sequence attenuation is

$$E = \exp\left(-b_{sin}D_0 \right) \tag{3.81}$$

which is the same attenuation we calculated in Eq. 3.53 using the Bloch-Torrey equation.

For a cosine gradient [43],

$$|F(\omega)|^2 = \frac{\pi^2 G^2}{\omega_m^2} \left[\delta(\omega + \omega_m) + \delta(\omega - \omega_m) \right] \tag{3.82}$$

and the signal attenuation is

$$E = \exp\left(-b_{cos} D(\omega_m) \right) \tag{3.83}$$

The cosine sequence measures $D(\omega)$ at $\omega = \omega_m$. An advantage of the cosine sequence is that it only picks out one frequency in the diffusion spectrum, whereas the sine sequence also picks out a zero frequency component. Measuring the diffusion-weighted signal with different gradient frequencies and using Eq. 3.83 allows the diffusion spectrum to be measured. This is called temporal diffusion spectroscopy [140]. Figure 3.11 compares the frequency spectra for PGSE and OGSE sequences.

3.3.5   Apparent Diffusion Coefficient

Molecules diffusing in a uniform medium with no barriers experience unrestricted diffusion. In non-uniform media (e.g. porous samples and cellular tissues) barriers hinder or restrict molecular displacements so that the diffusion depends on the time scale of the study, on the size of pores, and on the permeability of the barriers. Instead of a diffusion coefficient, we measure an apparent diffusion coefficient, or *ADC*. The signal in Eq. 3.41 becomes

$$E = \exp\left(-b \cdot ADC \right) \tag{3.84}$$

where $ADC < D$. Taking measurements at different b-values allow one to calculate the *ADC* (similar to Eq. 3.43). Understanding the influence of the restrictions or hindrances on the *ADC* gives information about the structural properties and geometry of the boundaries of the surrounding medium [79, 197, 198].

For narrow gradient pulses separated by diffusion time Δ, the *ADC* can be related to the
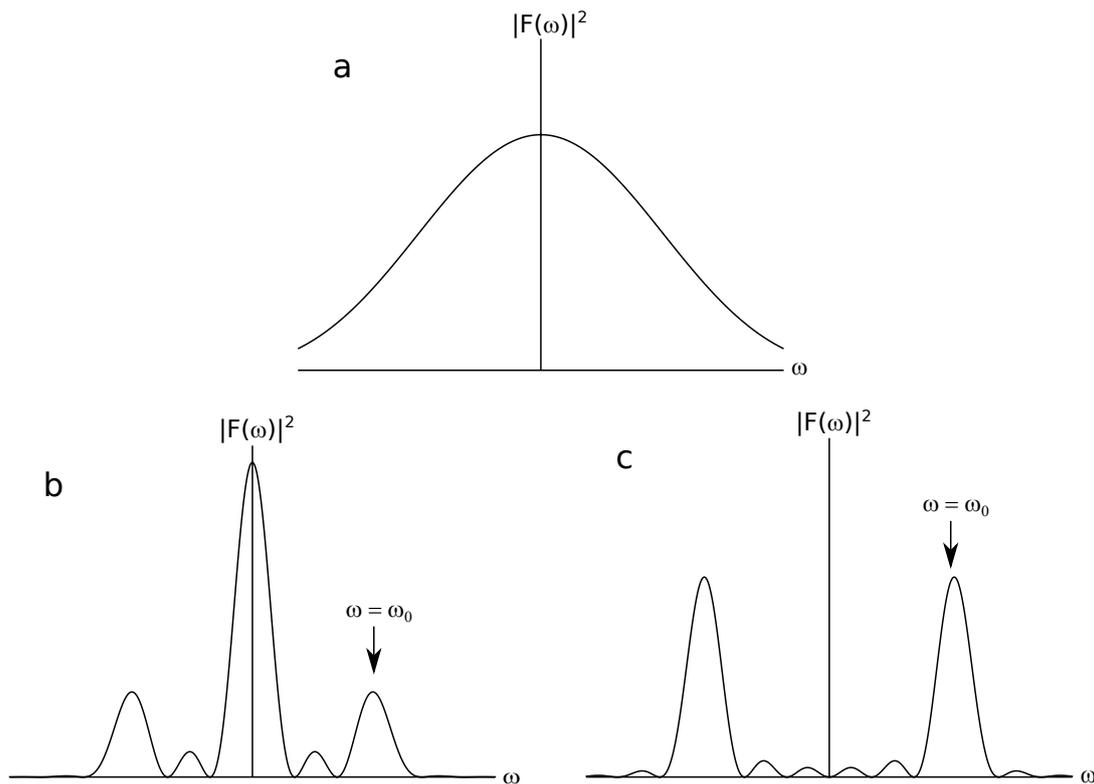
Figure 3.11: Frequency spectra for pulsed gradients and oscillating gradients. (a) Frequency spectrum for pulsed gradients. The spectrum has a broad lobe at $\omega = 0$. (b) Frequency spectrum for sine gradient with frequency ω_m. There are two smaller lobes at $\omega = \pm\omega_m$ and one larger lobe at $\omega = 0$. (c) Frequency spectrum for cosine gradient with frequency ω_m. There are two lobes at $\omega = \pm\omega_m$. Spectra in (b) and (c) were generated using the same pulse parameters (number of oscillations, pulse duration, angular frequency). An advantage of the cosine sequence is that it only picks out one frequency in the diffusion spectrum, whereas the sine sequence also picks out a zero frequency component.

mean squared displacement [52, 201]:

$$ADC_{NP}(\Delta) = \frac{\langle x^2(\Delta) \rangle}{2\Delta} \tag{3.85}$$

The quantity $\langle x^2(t) \rangle / 2t$ is the definition of a time-dependent diffusion coefficient $D(t)$ [157]. An *ADC* obtained from a narrow pulse experiment with diffusion time Δ is actually $D(t = \Delta)$. If the system is closed, then the mean squared displacement approaches a constant value at long times. According to Eq. 3.85, the *ADC* goes to zero at long measurement times. Comparing Eq. 3.84 to Eqs. 3.83 and 3.80, we also have the following relationships [43, 140]:

$$ADC_{sin}(\omega_m) = \tfrac{2}{3}D(0) + \tfrac{1}{3}D(\omega_m) \tag{3.86}$$

$$ADC_{cos}(\omega_m) = D(\omega_m) \tag{3.87}$$

The *ADC* measured with an idealized (infinite) sine gradient of angular frequency ω_m is a combination of diffusion spectrum components $D(0)$ and $D(\omega_m)$. The *ADC* measured with an idealized cosine gradient is equal to the diffusion spectrum component $D(\omega_m)$.

3.3.6 Diffusion Tensor Imaging

Diffusion is isotropic when the diffusivity is independent of direction. Anisotropic diffusion is when diffusivity is not the same in all directions. For instance, diffusion of water within a nerve axon (see Chapter 4) is more restricted across the axon than along the axon due to the axonal membrane. In anisotropic media, the measured *ADC* will depend on the direction of measurement [21].

Section 3.3.5 assumed a scalar *ADC*. This value is obtained from the projection of molecular displacements along one axis. In diffusion tensor imaging, we assume a three-dimensional Gaussian model for molecular displacements [17, 18],

$$P(\boldsymbol{R}, t) = \frac{1}{\sqrt{4\pi t |\mathbf{D}|}^3} \exp\left(\frac{-\boldsymbol{R}^T \mathbf{D}^{-1} \boldsymbol{R}}{4t} \right) \tag{3.88}$$

Instead of a scalar *ADC*, we use a nine element diffusion tensor **D** to describe diffusion [129]:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{3.89}$$

It is a symmetric tensor ($\mathbf{D} = \mathbf{D}^T$). The diagonal elements represent second displacement moments along the three axes ($\langle x^2 \rangle$, $\langle y^2 \rangle$, $\langle z^2 \rangle$), while the off-diagonal elements represent correlations between orthogonal displacements ($\langle xy \rangle$, $\langle xz \rangle$, $\langle yz \rangle$). To find the elements of the tensor, we need diffusion measurements made in multiple directions (at least 6).

The signal is now [79]

$$E = \exp\left(-b\hat{\boldsymbol{G}}^T \mathbf{D} \hat{\boldsymbol{G}}\right) \tag{3.90}$$

where $\hat{\boldsymbol{G}}$ is a unit vector giving the gradient direction. The amount of weighting is now represented by a b-matrix instead of a single *b*-value [114, 115]. The elements of the diffusion tensor can be found using multivariate linear regression [79, 83, 95]. If diffusion is isotropic, then all off-diagonal elements in the tensor are zero and the diagonal terms are all equal, with $ADC = (D_{xx} + D_{yy} + D_{zz})/3$.

Diffusion in each voxel can be visualized using ellipsoids. With isotropic diffusion, the diffusion propagator is spherically symmetric. Its value depends only on the displacement from the origin. Therefore, surfaces of constant probability for molecular displacements are concentric spherical shells. When diffusion is anisotropic, the surfaces of constant probability are ellipsoids. If diffusion is highly anisotropic, with diffusion higher along one axis, the ellipsoid is prolate. If diffusion is greater along two axes but smaller along the third, the ellipsoid will be oblate (see Figure 3.12).

3.3.6.1 Diffusion Tensor Metrics

The elements of the diffusion tensor depend on the relative orientation of the laboratory and principal axes. Scalar metrics that are independent of the orientation of the sample are therefore desirable. A few are described below [79].
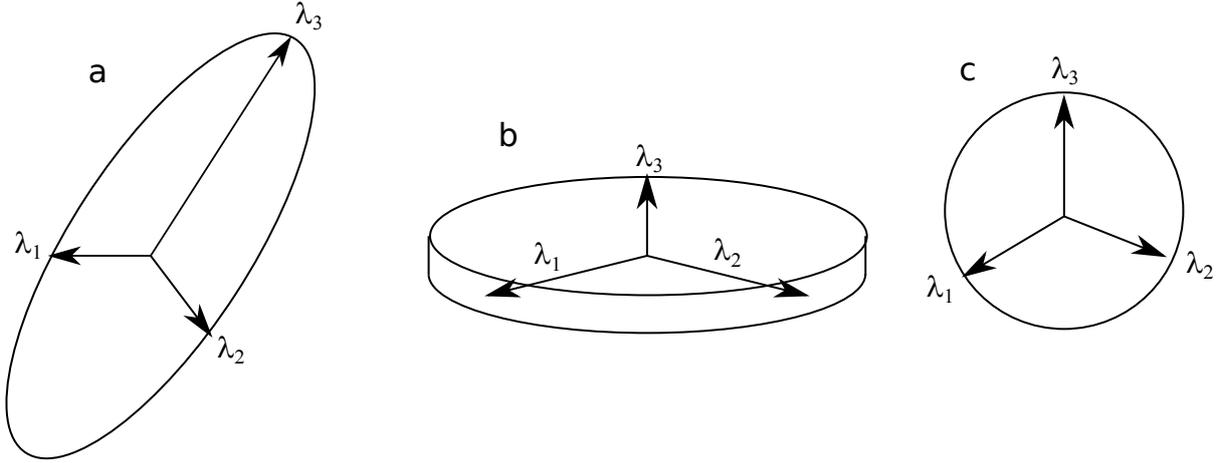
Figure 3.12: Three diffusion ellipsoids. (a) prolate ellipsoid ($\lambda_1 > \lambda_2 \approx \lambda_3$) (b) oblate ellipsoid ($\lambda_1 \approx \lambda_2 > \lambda_3$). (c) spherical ellipsoid ($\lambda_1 \approx \lambda_2 \approx \lambda_3$). These ellipsoids represent surfaces of constant probability. In (a) and (b) diffusion is anisotropic. In (c) the diffusion is isotropic.

3.3.6.2 Radial and Axial Diffusivity

The first set of scalar quantities are the three eigenvalues, λ_1, λ_2, and λ_3. They are the diffusivities along the principal axes of the tensor, sorted according to size ($\lambda_1 > \lambda_2 > \lambda_3$). They describe the degree of diffusion anisotropy. If the tensor is axially symmetric ($\lambda_1 \gg \lambda_2 \approx \lambda_3$), then λ_2 and λ_3 can be combined to give a mean radial diffusivity $\lambda_\perp = (\lambda_2 + \lambda_3)/2$. The diffusivity along the longitudinal axis is called the axial diffusivity, $\lambda_\parallel = \lambda_1$. In white matter, it is assumed that the direction of greatest diffusivity (λ_1) points in the direction of an axon (see Figure 3.13).

3.3.6.3 Mean diffusivity

The mean diffusivity represents an average *ADC* in a voxel. It is the average of the three principal diffusivities, or one-third the trace of the diffusion tensor, $Tr(\mathbf{D})/3$. It describes the size the ellipsoid. With *b*-values of around 1500 s/mm^2, the mean diffusivity is relatively uniform over the human brain (0.7×10^{-3} mm^2/s) [77].

3.3.6.4 Fractional Anisotropy

The sample variance Var(λ) of the principal diffusivities

$$Var(\lambda) = \frac{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}{3} \qquad (3.91)$$
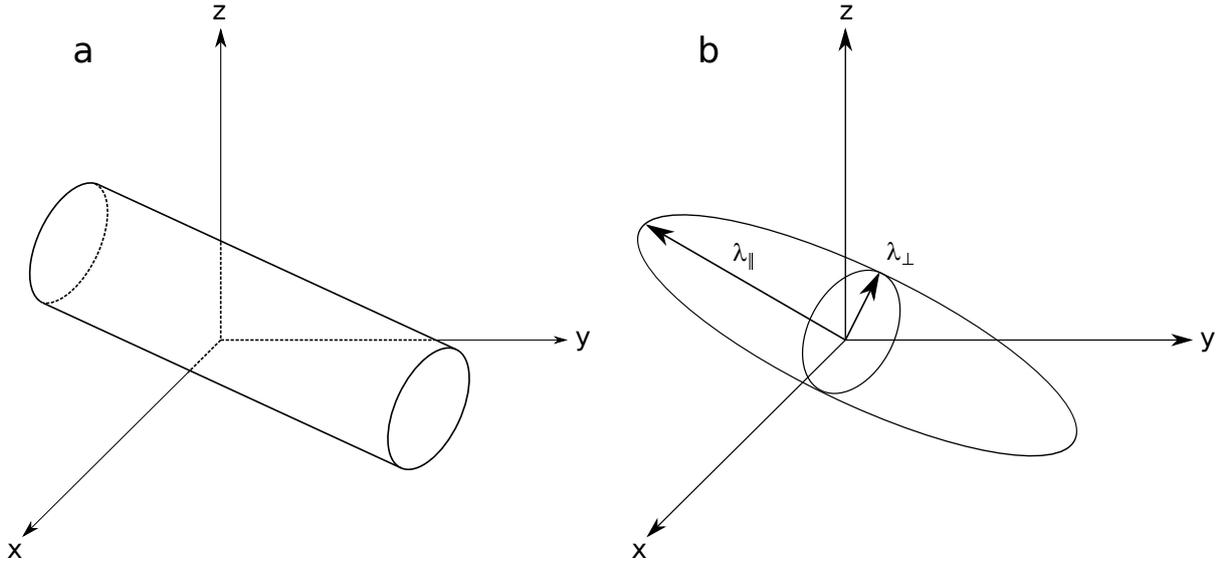
Figure 3.13: Ellipsoids are used to visualize diffusion. In highly anisotropic tissues such as white matter, the direction of greatest diffusivity is assumed to be parallel to the direction of the fibre. The direction of the fibre is used in diffusion tensor tractography methods [19, 37, 78, 92, 125] to study connectivity in the brain.

is an invariant quantity that can be used to create two anisotropy measures: relative anisotropy (*RA*) and fractional anisotropy (*FA*). The *RA* is given by

$$RA = \frac{Var(\lambda)}{\langle D \rangle},$$

(3.92)

and the FA given by

$$FA = \frac{3}{\sqrt{2}} \frac{\sqrt{Var(\lambda)}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

(3.93)

Fractional anisotropy gives the fraction of diffusion tensor magnitude that is anisotropic. Its value lies between 0 (completely isotropic diffusion) and 1 (diffusion along one direction only). Reported values in human white matter range from 0.5 to 0.8 [24,69,112]. Fractional anisotropy is highest in the corpus callosum [27].

3.4 More Sophisticated Models

The principles in the previous sections form the basis of more complicated models. Some of these are described below.

3.4.1 Multi-tensor Model

The multi-tensor model generalizes the single tensor model to include multiple tensors [79]. It assumes that each voxel has n populations of fibres, each described by its own Gaussian diffusion tensor. The displacement probability is

$$P(\boldsymbol{R}, \Delta) = \sum_{i=1}^{n} f_i \mathcal{G}(\boldsymbol{R}, \mathbf{D}_i, \Delta), \tag{3.94}$$

where f_i are population fractions and $\mathcal{G}(\boldsymbol{R}, \mathbf{D}, \Delta)$ is a Gaussian with zero mean and covariance $2\mathbf{D}_i\Delta$, with Δ as the diffusion time. Water exchange between populations is assumed to be zero. The normalized signal is

$$E = \sum_{i=1}^{n} f_i \exp(b\hat{\boldsymbol{G}}^{\boldsymbol{T}}\mathbf{D}_i\hat{\boldsymbol{G}}), \tag{3.95}$$

where $\hat{\boldsymbol{G}}$ is a unit vector pointing along the direction of the gradient. Model parameters are the f_i and the elements of each tensor. The principal eigenvectors of each tensor should give the direction of a different fibre population. Parameters are found using non-linear regression (for example using a Levenberg-Marquardt algorithm [95]).

The multi-tensor model has a large number of free parameters. While the single tensor model has 6 free parameters, a multi-tensor model with $n = 2$ has 13 free parameters. The model assumes that the number of populations is known beforehand. Due to limitations (such as acquisition time, number of measurements needed, etc), the number of compartments has usually been limited to $n = 2$. In voxels with one dominant fibre direction, the two tensor model estimates are less accurate. Some groups are working on ways to determine what model to use beforehand [5,67,138,180]. For $n = 2$, measurements are made with higher b-values compared to a single tensor model. One study found that b-values in the range $1800 - 2400$ s/mm^2 provided the best parameter estimates [138].

3.4.2 Ball and Stick model [22, 67]

The ball and stick model assumes two or more compartments. Molecules in one compartment (the ball) undergo isotropic Gaussian diffusion. Molecules in the other compartments only

diffuse in one direction (Gaussian displacements in one direction). These are the 'sticks,' which represent the fibres. The sticks can be oriented in any direction.

3.4.3 Composite Hindered and Restricted Model of Diffusion (CHARMED)

The CHARMED model [12] also assumes two or more compartments. One compartment, the "hindered" compartment, represents extra-axonal space where particle displacements are modeled by a Gaussian function. When diffusion is hindered, particles are slowed down by obstacles, but not totally trapped [189]. This was assumed to be the case in extra-axonal space. The *ADC* at long times will approach a constant value. In the other compartment(s), representing cylindrical axons, molecules undergo restricted diffusion. The model can accommodate multiple restricted compartments with the cylinders oriented in any direction. Images are acquired with a range of *b*-values (up to 10000 s/mm^2) and with directions arranged over a shell (up to 30 directions at the largest *b*-value).

3.4.4 Diffusion Spectrum Imaging (DSI)

This method is designed to measure the diffusion propagator $P(\boldsymbol{R}, \Delta)$ without making any assumptions about the microstructure [79, 102, 120, 187]. It uses the *q*-value model, which says that for narrow pulses, a set of PGSE measurements at fixed diffusion time Δ is the Fourier transform of $P(\boldsymbol{R}, \Delta)$ (Section 3.3.2). In DSI, signals are collected over a Cartesian grid in *q*-value and then Fourier transformed to get $P(\boldsymbol{R}, \Delta)$. Collecting many data points over three dimensional Cartesian space (often on the order of 500 *q*-values) requires long acquisition times, which makes it an inefficient sampling scheme.

3.4.5 High Angular Resolution Diffusion-weighted Imaging (HARDI)

This acquisition scheme involves collecting images with a single *b*-value sampled over a spherical shell in *q*-space [42, 79]. It provides information on the diffusion-weighted signals' angular dependence, which is important for resolving crossing fibres. It requires less time to acquire

data over the shell (typically on the order of 50 q-space samples) than over a full three dimensional grid. Since it is not possible to perform a three dimensional Fourier transform with a single shell, an assumption of the radial dependence of $E(\boldsymbol{q})$ needs to be made.

3.4.6 Q-ball Imaging

This method uses HARDI acquisition data to approximate a diffusion orientation distribution function ($dODF$) [120, 179, 181]. The $dODF$ describes the relative number of particles that have diffused along a given direction. It approximates the $dODF$ using a Funk-Radon transform of the q-value data. The method uses b-values of 4000 s/mm^2 or higher and was able to resolve fibre crossings of 45 degrees.

3.4.7 Spherical Deconvolution

The spherical deconvolution method is designed to estimate a fiber orientation distribution function ($fODF$) [8, 178]. The $fODF$ quantifies the fraction of fibres pointing along a certain direction in a voxel. The method works by assuming the final diffusion-weighted measurements are the sum of measurements for each fibre population orientation, weighted by the fraction of fibres with that orientation. Each measurement is a convolution of the $fODF$ with the measurement for a single fibre population. The goal of spherical deconvolution is to find $fODF$ through deconvolution. This requires a diffusion model for the fibres. For example, one could use the ball and stick model [8]. Originally, this method was susceptible to noise which produced spurious peaks and lobes, though later refinements to the model produced smoother output [79, 153, 177].

3.4.8 Diffusion Orientation Transform (DOT)

This model uses HARDI data to estimate the propagator $P(\boldsymbol{R})$ at a fixed radius R_0 [79, 137]. It can be evaluated at any R_0, which is useful because large R_0s will show increased separation between fibres. In order to calculate it, one needs to approximate how the signal decays with $|\boldsymbol{q}|$. Two possible approximations are monoexponential or bi-exponential decay [137], with the

more complex model more sensitive to crossing fibres.

3.5   Monte Carlo Simulation

A Monte Carlo method uses random number sampling to estimate a quantity [47]. Monte Carlo methods became very popular after the introduction of computers. With a computer, it is possible to generate large sequences of pseudorandom numbers in short amount of time.

In order to understand the effect of geometries, pulse sequences, or properties of the sample on diffusion weighted MR signals, we use Monte Carlo simulations. The geometry of the sample and diffusion properties of the sample, such as permeability, and diffusion coefficients, are programmed into the simulation. These parameters can be controlled so that their effects on the diffusion MR signal can be studied.

In this thesis a Monte Carlo simulation was used to study the diffusion spectrum from a variety of geometries using the cosine sequence. Before the simulation begins, N particles are randomly distributed over a lattice. At $t = 0$, the magnetization vectors of the particles are aligned in the transverse plane as if just tipped by a $90°$ RF pulse. During the simulation, ith particle's transverse magnetization vector is stored as a phase variable ϕ_i.

At each time step, the position of each particle is updated by generating a three dimensional step vector with random orientation in space. For three dimensions, the magnitude of the step vector is $\sqrt{6D\Delta t}$, where D is the diffusion coefficient of the region containing the particle, and Δt is the length of the time step. The new position of the particle then becomes

$$\boldsymbol{r}_i(t_{j+1}) = \boldsymbol{r}_i(t_j) + \sqrt{6D\Delta t}\,\hat{\boldsymbol{U}} \qquad (3.96)$$

where $\boldsymbol{r}_i(t_{j+1})$ is the updated position of the particle, $\boldsymbol{r}_i(t_j)$ is the original position, and $\hat{\boldsymbol{U}}$ denotes a random step vector. In practice, we usually choose the step direction by picking a random point on a sphere.

The samples modeled in this thesis all have impermeable boundaries. The implications of this assumption are discussed in Chapter 8. If the proposed trajectory of any particle intersects a boundary, then the particle is mirror reflected. The simulation allowed for the possibility of

multiple reflections in a single time step.

At time steps where gradients were applied, each particle has its transverse magnetization incremented through an angle $\gamma \boldsymbol{G}(t_j) \cdot \boldsymbol{r}_i(t_j) \Delta t$, where $\boldsymbol{r}_i(t_j)$ is the position of the ith particle and $\boldsymbol{G}(t_j)$ is the gradient vector. The $180°$ refocusing pulse is not explicitly included. In order to take into account the $180°$ refocusing pulse, we reverse the sign of $\boldsymbol{G}(t)$ for the first half of the simulation (Section 3.3.3).

The total signal at the echo time is taken to be the net transverse magnetization of the N particle ensemble, given by

$$E = \left| \frac{1}{N} \sum_i \boldsymbol{M}_{xy}^i \right| \tag{3.97}$$

where $\boldsymbol{M}_{xy}^i$ is the transverse magnetization vector of the ith particle.

Chapter 4

AxCaliber model

This chapter discusses the development of a method to measure micron-sized restrictions in samples and an application of the method for inferring axon diameter sizes in white matter in the brain.

4.1 Introduction

In this section, axons and the importance of knowing their diameters are explained followed by methods to determine axon diameters.

4.1.1 Axons

The CNS is made up of neurons and glial cells. There are about 100 billion neurons in the human brain, but almost 10 times more glial cells [20]. Neurons have three main regions: the cell body, axons, and dendrites. The cell body contains the nucleus and other organelles. Axons are long, slender fibres that lead away from the cell body. Axons carry electrical impulses away from the cell body. Many axons are surrounded by a myelin sheath. Dendrites are extensions that receive signals from other neurons. Information is transmitted across synapses. Glial cells are a class of cells that play a supporting role in the CNS. For example, oligodendrocytes are involved in myelin production [20, 121]. Figure 4.1 shows the main features of a neuron.
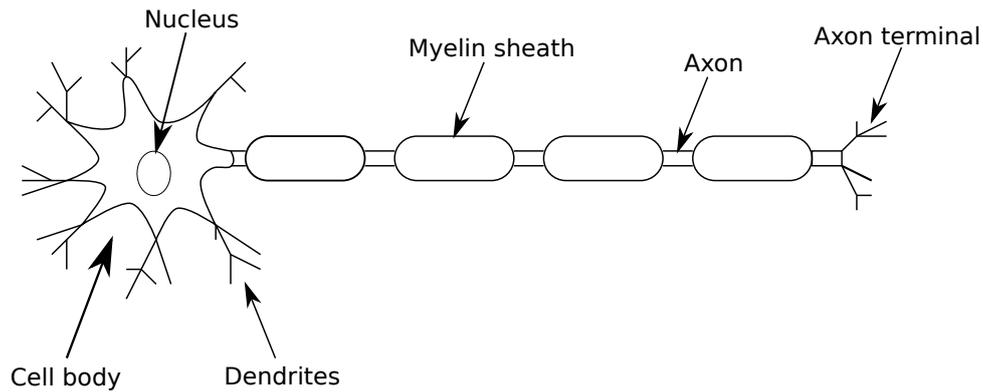
Figure 4.1: Diagram showing main features of a neuron. The cell body contains the nucleus. Dendrites receive signals from other neurons. The axon carries signals away from the cell body to the axon terminal. It is sometimes surrounded by a myelin sheath.

4.1.2 The Importance of Axon Diameters

Measuring axon diameter distributions is important for neuroscience because axon diameter is directly proportional to nerve conduction velocity [11,70,151,173,186]. Recent studies indicate possible changes in axon diameter distributions associated with diseases such as diabetes [107], Alzheimer's disease, [9] autism [68, 142], dyslexia [134], and schizophrenia [146, 150]. Currently, the only way to measure axon diameters is highly invasive, requiring sectioning of *ex vivo* tissue using histological procedures such as electron microscopy [185]. These measurements have limitations due to fixation and cutting artifacts, the need to use post mortem tissue, and the use of only small tissue sample sizes [16, 185]. These methods also do not allow repeated measures on the same subject over time to view development, damage due to disease or trauma, changes due to aging, or to monitor the effects of interventions/treatments. Recently there has been a push to develop models to infer axon diameter distributions from *in vivo* magnetic resonance images [7, 11, 16, 66] so that changes in axon diameters can be monitored over time.

4.1.3 Measurements of Axon Diameters

Axons are typically modeled as impermeable parallel cylinders with intra- and extra-axonal water each with their own diffusion coefficient. Diffusion perpendicular to the axons is measured as a function of diffusion time using magnetic resonance imaging. The diffusion signals are fit to models of diffusion in cylinders to infer the diameter of the cylinders or to obtain an

86

inferred axon diameter distribution (ADD). For instance, with the AxCaliber method [11], signal data are first collected at a range of diffusion times and gradient strengths. During analysis, signals are decomposed into intra-axonal and extra-axonal components. Water inside the axons is assumed to undergo restricted diffusion, while extra-axonal water is modeled as undergoing hindered diffusion.

Another method, ActiveAx [6], extends and optimizes these methods to determine the accuracy and precision with which this important new biomarker, axon diameter, can be estimated in live human subjects [7]. The work from this group combined a simplified version of composite hindered and restricted model of diffusion (CHARMED) [12] with high-angular-resolution diffusion imaging (HARDI) and a model with a single axon diameter [7]. It has been modified to be robust in the presence of orientation dispersion but is insensitive to axons < 2.5 μm in diameter [199].

These models have their limitations. Accurate models for extra-axonal water diffusion, permeable axon membranes, and exchange of water between compartments are still being developed [132]. Estimating axon diameters in the presence of orientation dispersion is challenging [152, 199]. AxCaliber assumes parallel axons and requires that the diffusion gradients be applied perpendicular to the axon. The need for many diffusion measurements can also lead to long scan times [11].

AxCaliber was used to infer axon diameters in human corpora callosa [66]. When results from this study were compared with results measured histologically by light microscopy, it was found that AxCaliber-inferred axon diameters were considerably larger than those found in the tissue being studied (by ≈ 2 μm) [73]. These axon diameters inferred from MRI are biased toward large axons. The comparison with tissue samples concluded that MRI derived axon diameter measurements do not yet resolve axons in the 1 to 2 μm range, which constitute the majority of cortical connections [73].

Both ActiveAx and AxCaliber use the pulsed gradient spin echo sequence [163]. The PGSE sequence limits the size of axons that can be inferred using these methods. The PGSE sequence cannot adequately access short enough diffusion times to distinguish small axon diameters in mice or humans. Axon diameters range from 0.16 to 9 μm [101], whereas most PGSE

sequences can only distinguish diameters 2.5 μm and larger [45]. Thus PGSE is irrelevant for rodent studies, because the majority of the axons in rodent brain are much smaller. For instance rat corpus callosum axon diameters are smaller than 4 μm with a mean around 1 μm [16] and mouse optic nerve axons range from 0.2 to 2.4 μm [77]. Distinguishing smaller axon diameters would provide important new information for neuroscience about the function of nerves because axon diameter is directly related to nerve conduction velocity [7, 70, 151, 173, 186].

Thus another pulse sequence should be used in order to infer smaller axon diameters allowing the ActiveAx and AxCaliber methods to be applied to mice. OGSE easily achieves diffusion times on the order of 1 ms and has even achieved a diffusion time of 0.375 ms [43]. Recently, the first demonstration of a trapezoidal OGSE sequence on humans showed its potential for investigating microstructure information on a human MR system [183].

Using OGSE sequences can improve the ability to distinguish smaller axon diameters. As Ref. [7] notes: "More significant improvements may come from replacing the standard PGSE sequence with other diffusion-sensitive sequences such as oscillating gradient...experiments, which may be more sensitive to microstructural parameters...Combination of the experiment design with these other pulse sequences should allow the *a priori* range of axon diameters to extend to include smaller diameters. This should provide protocols with sensitivity to wider ranges and provide more discriminative axon diameter indices." The same group performed experiments with optimized gradient waveforms (GEN) as a first step toward changing the gradient waveform in PGSE and were able to make axons with smaller radii more distinguishable with GEN than with PGSE [45, 160]. They suggest that oscillating waveforms could go even further than GEN [160].

Some studies have combined PGSE and OGSE for intermediate sizes using clinically feasible gradient strengths. These studies take advantage of the unique time scales available by each method to obtain intermediate size ranges that are wider than just using PGSE to measure cell sizes and volume densities [76, 147]. Using the Non-Uniform Oscillating Gradient Spin Echo (NOGSE) method, two oscillating frequencies are combined to benefit from both low frequency oscillations (similar to the time-scale from PGSE) and high frequency oscillations

(similar to the time-scales from OGSE) [159]. This method, which offers better sensitivity for smaller diameter measurements than PGSE alone, is not suitable as of yet for high gradient strengths but can be adapted for clinical imaging [141].

This chapter develops an important new method to distinguish smaller axon diameters (< 5 µm) by combining cosine OGSE sequences with AxCaliber [11] methods to distinguish smaller axons. This new method can be used in a research environment to study smaller axons in rodent models of diseases. We present Monte Carlo computer simulations of various geometries to show the feasibility of this method in pre-clinical settings with higher field strength imagers with large gradient strengths. For simplicity these studies also assume parallel axons with gradients perpendicular to the axons. The AxCaliber model assumes hindered diffusion of water in the extra-cellular space. We show this assumption breaks down for certain diffusion times and suggest improvements for the reliability of the extra-axonal diffusion in the model.

4.2 Methods

In this section, the methods used to infer axon diameters are discussed. The original CHARMED method and AxCaliber method are discussed first. Then the modifications we made from the use of PGSE to the OGSE pulse sequence are explained. Then the computer simulations using the method are explained.

4.2.1 CHARMED

As mentioned before (Section 3.4.3), CHARMED is a framework that has been used to describe the diffusion process in white matter [12]. Both restricted (non-Gaussian) and hindered (Gaussian) diffusion contribute to signal decay. It assumes that restricted diffusion occurs in the intra-axonal regions and hindered diffusion occurs in the extra-axonal region.

Suppose we have a bundle of parallel axons. One axon is modeled in Figure 4.2. The axons point in the $\hat{\boldsymbol{n}}$ direction. A gradient $\boldsymbol{G}$ (or $\boldsymbol{q}$) is turned on and a diffusion signal is collected. Diffusion of water occurs both parallel and perpendicular to the axons. Molecular displacements in each direction are independent, so the total propagator can be written as a product of parallel

and perpendicular propagators. This implies that the total signal is a product of signals due to parallel and perpendicular displacements:

$$E(\boldsymbol{q}, \Delta) = E_{\parallel}(\boldsymbol{q}_{\parallel}, \Delta) E_{\perp}(\boldsymbol{q}_{\perp}, \Delta) \qquad (4.1)$$

where $E_{\parallel}$ and $E_{\perp}$ are signals arising from parallel and perpendicular displacements, $\boldsymbol{q}_{\parallel}$ and $\boldsymbol{q}_{\perp}$ denote components of $\boldsymbol{q}$ that are parallel and perpendicular to the axons. We can write $\boldsymbol{q} = \boldsymbol{q}_{\perp} + \boldsymbol{q}_{\parallel}$.

Inside the axons, parallel molecular displacements are assumed to be Gaussian with diffusivity $D_{\parallel}$. This means that the parallel signal component is

$$E_{\parallel}(\boldsymbol{G}_{\parallel}, \Delta) = \exp\left(-\gamma^2 \boldsymbol{G}_{\parallel}^2 \delta^2 (\Delta - \delta/3) D_{\parallel}\right) \qquad (4.2)$$

where $\boldsymbol{G}_{\parallel}$ is the parallel component of the gradient $\boldsymbol{G}$. Other terms such as δ and Δ refer to standard pulsed sequence parameters.

Molecular displacements perpendicular to axons are assumed to be restricted. In their original paper [12], they assumed constant gradient pulses so that $\Delta \approx \delta$. This was partly because the mathematical formula is simpler and because the narrow pulse assumption was difficult to achieve. Therefore, signal decay was described by the Neuman formula [128],

$$E_{\perp}(\boldsymbol{G}_{\perp}, 2\tau) = \exp\left(-(a^4 \gamma^2 \boldsymbol{G}_{\perp}^2 / D_{\perp})(7/96)(2\tau - (99/112)(a^2/D_{\perp}))\right) \qquad (4.3)$$

where $D_{\perp}$ is the diffusivity perpendicular to the axons and a is the axon radius.

Diffusion in the extra-axonal compartment is Gaussian. Diffusion is anisotropic and described by a diffusion tensor $\mathbf{D}$,

$$E_{\mathrm{h}} = e^{-4\pi^2 (\Delta - \delta/3) \boldsymbol{q}^T \mathbf{D} \boldsymbol{q}} \qquad (4.4)$$

The model assumes that the principal axes of diffusion coincide with the parallel and perpendicular axes of the restricted compartment (i.e. the axons). The factor $\boldsymbol{q}^T \mathbf{D} \boldsymbol{q}$ can be expanded out and expressed in terms of perpendicular and parallel diffusivities $\lambda_{\perp}$ and $\lambda_{\parallel}$ in the

extra-axonal space and parallel and perpendicular components of $\boldsymbol{q}$,

$$E_{\text{h}}(\boldsymbol{q}, \Delta) = e^{-4\pi^2 |\boldsymbol{q}_\parallel|^2 \lambda_\parallel (\Delta - \delta/3)} \cdot e^{-4\pi^2 |\boldsymbol{q}_\parallel|^2 \lambda_\perp (\Delta - \delta/3)} \tag{4.5}$$

The CHARMED framework can be extended to include more than two compartments at the cost of increased complexity. In general, the total signal is

$$E(\boldsymbol{q}, \Delta) = \sum_{i=1}^{M} f_h^i E_h^i(\boldsymbol{q}, \Delta) + \sum_{i=1}^{N} f_r^j E_r^j(\boldsymbol{q}, \Delta) \tag{4.6}$$

where M is the number of hindered compartments, N is the number of restricted compartments, f_h^i is the volume fraction of the ith hindered compartment, and f_r^i is the volume fraction of the ith restricted compartment. Signals are measured with a range of different $\boldsymbol{q}$ magnitudes and orientations (concentric spherical shells). The $\boldsymbol{q}$ vectors in Eq. 4.6 can be written in spherical coordinates denoting fibre direction (θ_n, ϕ_n) and $\boldsymbol{q}$ direction (θ_q, ϕ_q). Then $\boldsymbol{q} = q(\theta_n, \phi_n, \theta_q, \phi_q)$, allowing for the extraction of fibre direction.

The number of free parameters depends on the number of compartments. For a single hindered compartment, there are 7 parameters. For one restricted and one hindered compartment, there are 12. For two hindered compartments (a dual tensor model), there are 13. For two restricted and one hindered compartment, there are 15. Multiple compartments were used to model crossing fibres. To reduce the number of free parameters, the original model fixed a and $D_\perp$. The parameters of interest were $\lambda_\parallel$ and $\lambda_\perp$ (hindered diffusivities), the orientation of the restricted compartments (θ_n^j, ϕ_n^j), restricted and hindered population fractions, $D_\parallel$, and a noise term.

Assaf tested the model using simulated data with a number of compartments to see how well parameters could be estimated [12]. Of particular interest was whether crossing fibres could be identified. They also calculated average propagators by Fourier transforming q-value data. It was found that with 90° crossings, two restricted compartments and a hindered compartment worked best ($M = 1$, $N = 2$), with 20 or more directions. With 30° crossings, at least 30 directions were needed to distinguish the two fibres. They also collected experiment data from excised pig spinal cord, including crossing fibres. The maximum gradient strength was

50 mT/m. Diffusion time was 150 ms, with pulse duration of 40 ms. Signals were measured in 31 directions. Results were close to actual values with both single and crossing fibres.

4.2.2 AxCaliber PGSE

AxCaliber extends the CHARMED model to account for a distribution of axon diameters [11]. In this protocol, gradients are applied perpendicular to the fibers. This simplifies some CHARMED expressions, since only displacements perpendicular to fibers are measured. This means that $\boldsymbol{q} = \boldsymbol{q}_\perp (= q)$ and $\boldsymbol{q}_\parallel = 0$. Parallel diffusivities $D_\parallel$ and $\lambda_\parallel$ disappear from the model. Only perpendicular terms remain. Since $\lambda_\perp$ is the only diffusivity that remains in the extra-axonal compartment, it is renamed the hindered diffusion coefficient D_h. Unlike CHARMED, AxCaliber assumes short gradients, so that the narrow pulse approximation can be used [11]. The signal from water inside an axon of radius R is modeled using Callaghan's expression [11, 32],

$$
\begin{aligned}
E_{\mathrm{r}}(q,\Delta) = &\sum_n 4\exp\left(-\nu_{n0}^2 D\Delta/a^2\right)\left[\frac{(2\pi qa)J_0'(2\pi qa)}{(2\pi qa)^2 - \nu_{n0}^2}\right] \\
&+ \sum_n\sum_k 8\exp\left(-\nu_{nk}^2 D\Delta/a^2\right) \times \frac{\nu_{nk}^2}{\nu_{nk}^2 - k^2} \times \left[\frac{(2\pi qa)J_k'(2\pi qa)}{(2\pi qa)^2 - \nu_{nk}^2}\right]
\end{aligned}
\tag{4.7}
$$

where J_k' are derivatives of kth order Bessel functions and ν_{nk} are solutions to $J_k'(\nu) = 0$.

The model uses one restricted and one hindered compartment, so that the total signal is

$$
E(q,\Delta) = f_{\mathrm{h}}E_h(q,\Delta) + (1 - f_{\mathrm{h}})\int f(a)E_r(a;q,\Delta)da
\tag{4.8}
$$

where $f(a)$ is a function that describes the distribution of diameters. In writing E_r, we included a to emphasize that it depends on axon radius. Since Eq. 4.7 is the restricted signal for a single axon, the full model (Eq. 4.8) includes a term for a range of diameters (described in Section 4.2.3). Based on earlier histological evidence, AxCaliber assumes that axon diameters follow a gamma distribution [1, 11].

The model was tested using data collected from porcine sciatic nerves, optic nerves, and porcine spinal cord. Diffusion times ranged from 10 ms to 150 ms (with gradient duration
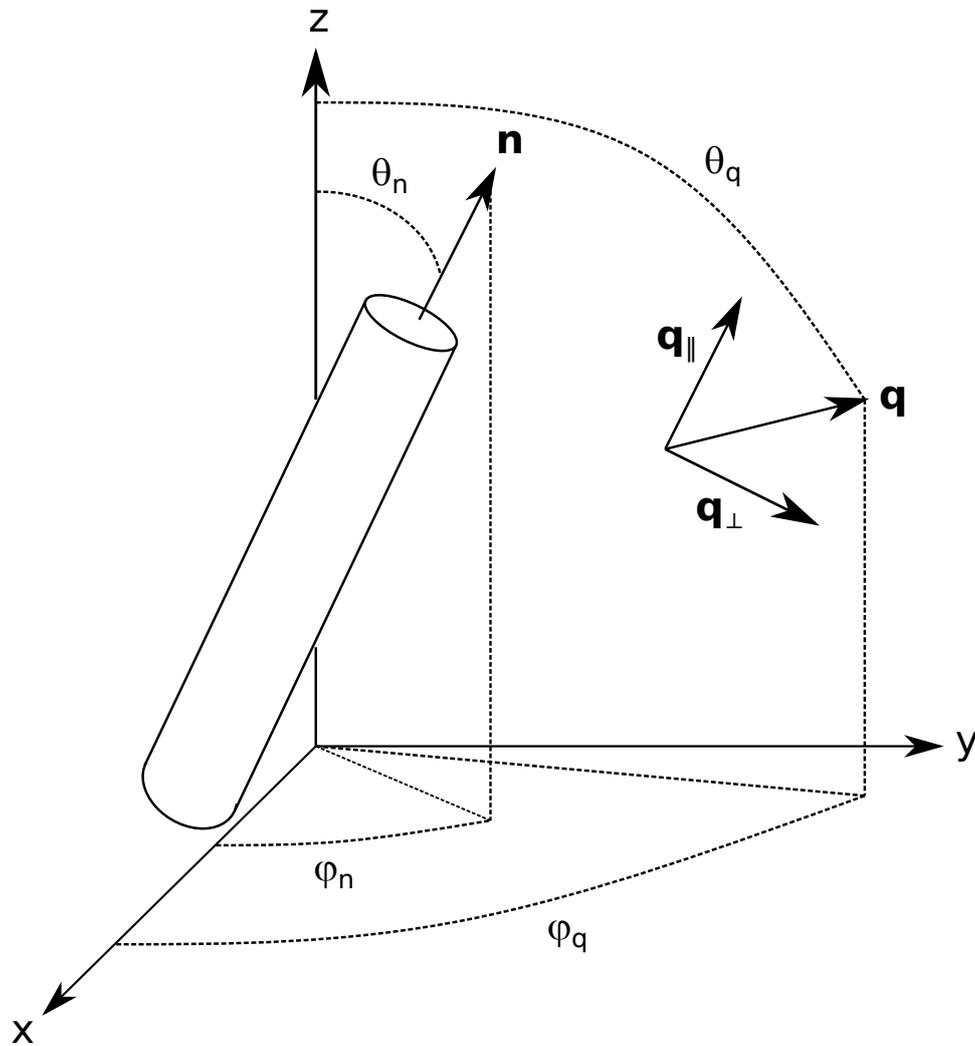
Figure 4.2: Diagram showing parameters used in the CHARMED model. A cylindrical axon points in the direction $\boldsymbol{n}$. Its orientation is described by azimuthal and polar angles ϕ_n and θ_n. A gradient $\boldsymbol{q}$ is applied. Its orientation is described by azimuthal and polar angles ϕ_q and θ_q. It can be decomposed into components parallel and perpendicular to the axon ($\boldsymbol{q}_{\parallel}$ and $\boldsymbol{q}_{\perp}$).

of 4 ms), with a maximum gradient strength of 300 mT/m. Extracted diameter distributions agreed well with histological measurements.

Barazany et al. [16] included an extra term to model a CSF compartment. The term is similar in form to that of the hindered compartment, but with diffusion coefficient D_{CSF} (≈ 3 µm^2/ms) instead of D_h. This addition expanded the number of free parameters to five (f_h, f_{CSF}, D_h, and two gamma distribution parameters). Their group measured axon diameter distributions along the length of the corpus callosum of a rat brain. Diffusion times ranged from 11 ms to 100 ms. The maximum gradient strength was 282 mT/m. The extracted distributions were similar in shape to those obtained from histological measurements.

4.2.3 OGSE and AxCaliber

For OGSE, the model in Eq. 4.8 is a function of frequency and gradient strength:

$$E(\omega_m, G) = f_{\text{axon}} E_r(\omega_m, G) + (1 - f_{\text{axon}}) E_h(\omega_m, G) \tag{4.9}$$

where f_{axon} is the relaxation-weighted volume fraction of axons, or intra-axonal space. Because we will be assuming that T_2 is the same in each compartment, f_{axon} will be referred to simply as the intra-axonal volume fraction from this point onwards. Eq. 4.9 has slightly different notation than Eq. 4.8 because we are using f_{axon} instead of f_h. Diffusion in the extra-axonal space is assumed to be hindered and is modeled with a Gaussian distribution

$$E_h = e^{-bD_h} \tag{4.10}$$

where D_h is the hindered diffusion coefficient of extra-axonal space. The b-value will depend on the pulse sequence. For a cosine-OGSE sequence of gradient strength G and gradient angular frequency ω_m, it is given by [43]

$$b_{\cos} = \frac{\gamma^2 G^2 T}{\omega_m^2} \tag{4.11}$$

94

For a sine-OGSE sequence, it is [43]

$$b_{\text{sin}} = \frac{3\gamma^2 G^2 T}{\omega_m^2} \tag{4.12}$$

The form of the restricted signal also depends on the pulse sequence used to collect data. The diffusion-weighted signal from a cosine-OGSE due to restricted water inside a pore is

$$E_r = \exp\left(-\beta(2\tau)\right) \tag{4.13}$$

with [192, 194]

$$\beta(2\tau) = 2\gamma^2 G^2 \sum_n \frac{B_n \lambda_n^2 D^2}{(\lambda_n^2 D^2 + \omega_m^2)^2} \left\{ \frac{\lambda_n^2 D^2 + \omega_m^2}{\lambda_n D} \left(\frac{T}{2} + \frac{\sin\left(2\omega_m T\right)}{4\omega_m} \right) - 1 + \exp\left(-\lambda_n D T\right) \right.$$
$$\left. + \exp\left(-\lambda_n D\tau\right)\left(1 - \cosh\left(\lambda_n D T\right)\right) \right\} \tag{4.14}$$

where D is the diffusion coefficient, 2τ is the echo-time, and λ_n and B_n are the structure dependent coefficients given below [192]. The coefficients λ_n and B_n depend on the radius a in the case of cylinders and spheres, and for planes, the separation distance L, so that Eq. 4.14 can be used to extract pore sizes. Therefore with OGSE, data can be collected with different gradient strengths and over a range of oscillation frequencies, to determine pore sizes. For comparison, the diffusion-weighted signal for a sine-OGSE gradient is [192]

$$\beta(2\tau) = 2\gamma^2 G^2 \sum_n \frac{B_n \omega_m^2}{(\lambda_n^2 D^2 + \omega_m^2)^2} \left\{ \frac{\lambda_n D T (\lambda_n^2 D^2 + \omega_m^2)}{2\omega_m^2} + 1 - \exp\left(-\lambda_n D T\right) \right.$$
$$\left. + \exp\left(-\lambda_n D\tau\right)\left(1 - \cosh\left(\lambda_n D T\right)\right) \right\} \tag{4.15}$$

We can give a brief overview of the derivation of Eqs. 4.14 and 4.15. The derivation starts with the Gaussian approximation (Section 3.3.3.1) [192],

$$\beta(2\tau) = \frac{1}{2}\gamma^2 \int_0^{2\tau} \int_0^{2\tau} \boldsymbol{G}(t_1) \cdot \langle \boldsymbol{r}(t_1)\boldsymbol{r}(t_2) \rangle \cdot \boldsymbol{G}(t_2) dt_1 dt_2 \tag{4.16}$$

95

If we replace the time averaged quantity $\langle \boldsymbol{r}(t_1)\boldsymbol{r}(t_2)\rangle$ with a spatial average $\int\int \rho(\boldsymbol{r}_1,t_1)P(\boldsymbol{r}_1,t_1|\boldsymbol{r}_2,t_2)\boldsymbol{r}_1\boldsymbol{r}_2 d\boldsymbol{r}_1 d\boldsymbol{r}_2$, we get

$$\beta(2\tau) = \frac{1}{2}\gamma^2 \int_0^{2\tau}\int_0^{2\tau} \boldsymbol{G}(t_1) \cdot \left[\int\int \rho(\boldsymbol{r}_1,t_1)P(\boldsymbol{r}_1,t_1|\boldsymbol{r}_2,t_2)\boldsymbol{r}_1\boldsymbol{r}_2 d\boldsymbol{r}_1 d\boldsymbol{r}_2\right] \cdot \boldsymbol{G}(t_2)dt_1 dt_2 \quad (4.17)$$

For restricted geometries, the propagator can be expanded in terms of its eigenfunctions (Eq. 3.14), so that Eq. 4.17 can be rewritten as [72, 192]

$$\beta(2\tau) = \frac{1}{2}\gamma^2 \sum_n B_n \int_0^{2\tau}\int_0^{2\tau} e^{-\lambda_n D|t_2 - t_1|} G(t_1)G(t_2)dt_1 dt_2 \quad (4.18)$$

where

$$B_n = \int\int (\hat{\boldsymbol{G}}\cdot\boldsymbol{r}_1)(\hat{\boldsymbol{G}}\cdot\boldsymbol{r}_2)u_n(\boldsymbol{r}_1)u_n(\boldsymbol{r}_2)d\boldsymbol{r}_1 d\boldsymbol{r}_2 \quad (4.19)$$

The exact form of B_n depends on the eigenfunctions $u_n(\boldsymbol{r})$, which in turn depend on the shape of the restricting geometry. For parallel planes separated by L [192],

$$B_n = \frac{8L^2}{(2n-1)^4\pi^4} \quad (4.20)$$

and λ_n is given in Eq. 3.21. For cylinders of radius a [192],

$$B_n = \frac{2(a/\nu_n)^2}{(\nu_n^2 - 1)} \quad (4.21)$$

where ν_n is the nth root of $J_1'(\nu) = 0$ and J_1 is a Bessel function of the first kind. The eigenvalues are

$$\lambda_n = \left(\frac{\nu_n}{a}\right)^2 \quad (4.22)$$

For spheres of radius a [192],

$$B_n = \frac{2(a/\nu_n)^2}{(\nu_n^2 - 2)} \quad (4.23)$$

where ν_n is the nth root of $\nu J_{3/2}'(\nu) - \frac{1}{2}J_{3/2}(\nu) = 0$. The eigenvalues are

$$\lambda_n = \left(\frac{\nu_n}{a}\right)^2 \quad (4.24)$$

The type of the gradient determines the form of the signal through Eq. 4.18. Solving the integral in Eq. 4.18 using cosine or sine and substituting the cylindrical versions of B_n and λ_n gives Eqs. 4.14 and 4.15.

The restricted signal from a large collection of axons is a volume-weighted sum of all cylinder signals. It is assumed that the axon diameters come from a distribution $w(a, \boldsymbol{p})$, parameterized by $\boldsymbol{p}$. Usually this is a gamma distribution [11,61,71]. The signal arising from the collection is

$$E_r = \frac{\sum_i w(a_i, \boldsymbol{p}) a_i^2 e^{-\beta(2\tau; a_i)}}{\sum_j w(a_j, \boldsymbol{p}) a_j^2} \tag{4.25}$$

where $e^{-\beta(2\tau; a_i)}$ is the signal from a cylinder of radius a_i [11].

To see where Eq. 4.25 comes from, we give a brief derivation. Suppose we have N axons. Including the extra-axonal space, there are $N+1$ compartments in total. If f_{axon}^i is the volume fraction of the ith axon, and S_i is the corresponding signal, then the total signal is the volume weighted sum of contributions from each compartment

$$S = f_{ex} S_{ex} + \sum_{i=1}^{N} f_{\text{axon}}^i S_i \tag{4.26}$$

where f_{ex} is the volume fraction of the extra-axonal compartment and S_{ex} is the extra-axonal signal. We can write $f_{ex} = 1 - f_{\text{axon}}$, where f_{axon} is the volume fraction of all axons taken together. Assume that there are n_i axons with radius a_i. If V is the volume (or surface area) of the substrate, then the fraction of volume taken up by all axons with radius a_i is

$$f_{\text{axon}}^i = \frac{n_i \pi a_i^2}{V} \tag{4.27}$$

The signal then becomes

$$S = (1 - f_{\text{axon}}) S_{ex} + \frac{n_1 \pi a_1^2}{V} S_1 + \frac{n_2 \pi a_2^2}{V} S_2 + \dots \tag{4.28}$$

After factoring out a term $\left(\frac{\sum n_k \pi a_k^2}{V}\right)$, it becomes

$$S = (1 - f_{\text{axon}})S_{ex} + \left(\frac{\sum_{i=1}^{N} n_k \pi a_k^2}{V}\right)\left(\frac{n_1 \pi a_1^2}{\sum_{i=1}^{N} n_k \pi a_k^2}S_1 + \frac{n_2 \pi a_2^2}{\sum_{i=1}^{N} n_k \pi a_k^2}S_2 + \ldots\right) \quad (4.29)$$

Since $\left(\frac{\sum n_k \pi a_k^2}{V}\right)$ is the fraction of space taken up by all axons, or f_{axon}, we can rewrite Eq. 4.29 as

$$S = (1 - f_{\text{axon}})S_{ex} + f_{\text{axon}}\frac{\sum_{i=1}^{N} n_i a_i^2 S_i}{\sum_i n_k a_k^2} \quad (4.30)$$

If the axon diameters come from a known distribution $w(a, \boldsymbol{p})$, the number of axons with diameters in the range $[a_i, a_i + \Delta a]$ is $w(a_i, \boldsymbol{p})\Delta a$. Substituting this into Eq. 4.30 and canceling N and Δa gives Eq. 4.25.

4.2.4 Monte Carlo Simulations

Monte Carlo simulations have been used to synthesize diffusion-weighted MR signals. We use a Monte Carlo simulation based on that developed by Szafer [170] to simulate our signals. This involves letting a set of spins diffuse over a lattice. While Szafer used a PGSE sequence, we use OGSE sequences [154]. Our code was adapted for use on a GPU in order to take advantage of increased speed and performance [64].

As described in Section 3.5, our simulations used impermeable barriers similar to other studies [192, 195]. Simulations described in Section 4.2.5.1) used 57344 particles, while simulations with the more complex geometries described in Section 4.2.5.2) used 114688 particles. For the simulations used in this thesis, these particle numbers give a relative uncertainty (standard deviation of simulated signals/mean of simulated signals) of less than 0.002. The time step used in all simulations was 1 µs. The maximum particle step distance with this time step will be approximately 0.12 µm, which should be smaller than most of the restriction sizes used in the simulations. Our simulation did not consider different T_2 in different tissues.

All simulations used two 20-ms ideal cosine gradient pulses, one before and after the 180° pulse, each applied perpendicular to the axons. We generated 400 cosine OGSE signals for each geometry. For initial simulations, we used very high frequencies and very high gradients

to obtain a sense of what would be needed to make the measurements and to confirm the restricted signal formulas in Eqs. 4.14 and Eq. 4.25. Cosine frequencies ranged from 0.05 kHz to 9.55 kHz. At each frequency there were 20 b-values (up to 2.09 ms/μm^2). This gives around 90% signal decay at the highest b-value. The echo time (2τ) was 42 ms.

For reasons to be discussed later, and to use experimentally feasible pulse sequence parameters, we repeated the previous simulations using a narrower and more feasible range of gradient frequencies and gradient amplitudes. There were 20 frequencies ranging from 50 to 1000 Hz in steps of 50 Hz. We will refer to the frequencies used in these simulations as the 'restricted range' of frequencies, in contrast to the 'full range' of frequencies described previously. Five gradient strengths were used for each gradient pulse, for a total of 100 signals. These pulse sequence parameters were chosen to be feasible for experiments using a BGA6 gradient set (Bruker Biospin) on our 7 T magnet (maximum gradient strength of 1.01 T/m) [63]. For $f_m > 200$ Hz, the gradient strengths were 0, 60, 70, 80 and 90% of maximum. The gradient strengths for 200 Hz or less were smaller. For $f_m = 50$ Hz, $G = 0, 1.5, 4, 7, 10$% of maximum; for $f_m = 100$ Hz, $G = 0, 4, 7, 10,$ and 15% of maximum; for $f_m = 150$ Hz, $G = 0, 15, 20, 25,$ and 30% of maximum; for $f_m = 200$ Hz, $G = 0, 30, 40, 50,$ and 60% of maximum. These are summarized in Table 4.1. We have achieved these values on our animal scanner [63]. The echo time was 44.52 ms.

The simulations were programmed in CUDA C/C++ and run on a HP Z240 workstation containing a Intel® Xeon® Processor E5-1650 6-core 3.20 GHz CPU. The HP Z240 workstation contained two graphics cards, an NVIDIA Tesla C2075 (Fermi 2.0) graphics card for dedicated CUDA computation and an NVIDIA Quadro 600 (Fermi 2.1) graphics card handling

OGSE frequency	Gradient strength (as percentage of 1.01 T/m)
50 Hz	0, 1.5, 4, 7, 10
100 Hz	0, 4, 7, 10, 15
150 Hz	0, 15, 20, 25, 30
200 Hz	0, 30, 40, 50, 60
250 Hz or higher	0, 60, 70, 80, 90

Table 4.1: Gradient strengths used in simulations with our MRI machine (the 'restricted' range). The maximum gradient strength is approximately 1.01 T/m. Values are given as percentages. The highest b-value used in the measurements is 0.47 ms/μm^2.

the display.

4.2.5 Cylinders

We used AxCaliber, modified to use OGSE, to fit diameter distributions for a variety of simulated distributions. White matter fibres were modeled as a collection of parallel non-overlapping cylinders. The AxCaliber models assumed no water exchange between intra- and extra-axonal regions [11]. This assumption has been made in later simulations [192, 199], but not in others [131, 161]. In the present study we assumed no water exchange for consistency with the AxCaliber model. Some studies have assumed different diffusivities in the intra- and extra-axonal regions [154, 170, 192], while others have assumed equal diffusivities [46, 61, 195]. We followed Ref. [192] in setting the intra-axonal diffusion coefficient (D_{in}) to 1.0 μm^2/ms and using a larger extra-axonal diffusion coefficient (D_{ex}), chosen here to be 2.5 μm^2/ms [154].

The full AxCaliber model takes both intra-axonal and extra-axonal signals into account. For studies performed here, each component (intra-axonal and extra-axonal) was simulated separately and then together to verify the validity of the model. For the "intra-axonal" simulation, water was located only inside the cylinders. In this case, the signal can be described by Eq. 4.25 alone. For the "extra-axonal" simulation, water was located only outside the cylinders and the signal should be described by Eq. 4.10 alone. We can test this assumption by calculating an *ADC* at each frequency using Eq. 3.84 and examining the frequency dependence of the resulting spectrum. If the *ADC* is independent of frequency, then Eq. 4.10 is a good approximation. For the "full" simulation, we distributed the water everywhere, both inside and outside the cylinders. For that simulation, the signal should be described by the full model given in Eq. 4.9.

4.2.5.1 Identical cylinders

The simplest geometry we consider is a collection of identical cylinders of diameter d arranged on a square lattice. This could be considered a special case of Eq. 4.25, with $w(a, \boldsymbol{p})$ equal to a δ-function centered at $a = d/2$.

For the intra-axonal simulation, we tested eight diameters {1, 2, 3, 4, 5, 6, 7, 8} μm as the

intra-axonal simulation. After each simulation, the cylinder diameter and intra-axonal diffusion coefficient were extracted by fitting the simulated signals to Eq. 4.13 and Eq. 4.14.

For the full simulation, we used packing fractions of 0.5, 0.6, and 0.7 with the same eight diameters. The signals were then fit to Eq. 4.9, with Eq. 4.10 and Eq. 4.13. Four parameters were extracted from each fit: cylinder diameter, intra-axonal diffusion coefficient, packing fraction, and the hindered diffusion coefficient.

4.2.5.2 Diameter distributions

The next two geometries consisted of a set of 100 cylinders randomly distributed over a square domain with periodic boundary conditions. The substrate was generated according to the algorithm in Ref. [61], briefly described by the following steps.

1. Choose 100 diameters from a diameter distribution and sort in descending order.

2. Starting with the largest cylinder:

   (a) Choose a random position on the substrate.

   (b) If the cylinder overlaps the edge(s) of the substrate, create copies of the cylinder and place them on the opposite side(s) so that lattice periodicity is maintained.

   (c) If either the cylinder or any of its copies overlap an existing cylinder, discard and choose a new position (go back to 2(a)).

   (d) Place the cylinder and its copies on the substrate.

3. Repeat the above steps until all 100 cylinders have been placed.

First, we allowed for a spread of diameters by sampling them from a Gaussian distribution. This means that $w(a, \boldsymbol{p})$ in Eq. 4.25 is a Gaussian with mean μ_N and standard deviation σ_N. There were four distributions tested, with mean diameters of 2, 3, 4, and 5 µm, and standard deviations of 0.6, 1, 1.34, and 1.6 µm, respectively. Packing fractions were between 0.6 and 0.75.

To model a more realistic situation, diameters were also sampled from a gamma distribution [61]. The gamma distribution has two parameters, α_Γ and β_Γ, and is given by

$$w(x, \alpha_\Gamma, \beta_\Gamma) = \frac{x^{-\alpha_\Gamma - 1} e^{x/\beta_\Gamma}}{\beta_\Gamma{}^{\alpha_\Gamma} \Gamma(\alpha_\Gamma)} \tag{4.31}$$

The mean of the distribution is $\alpha_\Gamma \beta_\Gamma$ and it has a maximum at $(\alpha_\Gamma - 1)\beta_\Gamma$, provided that $\alpha_\Gamma > 1$. We chose specific values for α_Γ and β_Γ with $\alpha_\Gamma > 1$ such that the maximum and mean were close to the same value (within 1 μm), for easier interpretation. We tested three distributions with mean diameters of 2, 2.1, and 4.2 μm. Packing fractions were greater than 0.6 for all distributions. Figure 4.3 shows example substrates for the two types of distributions. Because the placement of the cylinders could affect the results, we also generated three substrates for each diameter distribution.

The intra-axonal diffusion coefficient was held fixed to its actual value for all fits [11]. For the intra-axonal simulations, the two distribution parameters were extracted by fitting the signals to Eq. 4.25. The two distribution parameters were $\boldsymbol{p} = \{\alpha_\Gamma, \beta_\Gamma\}$ for the gamma distributions and $\boldsymbol{p} = \{\mu_N, \sigma_N\}$ for the Gaussian distributions. For the full simulation, signals were fit to Eq. 4.9, with Eq. 4.10 and Eq. 4.25. Four parameters were extracted from each fit: the two distribution parameters, the packing fraction, and the hindered diffusion coefficient.

4.2.6 Effects of Noise

We also look at the effects that noisy signals may have on fitted parameters in the model. Gaussian noise was added to each component of the transverse magnetization at the end of the simulations so that SNR values were 50, 100, and 200. This was repeated 1000 times and the fitted parameters were binned into histograms for analysis.

4.2.7 Data Fitting

Signal data were fitted to the model with nonlinear least-squares regression using the Optimization Toolbox in MATLAB [113]. All fitted parameters were constrained to be positive. Packing fractions were constrained to lie between 0 and 1. The fitting procedure was repeated 50 times
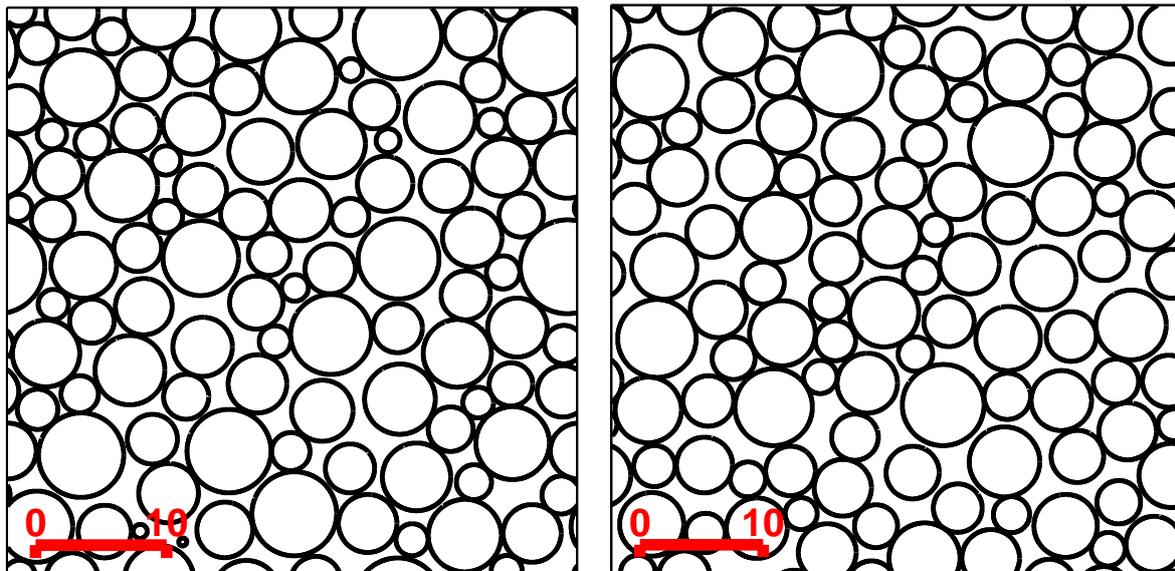
Figure 4.3: Two example substrates used in the simulations. The scale bar at bottom left of each substrate denotes 10 μm. Left: Cylinder diameters sampled from a Gaussian distribution with mean diameter 4 μm and standard deviation 1.33 μm. The lattice is 43.5 μm × 43.5 μm. Right: Cylinder diameters sampled from a gamma distribution ($\alpha_\Gamma = 21$, $\beta_\Gamma = 0.1$ μm) with mean diameter 4.2 μm and standard deviation 0.45 μm. The lattice is 46.5 μm × 46.5 μm.

with random initial parameters. At the end, the fitted parameters providing the smallest sum-of-squares were chosen as the best fit.

4.3 Results

The results of the intra-axonal, extra-axonal and full simulations are presented in this chapter. More reasons for the choice to use a restricted range of frequencies are provided. Because signal noise will not be added to the simulation data until Section 4.3.5, results show how the model performs in an almost ideal situation. Simulation uncertainty (due to finite number of particles) is more than an order of magnitude smaller than the noise we will add in Section 4.3.5, so we will postpone a more thorough error analysis until then (and continue in Chapter 5).

4.3.1 Intra-axonal Simulations

Fitted values for the diameter are shown in Figure 4.4 for intra-axonal simulations using identical cylinders. The fitted values for each cylinder are in good agreement with the actual values (to within 3 percent).

Simulated signals are comparable to analytical signals predicted with Eq. 4.25 for a distribution of diameters. An example of this comparison is shown in Figure 4.5 for one of the substrates. The simulated values agree very well with the predicted analytical values, differing at most by 5 percent. Similar results were obtained with simulated data from the other substrates.

Having established that Eq. 4.25 describes the simulated data, we can now use it to fit the simulated data and obtain fitted distribution parameters. Using distribution parameters obtained from fitting the signals to the intra-axonal term, a selection of fitted diameter distributions are plotted in Figure 4.6. For both types of distributions tested, the fitted distributions agree very well with the actual distributions. As an example, for the gamma distribution shown in Figure 4.6a, the predicted mean diameter differs from the true mean by only 1.2 percent.

4.3.2 Extra-axonal Simulations

We now consider the extra-axonal model (Eq. 4.10) in more detail. In the last section, we compared the simulated signals with the signal predicted by the model. Here, we perform this comparison in a slightly different way. Instead of comparing the signals, we will look at the *ADC* spectra. Eq. 4.10 assumes a flat *ADC* spectrum, independent of frequency in the extra-axonal space, because the extra-axonal water is assumed to be undergoing hindered diffusion. Using our simulation data, we calculated *ADC* spectra in order to see how well that assumption holds. *ADC*s were calculated from simulation data by performing linear regression on $-\ln(E_h)$ and $b_{\cos}$ (see Eq. 3.84).

Extra-axonal *ADC* spectra obtained from simulations on the square packed lattice are shown in Figure 4.7a. Each spectrum was generated from simulations with a single cylinder diameter. All spectra in Figs. 4.7a and 4.7b come from geometries with the same packing fraction, shown here is 0.6. Simulations using other packing fractions show similar trends and the discussion
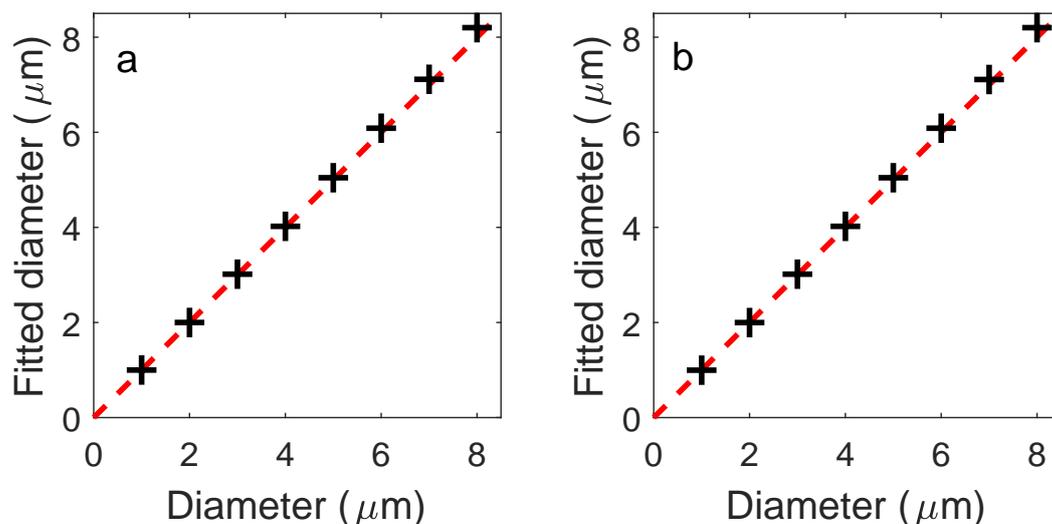
Figure 4.4: Comparison of fitted cylinder diameters with their actual values (intra-axonal model only). Fitted values are shown as '+' symbols. Dashed line indicates 'perfect' fit. (a) Fitted values obtained using the full frequency range. (b) Fitted values obtained using the restricted frequency range.
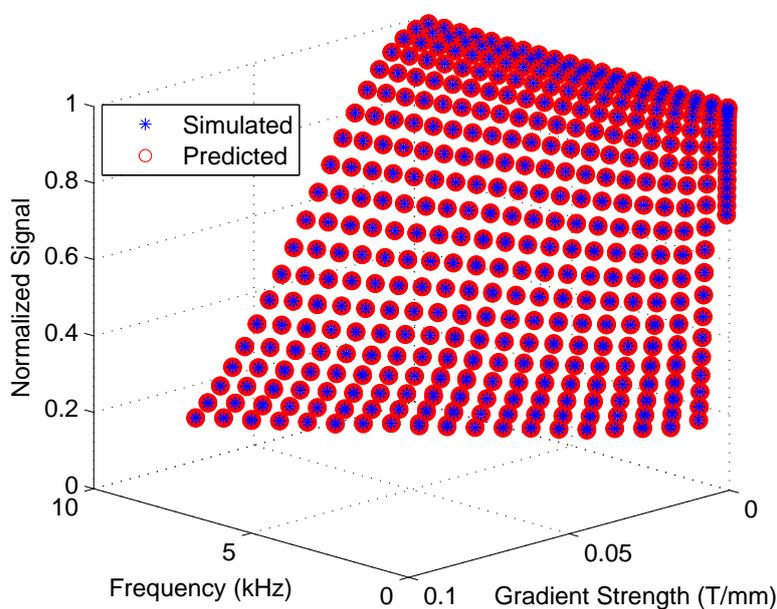
Figure 4.5: Comparison between simulated signals (blue) and predicted signals (red) for the intra-axonal model in Eq. 4.25 (using a gamma distribution of diameters). The signals are plotted against frequency and gradient strength used in each simulation. The simulated and predicted signals agree within 5 percent.
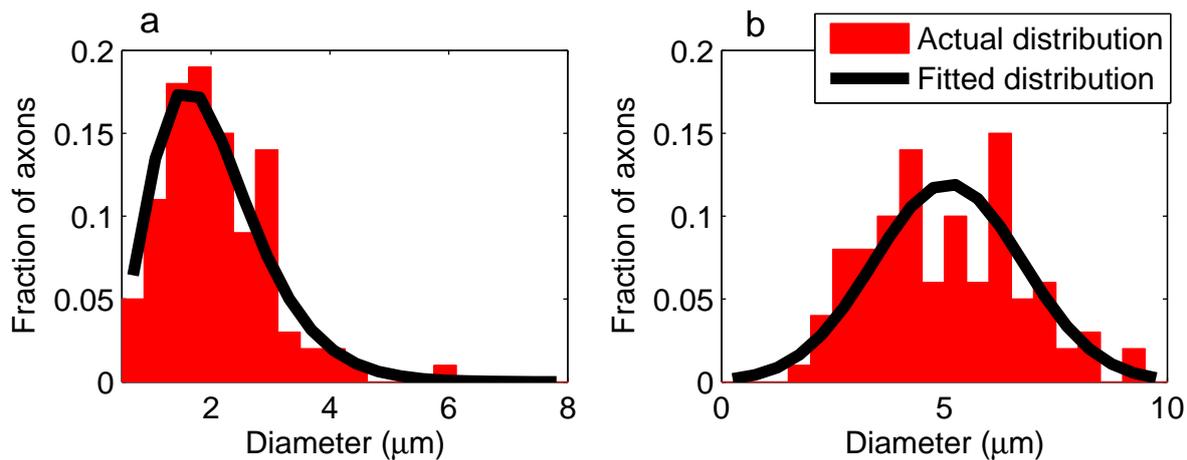
Figure 4.6: Two fitted diameter distributions for the intra-axonal model using the full range of frequencies. Red bars indicate the actual distributions, while the black curves indicate the fitted distributions. (a) Gamma distribution of diameters. (b) Gaussian distribution of diameters. Two fit parameters (out of two in total) were used to create the black curves indicating the fitted distributions.

that follows also applies to them.

For the same packing fraction, at a given frequency, the *ADC*s are smaller in lattices with smaller cylinders, as expected. In these smaller lattices, the distance between adjacent cylinders is smaller, and so these extra-axonal particles experience the effects of the boundaries at shorter diffusion times, leading to smaller *ADC*s. The *ADC* spectra are constant, as needed in the model, over the entire frequency range only when using the smallest diameter cylinder. When the lattice of the same packing fraction contains larger diameter cylinders, not all water molecules will experience the effects of the boundaries at all diffusion times and so the diffusion is no longer considered hindered for all frequencies. The *ADC* spectra rise dramatically as frequencies increase, indicating this breakdown of the model. Even when the diameter is only 2 μm, the spectrum rises from 1.4 to 2 μm^2/ms. Therefore, we would expect the full model to be adequate only for the most restricted geometries (for example, diameters of 1 μm).

One can attempt to overcome this limitation by using a restricted range of frequencies, up to only 1000 Hz as described previously. For this restricted range of frequencies, not only are the extra-axonal *ADC* spectra constant for the smallest diameter (1 μm), they are nearly constant for 2 μm as well (see Figure 4.7b). Comparing Figure 4.7b to 4.7a, all the spectra show less of an increase over the range of frequencies. None of the spectra shown in Figure 4.7b rise above 2 μm^2/ms. With this in mind, the full model should perform better over a wider range of
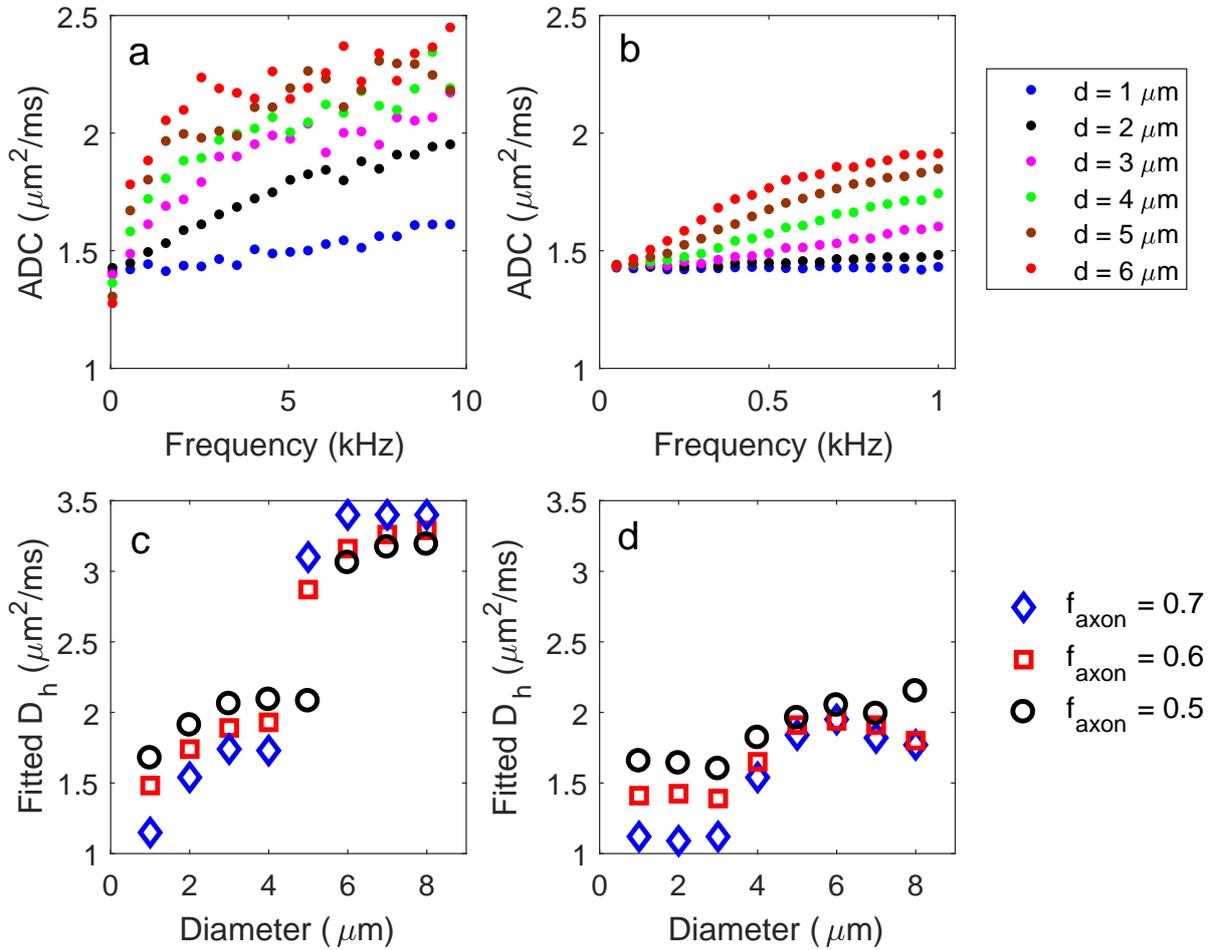
Figure 4.7: (a-b) Extra-axonal frequency dependent *ADC* spectra for lattices containing identical square packed cylinders with packing fractions of 0.6. Each spectrum was calculated using data generated from extra-axonal simulations using one cylinder diameter. To reduce clutter, only results for diameters up to 6 μm are shown. (a) Spectra generated from simulations utilizing the full frequency range. (b) Spectra generated from simulations utilizing the restricted frequency range. (c-d) Extracted hindered diffusion coefficients (D_h) as a function of cylinder diameter. Results for each of the three packing fractions (f_{axon} = 0.5, 0.6, and 0.7) are shown (as circles, squares, and diamonds, respectively). (c) Fitted values obtained using the full frequency range. (d) Fitted values obtained using the restricted frequency range.

diameters with this new range of frequencies. From the figure, this would include simulations with 1 and 2 μm (and possibly 3 μm) diameter cylinders.

In the long-time (low frequency) limit, the measured *ADC* is reduced from its free value by a factor λ_τ^2, called the tortuosity constant, [52, 88, 156, 157, 201]

$$ADC(\Delta \to \infty) = \frac{D_0}{\lambda_\tau^2} \tag{4.32}$$

There are tortuosity models, usually relating the tortuosity to other properties such as volume fraction. For example, in samples of packed spherical beads,

$$\lambda_\tau^2 = 1/\sqrt{\phi_{bead}} \tag{4.33}$$

where ϕ_{bead} is the volume fraction of fluid (the quantity ϕ_{bead} is also called the porosity) [156, 157]. That is, tortuosity is related to the fraction of space taken up by the pores. Tortuosity is sometimes used to relate long time *ADC*s to diffusion coefficients in other directions, based on previous experimental data. For example, if D_0 is the free diffusion coefficient and *ADC* is the long time apparent diffusion coefficient measured perpendicular to the axons, then [7]

$$ADC = D_0(1 - f_{\text{axon}}) \tag{4.34}$$

Eq. 4.34 has been used in certain models to eliminate parameters.

Figures 4.7a and 4.7b show that as the frequency goes to zero, each spectrum converges to around 1.4 μm^2/ms. Although we do not show them here, the *ADC* spectra obtained with the other packing fractions show the same behaviour. For a packing fraction of 0.5, the long-time *ADC* is around 1.6 μm^2/ms, while for a packing fraction of 0.7, it is around 1.1 μm^2/ms. Using Figure 4.7b, $\lambda_\tau \approx \sqrt{2.5/1.4} = 1.34$ for packing fraction 0.6. Because the long-time *ADC*s depend only on packing fraction (and not the size of the cylinders themselves), the tortuosity depends on the packing fraction, in line with previous studies [170].

4.3.2.1 Different substrates

Here we look into the possibility that cylinder placement on the substrate affects results. Figure 4.8 compares extra-axonal *ADC* curves from three different substrates. Each substrate uses cylinders drawn from the same diameter distribution. The only difference is the arrangement of the cylinders. The curves have similar shapes, rising from $ADC \approx 0.9$ μm^2/ms at 50 Hz to $ADC \approx 1.15 - 1.2$ μm^2/ms at 1000 Hz. At a given frequency, *ADC*s are within 5 percent of each other.

4.3.2.2 Apparent kurtosis

The extra-axonal *ADC* spectra in Figure 4.7a are noisy. The *ADC* spectra in Figure 4.7b are smoother. To see the reason why this occurs, we need to explain the concept of kurtosis.

Equation 4.10 is only valid for relatively small diffusion weightings, but the weightings used in the full frequency range lead to a 90% decrease in signal. This means that Eq. 4.10 needs to be modified. It can be shown that higher order terms in b need to be included, so that

$$\ln\left(E_h\right) = -bADC + \frac{1}{6}b^2 K_{app}ADC^2 \tag{4.35}$$

where K_{app} is called the apparent kurtosis [52, 75, 120, 144]. If diffusion is unrestricted, then $K_{app} = 0$. In the extra-axonal region, diffusion is hindered, so we expect that $K_{app} \neq 0$. If b is small, then only the first term in Eq. 4.35 contributes and we get Eq. 4.10.

We can use nonlinear regression on extra-axonal simulation data to extract both *ADC* and K_{app}. At each frequency, we fit the normalized signals $E(b)$ and b-values $b_{\cos}$ to Eq. 4.35 and extract *ADC* and K_{app}. Figure 4.9 shows the *ADC*s found using Eq. 4.35 as a function of frequency. The curves are smoother than before, as we see by comparing to Figure 4.9b to Figure 4.9a.
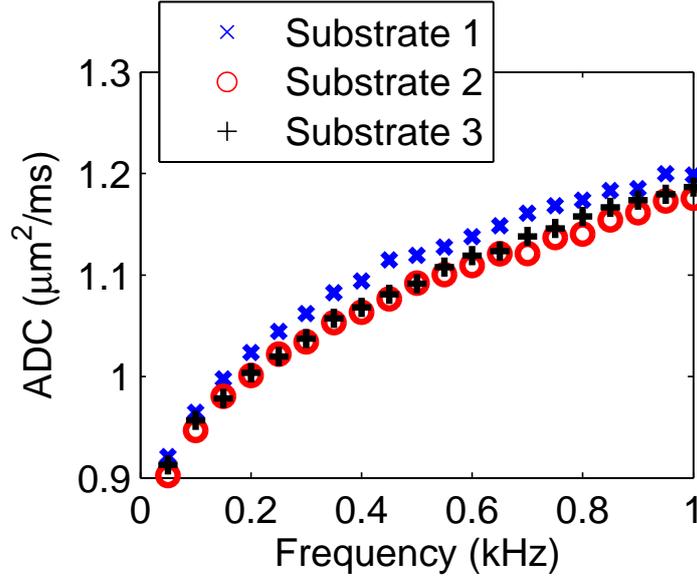
Figure 4.8: Extra-axonal *ADC* spectra for lattices containing gamma distributed diameters (packing fraction = 0.75). Each spectrum was calculated using data generated from extra-axonal simulations using a different cylinder placement. However, all used the same set of cylinders, so that the only difference was their locations on the substrate.

4.3.3 Full Simulation (Intra- and Extra-axonal water)

4.3.3.1 Full Frequency Range

In Section 4.3.2, we saw that the full model (Eqs. 4.9, 4.10, and 4.25) describes the actual extra-axonal data only when using the smallest cylinders in the simulations. When we collect simulated data of extra-axonal water over the full frequency range, we can only model systems with cylinders of 1 μm in diameter correctly. With this in mind, we present the results of fitting full simulation data to the full model in this section to show the breakdown of the model when both intra- and extra-axonal water is present. We start with the simplest situation: the square packed lattice of identical cylinders.

Fits of hindered diffusion coefficients for different packing fractions are shown in Figure 4.7c. When cylinder diameters reach 5 μm, fitted D_h values jump to 3 μm^2/ms. Because the extra-axonal diffusion coefficient is only 2.5 μm^2/ms, this suggests a breakdown in the model for larger diameters. Qualitatively, below 5 μm, D_h is larger at lower packing fractions, as one would expect. At higher packing fractions, the extra-axonal water comes into contact with barriers more often, causing the measured diffusion coefficient to decrease.
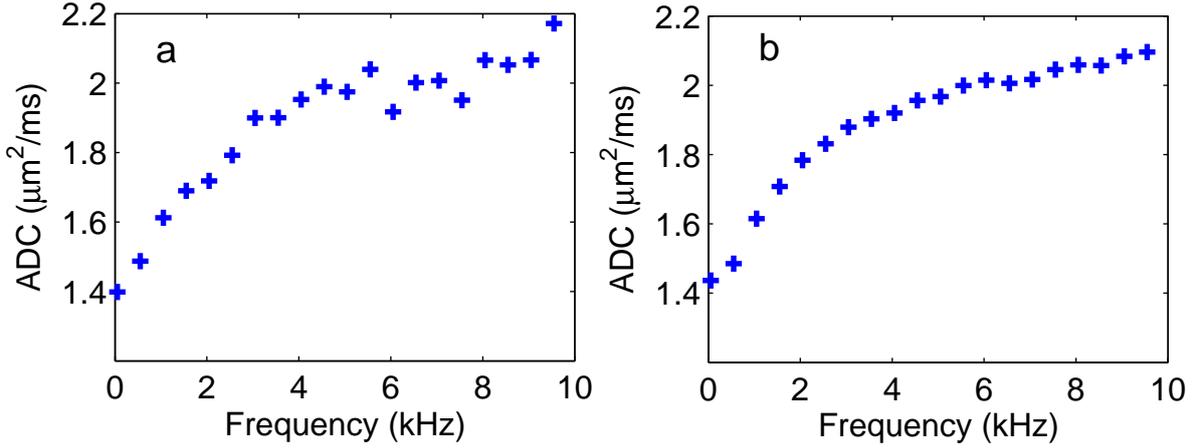
Figure 4.9: Extra-axonal frequency dependent *ADC* spectra for lattices containing identical square packed cylinders (diameter = 3 μm, packing fraction = 0.6). Simulations use the full frequency range, where maximum diffusion weightings are $\approx 90\%$. The spectrum in (a) was generated assuming $\ln(E_h)$ had no kurtosis term ($K_{app} = 0$ in Eq. 4.35). The spectrum in (b) was generated assuming $\ln(E_h)$ included a kurtosis term ($K_{app} \neq 0$ in Eq. 4.35). The spectrum in (b) is much smoother, showing that the kurtosis term should be included at very high weightings.

Figure 4.10a shows how fitted diameters compare to the actual diameters for a packing fraction of 0.6. Simulations with the smallest diameter cylinders produce the most accurate fits, while those with diameters above ≈ 2 μm produce poorer fits. The fitted diameters for the 3 μm and 4 μm diameter cylinders are smaller than the actual diameters. At 5 μm, a discontinuity occurs, which suggests a possible problem in the model. As mentioned in Section 4.2.7, we performed 50 fits with different randomly chosen initial parameters. Usually, the result with the smallest sum-of-squares was the "correct" solution. When the model gave more accurate results, 1 to 2 local minima in parameter space of the fit were found depending on the initial parameters. In other words, there were 1 or 2 unique solutions from the 50 different fits and the solution with the smaller sum-of-squares was the "correct" solution. More local minima (> 3) appeared for models that did break down. For each fit we chose the minima/minimum with the lowest sum-of-squares and physically plausible parameters but could have used other constraints to pick the minima and hence the fitted parameters. Thus the failing of the fitting method was more apparent in the increased minima found rather than the fitted values of the parameters.

Fitted volume fractions for different packing fractions are shown in Figure 4.10c. The fitted

packing fractions follow the trend of their true values. In other words, simulations with smaller packing fractions have smaller fitted packing fractions. Simulations with smaller diameter cylinders also provide more accurate fitted values. For example, lattices containing diameters larger than 5 μm produce fitted packing fractions up to 20 percent larger than the real values, while in simulations with diameters less than 5 μm, all fitted values are within 10 percent.

4.3.3.2   Restricted Frequency Range

We now present the simulation results for the previous geometries using the restricted range of frequencies with the goal of improving the model to avoid breakdown.

Figure 4.7d shows fits of hindered diffusion coefficients obtained from the restricted frequency range. Here the discontinuity at 5 μm no longer occurs (though there is some scatter starting at 4 μm). The values remain below 2.5 μm^2/ms, as they should. In fact, D_h is nearly constant for diameters up to 3 μm. Because the model begins to break down when larger diameter cylinders are present, the steady rise in fitted D_h values above 3 μm could be expected.

Figures 4.10b and 4.10d show fitted diameters plotted against the actual diameters at a packing fraction of 0.6 and a comparison of fitted packing fractions obtained with the restricted frequency range. The fitted packing fractions are all within 12 percent of the true packing fractions for the range of studied diameters. Also, diameters less than 4 μm can now be fit accurately. Once again, larger diameters pose a problem. When the diameters are larger than 4 μm, fitted diameters start to underestimate the true diameters. In fact, it appears that the fitted diameters reach a constant value of about 4.6 μm. This clearly indicates once again a breakdown in the model in that the extra-axonal water is not experiencing hindered diffusion over the range of frequencies used. Once again the breakdown in the model causes more minima to appear in parameter space. More emphasis should be made on the breakdown of the model and appearance of the extra minima rather than the plateau value of the fitted diameter because if other minima were used the fitted diameter would change.
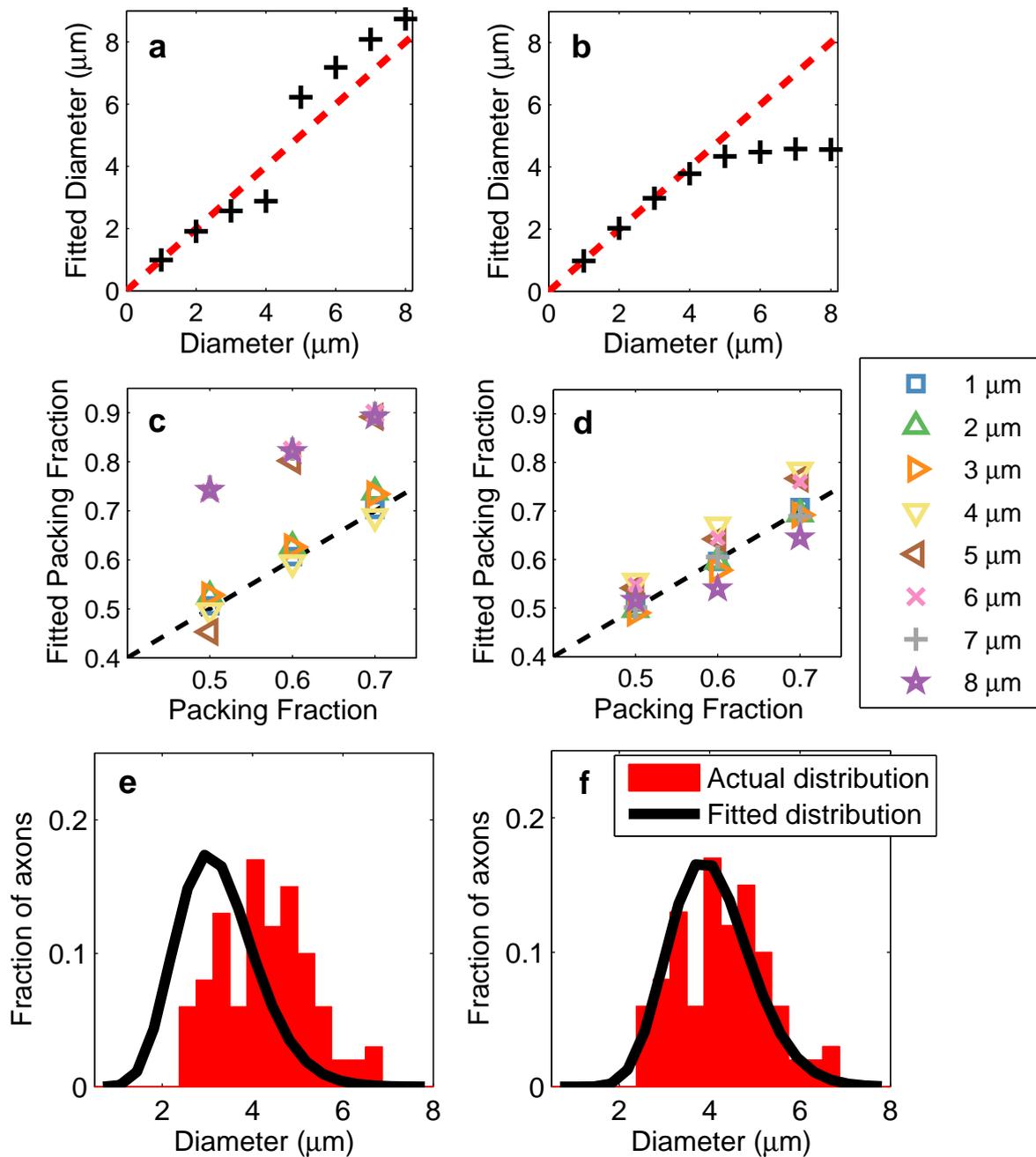
Figure 4.10: (a-b) Comparison of fitted cylinder diameters with their actual values using the total signal model. Actual fitted values shown as '+' symbols. Dotted line indicates ideal one-to-one correspondence. (a) Fitted values obtained using the full frequency range. (b) Fitted values obtained using the restricted frequency range. (c-d) Comparison of fitted packing fractions with their actual values. Results are shown for various cylinder diameters (see legend). Dotted line indicates 'perfect' fit. (c) Fitted values obtained using the full frequency range. (d) Fitted values obtained using the restricted frequency range. (e-f) Comparison of using a full frequencies range and the restricted frequency range for fitting diameter distributions with the total signal model. Distribution is a gamma distribution with mean diameter 4.2 μm. Red bars indicate the actual distributions, while the black curves indicate the fitted distributions. (e) Full range of frequencies. (f) Restricted range of frequencies. The restricted range of frequencies corrects for the underestimation of the mean axon diameter for the distribution for the larger diameters.

4.3.3.3 Diameter Distributions

Finally, we are ready to present the results of using our modified AxCaliber model on simulated axon diameter distributions. Because intermediate diameters can now be fitted accurately in the simple lattices for gradient frequencies less than 1000 Hz, one might expect that diameter distributions containing these diameters could also be accurately fit. This is the case, as seen in Figure 4.11. The top row shows a set of three Gaussian distributions, while the bottom row shows three gamma distributions. Previous fitted distributions which were shifted (as seen in Figure 4.10e) now fit the data much better. For five of the six distributions shown (Figure 4.11a, 4.11b, 4.11d-f), the fitted mean diameters are within 10 percent of the true means. The fitted mean of the distribution in Figure 4.11c is 15 percent smaller than the true mean. This distribution has a mean diameter of 5 μm, the largest we studied. In this regime, the extra-axonal space is probably too large for the current model to work and a different range of frequencies is needed. Thus careful selection of the frequency range is needed based on the expected range of diameters in the sample.

4.3.4 Simple Extra-axonal Model

We saw earlier that the extra-axonal diffusion spectrum was frequency dependent. At higher diameters, it was not described by Eq. 4.10. Because the actual form of the spectrum is complicated, some have sought to make approximations. Some have assumed that $D(\omega)$ is linear at short frequencies [195]. The extra-axonal signal is

$$E_h = e^{-b(A\omega + D_\infty)} \tag{4.36}$$

where D_∞ is the long-time hindered diffusion coefficient and A describes the rate of change of $D(\omega)$ with frequency. For this relationship to hold, we obviously cannot use all of our frequencies. We can see from the extra-axonal ADC spectra in Figure 4.7 that it might be better to use OGSE frequencies up to only 500 Hz. Above that, the ADC spectra begin to level off and the linear model might not be the best choice.

We tested this model on some of our simulation data (square-packed identical cylinders, for
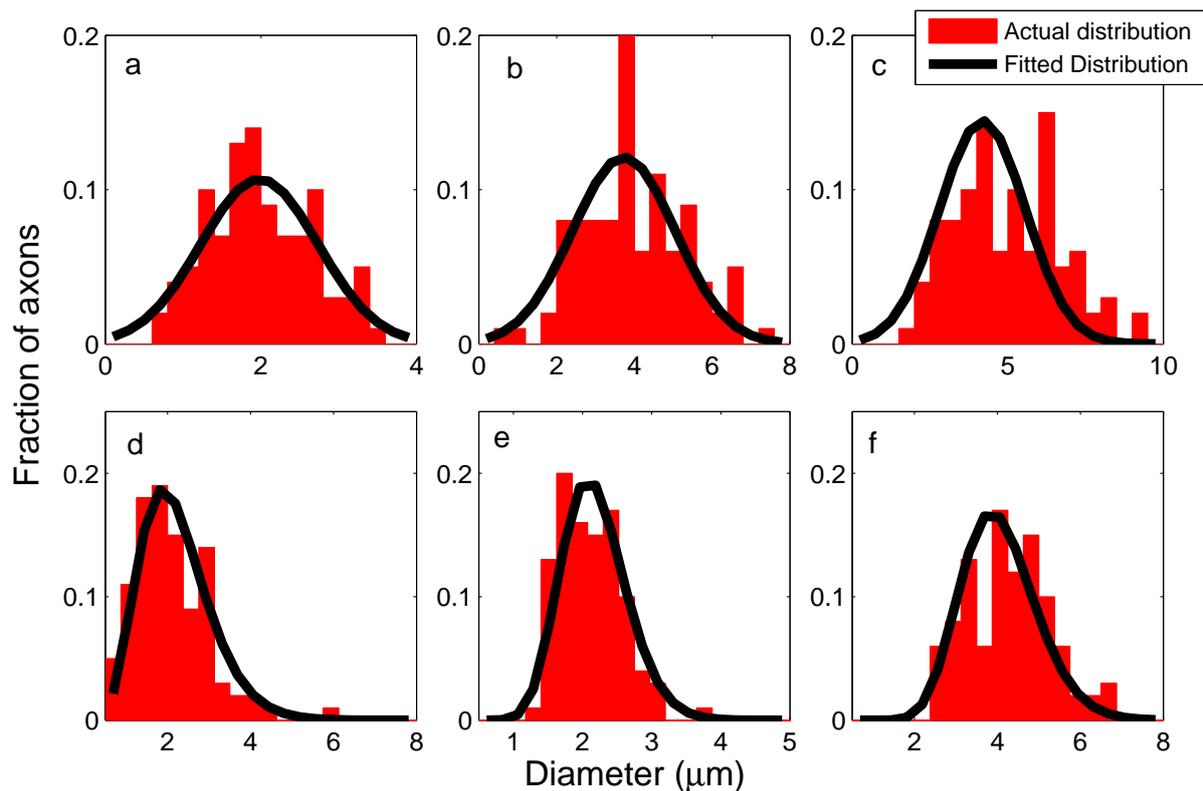
Figure 4.11: Six fitted diameter distributions for the total signal model. Red bars indicate the actual distributions, while the black curves indicate the fitted distributions. (a-c) Gaussian distributions with means of 2, 3, and 5 μm. (d-f) Gamma distributions with means of 2, 2.1, and 4.2 μm. The restricted range of frequencies corrects for the underestimation of the mean axon diameter for the distribution for the larger diameters. Two fit parameters (out of four in total) were used to create the black curves indicating the fitted distributions.

simplicity). Data were fitted separately to Eq. 4.9 with either Eq. 4.10 or Eq. 4.36 used as an extra-axonal term. Because there are five parameters in the new model, we repeated fitting with the intra-axonal diffusion coefficient fixed to its true value ($D_{in} = 1.0$ μm^2/ms).

Figure 4.12 compares fitted diameters with and without Eq. 4.36. Both models give similar results up to 4 μm. At cylinder diameters above 4 μm, diameters extracted using the original model start to level off at or below 5 μm, underestimating the true diameters. Results for the new model show some improvement. With a packing fraction of 0.5, the fitted diameter from the new model is more accurate than the old model for 5 μm (the fitted diameter is within 1 percent on the actual value), but fitted diameters level off afterwards. With a packing fraction of 0.7, extracted diameters from the new model do not seem to get "dragged down" until 8 μm (where the difference between true and fitted diameter is around 2.5 μm). However, fitted diameters at 6 and 7 μm are still underestimated by around 0.7 to 1 μm.
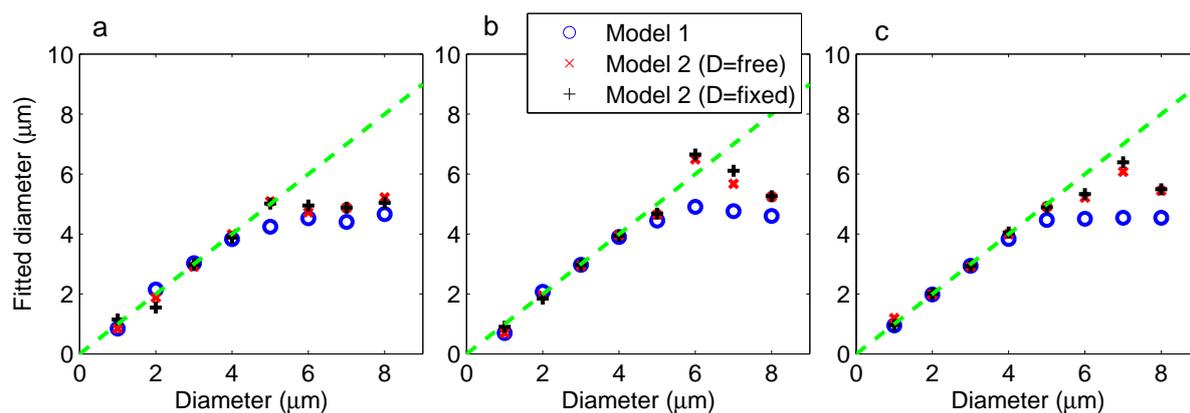
Figure 4.12: Comparison of fitted cylinder diameters using the simple extra-axonal model with a single diffusivity (Eq. 4.10) and the linear extra-axonal model (Eq. 4.36). For the linear model, fitting was performed twice. Once with the intra-axonal diffusion coefficient free and once with it fixed. Dashed line indicates ideal one-to-one correspondence. Subplots show results for different packing fractions. (a) Packing fraction = 0.5 (b) Packing fraction = 0.6 (c) Packing fraction = 0.7.

4.3.5 Effects of Noise

To determine the effects of noise on the estimations of axon diameter distributions, Gaussian noise was added to the transverse magnetization components in the simulations with the range of frequencies chosen to give the best axon diameter estimates.

Figure 4.13 shows two sets of comparisons for substrates with single diameter cylinders. The top row shows simulation results with 1 μm diameter cylinders, while the bottom row shows simulation results with 2 μm diameter cylinders. The three columns show increasing SNR values (50, 100, and 200). As one would expect, there is less variation in fitted diameters when SNR is high. The fitted diameter of the smaller cylinder is more affected by noise. Table 4.2 summarizes the results.

To illustrate how noise affects accuracy in diameter distributions, we give mean diameter estimates as a summary. As an example, Figure 4.14 shows fitted mean diameters obtained from noisy signals (Gaussian distributed diameters, mean diameter $\mu_N = 3$ μm). At a SNR of 50, many (≈ 40 percent) fitted values have pooled near zero, meaning that the fitting procedure has failed to find a physically plausible answer in some cases. As the SNR increases, the values cluster closer to the real value. At SNR = 50, approximately 50 percent of the fitted mean diameters lie within 0.5 μm of the true value, while at SNR = 200, this number rises to about 70 percent. Table 4.3 summarizes the results. Even at SNR = 200, the distribution of the fitted means is not as narrow as in Figure 4.13 (see bottom right panel) because there is a distribution of diameters rather than a single diameter like in Figure 4.13.

4.4 Discussion

We have used Monte Carlo simulated data to assess the model described by Eqs. 4.9, 4.10, and 4.25. Simulated signals were generated for both simple and complex geometries. Separate intra- and extra-axonal simulations were used to investigate each term in Eq. 4.9 before the

True Diameter	1 μm	2 μm
Mean diameter at SNR of 50	1.19 μm (34%)	2.28 μm (70%)
Mean diameter at SNR of 100	1.21 μm (64%)	2.12 μm (90%)
Mean diameter at SNR of 200	1.13 μm (92%)	2.04 μm (99%)

Table 4.2: Summary of results from fitting signals with added noise from a square packed lattice (packing fraction = 0.6) of identical axons: For each SNR, data were collected from a simulation and fitted to obtain the axon diameter. This process was repeated 1000 times. The mean of the distribution of fitted diameters are listed here. The percentage of fitted diameters within 0.50 μm of the true diameter for each case is given in brackets.
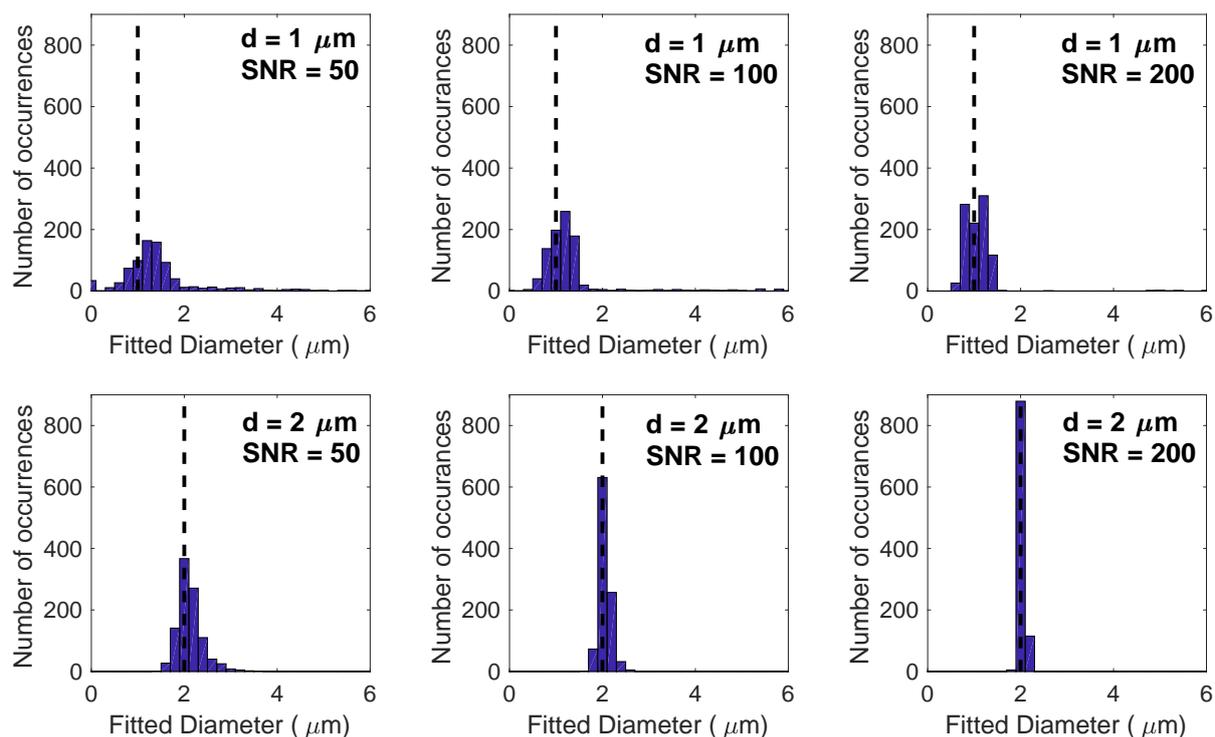
Figure 4.13: Fitted diameters using signals with added noise from a square packed lattice (packing fraction = 0.6) of identical axons: For each SNR, data were collected from a simulation and fitted to obtain the axon diameter. This process was repeated 1000 times. The distributions of fitted diameters are shown here. Dashed lines represent true values. As SNR is increased the distribution of fitted diameters converges closer to the actual diameter
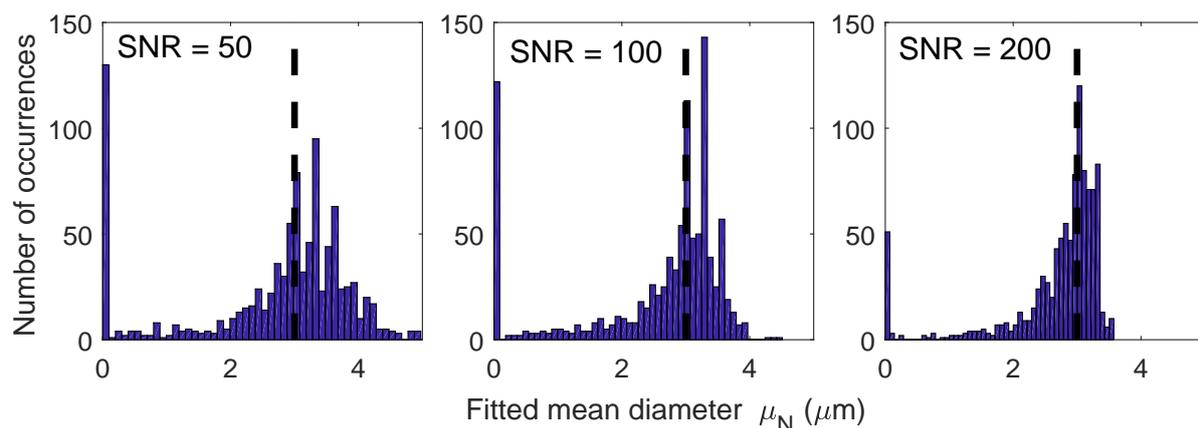
Figure 4.14: Fitted mean diameters using signals with added noise from 100 axons with Gaussian distributed diameters (actual mean diameter $\mu_N = 3$ μm): For each SNR, data were collected from a simulation and fitted to the modified AxCaliber model. The mean diameter μ_N was found. This process was repeated 1000 times. The distributions of fitted μ_N values are shown here. Dashed lines represent the true values. Again, as SNR is increased the distribution of fitted μ_N values converge closer to the actual mean diameter of 3 μm

True Mean Diameter	3 μm
Mean diameter at SNR of 50 (percent within 0.5 μm of actual value)	3.35 μm (44%)
Mean diameter at SNR of 100 (percent within 0.5 μm of actual value)	3.35 μm (60%)
Mean diameter at SNR of 200 (percent within 0.5 μm of actual value)	2.85 μm (77%)

Table 4.3: Summary of results from fitting signals with added noise from 100 axons with Gaussian distributed diameters (actual mean diameter $\mu_N = 3$ μm): For each SNR, data were collected from a simulation and fitted to the modified AxCaliber model. The mean diameter μ_N was found. This process was repeated 1000 times. The mean of the distribution of fitted μ_N values are listed here. The percentage of fitted mean diameters within 0.50 μm of the true mean diameter for each case is given in brackets.

whole model was investigated with combined intra- and extra-axonal simulations. As explained below, it was required that the range of gradient frequencies used in the simulation be restricted for Eq. 4.10 to be accurate. With restricted frequencies available on pre-clinical MRI machines, it was possible to measure axons as small as 1 μm and as large as 4 μm.

For the intra-axonal simulation, the acquired signals, combined with the analytical expressions, can be used to predict the diameters of small cylinders accurately. We also successfully fitted diameter distributions for Gaussian and gamma distributions.

For the full simulation, the model may not be adequate. For the collection of identical cylinders, larger diameters were underestimated. The model also had problems accurately fitting cylinder diameter distributions, especially with distributions containing large diameters. Certain fitted parameters, such as the mean diameter of a Gaussian distribution, underestimated the actual value. Similar behaviour has been noticed by other groups [195]. These results indicate that the model used in these simulations, the same model used in AxCaliber (Ref. [11]), does not accurately model the extra-axonal signal.

If a more accurate extra-axonal model is developed, such as the improved exchange model with variable radii developed by Lam et. al. [87], it could be combined with the existing intra-axonal model. Ideally, any new model should keep the number of free parameters as low as possible. Here we worked around some of the problems by choosing frequency ranges where the original model was approximately true, keeping the number of parameters to a minimum.

The MR signal from white matter is due to the combined contributions of intra-axonal and extra-axonal water. Close packing of cylinders will result in a relatively large intra-axonal

signal. Conversely, in geometries with lower packing fractions, the signal should be mostly due to extra-axonal water. Also, as the extra-axonal space increases, water should experience fewer restrictions, and the hindered diffusion term should approach that of free diffusion. The simulations indicate that the intermediate range is causing problems. The hindered diffusion term also assumes that water molecules, though experiencing effects of barriers, are still able to move throughout the pores. When we have many randomly placed highly packed cylinders, we might have pockets of trapped extra-axonal water, which are not currently taken into account in the model. This would lead to a deviation from the simple exponential decay describing the extra-axonal signal. In a three dimensional simulation with non-parallel axons, the extra-axonal water might not be trapped in pockets because it might be able to move to another area using the third dimension. Thus, as more sophisticated models with orientation dispersion are made, the problem of the trapped extra-axonal may become less relevant.

An assumption of the extra-axonal model is that the *ADC* in the extra-axonal space has a single value at all frequencies. When the model tries to assign a single diffusion coefficient to a frequency-dependent spectrum, it might have adverse effects on other parameter estimates. We worked around this problem by simulating a narrower range of frequencies over which the *ADC* was almost constant and Eq. 4.10 was satisfied.

Originally, we used 400 signals to fit cylinder sizes and distributions. It would be time-consuming to collect this many signals on real samples. Later, we used only 100 signals, but this would still lead to long experimental times. Monte Carlo simulations will need to be used to find an optimum number of data points needed to obtain adequate results. Future simulations will also be used to determine which frequencies and gradients are the best choices. This will be discussed in Chapter 5.

We also looked at how noise affected the fitted parameters. As with other studies, we also find that low signal-to-noise ratio has an effect on the accuracy of fitted diameter values [192]. We found that the value for the fitted axon diameter for small axons would be more affected by low SNR, and that the effects could be mitigated with a higher SNR.

Our simulations assumed parallel axons and gradients applied perpendicular to these axons. For actual tissue measurements where there is axon orientation dispersion, our model could

cause an overestimation of axon diameters [132]. Future studies will require models with axon orientation dispersion for more accurate estimations of axon diameters and could require more sophisticated gradients similar to diffusion tensor gradients rather than single direction as used here.

Even when using the restricted range of frequencies, the model had difficulty estimating accurate diameters in the upper range of those studied. This is probably because the extra-axonal model does not describe the true extra-axonal diffusion spectrum in this regime. We expect that these can be extracted with lower OGSE frequencies than the ones studied here, or through a method combining OGSE and PGSE [76, 141]. a priori knowledge of the range of axon diameters in the sample or the range to be studied could be necessary to obtain accurate results.

4.5 Conclusion

Using OGSE pulse sequences with the AxCaliber fit model allows measurements of axon diameter distributions containing smaller axons. Careful selection of the gradient frequency range (corresponding to the diffusion time) is necessary to ensure a constant extra-axonal *ADC* with frequency allowing for a more correct model of hindered diffusion in the extra-axonal space and thus a more accurate estimation of axon diameters. With frequencies available on pre-clinical MRI machines, it was possible to measure axons between 1 μm to 4 μm.

Chapter 5

Optimization of Parameters

5.1 Introduction

In Chapter 4, we looked at how well the AxCaliber model works. Now we are going to do a more thorough investigation of what happens in the presence of noise. We will also look at how parameter accuracy and precision change depending on the number of OG frequencies and gradient strengths. This analysis will be repeated using three different sets of gradient measurements (clinical to pre-clinical strengths).

The experiments that the Monte Carlo simulations model require the collection of many images at many different frequencies and using many different gradient strengths. This is time consuming. Making fewer measurements can reduce the imaging time. For phantoms and *ex vivo* imaging, this is less of an issue (only the cost of scanner time). For live imaging, it can become a problem for the health and comfort of the subject. Therefore this chapter focuses on reducing the number of images required while maintaining as much precision as possible in order to make a method for live animals, with the goal of moving toward a clinically relevant method. In practice, the desired accuracy and precision will depend on a number of factors, many of which are unknown at this time. Some of these factors include on the variability of the parameter of interest (e.g. axon diameter, axon density, etc) within the general population and the amount of change occurring with disease.

Other groups have looked into optimization of their pulse sequences as another means to

optimize data collection. These sequences have included GEN and trapezoidal gradients and optimization sometimes involves comparing signal differences using simulations or minimizing some cost function, as will be described below.

Alexander [6] developed an optimization procedure for measurements with rectangular gradients. Their protocol minimizes an objective function to determine ideal experimental parameters (such as gradient strength and direction, pulse separation, pulse duration). Given the number of desired measurements, optimization gives the experimental parameters in each measurement that maximize sensitivity to model parameters. Optimization requires choosing a signal model, so they used a two compartment model with parallel axons and impermeable boundaries. Optimization also requires *a priori* parameter estimates, so that final optimized experimental parameters depend on expected situations. During optimization, experimental parameters were constrained to feasible values if necessary. In their protocol, they assumed $N = 30$ directions with $M = 4$ measurements each. Simulations were performed to compare different protocols (with maximum gradient strengths of 70 or 200 mT/m), finding that axons with diameters between 10 and 20 μm were distinguishable in each protocol, though smaller diameters ($d = 2$ to 4 μm) were only distinguishable at higher gradient strengths (200 mT/m). Advantages of this method and the ones that follow are that they should provide optimized measurement parameters for PGSE experiments, if given *a priori* information on the sample.

Drobnjak et al. [44, 45] applied the experimental optimization protocol from Alexander [6] to arbitrary gradient waveforms. Instead of rectangular gradients, they use generalized gradients, calculating the required analytical cylinder signal expressions using a matrix formalism [33, 44]. The signal model itself was the same as Ref. [6], though limited to one dimension (assuming that gradients are applied perpendicular to axons). They determined the shape of the optimized waveforms by minimizing an objective function. Each measurement, the number of measurements being chosen beforehand, has its own separate optimized waveform. Like Ref. [6], the optimization process uses *a priori* parameter estimates based on expected values and allows for certain constraints, such as those based on hardware. The protocol was tested using maximum gradient strengths of 40, 70, and 400 mT/m. Simulation data acquired with the optimized measurements (called GEN) were compared to data acquired with rectangular

gradients, showing improvement in precision of estimated parameters (for radii up to 5 μm). They also found that optimal waveforms tended to include some form of oscillation. Their later work [44] explored whether varying gradient orientation during a single measurement improved sensitivity, finding little improvement over fixed orientation measurements, at least for cylindrical pores.

Afterwards, Siow et al. [160] performed experiments on water-filled microcapillary tubes (inner radii = 1 to 10 μm) using the GEN protocol. They used three sets of maximum gradient strengths (40, 80, and 200 mT/m). Experimental results from GEN were compared with results from PGSE, confirming that GEN waveforms provided an improvement over PGSE, especially at gradient strengths of 40 mT/m.

Perrault et al. [141] used non-uniform oscillating gradient spin echo sequences which combine features of OGSE and PGSE to allow greater sensitivity to small structures. After validating the OGSE analytical signals using simulations, they used a bootstrap procedure (500 runs) to compare sensitivity of PGSE and OGSE at different maximum gradient strengths. Maximum gradient strengths were 40 (typical clinical imaging maximum gradient strengths), 80 (the upper-end of clinical imaging maximum gradient strengths), and 300 (NIH Human Connectome Project maximum gradient strength, the highest for human imaging) mT/m. They simulated 100 signals for separate sets of diameters (d = 4 to 8 μm) using the OGSE analytical expressions and added 10% noise for the bootstrap analysis. Precision and accuracy were expressed in terms of the mean and standard deviation of parameters from the bootstrap trials. They found that with maximum gradient of 40 mT/m, diameter estimates for both PGSE and OGSE were low in precision and accuracy. At gradients of 80 mT/m, larger diameters had an improvement for OGSE in comparison to PGSE. As expected, increasing the maximum gradient strength to 300 mT/m improved precision and accuracy for both sequences.

Li et al. [97] compared fitted diameters and diffusion coefficients from filled capillary tubes (d = 1.5 to 19 μm) for different numbers of OGSE frequencies. *ADC* frequency spectra were modeled using a single compartment model. Extracted diameters showed almost no dependence on number of OGSE frequencies (N = 2 to 12). Extracted diffusion coefficients were independent of number of frequencies down to the 4.5 μm tube, where it showed dependence

on N. All diffusion coefficients were poorly estimated with the 1.5 μm tube.

While Li et al. compared estimated parameters acquired with experimental measurements of capillary tubing, modeling water inside the tubes, the work presented here uses Monte Carlo simulations on a variety of simulated geometries, including identical square packed cylinders and diameter distributions. In this study, the simulations are used to find how the precision of parameter estimates change depending on the number of frequencies and gradient strengths used in the experiment. Three sets of gradient strengths (with maximum strengths of 80 mT/m, 300 mT/m, and ∼900 mT/m) are used to study the method for clinical, Connectome, and pre-clinical situations. In addition, results are compared when multiple gradient measurements (or b-values) at a given frequency are spaced out over a range or collected at two points, one zero and the other nonzero (with the nonzero gradient repeated). The aim of these experiments is to reduce imaging time to a clinically useful sequence using clinically available gradient strengths.

5.2 Methods

The model and Monte Carlo simulation used in this study were similar to the previous study in Chapter 4. Briefly, a two compartment model like Eq. 4.9 was used. However, the presence of an axon diameter distribution was not incorporated into the model itself. Therefore, the hindered signal was given by Eq. 4.10 and the restricted signal by Eq. 4.13 (not Eq. 4.25).

For this model, we are assuming that all axons have the same diameter, even if they do not. The diameter that we extract is called the "effective mean axon diameter," or AxD. We can think of AxD as being the single axon diameter that would give the same diffusion signal as the entire collection of axons [7]. It has been shown that AxD correlates with the quantity

$$AxD = \frac{\sum_i d_i^3}{\sum_i d_i^2},$$

(5.1)

where d_i is the diameter of the ith axon [7, 48, 195].

5.2.1 Monte Carlo Simulations

Monte Carlo simulation parameters were the same as in Chapter 4. We assumed impermeable barriers [192, 195] and ignored T_2 differences between tissues. For simulations in simple lattices, we used N= 57344 particles. For diameter distribution simulations, we used N= 114688 particles. The time step used in all simulations was 1 μs.

5.2.2 Cylinders

We performed two types of simulations. Initially, we performed "intra-axonal" simulations, where water is contained inside the axons, but not outside. In this case, the signal can be described by Eq. 4.13 alone. We then distributed water both inside and outside the axons. The signal should be described by Eq. 4.9 (a "full" simulation). These simulations ("intra-axonal" and "full") were used together with the geometries given below.

The intra-axonal diffusion coefficient (D_{in}) was set to 1.0 μm^2/ms [192]. In simulations with intra- and extra-axonal water, we first set the extra-axonal diffusion coefficient to 2.5 μm^2/ms [154]. We repeated the simulations with $D_{ex} = 1.0$ μm^2/ms to see if using equal diffusivities affected results.

5.2.2.1 Geometry 1: Single cylinder

The simplest situation considered is diffusion inside a single cylinder of diameter d. Ten diameters $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ μm were tested. After each simulation, the cylinder diameter and intra-axonal diffusion coefficient were extracted by fitting the signal to 4.13 (together with 4.14).

5.2.2.2 Geometry 2: Identical cylinders on a square lattice

The next geometry considered is a collection of identical cylinders of diameter d arranged on a square lattice. Water was distributed everywhere, both inside and outside the cylinders (a "full" simulation). Five diameters $\{1, 2, 3, 4, 5\}$ μm were tested with packing fractions of 0.5, 0.6, and 0.7. The signals were then fitted to Eq. 4.9, with Eq. 4.13 and 4.10. Four parameters were

extracted from each fit: cylinder diameter, intra-axonal diffusion coefficient, packing fraction, and the hindered diffusion coefficient.

5.2.2.3 Geometry 3: Distribution of cylinder diameters

To model a more realistic situation, diameters were sampled from a gamma distribution [61]. A set of 100 cylinders were randomly distributed over a square domain with periodic boundary conditions for this model. Substrates were generated according to the algorithm in Ref. [61]. Water was distributed everywhere, both inside and outside the cylinders (a "full" simulation).

Five distributions with mean diameters of 1.5, 2.0, 2.1, 4.1, and 4.3 μm were simulated, with associated *AxD*s of 2.1, 2.3, 2.9, 4.8, and 6.1 μm. The *AxD*s were calculated using Eq. 5.1. Some of these distributions were used in Chapter 4. When choosing a distribution, we have some control over the mean diameter, but not necessarily *AxD*, which is the important quantity of interest. Therefore, covering equal intervals in *AxD* is not easily achievable. Packing fractions were between 0.63 and 0.78.

Signals were fitted to Eq. 4.13, with Eq. 4.13 and 4.10. Because the geometry was more complex, the intra-axonal diffusion coefficient was held fixed to its actual value for all fits [11]. We found that this reduced the number of local minima when fitting. Three parameters were extracted from each fit: *AxD*, the packing fraction, and the hindered diffusion coefficient.

5.2.3 Three Gradient Sets

Twenty frequencies ranging from 50 to 1000 Hz in steps of 50 Hz were used with an echo time of 44.52 ms. There were three different sets of gradients used.

1. Five gradient strengths were used at each frequency, for a total of 100 signals. These pulse sequence parameters were chosen to be feasible for experiments using a BGA6 gradient set (Bruker Biospin) on our 7 T magnet (maximum gradient strength of 1.01 T/m). For $f_m > 200$ Hz, the gradient strengths were 0, 60, 70, 80 and 90% of maximum. The gradient strengths for 200 Hz or less were smaller. For $f_m = 50$ Hz, $G = 0$, 1.5, 4, 7, 10% of maximum; for $f_m = 100$ Hz, $G = 0$, 4, 7, 10, and 15% of maximum; for

$f_m = 150$ Hz, $G = 0$, 15, 20, 25, and 30% of maximum; for $f_m = 200$ Hz, $G = 0$, 30, 40, 50, and 60% of maximum (see Table 4.1).

2. Seven gradient strengths were used at each frequency, for a total of 160 signals. Gradients were {0, 50, 100, 150, 200, 250, 300 mT/m}. The highest b-value in the set is 1.31 ms/μm^2.

3. Six gradient strengths were used at each frequency, for a total of 120 signals. Gradients were {0, 40, 50, 60, 70, 80 mT/m}. The highest b-value in the set is 0.09 ms/μm^2.

5.2.4 Data Fitting

At the end of the simulations, Gaussian noise was added to each component of the transverse magnetization so that SNR was 100. Previous studies with OGSE on pre-clinical scanners have reported SNR values from 50 [143] to 120 [149]. This was repeated 1000 times and the fitted parameters were saved for analysis. The means and standard deviations of each of the 1000 results were calculated and used to assess accuracy and precision of the results. We denote the mean and standard deviation for a particular variable by a subscript. The distribution means for the intra-axonal diffusion coefficient, diameter (AxD), and packing fraction (if applicable) are denoted by μ_D, μ_d (μ_{AxD}), and μ_f. Their associated standard deviations are denoted by σ_D, σ_d (or σ_{AxD}), and σ_f.

Signal data were fitted to the model with nonlinear least-squares regression using the Optimization Toolbox in MATLAB [113]. Packing fractions were constrained to lie between 0 and 1. Diffusion coefficients were constrained to lie in the range [0, 2.5] μm^2/ms. Diameters (or AxD) were constrained to the range [0, 20] μm. Before adding noise, the fitting procedure was repeated 50 times with random initial parameters. At the end, the fitted parameters providing the smallest sum-of-squares were chosen as the best fit. The best fit parameters were then used as the initial parameters when fitting the noisy signals. We were mainly interested in the effects of noise on known "true" solutions and therefore wanted to keep the problem of local minima separate. We also wanted to reduce computation time as much as possible. If the solutions are truly unknown, it would be better to repeat the fitting procedure with random initial parameters

like before.

5.2.5 Number of Frequencies

Data were fitted using different numbers of frequencies. Initially, all 20 frequencies were used (50 – 1000 Hz). This serves as a baseline. Higher frequencies were then successively removed before refitting the data to the model. The number of frequencies used will be denoted by n_f. Data were fitted with n_f = 5, 10, and 15 frequencies, corresponding to maximum frequencies of 250 Hz, 500 Hz, and 750 Hz.

5.2.6 Gradient Subsets

For each set of measurements, the highest gradients were successively removed from the data before refitting. We also fitted data using one nonzero gradient (e.g. $G = \{0, G_i\}$). Finally, fitting was performed using one nonzero gradient repeated multiple times. This was done by replicating the simulation results before adding noise. The subsets are summarized in Table 5.1. When necessary, the highest gradient strength in a subset is termed G_{high} to distinguish it from the maximum gradient strength for the entire set of gradient strengths (G_{max}). For the $G_{max} = 80$ mT/m gradient set, only the subsets with all six gradients or one nonzero gradient (at $G = 80$ mT/m) were studied. A focus on pre-clinical gradient strengths was chosen for this chapter given the difficulty of all methods to use gradient strengths below $G = 80$ mT/m.

5.3 Results

Simulations were performed with the goal of moving toward a clinically relevant imaging method by reducing imaging time. The results from these simulations are presented here. Many of the fits contain results that are pinned at the upper or lower bounds. This tends to happen with smaller gradient strengths where there is less signal decay. Because this will sometimes affect the sample statistics (mean and standard deviation), we will also show examples of the actual parameter distributions and make a note of when it occurs. If they were removed, results would appear better than they actually are.

Subset #	$G_{max} = 900$ mT/m	$G_{max} = 300$ mT/m	$G_{max} = 80$ mT/m
1	$G_0, G_1, G_2, G_3,$ G_4	$G_0, G_1, G_2, G_3,$ G_4, G_5, G_6	$G_0, G_1, G_2, G_3,$ G_4, G_5
2	G_0, G_4	G_0, G_6	G_0, G_5
3	G_0, G_1, G_2, G_3	$G_0, G_1, G_2, G_3, G_4, G_5$	
4	G_0, G_3	G_0, G_5	
5	G_0, G_1, G_2	G_0, G_1, G_2, G_3, G_4	
6	G_0, G_2	G_0, G_4	
7	G_0, G_1	G_0, G_1, G_2, G_3	
8		G_0, G_3	
9		G_0, G_1, G_2	
10		G_0, G_2	
11		G_0, G_1	

Table 5.1: Gradient subsets used in data fitting. The first column lists subsets for the first gradient set (7 subsets). The second column lists subsets for the second gradient set (11 subsets). The third column lists subsets for the third gradient set (2 subsets). Every second subset uses only one nonzero gradient. As defined here, higher subset numbers correspond to using fewer and smaller gradient strengths.

5.3.1 Intra-axonal Simulations

In this section, results of the intra-axonal simulations are presented. We first give an overview of the two fitted parameters (d and D_{in}) when using all measurements.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.1a shows μ_d from the intra-axonal simulations as a function of diameter. The μ_d from each diameter are in good agreement with the actual values. For $d > 1$ μm, differences are under 4 percent. For $d = 1$ μm, the difference is approximately 30 percent ($\mu_d = 1.34$ μm).

We should look at some of the distributions themselves. Figure 5.2a and 5.2b show the fitted diameter distributions for $d = 1$ μm and $d = 3$ μm. The distribution for $d = 3$ μm was relatively symmetric (this feature cannot be seen in Figure 5.4 because of the scale used). In the case of $d = 1$ μm, there are a few fitted diameters that have clustered near the upper bound ($17/1000 > 19$ μm). This skews the mean and standard deviation for the $d = 1$ μm cylinder. Otherwise, the peak in the distribution is around 1 μm.

Figure 5.3a shows μ_D found using each cylinder diameter. For $d > 1$ μm, the values are in good agreement with the actual value (within 1 percent). Error decreases at higher diameters.
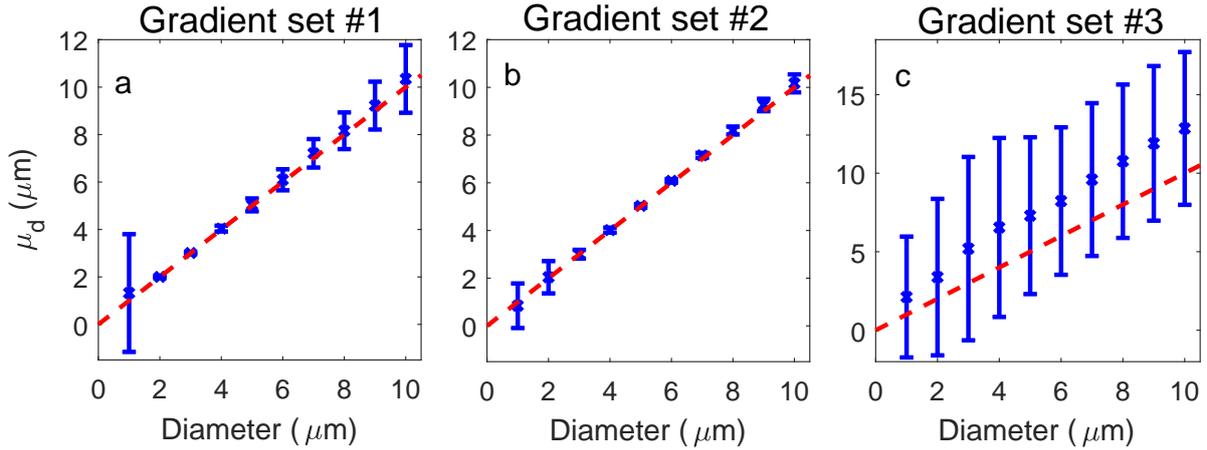
Figure 5.1: Intra-axonal simulation: μ_d as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted diameters (σ_d). Dashed line indicates a perfect fit. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All measurements in each set were kept when fitting to the model. For $G_{max} = 80$ mT/m, μ_d are all higher than the actual diameters. **Note**: (c) uses different y-axis limits.

For $d = 1$ μm, μ_D is 1.3 μm^2/ms and has a large error.

Figures 5.2c and 5.2d show respective D_{in} distributions for $d = 1$ μm and $d = 3$ μm. For $d = 1$ μm, almost half of the fitted D_{in} are clustered near the upper bounds ($460/1000 > 2.4$ μm^2/ms). Most of the rest are below 0.5 μm^2/ms, with a peak near 0.1 μm^2/ms.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.1b shows μ_d from the intra-axonal simulations as a function of diameter. For $d > 1$ μm, the μ_d for each diameter are within 3 percent of the true values. For $d = 1$ μm, the difference is 16 percent ($\mu_d = 0.84$ μm).

Figure 5.4a and 5.4b show the fitted diameter distributions for $d = 1$ μm and $d = 3$ μm. The diameter distribution for $d = 3$ μm is relatively symmetric. In the case of $d = 1$ μm, many fitted diameters have clustered near the lower bound ($366/1000 < 0.02$ μm). The peak of the distribution is between 1.1 to 1.3 μm.

Figure 5.3b shows μ_D for as a function of diameter. Like those of the first gradient set, the μ_D have better agreement with the true value at higher diameters ($d > 2$ μm). For $d = 1$ μm and $d = 2$ μm, μ_D are higher than the true value ($D_{in} = 1.8$ μm^2/ms and $D_{in} = 1.3$ μm^2/ms, respectively). As we might expect, error decreases as diameters become larger.
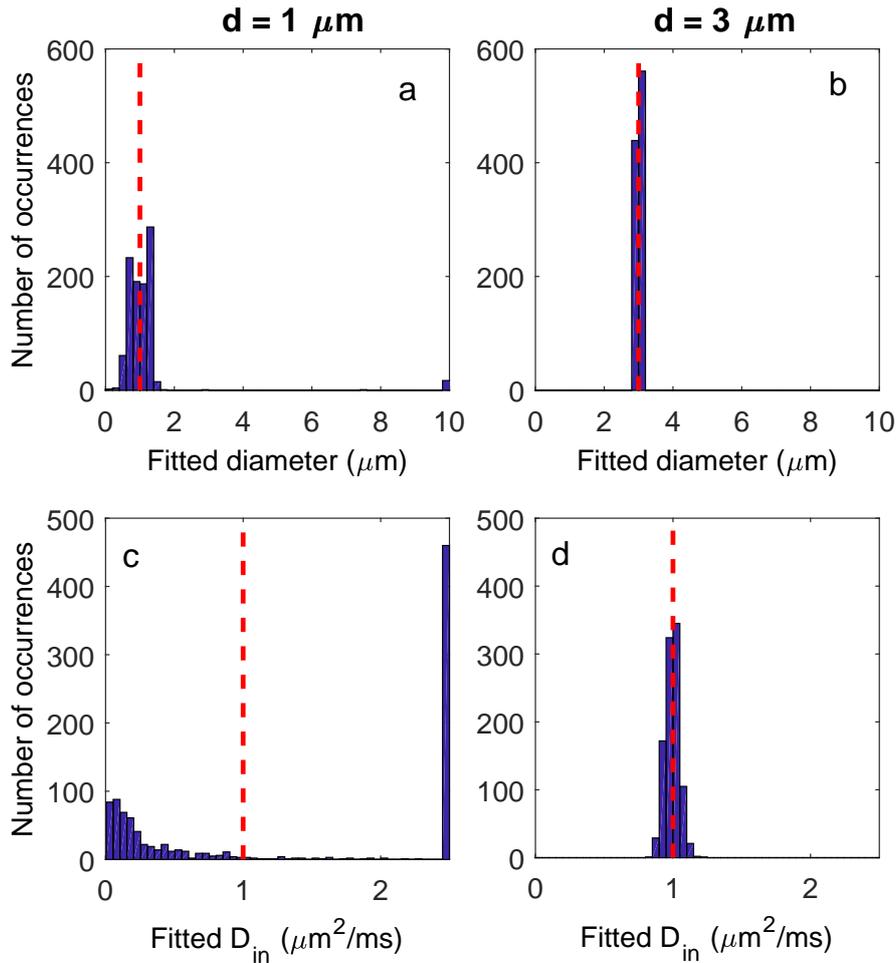
Figure 5.2: Intra-axonal simulations: Distributions of fitted diameters for (a) $d = 1$ μm and (b) $d = 3$ μm. Distributions of fitted intra-axonal diffusion coefficients for (c) $d = 1$ μm and (d) $d = 3$ μm. Simulations use the $G_{max} = 900$ mT/m set of measurements. All measurements were kept when fitting to the model. The fitted D_{in} for $d = 1$ μm are inaccurate. Both distributions are narrower with $d = 3$ μm.

Figures 5.4c and 5.4d show D_{in} distributions for $d = 1$ μm and $d = 3$ μm. For $d = 1$ μm, over half of the fitted D_{in} are clustered near the upper bound ($698/1000 > 2.4$ μm^2/ms). The rest are pushed against the lower bound. The distribution for $d = 3$ μm is wider than in Figure 5.2d.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.1c shows μ_d from the intra-axonal simulations. All μ_d are larger than the actual values. The relative difference between μ_d and the true diameter becomes smaller as d increases. At $d = 1$ μm, the difference is around 100 percent, decreasing to around 30 percent when $d = 10$ μm.
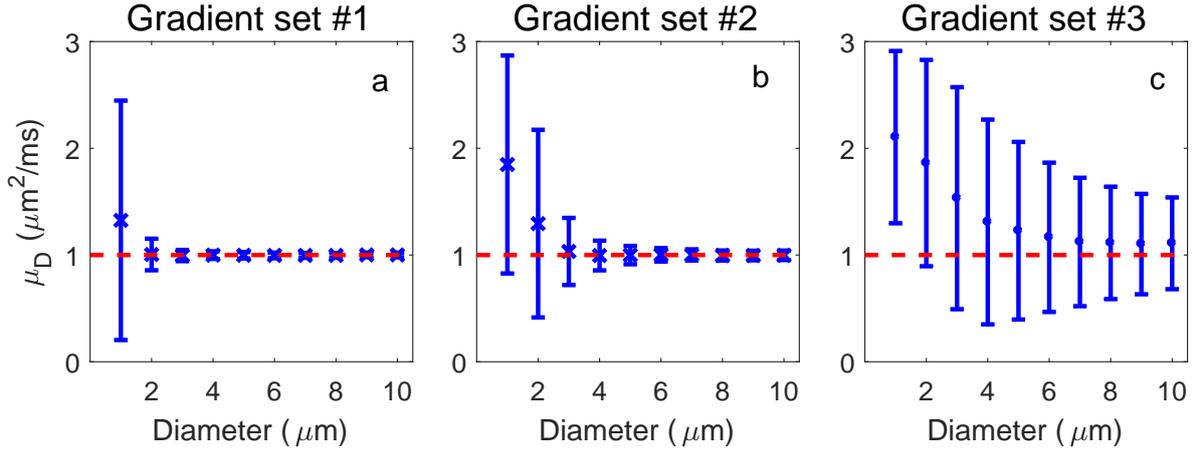
Figure 5.3: Intra-axonal simulations: μ_D as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted intra-axonal diffusion coefficients (σ_D). Dashed line indicates a perfect fit. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All measurements in each set were kept when fitting to the model. The μ_D are more accurate at larger diameters and with higher gradient strengths.

Figures 5.5a, 5.5b and 5.5c show diameter distributions for $d = 1$ μm, $d = 3$ μm, and $d = 5$ μm. We see the wide spread in fitted diameters. For $d = 1$ μm, large numbers of fitted diameters have clustered near $d = 0$ μm (353/1000 are < 0.02 μm), with some clustered at $d = 20$ μm (38/1000 > 19 μm). The distribution has a broad peak between 2 to 4 μm. As we move to $d = 3$ μm, there are less values at $d = 0$ (75/1000 < 0.02 μm), though more have appeared at $d = 20$ μm (127/1000 > 19 μm). The peak of the distribution is around 3.5 μm. By the time we reach $d = 5$ μm, fitted values at $d = 0$ μm have disappeared, but those near $d = 20$ μm remain (127/1000 > 19 μm). The peak is between 5.3 and 5.7 μm.

Figure 5.3c shows μ_D as a function of diameter. The μ_D are closer to the actual value with larger diameters, with errors decreasing at higher diameters. Figures 5.5d, 5.5e, and 5.5f show parameter histograms for $d = 1$ μm, $d = 3$ μm, and $d = 5$ μm. We can see that the diffusion coefficient is poorly estimated. For $d = 1$ μm, a majority of fitted parameters are near the upper bound (768/1000 > 2.4 μm^2/ms). For $d = 3$ μm, nearly half are near the upper bound (466/1000 > 2.4 μm^2/ms). As we move to $d = 5$ μm, there are less at the upper bound (189/1000 > 2.4 μm^2/ms). The distributions themselves are skewed towards smaller D_{in}, but less so with $d = 5$ μm.
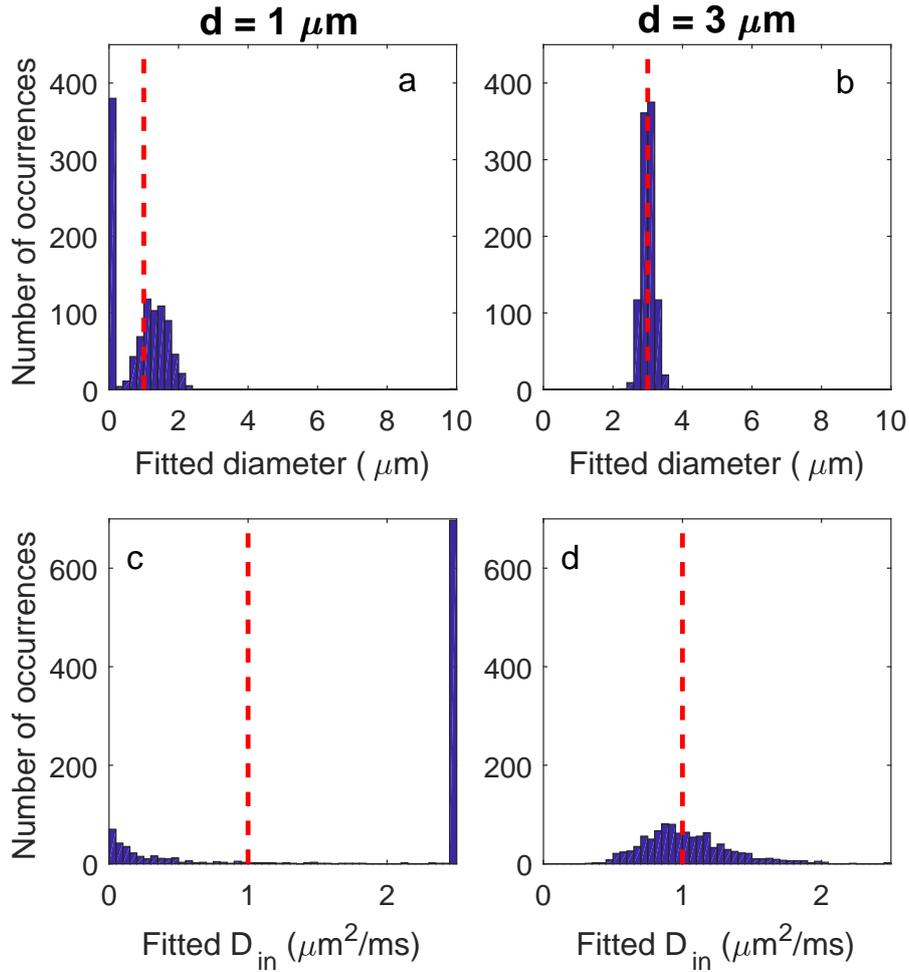
Figure 5.4: Intra-axonal simulations: Distributions of fitted diameters for (a) $d = 1$ μm and (b) $d = 3$ μm. Distributions of fitted intra-axonal diffusion coefficients for (c) $d = 1$ μm and (d) $d = 3$ μm. Simulations use the $G_{max} = 300$ mT/m set of measurements. All measurements were kept when fitting to the model. The fitted D_{in} for $d = 1$ μm are inaccurate. Both distributions are narrower using $d = 3$ μm, but wider than in Figure 5.2.

5.3.1.1 Gradient Subsets

We now show how μ_d and σ_d change when gradients are successively removed from the data. In this section, all frequencies ($n_f = 20$) are used.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.6a shows μ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. For $d = 2$ and 3 μm, μ_d show little change over the subsets and agree well with the diameter of the cylinders used in the simulations indicating the accuracy of the simulations is very good. It is therefore important to study the precision of the results. The μ_d for 4 and 5 μm show little change until

135
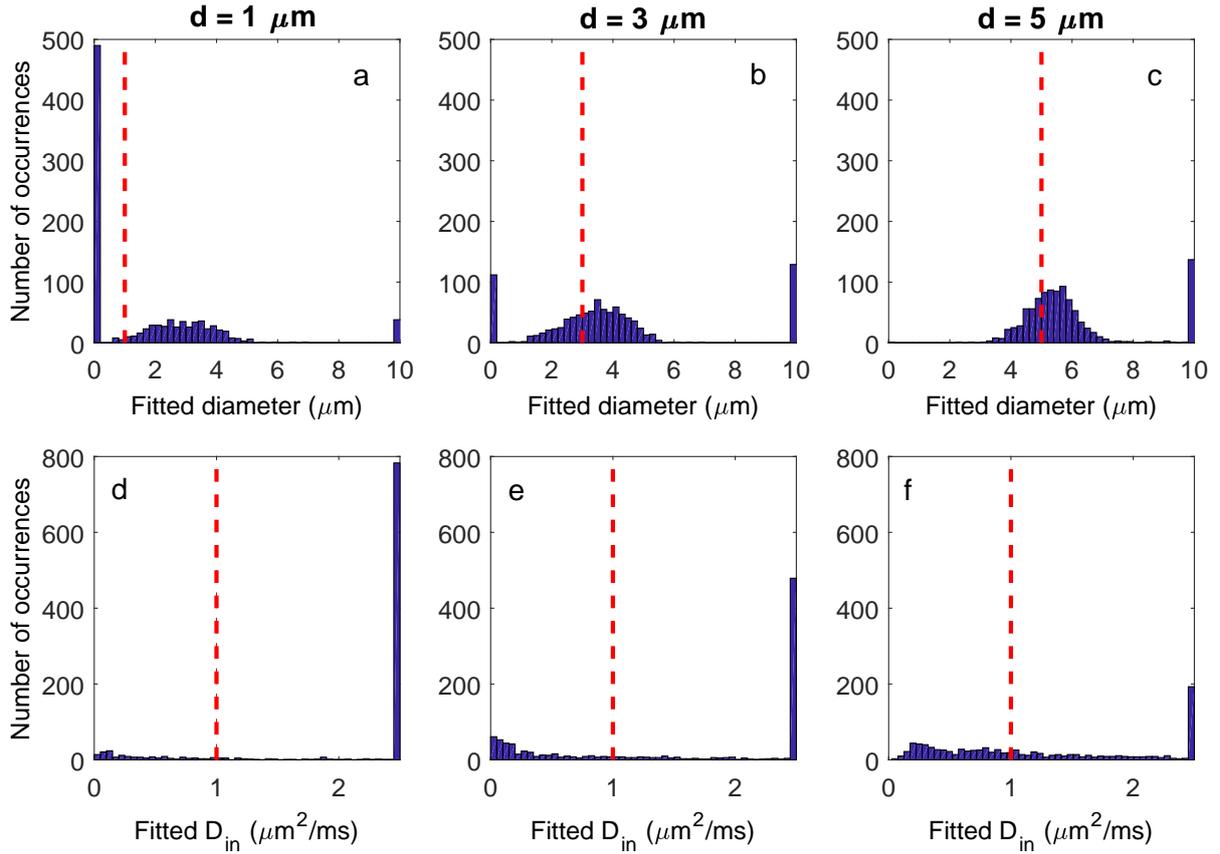
Figure 5.5: Intra-axonal simulations: Distributions of fitted diameters for (a) $d = 1$ μm, (b) $d = 3$ μm, and (c) $d = 5$ μm. Distributions of fitted diffusion coefficients for (d) $d = 1$ μm, (e) $d = 3$ μm, and (f) $d = 5$ μm. Simulations use the $G_{max} = 80$ mT/m gradient set. All measurements were kept when fitting to the model. Fitted diameters are more accurate and distributions are narrower for larger diameters. Fitted D_{in} are inaccurate.

the 4th subset. The μ_d for $d = 1$ μm increases from just above 1 μm to 2 μm.

Figure 5.7 shows σ_d for $d = 1$ to 5 μm and shows many interesting features. The first is that subsets with smaller gradients, whether using just two gradients or more, tend to have higher σ_d. For some diameters ($d = 1$ μm), there is a steady rise over the subsets (from around 2.4 μm to 5 μm) . In others, σ_d seems to undergo a more exponential increase, with larger increases after the 4th ($d = 5$ μm) or 5th ($d = 3$ μm) subset. Exceptions are $d = 2$ μm and $d = 3$ μm. In these, σ_d remains relatively steady compared to the others ($\sigma_d \leqslant 0.4$ μm).

The second is that multiple gradients spaced up to G_{high} cause a slight improvement over just two gradients. In addition in almost all cases, two gradients (e.g. $G = \{0, G_{high,1}\}$) are better than multiple gradients with a smaller G_{high} (e.g. $G = \{0, G_1, G_2,\ldots, G_{high,2}\}$, where $G_{high,2} < G_{high,1}$). This result is likely due to the fact that most gradients in this set are spaced

relatively far apart ($\Delta G \approx 100$ mT/m).
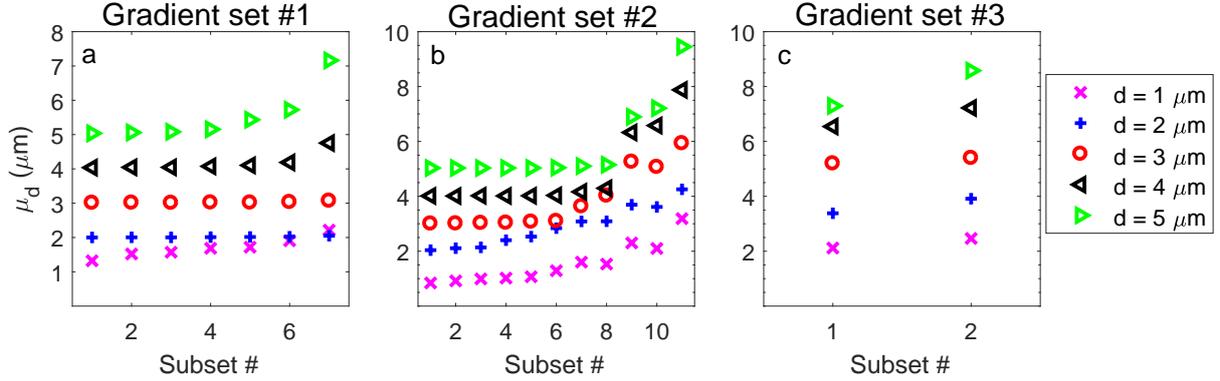
Figure 5.6: Intra-axonal simulations: μ_d for different gradient subsets. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m and $G_{max} = 300$ mT/m, most μ_d show little change until the higher subsets. **Note**: (b) and (c) use different y-axis limits.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.6b shows μ_d for $d = 1$ to 5 µm using the gradient subsets in Table 5.1. With diameters of 1 and 2 µm, μ_d increase over the subsets. With diameters of 3 to 5 µm, μ_d show little change up to the 6th subset. At the 7th subset ($G_{high} = 150$ mT/m), μ_d start to increase.

Figure 5.7b shows σ_d for $d = 1$ to 5 µm. The plot shares several features with Figure 5.7a. In subsets which include higher gradients, σ_d decreases. Also, using two gradients ($G = \{0, G_{high,1}\}$), if $G_{high,1}$ is high enough, produces better results than multiple gradients spaced up to a smaller gradient ($G = \{0, G_1, \ldots, G_{high,2}\}$, where $G_{high,2} < G_{high,1}$). There is a steady increase in σ_d with gradient subset number for $d = 1$ µm and $d = 2$ µm (from approximately 0.6 µm to 5.6 µm). There is a faster increase in σ_d with gradient subset number for $d > 2$ µm. All σ_d begin under 0.2 µm and eventually begin to spread apart at the higher subsets where the largest gradient strengths used were relatively small. At $G_{high} = 250$ mT/m, the σ_d for $d = 3$ µm split from the rest, followed by those of $d = 4$ µm at $G_{high} = 150$ mT/m. There appears to be a change at the 9th subset, corresponding to $G_{high} = 100$ mT/m, where there is a jump in σ_d. This was the point where mean fitted parameters also jumped (Figure 5.6).
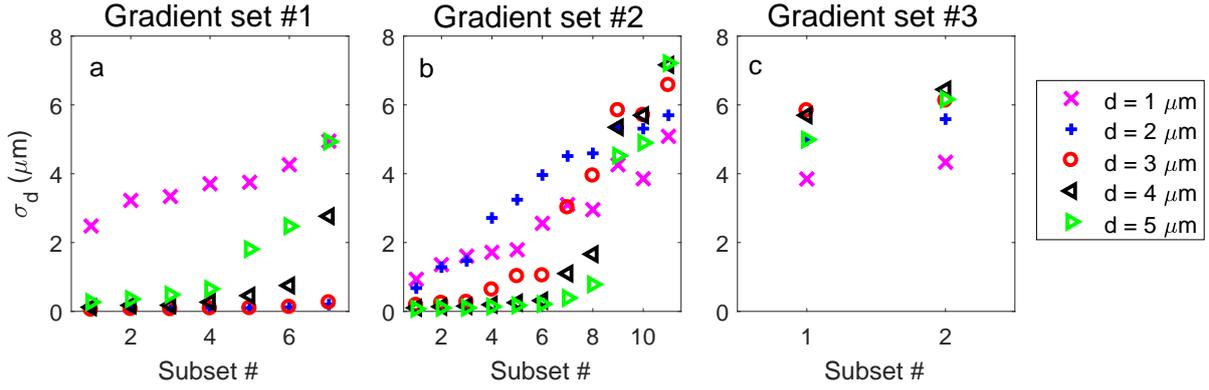
Figure 5.7: Intra-axonal simulations: σ_d for different gradient subsets. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ T/m. All frequencies in each set were kept when fitting to the model. Note that for G_{max} of 900 mT/m for diameters of 2 μm and 3 μm the σ_d does not depend much on the gradient subset. As the highest gradient strength used in the subset decreases, the σ_d for diameters of 4 μm and 5 μm increase. The σ_d for diameter of 1 μm is larger than for all other diameters and also increases when the highest gradient strength in the subset decreases. For G_{max} of 300 mT/m σ_d for all diameters increased as the highest gradient strength in the subset decreased. For G_{max} of 80 mT/m σ_d was larger for all gradient subsets than the other values of G_{max} and changing from using all gradients to two gradients caused an increase in σ_d.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.6c shows μ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. The μ_d are approximately 1 to 3 μm higher than the true diameter when using all the gradients. The μ_d are larger in the two gradient subset.

Figure 5.7c shows σ_d for $d = 1$ to 5 μm, using the $G_{max} = 80$ mT/m set of gradients. The σ_d are 0.5 to 1 μm larger in the two gradient subset.

5.3.1.2 Frequency Removal

Because the fitted values tended to be more precise when using both higher and more gradients, the effect of varying the number of frequencies was studied using all of the gradients for each of the three gradient sets.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.8a shows μ_d as a function of n_f for $d = 1$ to 5 μm. For $d > 1$ μm, μ_d show little change over n_f. With $d = 1$ μm, μ_d shows some increase towards lower n_f. The μ_d for $d = 1$ μm at

$n_f = 10$ is higher than that of $d = 2$ μm.

Figure 5.9a shows σ_d as a function of n_f for $d = 1$ to 5 μm. For $d > 2$ μm, σ_d show little change over n_f (all σ_d are under 0.4 μm). For $d \leqslant 2$ μm, σ_d for $n_f = 5$ are worse at small diameters as compared to those with more frequencies. The 1 μm diameter cylinder has relatively large σ_d for all frequencies, though it decreases at $n_f = 15$ and $n_f = 20$ (dropping by approximately 1 μm). The large σ_d in comparison to the other diameters was found to be due to outliers in its parameter distribution. For $d = 2$ μm, σ_d drops from 1 μm at $n_f = 5$ to 0.1 μm at $n_f = 10$ before leveling off. For $n_f \geqslant 10$, its σ_d is similar in magnitude to those of $d > 2$ μm.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.8b shows μ_d as a function of n_f for $d = 1$ to 5 μm. The μ_d show little change over n_f.

Figure 5.9b shows σ_d as a function of n_f for $d = 1$ to 5 μm. The σ_d for $d > 2$ μm remain under 0.4 μm over all n_f. These σ_d are similar in magnitude to those of the first gradient set. The 1 μm diameter cylinder has a larger σ_d than the others over all n_f. The 2 μm diameter cylinder has a larger σ_d at $n_f = 5$, which levels off at $n_f = 10$ (σ_d drops from 1.4 μm to ≈ 0.5 μm). Unlike the first gradient set, the σ_d for $d = 2$ μm remain higher than those for $d > 2$ μm even at higher n_f.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.8c shows μ_d as a function of n_f for $d = 1$ to 5 μm. For $d > 1$ μm, μ_d show little change over n_f. They are also all higher than the true diameters. For $d = 1$ μm, μ_d shows some variability over n_f.

Figure 5.9c shows σ_d as a function of n_f for $d = 1$ to 5 μm. There does not appear to be a clear trend in σ_d over n_f. For $d > 1$ μm, they change less than 5 percent over n_f, but the $d = 1$ μm σ_d declines by ≈ 25 percent). Each σ_d is much larger than the first or second gradient set ($\sigma_d > 4$ μm vs $\sigma_d \leqslant 1$ μm with the second gradient set).
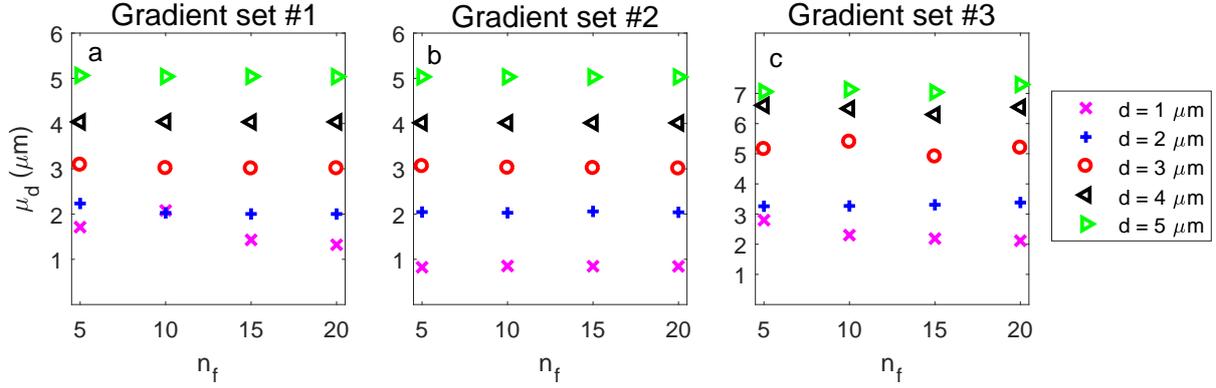
Figure 5.8: Intra-axonal simulations: μ_d as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For diameters larger than 1 μm, simulations at both G_{max} of 900 mT/m and 300 mT/m produce accurate μ_d values for trials with any number of frequencies, n_f, used. The values of μ_d at G_{max} of 80 mT/m are all biased upward by 1 to 2 μm. **Note**: (c) uses different y-axis limits.
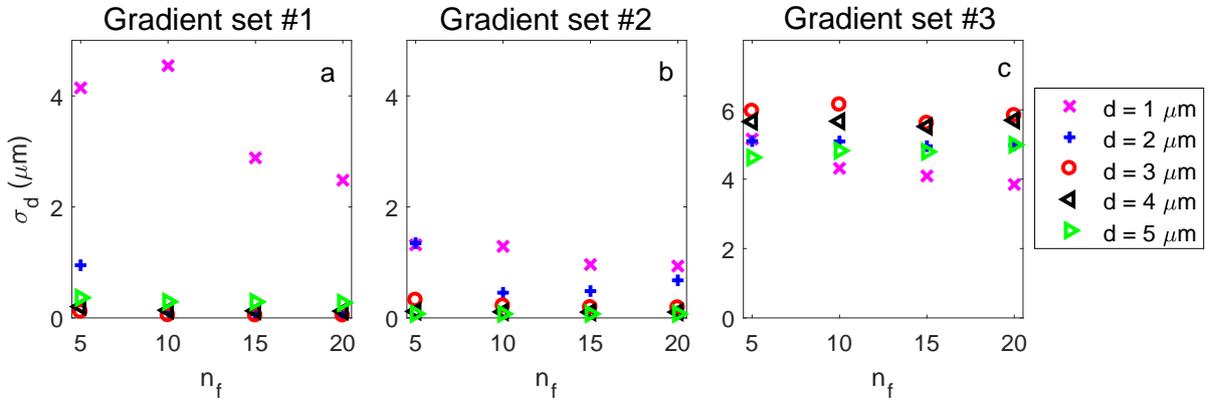
Figure 5.9: Intra-axonal simulations: σ_d as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For diameters larger than 2 μm, simulations at both G_{max} of 900 mT/m and 300 mT/m produce small σ_d values for trials with any number of frequencies, n_f, used. For 2 μm cylinders, the simulations with G_{max} of 900 mT/m produced results such that σ_d values were small for $n_f > 5$ and relatively small for $n_f = 5$. The simulations for 300 mT/m produced slightly larger σ_d values for 2 μm cylinders than the larger diameters for $n_f > 5$. Simulations of cylinders of diameters of 1 μm produced the largest values of σ_d at both G_{max} of 900 mT/m and 300 mT/m. The values of σ_d at G_{max} of 80 mT/m were all large. **Note**: (c) uses different y-axis limits.

5.3.2 Identical Cylinders (Intra- and Extra-axonal water) – Unequal Diffusivities

This section discusses the case of identical cylinders on a square lattice with $D_{in} \neq D_{ex}$. We first give an overview of the fitted parameters (packing fraction, d, and D_{in}) when using all the measurements.

Gradient set #1 (G_{max} = 900 mT/m)

Figures 5.10a, 5.10d, and 5.10g show μ_d for d = 1 to 5 μm. As d goes from 1 μm to 4 μm, the μ_d become closer to the true values. For $d > 1$ μm, all μ_d are within 10 percent of the true values. At $d = 5$ μm, μ_d is slightly worse than at $d = 4$ μm. This is probably because the extra-axonal model is less accurate here (Eq. 4.10).

Accuracy sometimes depends on packing fraction. The relative difference between μ_d and actual diameter for $d = 1$ μm decreases at higher packing fractions, dropping from 47 percent at packing fraction = 0.5 to around 15 percent at packing fraction = 0.7. The error also decreases with packing fraction, dropping from 2 μm to 0.7 μm. There does not appear to be as visible of a change with the other diameters.

We should also look at some of the distributions of fitted diameters, like we did for the intra-axonal model. Figures 5.11a and 5.11b show the distributions of fitted diameters for $d = 1$ μm and $d = 3$ μm (packing fraction = 0.6). For $d = 1$ μm, some fitted diameters have clustered near the upper bound, skewing the mean and standard deviation of the distribution ($2/1000 > 19$ μm). The peak of the distribution is around 1.3 μm.

Figures 5.12a, 5.12d, and 5.12g show μ_D as a function of diameter. The μ_D are 20 to 40 percent higher than the actual value (1.0 μm^2/ms). Error decreases at larger diameters. For example, with a packing fraction of 0.7, σ_D drops from around 1 μm^2/ms at $d = 1$ μm to 0.4 μm^2/ms at $d = 5$ μm.

In the intra-axonal simulations, D_{in} was the worst parameter to fit, so we should expect the same here. Figures 5.11c and 5.11d show the fitted D_{in} distributions for $d = 1$ μm and $d = 3$ μm (packing fraction = 0.6). For $d = 3$ μm, there are ($118/1000 > 2.4$ μm^2/ms) fitted D_{in} that have clustered at the upper bound. The rest of the fitted D_{in} are gathered around
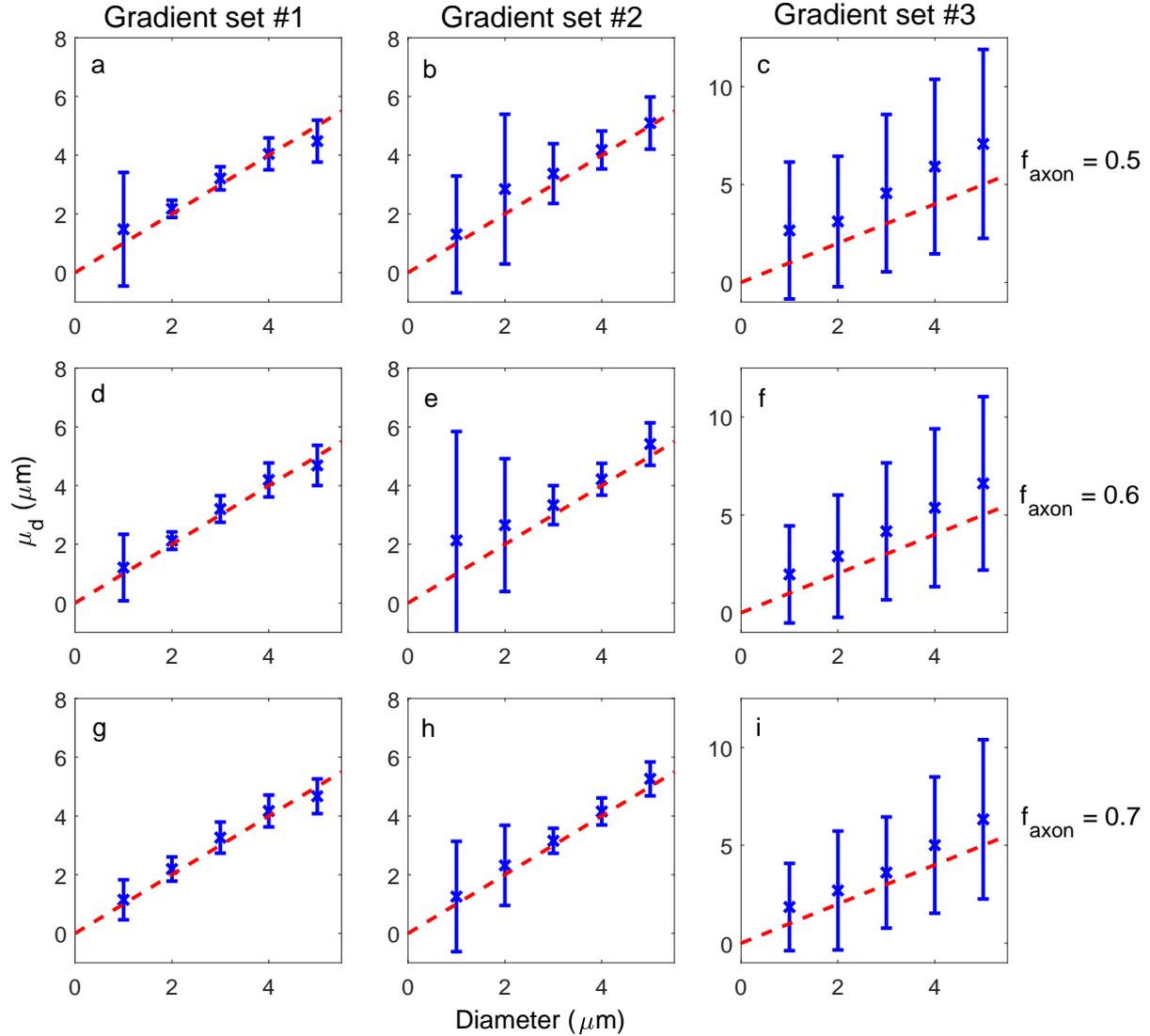
Figure 5.10: Full simulation: μ_d as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted diameters (σ_d). Dashed line indicates a perfect fit. Three sets of measurements are shown: (a-c) $G_{max} = 900$ mT/m (d-f) $G_{max} = 300$ mT/m (g-f) $G_{max} = 80$ mT/m. (a, d, g) are from simulations with packing fraction 0.5. (b, e, h) are from simulations with packing fraction 0.6. (c, f, i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m and $G_{max} = 300$ mT/m, μ_d are less accurate at smaller diameters. For $G_{max} = 80$ mT/m, μ_d are higher than the actual diameters. **Note**: (c), (f), and (i) use different y-axis limits.
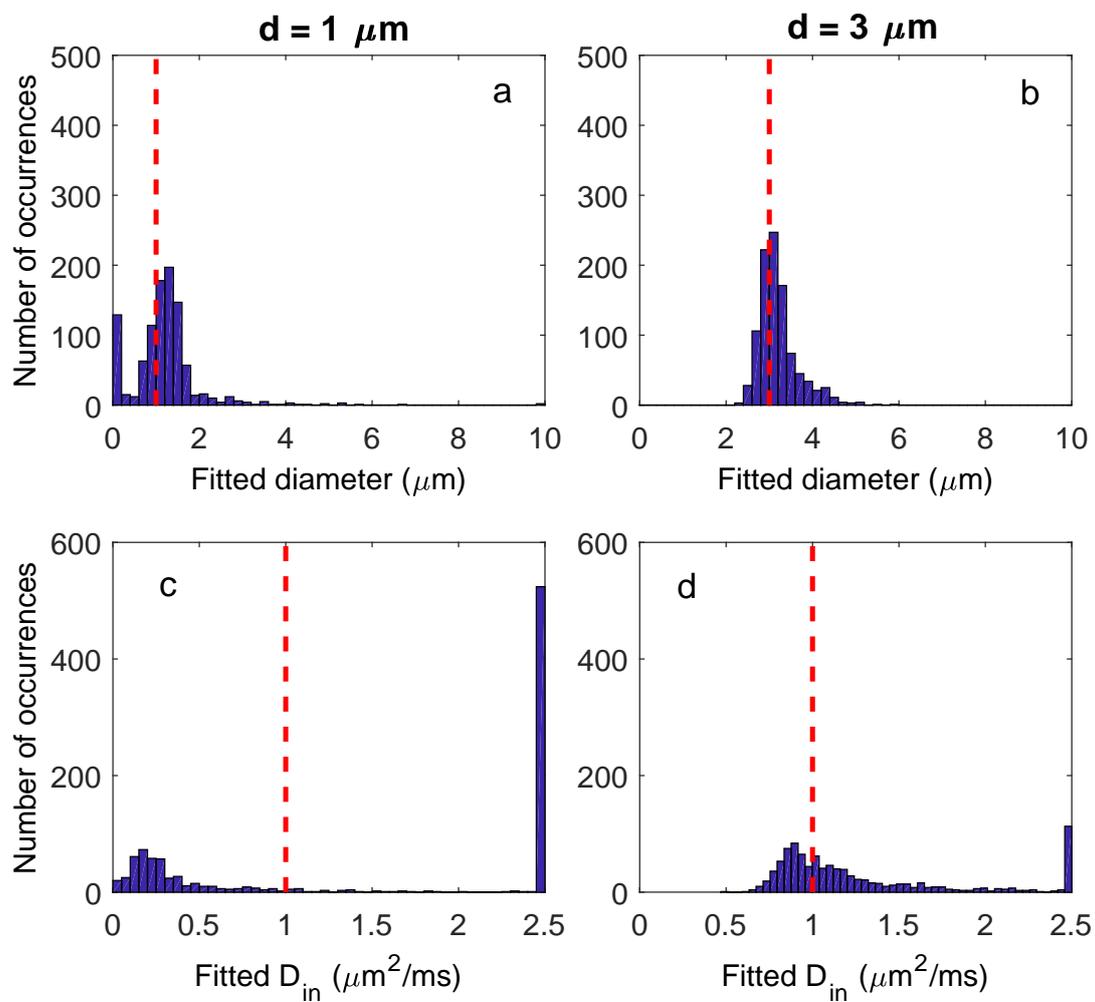
Figure 5.11: Full simulation: Distribution of fitted cylinder diameters for (a) $d = 1$ μm and (b) $d = 3$ μm. Distribution of fitted intra-axonal diffusion coefficients for (c) $d = 1$ μm and (d) $d = 3$ μm. Simulations use the $G_{max} = 900$ mT/m gradient set. All measurements were kept when fitting to the model. The fitted D_{in} for $d = 1$ μm are inaccurate. Both distributions are narrower with $d = 3$ μm.
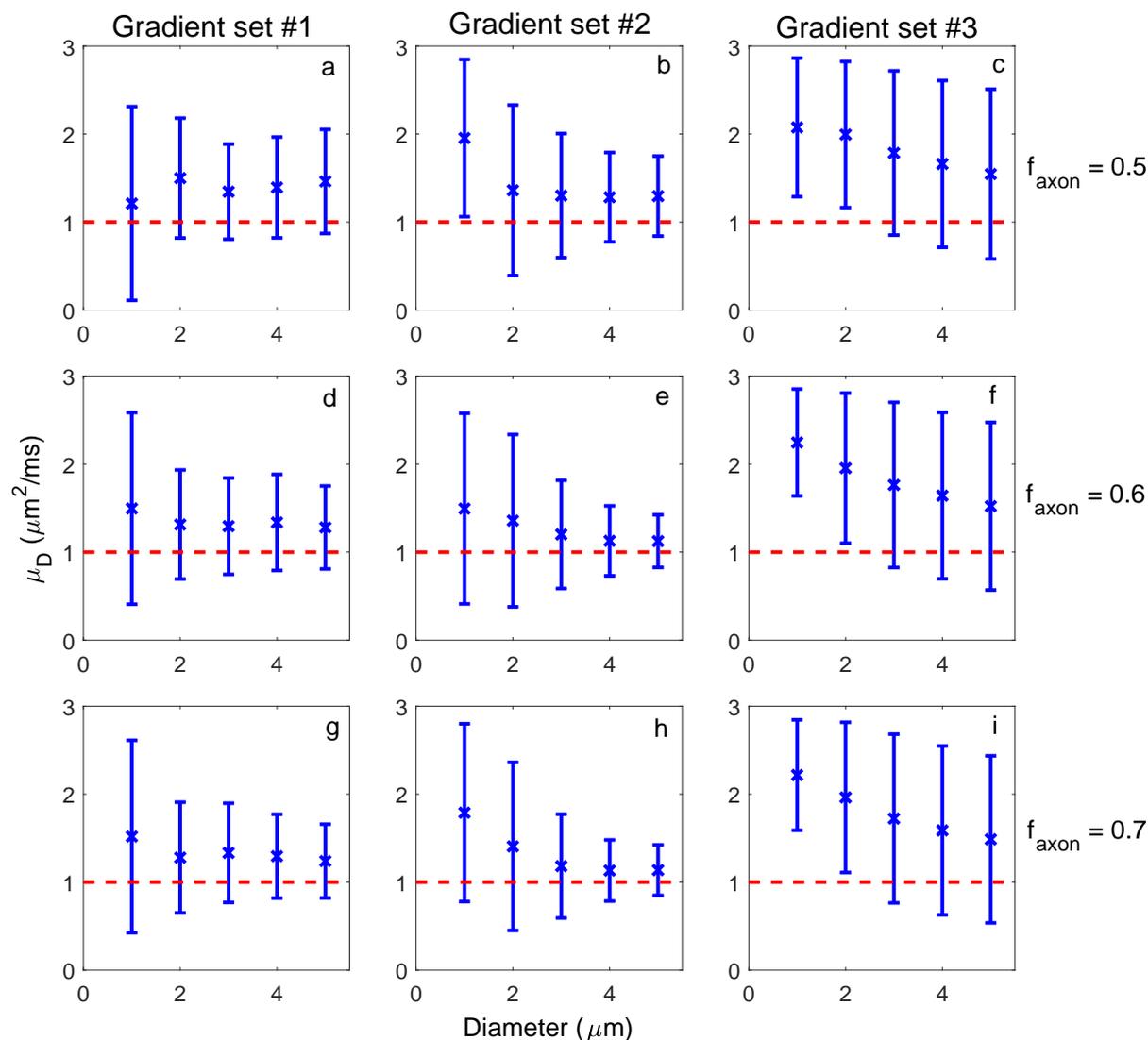
Figure 5.12: Full simulation: μ_D as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted intra-axonal diffusion coefficients (σ_D). Dashed line indicates a perfect fit. Three sets of measurements shown: (a-c) $G_{max} = 900$ mT/m (d-f) $G_{max} = 300$ mT/m (g-i) $G_{max} = 80$ mT/m. (a, d, g) are from simulations with packing fraction 0.5. (b, e, h) are from simulations with packing fraction 0.6. (c, f, i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. The μ_D tend to be more accurate at larger diameters.

1 μm^2/ms. For $d = 1$ μm, most of the fitted D_{in} are clustered near the lower and upper bounds (525/1000 > 2.4 μm^2/ms). The peak of the distribution is around 0.2 μm^2/ms.

Mean fitted packing fractions are shown in Figures 5.13a, 5.13d, and 5.13g. All mean fitted values are within 10 percent of the actual values.

Gradient set #2 (G_{max} = 300 mT/m)

Figures 5.10b, 5.10e, and 5.10h show μ_d for $d = 1$ to 5 μm. As diameters get larger, μ_d become closer to the actual values and σ_ds get smaller, at least up to $d = 5$ μm. Like in the last section, fitting for $d = 5$ μm may be affected by poor modeling (Eq. 4.10).

The σ_d for $d = 1$ μm and $d = 3$ μm look quite large, so Figures 5.14a and 5.14b compare the distributions of fitted diameters for $d = 1$ μm and $d = 3$ μm (packing fraction = 0.6). For $d = 1$ μm, some fitted diameters have clustered near the upper bound (33/1000 > 19 μm). There is also an upward bias in the distribution. The main peak is around 1.7 μm. For $d = 3$ μm, the peak is shifted upwards to around 3.3 μm.

Figure 5.12b, 5.12e, and 5.12h show μ_D as a function of diameter. Error decreases with larger diameters. For $d = 4$ μm and $d = 5$ μm, μ_D are around 10 to 13 percent higher than the true value. For $d = 1$ μm, μ_D are 50 to 100 percent larger than the true value.

Mean fitted packing fractions are shown in Figures 5.13b, 5.13e, and 5.13h. All mean values are within 10 percent of the true values.

Gradient set #3 (G_{max} = 80 mT/m)

Figures 5.10c, 5.10f, and 5.10i show μ_d for $d = 1$ to 5 μm. All μ_d are higher than the true diameters, which is what we saw with the intra-axonal simulations. Differences between μ_d and true diameters tend to decrease at larger diameters. For a packing fraction = 0.6, the difference drops from 100 percent at $d = 1$ μm to 30 percent at $d = 5$ μm. The relative difference also seems to be smaller at higher packing fractions. For example, using $d = 3$ μm, relative difference drops from 50 percent with a packing fraction of 0.5 to 20 percent with a packing fraction of 0.7.

Figures 5.12c, 5.12f, and 5.12i shows μ_D as a function of diameter. At smaller diameters, μ_D
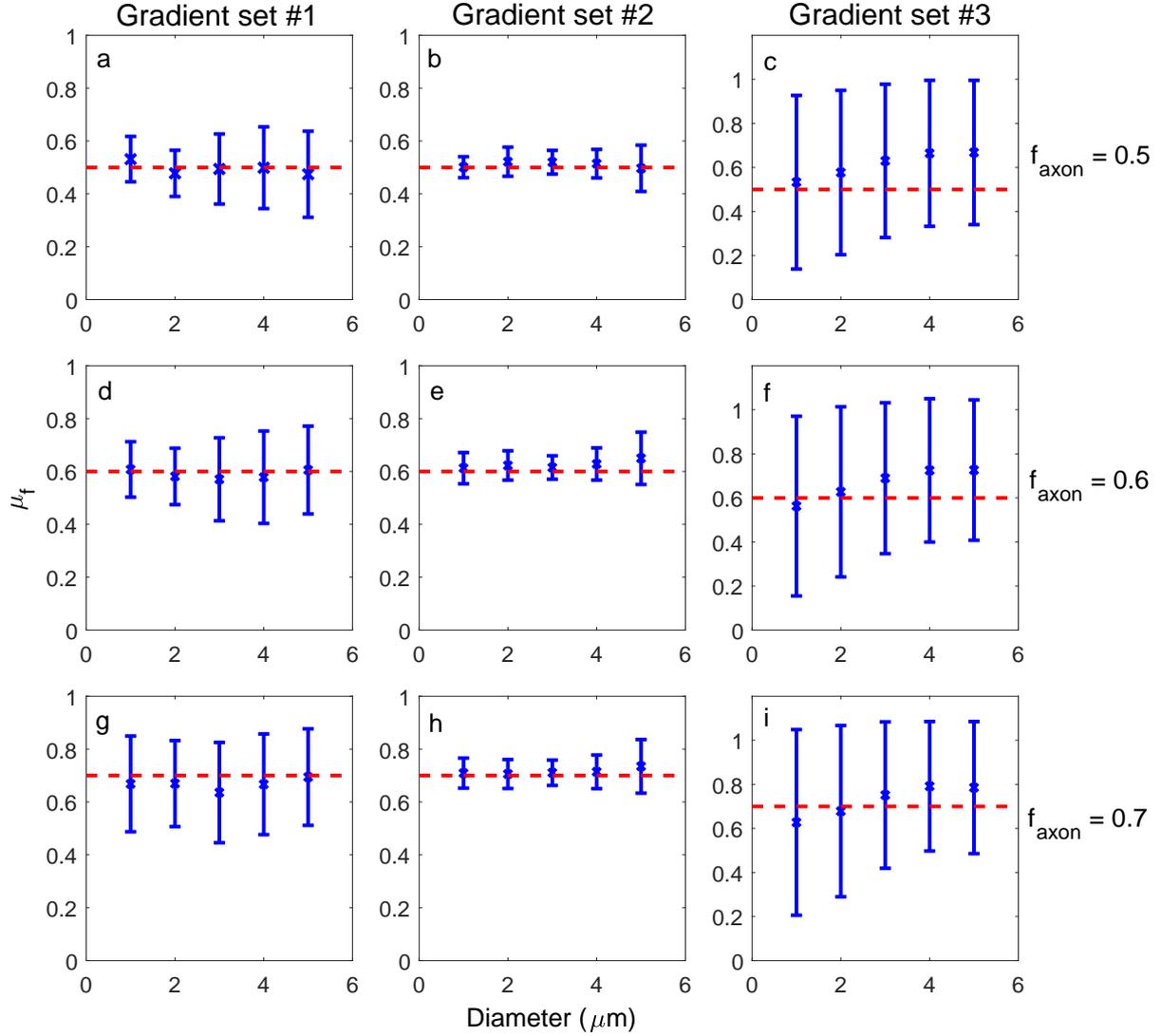
Figure 5.13: Full simulation: μ_f as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted packing fractions (σ_p). Dashed line indicates a perfect fit. Three sets of measurements shown: (a-c) $G_{max} = 900$ mT/m (d-f) $G_{max} = 300$ mT/m (g-i) $G_{max} = 80$ mT/m. (a, d, g) are from simulations with packing fraction 0.5. (b, e, h) are from simulations with packing fraction 0.6. (c, f, i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m and $G_{max} = 300$ mT/m, μ_f are accurate. Error bars are very large for $G_{max} = 80$ mT/m (see Figure 5.15). **Note**: (c), (f), and (i) use different y-axis limits.
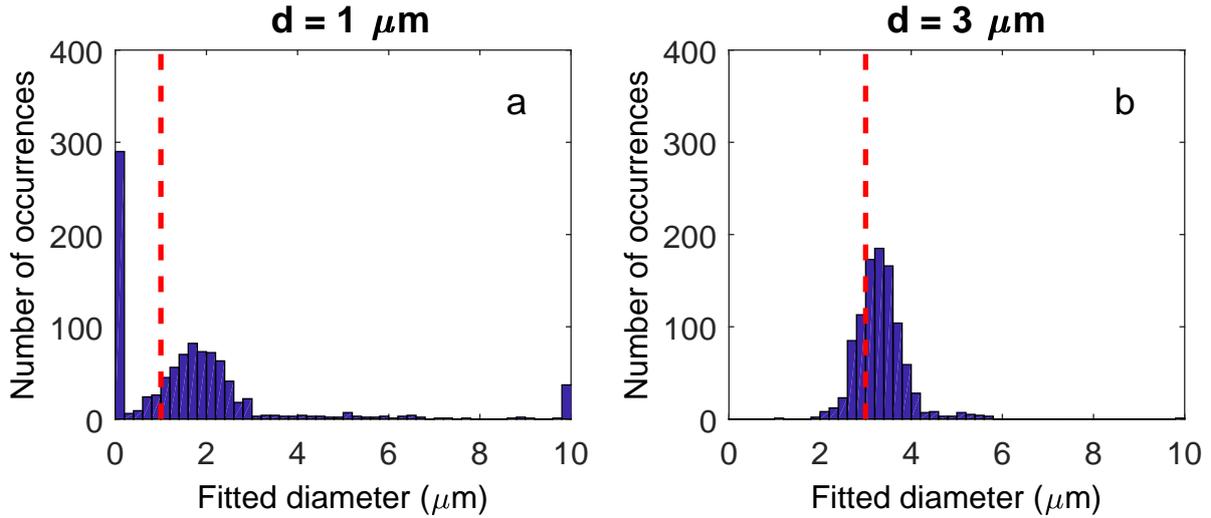
Figure 5.14: Full simulation: Distribution of fitted cylinder diameters for (a) $d = 1$ μm and (b) $d = 3$ μm. Simulations use the $G_{max} = 300$ mT/m gradient set. All measurements in each set were kept when fitting to the model. At $d = 1$ μm, the distribution is shifted upwards by ≈ 1 μm. At $d = 3$ μm, the distribution is narrower, but still shifted upwards by ≈ 0.5 μm.

are higher than the true value. The difference drops from approximately 100 percent to around 50 percent as d goes from 1 μm to 5 μm. Based on what we saw in the intra-axonal simulation, we should be cautious when interpreting these numbers, as the distributions themselves are quite wide.

Mean fitted packing fractions are shown in Figures 5.13c, 5.13f, and 5.13i. Unlike other gradient sets, mean packing fractions are inaccurate. Figure 5.15 shows the distribution of fitted packing fractions for $d = 3$ μm (packing fraction = 0.6). The fitted values are spread out with many clustered near the lower and upper bounds. Packing fraction, like the intra-axonal diffusion coefficient, is a parameter that cannot be fitted accurately with gradients this small, at least for cosine-OGSE.

5.3.2.1 Gradient Subsets

First changes in μ_d and σ_d will be studied when gradients are successively removed from the data. For these simulations, all frequencies ($n_f = 20$) are used and the packing fraction for the cylinders was chosen to be 0.6.
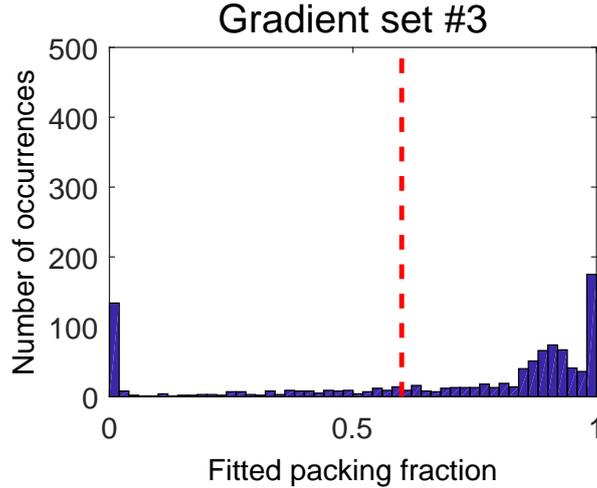
Figure 5.15: Full simulation: Distribution of fitted packing fractions for $d = 3$ μm (packing fraction of 0.6). Simulations use the $G_{max} = 80$ mT/m gradient set. All measurements were kept when fitting to the model. Fitted packing fractions tend to be inaccurate.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.16a shows μ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. Except for $d = 5$ μm, all are initially higher (< 20 percent) than the true diameters. The μ_d for $d = 5$ μm is smaller ($\mu_d \approx 4.7$ μm), possibly because the extra-axonal term (Eq. 4.10) is not accurate. There is an overall increase in μ_d as we move to smaller gradients, but the μ_d for $d = 4$ μm and 5 μm drop at the highest subset.

Figure 5.17a shows σ_d for $d = 1$ to 5 μm. There is a steady increase in σ_d as we move through the subsets and smaller gradients are used. For some diameters, σ_d increases at a steady rate, other diameters have a point where the rate of increase becomes larger. For example, σ_d for $d = 4$ μm and $d = 5$ μm increase at a faster rate starting at the 4th or 5th subset, while the rate of σ_d for $d = 3$ μm increases faster around the 7th. Compared to the intra-axonal model, fitted parameters here tend to be less precise (σ_d is approximately 0.3 to 1.5 μm higher here). The exception is $d = 1$ μm, which is likely because of outliers in its distribution, like Figure 5.2a.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.16b shows μ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. All μ_d are initially higher than the true diameters, with the biggest difference at $d = 1$ μm. The μ_d increase
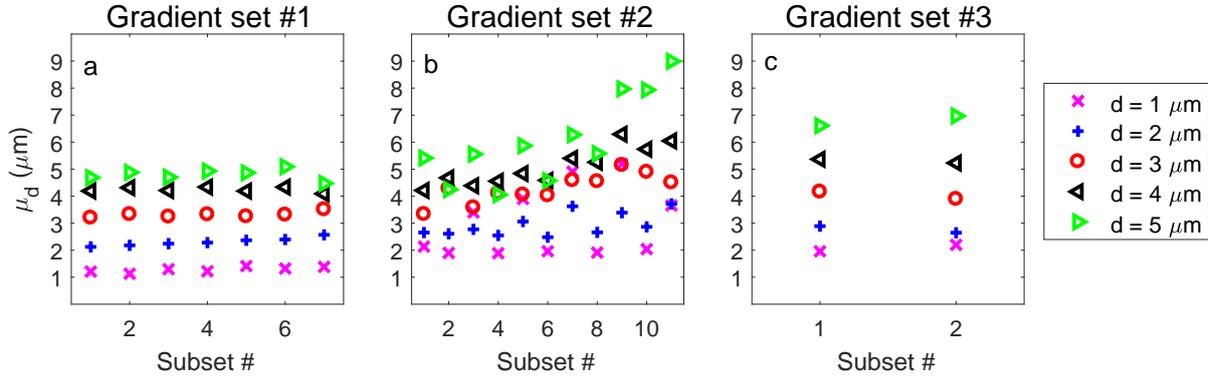
Figure 5.16: Full simulation: μ_d for different gradient subsets. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m and $G_{max} = 300$ mT/m, there is an overall increase in μ_d as we move to smaller gradients. There is less variation with $G_{max} = 900$ mT/m.
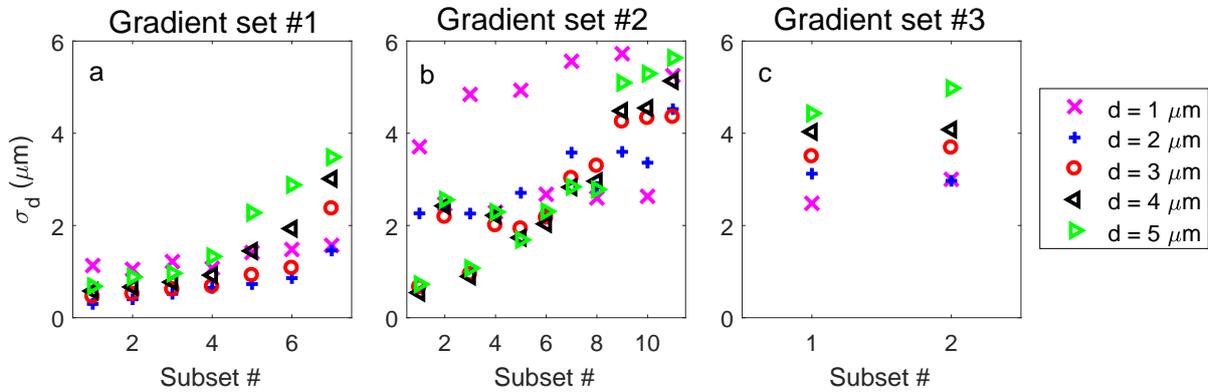
Figure 5.17: Full simulation: σ_d for different gradient subsets. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m, σ_d increases as the maximum gradient strength in the subset decreases. A consistent trend in σ_d with gradient subset is not seen for $G_{max} = 300$ mT/m. There appears to be a discontinuity at $G_{high} = 100$ mT/m where errors jump, like what was seen with the intra-axonal model (Figure 5.7b). This is also where the μ_d become more inaccurate due to small gradients. For the simulations with $G_{max} = 80$ mT/m, with exception of $d = 2$ µm, σ_d are slightly larger for the two gradient subset.

moving to higher subsets, but not smoothly. The subsets with multiple gradients tend to have higher mean fitted diameters. The subsets with two gradients tend to have smaller mean fitted diameters. This is mainly due to more fitted diameters at $d = 0$ μm.

Figure 5.17b shows σ_d for $d = 1$ to 5 μm. The plot shows most of the same features as in the intra-axonal model. However, this time it is not necessarily the case that increasing G_{high} alone leads to better results. Although there is an overall increase in σ_d as the subset number increases (as gradients become smaller), there is not the same smooth increase as before. There appears to be a discontinuity at $G_{high} = 100$ mT/m where errors jump, like what was seen with the intra-axonal model (Figure 5.7b). This is also where the μ_d become more inaccurate due to small gradients.

Using the largest gradients (the smaller subsets), the σ_d are larger compared to the intra-axonal model. However, there is less of an effect when moving to smaller gradients, at least for some diameters ($d = 2$ μm, $d = 3$ μm) as compared to the intra-axonal model. For example, in the intra-axonal model, errors for $d = 2$ μm rose by approximately 5 μm, while here they rise 2 μm. For $d > 2$ μm, the results for the first gradient set and the second gradient set (with $G_{high} = 250$ mT/m or 300 mT/m) are similar, at least for subsets with multiple gradients (subsets 1 and 3). With one nonzero gradient, this is no longer true (compare subsets 2 and 4 here with those in Figure 5.17a).

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.16c shows μ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. The μ_d are approximately 1 to 2 μm higher than the true diameter when using all the gradients. The μ_d in the second subset are within 12 percent of those in the first subset.

Figure 5.17c shows σ_d for $d = 1$ to 5 μm. With the exception of $d = 2$ μm, where errors drop by approximately 0.1 μm, σ_d are slightly larger in the two gradient subset.

5.3.2.2 Frequency Removal

The results presented here are for simulations with packing fractions of 0.6 that use all the gradients and fewer frequencies.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.18a shows μ_d as a function of n_f for $d = 1$ to 5 µm. The μ_d for $n_f = 15$ and $n_f = 20$ are very similar. However, μ_d begins to increase at $n_f = 10$ and $n_f = 5$.

Figure 5.19a shows σ_d as a function of n_f for $d = 1$ to 5 µm. The uncertainty for $d = 1$ µm is high at smaller n_f but levels off after $n_f = 15$, dropping from around 6 µm to 4 µm). For $d = 2$ µm to $d = 5$ µm, σ_d at the $n_f = 5$ are larger than those with $n_f > 5$ (by approximately 1 µm). After leveling off, σ_d at $n_f = 10$, $n_f = 15$, and $n_f = 20$ are all very similar to each other (within 0.5 µm of each other).

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.18b shows μ_d as a function of n_f for $d = 1$ to 5 µm. For $d > 2$ µm, there is a small increase in μ_d at lower n_f. For $d = 1$ and 2 µm, μ_d increases at smaller n_f, but there is some fluctuation. For example, at $n_f = 10$ and $n_f = 15$, μ_d for $d = 1$ µm are higher than the μ_d for $d = 2$ µm.

Figure 5.19b shows σ_d as a function of n_f for $d = 1$ to 5 µm. For $d = 3$, 4, and 5 µm, σ_d show little change over n_f (changes are < 0.1 µm). The $d = 1$ µm and $d = 2$ µm errors become smaller as n_f increases, with those of $d = 2$ µm dropping by 2 µm as n_f goes from 5 to 10. Unlike the first gradient set, errors for $d = 2$ µm remain around 1 to 2 µm higher than those of $d > 2$ µm when $n_f \geqslant 10$ (compare Figure 5.19a to 5.19b). For $d > 2$ µm, σ_d are similar to those of first gradient set (within 0.2 µm).

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.18c shows μ_d as a function of n_f for $d = 1$ to 5 µm. For $d > 1$ µm, there is little overall change in μ_d ($\leqslant 2$ µm above the true diameter). The $d = 1$ µm, $n_f = 5$ data point was excluded because even fitting without noise gave a result of zero.

Figure 5.19c shows σ_d as a function of n_f for $d = 1$ to 5 µm. The σ_d for a given diameter remain within 20 percent of each. There does not seem to be a downward trend over n_f.
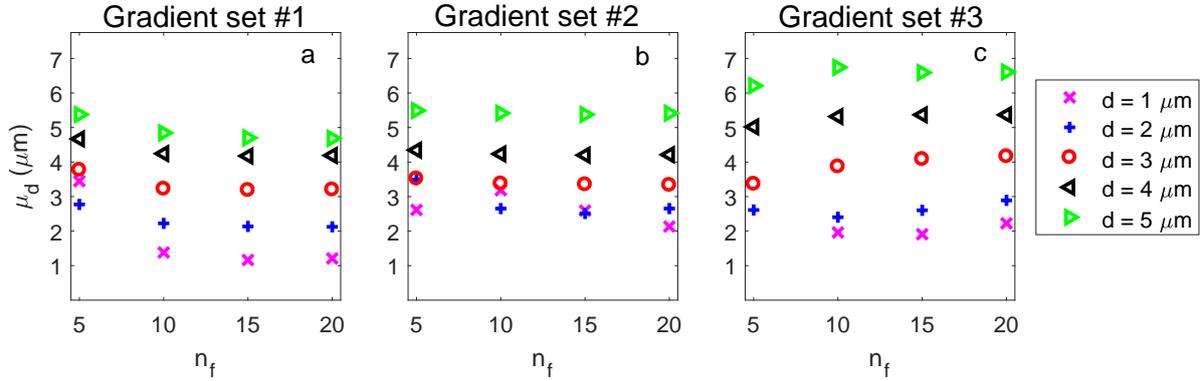
Figure 5.18: Full simulation: μ_d as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m, μ_d are highest at $n_f = 5$. For $G_{max} = 300$ mT/m and $G_{max} = 80$ mT/m, μ_d are larger than the actual diameters. With $d = 3$ to 5 μm, there is little change over n_f.
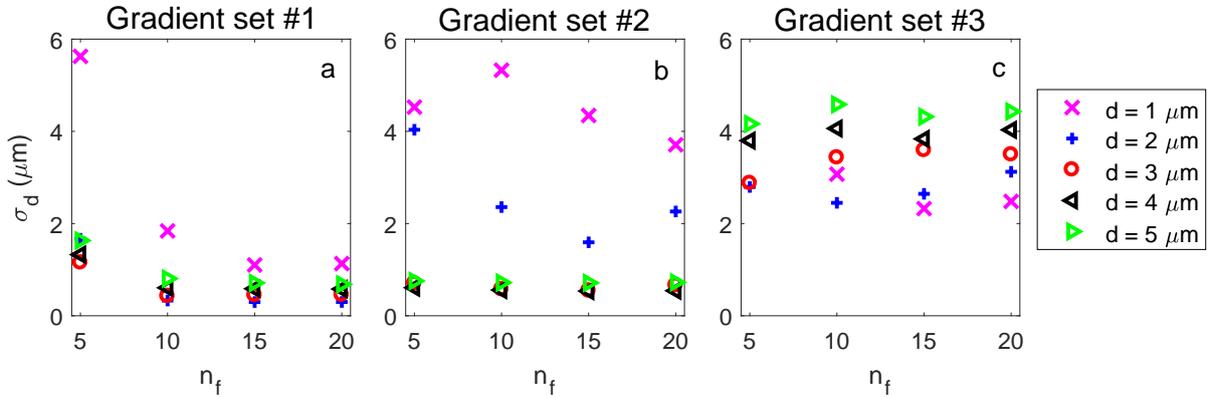
Figure 5.19: Full simulation: σ_d as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For simulations with $G_{max} = 900$ mT/m, the uncertainty for $d = 1$ μm is large at smaller n_f but levels off after $n_f = 15$. For $d = 2$ μm to $d = 5$ μm, σ_d at $n_f = 5$ are larger than those for $n_f > 5$. For simulations with $G_{max} = 300$ mT/m, σ_d for $d = 3$, 4, and 5 μm show little change over n_f. The $d = 1$ μm and $d = 2$ μm errors become smaller as n_f increases. For $d > 2$ μm, σ_d are similar to those of first gradient set. For simulations with $G_{max} = 80$ mT/m, σ_d are mostly independent of n_f.

5.3.3 Diameter Distributions – Unequal Diffusivities

This section presents results from simulations containing cylinders of different diameters.

5.3.3.1 Gradient Subsets

The results presented first are for simulations using all the frequencies ($n_f = 20$) with different gradient subsets.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.20a shows μ_{AxD} using the gradient subsets in Table 5.1. The μ_{AxD} for the two smallest AxD (2.1 µm and 2.3 µm) are larger than the true AxD, while the μ_{AxD} for the three highest AxD are smaller than the actual AxD. The highest AxD are underestimated the most ($AxD = 4.8$ µm, 6.1 µm). This may be due to poor modeling of the extra-axonal term (Eq. 4.10). The μ_{AxD} for all AxDs show little change over the subsets.

Figure 5.21a shows σ_{AxD} for the gradient subsets in Table 5.1. There is a steady increase in σ_{AxD} as we move through the subsets and use smaller gradients. While some σ_{AxD} increase at a steady rate, others have a point where the rate of increase becomes larger. For example, σ_{AxD} for $AxD = 4.8$ µm and $AxD = 6.1$ µm increase at a faster rate starting at the 5th subset, while the rate of that of $AxD = 3$ µm increases at the 7th (increasing by approximately 0.6 µm). The σ_{AxD} for $AxD = 2.1$ µm and $AxD = 2.3$ µm experience the slowest change over the subsets.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.20b shows μ_{AxD} using the gradient subsets in Table 5.1. There is not much of an overall trend over the subsets, but the μ_{AxD} do oscillate. The subsets with multiple gradients tend to have higher fitted AxD. The subsets with two gradients tend to have lower fitted AxD. This was partly due to more outliers at $AxD = 0$ µm in the parameter distribution.

Figure 5.21b shows σ_{AxD} for the gradient subsets in Table 5.1. There is an overall increase in σ_{AxD} as the gradients become smaller. Some oscillation in σ_{AxD} occurs as we move through the subsets. For $AxD = 6.1$ µm, there seems to be a discontinuity at $G_{high} = 100$ mT/m where σ_{AxD} jump, like what happened in Figures 5.7b and 5.17b. However, unlike the previous cases,

there was no corresponding change in the μ_{AxD}. In both the first and second gradient set, σ_{AxD} for all the *AxD* start off with similar values. All σ_{AxD} are initially smaller than 1 μm. However, as we move through the subsets, the differences in σ_{AxD} for the two gradient sets separate, so that those of the first set tend to remain smaller than those of the second set.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.20c shows μ_{AxD} using the gradient subsets in Table 5.1. The μ_{AxD} for the three smallest *AxD* (2.1 μm, 2.3 μm, 2.9 μm) are larger than the true *AxD*, while the μ_{AxD} for the two highest *AxD* are smaller than the actual *AxD*. In the second gradient subset, μ_{AxD} drops a small amount (< 0.5 μm).

Figure 5.21c shows σ_{AxD} for the gradient subsets in Table 5.1. For $AxD = 2.1$ μm, 2.3 μm, and 2.9 μm, σ_{AxD} are 0.1 μm smaller with two gradients. For $AxD = 4.8$ μm and 6.1 μm, respective σ_{AxD} are 0.5 μm and 1.5 μm larger with one nonzero gradient. The small decrease in the smallest *AxD*s might just be due to noise from the 1000 trials.

5.3.3.2 Frequency Removal

Results from simulations using all the gradients and varying number of frequencies are presented here.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.22a shows μ_{AxD} as a function of n_f for the five *AxD*. There is some increase in μ_{AxD} at lower n_f.

Figure 5.23a shows σ_{AxD} as a function of n_f for five *AxD*. The σ_{AxD} are highest at $n_f = 5$ and decrease with n_f. After $n_f = 10$, they begin to level off. They decrease by less than 0.2 μm between $n_f = 10$ and $n_f = 20$. The $AxD = 6.1$ μm distribution has the largest errors (0.25 to 0.4 μm greater than the rest). This is probably because the extra-axonal model does not model distributions with larger diameters (or *AxD*) as well. These results cannot be compared directly to the single diameter case because the axons are not necessarily the same size, but it is still interesting to try. To use one example, comparing $AxD = 2.1$ μm to $d = 2$ μm, at $n_f > 5$,
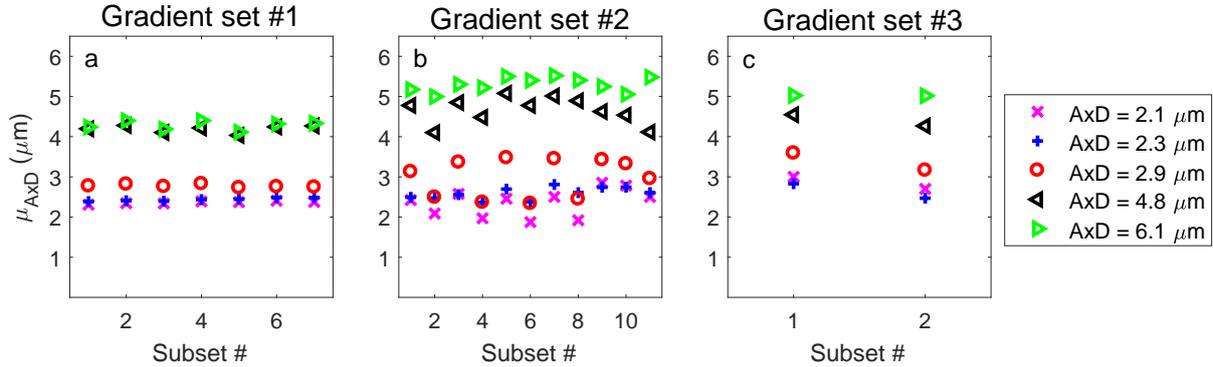
Figure 5.20: Full simulation: μ_{AxD} for different gradient subsets. (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m, μ_{AxD} for 4.8 μm and 6.1 μm are smaller than the true *AxD*. Otherwise, there is little overall change. For $G_{max} = 300$ mT/m, there is some fluctuation in μ_{AxD} between even and odd subsets. For $G_{max} = 80$ mT/m, μ_{AxD} are smaller using the two gradient subset.
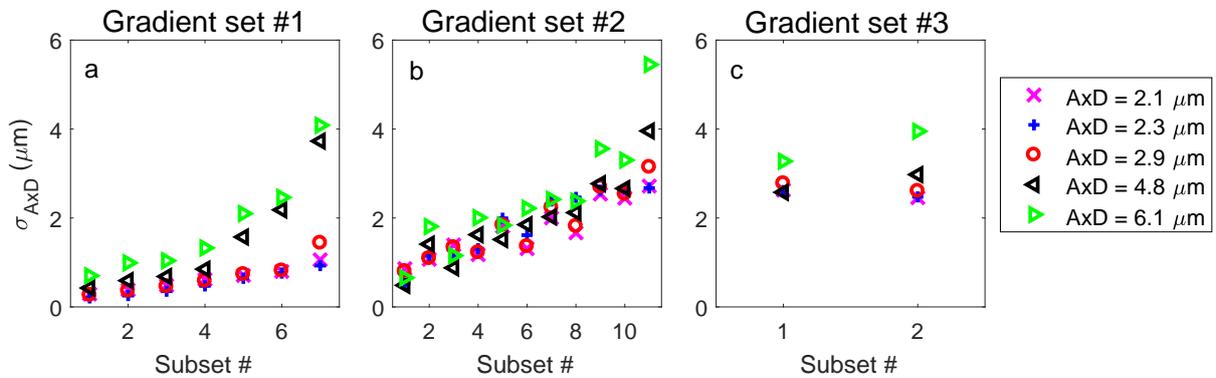
Figure 5.21: Full simulation: σ_{AxD} for different gradient subsets. (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For the simulations with $G_{max} = 900$ mT/m, there is an increase in σ_{AxD} as the highest gradient strength in a subset decreases. For distributions with smaller *AxD*, σ_{AxD} tends to increase at a slower rate. For the simulations with $G_{max} = 300$ mT/m, there is an increase in σ_{AxD} as the largest gradient in the subset decreases. In both (a) and (b), σ_{AxD} for all the *AxD* start off with similar values, but eventually separate, so that those of the first set tend to remain smaller than those of the second set. For the simulations with $G_{max} = 80$ mT/m, σ_{AxD} are similar in both subsets.

the σ_{AxD} (σ_d) are similar, with those of $AxD = 2.1$ μm slightly smaller. At $n_f = 5$, those of $AxD = 2.1$ μm are smaller, by ≈ 0.8 μm.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.22b shows μ_{AxD} as a function of n_f for the five AxD. The μ_{AxD} show little change over n_f.

Figure 5.23b shows σ_{AxD} as a function of n_f for the five AxD. The values of σ_{AxD} show little change over n_f, changing less than 0.2 μm. All σ_{AxD} are under 1 μm. Excluding $n_f = 5$, σ_{AxD} of $AxD = 2.1$ μm, 2.3 μm, and 2.9 μm are around 2 to 3 times larger than those of the first gradient set, while σ_{AxD} of $AxD = 4.8$ μm and $AxD = 6.1$ μm are approximately the same (within 10 to 20 percent). However, at $n_f = 5$, σ_{AxD} with the first gradient set are up to 2 times larger than σ_{AxD} here.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.22c shows μ_{AxD} as a function of n_f for the five AxD. The μ_{AxD} show little change over n_f.

Figure 5.23c shows σ_{AxD} as a function of n_f for the five AxD. The values of σ_{AxD} are obviously much larger than either the first or second gradient set. There does not appear to be a clear trend over n_f.

5.3.4 Identical Cylinders (Intra- and Extra-axonal water) – Equal Diffusivities

We now move on to the case of identical cylinders on a square lattice with $D_{in} = D_{ex}$. Since the results were similar to when $D_{in} \neq D_{ex}$, we will not give the overview of results of each parameter before beginning with removal of data. Only σ_d are presented because μ_d showed similar behavior compared to when $D_{in} \neq D_{ex}$.
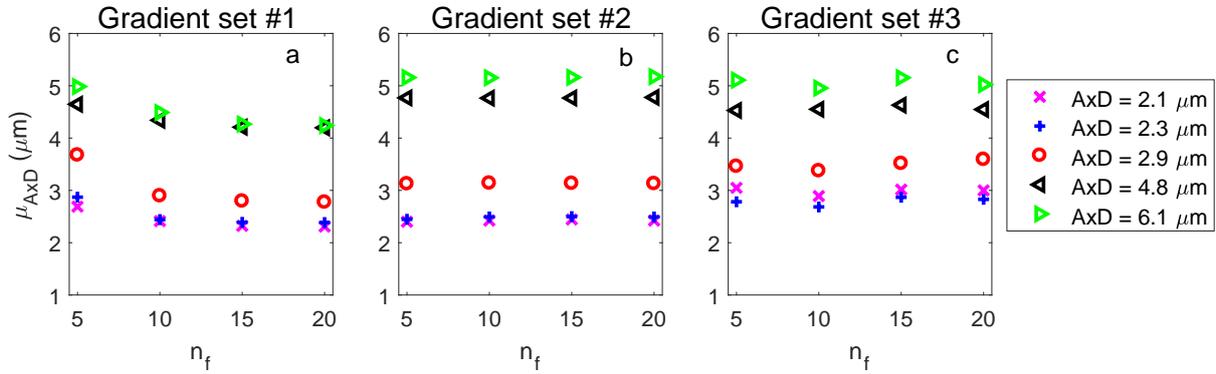
Figure 5.22: Full simulation: μ_{AxD} as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m, μ_{AxD} are highest at $n_f = 5$ and decrease with n_f, leveling off past $n_f = 10$. For $G_{max} = 300$ mT/m and $G_{max} = 80$ mT/m, there is little change in μ_{AxD}.
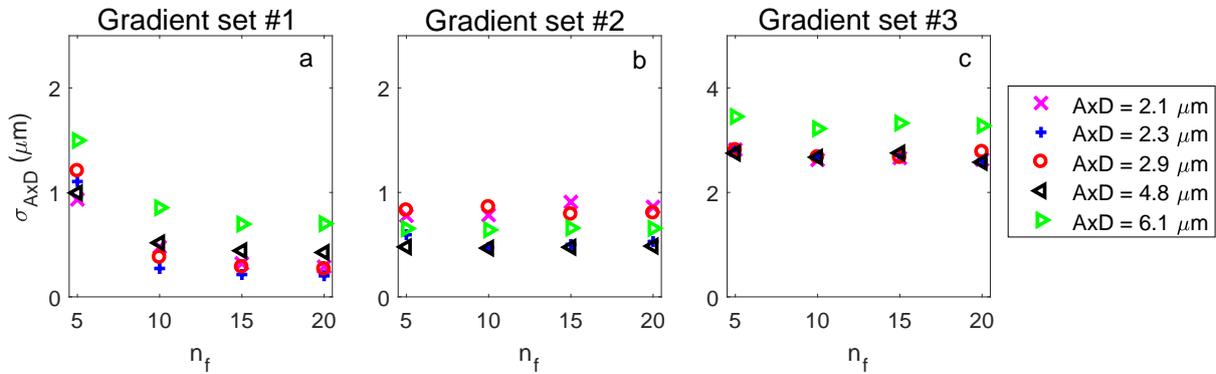
Figure 5.23: Full simulation: σ_{AxD} as a function of n_f. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For the simulations with $G_{max} = 900$ mT/m, the σ_{AxD} are highest at $n_f = 5$ and decrease with n_f. After $n_f = 10$, they begin to level off. For the simulations with $G_{max} = 300$ mT/m, the values of σ_{AxD} show little change over n_f. For the simulations with $G_{max} = 80$ mT/m, the values of σ_{AxD} are obviously much larger than either the first or second gradient set. There does not appear to be a clear trend over n_f. **Note**: (c) uses different y-axis limits.

5.3.4.1 Gradient Subsets

We begin by removing gradients. For these simulations, all frequencies ($n_f = 20$) are used and the packing fraction for the cylinders was chosen to be 0.6.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.24a shows σ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. There is a steady rise in σ_d as fewer and smaller gradients are used, like in the previous cases. Initially, σ_d for all diameters are within around 0.5 μm of each other (subset #1). Some σ_d increase at a steadier rate than others. Those of $d = 2$ μm or $d = 3$ μm steadily increase until the 7th subset, while $d = 5$ μm increases at a faster rate after the 4th or 5th subset. Overall, these are very similar to the corresponding $D_{in} \neq D_{ex}$ case (Figure 5.17a).

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.24b shows σ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. The main features are the same as Figure 5.17b. At $G_{high} = 100$ mT/m (subset #9), σ_d for $d = 4$ μm and $d = 5$ μm jump about 2 μm. One difference between this case and the $D_{in} \neq D_{ex}$ case is that σ_d for $d = 1$ μm are smaller here than in Figure 5.17b, but that could be explained by less outliers in the fitted diameter distributions here (in other words, by chance).

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.24c shows σ_d for $d = 1$ to 5 μm using the gradient subsets in Table 5.1. When using one nonzero gradient, σ_d are up to 1 μm higher.

5.3.4.2 Frequency Removal

The results presented here are for simulations with packing fractions of 0.6 that use all the gradients and fewer frequencies.
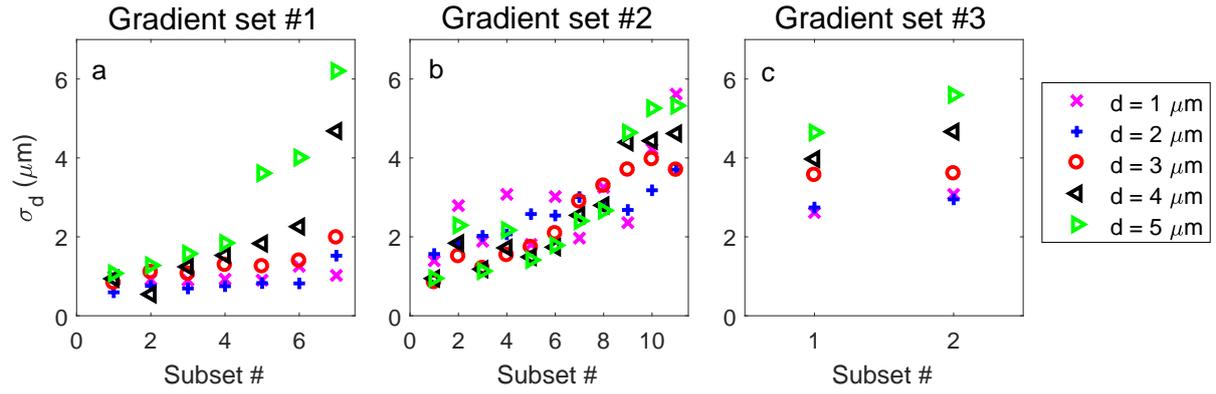
Figure 5.24: Full simulation: σ_d for different gradient subsets. Simulations used the same diffusivity for intra-axonal and extra-axonal compartments. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m, σ_d increases as the maximum gradient strength in the subset decreases. For $G_{max} = 300$ mT/m, there is an overall increase in σ_d with gradient subset. For the simulations with $G_{max} = 80$ mT/m, with exception of $d = 2$ μm, σ_d are slightly larger for the two gradient subset.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.25a shows σ_d as a function of n_f for $d = 1$ to 5 μm. For $d = 1$ μm, σ_d levels after $n_f = 10$ (where $\sigma_d < 1.2$ μm). For $d = 2$ to 5 μm, σ_d show little change, changing less than 0.5 μm over n_f. All σ_d remain under 1.2 μm.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.25b shows σ_d as a function of n_f for $d = 1$ to 5 μm. For $d > 2$ μm, the σ_d show little change over n_f. The σ_d for each diameter stay in a 0.2 μm range. The σ_ds for both $d = 1$ μm and $d = 2$ μm decrease over n_f, with those of $d = 1$ μm leveling at $n_f = 10$.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.25c shows σ_d as a function of n_f for $d = 1$ to 5 μm. There does not appear to be a consistent decrease of σ_d with n_f.

5.3.5 Diameter Distributions – Equal Diffusivities

Finally, we look at the results from diameter distributions when $D_{in} = D_{ex}$.
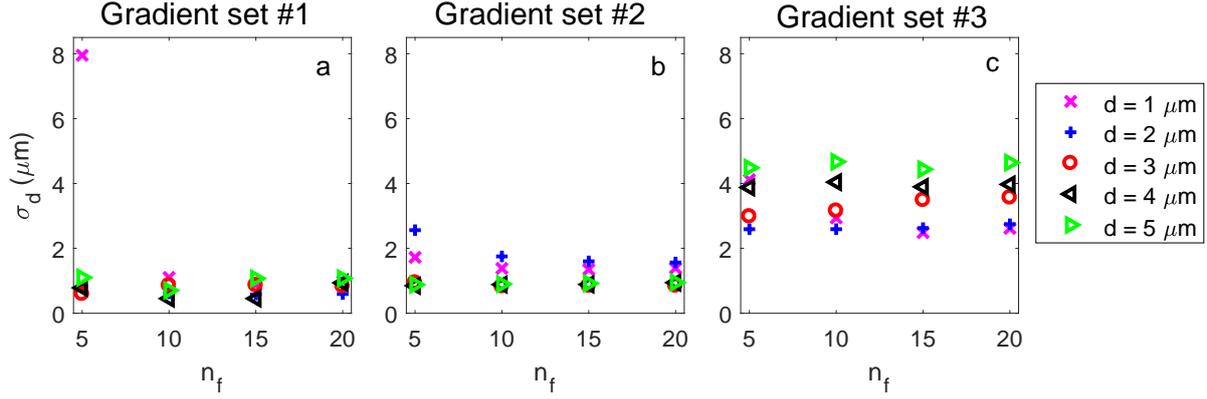
Figure 5.25: Full simulation: σ_d as a function of n_f. Simulations used the same diffusivity for intra-axonal and extra-axonal compartments. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For simulations with $G_{max} = 900$ mT/m, the uncertainty for $d = 1$ μm is large at smaller n_f but levels off after $n_f = 10$. For $d = 2$ μm to $d = 5$ μm, σ_d at $n_f = 5$ are larger than those for $n_f > 5$. For simulations with $G_{max} = 300$ mT/m, σ_d for $d = 3$, 4, and 5 μm show little change over n_f. The $d = 1$ μm and $d = 2$ μm errors become smaller as n_f increases. For $d > 2$ μm, σ_d are similar to those of first gradient set. For simulations with $G_{max} = 80$ mT/m, σ_d are mostly independent of n_f.

5.3.5.1 Gradient Subsets

The results presented first are for simulations using all the frequencies ($n_f = 20$) with different gradient subsets.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.26a shows σ_{AxD} using the gradient subsets in Table 5.1. We can see the same features as Figure 5.21a. There is an increase in σ_{AxD} as smaller and fewer gradients are used. Some increase at a steadier rate, while others have a point where the rate of increase becomes larger. The σ_{AxD} for $AxD = 4.8$ μm and $AxD = 6.1$ μm increase at a faster rate starting at the 4th subset, while σ_{AxD} of $AxD = 2.1$ μm, $AxD = 2.3$ μm, and $AxD = 3$ μm experience a small jump at the 7th (increasing by around 0.25 μm).

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.26b shows σ_{AxD} using the gradient subsets in Table 5.1. It looks similar to Figure 5.21b. There is an overall increase in σ_{AxD} at higher subsets. Otherwise, it cannot be seen from the figures whether using $D_{in} = D_{ex}$ gives smaller σ_{AxD} or not. For example, the σ_{AxD}s

of $AxD = 2.1$ µm are larger here than the other case ($D_{in} \neq D_{ex}$) when using subsets with $G_{\text{high}} > 100$ mT/m (up to 0.9 µm), whereas using subsets with $G_{\text{high}} \leqslant 100$ mT/m its error is smaller than the other case.

Gradient set #3 (G_{max} **= 80 mT/m**)

Figure 5.26c shows σ_{AxD} using the gradient subsets in Table 5.1. With the exception of $AxD = 4.8$ µm and 6.1 µm, where σ_{AxD} increases by ≈ 0.2 to 0.3 µm, σ_{AxD} decrease by less than 0.1 µm when using the subset with one nonzero gradient. This might just be due to noise (in the 1000 trials).
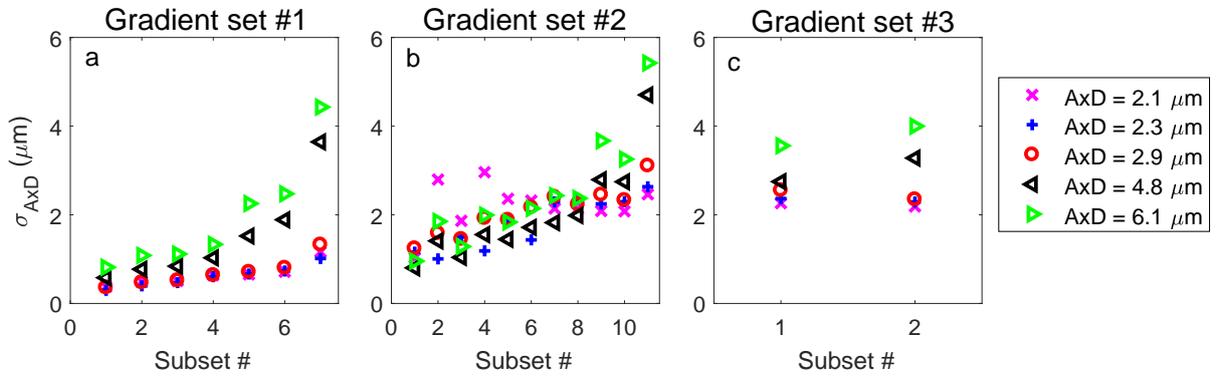
Figure 5.26: Full simulation: σ_{AxD} for different gradient subsets. Simulations used the same diffusivity for intra-axonal and extra-axonal compartments. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For the simulations with $G_{max} = 900$ mT/m, there is an increase in σ_{AxD} as the highest gradient strength in a subset decreases. For distributions with smaller AxD, σ_{AxD} tends to increase at a slower rate. For the simulations with $G_{max} = 300$ mT/m, there is an overall increase in σ_{AxD} as the largest gradient in the subset decreases. For the simulations with $G_{max} = 80$ mT/m, σ_{AxD} are similar in both subsets.

5.3.5.2 Frequency Removal

We now remove frequencies to see how σ_{AxD} changes. The results presented here use all the gradients and fewer frequencies.

Gradient set #1 (G_{max} **= 900 mT/m**)

Figure 5.27a shows σ_{AxD} as a function of n_f for the five AxD. The σ_{AxD} decrease with n_f, though more slowly past $n_f = 10$. Those of $n_f = 5$ are around 1.5 to 2 times larger than those

of $n_f \geqslant 10$. With some exceptions, mostly at $n_f = 5$, σ_{AxD} here are slightly smaller than when $D_{in} \neq D_{ex}$ (Figure 5.23a).

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.27b shows σ_{AxD} as a function of n_f for the five *AxD*. Each set of σ_{AxD} changes less than 0.1 μm over n_f. All σ_{AxD} are within 0.8 μm to 1.4 μm for the entire range of n_f. These σ_{AxD} are larger than when $D_{in} \neq D_{ex}$ (Figure 5.23b) by up to 0.5 μm, depending on the particular *AxD*.

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.27c shows σ_{AxD} as a function of n_f for the five *AxD*. There does not seem to be a consistent decrease of σ_{AxD} over n_f but rather a dip in σ_{AxD} going from $n_f = 5$ to $n_f = 20$ for the smaller *AxD*. For *AxD* = 2, 2.4, and 3 μm, the σ_{AxD} here are slightly smaller (by 0.1 to 0.2 μm) than the $D_{in} \neq D_{ex}$ case (for all n_f). For *AxD* = 4.8 μm and *AxD* = 6.1 μm, σ_{AxD} are approximately 0.1 μm larger here (see Figure 5.23c).
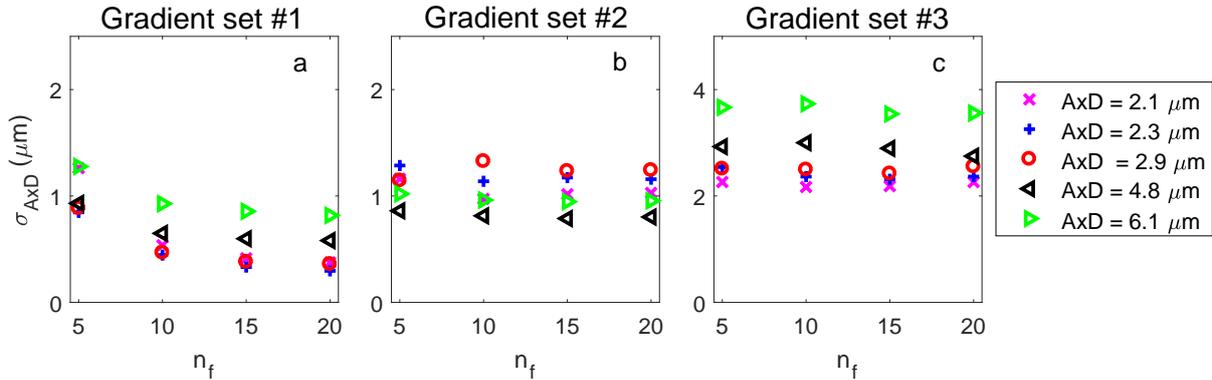
Figure 5.27: Full simulation: σ_{AxD} as a function of n_f. Simulations used the same diffusivity for intra-axonal and extra-axonal compartments. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths in each set were kept when fitting to the model. For the simulations with $G_{max} = 900$ mT/m, the σ_{AxD} are highest at $n_f = 5$ and decrease with n_f. After $n_f = 10$, they begin to level off. For the simulations with $G_{max} = 300$ mT/m, the values of σ_{AxD} show little change over n_f. For the simulations with $G_{max} = 80$ mT/m, the values of σ_{AxD} are obviously much larger than either the first or second gradient set. There does not appear to be a clear trend over n_f. **Note**: (c) uses different y-axis limits.

5.3.6 Repeated Measurements at G_{max}

In the previous sections, the best results were usually found when keeping all of the gradients. To study whether two gradients ($G = 0$ and G_{max}) with more measurements at G_{max} are better than one measurement at each of the original gradients, this section studies σ_d (σ_{AxD}) for repeated measurements at G_{max}. When comparing, the total number of measurements is kept constant, so that if there were four nonzero gradients before ($G = \{0, G_1, G_2, G_3, G_4\}$), there will still be four nonzero gradients ($G = \{0, G_4, G_4, G_4, G_4\}$). We only use the $D_{in} \neq D_{ex}$ simulation data here because the previous results from the $D_{in} = D_{ex}$ simulations were very similar.

5.3.6.1 Intra-axonal simulation

We begin the comparisons with the intra-axonal simulation data. For each gradient set, three gradient subsets were used to study σ_d. One choice uses all the original gradients, while another uses repeated measurements at $G = G_{max}$ (while keeping the number of measurements constant). For comparison, the subset with one nonzero gradient ($G = \{0, G_{max}\}$) is included.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.28a shows σ_d as a function of n_f for the three subsets (for $d = 2$ μm). Repeated measurements at G_{max} give the smallest σ_d over the entire range of n_f. At $n_f = 5$, the new σ_d is around 0.6 μm smaller than when using all gradient values. The difference shrinks as we move towards $n_f = 20$, as the σ_d for each subset start to converge.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.28b shows σ_d as a function of n_f for the three subsets (for $d = 2$ μm). Repeated measurements at G_{max} give the smallest σ_d at all n_f. At $n_f = 5$, the difference between using repeated measurements at G_{max} and the original measurements is the largest (around 0.8 μm). At $n_f = 10$, the difference is smallest (around 0.1 μm).

Gradient set #3 (G_{max} = 80 mT/m)

We have already shown that fitting smaller diameter cylinders may not be feasible with these gradient strengths. Thus, instead of $d = 2$ μm, the results from $d = 3$ μm were studied, because the fitted diameter distribution for $d = 3$ μm had its maximum near $d = 3$ μm.

Figure 5.28c shows σ_d as a function of n_f for the three subsets (for $d = 3$ μm). Repeated measurements at G_{max} once again produced the smallest σ_ds. At $n_f = 5$ and $n_f = 10$, the σ_d from repeated measurements is 1 μm smaller than from using all measurements, otherwise σ_d is around 0.5 μm smaller.

5.3.6.2 Full simulation – Identical cylinders

This section presents comparisons for the full model (intra- and extra-axonal).

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.29a shows σ_d as a function of n_f for the three subsets ($d = 2$ μm, packing fraction of 0.6). As before, repeated measurements at G_{max} give the smallest σ_d at each n_f. The improvement over using all measurements is greatest with $n_f = 5$, with a 0.5 μm difference. As n_f increases, σ_d for the subsets start to converge.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.29b shows σ_d as a function of n_f for the three subsets ($d = 2$ μm, packing fraction of 0.6). Repeated measurements at G_{max} give the smallest σ_d, with improvements over the other subsets ranging from 0.5 to 1 μm.

Gradient set #3 (G_{max} = 80 mT/m)

For illustration, a diameter with a more well-defined distribution ($d = 4$ μm) was studied with this gradient set. Figure 5.29c shows σ_d as a function of n_f for the three subsets (packing fraction of 0.6). Repeated measurements at G_{max} lead to smaller σ_d over all n_f, with improvements between 0.2 to 0.5 μm.
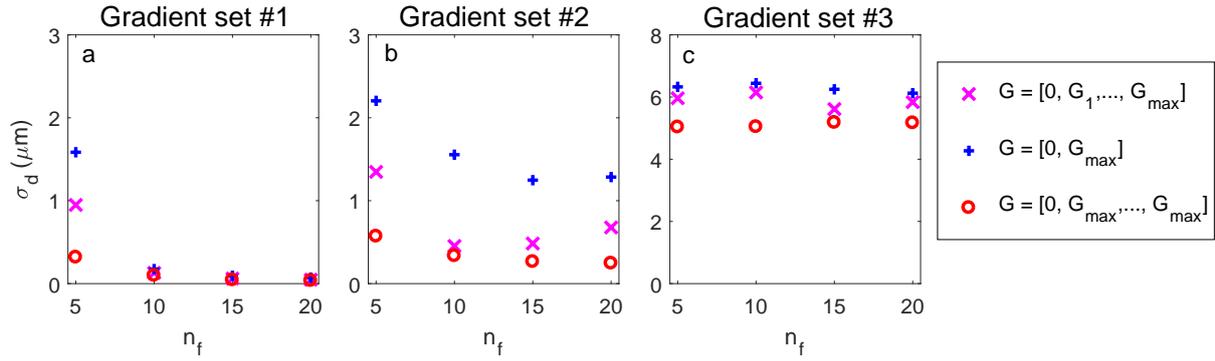
Figure 5.28: Intra-axonal model: σ_d as a function of n_f for three different gradient subsets. Subsets include those with ('+') one nonzero gradient ('x') all gradient strengths ('o') repeated nonzero gradients at G_{max}. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. In (a) and (b), σ_d are from $d = 2$ μm simulations, in (c) σ_d are from $d = 3$ μm simulations. For simulations with $G_{max} = 900$ mT/m repeated measurements at G_{max} give the smallest σ_d over the entire range of n_f. The improvement in σ_d shrinks with increasing n_f. For simulations with $G_{max} = 300$ mT/m repeated measurements at G_{max} give the smallest σ_d at all n_f. At $n_f = 5$, the difference between using repeated measurements at G_{max} and all measurements is the largest. At $n_f = 10$, the difference is smallest. For simulations with $G_{max} = 80$ mT/m repeated measurements at G_{max} once again produced the smallest σ_ds. **Note**: (c) uses different y-axis limits.
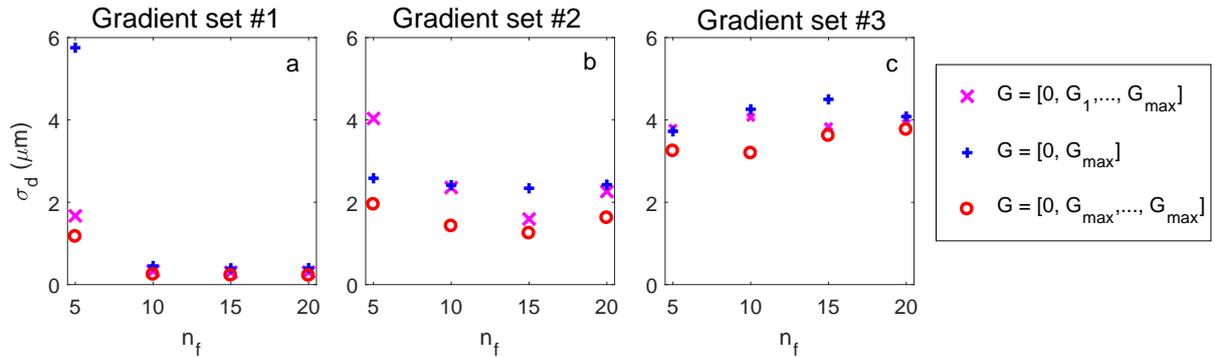
Figure 5.29: Full simulation: σ_d as a function of n_f for three different gradient subsets. Subsets include those with ('+') one nonzero gradient ('x') all gradient strengths ('o') repeated nonzero gradients at G_{max}. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. In (a) and (b), σ_d are from $d = 2$ μm simulations, in (c) σ_d are from $d = 4$ μm simulations. For simulations with $G_{max} = 900$ mT/m, as before, repeated measurements at G_{max} give the smallest σ_d at each n_f. As n_f increases, σ_d for the subsets start to converge. For simulations with $G_{max} = 300$ mT/m, repeated measurements at G_{max} give the smallest σ_d. For simulations with $G_{max} = 80$ mT/m, repeated measurements at G_{max} lead to smaller σ_d over all n_f.

5.3.6.3   Full simulation - diameter distribution

Finally, we also illustrate the comparison using the diameter distributions.

Gradient set #1 (G_{max} = 900 mT/m)

Figure 5.30a shows σ_{AxD} as a function of n_f for the three subsets ($AxD = 2.1$ μm). Repeated measurements at G_{max} give smaller σ_{AxD} for $n_f > 5$. Improvement in σ_{AxD} ranges from around 0.1 μm ($n_f = 15, 20$) to 0.2 μm ($n_f = 10$). Differences decrease with n_f, so that σ_{AxD} for each subset converge (or level off) at high n_f.

Gradient set #2 (G_{max} = 300 mT/m)

Figure 5.30b shows σ_{AxD} as a function of n_f for the three subsets ($AxD = 2.1$ μm). Repeated measurements at G_{max} give smaller σ_{AxD} for all n_f. The improvement over the second best subset ranges from 0.1 μm ($n_f = 5$) to 0.3 μm ($n_f = 10, 15, 20$).

Gradient set #3 (G_{max} = 80 mT/m)

Figure 5.30c shows σ_{AxD} as a function of n_f for the three subsets ($AxD = 2.9$ μm). There is a slight improvement in σ_{AxD} when using repeated gradients, but the σ_{AxD} jump around, making it more difficult to see a clear trend.

Figures 5.31a-c shows the fitted *AxD* distributions using each of the three subsets ($n_f = 20$). When using all the original gradients, the distribution looks bimodal, with one peak between 2 and 3 μm and another between 5 and 6 μm. There are just over 200 values near $AxD = 0$ μm ($213/1000 < 0.02$ μm). Using repeated gradients, there are fewer values near $AxD = 0$ μm ($179/1000 < 0.02$ μm) and the second peak between 5 and 6 μm has become smaller, so that the maximum near 3 μm is the only maximum left. With one nonzero gradient, there are more values near $AxD = 0$ μm than either of the others ($268/1000 < 0.02$ μm) and the distribution itself has become flatter over the range of 2 to 6 μm.
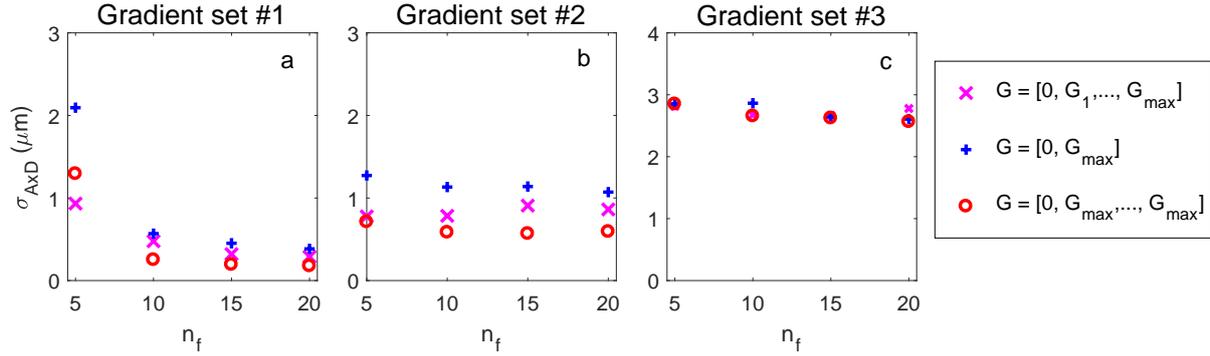
Figure 5.30: Full simulation: σ_{AxD} as a function of n_f for three different gradient subsets. Subsets include those with ('+') one nonzero gradient ('x') all gradient strengths ('o') repeated nonzero gradients at G_{max}. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. In (a) and (b), σ_{AxD} are from $AxD = 2.1$ µm simulations, in (c) σ_{AxD} are from $AxD = 2.9$ µm simulations. For simulations with $G_{max} = 900$ mT/m, repeated measurements at G_{max} give smaller σ_{AxD} for $n_f > 5$. The improvement in σ_{AxD} becomes smaller with increasing n_f, and the σ_{AxD} for each subset level off at high n_f. For simulations with $G_{max} = 300$ mT/m, repeated measurements at G_{max} give smaller σ_{AxD} for all n_f. For simulations with $G_{max} = 80$ mT/m, there is a slight improvement in σ_{AxD} when using repeated gradients, but the trend is not as clear. **Note**: (c) uses different y-axis limits.
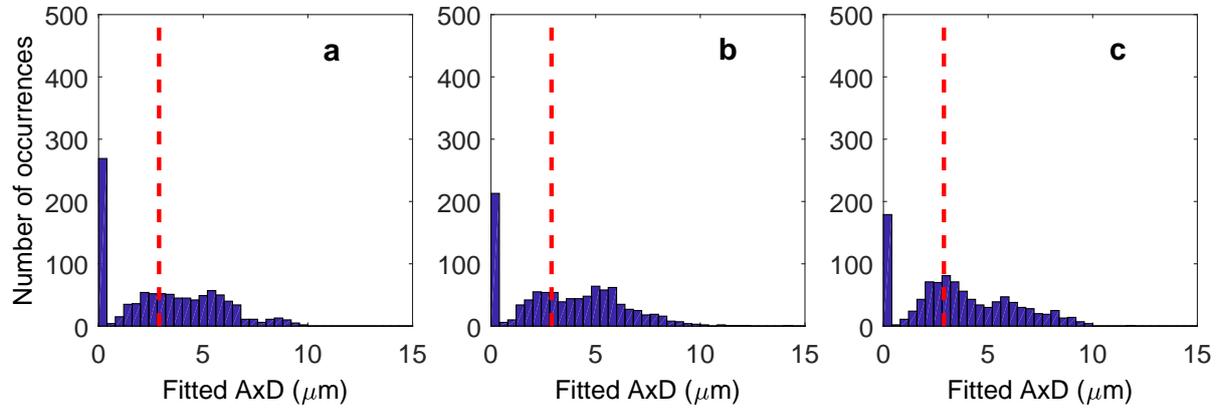
Figure 5.31: Full simulation: Distribution of fitted AxD from the substrate with $AxD = 2.9$ µm using different gradient subsets (here, $n_f = 20$) of the $G_{max} = 80$ mT/m gradient set. Subsets include those with (a) one nonzero gradient, (b) all gradient strengths, (c) six repeated nonzero gradients at $G_{max} = 80$ mT/m. When using all gradients, the distribution looks bimodal, with one peak between 2 and 3 µm and another between 5 and 6 µm. There are just over 200 values near $AxD = 0$. Using repeated gradients, there are fewer values near $AxD = 0$ and the second peak between 5 and 6 µm has become smaller, so that the maximum near 3 µm is the main peak. With one nonzero gradient, there are more values near $AxD = 0$ than in either (b) or (c) and the distribution itself has become flatter (ranging between 2 and 6 µm.

5.4 Discussion

The results of the simulations suggest that 5 frequencies were adequate for $d = 3$ to 5 μm and $AxD > 2$ μm. There was some improvement in precision for $d = 1$ and 2 μm with 10 frequencies (at least for $G = 300$ mT/m and $G = 900$ mT/m), but comparisons at these diameters might require a more refined technique. Similar to other diffusion MR studies [51], we found that it was better to repeat measurements at higher gradient strengths than to make single measurements over a range of gradient strengths. The improvement in σ_d/σ_{AxD} tended to be greatest at small n_f, especially noticeable at very high gradient strengths (= 900 mT/m). At high n_f, there was less improvement.

The measure we used here was a sample standard deviation over the fitted parameters. For fitted parameters with symmetric distributions, standard deviation may be an adequate description of uncertainty. But it was found that many fitted parameters have asymmetric distributions, especially for smaller gradients or smaller diameters. It was also found that σ_d/σ_{AxD} for these distributions tended to be skewed by outliers, especially at the smallest diameters (e.g. $d = 1$ μm), and even with high gradients.

Bayesian methods might be useful for fitting [6, 7, 160]. Some have shown that the use of prior distributions might make results converge for smaller gradients [34]. Clayden et al. have found this improved their results using gradient strengths as small as 35 mT/m.

In all the cases we studied, the intra-axonal diffusion coefficient was the most poorly fit parameter. Many of the fitted estimates clustered near the upper and lower bounds and the distribution itself was less clearly defined (in terms of peaks). It might be best just to fix it to a set value beforehand, like we did for the diameter distribution simulations. The fitted packing fraction was also a poorly estimated using the $G_{max} = 80$ mT/m measurements.

In Figure 5.19b, σ_d for $d = 3$, 4, and 5 μm are constant over n_f, whereas in Figure 5.19a, those σ_d increase at $n_f = 5$. The same can be seen in Figure 5.23, where Figure 5.23b shows that σ_{AxD} remain constant over n_f, but Figure 5.23a shows σ_{AxD} increase at $n_f = 5$. The reason is that the first gradient set ($G_{max} = 900$ mT/m) does not necessarily use the highest gradient strengths at each frequency. Some gradient strengths for $n_f \leqslant 5$ ($f_m \leqslant 250$ Hz) are smaller in this set of measurements than the $G_{max} = 300$ mT/m set of measurements. The fact that the

second set of measurements uses a single gradient over all frequencies, with higher gradient strengths at low frequencies, might be responsible for the smaller σ_d/σ_{AxD} at $n_f = 5$.

The σ_d/σ_{AxD} from the $G_{max} = 300$ mT/m and $G_{max} = 80$ mT/m (especially 80 mT/m) measurements showed less change over n_f compared to those from the $G_{max} = 900$ mT/m measurements. One possible cause is that the $G_{max} = 300$ mT/m and $G_{max} = 80$ mT/m measurements have less diffusion weighting at higher frequencies. Therefore including higher frequency measurements when fitting does not add very much new information, especially once noise is added. This would be especially true for $G_{max} = 80$ mT/m measurements, where the signal only decays to 95% at 50 Hz.

For the $G_{max} = 900$ mT/m measurements, where the gradient step increase is largest, the differences between σ_d/σ_{AxD} from different gradient subsets was noticeable. A clear increase in σ_d/σ_{AxD} is observed moving from multiple or higher gradient strengths to fewer and smaller gradient strengths. In the $G_{max} = 80$ mT/m measurements, it was not completely clear what was happening as the gradient subsets were changed from five nonzero gradients to a single nonzero gradient. For some diameters, σ_d/σ_{AxD} increased, for others σ_d/σ_{AxD} decreased. For a more accurate measure of how uncertainty changes, a better measure than just taking the standard deviations of the parameter distribution is needed. Alternatively, more than 1000 noisy trials could be used to find σ_d (for example, 5000 [192]).

With the intra-axonal model, the $G_{max} = 300$ mT/m and $G_{max} = 900$ mT/m measurements gave similar σ_ds for $d = 3$ to 5 µm, at least when all gradient strengths were included (in other words, there was little difference between using $G = 300$ mT/m and $G = 900$ mT/m). When going below $G = 300$ mT/m, the difference becomes apparent, with σ_d becoming worse. The $G_{max} = 300$ mT/m and $G_{max} = 900$ mT/m measurements also gave similar σ_ds for $d = 3$ to 5 µm with the two compartment model. The two sets of measurements also gave similar σ_{AxD}s for the substrates we studied ($AxD = 2.1$ µm to 6.1 µm). There did not appear to be a noticeable difference in results between simulations with $D_{in} \neq D_{ex}$ and simulations with $D_{in} = D_{ex}$.

5.5 Conclusion

Oscillating gradients have already shown to be useful in inferring small axon diameters. Here we have shown using Monte Carlo simulations that images can be collected with fewer gradient strengths (as few as 2) and fewer gradient frequencies (between 5 to 10) without sacrificing the precision of the measurements. This could be useful in reducing imaging time so that OG techniques can be used in clinical settings. More work is needed for the method to work with clinical gradient strengths.

Chapter 6

Short-time model

6.1 Introduction

In this chapter, we will study a short-time *ADC* model which can be used to infer surface to volume ratios in samples. First we will look at the accuracy of fitted parameters without signal noise. Then we will investigate how signal noise affects parameter estimates. We will also look at how parameter precision changes depending on the number of OG frequencies and gradients. This analysis will be repeated using three different sets of gradient measurements (clinical to pre-clinical strengths).

At short timescales, the interaction of molecules with the surfaces in porous samples results in some interesting behaviour. Mitra showed that at short times ($\sqrt{D_0 t} \ll$ pore size [89]), the diffusion coefficient is proportional to the square root of time and depends on the surface to volume ratio of restrictions in a sample [122, 123, 124]. The relationship between the time dependence of the diffusion coefficient, assumed to be measured with narrow gradient pulses, and surface to volume ratio allows for the possibility of extracting information about restrictions in porous media.

Latour [89] performed experiments using stimulated pulsed field gradients (PFG) on packed glass beads (48 μm, 96 μm, 194 μm in diameter) and a sample of onion (extracted cell diameter = 138 μm). The model incorporated the short time model and a long time tortuosity model covering a range of diffusion times, between 7 ms to 800 ms. Due to finite duration pulses,

the original model in Ref. [124] was modified to take into account movement occurring while the gradients were turned on. Their experiments yielded good results and were able to extract surface to volume ratios, free diffusion coefficients, and tortuosity constants.

With the potential for oscillating gradient sequences to probe shorter diffusion times, experiments began to be made with smaller samples using OGSE sequences. Schachter [154] measured surface to volume ratio in samples of packed beads using the short-medium time interpolation model from Latour [89]. Bead diameters were 3 μm, 9 μm, 25 μm, and 46 μm. Experimental OGSE frequencies ranged from 12.5 Hz to 87.5 Hz, with a maximum gradient strength of 157.5 mT/m. Diffusion was completely restricted in the 3 μm sample, so the long time diffusion coefficient was taken from it and used, through a tortuosity relation, to help extract the surface to volume ratio from the 9 μm bead sample. Extracted parameters were in good agreement with the true values.

Parsons [140] also estimated surface to volume ratios of packed bead samples using OGSE sequences. Bead samples ranged from 1 μm to 400 μm in diameter. Since the short time model as derived by Mitra is only strictly true for narrow pulses, where diffusion times are well defined, they used a modified form of the short time model to account for the fact that diffusion times are not well defined for oscillating gradients. However, their modification assumed oscillating gradients of infinite duration. Frequencies ranged from 33.33 Hz to 1000 Hz (30 measurements) and the maximum gradient strength was 1730 mT/m (or b-values of 400 s/mm^2). Bead diameters were extracted by fitting either the highest or lowest measured OGSE frequencies, depending on the restriction sizes in the sample. Surface to volume ratios for the 1 μm and 3 μm bead samples were unable to be properly extracted because of small restriction sizes. Diffusion times were too long for the short time model to apply. Instead, they extracted lower bounds on the surface to volume ratios by assuming a known free diffusion coefficient ($D_0 = 2.05$ μm^2/ms).

Novikov [135] derived the exact high frequency behaviour of the diffusion coefficient, a useful result for OGSE sequences and valid for a large number of oscillations. While previous groups such as Parsons had numerically derived the necessary modifications for the high frequency behaviour, Novikov's result was derived analytically. Sukstanskii [169] found the exact

172

high frequency behaviour for sine and cosine gradients of any duration, expressing results in terms of correction factors dependent upon number of lobes. As gradients become infinite in duration, the expressions of Sukstanskii, Novikov and Parsons all converge to the same result.

Reynaud [149] used the short time OGSE behaviour to estimate surface to volume ratios in tumours of mice. Measurements were taken both *in vivo* and *ex vivo*. Three *b*-values were used (0, 0.2, 0.4 ms/μm^2). Frequencies ranged from 65 to 225 Hz (10 measurements). Results were used to create parametric maps of D_0 and S/V. They found that using only higher frequencies ($f_m > 100$ Hz) when fitting caused confidence intervals of extracted parameters to increase. It was also found that neglecting the correction factors of Sukstanskii caused an overestimation of S/V (by 20%).

In this chapter, we use data from Monte Carlo simulations to test the ability of the short-time model to extract surface to volume ratios for a variety of geometries. Geometries include identical square packed cylinders and substrates where cylinders have a distribution of diameters. We also use simulations to find how the precision of parameter estimates change depending on the number of frequencies and gradient strengths used in the experiment. We will describe results for three sets of gradient strengths (with maximum strengths of 80 mT/m, 300 mT/m, and 900 mT/m). The aim of these experiments is to reduce imaging time to a clinically useful sequence using clinically available gradient strengths. In addition, we compare results when multiple gradient measurements (or *b*-values) at a given frequency are spaced out over a range or collected at two points, one zero and the other nonzero (with the nonzero gradient repeated).

6.2 Mitra model

Suppose we have a collection of molecules diffusing near a surface (shown as a cylinder in Figure 6.1, but the following ideas will hold for almost any shape). At very short timescales, most molecules do not come into contact with the boundary. These molecules have no information about the shape or overall size of the pore. A short time later, molecules within a distance ($\approx \sqrt{2D_0t}$) of the boundary will have interacted with it. During this time the surface seems to be completely flat, so molecules find no information about the curvature.

173

We can derive a simple formula for the time dependent diffusion coefficient near a boundary at short times [156, 157]. We first divide the molecules into two groups. The first group is a 'free' group, made up of particles too far from the boundary to interact with it. The second group is a 'wall' group, with molecules near enough to the boundary to interact. Assume that the boundary is flat and at the yz-plane. The total mean squared displacement in the direction perpendicular to the wall is the average of the mean squared displacements of the free molecules and the molecules near the boundaries. The mean squared displacement of the 'free' molecules is $2D_0 t$. The mean squared displacement of the 'wall' molecules is reduced by some fraction p. Due to the presence of the wall, their squared displacements are smaller than if they were free. The total mean squared displacement can be written as

$$\langle x^2 \rangle = f_{free} \langle x^2 \rangle_{free} + f_{wall} \langle x^2 \rangle_{wall} \tag{6.1}$$

where f_{free} and f_{wall} are the fractions of molecules far from the wall and close to wall, respectively. If we assume the entire pore system is closed, we also have $f_{\text{free}} = 1 - f_{\text{wall}}$. We expect that only molecules less than a distance $\sqrt{2D_0 t}$ from the wall will feel its effects. If the surface area of the boundary is S, then the volume taken up by these "wall" molecules is $S\sqrt{2D_0 t}$. If the entire pore has total volume V, f_{wall} is $\frac{S}{V} \cdot \sqrt{2D_0 t}$. The total mean squared displacement will be

$$\langle x^2 \rangle = \left(1 - \frac{S}{V} \cdot \sqrt{2D_0 t} \right) \cdot 2D_0 t + \frac{S}{V} \cdot \sqrt{2D_0 t} \cdot p \cdot 2D_0 t \tag{6.2}$$

Displacements parallel to the wall are unaffected:

$$\langle y^2 \rangle = \langle z^2 \rangle = 2D_0 t \tag{6.3}$$

The total mean squared displacement is

$$\langle r^2 \rangle = 4D_0 t + \left(1 - \frac{S}{V} \cdot \sqrt{2D_0 t} \right) \cdot 2D_0 t + \frac{S}{V} \cdot p \cdot \sqrt{2D_0 t} \cdot 2D_0 t \tag{6.4}$$

The time dependent diffusion coefficient $D(t)$ is related to the mean squared displacement

through the relationship $D(t) = \frac{\langle r^2 \rangle}{6t}$ (recall Section 3.3.5). Therefore,

$$D(t) = D_0 \left(1 - \frac{(1-p)}{6} \frac{S}{V} \sqrt{2D_0 t} \right) \tag{6.5}$$

The diffusion coefficient depends on the surface to volume ratio S/V and on $\sqrt{D_0 t}$, independent of the exact shape of the pore. All that remains is for the exact form of p to be determined. It has been shown that $D(t)$ is [89, 157, 201]

$$D(t) = D_0 \left(1 - \frac{4}{3n_d \sqrt{\pi}} \cdot \frac{S}{V} \sqrt{D_0 t} \right) \tag{6.6}$$

where n_d is the dimensionality of the system (whether diffusion takes place in 1, 2, or 3 dimensions).

The *ADC* measured in a typical experiment is equal to $D(t)$ only for narrow pulses [169] (Section 3.3.5). Therefore, Eq. 6.6 is only valid for narrow gradient pulses. For arbitrary gradients, Eq. 6.6 becomes

$$ADC(t) = D_0 \cdot \left[1 - c_G \cdot \frac{S}{n_d V} \sqrt{D_0 t} \right] \tag{6.7}$$

where c_G depends on the gradient waveform [169]. In principle, given $G(t)$, one can find c_G. It has been shown that c_G for a cosine-OGSE with N_{OG} periods is

$$c'_{\cos}(N_{OG}) = \frac{4\pi N_{OG} \cdot C(2\sqrt{N_{OG}}) + 3 \cdot S(2\sqrt{N_{OG}})}{2\sqrt{2}\pi N_{OG}} \tag{6.8}$$

where $C(x)$ and $S(x)$ are Fresnel functions[1] [3, 41, 169]. Previously, c_G had been found for the case of $N \to \infty$ [135]. For a sine-OGSE, c_G is

$$c'_{\sin}(N_{OG}) = \frac{32\pi N_{OG}^{3/2} + 12\pi N_{OG} \cdot C(2\sqrt{N_{OG}}) + 21 \cdot S(2\sqrt{N_{OG}})}{18\sqrt{2}\pi N_{OG}} \tag{6.9}$$

As N_{OG} increases, the coefficient $c'_{\cos}(N_{OG})$ approaches a constant value of $1/\sqrt{2}$. The limit is reached relatively quickly, the difference between $c'_{\cos}(N_{OG} \to \infty)$ and $c'_{\cos}(N_{OG} = 5)$ is less

---

[1]The Fresnel integrals are defined as follows: $S(x) = \int_0^x \sin(\pi t^2/2)dt$ and $C(x) = \int_0^x \cos(\pi t^2/2)dt$

than 5%. Because of the first term in the numerator of Eq. 6.9, $c'_{\sin}(N_{OG})$ diverges as $N_{OG} \to \infty$. Finally, rewriting Eq. 6.7 in terms of OGSE angular frequency ω_m, we have

$$ADC(t = 2\pi N_{OG}/\omega_m) = D_0 \cdot \left[1 - c_G \cdot \frac{S}{dV} \sqrt{D_0 \cdot \frac{2\pi N_{OG}}{\omega_m}} \right] \tag{6.10}$$

6.2.1 Cylinders

As usual, white matter fibres were modeled as a collection of parallel non-overlapping cylinders. We will use the short-time model in Eq. 6.10 to extract surface to volume ratios and diffusion coefficients from a variety of simulated geometries. For simplicity, we assume no water exchange between intra- and extra-axonal regions [7, 11, 195].

We performed two types of simulations. Initially, we performed "intra-axonal" simulations, where water is contained inside the axons, but not outside. We then distributed water both inside and outside the axons (a "full" simulation). In simulations with intra- and extra-axonal water, we first set the extra-axonal diffusion coefficient to 2.5 μm^2/ms [154]. We repeated the simulations with $D_{ex} = 1.0$ μm^2/ms. The *ADC* should be given by Eq. 6.10. These simulations ("intra-axonal" or "full") were used on the geometries given below.

6.2.1.1 Geometry 1: Single cylinder

The simplest simulation consists of water diffusing inside the cylinder of diameter d (an "intra-axonal" simulation). Ten diameters $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ μm were used.

6.2.1.2 Geometry 2: Identical cylinders on a square lattice

The next geometry considered is a collection of identical cylinders of diameter d arranged on a square lattice. Water was distributed everywhere, both inside and outside the cylinders (a "full" simulation). Ten diameters $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ μm were used, with packing fractions of 0.5, 0.6, and 0.7.
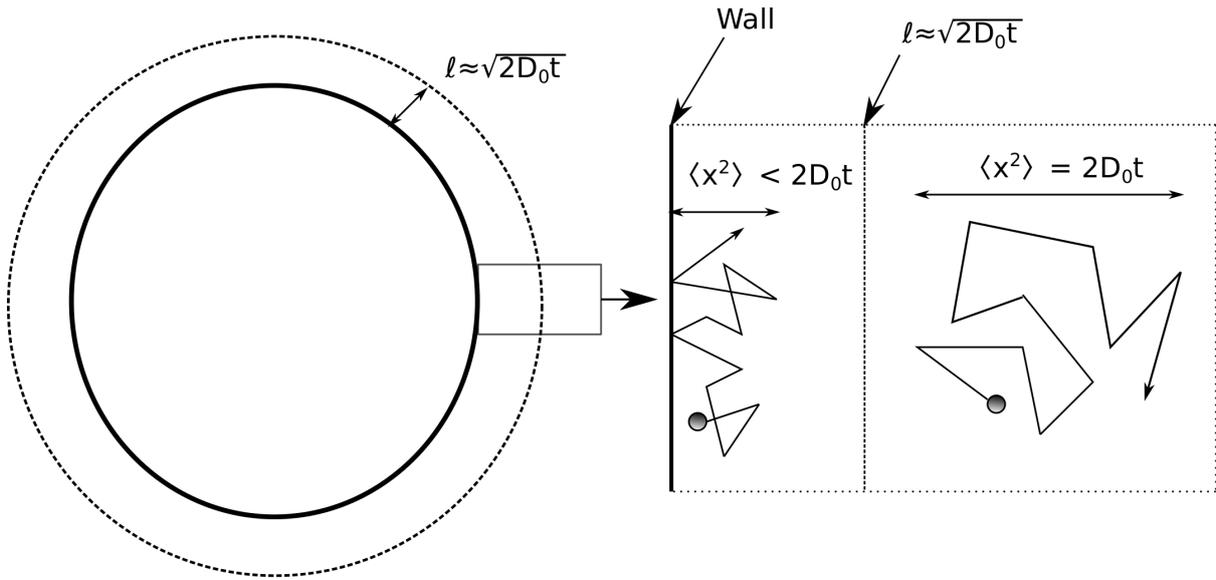
Figure 6.1: Short-time behaviour of particles surrounding a solid cylinder. At very short times, only particles within $\approx \sqrt{2D_0 t}$ can interact with the wall. The fraction of particles that this represents is approximately $S/V \cdot \sqrt{2D_0 t}$, where S is the total surface area and V is the volume of the entire container. On the right is a zoomed illustration of two particles inside the rectangular box. The particle far from the wall ($\ell^2 > 2D_0 t$) undergoes free diffusion with total mean-squared displacement $6D_0 t$. The x-component of the mean-squared displacement is $2D_0 t$. The particle near the wall ($\ell^2 < 2D_0 t$) has its mean-squared displacement in the x-direction reduced by a factor p. Displacements in other directions are unaffected.

6.2.1.3   Geometry 3: Distribution of cylinder diameters

We also used five diameter distributions with mean diameters of 1.5, 2.0, 2.1, 4.1, 4.3 μm. Packing fractions were between 0.63 and 0.78. Water was distributed everywhere, both inside and outside the cylinders (a "full" simulation).

6.2.2   Monte Carlo Simulations

Monte Carlo simulations were similar to those in Chapter 4. For simplicity, we assume impermeable barriers [192, 195] and ignore T_2 differences between tissues. For simulations in simple lattices, we used $N = 57344$ particles. For diameter distribution simulations, we used $N = 114688$ particles. The time step used in all simulations was 1 μs.

6.2.3 Three Gradient Sets

Twenty frequencies ranging from 50 to 1000 Hz in steps of 50 Hz were used with an echo time of 44.52 ms. There were three different sets of gradients used.

1. Five gradient strengths were used at each frequency, for a total of 100 signals. These pulse sequence parameters were chosen to be feasible for experiments using a BGA6 gradient set (Bruker Biospin) on our 7 T magnet (maximum gradient strength of 1.01 T/m). For $f_m > 200$ Hz, the gradient strengths were 0, 60, 70, 80 and 90% of maximum. The gradient strengths for 200 Hz or less were smaller. For $f_m = 50$ Hz, $G = 0, 1.5, 4, 7, 10\%$ of maximum; for $f_m = 100$ Hz, $G = 0, 4, 7, 10$, and 15% of maximum; for $f_m = 150$ Hz, $G = 0, 15, 20, 25$, and 30% of maximum; for $f_m = 200$ Hz, $G = 0, 30, 40, 50$, and 60% of maximum (see Table 4.1).

2. Eight gradient strengths were used at each frequency, for a total of 160 signals. Gradients were {0, 50, 100, 150, 200, 250, 300} mT/m.

3. Six gradient strengths were used at each frequency, for a total of 120 signals. Gradients were {0, 40, 50, 60, 70, 80} mT/m.

6.2.4 Data Fitting

Signal data were fitted to the model with nonlinear least-squares regression using the Optimization Toolbox in MATLAB [113]. The diffusion coefficient was constrained to the range [0, 3] μm^2/ms. The surface to volume ratio was constrained to the range [0, 2] μm^{-1}. The fitting procedure was repeated 50 times with different random initial parameters. At the end, the fitted parameters providing the smallest sum-of-squares were chosen as the best fit.

There are two ways to fit simulation data to the model. The first method involves calculating an *ADC* at each frequency and then fitting the *ADC*s to Eq. 6.10. The *ADC*s are calculated using linear regression on $\ln(E)$ and b_{cos} (also how we calculated the *ADC* spectra in Section 4.3.2). The second method is to fit the signal data directly to Eq. 3.84 with the *ADC* term in the exponent given by Eq. 6.10. The second method was used by Renaud [149]. We will use both methods and then compare the results.

6.2.5 Effects of Noise

At the end of the simulations, Gaussian noise was added to each component of the transverse magnetization so that SNR was 100. This was repeated 1000 times and the fitted parameters were saved for analysis. The means and standard deviations of each of the 1000 results were calculated and used to assess accuracy and precision of the results. The distribution mean and standard deviation for the diffusion coefficient are denoted by μ_D and σ_D. The distribution mean and standard deviation for the surface to volume ratio are denoted by $\mu_{S/V}$ and $\sigma_{S/V}$. We will use $\sigma_{S/V}$, the width of the S/V distribution, as a measure of uncertainty.

6.2.6 Number of Frequencies

Data were fitted using different numbers of frequencies. Initially, all 20 frequencies were used (50 to 1000 Hz). This serves as a baseline. Higher frequencies were then successively removed before refitting the data to the model. The number of frequencies used will be denoted by n_f. We fit the data with n_f = 5, 10, and 15 frequencies, corresponding to maximum frequencies of 250 Hz, 500 Hz, and 750 Hz.

Since Eq. 6.10 is applicable at short-times, another option is to remove lower frequency measurements. We chose to keep every second (or fourth) frequency when fitting the data. If every second frequency is kept, there are n_f = 10 frequencies (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 Hz). If every fourth frequency is kept, there are n_f = 5 frequencies (200, 400, 600, 800, 1000 Hz). We will compare this method of frequency removal with the above method to see if there is a difference in results.

6.2.7 Number of Gradients

For each set of measurements, the highest gradients were successively removed from the data before refitting. We also fit data using one nonzero gradient (e.g. $G = \{0, G_i\}$). Finally, fitting was performed using one nonzero gradient repeated multiple times (e.g. $G = \{0, G_i, G_i, G_i, \ldots\}$). This was done by replicating the simulation results before adding noise. The subsets are summarized in Table 5.1. For the $G_{max} = 80$ mT/m gradient set, only the subsets with all six

gradients or one nonzero gradient (at $G = 80$ mT/m) were used since we were less interested in the smallest gradients.

6.3 Results

In this section the results from the simulations of the short-time model are presented. Many of the fits contain results that are pinned at the upper or lower bounds. This tends to happen with smaller gradient strengths where there is less signal decay. Because this will sometimes affect the sample statistics (mean and standard deviation), we will also show examples of the actual parameter distributions and make a note of when it occurs. If they were removed, results would appear better than they actually are.

6.3.1 Model Accuracy

We begin by finding the situations were the model in Eq. 6.10 gives accurate results. For now, no noise will be added to the simulation data. After determining the situations where the model works best, noise will be added to assess the precision of the fitted parameters. In this section, we will use the first gradient set ($G_{max} = 900$ mT/m) because it gives the largest weighting to the signals.

When describing the data fitting procedure, we noted that there were two ways to fit the data to the model. The first involved calculating ADCs and fitting the ADCs to Eq. 6.10. The second involved fitting simulated signals directly to Eq. 3.84 (with the ADC in the exponent given by Eq. 6.10). Here we compare results from the two fitting methods.

For each fit, we choose 50 randomized starting parameters, giving 50 solutions. Solutions are sorted according to their associated sum-of-squares. Figure 6.2 shows 50 extracted surface to volume ratios from intra-axonal simulations with $d = 6$ µm. Figure 6.2a shows solutions obtained by fitting calculated ADCs to Eq. 6.10. Figure 6.2b shows solutions obtained by fitting signals to Eq. 3.84 (with Eq. 6.10 in the exponent). In Figure 6.2a, final solutions depend on the initial parameters. In Figure 6.2b, most of the solutions are the same and do not depend on the initial parameters. The "best" solutions in Figure 6.2a (those defined as having

180

the smallest sum-of-squares) appear to be converging towards the solution in Figure 6.2b. Since we want our solutions to depend as little as possible on the initial parameters, we are going to use the second method from now on and fit signal data directly to Eq. 3.84.

6.3.1.1   Intra-axonal Simulations

We begin with the intra-axonal simulations. Figure 6.3a shows fitted diffusion coefficients as a function of cylinder diameter. With small cylinder diameters ($d < 4$ µm), fitted D_0 are less than the actual value (1.0 µm^2/ms). In smaller pores, where the short-time condition is not met, Equation 6.10 does not hold. In larger diameters, fitted D_0 converge toward 1.0 µm^2/ms. At $d = 1$ µm, the fitted D_0 is only 0.25 µm^2/ms, while at $d = 4$ µm, fitted D_0 is around 1.07 µm^2/ms. For cylinders with $d > 3$ µm, the results have leveled off and the average fitted D_0 is around 1.07 µm^2/ms.

Figure 6.3b shows fitted S/V as a function of cylinder diameter. Fitted S/V at each diameter are all larger than the actual value. The average percent difference over all cylinder diameters is 36 percent. Absolute difference between fitted S/V and true S/V decreases from around 1 µm^{-1} at $d = 1$ µm to 0.12 µm^{-1} at $d = 10$ µm. Since fitted D_0 were underestimated for $d < 3$ µm, we should probably use caution interpreting the S/V results for those diameters, even if they are close. Figure 6.3c directly compares fitted S/V with true S/V. Note that all S/V are overestimated.

6.3.1.2   Full Simulation (Intra- and Extra-axonal water) - Unequal Diffusivities

Next, we consider the case of identical cylinders on a square lattice.

Figure 6.4a shows fitted D_0 as a function of diameter for each packing fraction. Fitted D_0 increase with diameter before starting to level off at higher diameters.

Figure 6.4c shows fitted S/V as a function of diameter with each packing fraction. All fitted S/V are underestimated, especially at small diameters. At larger diameters ($d \rightarrow 10$ µm), fitted S/V become closer to the actual values, almost converging (percent differences are 5 to 10 percent at $d = 10$ µm). Figure 6.4e compares fitted S/V with true S/V for the entire collection of simulations. Fitted S/V for each packing fraction have slightly different behaviour. Fitted
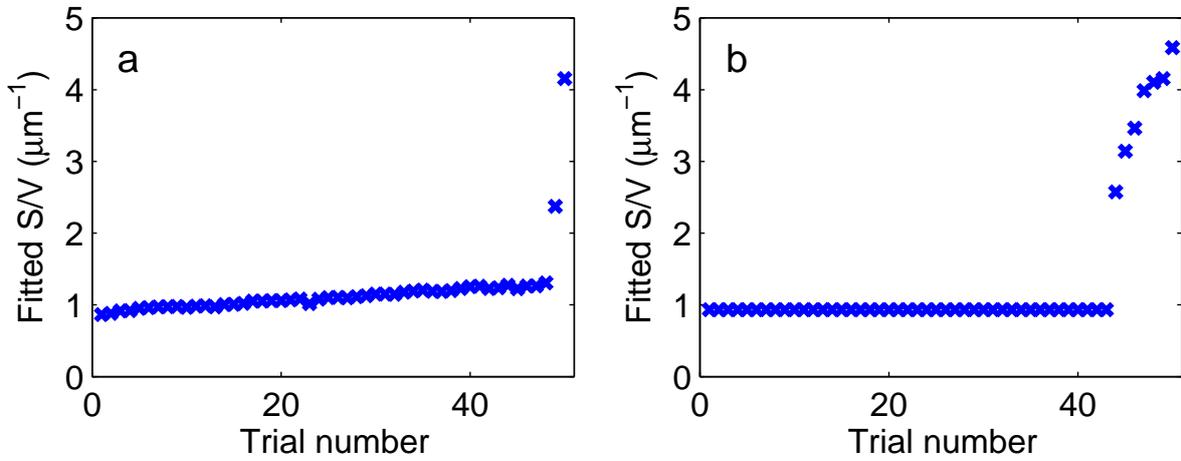
Figure 6.2: Fitted surface to volume ratios obtained using 50 randomized initial parameters, sorted by sum-of-squares (shown in ascending order), using intra-axonal simulations ($d = 6$ μm). Solutions were obtained by (a) fitting calculated *ADC*s to Eq. 6.10 or (b) fitting signals directly to Eq. 3.84. Results obtained by fitting signals to Eq. 3.84 show less dependence on initial parameters.
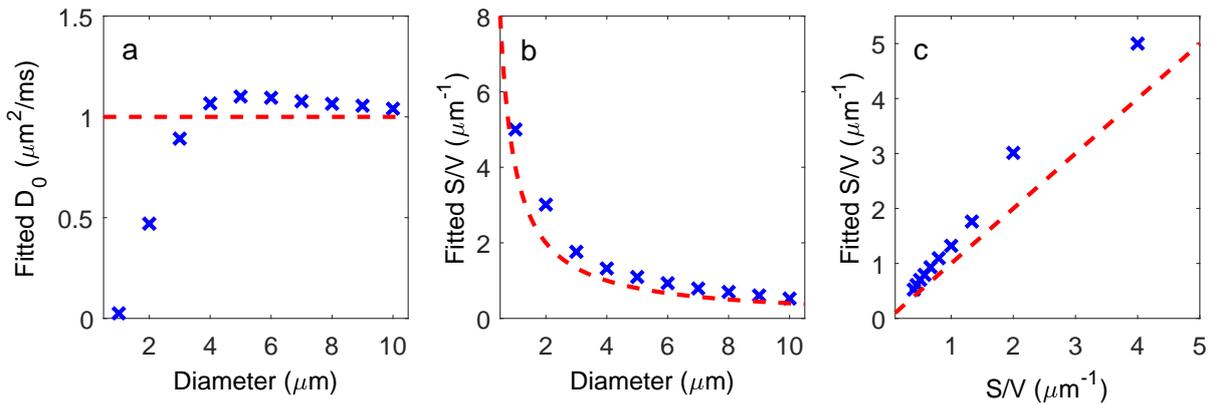
Figure 6.3: Fitted model parameters using intra-axonal simulation data. (a) Fitted diffusion coefficients as a function of cylinder diameter. (b) Fitted surface to volume ratios as a function of cylinder diameter. (c) Comparison of fitted surface to volume ratios with actual surface to volume ratios. Dashed line indicates a perfect fit. Fitted D_0 and fitted S/V are more accurate with larger diameters.

S/V from higher packing fractions (= 0.7) are more accurate and converge towards the true value more quickly. For example, fitted S/V with packing fractions of 0.7 are on average 7 percent closer to the true S/V compared to those with packing fraction = 0.6. Since the model in Eq. 6.10 does not account for compartments with different diffusivities, these are not as accurate.

Figure 6.5c compares fitted S/V with true S/V for the five diameter distributions. All S/V are underestimated by at least 25 percent.

6.3.1.3 Full Simulation (Intra- and Extra-axonal water) - Equal Diffusivities

We now apply the model on simulations with equal diffusivities in the intra- and extra-axonal regions ($D_{in} = D_{ex}$).

Figure 6.4b shows fitted D_0 as a function of diameter for each packing fraction. Fitted D_0 increase with diameter before leveling off above $d = 4$ µm. When $d \geqslant 5$ µm, relative differences between fitted and actual D_0 are less than 1 percent.

Figure 6.4d shows fitted S/V as a function of diameter for each packing fraction. Fitted S/V are more accurate than in the last section (when $D_{in} \neq D_{ex}$). At smaller diameters, fitted S/V underestimate the true values, while at larger diameters they overestimate them. The crossover point for packing fractions of 0.5 and 0.6 occurs at $d = 5$ µm. When the packing fraction is 0.7, the crossover occurs at $d = 8$ µm. Figure 6.4f compares fitted S/V with true S/V. For $S/V < 1.5$ µm^{-1}, the fitted S/V are within 0.2 µm^{-1} of the true values.

We now look at simulations with diameter distributions. Figure 6.5a-b shows fitted diffusion coefficients from each substrate. All fitted D_0 underestimate the actual D_0. The exact difference depends on the particular substrate. Substrates with higher S/Vs (with more restricted geometries) have smaller fitted D_0, while substrates with a lower S/V have fitted D_0 that are closer to 1.0 µm^2/ms. The fitted diffusion coefficient is around 0.6 µm^2/ms for the distribution with the highest S/V (= 3.3 µm^{-1}), so that the difference between fitted and actual S/V is approximately 40 percent. The fitted diffusion coefficients for the distributions with smaller S/V (= 1.21 µm^{-1} and 1.24 µm^{-1}) are 0.97 µm^2/ms and 0.95 µm^2/ms, respectively, differences of less than 5 percent.
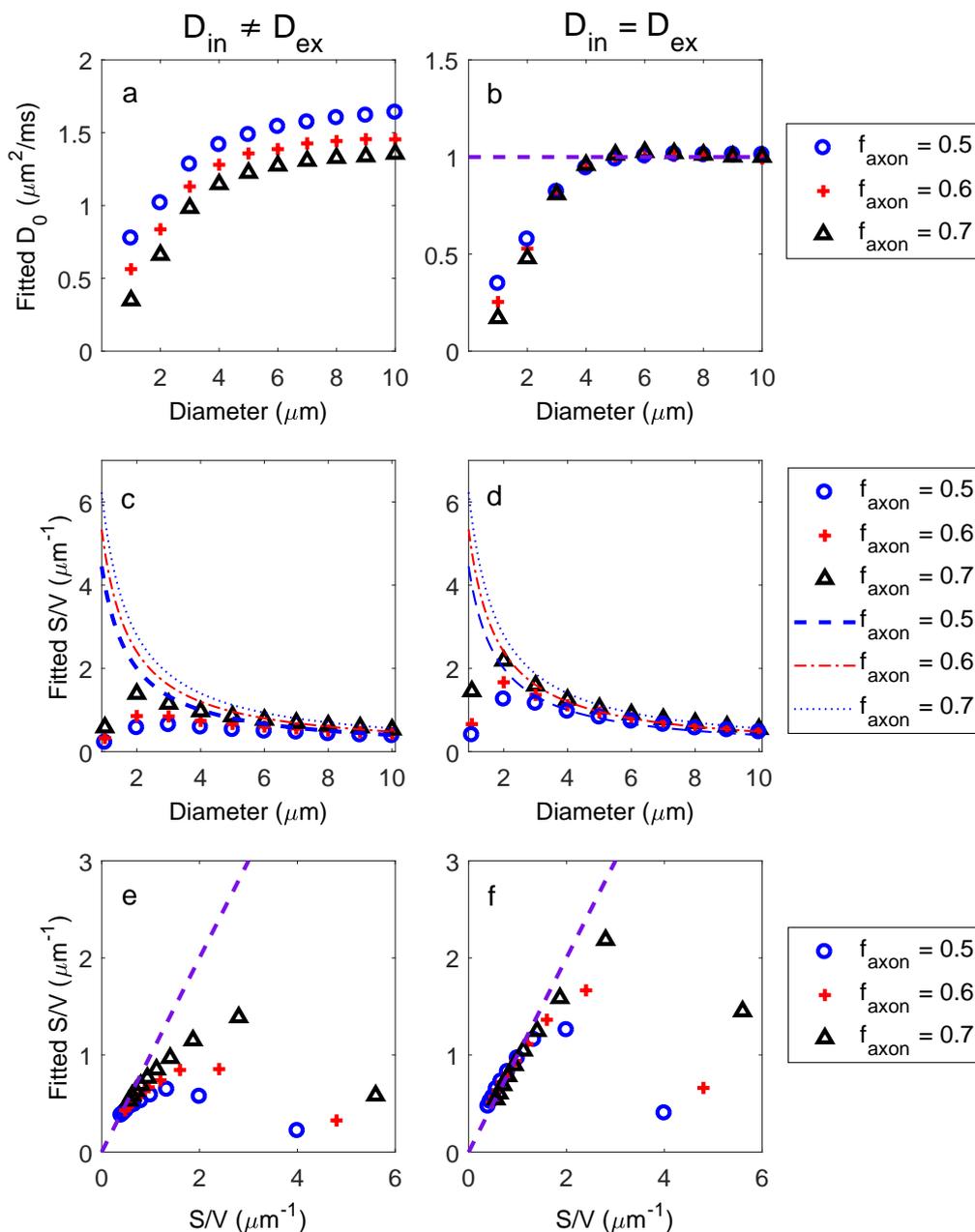
Figure 6.4: Fitted model parameters using simulations with identical cylinders on a square lattice. In (a, c, e), $D_{in} \neq D_{ex}$. In (b, d, f), $D_{in} = D_{ex}$. (a-b) Fitted D_0 as a function of cylinder diameter. (c-d) Fitted S/V as a function of cylinder diameter. (e-f) Comparison of fitted S/V with actual S/V (with all simulations included). Dashed line indicates a perfect fit. Fitted S/V are more accurate at lower S/V and when $D_{in} = D_{ex}$.

Figure 6.5c-d compares fitted S/V with true S/V. The fitted S/Vs all underestimate the true values. The difference between fitted and true S/V is smaller in distributions having lower S/V. For the distribution with the highest S/V (= 3.3 μm^{-1}), where D_0 was underestimated by 40 percent, the difference between fitted and true S/V is around 36 percent. For the distribution with the lowest S/V (= 1.21 μm^{-1}), where there was a 3 percent difference between fitted and true D_0, the difference between fitted and true S/V is around 13 percent. Note the overall similarity with Figure 6.4f. Fitted and true S/V start to converge as the true S/V approaches 1 μm^{-1}.

6.3.2 Effects of Noise

To explore the effects that noise may have on fitted diffusion coefficients and surface to volume ratios, Gaussian noise (SNR = 100) was added to each transverse magnetization component in the simulation data. We saw in the previous section that in some situations (different diffusivities and small diameters), the model (Eq. 6.10) gives inaccurate results. We want to know how noise affects fitted parameters with as little confounding as possible, so we will only use results for the situations that worked well in the previous section. These are the square packed medium-large cylinders ($d > 5$ μm) with a single diffusivity ($D_{in} = D_{ex}$).

6.3.2.1 Overview of Parameters

We begin by presenting an overview of the parameters and their uncertainties. All measurements have been used.

Gradient Set #1 (G_{max} = 900 mT/m)

Figures 6.6a, 6.6d, and 6.6g show mean fitted diffusion coefficients (μ_D) for $d = 6$ to 10 μm. All μ_D are within 2 percent of the actual value. All σ_D are less than 0.04 μm^2/ms. The distributions of fitted D_0 for $d = 6$ μm and $d = 10$ μm (packing fraction = 0.6) are shown in Figures 6.7a and 6.7d. Both distributions are relatively symmetric about $D_0 = 1.0$ μm^2/ms with no outliers (all fitted D_0 are within 0.15 μm^2/ms of 1.0 μm^2/ms).

Figures 6.8a, 6.8d, and 6.8g show mean fitted S/V ($\mu_{S/V}$) for $d = 6$ to 10 μm. Figures 6.9a,
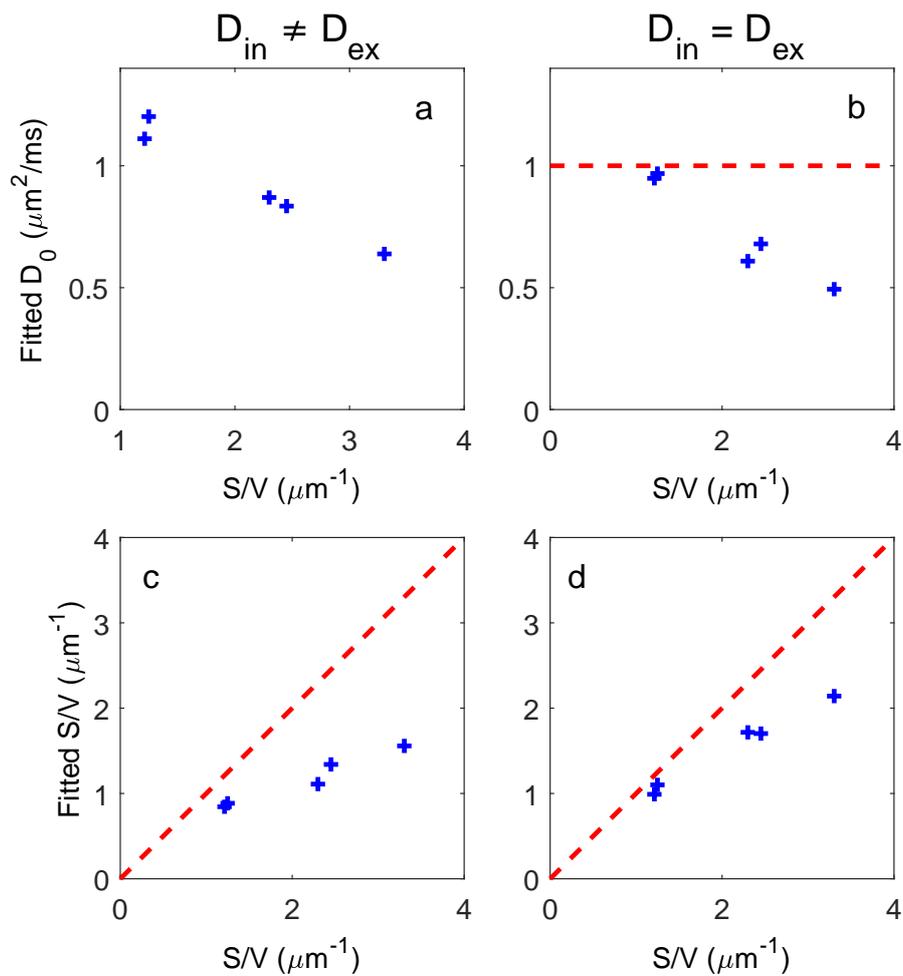
Figure 6.5: Fitted model parameters from simulations with diameter distributions. In (a, c), $D_{in} \neq D_{ex}$. In (b, d), $D_{in} = D_{ex}$. (a-b) Fitted D_0 for the five distributions. (c-d) Comparison of fitted S/V with actual S/V for the five diameter distributions. Dashed line indicates a perfect fit. Fitted S/V are more accurate at lower S/V and when $D_{in} = D_{ex}$.
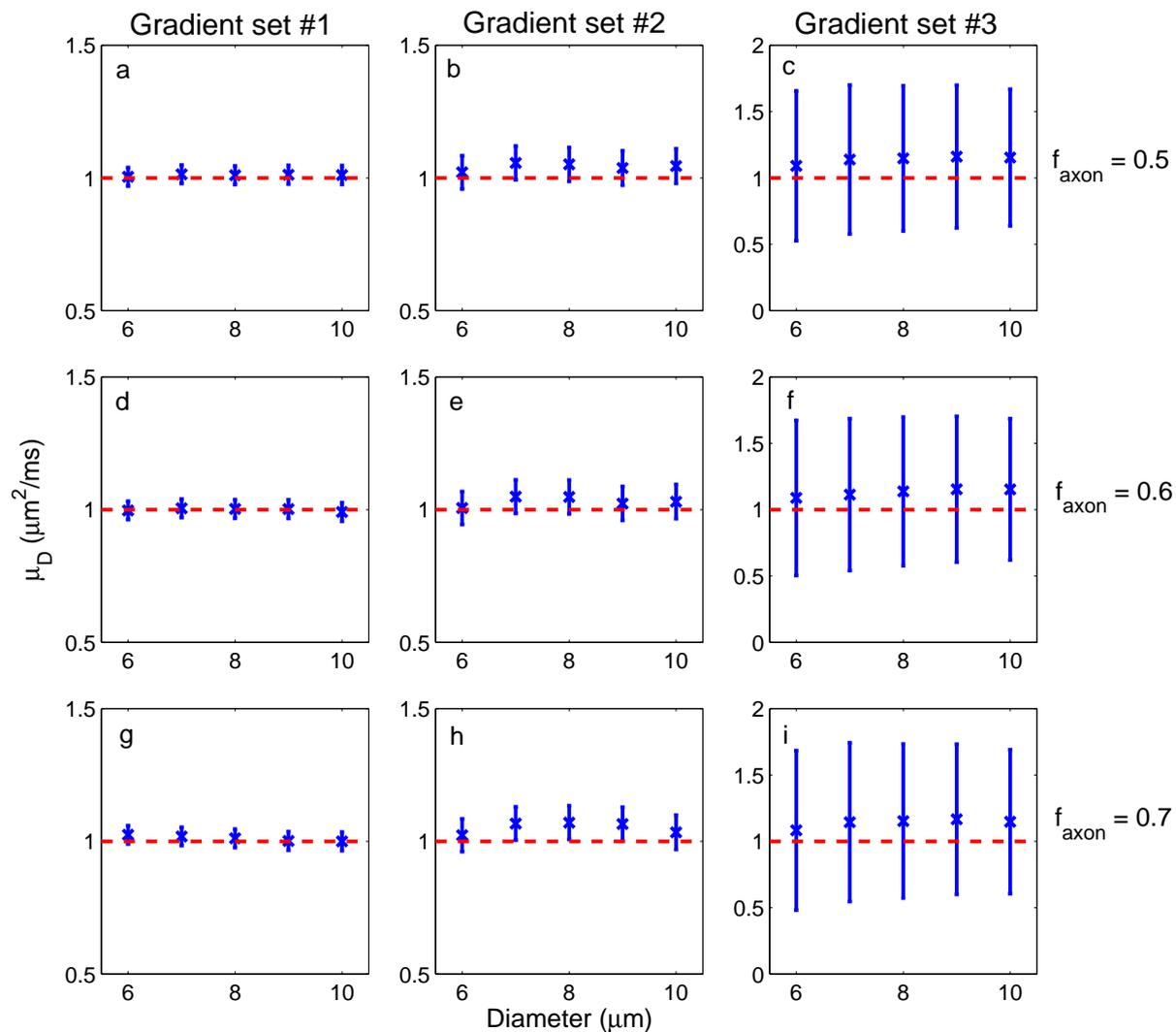
Figure 6.6: μ_D as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted diffusion coefficients (σ_D). Dashed line indicates a perfect fit. Three sets of measurements shown: (a, d, g) $G_{max} = 900$ mT/m (b, d, h) $G_{max} = 300$ T/m (c, f, i) $G_{max} = 80$ mT/m. (a-c) are from simulations with packing fraction 0.5. (d-f) are from simulations with packing fraction 0.6. (g-i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. Higher gradient strengths give more accurate and precise results. **Note**: (c), (f), and (i) use different y-axis limits.
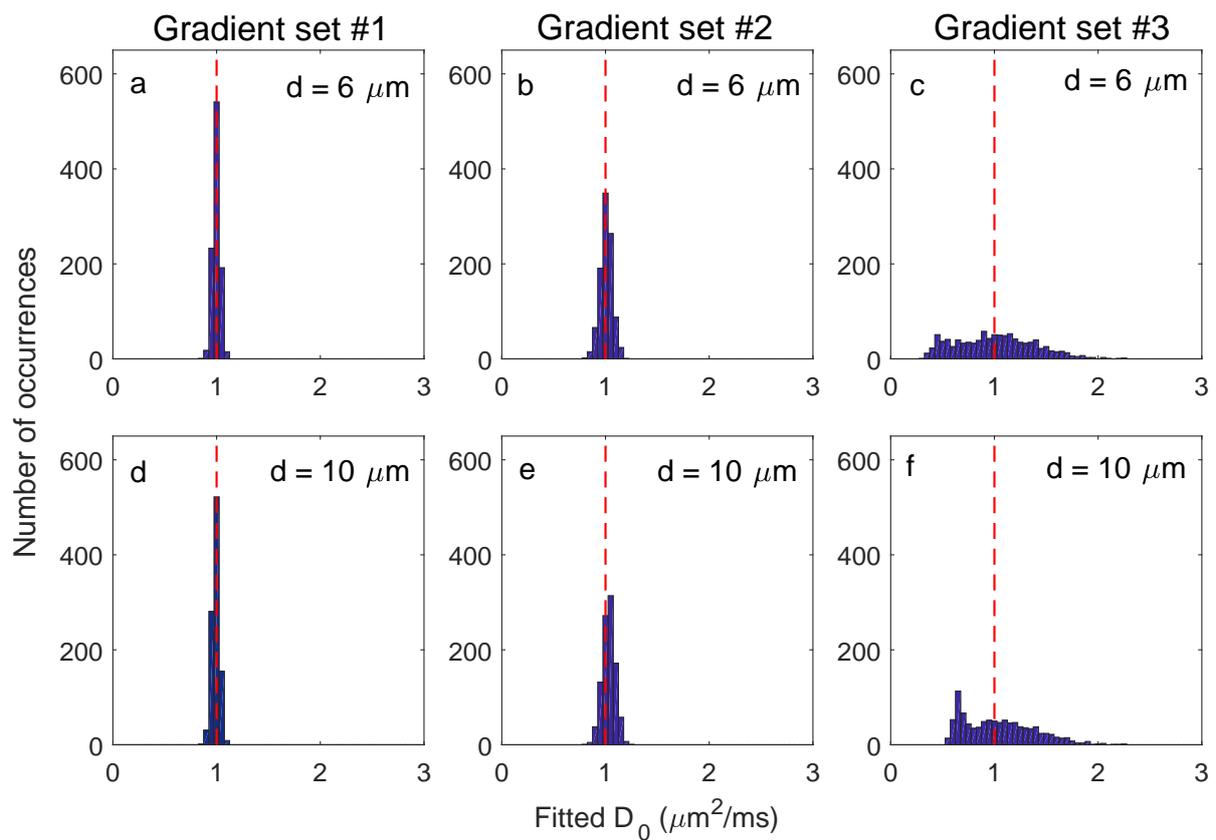
Figure 6.7: Distributions of fitted diffusion coefficients for (a-c) $d = 6$ μm and (d-f) $d = 10$ μm. Simulations use a packing fraction of 0.6. Three sets of measurements shown: (a, d) $G_{max} = 900$ mT/m (b, e) $G_{max} = 300$ mT/m (c, f) $G_{max} = 80$ mT/m. All measurements were kept when fitting to the model. Distributions become narrower with higher gradient strengths.

6.9d, and 6.9g compare $\mu_{S/V}$ to true S/V. Figures 6.10a and 6.10d show the S/V distributions for $d = 6$ μm and $d = 10$ μm. The distributions are slightly skewed towards lower S/Vs (skewness = -0.44 [$d = 10$ μm] and -0.43 [$d = 6$ μm]). In the distribution for $d = 6$ μm, 95% of fitted S/V lie in the range [0.63, 0.94] μm^{-1}. In the distribution for $d = 10$ μm, 95% of fitted S/V lie in the range [0.30, 0.67] μm^{-1}.

Gradient Set #2 (G_{max} = 300 mT/m)

Figures 6.6b, 6.6e, and 6.6h show μ_D for $d = 6$ to 10 μm. All μ_D are within 7 percent of the actual value, though all overestimate it. All σ_D are less than 0.07 μm^2/ms. The distributions of fitted D_0 for $d = 6$ μm and $d = 10$ μm (packing fraction = 0.6) are shown in Figures 6.7b and 6.7e. Both distributions are symmetric about $D_0 = 1.0$ μm^2/ms.

Figures 6.8b, 6.8e, and 6.8h show $\mu_{S/V}$ for $d = 6$ to 10 μm. Figures 6.9b, 6.9e, and 6.9h compare $\mu_{S/V}$ to true S/V. Figures 6.10b and 6.10e show the S/V distributions for $d = 6$ μm and $d = 10$ μm. Both distributions are more skewed towards lower S/Vs than the $G_{max} = 900$ mT/m results (skewness = -0.57 [$d = 10$ μm] and -0.56 [$d = 6$ μm]). In the distribution for $d = 6$ μm, 95% of fitted S/V lie in the range [0.80, 0.86] μm^{-1}. In the distribution for $d = 10$ μm, 95% of fitted S/V lie in the range [0.44, 0.61] μm^{-1}. This means that the S/V distributions for the $G_{max} = 300$ mT/m measurements are actually narrower than those from the $G_{max} = 900$ mT/m measurements.

Gradient Set #3 (G_{max} = 80 mT/m)

Figures 6.6c, 6.6f, and 6.6i show μ_D for $d = 6$ to 10 μm. All μ_D overestimate D_0 by 8 to 16 percent. The σ_D are between 0.5 and 0.6 μm^2/ms. The distributions of fitted D_0 for $d = 6$ μm and $d = 10$ μm (packing fraction = 0.6) are shown in Figures 6.7c and 6.7f. Both distributions are skewed to the right. Although both distributions are much wider than the ones in Figures 6.7b and 6.7e, all fitted D_0 still lie within the parameter bounds ([0, 3] μm^2/ms). The distribution for $d = 6$ μm is bimodal, with a bigger peak around 0.5 μm^2/ms and a smaller one near 1 μm^2/ms. The distribution for $d = 10$ μm shows both peaks beginning to merge, with the combined peak near $D_0 = 0.6$ μm^2/ms.
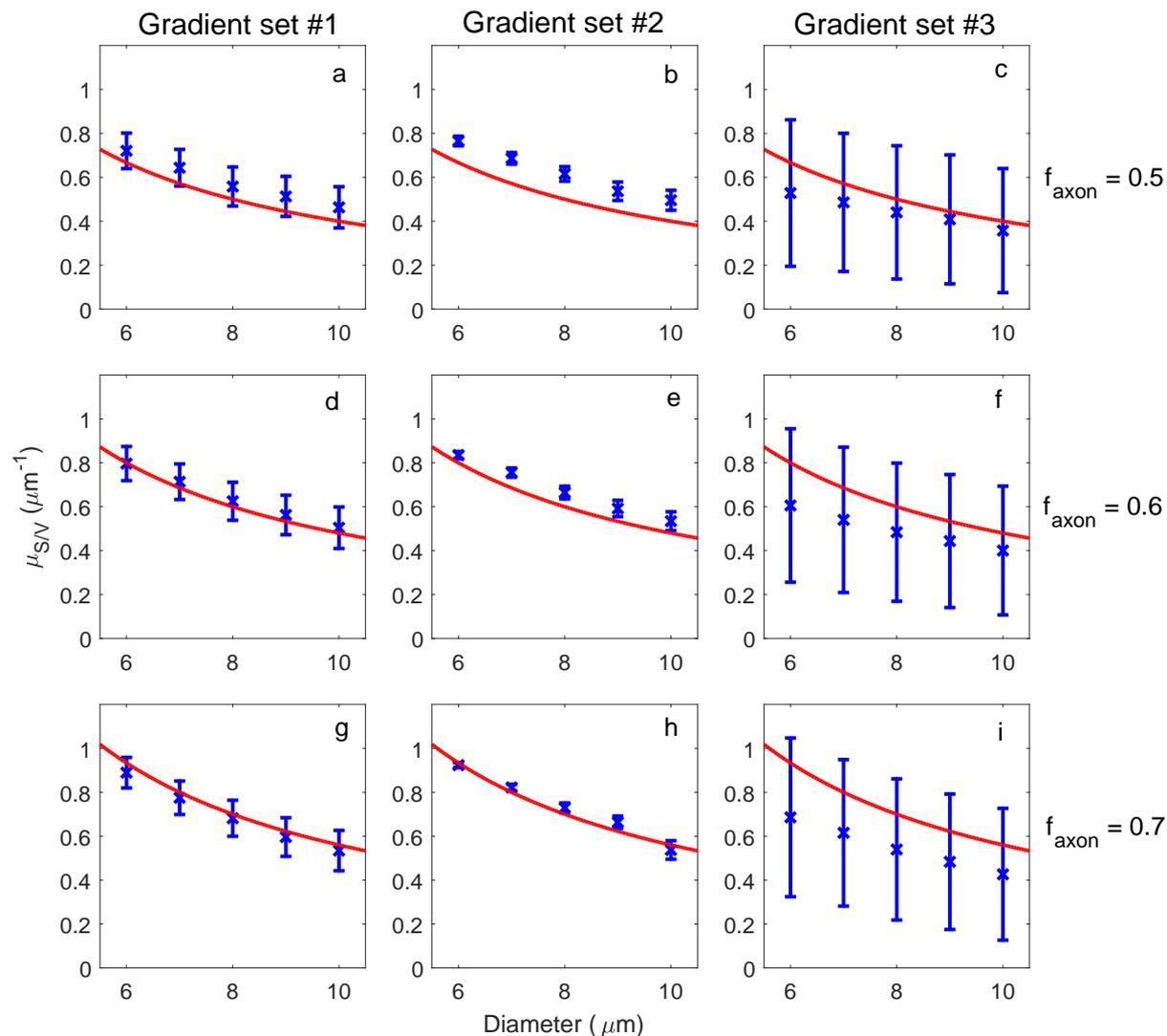
Figure 6.8: $\mu_{S/V}$ as a function of cylinder diameter. Error bars represent the standard deviation of the distribution of fitted surface to volume ratios ($\sigma_{S/V}$). Dashed lines denote the actual S/V. Three sets of measurements shown: (a, d, g) $G_{max} = 900$ mT/m (b, d, h) $G_{max} = 300$ mT/m (c, f, i) $G_{max} = 80$ mT/m. (a-c) are from simulations with packing fraction 0.5. (d-f) are from simulations with packing fraction 0.6. (g-i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. Fitted S/V for $G_{max} = 300$ mT/m are most precise, but sometimes have an upward bias. For $G_{max} = 80$ mT/m, $\mu_{S/V}$ are smaller than the true S/V, but Figure 6.10 shows that the peak of the distribution may sometimes be higher than the true S/V.
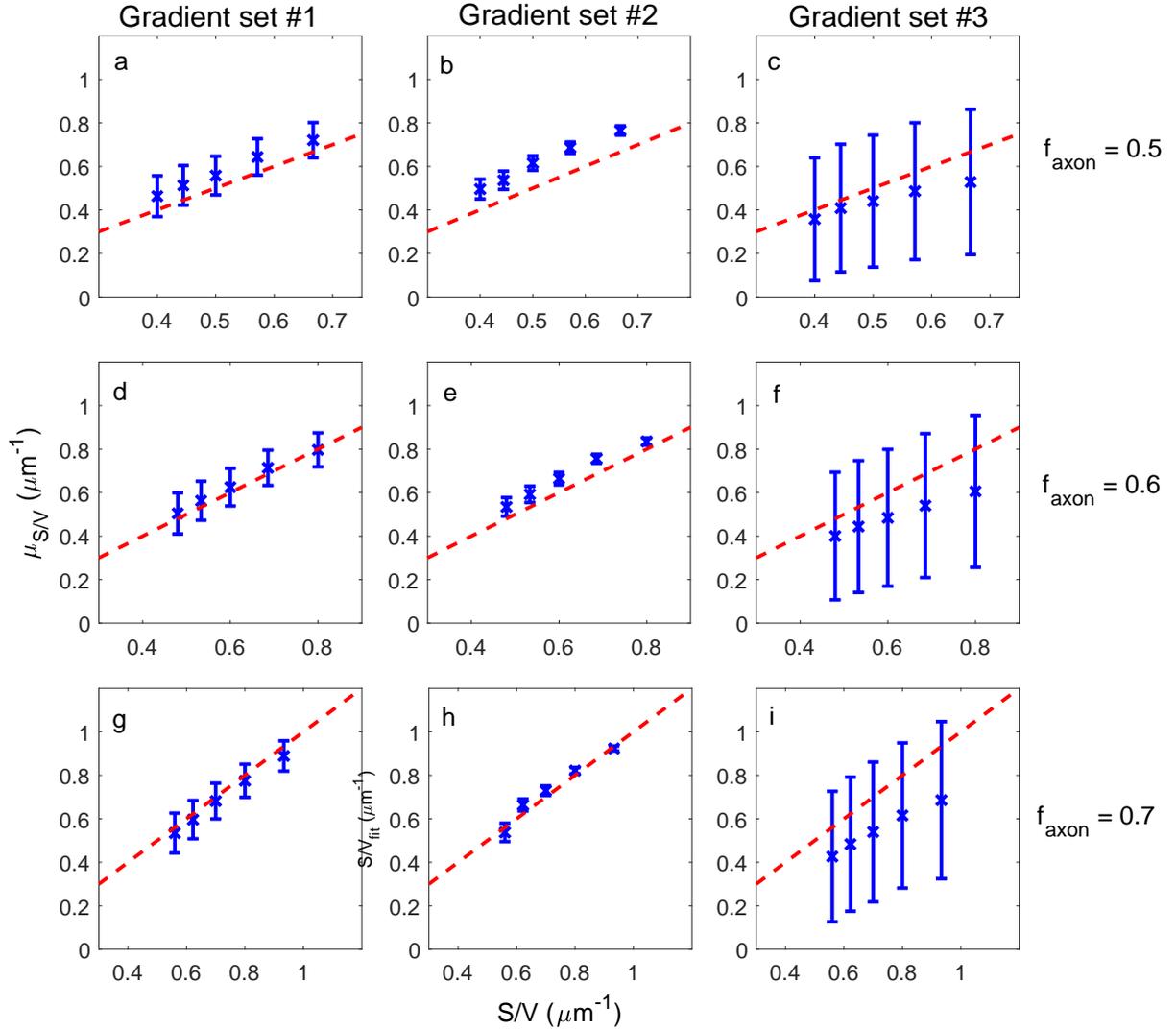
Figure 6.9: $\mu_{S/V}$ as a function of true S/V. Error bars represent the standard deviation of the distribution of fitted surface to volume ratios ($\sigma_{S/V}$). Dashed lines denote actual S/V. Three sets of measurements shown: (a, d, g) $G_{max} = 900$ mT/m (b, d, h) $G_{max} = 300$ mT/m (c, f, i) $G_{max} = 80$ mT/m. (a-c) are from simulations with packing fraction 0.5. (d-f) are from simulations with packing fraction 0.6. (g-i) are from simulations with packing fraction 0.7. All measurements in each set were kept when fitting to the model. Fitted S/V for $G_{max} = 300$ mT/m are most precise, but sometimes have an upward bias.
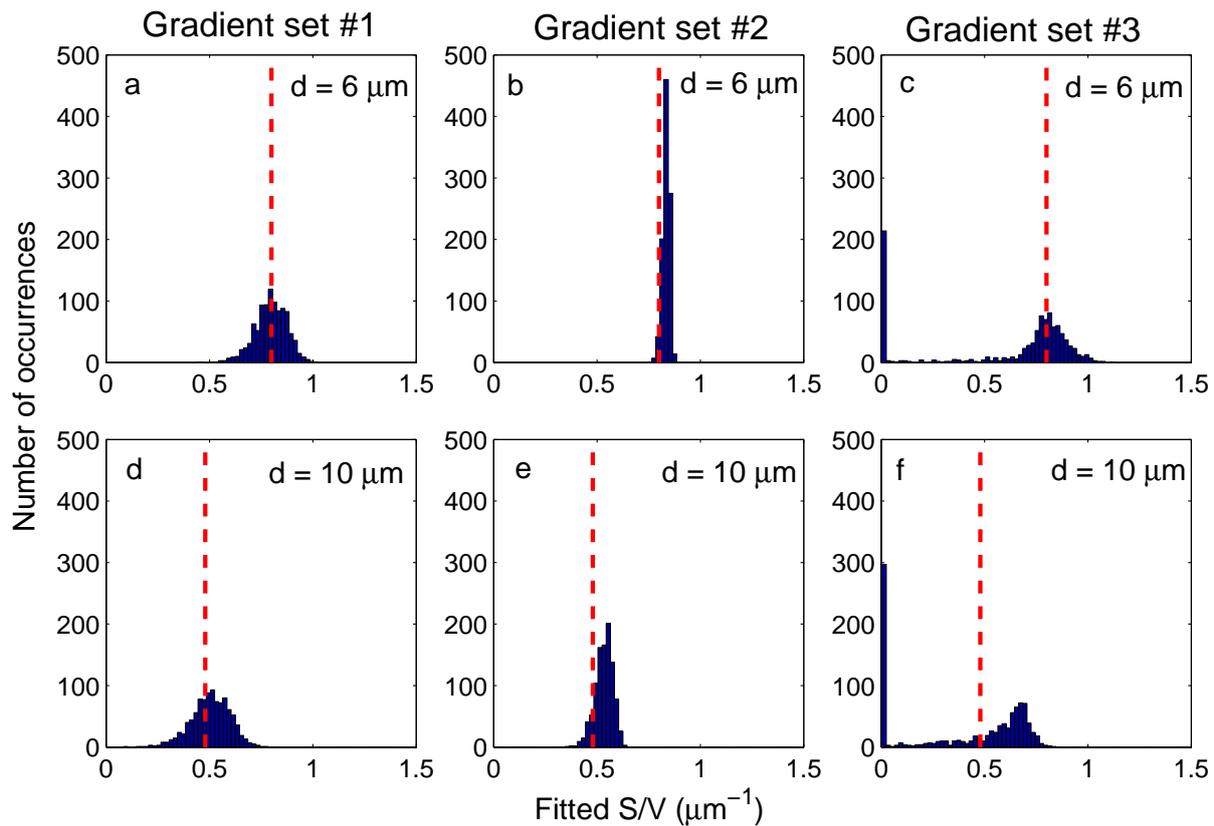
Figure 6.10: Distributions of fitted surface to volume ratios for (a-c) $d = 6$ µm and (d-f) $d = 10$ µm. Simulations use a packing fraction of 0.6. Three sets of measurements shown: (a, d) $G_{max} = 900$ mT/m (b, e) $G_{max} = 300$ mT/m (c, f) $G_{max} = 80$ mT/m. All measurements were kept when fitting to the model. Distributions are accurate at $d = 6$ µm, but the peaks are shifted to higher S/V at $d = 10$ µm. Distributions are narrowest with $G_{max} = 300$ mT/m.

Figures 6.8c, 6.8f, and 6.8i show $\mu_{S/V}$ for $d = 6$ to 10 μm. Figures 6.9c, 6.9f, and 6.9i compare $\mu_{S/V}$ to true S/V. Figures 6.10c and 6.10f show the S/V distributions for $d = 6$ μm and $d = 10$ μm. Both distributions are skewed towards smaller S/V. In the distribution for $d = 6$ μm, some fitted S/Vs have clustered near $S/V = 0$ ($209/1000 < 0.001$ μm^{-1}). The peak of the distribution is around 0.8 μm^{-1}. In the distribution for $d = 10$ μm, many fitted S/Vs are clustered near $S/V = 0$ ($291/1000 < 0.001$ μm^{-1}). The peak is between 0.6 and 0.7 μm^{-1}. For $d = 6$ μm, the peak is located near the true S/V. But for $d = 10$, the peak is ≈ 0.1 to 0.2 μm^{-1} to the right of the true S/V.

6.3.2.2 Gradient Subsets

Here we show how $\sigma_{S/V}$ changes when gradients are removed from the data. For the next three sections, we keep all frequencies ($n_f = 20$) and work with data from simulations with packing fraction = 0.6.

Gradient Set #1 (G_{max} = 900 mT/m)

Figure 6.11a shows $\sigma_{S/V}$ for $d = 6$ to 10 μm using the gradient subsets in Table 5.1. There is a steady rise in $\sigma_{S/V}$ as fewer and smaller gradients are used. One nonzero gradient measurement ($G = \{0, G = G_n\}$) can potentially have a smaller $\sigma_{S/V}$ than using multiple gradient strengths ($G = \{0, G_1, \ldots, G_{n-1}\}$), provided that $G_n > G_{n-1}$. At a given subset, $\sigma_{S/V}$ for different diameters are within approximately 0.04 μm^{-1} of each other. At the lower subsets, $\sigma_{S/V}$ increases with diameter, so that $\sigma_{S/V}$ for $d = 6$ μm is the smallest and $\sigma_{S/V}$ for $d = 10$ μm is the largest. At higher subsets, the opposite occurs. In subsets #5 to 7 and above, $d = 6$ μm has the largest $\sigma_{S/V}$ and $d = 10$ μm has the smallest.

Gradient Set #2 (G_{max} = 300 mT/m)

6.11b shows $\sigma_{S/V}$ for $d = 6$ to 10 μm using the gradient subsets in Table 5.1. There is a steady rise in $\sigma_{S/V}$ as fewer and smaller gradients are used. Again, one nonzero gradient measurement ($G = 0, G = G_n$) can potentially give smaller $\sigma_{S/V}$ compared to using multiple gradient strengths ($G = \{0, G_1, \ldots, G_{n-1}\}$), provided that $G_n > G_{n-1}$. We also saw this with the

$G_{max} = 900$ mT/m measurements. The difference in $\sigma_{S/V}$ between consecutive subsets appears to get larger as we move to higher subsets (with exceptions for subsets in the #9 to 11 range). In subsets #1 to 7, $d = 6$ μm has the smallest $\sigma_{S/V}$ and $d = 10$ μm has the largest. In subsets #8 to 11 (maximum gradient strengths $\leqslant 150$ mT/m), the order is reversed: $d = 6$ μm has the largest $\sigma_{S/V}$ and $d = 10$ μm has the smallest.

We saw in the last section that the width of the S/V distributions for $d = 6$ and 10 μm cylinders for the $G_{max} = 300$ mT/m measurements were actually narrower than those for the $G_{max} = 900$ mT/m measurements (Figures 6.10a and 6.10d). Here we see that it is true for more than one subset. Subsets #1 to 4 (possibly even #5) have smaller $\sigma_{S/V}$ than subset #1 of the first gradient set (Figure 6.11a). Subset #5 here uses a maximum gradient strength of 200 mT/m. Very high gradient strengths (≈ 900 mT/m) do not provide any improvement and may even be worse.

Gradient Set #3 (G_{max} = 80 mT/m)

Figure 6.11c shows $\sigma_{S/V}$ for $d = 6$ to 10 μm using the gradient subsets in Table 5.1. All $\sigma_{S/V}$ are larger when using only one nonzero gradient, although the difference is only 0.02 to 0.04 μm^{-1}. For the full picture, we should view the S/V distributions for each subset. Figures 6.12a and 6.12b show the S/V distribution for $d = 8$ μm. With one nonzero gradient, around 30 percent of the fitted S/Vs are near zero ($327/1000 < 0.001$ μm^{-1}). Compared with the distribution in Figure 6.12a, which uses all the gradients, the distribution peak here is less pronounced and the distribution is a little wider, extending ≈ 0.1 μm^{-1} farther to the right. Otherwise, the peaks of each distribution are in similar locations (≈ 0.7 to 0.8 μm^{-1}).

6.3.2.3   Frequency Removal

Since the fitted values tended to be more precise when using both higher and more gradients, we will study the effect of varying the number of frequencies while keeping all of the gradient strengths in each of the three sets.
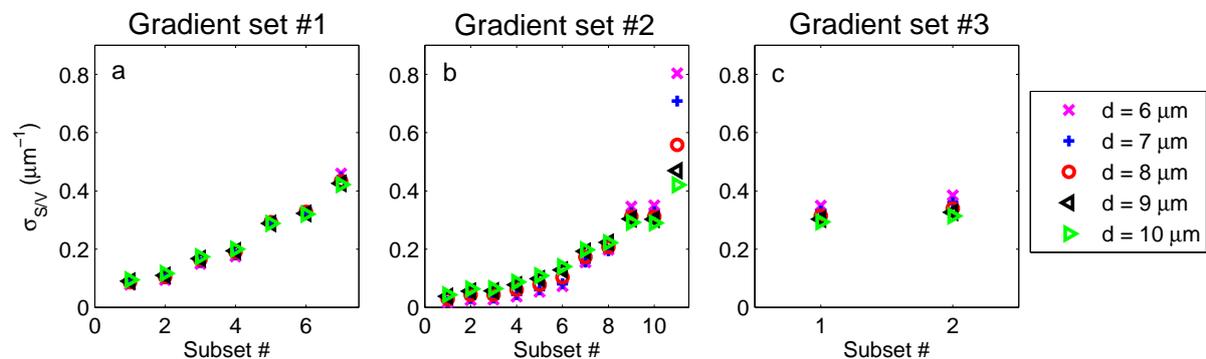
Figure 6.11: $\sigma_{S/V}$ for different gradient subsets. Simulations use a packing fraction of 0.6. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All frequencies in each set were kept when fitting to the model. For $G_{max} = 900$ mT/m and $G_{max} = 300$ mT/m, $\sigma_{S/V}$ increases as the maximum gradient strength in the subset decreases. For the simulations with $G_{max} = 80$ mT/m, $\sigma_{S/V}$ are slightly larger for the two gradient subset.
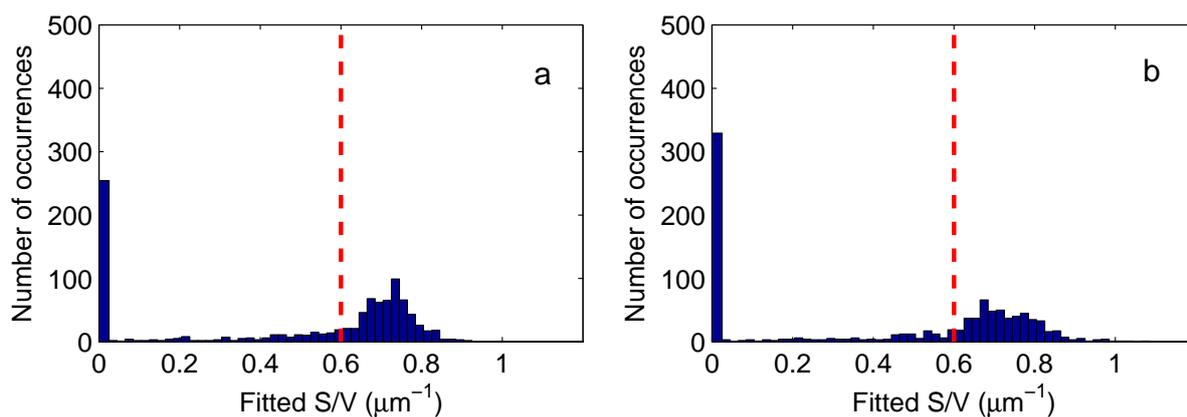
Figure 6.12: Distribution of fitted surface to volume ratios for $d = 8$ μm using (a) all gradient strengths and (b) two gradients ($G = 0$, $G_{max} = 80$ mT/m). Simulations use packing fraction of 0.6. All frequencies were kept when fitting to the model. Both distributions are shifted higher than the true S/V, but the distribution from the two gradient subset is flatter.

Gradient Set #1 (G_{max} = **900 mT/m**)

There are a few ways to remove frequencies from the data. One involves keeping every *i*th frequency (for example, keeping every second frequency measurement and removing the rest). Another is to successively remove the highest frequencies first. We begin with the first of these methods.

Figure 6.13 shows $\sigma_{S/V}$ as a function of n_f. For $n_f = 10$, every second frequency was kept, meaning f_m = {100, 200, 300, 400, 500, 600, 700, 800, 900, 1000} Hz. For $n_f = 5$, every fourth frequency was kept. This means f_m = {200, 400, 600, 800, 1000} Hz. The $\sigma_{S/V}$ are larger at smaller n_f. At $n_f = 20$, all $\sigma_{S/V}$ are around 0.1 μm^{-1}, while at $n_f = 5$, they are almost 0.3 μm^{-1}. Since this is a difference of nearly 0.2 μm^{-1}, we should see how the fitted S/V distributions change. Figure 6.14a shows the S/V distribution for $d = 8$ μm using $n_f = 20$. Most fitted S/V (95%) are between 0.44 μm^{-1} and 0.78 μm^{-1}. Figure 6.14b shows the distribution using $n_f = 10$. The distribution is much wider, with 95% of fitted S/V between 0.06 μm^{-1} and 0.94 μm^{-1}. The distribution using $n_f = 5$ is much worse (Figure 6.14c). In this distribution, 95% of fitted S/V lie in the range [0, 1.0] μm^{-1}, with some clustering near $S/V = 0$ (52/1000 < 0.001 μm^{-1}).

Next, we compare the previous method with the method of only removing the highest frequencies. Figure 6.15a shows $\sigma_{S/V}$ as a function of n_f. In this case, the highest frequencies were removed. The $\sigma_{S/V}$ decrease 0.01 to 0.03 μm^{-1} between $n_f = 5$ and $n_f = 20$, which is a much smaller change than in Figure 6.13. To make sure that nothing else has changed in the S/V distributions themselves, we will compare them. Figures 6.14d and 6.14e show the distributions for $d = 8$ μm with $n_f = 10$ and $n_f = 5$. Both distributions have similar shapes and similar widths. With $n_f = 10$, 95% of fitted S/V lie in the range [0.43, 0.81] μm^{-1}. With $n_f = 5$, 95% lie in the range [0.44, 0.82] μm^{-1}. Both distributions are much narrower than those in Figures 6.14b and 6.14c.

Gradient Set #2 (G_{max} = **300 mT/m**)

Since we found that removing higher frequencies is the preferred method of lowering n_f, we will not show $\sigma_{S/V}$ as a function n_f from the first method here. Instead, we will give an example
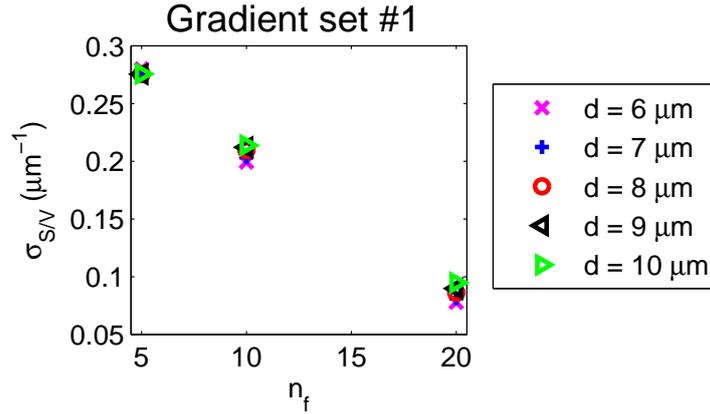
Figure 6.13: $\sigma_{S/V}$ as a function of n_f. For $n_f = 10$, every second frequency was kept. The ten frequencies were 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 Hz. For $n_f = 5$, every fourth frequency was kept. The five frequencies were 200, 400, 600, 800, and 1000 Hz. Simulations use the $G_{max} = 900$ mT/m set of measurements with a packing fraction of 0.6. All gradient strengths in each set were kept when fitting to the model. Each $\sigma_{S/V}$ is much higher at $n_f = 10$ and $n_f = 5$ compared to $n_f = 20$.

of how S/V distributions change. Figures 6.16a-c compare S/V distributions for $d = 8$ μm using $n_f = 5$, 10, and 20. In the distribution for $n_f = 10$, most (95%) of the fitted S/V are in the range [0, 0.90] μm^{-1}. In the distribution for $n_f = 5$, many fitted S/V are near zero ($445/1000 < 0.001$ μm^{-1}). The distribution is also more spread out (95% of S/Vs in the range [0, 1.43] μm^{-1}) and its peak is now at 1.3 μm^{-1}.

Figure 6.15b shows $\sigma_{S/V}$ as a function of n_f. In this case, the highest frequencies were removed. For a given diameter, $\sigma_{S/V}$ remains within a range of 0.001 μm^{-1} over n_f. If we compare Figure 6.15b with Figure 6.15a, we see that $\sigma_{S/V}$ are smaller here ($\sigma_{S/V} = 0.02$ to 0.04 μm^{-1}) than they are with the $G_{max} = 900$ mT/m measurements ($\sigma_{S/V} = 0.08$ to 0.12 μm^{-1}). Looking at the S/V distributions themselves, we see that there is little difference between using $n_f = 5$ and $n_f = 20$ (see Figure 6.16d-e for examples).

Gradient Set #3 (G_{max} = 80 mT/m)

Figures 6.17a-c compare S/V distributions for $d = 8$ μm using $n_f = 5$, 10, and 20. In the distribution for $n_f = 20$, 95% of the fitted S/V are in the range [0, 0.83] μm^{-1}, with a peak at approximately 0.7 μm^{-1}. In the distribution for $n_f = 10$, most (95%) of the fitted S/V are in the range [0, 1.38] μm^{-1}. The distribution is also shifted to higher S/V, with a peak near 1 μm^{-1}. Many fitted S/V have clustered near $S/V = 0$ ($445/1000 < 0.001$ μm^{-1}). In the distribution
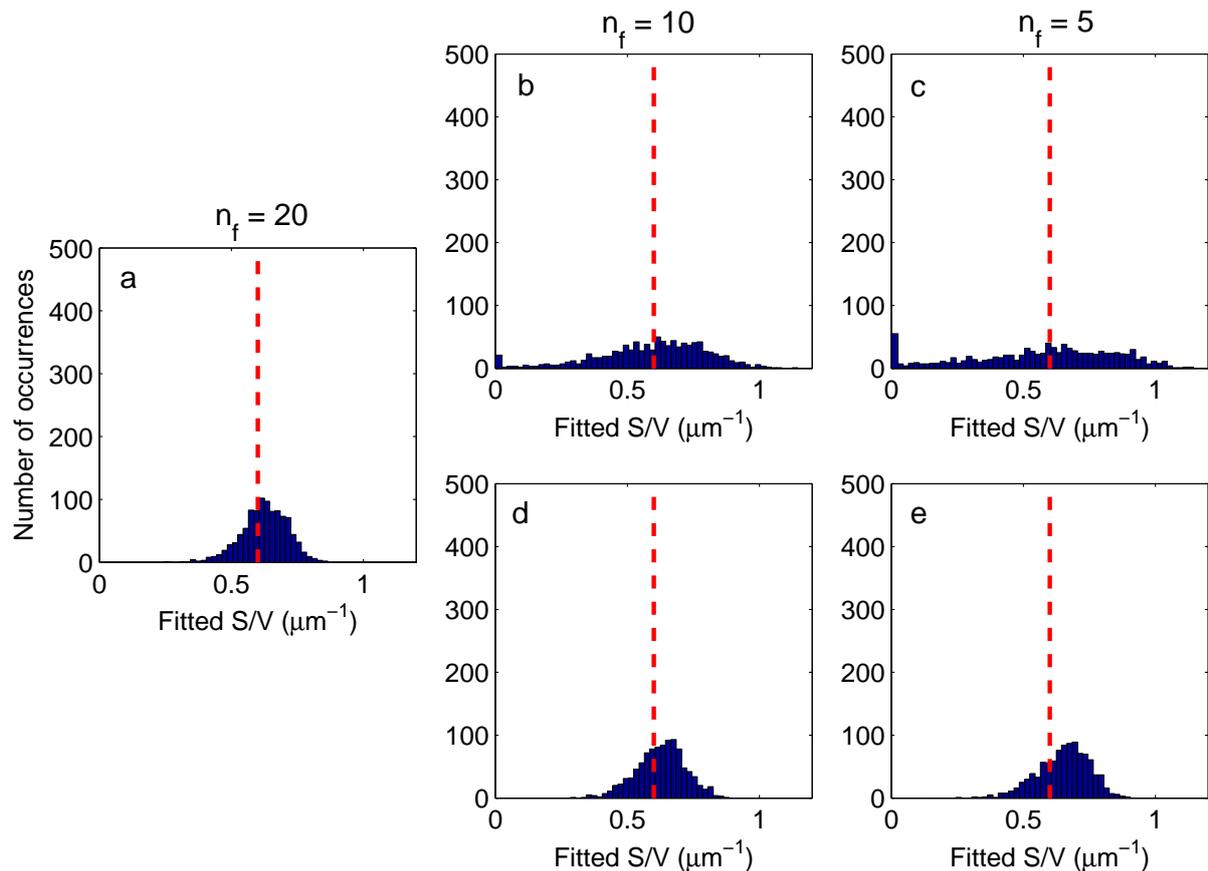
Figure 6.14: Distribution of fitted surface to volume ratios for $d = 8$ µm with (a) $n_f = 20$ (b) $n_f = 10$ (c) $n_f = 5$ (d) $n_f = 10$ (e) $n_f = 5$. In (b), the ten frequencies were 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 Hz. In (c), the five frequencies were 200, 400, 600, 800, and 1000 Hz. In (d-e), the highest frequencies were removed. In (d), the ten frequencies were 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500 Hz. In (e), the five frequencies were 50, 100, 150, 200, and 250 Hz. Simulations used a packing fraction of 0.6 and the $G_{max} = 900$ mT/m set of measurements. All gradient strengths were kept when fitting to the model. Distributions obtained when high frequencies were removed (d-e) are much narrower than when intermediate frequencies were removed (b-c).
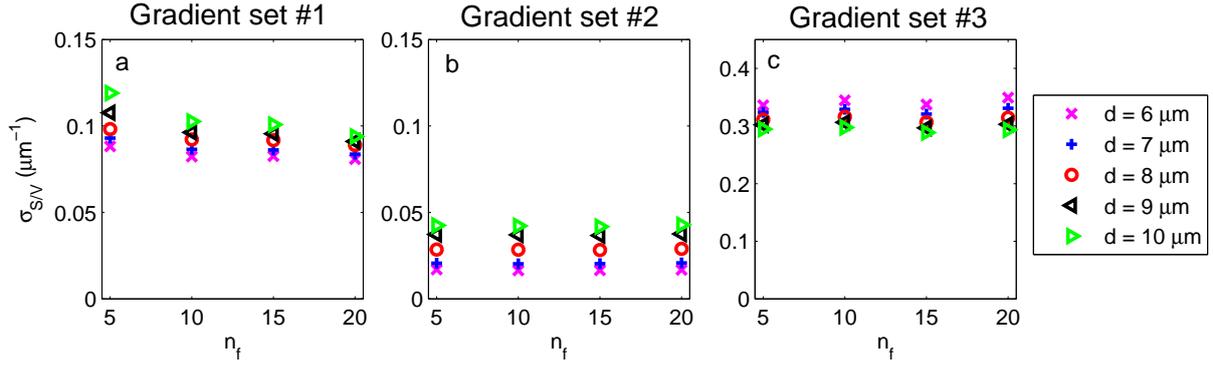
Figure 6.15: $\sigma_{S/V}$ as a function of n_f. Unlike Figure 6.13, the highest frequencies were removed. For $n_f = 10$, $f_m = \{50, 100, 150, 200, 250, 300, 350, 400, 450, 500\}$ Hz. For $n_f = 5$, $f_m = \{50, 100, 150, 200, 250\}$ Hz. Simulations use a packing fraction of 0.6. Three sets of measurements shown: (a) $G_{max} = 900$ mT/m (b) $G_{max} = 300$ mT/m (c) $G_{max} = 80$ mT/m. All gradient strengths were kept when fitting to the model. In contrast to Figure 6.13, $\sigma_{S/V}$ shows little dependence on n_f. **Note**: (c) uses different y-axis limits.

for $n_f = 5$, the peak is closer to 1.5 μm^{-1}. The spread in fitted S/V goes above 2 μm^{-1}, with some hitting the upper bound at $S/V = 5$ μm^{-1} ($161/1000 > 4.9$ μm^{-1}).

Figure 6.15c shows errors as a function of n_f. In this case, the highest frequencies were removed. For a given diameter, $\sigma_{S/V}$ remains within a range of 0.02 μm^{-1} over n_f. Looking at the S/V distributions themselves, we see that there is little difference between using $n_f = 5$ and $n_f = 20$ (Figures 6.17d-e).

6.3.2.4 Repeated Gradients

In the previous sections, the best results were found when keeping all of the gradients. We now compare what happens when instead of spacing gradient measurements out between $G = 0$ and $G = G_{max}$, we repeat measurements at G_{max}. When comparing, we will keep the total number of measurements constant, so that if there were four nonzero gradients before ($G = \{0, G_1, G_2, G_3, G_4\}$), there will still be four nonzero gradients ($G = \{0, G_4, G_4, G_4, G_4\}$). We will use $\sigma_{S/V}$ to compare the different results. For each gradient set, we will show $\sigma_{S/V}$ for three gradient subsets. One choice uses all the original gradients, while another uses repeated measurements at $G = G_{max}$ (while keeping the number of measurements constant). For comparison, we also include the subset with one nonzero gradient ($G = \{0, G_{max}\}$).
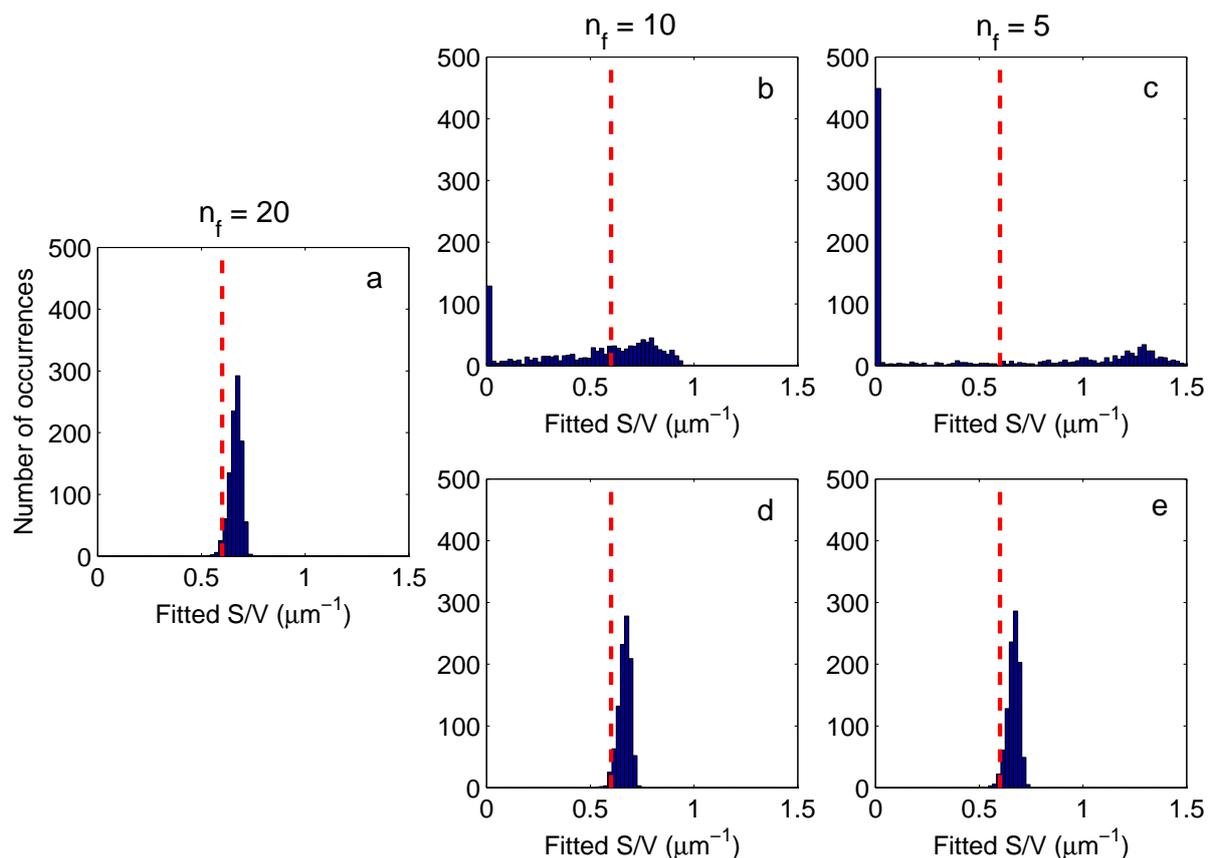
Figure 6.16: Distribution of fitted surface to volume ratios for $d = 8$ μm with (a) $n_f = 20$, (b) $n_f = 10$, (c) $n_f = 5$, (d) $n_f = 10$, (e) $n_f = 5$. In (b), the ten frequencies were 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 Hz. In (c), the five frequencies were 200, 400, 600, 800, and 1000 Hz. In (d-e), the highest frequencies were removed. In (d), the ten frequencies were 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500 Hz. In (e), the five frequencies were 50, 100, 150, 200, and 250 Hz. Simulations used a packing fraction of 0.6 and the $G_{max} = 300$ mT/m set of measurements. All gradient strengths were kept when fitting to the model. Distributions obtained when high frequencies were removed (d-e) are much narrower than when intermediate frequencies were removed (b-c).
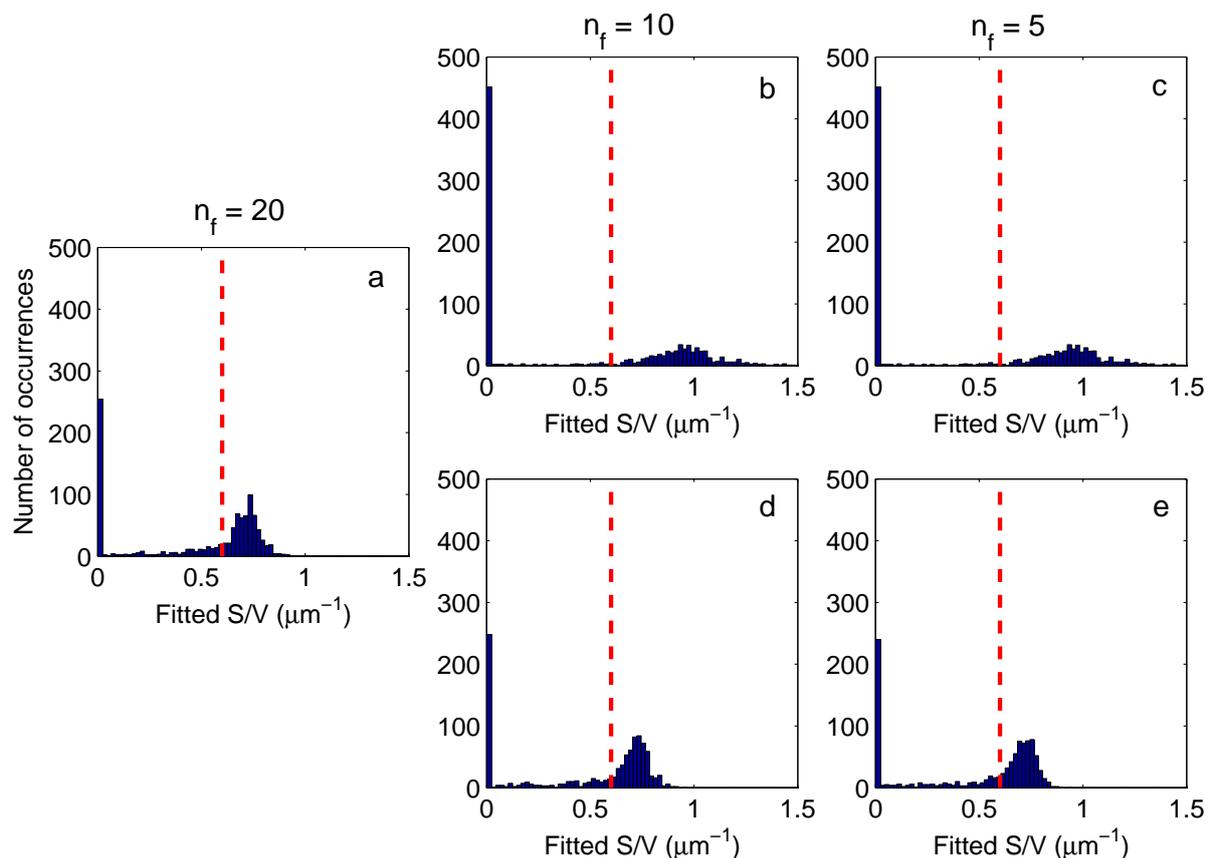
Figure 6.17: Distribution of fitted surface to volume ratios for $d = 8$ μm with (a) $n_f = 20$, (b) $n_f = 10$, (c) $n_f = 5$, (d) $n_f = 10$, (e) $n_f = 5$. In (b), the ten frequencies were 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 Hz. In (c), the five frequencies were 200, 400, 600, 800, and 1000 Hz. In (d-e), the highest frequencies were removed. In (d), the ten frequencies were 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500 Hz. In (e), the five frequencies were 50, 100, 150, 200, and 250 Hz. Simulations used a packing fraction of 0.6 and the $G_{max} = 80$ mT/m set of measurements. All gradient strengths were kept when fitting to the model. Distributions obtained when high frequencies were removed (d-e) are much narrower than when intermediate frequencies were removed (b-c).

Gradient Set #1 (G_{max} = 900 mT/m)

Figure 6.18a shows $\sigma_{S/V}$ as a function of n_f for the three subsets (using $d = 6$ μm, packing fraction = 0.6). Repeated measurements at G_{max} give the smallest $\sigma_{S/V}$ over the entire range of n_f. There is approximately a 0.03 to 0.04 μm^{-1} decrease in $\sigma_{S/V}$ when repeating the highest gradient compared to using all original measurements.

Gradient Set #2 G_{max} = 300 mT/m

Figure 6.18b shows $\sigma_{S/V}$ as a function of n_f for the three subsets (using $d = 6$ μm, packing fraction of 0.6). Repeated measurements at G_{max} give the smallest $\sigma_{S/V}$ over the entire range of n_f. There is approximately a 0.005 μm^{-1} decrease in $\sigma_{S/V}$ when repeating the highest gradient compared to using all original measurements. The amount of improvement does not depend on n_f.

Gradient Set #3 G_{max} = 80 mT/m

Figure 6.18c shows $\sigma_{S/V}$ as a function of n_f for the three subsets (using $d = 6$ μm, packing fraction of 0.6). Repeated measurements at G_{max} give the smallest $\sigma_{S/V}$ over the entire range of n_f. There is approximately a 0.06 μm^{-1} decrease in $\sigma_{S/V}$ when repeating the highest gradient compared to using all original measurements. The amount of improvement does not depend on n_f.
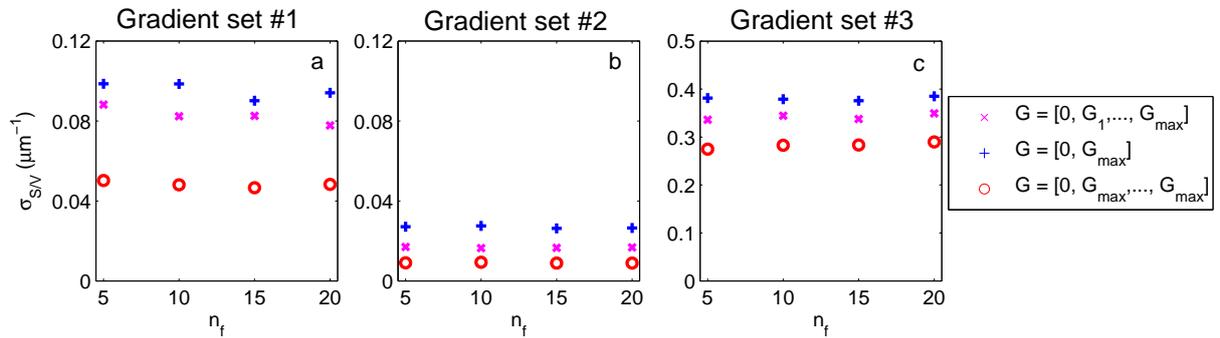
Figure 6.18: $\sigma_{S/V}$ as a function of n_f for three different gradient subsets. Subsets include those with (a) one nonzero gradient (b) all gradient strengths (c) repeated nonzero gradients at G_{max}. Three sets of measurements shown: (a) G_{max} = 900 mT/m (b) G_{max} = 300 mT/m (c) G_{max} = 80 mT/m. Data and $\sigma_{S/V}$ are from $d = 6$ μm simulations with a packing fraction of 0.6. Repeating the highest gradient strength gives the smallest $\sigma_{S/V}$. **Note**: (c) uses different y-axis limits.

6.4   Discussion

We found that the model in Eq. 6.10 gives less accurate results if diffusivities in the intra- and extra-axonal compartments are unequal compared to when diffusivities are equal. This is expected because the derivation of the *ADC* model in Eq. 6.10 assumes one diffusivity. When studying the case of different diffusivities, we used diffusivity of 1.0 μm^2/ms in the intra-axonal space and 2.5 μm^2/ms in the extra-axonal space. As the extra-axonal diffusivity approaches 1.0 μm^2/ms, we would expect the results to converge. Further study is needed to know how similar the intra- and extra-axonal diffusivities need to be for the results to adequately converge. Since diffusivities are out of our control in real tissue, this is a potential source of error we need to keep in mind. If diffusivities are different, we need to use caution when interpreting results.

In both cases, surface to volume ratios are more accurate when S/V is smaller. When S/V was too high, or the geometry too restricted, then fitted S/V underestimated the true S/V. The diameter distributions we used had higher surface to volume ratios ($> 1 \, \mu m^{-1}$), so fitted S/V underestimate the actual values. The amount of underestimation was greater in distributions having higher S/V (in other words, those having smaller cylinders). An important consequence of all this is that we might not be able to detect changes in surface to volume ratio in geometries with high S/V.

We also extracted diffusion coefficients. We found that in restricted geometries with high surface to volume ratios, fitted diffusion coefficients were underestimated. As surface to volume ratios decreased (and pore size increased), fitted diffusion coefficients rose and converged to their actual values.

We looked into how the number of frequency measurements affects the precision of estimated parameters. We compared two methods of doing this. One method involved removing high frequencies. The other method involved only keeping every second or every fourth frequency and removing the others. We found that if high frequencies are removed, results remain mostly unchanged. However, results become much worse when removing frequencies using the other method. This might be because the most important information of the diffusion spectrum $D(\omega)$ is contained at lower frequencies (< 250 Hz), at least for the situations (or axon diameters) we were interested in. The highest frequencies also have lower signal weighting

and are more affected by the addition of noise. If only smaller frequencies are used, we noticed almost no change when moving down to five frequencies ($\leqslant 250$ Hz).

Measurements with maximum gradient strengths of 300 mT/m produced more precise S/V results than measurements with maximum gradient strengths of 900 mT/m. This can be seen in both the frequency comparisons and the gradient comparisons (Figures 6.11 and 6.15). It was not until the highest gradient strength in the $G_{max} = 300$ mT/m set of measurements was reduced to 200 mT/m that very high gradient strengths ($G = 900$ mT/m) were any better.

We also considered whether it was better to repeat measurements at the high gradient strengths or to space the gradient measurements out. It was found that repeated measurements at higher gradients resulted in more precise surface to volume estimates. We saw a similar result in Chapter 5, where repeated gradient measurements led to more precise diameter estimates.

6.5 Conclusion

Here we have shown that surface to volume ratios can be extracted accurately at pre-clinical gradient strengths using OGSE. We have also shown using Monte Carlo simulations that images can be collected with fewer gradient strengths and fewer gradient frequencies without sacrificing the precision of the measurements. This could be useful in reducing imaging time so that OG techniques can be used in clinical settings. More work is needed for the method to work with clinical gradient strengths.

Chapter 7

Experimental data

So far this thesis has discussed a new method to infer micron-sized axon diameters and surface to volume ratios using temporal diffusion spectroscopy. The method was developed and optimized using Monte Carlo computer simulations. In this chapter, the method is tested on a human corpus callosum. Data were collected using 15 frequencies and six gradient amplitudes and analysis was performed leaving out images in a similar way to the Monte Carlo simulation studies in Chapters 5 and 6.

7.1 Introduction

Many OGSE experiments have used phantoms, such as beads [140, 154] and tubing [94, 97]. Experiments have also been performed with cancer cells in mice and rats [35, 36, 148, 149, 196], yeast cells [76], and rat grey matter [4, 43]. One study examined rat spinal cord with OGSE and found mean effective diameters of 1.27 to 5.54 μm [195]. A more recent study examined rat sciatic nerve and found mean effective diameters of 4.2 to 6.5 μm [80]. These values were found to be in good agreement with histology.

7.2 Ethics Statement

Pieces of tissue were collected from the autopsy specimen, under the protocol approved by the institutional health research ethics board along with the consent obtained from the family

members.

7.3 Methods

This section describes the methods used to collect data on human corpus callosum. The sample is described, and the imaging method and analysis method are described.

7.3.1 Sample

A portion of normal-appearing corpus callosum from an autopsy human brain which did not demonstrate any pathological changes was obtained from the Matsuda lab in Pathology (University of Manitoba). The sample was kept in formalin for 10-14 days before imaging. The sample was embedded in agarose within a 15 mL sample tube for imaging.

7.3.2 MRI

Images were acquired with a 2.5 cm diameter RF bird cage coil (Bruker Biospin), using a 7 T Bruker Avance III NMR system (Paravision 5.0), with a BGA6 gradient insert (maximum gradient strength of 1.01 T/m).

Two 20 ms sine gradient pulses were used, separated by 24.52 ms. There were 15 frequencies ranging from 50 to 750 Hz in steps of 50 Hz. Six gradient strengths were used for each gradient pulse, for a total of 90 signals. For $f_m > 50$ Hz, the gradient strengths were 0, 44, 61, 76, 88, 99% of maximum. For $f_m = 50$ Hz, G = 0, 22, 32, 39, 45, 50%. A 1 mm thick slice taken to be approximately perpendicular to the direction of the axons within the corpus callosum was taken. The field of view was 2.56 cm^2. A 128×128 matrix was used for 200 μm in-plane resolution. The following imaging parameters were used: number of averages (*NA*) = 4, $T_R = 1250$ ms, $T_E = 90$ ms. A total of 14 sets of images were collected. The total imaging time was 112 hours (4.67 days). The SNR in the corpus callosum in each image was 32 without diffusion-weighting. The average SNR when using the highest gradient strength was 26.

7.3.3 Histology

After imaging, the region just above the MRI slice was analyzed with electron microscopy (EM). Ten EM slices were collected. Axon diameters were measured using ImageJ software (National Institutes of Health, Bethesda, Maryland). To measure axon diameter, lines were drawn across the smallest diameter of all cells identified as axons. Each slice contained 500 to 700 identifiable axons. Axon diameters ranged from 0.14 μm to 6 μm. The mean axon diameter over all slices was 0.8±0.4 μm. Using Eq. 5.1, the average *AxD* was found to be 1.4±0.2 μm. The fitted *AxD* from Eq. 4.9 should correlate with this quantity and not the mean diameter [7, 195].

7.3.4 Analysis

Images were registered to the first scan's $b_0 = 0$ image [174]. The mean $\pm$ standard deviation of the signal in the ROIs, shown in Figure 7.1, was calculated using a custom written MATLAB script. The signal was assumed to be described by a two compartment model given in Eq. 4.9. The extra-axonal component E_h was described by Eq. 4.10. The restricted component E_r was given by Eq. 4.13, with $\beta(2\tau)$ given by Eq. 4.15. This time, we use the restricted signal for a sine gradient. There are four model parameters (intra-axonal diffusion coefficient D_{in}, *AxD*, packing fraction f_{axon}, hindered diffusion coefficient D_h), but we are mainly interested in *AxD*.

Signals were fitted to the two compartment model using least squares minimization to extract *AxD*. We used three different methods of removing data. Higher OGSE frequencies were removed and the remaining data was refitted to the model to see how fitted parameters changed. Model fitting was repeated using all possible combinations of the gradient strengths (30 combinations). Finally, we removed repeated measurements. There were initially 1260 images (15 OGSE frequencies $\times$ 6 gradient strengths $\times$ 14 sets of images).

7.4 Results

Figure 7.2 shows the variation in fitted *AxD* as a function of number of frequencies for each corpus callosum ROI. In Figure 7.2a, which has the largest *AxD* (2.35 ± 0.01 μm with all
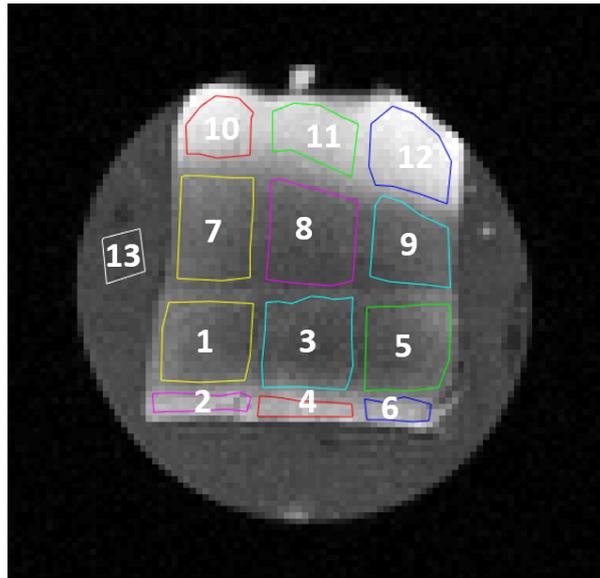
Figure 7.1: Image of sample showing the 13 regions of interest. Analysis ROIs were created in the corpus callosum (ROI #1, 3, 5, 7-9), ependymal layer (ROI #2, 4, 6), cortex (ROI #10-12), and in the agarose (ROI #13). We expect to find axon diameters in the corpus callosum, and cell diameters in the ependymal layer given that they can be modeled as elongated parallel cells.

measurements), *AxD* are within 1% of each other if there are at least 7 frequencies. The highest and lowest *AxD* occur when using 2 or 3 frequencies. The smallest fitted *AxD* is 1.9 ± 0.1 μm (3 frequencies). The highest fitted *AxD* is 2.6 ± 0.2 μm (2 frequencies). Compared to *AxD* with 15 frequencies, these are respective differences of 17% and 3%. Error decreases when more frequencies are used. In Figures 7.2b-f, *AxD* become smaller when the number of frequencies is above 5. There is some variability in *AxD* for small numbers (< 6) of frequencies. Confidence intervals also tend to become much wider when there are fewer frequencies.

Figure 7.3 shows variation in fitted *AxD* when using only two gradient strengths (ROI #1). With the exception of the first gradient, fitted *AxD* are within 5% of each other. The error also increases when smaller gradients are used, with fitted *AxD* values ranging from 2.40 ± 0.08 μm with the highest gradient strength to 4 ± 6 μm with the smallest gradient strength.

Figure 7.4 shows fitted *AxD* for the best combinations with each number of gradients (the best combinations for ROI #1 are shown in Table 7.1). The fitted values in Figures 7.4a, 7.4b, and 7.4c are within a few percent of each other, while error when using just two gradients is about 2 times larger than when using all six gradients. The fitted values in Figures 7.4d, 7.4e, and 7.4f are within a few percent ($< 5\%$) of each other, provided there are at least 3 gradient
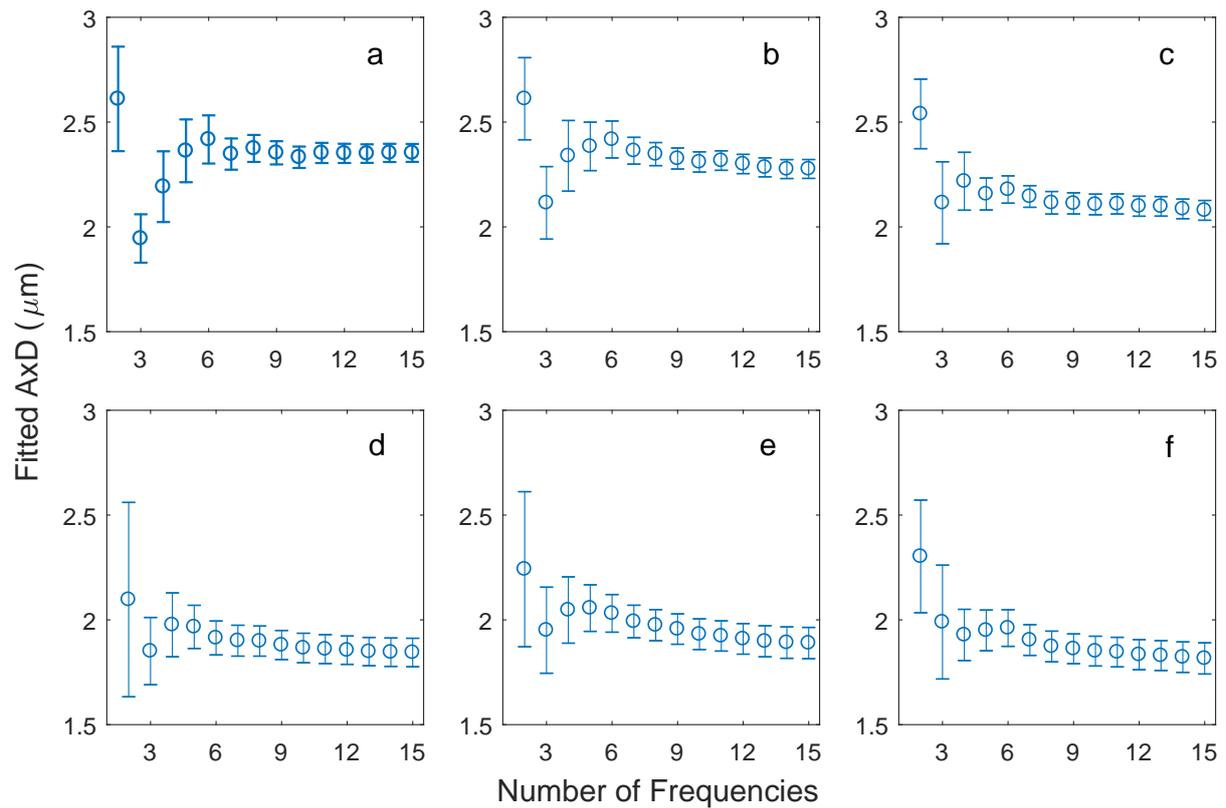
Figure 7.2: Fitted *AxD* ($\pm$95% confidence bounds) as a function of number of frequencies for each corpus callosum ROI. There is some variability in *AxD* for small numbers (< 6) of frequencies. In (b-f), *AxD* appear to decrease in as more frequencies are included. (a) ROI #1 (b) ROI #3 (c) ROI #5 (d) ROI #7 (e) ROI #8 (f) ROI #9.
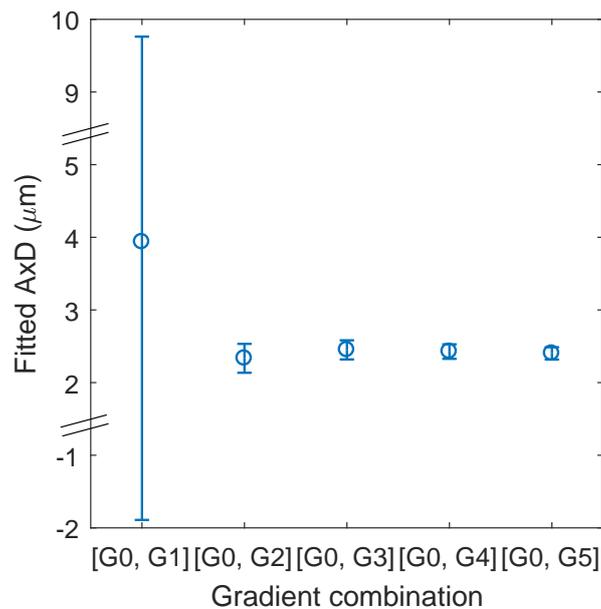
Figure 7.3: Fitted *AxD* ($\pm$95% confidence bounds) using two gradient strengths, $G_0 = 0$ and another chosen from between G_1 through G_5. The error in *AxD* decreases as the nonzero gradient strength increases. Data shown here are from ROI #1 in the corpus callosum.

209

strengths.

Number of gradient strengths	Optimal choice of gradients
2	$G_0 = 0, G_5$
3	$G_0 = 0, G_2, G_5$
4	$G_0 = 0, G_2, G_4, G_5$
5	$G_0 = 0, G_1, G_2, G_4, G_5$
6	$G_0 = 0, G_1, G_2, G_3, G_4, G_5$

Table 7.1: The gradient combinations for 2, 3, 4, 5, and 6 gradient strengths that produced the smallest error after fitting (ROI #1). Corresponding *AxD* shown in Figure 7.4a. Note that all combinations include the highest gradient strength (G_5).

Figure 7.5 shows fitted *AxD* as a function of number of repeated measurements for each corpus callosum ROI. Error decreases as more repeated measurements are included. There appears to be less variability in Figure 7.5a-c than Figure 7.5d-f. Error also tends to be higher in Figure 7.5d-f and increases faster with fewer repeats. In Figures 7.5a, 7.5b, and 7.5c, *AxD* are within 5% percent as long as the number of repeats is greater than 2 (though the *AxD* in Figure 7.5a differ by less than 4% even with one repeat).

7.5 Discussion

We separately reduced the number of frequencies and number of gradients. For example, in ROI #1, reducing the number of measurements such that the imaging time is shortened by a factor of 1.5 increases error by a factor of 1.2 (by reducing the number of frequencies), while changing *AxD* by 1%. Reducing the number of measurements such that the imaging time is shortened by a factor of 3 increased error by a factor of 2 (by reducing the number of gradients), while changing *AxD* by 3%. Reducing the number of measurements such that the imaging time is shortened by a factor of 14 increased error by a factor of 3 (by reducing the number of repeated measurements), while changing *AxD* by 3%. However, the trade-off seems to depend on the ROI (or fitted *AxD*). For example, in ROIs with a smaller fitted *AxD* ($\#7 - 9$), using just two gradients changes *AxD* by 10%, but using three gradients only changes *AxD* by 1%. In ROI #9 (fitted $AxD \approx 1.8$ μm), we could have reduced number of measurements by a factor of two (using fewer repeats) and kept *AxD* within 2%.
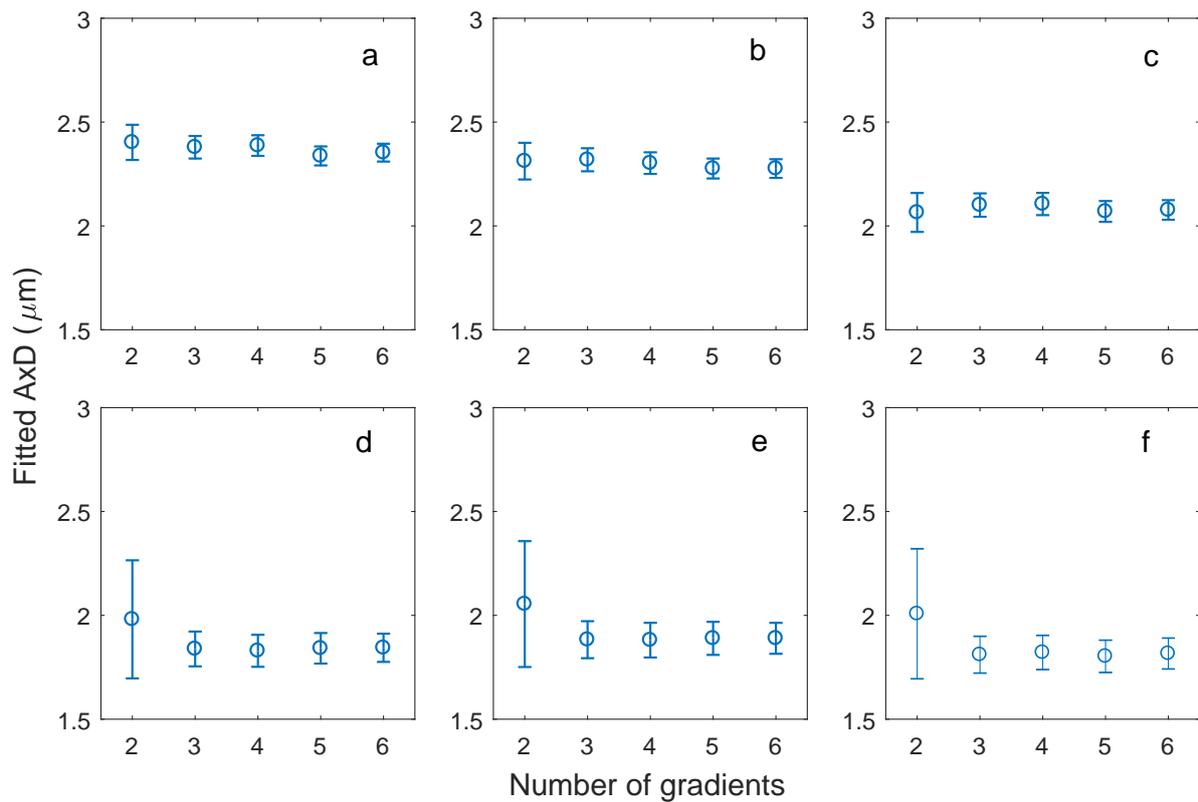
Figure 7.4: Best fitted *AxD* (±95% confidence bounds) (as determined from smallest fitted error) when using 2, 3, 4, 5, or 6 gradient strengths for data fitting. The gradient subsets are shown in Table 7.1. If 3 or more gradients are used, *AxD* values are all consistent with each other. (a) ROI #1 (b) ROI #3 (c) ROI #5 (d) ROI #7 (e) ROI #8 (f) ROI #9.
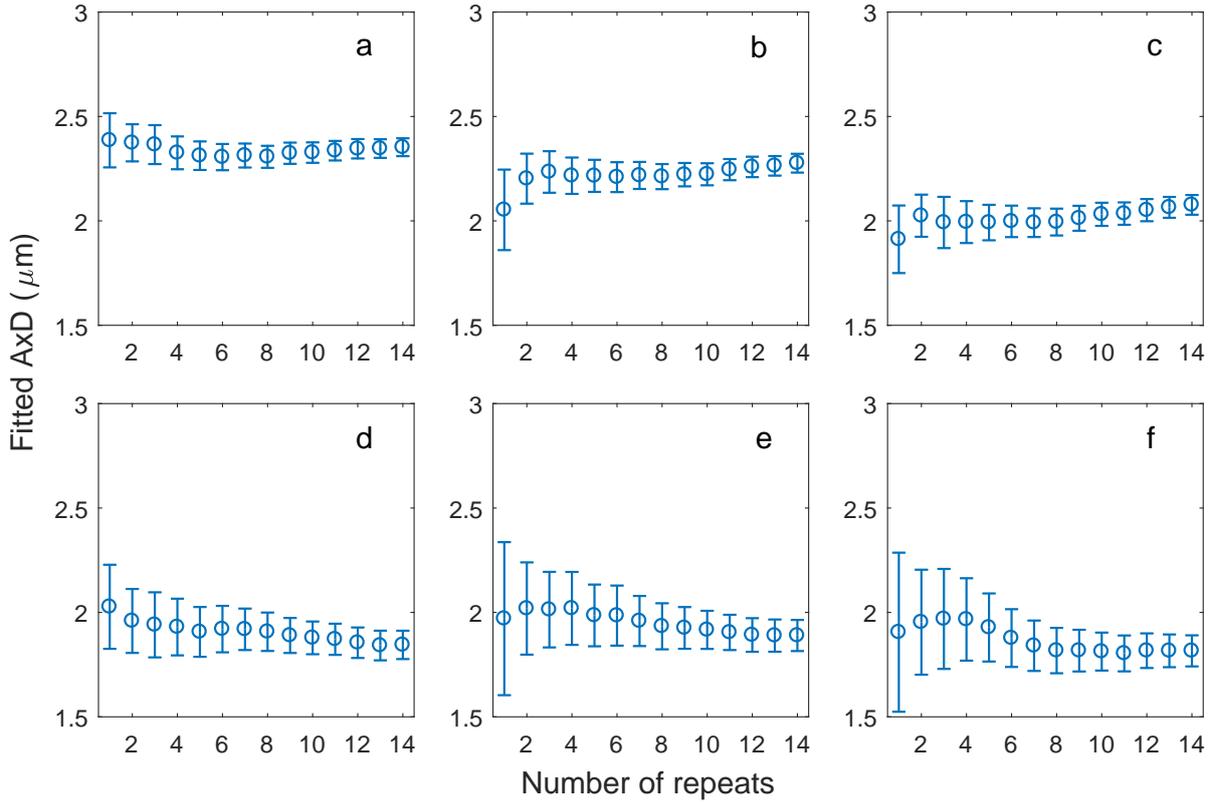
Figure 7.5: Fitted *AxD* (±95% confidence bounds) as a function of number of repeats for all corpus callosum ROIs. (a) ROI #1 (b) ROI #3 (c) ROI #5 (d) ROI #7 (e) ROI #8 (f) ROI #9.

The largest fitted *AxD* from Figure 7.2 using all 1260 measurements was 2.35 ± 0.04 μm. This fitted *AxD* is almost 1 μm higher than the *AxD* measured from histology (Section 7.3.3). Fitted *AxD* from the other ROIs with all measurements were smaller, but still higher than the histological *AxD*. The smallest fitted *AxD* was 1.81 ± 0.07 μm, which is only around 0.4 μm higher than the histological *AxD*. The average fitted *AxD* over the corpus callosum ROIs was 2.0 ± 0.2 μm.

In Chapter 5, we found that simulations with 5 to 10 frequencies gave similar results as simulations with 20 frequencies. Here we found that there was more variability in fitted *AxD* below 7 frequencies. However, in most ROIs, fitted *AxD* continue to decrease beyond 15 frequencies. Simulations predicted that error should increase if smaller gradients were used. In Figure 7.3, we see that error in *AxD* does increase as smaller gradient strengths are used. Simulations also predicted that fitted *AxD* errors from two gradients should be similar, but slightly larger, compared to multiple gradients, especially if higher gradient strengths were used. For ROIs with fitted $AxD > 2$ μm, fitted *AxD* using two gradients were similar to those using six

gradients. For ROIs with fitted $AxD < 2$ μm, fitted AxD using 3 gradients were similar to those using six gradients.

The maximum SNR in each set of images was around 30. We tried to compensate by collecting many sets ($\times14$) of images and including all of them in the analysis. This means that the first data point (number of repeats = 1) in each subplot in Figure 7.5 uses one set of $SNR \approx 30$ images. Fitted AxD from ROIs #8 and #9 appear to be noticeably worse in terms of uncertainty with just one set of images included. Even when including multiple sets of images, these ROIs tend to be worse in terms of uncertainty and variability. These ROIs also have the smallest fitted AxD even when including all sets of images. We know from Chapter 4 that lower SNR affects smaller diameters more than larger diameters and that it forces axon diameters to be overestimated. If there happened to be heterogeneity in axon diameter over the slice, and these ROIs actually contained smaller axons than the others, then this might be something we would expect. The smallest AxD of the simulated axon distributions in Chapter 5 was 2.1 μm. The mean fitted AxD was 2.31 μm. With a more realistic clinical SNR of 20, the mean fitted AxD was found to be 2.50 μm. Therefore, we might expect fitted AxDs to be around 0.2 to 0.4 μm higher than the true values when $AxD \approx 2$ μm. The difference between true and fitted AxD for the simulated distribution with $AxD = 2.9$ μm was only 0.1 μm. If the true AxD is smaller than 2 μm, then there might be a larger difference. If the difference is large enough (if the difference was ≈ 1 μm), then 1.4 μm might be indistinguishable from 2 μm.

This study presents the first step toward reducing imaging time toward feasible axon diameter measurements *in vivo*. Higher frequencies may be needed for sensitivities toward smaller axons. Figure 7.2 shows that AxD might continue to decrease past 15 frequencies for most ROIs. At $f_m = 750$ Hz, the diffusion signal only decays to around 97% its original value. If higher frequencies are used, it might also be necessary to use higher gradient strengths. Permeability might need to be included in the model. One study [99] found that the influence of permeability was negligible in OGSE measurements if cells were larger than 5 μm in diameter. The cells here are much smaller than 5 μm, so permeability might play a role.

7.6 Conclusion

Here we imaged a portion of human corpus callosum and fit the data to a two compartment model. We found that fitted *AxD* overestimated the histological *AxD* by 0.4 to 0.9 μm. We also suggested some possibilities for reducing the number of images needed for accurate results.

Chapter 8

Conclusion

We used Monte Carlo simulation data to test a few different microstructural models. We successfully found models that, when combined with the OGSE sequences, can be used to measure axon diameters and surface to volume ratios. We studied the effects of using different numbers of gradient frequencies for measurements and in general found that 5 – 10 frequencies could be used to give accurate results. We also found that collecting one image without added diffusion weighting and multiple images with the same added diffusion weighting provided the best results. A more detailed summary is presented in this chapter.

8.1 AxCaliber Model

The first microstructural model studied here was a modified AxCaliber model used to find axon diameter distributions (Chapter 4). The model was modified for use with cosine OGSE sequences. We started with intra-axonal simulations ($d = 1$ μm to 8 μm). This verified that the intra-axonal OGSE model was accurate and could extract axon diameters. These were followed by simulations with identical square packed axons. This situation corresponds to a two compartment model. Extra-axonal diffusion was assumed to be hindered (*ADC* constant with OGSE frequency). Because of the frequency dependence of the *ADC* in the extra-axonal region, not all diameters were accurately estimated. For realistic frequencies of 50 – 1000 Hz, diameters above 4 or 5 μm were underestimated. Finally, we ran simulations that included a distribution of diameters and fitted the data to the AxCaliber model. Gaussian and gamma

215

distributions, having diameters in the range 0.25 μm to 8 μm, were used. After data fitting, we were able to recover the shape of the diameter distributions. In addition, the frequency independent extra-axonal *ADC* model was replaced with a model where *ADC* linearly varied with frequency. For diameters smaller than 4 or 5 μm, the original extra-axonal model gave similar results as the more complicated model. Above 4 and 5 μm, there was little improvement when using the new extra-axonal model. Noise (SNR = 50, 100, 200) was added to simulation data to see how parameter estimates were affected. As expected, higher SNR led to more accurate and precise parameters. We also found that lower SNR affected smaller diameters (e.g. $d = 1$ μm) more than larger diameters. This limits the ability to infer smaller diameters.

8.2 Two Compartment Model

The second model was a simplified two compartment model (Chapter 5). The main objective was to see how the number of measurements (e.g. number of OGSE frequencies) affected parameter estimates (such as axon diameter). We were especially interested in parameter precision. Noise was added to all simulation data before fitting. We performed intra-axonal simulations, simulations of square packed axons, and simulations with axon diameter distributions. For intra-axonal simulations, axon diameters ranged from 1 μm to 10 μm. For identical square packed axons, diameters ranged from 1 μm to 5 μm. For axon diameter distributions, diameters ranged from approximately 0.25 μm to 8 μm. With an axon diameter distribution present, the two compartment model gives a quantity *AxD*, similar to a mean diameter. Simulations were performed with maximum gradient strengths of 80, 300, and 900 mT/m.

Precision of parameters was higher with larger axon diameter. Accuracy and precision of estimated parameters were worse when using more complicated models (e.g. identical square packed axons and diameter distributions). Measurements with higher gradient strengths gave more accurate and precise results. With the smallest gradient strengths (80 mT/m), there was an upward bias in estimated axon diameter. Intra-axonal diffusivity was poorly estimated at the smallest diameters ($d = 1$ μm) and with smaller gradient strengths. There was little change in fitted diameter down to OGSE frequencies of 250 Hz. This represents four times fewer

measurements compared to the original number. There was also some decrease in parameter precision when using two gradient strengths as compared to four or more. We also saw that repeating measurements at a single gradient strength was better than spreading out the measurements.

8.3 Surface to Volume Model

The third model was a short time *ADC* model used to find surface to volume ratios. Here there were two objectives. The first was to assess the accuracy of the model for different axon diameters. The second was to see how the number of measurements (e.g. number of OGSE frequencies or number of gradients) affected parameter estimates. We performed intra-axonal simulations, simulations of square packed axons, and simulations with axon diameter distributions. Axon diameters ranged from 1 to 10 μm. Initially, simulations used different diffusivities in the intra- and extra-axonal compartments. Later simulations used equal diffusivities. In intra-axonal simulations, it was found that diffusion coefficients were underestimated at smaller diameters ($d < 5$ μm), while surface to volume ratios were overestimated. Results began to converge at lower surface to volume ratios (corresponding to larger axon diameters). The same was true for simulations of square packed axons and simulations with axon diameter distributions. At low surface to volume ratios, extracted surface to volume ratios and extracted diffusion coefficients were more accurate ($S/V < 1$ μm^{-1}). At high surface to volume ratios, diffusion coefficients and surface to volume ratios were underestimated. With different diffusivities, fitted parameters started to converge at axon diameters of $d = 10$ μm. With equal diffusivities, fitted parameters converged to the actual values at around 5 μm. Since axons in the diameter distributions were relatively small, S/V was always underestimated by some amount, though less so with distributions with larger diameters.

Simulations used maximum gradient strengths of 80, 300, and 900 mT/m. Fitted diffusion coefficients were more precise with larger diameters. Measurements with higher maximum gradients tended to give more precise estimates. Maximum gradient strengths of 300 mT/m gave more precise results than measurements at 900 mT/m. There was little change in fitted

S/V down to OGSE frequencies of 250 Hz, which is equivalent to using four times fewer measurements than the original number. However, this benefit only occurred when the highest frequencies were removed. Removal of higher gradient strengths from the data led to a decrease in precision. There was also some decrease in parameter precision when using two gradient strengths as compared to four or more. We also saw that repeating measurements at a single gradient strength was better than spreading out the measurements.

8.4   Extensions to Previous Work

One previous study had shown that just two OG frequencies were sufficient to estimate inner diameters of capillary tubing [97]. In a capillary tube experiment, the signal comes from a single compartment. This thesis used simulations that included both intra-axonal and extra-axonal compartments. In addition to using strong pre-clinical gradient strengths (≈ 900 mT/m), we used simulations with maximum gradient strengths of 300 mT/m and 80 mT/m. In Chapter 5, we found that estimated axon diameters were worse with a two compartment model, but 5 to 10 OG frequencies were sufficient to estimate axon diameters. We used the same methodology on *ex vivo* diffusion-weighted images of a human corpus callosum and found that 7 or more frequencies were needed for consistent axon diameter estimates. The same methodology was used to show that 2 or 3 gradient strengths were sufficient.

8.5   Recommendations for Experiments

Based on the results of the simulations and analysis of the brain data, I recommend the following parameters be used for data collection for samples with expected diameters of 2 to 4 μm.

If an *AxD* is desired, then the two compartment model in Chapter 5 should be used. Five frequencies, equally spaced between 50 and 250 Hz, should be used. If diameters are closer to 2 μm, then 10 frequencies should be used ($50 - 500$ Hz). Gradient strengths should be as high as possible. Two gradient strengths should be sufficient ($G = \{0, G_n\}$). But if there is time, more measurements should be taken at the higher gradient strength ($G = G_n$).

If a diameter distribution is desired, the AxCaliber method using OGSE should be used, as

in Chapter 4. We did not look at optimizing the choice of frequencies or gradients that could be used to acquire fewer images in a shorter amount of time. But based on the results in Chapter 5, ten frequencies ($\leqslant 500$ Hz) are probably sufficient.

If the sample has a relatively uniform diffusivity and the expected surface to volume ratio is less than 2 μm^{-1}, then the short-time model in Chapter 6 can be used to calculate S/V. Based on the simulations, these conditions require the sample to have larger diameter axons. Five frequencies, equally spaced between 50 and 250 Hz, should be used. Two gradient strengths should be sufficient ($G = \{0, G_n\}$). If there is time, more measurements should be taken at the higher gradient strength ($G = G_n$).

Future students should be aware of slew rate considerations when adapting the method for human use. Slew rate is restricted to less than 200 $\frac{\text{T}}{\text{m·s}}$ in humans [171]. The gradients used in this thesis do not meet slew rate limitations.

For the brain experiment presented in Chapter 7, 1260 images were collected over a period of 112 hours from 6 gradient strengths and 15 frequencies repeated 14 times. Through the analysis presented in section 7.4 I found that the number of frequencies could safely be reduced to around 7 or 8, the number of gradient strengths to 2 or 3, and the number of repetitions to 7. This would require $1260 \times (3/6) \times (8/15) \times (7/14) = 168$ images to be collected over a period of 15 hours. If spatial variance in the axon diameter was not required, the SNR of the image could be increased by decreasing the resolution. For instance, increasing the voxel size by a factor of 2 in each of the in plane directions increases the voxel size by a factor of 4 while decreasing the data collection time by a factor of 2 (2 times fewer number of k-space lines). We can further reduce the imaging time by a factor of 2 by reducing the number of repetitions to keep the SNR the same as the images collected in chapter 7. This will result in an image collection time of 3.75 hours.This makes live mouse imaging possible. Further means to reduce imaging time are needed to make the method suitable for research on humans and even further reductions are needed to make this into a clinically relevant method.

8.6 Future work

There are many ways we could move forward.

The first possibility is to explore the effects of using more complex geometries in the simulations. This could mean including axon dispersion, or non-parallel axons [46, 158, 199]. It could also mean including permeable membranes. Permeability has been studied in the case of pulsed gradients [39, 57], but it has recently been shown that permeable boundaries may have a negligible effect on *ADC*s at timescales accessible to OGSE [99]. Simulation data could then be used in the models described here to see how much extracted parameters, like axon diameter, are affected. In the short time S/V model, we saw that different intra- and extra-axonal diffusivities ($D_{in} = 1.0 \ \mu\mathrm{m}^2/\mathrm{ms}$, $D_{ex} = 2.5 \ \mu\mathrm{m}^2/\mathrm{ms}$) led to inaccurate S/V estimates for $d < 10 \ \mu\mathrm{m}$. Simulations with equal diffusivities ($D_{in} = D_{ex} = 1.0 \ \mu\mathrm{m}^2/\mathrm{ms}$) gave relatively accurate S/V estimates for $d > 5 \ \mu\mathrm{m}$. Further simulations could be used to slowly increase the difference in diffusivity to see any changes.

A second possibility is to develop more complicated microstructure models. For example, axon dispersion models developed for PGSE could be adapted for OGSE [199]. Models that incorporate crossing fibres, like CHARMED [12], are another possibility. We could also include axons with unknown orientations. More complicated modeling may require more complex acquisition procedures, such as measurements with different gradient directions. Increasing model complexity introduces more parameters. We may need to reduce the number of free parameters, either by keeping some fixed or combining some of them together. New acquisition schemes would also increase imaging times. We might have to reduce the number of measurements again.

Another possibility is to use more sophisticated experimental design procedures and analysis. This might involve introducing an ActiveAx-style experimental optimization [6]. It could also involve incorporating prior information before data fitting [34].

The use of oscillating gradient temporal diffusion spectroscopy is gaining more attention. A recent article in Nature [130] used elliptically polarized oscillating gradient spin echo in order to separate tissue compartment signals in disordered systems in a monkey brain. Future work could combine this method with the method developed in this thesis for a better inference of

tissue properties.

References

[1] F. Aboitiz, A. B. Scheibel, R. S. Fisher, and E. Zaidel. Fiber composition of the human corpus callosum. *Brain research*, 598(1):143–153, 1992.

[2] A. Abragam. *The principles of nuclear magnetism*. Number 32. Oxford university press, 1961.

[3] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1964.

[4] M. Aggarwal, M. V. Jones, P. A. Calabresi, S. Mori, and J. Zhang. Probing mouse brain microstructure using oscillating gradient diffusion mri. *Magnetic resonance in medicine*, 67(1):98–109, 2012.

[5] D. Alexander, G. Barker, and S. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.

[6] D. C. Alexander. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine*, 60(2):439–448, 2008.

[7] D. C. Alexander, P. L. Hubbard, M. G. Hall, E. A. Moore, M. Ptito, G. J. Parker, and T. B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *Neuroimage*, 52(4):1374–1389, 2010.

[8] A. Anderson and Z. Ding. Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging. In *Book of abstracts: Tenth Annual Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, CA: ISMRM*, volume 10, page 440, 2002.

[9] R. A. Armstrong. Alzheimer's disease and the eye. *Journal of Optometry*, 2(3):103–111, 2009.

[10] I. Åslund and D. Topgaard. Determination of the self-diffusion coefficient of intracellular water using PGSE NMR with variable gradient pulse length. *Journal of Magnetic Resonance*, 201(2):250–254, 2009.

[11] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6):1347–1354, 2008.

[12] Y. Assaf, R. Z. Freidlin, G. K. Rohde, and P. J. Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5):965–978, 2004.

[13] Y. Assaf, A. Mayk, and Y. Cohen. Displacement imaging of spinal cord using q-space diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 44(5):713–722, 2000.

[14] A. Baert, R. Günther, and G. von Schulthess. *Interventional Magnetic Resonance Imaging*. Springer Science & Business Media, 2012.

[15] B. Balinov, B. Jonsson, P. Linse, and O. Soderman. The NMR self-diffusion method applied to restricted diffusion. simulation of echo attenuation from molecules in spheres and between planes. *Journal of Magnetic Resonance, Series A*, 104(1):17–25, 1993.

[16] D. Barazany, P. J. Basser, and Y. Assaf. *In vivo* measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain*, 132(5):1210–1220, 2009.

[17] P. J. Basser, J. Mattiello, and D. LeBihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B*, 103(3):247–254, 1994.

[18] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.

[19] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. *in vivo* fiber tractography using dt-MRI data. *Magnetic Resonance in Medicine*, 44(4):625–632, 2000.

[20] M. F. Bear, B. W. Connors, and M. A. Paradiso. *Neuroscience*, volume 2. Lippincott Williams & Wilkins, 2007.

[21] C. Beaulieu. The basis of anisotropic water diffusion in the nervous system–a technical review. *NMR in Biomedicine*, 15(7-8):435–455, 2002.

[22] T. E. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5):1077–1088, 2003.

[23] M. A. Bernstein, K. F. King, and X. J. Zhou. *Handbook of MRI pulse sequences*. Elsevier, 2004.

[24] S. Bisdas, D. Bohning, N. Bešenski, J. Nicholas, and Z. Rumboldt. Reproducibility, interrater agreement, and age-related changes of fractional anisotropy measures at 3t in healthy subjects: effect of the applied b-value. *American Journal of Neuroradiology*, 29(6):1128–1133, 2008.

[25] F. Bloch. Nuclear induction. *Physical review*, 70(7-8):460, 1946.

[26] N. Bloembergen, E. M. Purcell, and R. V. Pound. Relaxation effects in nuclear magnetic resonance absorption. *Physical review*, 73(7):679, 1948.

[27] A. Brander, A. Kataja, A. Saastamoinen, P. Ryymin, H. Huhtala, J. Öhman, S. Soimakallio, and P. Dastidar. Diffusion tensor imaging of the brain in a healthy adult population: Normative values and measurement reproducibility at 3 t and 1.5 t. *Acta Radiologica*, 51(7):800–807, 2010.

[28] M. A. Brown and R. C. Semelka. MR imaging abbreviations, definitions, and descriptions: a review. *Radiology*, 213(3):647–662, 1999.

[29] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan. *Magnetic resonance imaging: physical principles and sequence design.* John Wiley & Sons, 2014.

[30] G. Bydder and I. Young. MR imaging: clinical use of the inversion recovery sequence. *Journal of computer assisted tomography*, 9(4):659–675, 1985.

[31] P. T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy.* Oxford University Press on Demand, 1993.

[32] P. T. Callaghan. Pulsed-gradient spin-echo NMR for planar, cylindrical, and spherical pores under conditions of wall relaxation. *Journal of Magnetic Resonance, Series A*, 113(1):53–59, 1995.

[33] P. T. Callaghan. A simple matrix formalism for spin echo analysis of restricted diffusion under generalized gradient waveforms. *Journal of Magnetic Resonance*, 129(1):74–84, 1997.

[34] J. D. Clayden, Z. Nagy, N. Weiskopf, D. C. Alexander, and C. A. Clark. Microstructural parameter estimation *in vivo* using diffusion MRI and structured prior information. *Magnetic Resonance in Medicine*, 75(4):1787–1796, 2016.

[35] D. C. Colvin, M. E. Loveless, M. D. Does, Z. Yue, T. E. Yankeelov, and J. C. Gore. Earlier detection of tumor treatment response using magnetic resonance diffusion imaging with oscillating gradients. *Magnetic Resonance Imaging*, 29(3):315–323, 2011.

[36] D. C. Colvin, T. E. Yankeelov, M. D. Does, Z. Yue, C. Quarles, and J. C. Gore. New insights into tumor microstructure using temporal diffusion spectroscopy. *Cancer research*, 68(14):5941–5947, 2008.

[37] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences*, 96(18):10422–10427, 1999.

[38] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.

[39] E. Davoodi-Bojd, M. Chopp, H. Soltanian-Zadeh, S. Wang, G. Ding, and Q. Jiang. An analytical model for estimating water exchange rate in white matter using diffusion MRI. *PLOS ONE*, 9(5):1–15, 05 2014.

[40] T. M. de Swiet and P. N. Sen. Decay of nuclear magnetization by bounded diffusion in a constant field gradient. *The Journal of chemical physics*, 100(8):5597–5604, 1994.

REFERENCES

[41] J. D'Errico. FresnelS and FresnelC. URL https://www.mathworks.com/matlabcentral/fileexchange/28765-fresnels-and-fresnelc, 2012. The Math-Works, Natick, MA, USA.

[42] M. Descoteaux, R. Deriche, D. Le Bihan, J.-F. Mangin, and C. Poupon. Multiple q-shell diffusion propagator imaging. *Medical image analysis*, 15(4):603–621, 2011.

[43] M. D. Does, E. C. Parsons, and J. C. Gore. Oscillating gradient measurements of water diffusion in normal and globally ischemic rat brain. *Magnetic Resonance in Medicine*, 49(2):206–215, 2003.

[44] I. Drobnjak and D. C. Alexander. Optimising time-varying gradient orientation for microstructure sensitivity in diffusion-weighted MR. *Journal of Magnetic Resonance*, 212(2):344–354, 2011.

[45] I. Drobnjak, B. Siow, and D. C. Alexander. Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. *Journal of Magnetic Resonance*, 206(1):41–51, 2010.

[46] I. Drobnjak, H. Zhang, A. Ianuş, E. Kaden, and D. C. Alexander. PGSE, OGSE, and sensitivity to axon diameter in diffusion MRI: insight from a simulation study. *Magnetic Resonance in Medicine*, 75(2):688–700, 2016.

[47] W. L. Dunn and J. K. Shultis. *Exploring monte carlo methods*. Elsevier, 2011.

[48] T. B. Dyrby, M. G. Hall, M. Ptito, D. Alexander, et al. Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic Resonance in Medicine*, 70(3):711–721, 2013.

[49] W. A. Edelstein, J. M. Hutchison, G. Johnson, and T. Redpath. Spin warp NMR imaging and applications to human whole-body imaging. *Physics in medicine and biology*, 25(4):751, 1980.

[50] A. Einstein. *Investigations on the Theory of the Brownian Movement*. Courier Corporation, 1956.

[51] M. Eis and M. Hoehnberlage. Correction of gradient crosstalk and optimization of measurement parameters in diffusion MR imaging. *Journal of Magnetic Resonance, Series B*, 107(3):222–234, 1995.

[52] E. Fieremans, Y. De Deene, S. Delputte, M. S. Özdemir, Y. D'Asseler, J. Vlassenbroeck, K. Deblaere, E. Achten, and I. Lemahieu. Simulation and experimental verification of the diffusion in an anisotropic fiber phantom. *Journal of Magnetic Resonance*, 190(2):189–199, 2008.

[53] J. Frahm, K. Merboldt, W. Hänicke, and A. Haase. Stimulated echo imaging. *Journal of Magnetic Resonance (1969)*, 64(1):81–93, 1985.

[54] R. Freidlin, J. Kakareka, T. Pohida, M. Komlosh, and P. Basser. A spin echo sequence with a single-sided bipolar diffusion gradient pulse to obtain snapshot diffusion weighted images in moving media. *Journal of Magnetic Resonance*, 221:24–31, 2012.

[55] A. F. Frøhlich, S. N. Jespersen, L. Østergaard, and V. G. Kiselev. The effect of impermeable boundaries of arbitrary geometry on the apparent diffusion coefficient. *Journal of Magnetic resonance*, 194(1):128–135, 2008.

[56] J. C. Gore, J. Xu, D. C. Colvin, T. E. Yankeelov, E. C. Parsons, and M. D. Does. Characterization of tissue structure at varying length scales using temporal diffusion spectroscopy. *NMR in Biomedicine*, 23(7):745–756, 2010.

[57] D. S. Grebenkov, D. Van Nguyen, and J.-R. Li. Exploring diffusion across permeable barriers at high gradients. I. narrow pulse approximation. *Journal of Magnetic Resonance*, 248:153–163, 2014.

[58] B. Gross and R. Kosfeld. Anwendung der spin-echo-methode der messung der selbstdiffusion. *Messtechnik*, 77:171–177, 1969.

[59] E. Hahn. Nuclear induction due to free larmor precession. *Physical Review*, 77(2):297, 1950.

[60] E. L. Hahn. Spin echoes. *Physical review*, 80(4):580, 1950.

[61] M. G. Hall and D. C. Alexander. Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI. *IEEE transactions on medical imaging*, 28(9):1354–1364, 2009.

[62] L. Hebel and C. P. Slichter. Nuclear spin relaxation in normal and superconducting aluminum. *Physical Review*, 113(6):1504, 1959.

[63] S. Herrera, M. Mercredi, T. Vincent, B. R, and M. Martin. Using oscillating gradient spin-echo sequences to infer micron-sized bead and pore radii. In *Proceedings of the 23rd Annual meeting of the ISMRM*, page 3027, Toronto, Canada, 2015.

[64] S. Herrera, T. Vincent, M. Mercredi, T. Vincent, B. R, C. Bidinosti, and M. Martin. Inferring axon diameter sizes using Monte Carlo simulations and oscillating gradient spin echo sequences. In *Proceedings of the 22rd Annual meeting of the ISMRM*, page 4508, Milan, Italy, 2014.

[65] R. K. Hobbie and B. J. Roth. *Intermediate physics for medicine and biology*. Springer Science & Business Media, 2007.

[66] A. Horowitz, D. Barazany, I. Tavor, M. Bernstein, G. Yovel, and Y. Assaf. *In vivo* correlation between axon diameter and conduction velocity in the human brain. *Brain Structure and Function*, 220(3):1777–1788, 2015.

[67] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.

[68] J. R. Hughes. Autism: the first firm finding= underconnectivity? *Epilepsy & Behavior*, 11(1):20–24, 2007.

[69] T. A. Huisman, T. Loenneker, G. Barta, M. E. Bellemann, J. Hennig, J. E. Fischer, and K. A. Il'yasov. Quantitative diffusion tensor MR imaging of the brain: field strength related variance of apparent diffusion coefficient (ADC) and fractional anisotropy (FA) scalars. *European radiology*, 16(8):1651, 2006.

[70] J. Hursh. The properties of growing nerve fibers. *American Journal of Physiology–Legacy Content*, 127(1):140–153, 1939.

[71] A. Ianus, I. Drobnjak, and D. C. Alexander. Model-based estimation of microstructure parameters from diffusion MRI data in a substrate with microscopic anisotropy and a distribution of pore sizes. In *Proc. Intl. Soc. Mag. Reson. Med*, volume 23, page 3035, 2015.

[72] A. Ianuş, B. Siow, I. Drobnjak, H. Zhang, and D. C. Alexander. Gaussian phase distribution approximations for oscillating gradient spin echo diffusion MRI. *Journal of Magnetic Resonance*, 227:25–34, 2013.

[73] G. Innocenti, R. Caminiti, and F. Aboitiz. Comments on the paper by horowitz et al.(2014). *Brain structure and function*, 220(3):1789, 2015.

[74] S. Jbabdi and H. Johansen-Berg. Tractography: where do we go from here? *Brain connectivity*, 1(3):169–183, 2011.

[75] J. H. Jensen and J. A. Helpern. MRI quantification of non-gaussian water diffusion by kurtosis analysis. *NMR in Biomedicine*, 23(7):698–710, 2010.

[76] X. Jiang, H. Li, J. Xie, P. Zhao, J. C. Gore, and J. Xu. Quantification of cell size using temporal diffusion spectroscopy. *Magnetic Resonance in Medicine*, 75(3):1076–1085, 2016.

[77] H. Johansen-Berg and T. E. Behrens. *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy*. Academic Press, 2013.

[78] D. K. Jones. Studying connections in the living human brain with diffusion MRI. *cortex*, 44(8):936–952, 2008.

[79] D. K. Jones. *Diffusion MRI*. Oxford University Press, 2010.

[80] L. S. Kakkar, O. F. Bennett, B. Siow, S. Richardson, A. Ianuş, T. Quick, D. Atkinson, J. B. Phillips, and I. Drobnjak. Low frequency oscillating gradient spin-echo sequences improve sensitivity to axon diameter: An experimental study in viable nerve tissue. *NeuroImage*, 2017.

[81] J. Kershaw, C. Leuze, I. Aoki, T. Obata, I. Kanno, H. Ito, Y. Yamaguchi, and H. Handa. Systematic changes to the apparent diffusion tensor of *in vivo* rat brain measured with an oscillating-gradient spin-echo sequence. *Neuroimage*, 70:10–20, 2013.

[82] F. Kiessling, B. J. Pichler, and P. Hauff. *Small animal imaging*. Springer, 2016.

[83] P. B. Kingsley. Introduction to diffusion tensor imaging mathematics: Part III. tensor calculation, noise, simulations, and optimization. *Concepts in Magnetic Resonance Part A*, 28(2):155–179, 2006.

[84] A. J. Kiruluta. Probing restrictive diffusion dynamics at short time scales. *Journal of Magnetic Resonance*, 192(1):27–36, 2008.

[85] A. P. Koretsky. New developments in magnetic resonance imaging of the brain. *NeuroRx*, 1(1):155–164, 2004.

[86] K. S. Krane and D. Halliday. *Introductory nuclear physics*, volume 465. Wiley New York, 1988.

[87] W. W. Lam, S. Jbabdi, and K. L. Miller. A model for extra-axonal diffusion spectra with frequency-dependent restriction. *Magnetic Resonance in Medicine*, 73(6):2306–2320, 2015.

[88] L. Latour, R. L. Kleinberg, P. P. Mitra, and C. H. Sotak. Pore-size distributions and tortuosity in heterogeneous porous media. *Journal of Magnetic Resonance, Series A*, 112(1):83–91, 1995.

[89] L. L. Latour, P. P. Mitra, R. L. Kleinberg, and C. H. Sotak. Time-dependent diffusion coefficient of fluids in porous media as a probe of surface-to-volume ratio. *Journal of Magnetic Resonance, Series A*, 101(3):342–346, 1993.

[90] L. L. Latour, K. Svoboda, P. P. Mitra, and C. H. Sotak. Time-dependent diffusion of water in a biological model system. *Proceedings of the National Academy of Sciences*, 91(4):1229–1233, 1994.

[91] P. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242(190):191, 1973.

[92] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, et al. White matter tractography using diffusion tensor deflection. *Human brain mapping*, 18(4):306–321, 2003.

[93] D. Le Bihan, R. Turner, C. T. Moonen, and J. Pekar. Imaging of diffusion and microcirculation with gradient sensitization: design, strategy, and significance. *Journal of Magnetic Resonance Imaging*, 1(1):7–28, 1991.

[94] G. Lemberskiy, S. H. Baete, M. A. Cloos, D. S. Novikov, and E. Fieremans. Validation of surface-to-volume ratio measurements derived from oscillating gradient spin echo on a clinical scanner using anisotropic fiber phantoms. *NMR in Biomedicine*, 30(5):e3708, 2017. DOI http://dx.doi.org/10.1002/nbm.3708.

[95] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.

[96] M. H. Levitt. *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons, 2001.

[97] H. Li, J. C. Gore, and J. Xu. Fast and robust measurement of microstructural dimensions using temporal diffusion spectroscopy. *Journal of Magnetic Resonance*, 242:4–9, 2014.

[98] H. Li, X. Jiang, F. Wang, J. Xu, and J. C. Gore. Structural information revealed by the dispersion of ADC with frequency. *Magnetic Resonance Imaging*, 33(9):1083–1090, 2015.

[99] H. Li, X. Jiang, J. Xie, J. O. McIntyre, J. C. Gore, and J. Xu. Time-dependent influence of cell membrane permeability on MR diffusion measurements. *Magnetic Resonance in Medicine*, 75(5):1927–1934, 2016.

[100] Z.-P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. SPIE Optical Engineering Press, 2000.

[101] D. Liewald, R. Miller, N. Logothetis, H.-J. Wagner, and A. Schüz. Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biological cybernetics*, 108(5):541–557, 2014.

[102] C.-P. Lin, V. J. Wedeen, J.-H. Chen, C. Yao, and W.-Y. I. Tseng. Validation of diffusion spectrum magnetic resonance imaging with manganese-enhanced rat optic tracts and *ex vivo* phantoms. *Neuroimage*, 19(3):482–495, 2003.

[103] S. Ljunggren. A simple graphical representation of fourier-based imaging methods. *Journal of Magnetic Resonance (1969)*, 54(2):338–343, 1983.

[104] N. F. Lori, T. E. Conturo, and D. Le Bihan. Definition of displacement probability and diffusion time in q-space magnetic resonance measurements that use finite-duration diffusion-encoding gradients. *Journal of Magnetic Resonance*, 165(2):185–195, 2003.

[105] D. Lurie. A systematic design procedure for selective pulses in NMR imaging. *Magnetic Resonance Imaging*, 3(3):235–243, 1985.

[106] M. Mallett and J. H. Strange. Diffusion measurements using oscillating gradients. *Applied Magnetic Resonance*, 12(2-3):193–198, 1997.

[107] T. Malysz, J. Ilha, P. S. d. Nascimento, K. D. Angelis, B. D. Schaan, and M. Achaval. Beneficial effects of treadmill training in experimental diabetic nerve regeneration. *Clinics*, 65(12):1329–1337, 2010.

[108] P. Mansfield. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics*, 10(3):L55, 1977.

[109] P. Mansfield and P. K. Grannell. NMR 'diffraction' in solids? *Journal of Physics C: solid state physics*, 6(22):L422, 1973.

[110] P. Mansfield and A. A. Maudsley. Medical imaging by NMR. *The British Journal of Radiology*, 50(591):188–194, 1977.

[111] P. Mansfield and I. Pykett. Biological and medical imaging by NMR. *Journal of Magnetic Resonance (1969)*, 29(2):355–373, 1978.

[112] S. Marenco, R. Rawlings, G. K. Rohde, A. S. Barnett, R. A. Honea, C. Pierpaoli, and D. R. Weinberger. Regional distribution of measurement error in diffusion tensor imaging. *Psychiatry Research: Neuroimaging*, 147(1):69–78, 2006.

[113] Matlab optimization toolbox, 2017. The MathWorks, Natick, MA, USA.

[114] J. Mattiello, P. J. Basser, and D. Le Bihan. The b matrix in diffusion tensor echo-planar imaging. *Magnetic Resonance in Medicine*, 37(2):292–300, 1997.

[115] J. Mattiello, P. J. Basser, and D. LeBihan. Analytical expressions for the b matrix in NMR diffusion imaging and spectroscopy. *Journal of Magnetic Resonance, Series A*, 108(2):131–141, 1994.

REFERENCES

[116] K.-D. Merboldt, W. Hanicke, and J. Frahm. Self-diffusion NMR imaging using stimulated echoes. *Journal of Magnetic Resonance (1969)*, 64(3):479–486, 1985.

[117] M. Mercredi and M. Martin. Toward faster inference of micron-scale axon diameters using monte carlo simulations. *Magnetic Resonance Materials in Physics, Biology and Medicine*, pages 1–20, 2018.

[118] M. Mercredi, T. J. Vincent, C. P. Bidinosti, and M. Martin. Assessing the accuracy of using oscillating gradient spin echo sequences with axcaliber to infer micron-sized axon diameters. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 30(1):1–14, 2017.

[119] R. Mezrich. A perspective on k-space. *Radiology*, 195(2):297–315, 1995.

[120] L. Minati and W. P. Weglarz. Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain: A review. *Concepts in Magnetic Resonance Part A*, 30(5):278–307, 2007.

[121] V. E. Miron, T. Kuhlmann, and J. P. Antel. Cells of the oligodendroglial lineage, myelination, and remyelination. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1812(2):184–193, 2011.

[122] P. P. Mitra and P. N. Sen. Effects of microgeometry and surface relaxation on NMR pulsed-field-gradient experiments: simple pore geometries. *Physical Review B*, 45(1):143, 1992.

[123] P. P. Mitra, P. N. Sen, and L. M. Schwartz. Short-time behavior of the diffusion coefficient as a geometrical probe of porous media. *Physical Review B*, 47(14):8565, 1993.

[124] P. P. Mitra, P. N. Sen, L. M. Schwartz, and P. Le Doussal. Diffusion propagator as a probe of the structure of porous media. *Physical review letters*, 68(24):3555, 1992.

[125] S. Mori, B. J. Crain, V. P. Chacko, and P. Van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45(2):265–269, 1999.

[126] K. W. Moser and J. G. Georgiadis. Extraction and validation of correlation lengths from interstitial velocity fields using diffusion-weighted MRI. *Magnetic Resonance Imaging*, 22(2):257–268, 2004.

[127] J. Murday and R. M. Cotts. Self-diffusion coefficient of liquid lithium. *The Journal of Chemical Physics*, 48(11):4938–4945, 1968.

[128] C. Neuman. Spin echo of spins diffusing in a bounded medium. *The Journal of Chemical Physics*, 60(11):4508–4511, 1974.

[129] K. Nicolay, K. P. Braun, R. A. d. Graaf, R. M. Dijkhuizen, and M. J. Kruiskamp. Diffusion NMR spectroscopy. *NMR in Biomedicine*, 14(2):94–111, 2001.

[130] J. S. Nielsen, D. T. B, and H. Lundell. Magnetic resonance temporal diffusion tensor spectroscopy of disordered anisotropic tissue. *Scientific Reports*.

[131] M. Nilsson, J. Lätt, E. Nordh, R. Wirestam, F. Ståhlberg, and S. Brockstedt. On the effects of a varied diffusion time *in vivo*: is the diffusion in white matter restricted? *Magnetic Resonance Imaging*, 27(2):176–187, 2009.

[132] M. Nilsson, J. Lätt, F. Ståhlberg, D. Westen, and H. Hagslätt. The importance of axonal undulation in diffusion MR measurements: a Monte Carlo simulation study. *NMR in Biomedicine*, 25(5):795–805, 2012.

[133] L. Ning, K. Setsompop, C.-F. Westin, and Y. Rathi. New insights about time-varying diffusivity and its estimation from diffusion MRI. *Magnetic Resonance in Medicine*, 78(2):763–774, 2017.

[134] C. Njiokiktjien, L. De Sonneville, and J. Vaal. Callosal size in children with learning disabilities. *Behavioural brain research*, 64(1):213–218, 1994.

[135] D. S. Novikov, E. Fieremans, J. H. Jensen, and J. A. Helpern. Characterizing microstructure of living tissues with time-dependent diffusion. *arXiv preprint arXiv:1210.3014*, 2012.

[136] R. Ordidge, P. Mansfield, and R. Coupland. Rapid biomedical imaging by NMR. *The British journal of radiology*, 54(646):850–855, 1981.

[137] E. Özarslan, T. M. Shepherd, B. C. Vemuri, S. J. Blackband, and T. H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31(3):1086–1103, 2006.

[138] G. J. Parker and D. C. Alexander. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *IPMI*, volume 18, pages 684–695. Springer, 2003.

[139] E. C. Parsons, M. D. Does, and J. C. Gore. Modified oscillating gradient pulses for direct sampling of the diffusion spectrum suitable for imaging sequences. *Magnetic Resonance Imaging*, 21(3):279–285, 2003.

[140] E. C. Parsons, M. D. Does, and J. C. Gore. Temporal diffusion spectroscopy: theory and implementation in restricted systems using oscillating gradients. *Magnetic Resonance in Medicine*, 55(1):75–84, 2006.

[141] W. Perrault, T. Duval, and J. Cohen-Adad. Comparison of NOGSE and PGSE sequences for axon diameter estimation. In *Proceedings of the 23rd Annual meeting of the ISMRM*, page 2884, Toronto, Canada, 2015.

[142] J. Piven, J. Bailey, B. J. Ranson, and S. Arndt. An MRI study of the corpus callosum in autism. *American Journal of Psychiatry*, 154(8):1051–1056, 1997.

[143] S. Portnoy, N. Fichtner, C. Dziegielewski, M. Stanisz, and G. Stanisz. In vitro detection of apoptosis using oscillating and pulsed gradient diffusion magnetic resonance imaging. *NMR in biomedicine*, 27(4):371–380, 2014.

[144] S. Portnoy, J. Flint, S. Blackband, and G. Stanisz. Oscillating and pulsed gradient diffusion magnetic resonance microscopy over an extended b-value range: Implications for the characterization of tissue microstructure. *Magnetic Resonance in Medicine*, 69(4):1131–1145, 2013.

[145] W. S. Price. Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part 1. basic theory. *Concepts in Magnetic Resonance Part A*, 9(5):299–336, 1997.

[146] P. Randall. Schizophrenia, abnormal connection, and brain evolution. *Medical hypotheses*, 10(3):247–280, 1983.

[147] O. Reynaud, K. Winters, D. Hoang, Y. Wadghiri, D. Novikov, and S. Kim. *In vivo* and *ex vivo* characterization of extracellular space (ECS) in mouse GBM using PGSE and OGSE. In *Proceedings of the 23rd scientific meeting, International Society for Magnetic Resonance in Medicine, Toronto*, volume 344, 2015.

[148] O. Reynaud, K. V. Winters, D. M. Hoang, Y. Z. Wadghiri, D. S. Novikov, and S. G. Kim. Pulsed and oscillating gradient mri for assessment of cell size and extracellular space (pomace) in mouse gliomas. *NMR in Biomedicine*, 29(10):1350–1363, 2016.

[149] O. Reynaud, K. V. Winters, D. M. Hoang, Y. Z. Wadghiri, D. S. Novikov, and S. G. Kim. Surface-to-volume ratio mapping of tumor microstructure using oscillating gradient diffusion weighted imaging. *Magnetic Resonance in Medicine*, 76(1):237–247, 2016.

[150] D. Rice and S. Barone Jr. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environmental health perspectives*, 108(Suppl 3):511, 2000.

[151] J. Ritchie. On the relation between fibre diameter and conduction velocity in myelinated nerve fibres. *Proceedings of the Royal Society of London B: Biological Sciences*, 217(1206):29–35, 1982.

[152] I. Ronen, M. Budde, E. Ercan, J. Annese, A. Techawiboonwong, and A. Webb. Microstructural organization of axons in the human corpus callosum quantified by diffusion-weighted magnetic resonance spectroscopy of n-acetylaspartate and post-mortem histology. *Brain Structure and Function*, 219(5):1773–1785, 2014.

[153] K. E. Sakaie and M. J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted MRI. *NeuroImage*, 34(1):169–176, 2007.

[154] M. Schachter, M. Does, A. Anderson, and J. Gore. Measurements of restricted diffusion using an oscillating gradient spin-echo sequence. *Journal of Magnetic Resonance*, 147(2):232–237, 2000.

[155] P. N. Sen. Time-dependent diffusion coefficient as a probe of the permeability of the pore wall. *The Journal of Chemical Physics*, 119(18):9871–9876, 2003.

[156] P. N. Sen. Diffusion and tissue microstructure. *Journal of Physics: Condensed Matter*, 16(44):S5213, 2004.

[157] P. N. Sen. Time-dependent diffusion coefficient as a probe of geometry. *Concepts in Magnetic Resonance Part A*, 23(1):1–21, 2004.

[158] F. Sepehrband, D. C. Alexander, N. D. Kurniawan, D. C. Reutens, and Z. Yang. Towards higher sensitivity and stability of axon diameter estimation with diffusion-weighted MRI. *NMR in Biomedicine*, 29(3):293–308, 2016.

[159] N. Shemesh, G. A. Álvarez, and L. Frydman. Measuring small compartment dimensions by probing diffusion dynamics via non-uniform oscillating-gradient spin-echo (NOGSE) NMR. *Journal of Magnetic Resonance*, 237:49–62, 2013.

[160] B. Siow, I. Drobnjak, A. Chatterjee, M. F. Lythgoe, and D. C. Alexander. Estimation of pore size in a microstructure phantom using the optimised gradient waveform diffusion weighted NMR sequence. *Journal of Magnetic Resonance*, 214:51–60, 2012.

[161] G. J. Stanisz, G. A. Wright, R. M. Henkelman, and A. Szafer. An analytical model of restricted diffusion in bovine optic nerve. *Magnetic Resonance in Medicine*, 37(1):103–111, 1997.

[162] E. Stejskal. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, 43(10):3597–3603, 1965.

[163] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292, 1965.

[164] J. Stepišnik. Measuring and imaging of flow by NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 17:187–209, 1985.

[165] J. Stepišnik. Time-dependent self-diffusion by NMR spin-echo. *Physica B: Condensed Matter*, 183(4):343–350, 1993.

[166] J. Stepišnik. A new view of the spin echo diffusive diffraction in porous structures. *EPL (Europhysics Letters)*, 60(3):453, 2002.

[167] J. Stepišnik, A. Mohorič, and A. Duh. Diffusion and flow in a porous structure by the gradient spin echo spectral analysis. *Physica B: Condensed Matter*, 307(1):158–168, 2001.

[168] P. Stilbs. Fourier transform pulsed-gradient spin-echo studies of molecular diffusion. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 19(1):1–45, 1987.

[169] A. L. Sukstanskii. Exact analytical results for ADC with oscillating diffusion sensitizing gradients. *Journal of Magnetic Resonance*, 234:135–140, 2013.

[170] A. Szafer, J. Zhong, and J. C. Gore. Theoretical model for water diffusion in tissues. *Magnetic Resonance in Medicine*, 33(5):697–712, 1995.

[171] E. T. Tan, S.-K. Lee, P. T. Weavers, D. Graziani, J. E. Piel, Y. Shu, J. Huston III, M. A. Bernstein, and T. K. Foo. High slew-rate head-only gradient for improving distortion in echo planar imaging: Preliminary experience. *Journal of Magnetic Resonance Imaging*, 44(3):653–664, 2016.

[172] J. E. Tanner and E. O. Stejskal. Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. *The Journal of Chemical Physics*, 49(4):1768–1777, 1968.

REFERENCES

[173] I. Tasaki, K. Ishii, and H. Ito. On the relation between the conduction-rate, the fiber-diameter and the internodal distance of the medullated nerve fiber. *Jpn J Med Sci III, Biophysics*, 9(1):189–99, 1943.

[174] J. D. Thiessen. *Development and Application of Quantitative MRI Methods for Assessing White Matter Integrity in the Mouse Brain*. PhD thesis, University of Manitoba, Manitoba, Canada, 2012.

[175] J. D. Thiessen, Y. Zhang, H. Zhang, L. Wang, R. Buist, M. R. Del Bigio, J. Kong, X.-M. Li, and M. Martin. Quantitative MRI and ultrastructural examination of the cuprizone mouse model of demyelination. *NMR in Biomedicine*, 26(11):1562–1581, 2013.

[176] H. C. Torrey. Bloch equations with diffusion terms. *Phys. Rev.*, 104:563–565, Nov 1956.

[177] J.-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage*, 35(4):1459–1472, 2007.

[178] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.

[179] D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, 2004.

[180] D. S. Tuch et al. *Diffusion MRI of complex tissue structure*. PhD thesis, Massachusetts Institute of Technology, 2002.

[181] D. S. Tuch, T. G. Reese, M. R. Wiegell, and V. J. Wedeen. Diffusion MRI of complex neural architecture. *Neuron*, 40(5):885–895, 2003.

[182] D. B. Twieg. The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Medical physics*, 10(5):610–621, 1983.

[183] A. T. Van, S. J. Holdsworth, and R. Bammer. *In vivo* investigation of restricted diffusion in the human brain with optimized oscillating diffusion gradient encoding. *Magnetic Resonance in Medicine*, 71(1):83–94, 2014.

[184] P. Vangelderen, D. DesPres, P. Vanzijl, and C. Moonen. Evaluation of restricted diffusion in cylinders. phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance, Series B*, 103(3):255–260, 1994.

[185] J. Virtanen, H. Uusitalo, A. Palkama, and H. Kaufman. The effect of fixation on corneal endothelial cell dimensions and morphology in scanning electron microscopy. *Acta ophthalmologica*, 62(4):577–585, 1984.

[186] S. G. Waxman. Determinants of conduction velocity in myelinated nerve fibers. *Muscle & nerve*, 3(2):141–150, 1980.

[187] V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.

[188] C.-F. Westin, H. Knutsson, O. Pasternak, F. Szczepankiewicz, E. Özarslan, D. van Westen, C. Mattisson, M. Bogren, L. J. O'donnell, M. Kubicki, et al. Q-space trajectory imaging for multidimensional diffusion MRI of the human brain. *Neuroimage*, 135:345–362, 2016.

[189] N. S. White, C. R. McDonald, N. Farid, J. Kuperman, D. Karow, N. M. Schenker-Ahmed, H. Bartsch, R. Rakow-Penner, D. Holland, A. Shabaik, et al. Diffusion-weighted imaging in cancer: physical foundations and applications of restriction spectrum imaging. *Cancer research*, 74(17):4638–4652, 2014.

[190] D. Wu, L. J. Martin, F. J. Northington, and J. Zhang. Oscillating gradient diffusion MRI reveals unique microstructural information in normal and hypoxia-ischemia injured mouse brains. *Magnetic Resonance in Medicine*, 72(5):1366–1374, 2014.

[191] Y.-C. Wu, A. S. Field, and A. L. Alexander. Computation of diffusion function measures in q-space using magnetic resonance hybrid diffusion imaging. *IEEE transactions on medical imaging*, 27(6):858–865, 2008.

[192] J. Xu, M. D. Does, and J. C. Gore. Quantitative characterization of tissue microstructure with temporal diffusion spectroscopy. *Journal of Magnetic Resonance*, 200(2):189–197, 2009.

[193] J. Xu, M. D. Does, and J. C. Gore. Sensitivity of MR diffusion measurements to variations in intracellular structure: effects of nuclear size. *Magnetic Resonance in Medicine*, 61(4):828–833, 2009.

[194] J. Xu, M. D. Does, and J. C. Gore. Dependence of temporal diffusion spectra on microstructural properties of biological tissues. *Magnetic Resonance Imaging*, 29(3):380–390, 2011.

[195] J. Xu, H. Li, K. D. Harkins, X. Jiang, J. Xie, H. Kang, M. D. Does, and J. C. Gore. Mapping mean axon diameter and axonal volume fraction by MRI using temporal diffusion spectroscopy. *NeuroImage*, 103:10–19, 2014.

[196] J. Xu, K. Li, R. A. Smith, J. C. Waterton, P. Zhao, H. Chen, M. D. Does, H. C. Manning, and J. C. Gore. Characterizing tumor response to chemotherapy at various length scales using temporal diffusion spectroscopy. *PLOS One*, 7(7):e41714, 2012.

[197] D. A. Yablonskiy, G. L. Bretthorst, and J. J. Ackerman. Statistical model for diffusion attenuated MR signal. *Magnetic Resonance in Medicine*, 50(4):664–669, 2003.

[198] D. A. Yablonskiy and A. L. Sukstanskii. Theoretical models of the diffusion weighted MR signal. *NMR in Biomedicine*, 23(7):661–681, 2010.

[199] H. Zhang, P. L. Hubbard, G. J. Parker, and D. C. Alexander. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *Neuroimage*, 56(3):1301–1315, 2011.

[200] L. J. Zielinski and M. D. Hürlimann. Probing short length scales with restricted diffusion in a static gradient using the CPMG sequence. *Journal of Magnetic Resonance*, 172(1):161–167, 2005.

[201] L. J. Zielinski and P. N. Sen. Effects of finite-width pulses in the pulsed-field gradient measurement of the diffusion coefficient in connected porous media. *Journal of Magnetic Resonance*, 165(1):153–161, 2003.