Validation of MRtrix tractography for clinical use

Olivia Lannan

A thesis submitted in fulfilment of the requirements for the degree of
Master of Philosophy

June 2021

School of Physical Sciences
The University of Adelaide
North Terrace Campus
Adelaide, South Australia 5005

Contents

Abstract

Tractography is a technique which uses non-invasive diffusion magnetic resonance imaging (dMRI) to model the structural connections of the brain *in vivo*. Since it is not known to what extent the exact positions of white matter tracts vary between individuals, and how scan parameters affect the reconstructed location of tracts, the performance of tractography algorithms cannot be assessed by comparing them against a standardised model of white matter in the human brain. Until technology has advanced to a stage where these issues can be resolved, if this is indeed possible, the quality of tractograms must be assessed via alternative means if they are to be used in a clinical setting.

MRtrix is a software package which offers a suite of tools for tractography. To investigate the clinical viability of tractography, the effect of adding artificial noise and patient movement to the input dMRI scans was examined. The performance of the following three MRtrix tractography algorithms was considered: the second-order integration over fibre orientation distributions (iFOD2) algorithm, the spherical deconvolution streamlines tractography (SD_STREAM) algorithm, and the probabilistic tractography (Tensor_Prob) algorithm.

The accuracy of each tractogram and its robustness to added noise and motion was assessed through both quantitative and qualitative means. Streamline length histograms, polar plots, a resemblance metric, the RMS difference between tractmaps and polar plots, and false positive and negative rates were used to quantitatively measure tractogram quality. Each of these techniques provide measures of quality relative to an initial unmodified scan. Potentially, these techniques could be extended to provide a stand-alone metric of tractogram quality to eliminate operator dependence in diagnosing, treating and managing conditions associated with the integrity of white matter pathways in the brain.

The tractography algorithm found to be most suitable for clinical applications was iFOD2, which provided the highest level of detail compared with the SD_STREAM and Tensor_Prob algorithms. iFOD2 was capable of consistently resolving tracts with added noise of up to about an added white Gaussian noise (WGN) of power 20, or a SNR of about 8 for the majority of patients (calculated using the single image method). Additionally, the iFOD2 algorithm successfully reconstructed tractograms with added translations up to 5 mm, and rotations up to 10 degrees, despite the reconstructed tracts being spatially shifted.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Olivia Lannan
June 2021

Acknowledgements

Throughout my research I have received a great deal of guidance and support. I would like to express my sincere appreciation to those who have contributed to my thesis in some way, whether it be through providing inexhaustible encouragement, supporting me in my research or shaping it through input of time and ideas.

I would first like to thank my supervisors, Dr. Alexandre Santos, Dr. Donald McRobbie and Dr. Judith Pollard, who have helped shape my thesis by facilitating productive discussions and providing insightful suggestions and feedback every step of the way. I cannot thank them enough for all the invaluable knowledge and experience they have shared and for their patience, support and understanding.

I would also like to acknowledge Dr. Mark Agzarian for sharing his ideas, assisting me with qualitative tractogram scoring and offering helpful suggestions for my research. Having the opportunity to meet and discuss tractography with you was illuminating, and I strongly admire the passion he has for this area of research.

A heartfelt thanks to my family and close friends for their endless encouragement and continual support. They have all inspired and motivated me in different ways throughout my studies and I appreciate them more than words can express.

Lastly, to my other half, Justin, thank you for being my biggest source of encouragement, and for providing me with food, nourishment, regular motivational speeches and open arms.

Abbreviations

3D three-dimensional

AC anterior commissure

ADC apparent diffusion coefficient

BF brute-force

CC corpus callosum

CHARMED the composite hindered and restricted model of diffusion

CP cerebellar penduncles

CS corticospinal tracts

CSD constrained spherical deconvolution

CSF cerebrospinal fluid

DKI diffusion kurtosis imaging

dMRI diffusion magnetic resonance imaging

dODF diffusion orientation density function

DTI diffusion tensor imaging

dtMRI diffusion tensor magnetic resonance imaging

DW diffusion-weighted

DWI diffusion-weighted magnetic resonance imaging

FA fractional anisotropy

FACT fibre assignment by continuous tracking

fMRI functional MRI

FOD fibre orientation distribution

fODF fibre orientation density function

FOV field of view

LF longitudinal fasciculi

MAP-MRI mean apparent propagator MRI

ML medial lemnisci

MR magnetic resonance

MRI magnetic resonance imaging

NMR nuclear magnetic resonance

NODDI neurite orientation dispersion and density imaging

NPA narrow pulse approximation

OR optic radiations

OT optic tracts

PET positron emission tomography

PF medial lemnisci

PMC prospective motion correction

RF radio frequency

RMS root-mean-square

ROI region of interest

ROIs regions of interest

SD spherical deconvolution

SNR signal to noise ratio

TE echo time

UF U fibres

WGN white Gaussian noise

List of Figures

CHAPTER 1

Introduction

Diffusion magnetic resonance imaging (dMRI) harnesses the principles of diffusion to non-invasively probe the white matter fibres of the brain in vivo. This voxellized information about the brain's fibre orientations can be used by computer algorithms to reconstruct pathways of brain connectivity. This process is referred to as tractography, and results in the generation of three-dimensional (3D) white matter maps or 'tractograms', which can reveal microstructural details about the composition, integrity and orientation of white matter fibres in the brain.

This information, often inaccessible to other imaging modalities, has proven instrumental for enabling early diagnosis, faster intervention and more effective management of specific neurological conditions for improved patient outcomes. Other potential applications of tractography include aiding neurosurgeons in preoperative planning, and informing studies of neuronal development, activation and aging.

However, there is some controversy surrounding the interpretation of tractography results. In particular, tractography has increasingly been used to derive measures of the the strength of connection between different areas of the brain, but this approach has been shown to be ill-informed (Jeurissen et al., 2017). Furthermore, the accuracy of a tractogram is not simple to measure given that there is no gold standard model of white matter within the brain to compare against, nor an ideal model of signal generation in tractography (Hubbard and Parker, 2009, Wakana et al., 2007). A number of methods for tractography have been proposed, each with its own inherent limitations (Basser and Pierpaoli, 1996, DellAcqua and Tournier, 2018). However, some approaches have proven to be more robust than others (DellAcqua and Tournier, 2018, Tournier et al., 2008). Here, we investigate the performance of a small subset of tractography algorithms available in the software package MRtrix, namely the second-order integration over fibre orientation distributions (iFOD2) algorithm, the spherical deconvolution streamlines tractography (SD_STREAM) algorithm, and the probabilistic tractography (Tensor_Prob) algorithm.

1.1 What diffusion tells us about the brain

Diffusion-weighted magnetic resonance imaging (DWI) is an imaging method which utilises measurements of the Brownian motion of water molecules in tissues to generate contrast. In 1905, Albert Einstein laid the groundwork for this technique by developing a theoretical model to explain Brownian motion (Einstein, 1956). Brownian motion is the random motion of molecules which results from their thermal energy. The net movement of molecules over time as a result of this random motion is referred to as molecular diffusion. At the most fundamental level all biological processes are driven by interactions of molecules. For this to arise molecules must be brought into close proximity to one another, primarily through diffusion. Although the exact mechanisms which govern this process have not yet been elucidated, it is possible to harness the principles of diffusion to characterise tissue microstructure *in vivo* noninvasively. This is the basis upon which DWI is founded (Chenevert et al., 1990, Le Bihan, 2014, Pierpaoli and Basser, 1996, Reese et al., 1996).

Free diffusion of water is isotropic in character, in that there is no favoured direction for net diffusion. If diffusion has a directional dependence it is termed anisotropic. Anisotropic diffusion can occur in biological tissues, since molecular motion can be restricted by structures at the cellular level such as cell membranes, macromolecules and fibres, and may also be affected by active processes within the tissues (Beaulieu, 2002). Diffusion in biological tissues is referred to as apparent diffusion due to these external influences. Water diffusion may be restricted within cells, traverse permeable cell membranes, or follow erratic pathways through extracellular space.

One region of the body within which diffusion plays a key role is the brain. The grey matter of the brain consists of neural cell bodies, axon terminals, dendrites and nerve synapses. Within the grey matter, molecules are expected to exhibit isotropic diffusion since these formations do not have an oriented fibre structure. The white matter of the brain is primarily composed of myelinated fibres bundled into tracts which conduct nerve signals and connect different regions of grey matter. Consequently, molecules diffuse through white matter tracts preferentially in the direction parallel to the orientation of axonal fibres due to axon membranes restricting perpendicular movement. This means that white matter yields anisotropic diffusion signals when imaged with a diffusion-weighted (DW) magnetic resonance imaging (MRI) sequence, with the directions of fibre bundles corresponding to local maxima in the diffusion orientation density function (dODF) (Glenn et al., 2016).

By measuring this anisotropic diffusion and mapping the white matter bundles, it is possible to evaluate white matter integrity and gain an insight into the connectivity of the brain. This could supply information for various clinical applications such as neurosurgery and illuminate how brain dysfunction occurs in abnormal development, ageing, mental health disorders, addiction or neurological disease. DWI also has promise for use in oncological applications, with the potential to aid tumour detection, staging, monitoring and prediction of treatment responses.

1.2 Diffusion tensor magnetic resonance imaging

MRI harnesses the principles of nuclear magnetic resonance (NMR) to generate high quality images of human anatomy and physiology. Elementary particles carry an intrinsic form of angular momentum known as spin. The particles which comprise the human body, electrons, protons and neutrons, each have a spin of 1/2. The total magnetic field of a nucleus is referred to as its magnetic moment. In atoms with even numbers of protons and even numbers of neutrons the nucleons pair with opposite spins, such that the nucleus has no net magnetic moment. For atoms with odd numbers of nucleons, the unpaired proton and/or neutron contribute(s) to the net magnetic moment of the nucleus.

If an external magnetic field is applied to a proton, it will assume either a low energy state with its spin parallel to the magnetic field, or a higher energy state with its spin antiparallel to the field (Bushberg et al., 2012). By supplying a discrete amount of energy to the proton, it can be induced to transition from the low to the high energy state. The change from the high to the low energy state will result in the emission of a photon of equivalent energy. The frequency of the photon absorbed or emitted in such a transition is the Larmor frequency.

The relative number of spins in the lower energy state is dependent on the strength of the external magnetic field. At body temperature, the lower energy state is slightly favoured, with approximately 4 spins per million more in the lower energy state at 1.5 Tesla (McRobbie et al., 2006). Since a typical voxel contains about 10^{21} protons, the net magnetic moment is sufficient to generate an observable NMR signal. The most prominent source of this signal is hydrogen, making up over 60% of the atoms in the human body, 99.985% of these being ^{1}H.

The net magnetisation vector due to the spin packets in a voxel may be resolved into three components. Assuming the applied magnetic field is in the z direction, M_z is the longitudinal magnetisation which peaks at equilibrium. M_{xy} is the magnitude of the magnetisation vector perpendicular to the applied magnetic field. This quantity is referred to as the transverse magnetisation, and has a value of zero at equilibrium. The M_z component may be reduced through the application of a radio frequency (RF) pulse at the Larmor frequency. This causes more protons to transition to the high energy state and spin packets to become coherent in phase. The time taken for M_z to recover from zero to $(1 - 1/e)M_0$ is the T1 relaxation constant. Once the RF pulse is applied the spin packets gradually lose their phase coherence due to inhomogeneity of the magnetic field or variations in molecular structure that bring about local magnetic field changes. As the Larmor frequency is dependent on the magnetic field strength, this causes spin packets to precess at different rates. The rotation of M_{xy} induces a current in the receiver coil, producing a damped sinusoidal signal. The time taken for the signal to fall to $1/e$ of the initial signal is termed the free induction decay or T2* signal, which incorporates the effects of inhomogeneities in the external magnetic field or contrast agents and dephasing from micromagnetic inhomogeneities. The decay time excluding these external factors is T2, which can be acquired in spin-echo pulse sequences. T1 and T2 are affected by molecular motion, size and interaction and chemical shift. T1 is also strongly affected by the magnetic field strength. The relaxation times and spin density are fundamental tissue properties and MRI exploits the differences in these properties between different types of tissues to generate contrast.

A pulse sequence is a series of RF pulses that are applied to generate a certain type of NMR signal. The timing, order, polarity and repetition frequency are selected in a particular pulse sequence to modify the effect of T1, T2 or spin density on the signal. The three main pulse sequences used in MRI are the spin-echo, inversion recovery and gradient recalled echo sequences. In a spin-echo pulse sequence, a 90 degree pulse is applied, where a suitable RF is used to reduce M_z to zero and rotate the net magnetisation into the transverse plane, such that M_{xy} is maximised. The magnitude of the net magnetisation vector decays over time as it precesses about the z-axis, due to local magnetic field inhomogeneities that cause spin packets to precess at slightly different rates. After a time delay a 180 degree pulse is applied which inverts the spin system. The net magnetisation decays, passing through zero to return to equilibrium, and the transverse magnetisation rephases due to the external non-homogeneous magnetic field acting on it in an opposing direction. Local magnetic field fluctuations that are related to the T2 effect remain unaffected, so the maximum value of the free induction decay signal is dependent on T2. The time of echo (TE) gives the time for the signal to fall then rise again after applying the 180 degree pulse. The 180 degree pulse is therefore applied at TE/2.

MRI can be made sensitive to molecular diffusion through the application of gradient pulses to pulse sequences such as spin-echo. In the mid-1980s, Denis Le Bihan first utilised this technique to visualise water diffusion in the brain (Le Bihan, 2014). Two diffusion gradients of equal amplitude and opposite effect are added to the MRI spin echo sequence symmetric to the 180 degree RF pulse. The first gradient dephases the protons. The second reverses the changes imposed by the first gradient (rephasing). If a proton is stationary during the time spanned by the application of the two gradient pulses, then the effects of the gradients will cancel and there will be no signal attenuation, corresponding to a sustained signal (Chilla et al., 2015). If a proton undergoes diffusion during application of the gradients, there will be incomplete rephasing of proton spins, which causes the signal to be attenuated according to the Stejskal-Tanner equation:

$$S(b) = S_0 \exp(-bD) \tag{1.1}$$

and

$$S_0 = P \exp(-TE/T_2) \tag{1.2}$$

where $S(b)$ is the signal received for the gradient value b, which reflects the strength, duration, and spacing of the gradients used to generate DW images. S_0 is the signal intensity without diffusion sensitizing gradients (i.e. $b = 0$), P is a function of the proton density, D is the apparent diffusion coefficient, and b is given by:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \tag{1.3}$$

where γ is the gyromagnetic ratio of a hydrogen proton (42.58 MHz/T), G is the magnitude of the applied gradient, δ is the duration of the applied gradient and Δ is the time between the two gradient applications. Signal loss therefore depends on the time between pulses and the strength and duration of applied gradients. For a given set of these parameters, in regions where there is rapid diffusion the signal is significantly reduced, whereas in slow diffusion regions it is only slightly reduced. It is these differences in diffusion rates for different tissues which generate contrast in DWI, which is not visible in other MRI sequences.

Through the acquisition of multiple MR images with varied diffusion-sensitising gradient directions, the orientations of the coherent neuronal axon populations can be determined. A map of the anatomical connections in the brain can then be reconstructed by using these orientations to define tangents along continuous trajectories called streamlines. This is the basis of tractography.

1.3 Tractography

1.3.1 Models for white matter

The first proposed and most common model used to derive fibre orientations within elements of three-dimensional space or ‘voxels’ to map brain connectivity is the diffusion tensor model (Basser et al., 1994). This model assumes that diffusion is free or unhindered, which is an approximation for the *in vivo* case. Since the diffusion tensor has only one major eigenvector it is only capable of characterising the direction of one coherently oriented fibre population in each voxel with highest diffusivity. This provides a reasonable description for fibres with a unique orientation. However, this model has been demonstrated to perform poorly in voxels containing complex fibre architectures, such as interdigitating, diverging, bending, crossing and kissing fibres, or partial volume contamination between adjacent fibre populations (Smith et al., 2012).

Given that voxel sizes in diffusion images are typically large, the proportion of voxels containing complex fibre architecture is significant, reportedly ranging between 63% and 90% (Jeurissen et al., 2012). Based on this finding, it is evident that higher order models are needed to accurately resolve fibre populations in the brain.

A variety of alternative models have been proposed for the determination of fibre orientations, particularly in voxels containing crossing fibres. These include multiple tensor fitting, q-space models and Q-ball imaging (Mori and J.D., 2014). The multiple tensor fitting model fits several diffusion tensors to high angular resolution DWI data. However, the number of fibre populations in each voxel must be input into the model. It also becomes unstable if more than two populations are present. In the q-space model, diffusion is represented by a probability density function over the spin displacement, which requires a large number of data samples. This is impractical in a clinical setting. Q-ball imaging reduces the number of data samples needed. However, these techniques do not provide actual fibre orientations, as the relationship between the spin displacement profile and fibre orientation has not yet been elucidated. Instead, these techniques rely on the assumption that the spin displacement profile peaks correspond to the direction of the fibre population.

It is evident that each of these methods has its inherent limitations, due to the assumptions and approximations made in each. Most of these algorithms also require lengthy acquisition or computation times that are not optimal for clinical use. An alternative means of resolving crossing fibres is through constrained spherical deconvolution (CSD), which has been shown to produce more anatomically realistic tracts than diffusion tensor imaging (DTI) (Toselli et al., 2017). This method has been implemented for single-tissue and multi-tissue tractography applications in the MRtrix algorithms tournier and dhollander, which have been employed in this study (Tournier et al., 2012). The details of these techniques are discussed further in Chapter 2.

1.3.2 MRtrix

The MRtrix software package was developed by Tournier et al. (Tournier et al., 2012) for diffusion MRI processing, analysis and visualisation, with a core focus on optimising consistency, performance and stability. This software provides an array of tools to perform various kinds of dMRI analysis. MRtrix version 3.0 was utilised to generate all tractograms used in this study.

MRtrix estimates the fibre orientation distribution directly from high angular resolution diffusion-weighted MR data. It is assumed that white matter fibre bundles have identical diffusion characteristics, such that any difference in diffusion anisotropy can be attributed to partial volume effects. The average displacement of water molecules in a diffusion-weighted experiment is of the order of 10 micrometres, so it can also be assumed that there is negligible exchange between spatially distinct fibre bundles. Since the total signal is given by the sum of the independent diffusion-weighted signals emanating from orientationally distinct regions, there is effectively no exchange between orientationally distinct sections of a fibre bundle. Along with the first assumption this suggests that the diffusion-weighted signal attenuation measured from a single coherently oriented fibre population can be represented by an axially symmetric response function. This response function represents the diffusion-weighted attenuation profile for a typical fibre bundle. Since the diffusion weighted signal attenuation measured over a sphere's surface can be expressed as the convolution over the sphere of a response function with the fibre orientation density function (ODF), the fibre ODF can be acquired through spherical deconvolution. MRtrix therefore recovers the fibre ODF in regions of multiple fibre crossing.

MRtrix then performs tractography, a modelling technique used to visually represent neural tracts using diffusion-weighted image data. Since DWI is an inherently low signal-to-noise ratio (SNR) technique, the precision of estimated fibre orientations is limited. As a result, tractography algorithms may delineate pathways that do not exist (false positives) or fail to identify existing tracts (false negatives). Most approaches use deterministic streamlines tracking algorithms that delineate one pathway of best fit. However, several studies have demonstrated that approaches using probabilistic tractography algorithms are superior

(Moldrich et al., 2010). Probabilistic algorithms trace out multiple possible paths, visualising the uncertainty in the fibre orientation estimates. For further information refer to Chapter 2. In MRtrix both deterministic and probabilistic tracking algorithms are available, enabling more effective delineation of white matter tracts. An additional concern in clinical applications is the necessity to manually seed the starting and/or target point for tracking. This form of operator dependence can be eliminated in MRtrix through whole-brain tracking using CSD.

1.3.3 MRtrix's tractography algorithms

MRtrix offers a diverse selection of tracking algorithms to perform tractography, each requiring different inputs from the user. The default algorithm, titled Second-order Integration over Fibre Orientation Distributions (iFOD2), is a probabilistic algorithm which accepts a Fibre Orientation Distribution (FOD) image represented in the Spherical Harmonic (SH) basis as input. This algorithm constructs candidate streamline paths and samples the underlying (trilinear-interpolated) FOD amplitudes along them. A streamline is more likely to follow a path where FOD amplitudes are large, though they may occasionally traverse orientations where FOD amplitudes are small, provided the amplitude is above the FOD threshold along the entire path.

Another available algorithm is the Tensor_Prob algorithm, which is also probabilistic but requires a 4D diffusion-weighted image (DWI) series as input. Within each image voxel, the algorithm performs a residual bootstrap. These data are then sampled via trilinear interpolation at each streamline step, the diffusion tensor model is fitted and the streamline follows the orientation of the principal eigenvector of that tensor.

MRtrix's SD_STREAM algorithm performs streamlines tractography based on Spherical Deconvolution (SD). Unlike iFOD2 and Tensor_Prob, this algorithm is deterministic, but similarly to iFOD2 it accepts a FOD image represented in the SH basis as input. At each streamline step, the local (trilinear-interpolated) FOD is sampled, and a Newton optimisation on the sphere is performed on the current streamline tangent orientation. This enables the orientation of the nearest FOD amplitude peak to be located.

A tractogram seeded at the base of the brain stem produced by the three algorithms, iFOD2, SD_STREAM and Tensor_Prob, from the same input scan are shown in Figure 1.1. A whole-brain tractogram for each algorithm is also shown in Figure 1.2.

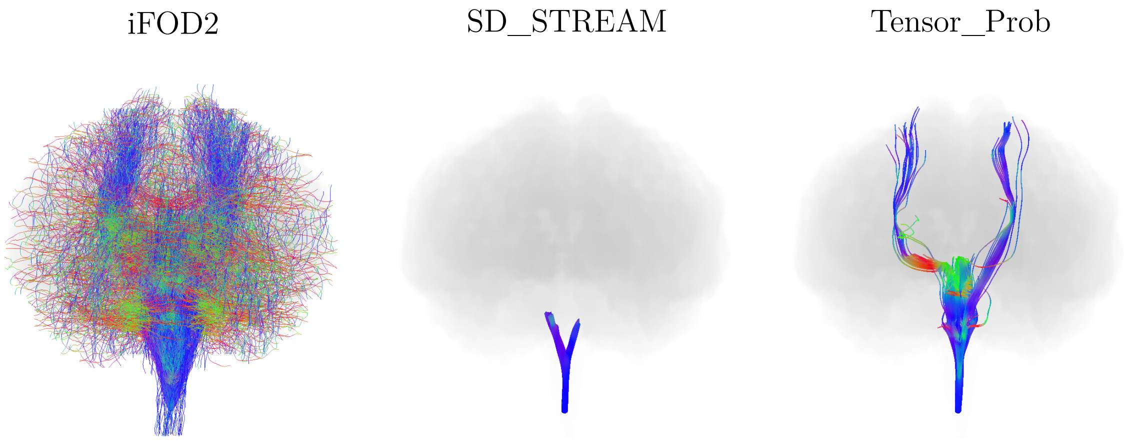


Figure 1.1: 5k streamline tractograms seeded from a point at the base of the brain stem for patient 1 using the MRtrix algorithms iFOD2 (left), SD_STREAM (centre) and Tensor_Prob (right). The patient 1 scan had a b value of 1000, with 2 b=0 images. The colours of the streamlines indicate their orientations, with red being left-right, green being dorsal-ventral and blue being cranial-caudal. It is evident that the iFOD2 algorithm provides the highest level of detail. However, further analysis is required to validate whether these complex structures are reproducible or a result of noise in the input dMRI. The Tensor_Prob produces a tract with a more refined structure, which may be of use for segmentation or extraction of particular features from a dMRI scan. The SD_STREAM algorithm does not delineate intricate structures, but could be applied to search for regions of high connectivity or strong fibre coherence, provided that the algorithm is similarly validated.

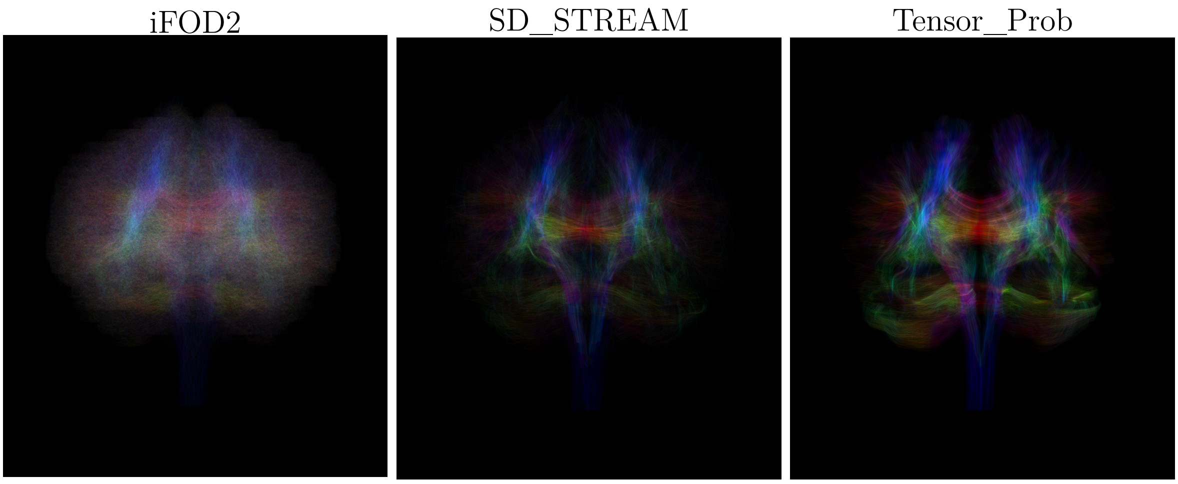


Figure 1.2: 50k streamline whole brain tractograms seeded at random points within a mask of the brain for patient 1 using the MRtrix algorithms iFOD2 (left), SD_STREAM (centre) and Tensor_Prob (right). The streamlines have been set to the same level of transparency, to better visualise the density in different regions of the brain and compare the overall distribution of streamlines between the three algorithms.

1.3.4 Tractography algorithm options

Under default settings, MRtrix produces a fixed number of streamlines by seeding streamlines in random positions until the target number of streamlines has been successfully reconstructed, or until the maximum number of seeds has been exceeded (this value is by default 1000 times the desired number of selected streamlines). As a compromise between obtaining a statistically significant set of data (i.e. enough streamlines to provide a reasonable representation of the structure of a tract) and ensuring that the computational time was reasonable, a target number of 5000 streamlines was selected. The option to select a maximum number of seeds was also disabled so that the full 5000 streamlines were generated for every tractogram. This provides some level of consistency between the data sets. However, given that both the computational time required to generate a tract (a clinically relevant factor) and the density of streamlines within that tract for a given level of artificial noise or motion vary, this may not be optimal.

The same streamline termination criteria were used for each tractography algorithm. For iFOD2 and SD_STREAM, the cutoff value for the FOD amplitude was set to 0.1. For Tensor_Prob, the tensor FA cutoff was 0.1.

An option to track from the seed points in one direction only was implemented for generating tractograms seeded from a small region of the brain stem. The default is to track in both directions to reach the diffusive limits of the tract or brain. However, the unidirectional tracking option enabled comparison of tracts on a “point-by-point” basis. With both tracts having an initial random spherical distribution of seed points, the distribution of progressive points of all the streamlines may be compared (see section 3.10 and section 4.8).

The option to seed from a single spherical point was utilised to generate “singular” tracts from each scan starting in approximately the same position at the base of the brain stem. The spherical point was assigned a radius of 1 pixel.

There are various other options which allow other parameters of the algorithm to be tweaked. For algorithms that require processing of the dicom image prior to tractography, a range of options is also available at each of the processing steps. All these parameters are likely to have some contribution to the resulting tractogram. However, it is not feasible to explore all possible variations in this study, so modifications were limited to options relevant for the analysis of noise and patient motion only.

1.4 Applications and clinical relevance

Most standard imaging modalities obtain scalar measurements for each voxel. Therefore, the amount of structural information that can be gained via these methods is limited by spatial resolution (i.e. the size of the voxels comprising the image) (Minati and Weglarz, 2007). dMRI, on the other hand, allows more information to be extracted for individual voxels, by using the diffusion of water as a means of probing microstructural characteristics of tissue. For this reason, dMRI is now recognised as a useful tool for clinical imaging of the brain. However, image quality issues in dMRI exacerbated the technique’s high sensitivity to physiological motion. Given that dMRI is sensitive to the diffusion of water on the scale of 5-15 μm assuming typical acquisition times, a small degree of subject movement, including involuntary movements, can result in substantial signal phase shifts or signal loss (Tournier et al., 2011). It is therefore important to consider means of mediating the amount of noise or patient movement during a scan, or accounting for the changes in post-processing.

To reduce the sensitivity to motion, echo-planar imaging is frequently employed, though this technique is susceptible to artifacts and other limitations such as B_0 susceptibility effects, eddy current induced distortions, and blurring and limited spatial resolution from the prolonged T_2^* signal decay during the echo train (Tournier et al., 2011). To reduce these artifacts, methods such as parallel imaging can be applied. Parallel imaging is now standard on most MRI scanners, and enables both the echo train length and echo spacing to be reduced while maintaining improved robustness to patient motion. The parallel imaging factor, however, is limited to 2-4 in most practical applications depending on the available number of receiver channels and the coil geometry.

dMRI has proven to be particularly useful in some cases where conventional MR sequences show no significant change. In particular, dMRI provides the earliest and most sensitive marker of acute brain ischemia (stroke) and has therefore enabled faster intervention by visualising abnormalities that cannot be resolved by other imaging modalities until later stages (Gonzalez, 2012, Minati and Weglarz, 2007, Moseley et al., 1990). During acute brain ischemia, water diffusion has been shown to decline by between 30% and 50% in infarcted areas where cytotoxic oedema is occurring. In DWI these regions appear as bright spots. When this was discovered, an intravenous recombinant tissue plasminogen activator (rt-PT) drug was simultaneously being developed for thrombolytic therapy for acute stroke patients. The combination of these diagnostic and therapeutic developments has revolutionized the management of this condition, allowing patients to receive thrombolytic treatment at a stage where brain tissue can still be salvaged. Although the use of DWI has vastly improved prediction of stroke patient outcomes (Leiva-Salinas and Wintermark, 2010), the reason why diffusion declines significantly in acute ischemia and the exact mechanisms through which cell swelling leads to reduced diffusion have not yet been elucidated.

dMRI is also effective for the study of brain activity. Currently, positron emission tomography (PET) and functional MRI (fMRI), which is based on the Blood Oxygen Level Dependant or BOLD contrast mechanism, are employed for this purpose (Le Bihan, 2014). Both these techniques depend on the concept that neuron activation is coupled to blood flow through metabolism, meaning that brain activity can be visualised indirectly through local variations in blood flow. However, molecular diffusion is also modulated by brain activity and may be measured directly using diffusion fMRI. The diffusion signal response possesses a sharp peak which has a more rapid onset than the indirect hemodynamic response. Additionally, the diffusion response continues even after neurovascular coupling is inhibited. Consequently, diffusion fMRI is considered to be more directly associated with neuronal function than PET or conventional fMRI. During neuronal activation, diffusion slows, which could indicate the swelling of cells within an activated cortical ribbon. This suggests that neural cells can be modelled as piezoelectric sensors, with variations in cell shape inducing cell depolarisation which activates a rapid non-synaptic transmission mechanism within neural clusters of the cortical ribbon.

Further to this, tractography can also be used for neurosurgery guidance or for quantifying white matter damage from various other pathologies.

In order to assess patients who have suffered acute brain ischemia, or study the activity of the brain, it is pivotal to first understand what other factors can affect the signal in a dMRI and how these can be accounted for or measured post-scan. The presence of additional noise or subject motion are two such important factors for consideration.

For example, the use of single-shot EPI readout is prone to distortions as a result of eddy currents that are formed by the rapid switching of the diffusion gradient pulses. The development of eddy-current compensation systems and improved hardware has significantly mitigated this issue (Tournier et al., 2011). The choice of appropriate pulse sequence and use of bipolar gradients can also minimise the effects of eddy currents. However, it is not yet clear what baseline level of noise a given tractography algorithm can be reasonably expected to handle.

Several studies have also investigated how cardiac pulsation results in nonlinear motion and local deformations of the brain parenchyma (Tournier et al., 2011) and consequent distortion of the measured diffusion signal. The signal losses and residual misalignments between images can result in flawed diffusion tensor estimates. For this reason, approaches of triggering the acquisition sequence in time with the cardiac cycle have been explored. While this form of motion is common to all subjects, certain patients may also experience involuntary movements on a larger scale that may need to be accounted for. For this reason it is also important to study the extent of the effects of movement on tractography outputs.

1.5 Quality control for tractography

One of the major difficulties in analysing tractography performance lies in the fact that there is no ground truth with which to compare tractograms obtained from patient data. It is likely that gross neuroanatomical features will be generally conserved across the population, including major white matter tracts, but other structures may vary.

It is crucial to assess the quality of acquired tractography data in order to make informed decisions about treating a patient or inferring information about the activity in the brain. Therefore, a key stage in the analysis of MRtrix's performance will be to demonstrate that the expected structures can be identified in the tractograms. The qualitative quality control method for clinical MR tractography using whole-brain CSD devised by McRobbie and Agzarian (McRobbie and Agzarian, 2016) is used for this purpose. Their method uses a scoring system to rate the presence of expected white matter tracts, and therefore provide an evaluation of how well MRtrix is able to reproduce anatomically realistic tracts. This process has been applied in chapters 3 and 4 to a range of MR scanners, acquisitions and with varied raw image quality to provide a qualitative means of assessing whether the MRtrix software is successful in delineating the major white matter tracts. CSD tractograms have also been validated theoretically and with phantoms. The scoring conditions used include demonstration of the left and right corticospinal tracts as two discrete structures in the anterior medulla and pons (CS); the left and right medial lemnisci as two discrete structures in the posterior medulla and pons (ML); the transverse pontine fibres separating the corticospinal tracts from the medial lemnisci (PF); decussation of the superior cerebellar peduncles in the midbrain (CP); the left and right optic tracts extending from the optic chiasm to the left and right lateral geniculate bodies (OT); the left and right optic radiations extending from the left and right lateral geniculate bodies to the left or right primary visual cortex including Meyer loops (OR); the anterior commissure (AC); the left and right superior longitudinal fasciculi (LF); the corpus callosum (CC); and bilateral subcortical U-fibres (UF). Potentially the presence of these expected tracts in the tractogram could also be distinguished quantitatively if they are located in consistent regions of the brain for different patients. However, this is outside the scope of the project.

Quantitative techniques of assessing whether a tractogram is reliable have not been explored in depth at this stage. Therefore, the focus of this research will be to further investigate quantitative methods of determining whether MRtrix tractograms are viable for clinical use. If we assume that the original patient scans exemplify reliable and anatomically accurate visualisations of the white matter of these patients, then by measuring the extent of degradation of the existing features of these tractograms which arises from the addition of noise or motion, the relative quality of the tractogram can be measured. These methods will be expounded upon in Chapter 3.

1.6 Patient data

Twelve dMRI anonymised scans of former patients were utilised in this study. The use of these scans was approved by the chair of the human research ethics committee at Flinders Medical Centre. The subjects of these scans were affected by various neurological conditions, with each white matter map exhibiting unique abnormalities. These dMRI scans were obtained from MRI scanners at the Flinders Medical Centre and SAHMRI, with the total number of dicom images ranging from 124 to 260, number of slices between 20 and 28, and with all but one having 64 directions. Further information including basic patient demographics and scan parameters is provided in 3.1.

1.7 Chapter summary

A quantitative measure of tractogram accuracy would be highly valuable for clinical tractography, as this would remove any operator dependence or subjectivity that might be associated with determining a disease state qualitatively. Various studies have contributed to the development of a reliable means of quantifying the quality of a tractogram, but significant work remains to isolate a single metric that is robust to all potential variables present in a clinical setting, the most prominent of which being noise, patient movement and variations in scan parameters.

Quantitative and qualitative metrics were devised to evaluate the sensitivity of the considered tractography algorithms to artificially added noise and varied forms of motion, intended to resemble possible clinical scenarios. The three tractography algorithms considered in this thesis are the iFOD2, SD_STREAM and Tensor_Prob algorithms offered by MRtrix. The impact of noise on the performance of these algorithms is expounded upon in Chapter 3, and the effects of motion, specifically translation or rotation between volumes of a scan, are explored in Chapter 4.

To assess the impacts of noise and motion, a simulated set of dMRI scans were generated by adding increments of artificial noise, translation and rotation to 12 patient dMRI scans.

The dicom images supplied were converted to the .mif file format utilised by MRtrix. MRtrix generates tractograms with the file extension .tck, which consist of a three-dimensional matrix, and selected scan and tractography parameters. In order to add noise or motion to the original dMRI images, the .mif files were loaded into Matlab and manipulated using the processes described in sections 3.4 and 4.1. The three MRtrix algorithms were then used to generate tractograms of the whole brain, and tractograms of an individual tract structure for analysis. Matlab was then used to load in the .tck files and examine the effects of noise and motion on the reliability of the resulting tractograms.

A qualitative scoring system devised by Mark Agzarian and Donald McRobbie (McRobbie and Agzarian, 2016), was used to characterise how noise and motion degrade the visual appearance of the tractograms generated using a given MRtrix algorithm, presented in sections 3.5 and 4.3, respectively. Metrics were used and derived to quantify the resemblance of modified tractograms to those generated from the original patient scans, to provide an objective measure of the relative quality of the tractograms (see sections 3.10 and 3.11 for the results of these metrics applied to the noise analysis, and sections 4.8 and 4.9 for those applied to the motion analysis). The effects of noise and motion addition on the lengths of reconstructed streamlines comprising each tractogram were also assessed in section 3.8 for noise, section 4.4 for translation and section 4.5 for rotation. Through these means, a threshold noise or motion level was identified for each algorithm, beyond which a tractogram should be considered unreliable. The qualitative tractogram scores were compared between patients. The corresponding metrics were also compared, to determine whether consistent trends with noise and motion were observed across patients. These assessments gave an overall indication of the viability of each method for clinical implementation.

Together, the methods applied in this study provide a way of comparing both the performance of different tractography algorithms, and the tractograms constructed by a single algorithm for different patients. Through this research, we have both furthered an understanding of the program's capabilities and limitations, and contributed to its continued development through designing a framework for testing and comparing different tractography algorithms.

CHAPTER 2

Literature Review

2.1 Overview

In this section a selection of common approaches to tractography and their advantages and limitations are expounded upon. An evolution in the sophistication of tractography algorithms is evident in this discussion, as approaches have progressed from a simple diffusion tensor model, to constrained spherical deconvolution with probabilistic streamline tracing. This provides both justification for and a description of the processes involved in the tractography approaches employed in this thesis.

2.2 The diffusion tensor model

The diffusion tensor model was the first widely adopted model for neuroimaging. Diffusion tensor imaging (DTI) uses eigenvalues (λ_1, λ_2 and λ_3) and eigenvectors (ϵ_1, ϵ_2 and ϵ_3) to characterize the diffusion properties of each voxel comprising the brain (DellAcqua and Tournier, 2018). Summing the eigenvalues in each voxel (i.e. finding the trace) produces a diffusion-weighted (DW) image. The mean diffusivity, also known as the apparent diffusion coefficient (ADC), can also be obtained by averaging the eigenvalues.

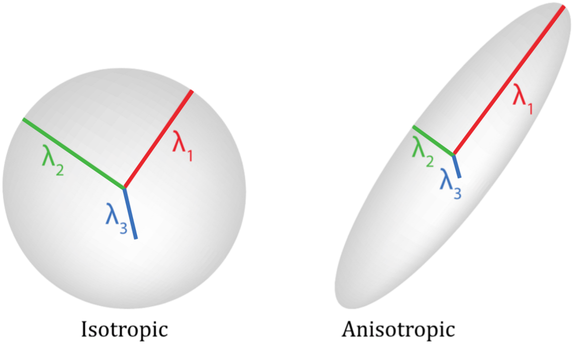


Figure 2.1: A single ellipsoid representation of the diffusion tensor demonstrating isotropic diffusion (left), where $\lambda_1 \cong \lambda_2 \cong \lambda_3$, and anisotropic diffusion (right), where $\lambda_1 \gg \lambda_2, \lambda_3$.

Since DTI is only capable of describing the mean diffusion properties of voxels, it is limited in its ability to visualise and resolve crossing fibres (Figure 2.2) (DellAcqua and Tournier, 2018). This can result in the incorrect delineation of white matter pathways which traverse regions of crossing fibres in the brain. The prominence of crossing fibres was formerly considered to be limited to a few problematic regions, such as the pons and centrum semiovale. Behrens et al. (Behrens et al., 2007) and Jeurissen et al. (Jeurissen et al., 2013) have since demonstrated that between 70% and 90% of the white matter in the human brain contains at least two crossing fibre populations. The limited capacity of DTI to characterise crossing fibres therefore poses a major problem, since crossing fibres are effectively present in all voxels of the human brain. Tractography and microstructure metrics must therefore be able to account for crossing fibres in order to provide an accurate representation of the brain's white matter.

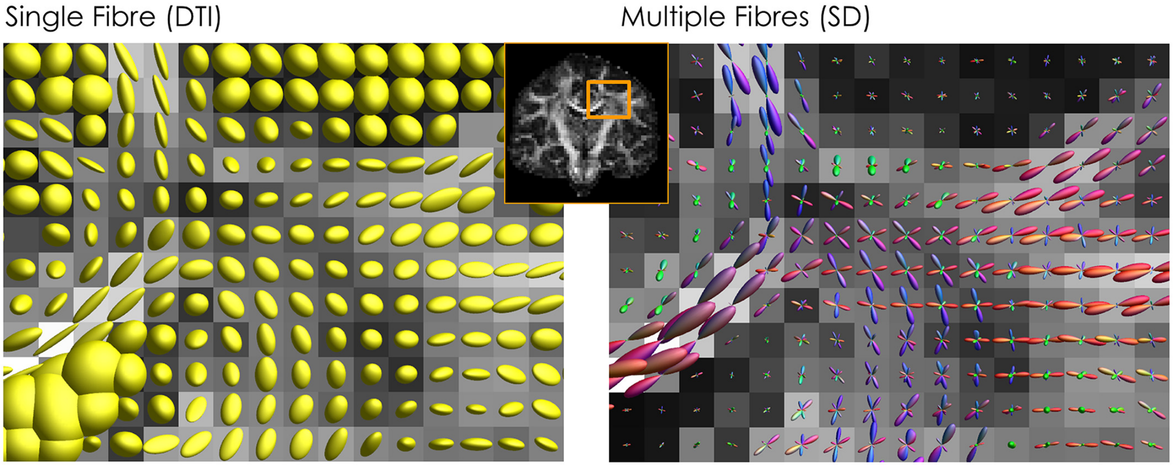


Figure 2.2: The DTI model characterises the diffusion in a given voxel based on the mean fibre orientation in the voxel (left). Spherical deconvolution (SD), on the other hand, utilises fODFs to estimate the contributions from multiple fibres to the diffusion signal. This ability to resolve multiple fibre orientations within a voxel means that SD can delineate crossing fibres (DellAcqua and Tournier, 2018).

Given that DTI cannot delineate crossing fibres, it follows that metrics derived from the diffusion tensor model cannot distinguish between fibre coherence and intrinsic white matter properties (Figure 2.2) (DellAcqua and Tournier, 2018). This has resulted in the misuse of the term 'white matter integrity' to mean fractional anisotropy (FA), as changes in FA are driven by variations in fibre coherence (Alexander et al., 2000, Jones et al., 2013, Wheeler-Kingshott and Cercignani, 2009). These limitations were first identified well before diffusion magnetic resonance imaging (dMRI) was used for tractography (Basser and Pierpaoli, 1996). Despite this, DTI remained prominent due to its unique ability to extract orientational information non-invasively from living biological tissues.

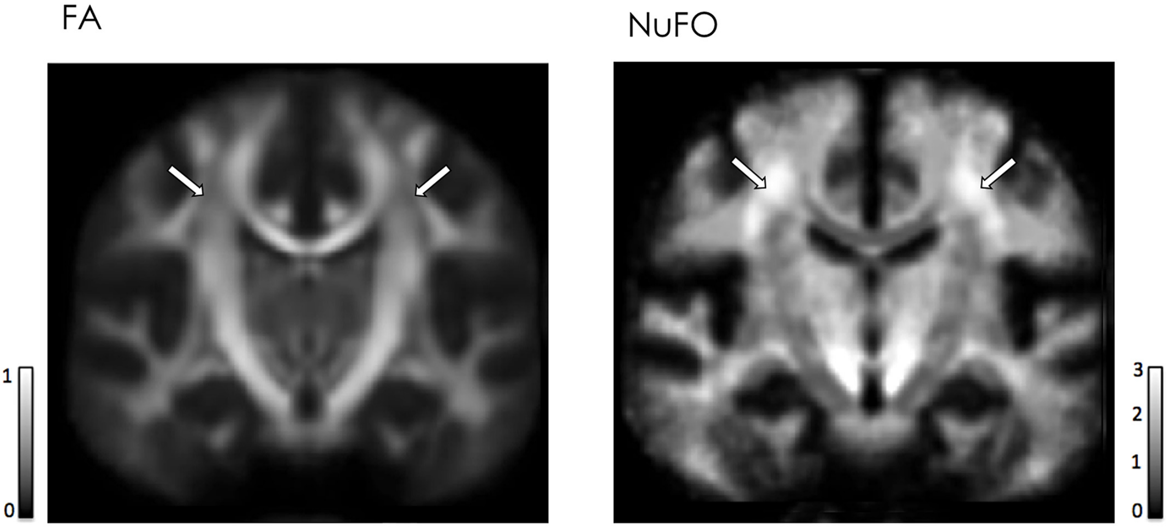


Figure 2.3: Crossing fibres can significantly affect metrics derived from diffusion data. The highlighted regions of the centrum semiovale correspond to low FA values (shown on the left), but high numbers of crossing fibres (shown on the right as a map of the number of fibre orientations (NuO)(DellAcqua and Tournier, 2018).

2.3 Alternative models

At the start of the 21st century, a number of techniques were proposed to resolve complex microstructural tissue organisation, including 6D q-space imaging to map the full diffusion propagator; multi-tensor fitting (Tuch et al., 2002); Persistent Angular Structure MRI (PAS-MRI) (Jansons and Alexander, 2003); Q-ball imaging (Tuch, 2004); spherical deconvolution (Tournier et al., 2004); ball and sticks (Behrens et al., 2007); and modern contrained variants of spherical deconvolution (Tournier et al., 2007). The last of these techniques have since been shown to perform reliably on scans representative of those that would be acquired in a clinical scenario (Tournier et al., 2011). Their success has led to a transition from the use of DTI to fibre orientation density functions as a more effective method of describing white matter complexities, with spherical deconvolution being one of the main contenders for modelling multiple fibre orientations.

A number of approaches have been used to estimate fibre orientations. These can be divided into two subsets: q-space or 'model-free' methods; and methods using mixture models or 'model-based' approaches. Q-space methods tend to estimate the diffusion orientation density function (dODF), while mixture model methods estimate the fibre orientation density function (fODF) (Tournier et al., 2011).

Fibre orientation estimation approaches can also be differentiated based on what form of output they yield. 'Parametric' approaches generally produce discrete estimates of parameters such as the orientations and volume fractions of particular fibres. 'Non-parametric' approaches provide a continuous distribution of orientation estimates such as the fibre orientation distribution (FOD). This distinction is significant since it influences which optimisation methods are required to trace out white matter tracts, which in turn determines the computational requirements for performing tractography (Tournier et al., 2011).

Despite the ability of spherical deconvolution and other modern techniques to extract more accurate fibre orientations for tractography, a strong reliance on the diffusion tensor model for measuring microstructural details such as FA, radial and axial diffusivities remains

(Assaf et al., 2017, Douaud et al., 2011). A number of the proposed techniques for estimation of fibre orientations depend on the assumption that diffusion properties are conserved across all white matter tracts, which implies that any variations that emerge in tensor-derived measures of microstructure are the result of partial volume effects, such as crossing fibres (Tournier et al., 2004). This contradicts the intended function of dMRI, which is to non-invasively probe white matter microstructure and integrity. Consequently, tensor-derived metrics are still preferred by a number of researchers, but applied with greater caution and more informed interpretations.

2.4 *Q*-space approaches

Q-space methods are based on estimating the ensemble spin propagator, $P(\mathbf{r}, \Delta)$, as a function of the diffusion time Δ and the displacement $\mathbf{r}$, or, as in the case of fibre orientation estimation, its angular dependence (DellAcqua and Tournier, 2018). The propagator can be obtained by taking the Fourier transform of the diffusion-weighted signal $S(q, \Delta)$ recorded by sampling q-space at regular intervals. This relationship can be written as:

$$P(\mathbf{r}) \propto \int S(\mathbf{q}) \exp\left(-i2\pi\mathbf{q}\cdot\mathbf{r}\right) dq \tag{2.1}$$

The q-vector $\mathbf{q} \propto \mathbf{G}\delta$, where $\mathbf{G}$ is the diffusion encoding gradient vector and δ is its duration.

The angular dependence of the propagator is characterised by the dODF, which is the radial projection of the ensemble spin propagator. Peaks in the dODF indicate the directions of the highest spin displacement.

Q-space methods are purportedly ‘model free’, since *q*-space does not assume a model. This claim is true for the propagator and hence the dODF. However, it does not apply when the dODF is used to estimate the orientations of fibres. Since the dODF is distinct, a model must be used to extract the fibre orientations from it. Generally, fibre orientations are taken to be the directions corresponding to the peaks in the dODF. However, this method is ineffective when fibres cross at non-orthogonal angles (Tournier et al., 2008, Zhan and Yang, 2006). Several alternatives have been proposed, such as a *Q*-ball sharpening transform, which reduces to a spherical deconvolution (Descoteaux et al., 2009).

Another significant downfall of the *q*-space method is that it employs a narrow pulse approximation (NPA), which is not physical. In order to meet this criterion, the diffusion gradient pulses would need to be sufficiently short as to make the diffusion of water molecules negligible compared with the axonal structures of interest. This would in turn require the use of gradient strengths which are orders of magnitude larger than what is currently achieveable, to achieve the sub-millisecond pulse duration which is necessary for the root-mean-square displacement of water molecules to be less than a typical axonal radius of 1 micrometre. However, the benefit of violating the NPA is that higher angular contrast can be achieved (Yeh et al., 2010).

Several other approaches have been derived from the *q*-space model, each with its unique advantages and limitations (DellAcqua and Tournier, 2018). Some have enabled improved reconstruction speed and robustness, improved time efficiency and angular resolution, and relaxed acquisiton requirements. Factors which generally result in these techniques not becoming widely adopted include reduced time efficiency or complicated acquisition sequences, and reduced angular resolution.

2.5 Mixture models

Conversely to q-space methods, mixture or multi-compartmental models aim to estimate fibre orientations and their volume fractions directly, by assuming a model for the signal corresponding to a single fibre population (DellAcqua and Tournier, 2018). The core assumption of mixture models is that the signal in any given voxel is given by the sum of DW signals for each individual fibre population contained in that voxel, meaning that the exchange between voxels is negligible. Mixture models vary in the following respects:

- The signal model assumed for the fibres
- The use of compartments to account for partial volume with cerebrospinal fluid (CSF) or grey matter
- How the orientational information is represented (i.e. discrete or continuous)
- The use of regularization or constraints; and
- The type of algorithm used.

2.6 Discrete or continuous representations of fibre orientations

There are two general approaches to orientation parametrization, namely discrete and continuous representations of fibre populations. Several approaches assume there are a discrete quantity of fibre populations or ‘fixels’ (i.e. a fibre population within a voxel) with well-defined orientations, which can be mapped with a small, variable number of parameters. Generally three parameters are used to describe each fibre orientation (two for the θ, ϕ orientation, one for the volume fraction) (Behrens et al., 2007, Tuch et al., 2002). The alternative approach is to represent the orientation information as a continuous fODF or FOD. This form of parameterization involves a fixed, large number of coefficients. Examples include the use of spherical harmonics (Tournier et al., 2004) , or dictionary-based approaches (Jansons and Alexander, 2003, Tuch, 2004). Some methods have also combined aspects of both discrete and continuous parametrizations, by modelling the fODF as a discrete set of fibre populations, with unique dispersions, yielding a continuous FOD.

Discrete representations of fibre orientations have the advantage of simplicity, assuming a low number of fixels per voxel and few parameters per fixel. However, the fibre orientations are not linearly related to the signal, which means that computationally expensive non-linear fitting methods are necessary to determine the parameters (Tuch et al., 2002). The number of fixels used to ascertain the fit depends on the model comparison technique used and level of detail in the data. Common methods such as an F-test , Bayes factor and automatic relevance determination (Behrens et al., 2007) favour simpler models with fewer fixels for higher noise data. If the number of fixels is guessed incorrectly, the orientations will not be accurately estimated, possibly leading to neighbouring voxels having inconsistent orientations. This is unfavourable for applications like tractography and microstructure modelling.

The advantage of a continuous representation of orientation information is that fibre dispersion, curvature or fanning can be correctly accounted for since any organisation of fibre orientations within a voxel can be characterised (Budde and Annese, 2013, Schilling et al., 2016). The spherical harmonic and dictionary-based approaches are also favourable as they are linear with signal, enabling reconstructions to be run with linear methods, which marketedly improves computational efficiency (Frank, 2002, Hess et al., 2006, Roine et al., 2014, n.d., Tournier et al., 2004, Tuch, 2004).

The spherical harmonic approach is the spherical equivalent of a Fourier series, and it can be similarly affected by the spherical equivalent of Gibbs ringing (DellAcqua and Tournier, 2018, Hess et al., 2006). Dictionary-based methods estimate the fODF as a weighted sum of direction-specific kernels, with each kernel being the angular point spread function or the signal expected for a fibre population aligned in that direction. This allows constraints like non-negativity and sparsity to be applied directly to the coefficients. This is often necessary to obtain a full reconstruction of the fODF (DellAcqua and Tournier, 2018).

The diffusion tensor model can be used to model the signal originating from each fibre population. Simplified variations of the model also exist, such as the axial symmetric tensor, which assumes the second and third eigenvalues are equal (Behrens et al., 2007, Hosey et al., 2005); and the totally anisotropic tensor or ‘stick’ model, where the second and third eigenvalues are set to zero (DellAcqua and Tournier, 2018). This is employed in multi-tensor approaches, which assume the signal to be a weighted sum of signals corresponding to individual tensors. In order to estimate relevant parameters, an independent estimation of the number of fibre populations present is required as input to non-linear fitting routines, or Markov chain Monte Carlo sampling approaches. This can be reduced to a spherical deconvolution if the axial symmetric tensor or stick model is used.

Some approaches have refined this to a microstructural scale, attempting to model different types of diffusion within fibres. An example of this is the composite hindered and restricted model of diffusion (CHARMED), which uses analytical models of restricted diffusion in impermeable cylinders to infer diffusion within axons (Assaf and P.J., 2005).

2.7 Advantages and disadvantages of the dODF and fODF

The dODF, estimated by radially projecting the q-space ensemble spin propagator, is considered superior to the fODF due to its ability to describe diffusion in an unbiased, model-independent fashion (DellAcqua and Tournier, 2018). In order to extract the fODF directly through use of a mixture model, assumptions must be made about how the magnetic resonance imaging (MRI) signal is affected by tissue microstructure.

From an opposing point of view, the dODF offers only an indirect, obscure representation of the true fibre orientation. If only one coherently orientated fibre population is present, the dODF will still produce a smooth profile, implying an array of possible orientations centred about the primary orientation, rather than a single clear fibre population (DellAcqua and Tournier, 2018). Given that the propagator is a smooth function, the angular resolution is intrinsically limited due to angular blurring introduced by the diffusion process. The dODF is thus less able to resolve multiple distinct fibre populations.

Conversely, the fODF can use models to extract more precise orientational information about constituent fibre populations. Recent studies have also demonstrated that fODF methods consistently produce sharper functions in comparison with dODF methods (Descoteaux et al., 2009, Jones et al., 2013, Tournier et al., 2011, 2008). The fODF for a single coherently oriented fibre population can be represented by a delta function oriented in the same direction. However, a smoother fODF is generally used for regularisation procedures or to resolve noise instabilities and numerical errors.

In tractography, a fibre-tacking algorithm traces out fibre paths based on estimates of the fibre orientations provided by the FOD (Conturo et al., 1999). Precise characterisation of individual fibre orientations is therefore essential, which can be best provided by estimating the fODF directly, rather than the dODF. Additionally, the sole merit of the dODF, that is, its lack of model dependence, is irrelevant for tractography, as an appropriate model must be used to extract fibre orientations from it. dODF approaches typically use the most simplistic model, using the peaks in the dODF to estimate fibre orientations, which is unreliable when fibres meet at angles less than 90 degrees (Tournier et al., 2008, Zhan and Yang, 2006). More intricate models can be used to convert the dODF into a fODF, but these are essentially the same as directly computing the fODF (Descoteaux et al., 2009).

High accuracy in the fODF is also important for microstructure models which endeavour to map crossing fibre voxels or voxels with variable dispersion at the microscale. fODF-based methods are therefore more favourable in most practical applications.

2.8 Inferring tract paths from diffusion signals

Due to the high level of noise often present in diffusion MRI data, inferring fibre orientations from a profile of water displacement in the brain is a complex and poorly conditioned inverse problem. However, more robust solutions can be obtained in the presence of noise through the use of constraints and regularisers (DellAcqua and Tournier, 2018).

Various linear constraints can be used to reduce the impact of noise on a tractogram. For example, low pass filtering may be used to attenuate high frequency noise (Tournier et al., 2004). However, this may result in loss of angular resolution.

Non-negativity and sparsity are two effective non-linear regularlisers that can be applied to a fODF. These constraints are not applicable to the dODF as they require the ODF to be positive semi-definite with voxels containing a low number of well-defined orientations. Non-negativity is always a reasonable condition to impose, since there is no plausible circumstance where the fibre density should be negative. Sparsity constraints, conversely, can create biases in regions of significant dispersion or curvature as they have the effect of reducing broad distributions into singular well-defined peaks. They can also supress small secondary fibre orientations if applied too strongly (DellAcqua and Tournier, 2018).

2.9 Spherical deconvolution

A successful fibre orientation estimation method should provide an estimate of the fODF rather than the dODF, as most applications require fibre orientations as input. This also allows more powerful constraints to be applied. The ODF should also have a continuous representation to characterise any fibre configuration. This avoids dependence on models to determine a discrete number of fibre populations in each voxel, and also allows fast, linear solvers to be used. A non-negativity constraint should also be imposed, though a unit integral contraint should not.

Spherical deconvolution is the ideal approach to satisfy these criteria. Spherical deconvolution relies on DW signals being sufficiently similar within any given fibre population that any variations can be ignored. In general, the DW signal is low along the main axis of a fibre, and preserved across the fibres. If a reasonable estimate of the DW signal for a fibre can be obtained, the problem becomes a linear sum of the signals for all fibre populations present in a given voxel. When the fibre populations are represented by an fODF, this mixture of signals becomes a spherical deconvolution. The fibre orientations can then be determined through inversion,where an appropriately calibrated response for a representative fibre population is used to infer the fODF from measured signals.

Various spherical deconvolution methods have been put forth, which vary in how they derive the response function, which contraints are chosen, how they represent the fODF and the techniques they use to solve the inverse problem.

2.10 Constrained spherical deconvolution (CSD)

Constrained spherical deconvolution is one form of spherical deconvolution which represents the fODF using the spherical harmonic basis and imposes a non-negativity contraint as a soft regulariser (Tournier et al., 2007). CSD differs from other spherical deconvolution approaches in that the solution is not necessarily non-negative, as negative values are only penalised not removed. While negative values are not physical, applying non-negativity as a soft rather than a hard constraint improves computational efficiency and prevents some loss of angular resolution. Additionally, in CSD the response function is empirically determined rather than derived from a model, and the deconvolution is performed on the raw DW signal without normalising to the b=0 signal (i.e. the signal without diffusion weighting). Operating on the raw DW signal yields more robust fibre density estimates, especially for larger b-values, as the linear relationship of the fODF to the DW signal is maintained even in regions where there are large changes in T_2, such as within the CSF.

All three tractography algorithms considered in this study utilise contrained spherical deconvolution, as this technique has been shown to outperform DTI and other alternatives in regions of crossing fibres (Tournier et al., 2008). It is therefore the approach upon which MRtrix is based.

2.11 The fibre response function

Spherical deconvolution requires an estimate of the single-fibre response function. The optimal single-fibre response function is the DW signal that would be obtained for a unit volume of white matter coherently aligned along one axis. Various methods can be used to approximate this function, such as the axial symmetric diffusion tensor model. Tensor models offer close approximations to more complex fibre response models that can characterise such complexities as restricted diffusion.

For a given b-value, the shape of the fibre response signal profile is governed by:

$$\alpha = D_{\|} - D_{\perp} \tag{2.2}$$

A fibre response based on a model can have the same effect as a regularization parameter: balancing the FOD's angular resolution and noise stability. The alternative is to employ an empirical fibre response which has no model dependence. This is advantageous for cases where the response function cannot be determined, such as for *ex vivo* studies. It also allows the inherent dispersion in axon directions within a fibre bundle to be accounted for. The inclusion of this factor produces a sharp fODF, while model-based approaches yield a more dispersed fODF (DellAcqua and Tournier, 2018).

Spherical deconvolution assumes that the response function is constant throughout the brain. Not all research supports this assumption, with evidence that axonal diameter distributions of white matter tracts vary throughout the brain, which would be expected to influence the diffusion signal (Aboitiz et al., 1992). However, a recent study utilising spherical tensor encoding has demonstrated that microstructural anisotropy is extraordinarily uniform throughout the brain (Lampinen et al., 2016). Various studies have also demonstrated that the underlying assumptions for this to be true are consistent with observation (DellAcqua et al., 2013, Tournier et al., 2004). The use of a single representative fibre response function is therefore widely considered to be justified. The advantage of this assumption is that it ensures the problem is linear and can be readily solved. Additionally, variations in axial and radial diffusivity only have significant effects on the scale or amplitude of the response, only marginally impacting its shape. Therefore, while large variations in diffusivity may result in the fibre density being poorly estimated in a particular region, the orientation estimates for the fibres will remain accurate.

MRtrix derives the fibre response function directly from image data, generally by averaging the diffusion signal from a set of empirically-determined ‘single-fibre (SF)’ voxels (Tournier et al., 2019). MRtrix initially supplied a command to automatically determine the response function. However, there were a number of documented edge cases for which this algorithm produced an undesired result Consequently, the response function determination is now performed using a script with a variety of available mechanisms for response function estimation. In general, the response function is derived by selecting voxels from the image which represent the ‘single-fibre mask’ (i.e. a set of voxels where the diffusion signal is expected to be coherently oriented in one direction). The directions of the underlying voxels in these fibres are then estimated. The signal measured in each single-fibre voxel is rotated such that the estimated fibre direction coincides with the z-axis. The signals are then combined to produce a single response function.

2.12 Deterministic versus probabilistic tractography

Tractography algorithms commonly depend on line propagation as a means of delineating white matter pathways. This type of tractography is termed deterministic streamline fibre tractography (DellAcqua and Tournier, 2018). Deterministic tractography initially determines a seed point from which to initiate the algorithm. From this point the algorithm uses the diffusion information to track the estimated orientation of the fibre then terminates once its termination criteria are satisfied. MRtrix’s SD_STREAM algorithm performs deterministic tractography using trilinear interpolation to trace the path of each streamline. This method uses the weighted sum of a given measurement from the 8 voxels nearest to the point of interest to determine the orientation of the next step in the path.

iFOD2 and Tensor_Prob also utilise trilinear interpolation to trace out streamline paths. However, these algorithms are probabilistic rather than deterministic. The inherent noise in diffusion-weighted magnetic resonance imaging (DWI) scans produces uncertainties in the estimated fibre orientations, which can cause the delineated pathway to veer from true, since a small error can result in the algorithm proceeding along a different white matter path. Deterministic tractography produces only a single estimate for each white matter path from each seed point, without any quantification of the error in these estimates (Tournier et al., 2011). Conversely, probabilistic tractography introduces a means of accounting for the errors in diffusion weighted measurements, by delivering outputs with a probability distribution.

Given that both probabilistic and deterministic algorithms may use the same underlying model to perform tractography, one method cannot be said to be more accurate than the other. Additionally, the computed probabilities produced by probabilistic algorithms provide no insight into the connectivity of the corresponding white matter structures, rather they indicate the degree of likelihood that a given connection exists. However, the advantage of probabilistic approaches is that they can provide an estimated precision for the results of tractography (Jones, 2008). Furthermore, they are not restricted to reconstructing a single tract per seed point. Instead, probabilistic approaches assess connections between the seed point and any other point in the brain, which could only be achieved by deterministic tractography through repeated tracking from the same seed point.

2.13 Effects of noise and motion on tractograms

A limited number of studies have been performed on analysing the effects of noise or patient motion on tractography. Anderson (2001) published a theoretical framework for assessing the effects of noise on estimated diffusion tensor eigenvalues and eigenvectors at mid to high signal to noise ratio (SNR). Through this framework Anderson determined that a change in noise resulted in a shift in eigenvalues independent of sorting bias. Averaging the results of many independent measurements reduced this bias. The presence of noise in a given eigenvector produced a random walk of trajectories, indicating that there is a fundamental limit to the accuracy of tractography.

Huang et al. (2004) analysed the effects of noise and the size and position of seed regions of interest (ROIs) on a high resolution, high SNR DTI scans of a fixed mouse brain. A similar approach as taken in this study was used to investigate the effects of noise on a DTI image, where Gaussian noise was artificially added and the tractography reapplied. However, in this case three groups of Gaussian noise, generated using the Box-Muller method, were added to the time domain data (real and imaginary channels of the k-space raw data) rather than a single source of Gaussian noise being added to the axial slices of the dMRI as done here. Huang et al. (2004) employed fibre assignment by continuous tracking (FACT) tractography in this study, where a given streamline is traced in both directions. To evaluate the tractography performance, they used a one-region of interest (ROI) and two-ROI and brute-force (BF) approach to select for tracts in the anterior commissure (AC). Using the one-ROI approach, 17 tracts were obtained, irrelevant of the SNR. For the two-ROI plus BF approach, the number of tracts was strongly correlated with SNR. Few streamlines remained within the AC using the one-ROI approach, while all streamlines remained within the AC using the two-ROI and BF approach regardless of the SNR. Huang et al. (2004) concluded that the BF and multi-ROI approach provides reliable and reproducible results irrelevant of SNR, but should be used with caution for quantitative assessments as the number of reconstructed streamlines is strongly affected by SNR.

A broader study of the impacts of experimental design and noise on tractography of a fixed mouse brain was performed by Hutchinson et al. (2017). DTI, diffusion kurtosis imaging (DKI), mean apparent propagator MRI (MAP-MRI), and neurite orientation dispersion and density imaging (NODDI) were the four dMRI models assessed in this study. Dependence on a variety of experimental factors, including noise, was observed for each of these approaches. Hutchinson et al. (2017) noted a correlation between the model complexity and sensitivity to DWI sampling and noise, particularly for metrics representing non-Gaussian data.

A study conducted by McClymont et al. (2017) looked at the effects of SNR, diffusion-weighted directions and image resolution on tractography. The application considered in this study, however, was cardiac imaging, where scan time is a major constraint and SNR therefore limited. The results of this research emphasise the importance of maximising SNR, with the optimal accuracy and precision in the mean diffusivity being achieved when SNR is maximised at the expense of a reduced number of diffusion-weighted directions.

A review by Havsteen et al. (2017) examined the impacts of patient movement on MRI, with a particular emphasis on applications which often contend with low SNR such as functional MRI (fMRI) and dMRI. Physiological sources of motion that may contribute significant signal change include respiration, flow and pulse related to cardiac cycles, swallowing, and small spontaneous head movements. In relation to DTI, the main compromise is between resolution and acquisition time. The long acquisition (on the order of minutes) and repetition times for DTI increase sensitivity to motion artifacts. Patient motion in DTI can produce noise in the image and cause structures to become misaligned, ultimately compromising the reliability of the scan. Havsteen et al. (2017) present a range of approaches that can be taken, depending on the magnetic resonance (MR) application, to mitigate the impacts of patient movement on a scan, including the minimisation of acquisition times, shielding to suppress MR signal from moving tissues outside the targeted structure, propeller-shaped data acquisition patterns to enable motion correction, respiratory gating, triggering, compensation or phase re-ordering, navigator echoes, and external motion tracking strategies. External tracking is purportedly the best strategy, but still faces challenges like real-time data integration.

More focused reviews of motion correction for MRI images were conducted by Godenschweger et al. (2016) and Zaitsev et al. (2015). Patient movement compromises clinical diagnostic MRI images to the extent that they become non-diagnostic in up to 30% of patients, necessitating repeated examinations and imaging costs. To date, there is no single method of motion correction that can be applied to all MR applications (Godenschweger et al., 2016, Zaitsev et al., 2015). Single shot navigation is one trialled method of correcting for motion in DTI, where information about patient movement is inferred by obtaining additional data (navigator echoes) at the start or end of the repetition loop. This approach can account for translation and rotation along the readout direction, but reduces efficiency by prolonging the repetition time. It can be extended to measure motion along the phase-encoding direction by the use of orthogonal navigational echoes in multi-shot diffusion-weighted echo planar sequences. A range of other extensions of this technique are also discussed.

Godenschweger et al. (2016) also discuss the use of registration methods for determining motion between volumes in DWI. This approach assumes that patient movement occurs solely between volume acquisitions. More sophisticated slicewise registration approaches for correcting motion are discussed, though these techniques are sensitive to noise and image artifacts. A registration approach has been employed in this thesis, as motion was added artificially, meaning there is no possibility of noise or image artifacts differing between a modified and unmodified dMRI for a given patient.

Godenschweger et al. (2016) also present prospective motion correction (PMC) methods which provide a real-time solution for motion correction. PMC can effectively prevent motion artifacts due to displacements between acquisition steps and signal dropouts due to dephasing or misalignment of the encoding direction. It can also correct spin-history effects caused by through-plane motion, which would not be possible to implement retrospectively.

While research on the effects of noise and motion and methods to correct for these effects have been developing over the past decade, it is clear that the research is limited and there is a great deal of progress yet to be made. In addition, the effects of noise and motion have not been analysed in detail for the algorithms offered by MRtrix. Here, we seek to obtain a deeper insight into the effects of these factors on MRtrix tractography.

CHAPTER 3

The effect of noise on tractography performance

3.1 Scan Parameters

Twelve patient diffusion magnetic resonance imaging (dMRI) scans performed at the Flinders Medical Centre were provided for analysis. The scan parameters used for acquisition of the diffusion-weighted images are tabulated in Table 1.

Table 1. Table of patient scan parameters. From left to right: patient number corresponding to plots, age of patient, patient sex (m for male, f for female), field strength B_0 (T), b value, number of directions, number of signals averaged, harmonic order of the response function, anatomical quality control score, number of slices, width (slice thickness in mm), pixel size (mm), number of images, repetition time (ms), echo time (ms), image acceleration factor, receiver bandwidth (Hz), and number of b=0 images. No quality control score was provided for patients 10 to 12. The tabulated quality control scores were calculated based on the method described in section 3.5.

Patient	**Age**	**Sex**	B_0 **(T)**	**b**	**n**	**NSA**	**Lmax**	**QC anat**	**Slices**	**Width**	**Pixel**	**Images**	**TR**	**TE**	**iPAT**	**bw**	**Num b=0**
1	20	m	1.5	1000	64	3	8	9.5	26	5	1.8	194	4200	105	2	1500	2
2	33	m	3	1000	64	3	8	9	22	5	1.8	194					2
3	9	m	3	1000	64	4	8	8	23	5	1.8	260	3100	93	2	1395	4
4	3	m	1.5	1000	30	4	6	9	20	4	1.8	124	3300	83	2	1500	4
5	22	m	3	1000	64	4	8	9	20	5	1.8	260	3000	93	2	1395	4
6	35	f	3	3000	64	4	8	10	20	5	1.8	260	3200	114	2	1395	4
7	29	m	1.5	1000	64	3	8	7.5	28	5	1.8	194	4200	105	2	1500	2
8	58	m	1.5	1000	64	3	8	9	24	5	1.8	194	4200	105	2	1500	2
9	19	f	3	3000	64	3	8	9	20	5	1.8	260	3200	114	2	1395	4
10	70	f	3	1000	64	4	8	-	22	5	1.8	260	3000	93	2	1395	4
11	29	m	3	3000	64	4	8	-	23	5	1.8	260	3100	93	2	1395	4
12	59	f	1.5	1000	64	3	8	-	20	5	1.8	194	4200	105	2	1500	2

3.2 The signal to noise ratio (SNR)

Noise in an MRI scan is a random signal produced by the resistive components of the imaging system. Sources of noise in an MRI scan include the receiver coils, and other electronic components of the system, which contribute electrical resistance; and the patient, who possesses some finite resistance. The noise which is produced in an MRI scan is dependent on a number of factors including the strength of the magnetic field, the proximity of radio frequency (RF) coils to the anatomy being imaged, the magnetic properties of the tissue, and the pulse sequence design (i.e. choice of TR, TE, receiver bandwidth, voxel size, number of signal averages, field of view, etc.).

The magnitude of noise in a signal can be quantified by its standard deviation (SD). When the magnitude of the noise approaches the signal magnitude, the signal becomes increasingly difficult to discern. The SNR provides a means of characterising image quality by comparing the desired signal with the level of backgroung noise. SNR is typically calculated as the quotient of the mean signal intensity in a region of interest (ROI) and the SD of the background signal intensity.

$$\mathrm{SNR} = \frac{\mathrm{signal}}{\mathrm{SD}_{\mathrm{noise}}} \tag{3.1}$$

Summing the signals from N measurements with the same SD increases the signal by a factor of N, and the SD of the noise by a factor of $\sqrt{\mathrm{N}}$. Therefore, averaging the signal over N acquisitions leads to an improvement in the SNR by a factor of $\sqrt{\mathrm{N}}$. This process of summing repeat measurements, or equivalently extending the acquisiton time, is referred to as signal averaging.

3.3 Measuring intrinsic noise

Long acquisition times are often necessary when performing dMRI scans, in order to obtain data at sufficiently high spatial resolutions and with a sufficient number of diffusion-weighted (DW) directions and signal averages to improve the SNR. It has been well established that increasing the number of DW directions reduces bias in dMRI parameters. (McClymont et al., 2017). However, factors such as patient comfort and high demand for use of the scanner may place limits on the amount of time available for a scan. These trade-offs often compromise SNR in dMRI, with the total scan time being directly proportional to the number of DW directions and number of signal averages, and SNR being proportional to the square root of the number of signal averages. Additionally, increasing the b-value to achieve stronger diffusion weighting generally decreases SNR and promotes mechanical vibration artifacts, due to the application of larger diffusion gradients (Burdette et al., 2001). It is therefore crucial for any tractography algorithm to be capable of handling a base level of noise in a dMRI image.

The level of intrinsic noise in a dMRI scan can be measured using the SNR. There are various approaches that can be used to calculate the SNR in a dMRI scan. These approaches can be distinguished based on whether they utilise a single image, a pair of images or a sequence of images to ascertain the SNR (Dietrich et al., 2007).

If a single image is used, the mean in a ROI and the SD or mean with a correction factor in a second, noise dominated, region are used to calculate the SNR. This technique is not affected by system drift artifacts but unidentified signal artifacts can influence noise statistics. The single image method is applicable to any n-channel receive system with appropriate correction factors.

Alternatively, the SNR can be calculated from a difference image of two repeated acquisitions. However, this method is sensitive to system drift artifacts and is recommended in the NEMA MS 1-2008 for single-channel receive systems only (NEMA, 2008). Another NEMA recomended two-image method involves the acquisition of an image with no nuclear magnetic resonance (NMR) signal. This technique is insensitive to system drift but affected by distortions in the expected noise distribution. It can be applied to both single-channel receive systems and n-channel receive systems with the use of appropriate correction factors. A pair of images can also be derived from a single scan, and the difference method applied to obtain the SNR. This technique is effectively insensitive to system drift as the time difference in the extracted images is the acquisition sampling interval.

In this study, the following procedure was applied to compute the SNR for each input dMRI scan:

- The level of background noise was estimated by calculating the SD of the signal intensity in pixels external to the brain (i.e. areas known to contain no biological diffusion signal).
- The mean signal intensity in the brain was then computed by averaging pixels within the brain mask.
- The SNR was computed as the quotient of the mean signal intensity and the estimated background noise.

This approach is the same as that used by Huang et al. (2004) to analyse the effect of varying noise on diffusion tensor imaging (DTI). The process was then repeated using the difference of two volumes for each dMRI (i.e. the difference method). For both methods only the diffusion-weighted images were utilised for the calculation. Using the single image method, the standard deviation of all background pixels in all diffusion-weighted images was used as the noise value, and the root-mean-square (RMS) of all pixels in the brain mask of the diffusion-weighted images was used as the signal . Using the difference method, the standard deviation of the differences in the background pixels of the first two volumes was used as the noise value, and the RMS of differences in pixels within the brain mask was used as the signal.

The SNR of a data acquisition is dependent on a number of factors other than imaging noise, including the number of signal averages, diffusion directions, and geometric scan parameters such as slice thickness, matrix and field of view. These factors must be taken into consideration for the conclusions derived from this data to be used in practical applications.

3.4 Addition of White Gaussian Noise

In order to examine how each algorithm responds to different levels of noise, artificial noise was added using Matlab to each input .mif file prior to tractography, then the output tractograms were compared. How the distribution of streamlines comprising a tract and their orientations change as noise is added provides an insight into how much noise an input scan can possess without compromising the reconstructed map of the brain's white matter.

A sequence of white noise is one in which each random element is not correlated with any other element.

$$E[y[n]y[m]] = \begin{cases} 0 & n \neq m \\ \sigma_y^2 & n = m \end{cases} \tag{3.2}$$

where σ_y^2 is the variance of y.

A sequence of white noise which follows a Gaussian distribution is known as white Gaussian noise (WGN). Matlab's wgn function was used to generate WGN of powers 0 to 40 in increments of 5, which were added to each .mif file to form a set of 9 noise variations for each patient. An example of a dMRI scan with no noise, and one with a power of 40 WGN is shown in 3.1.

If a matrix of WGN has power P, this implies that is has a mean of 0 and variance of $10^{P/10}$ (i.e. $\mu = 0$, $\sigma^2 = 10^{P/10}$), based on the following equation:

$$P_{dB} = 10 \log_{10} \sigma^2 \tag{3.3}$$

The 9 levels of added WGN are therefore evenly spaced on a logarithmic scale.

Tractograms and tract-weighted maps were generated for each input file with artifically added noise using the iFOD2, SD_STREAM and Tensor_Prob tractography algorithms.

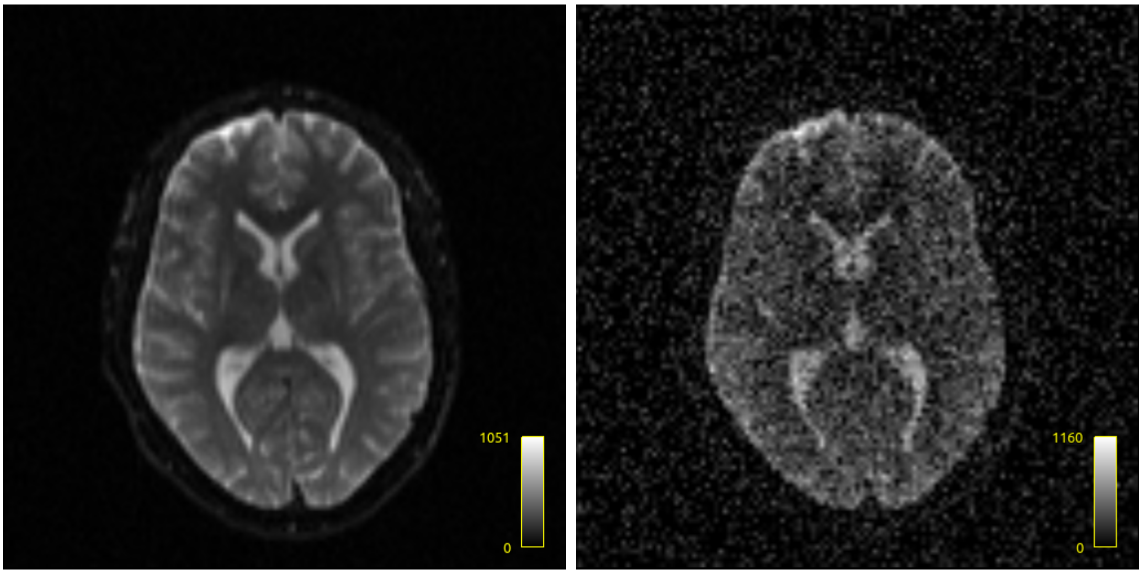


Figure 3.1: To illustrate the effect of adding WGN to a slice of a dMRI scan, the above images display slice 14 of the dMRI scan belonging to patient 1 prior to noise being added (left) and after the addition of WGN of power 40 (right).

3.4.1 Caveat

The signal intensity in a magnetic resonance image is known to be governed by a Rician distribution (Gudbjartsson and Patz, 1995). The noise in MRI images is typically assumed to follow a Gaussian distribution, which allows the noise power to be estimated by taking the standard deviation of the pixel intensity in the background signal. However, this can result in underestimation of the true noise power by up to 60% (Gudbjartsson and Patz, 1995).

To produce a magnitude image, the magnetic resonance imaging (MRI) signal is measured through a quadrature detector that yields a real and imaginary signal. Assuming that the noise in each signal is Gaussian distributed with a mean of zero, and each channel contains white noise, a Fourier transform can be used to extract the real and imaginary images from the acquired signal. Since this is a linear and orthogonal transform, the Gaussian properties of the noise are preserved. The noise in the real and imaginary voxels is uncorrelated, and the noise variance is uniform over the field of view (FOV). A magnitude image can then be generated by using the real and imaginary images to compute the magnitude in each pixel. This process is nonlinear and therefore does not preserve the Gaussian form of the noise distribution.

Let the image pixel intensity in the absence of noise be A, and M be the measured pixel intensity. The probability distribution for M, known as the Rice density, is then given by:

$$p_M(M) = \frac{M}{\sigma^2} e^{-(M^2+A^2)/2\sigma^2} I_0(\frac{A \cdot M}{\sigma^2}) \tag{3.4}$$

Here, I_0 is the zeroth order Bessel function of the first kind and σ is the standard deviation of the Gaussian noise in the real and imaginary images.

For larger SNRs ($A/\sigma \geq 3$) the Rician distribution begins to approximate a Gaussian distribution. However, the mean of the distribution is biased due to the nonlinear transform of the noisy data.

In background regions where only noise contributes to the signal (i.e. $A = 0$) this simplifies to the Rayleigh distribution.

$$p_M(M) = \frac{M}{\sigma^2} e^{-M^2/2\sigma^2} \tag{3.5}$$

To produce a phase image, the real and imaginary images are reconstructed by computing the arctangent of their ratio in each pixel. Since this function is nonlinear, the noise distribution is non-Gaussian and given by:

$$p_{\Delta\theta}(\Delta\theta) = \frac{1}{2\pi} e^{-A^2/2\sigma^2} [1 + \frac{A}{\sigma} \sqrt{2\pi} \cos\Delta\theta e^{A^2 \cos^2\Delta\theta/2\sigma^2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\inf}^{\frac{\Delta\cos\Delta\theta}{\sigma}} e^{-x^2/2} dx] \tag{3.6}$$

In regions where only noise is present (i.e. $A = 0$), the function reduces to:

$$p_{\Delta\theta}(\Delta\theta) = \begin{cases} \frac{1}{2\pi} & if - \pi < \Delta\theta < \pi \\ 0 & otherwise \end{cases} \tag{3.7}$$

This reflects the fact that the noise data points in all directions with equal probability.

For large SNR, where $A >> \sigma$, the noise distribution in the phase angle approximates a zero mean Gaussian distribution. In the case of phase images, the Gaussian distribution is a reasonable approximation even at relatively small SNR. This provides justification for the use of a white Gaussian noise function to generate simulated noise for the purpose of testing the performance of the tractography algorithms.

3.5 Qualitative results

Due to the lack of a gold standard with which to compare clinical tractograms, the evaluation of tractograms is performed primarily through qualitative means. These methods rely on the subjective evaluation of known structures in a tractogram by trained professionals, and are therefore prone to human error.

McRobbie and Agzerian (McRobbie and Agzarian, 2016) devised a qualitative scoring system to evaluate the presence of major tracts in a tractogram. A set list of structures are scored using a weighted point system to obtain a score out of 10 which describes the overall tractogram quality. A list of these structures, shown in 3.2 is provided below.

- Left and right corticospinal tracts (CS) as two discrete structures in the anterior medulla and pons.
- Left and right medial lemnisci (ML) as two discrete structures in the posterior medulla and pons.
- Transverse medial lemnisci (PF) separating the corticospinal tracts from the medial lemnisci.

- Decussation of the superior cerebellar pendunces (CP) in the midbrain.
- Left and right optic tracts (OT) extending from the optic chiasm to the left and right lateral geniculate bodies.
- Left and right optic radiations (OR) extending from the left and right lateral geniculate bodies to the left and right primary visual cortex including Meyer loops.
- Anterior commissure (AC).
- Corpus callosum (CC).
- Subcortical U fibres (UF) bilaterally.
- Left and right superior longitudinal fasciculi (LF).

Structures that are not bilateral receive a score of 1 if they are demonstrated clearly in the tractogram. Half a point is assigned to each side of a bilateral structure that is present in the tractogram. Structures that are absent receive a score of 0.

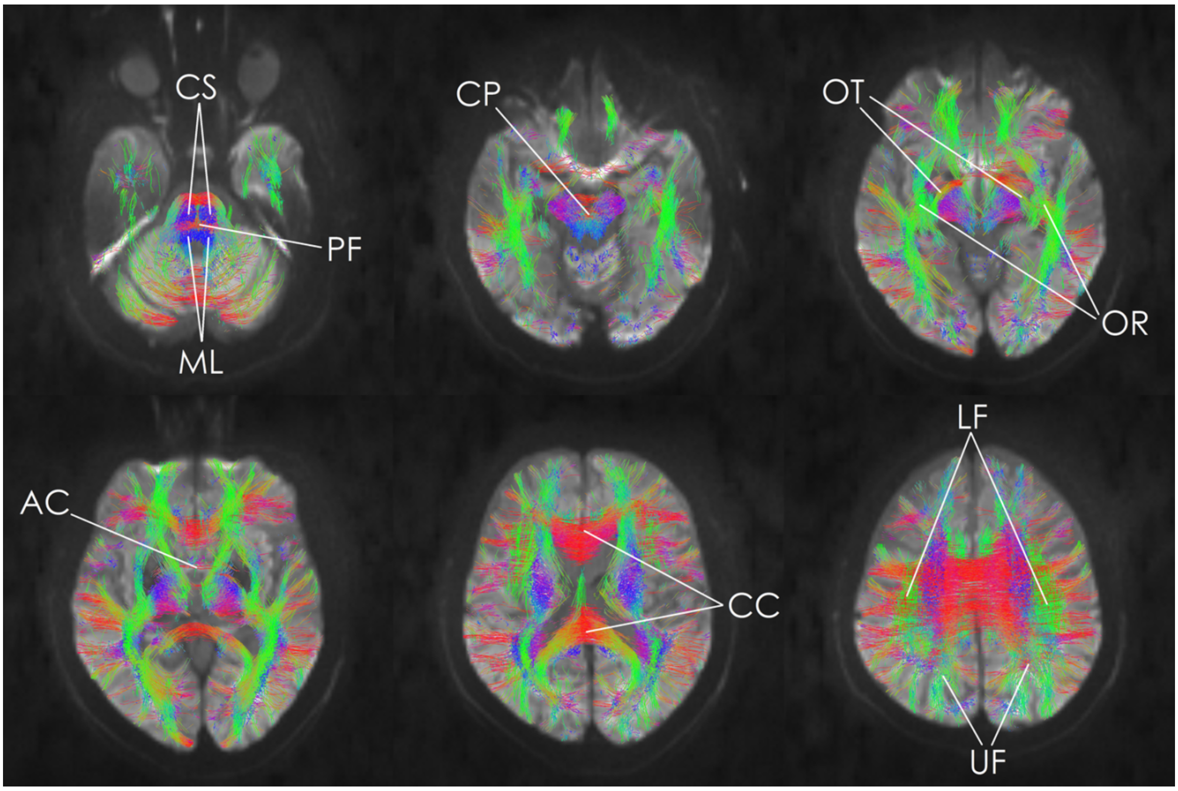


Figure 3.2: Structures scored through visual inspection of a tractogram.

A variation of this method was used to compute the qualitative score for each tractogram with varied levels of added noise. Bilateral structures were scored out of a total 2 points (1 for each individual structure), while unilateral structures were scored out of 1 point. This effectively gave bilateral structures a higher weighting than unilateral structures, with the maximum possible score for a given tractogram being 16. Individual structures received a score of 1 if they were clearly observed in the tractogram. If a structure was partially visible it received a score of 1/2. Structures that were visually absent from the tractogram were given a score of 0.

One individual performed the qualitative scoring for all patients, eliminating inter-scorer error. However, the scores are potentially biased as the scoring process is highly subjective. Scores for individual structures can be expected to vary by $\pm 1/2$.

For each tractogram generated for each patient with added WGN of powers between 0 and 40, this scoring system was applied to assess the presence of the major tracts. These scores provide a baseline measure with which alternative measures of the tractogram accuracy can be compared.

Figure 3.3 shows the change in the tractogram score (as a percentage) with added noise for each patient and each tractography algorithm. For the majority of patients analysed, the overall tractogram score starts to diminish around added WGN 25 for the probabilistic algorithms iFOD2 and Tensor_Prob, and WGN 25 for the deterministic algorithm, SD_STREAM. Patients 7 and 8 have lower initial scores as the clinical conditions of these two patients significantly compromised a number of major structures in the unmodified dMRI scans. The variations in the scores of the no-noise tractograms for these two patients between the three tractography algorithms demonstrates the algorithms' relative capacity to reproduce structures that are represented by signals of low intensity in the input dMRI scans. In particular, it appears that Tensor_Prob is the most capable at delineating structures with low diffusion signals, since about 66% and 63% of the structures were observed in the initial tractograms for patients 7 and 8 respectively, versus 53% and 67% for iFOD2, and 30% and 57% for SD_STREAM.

However, this apparent difference in initial tractogram scores for patients 7 and 8 may be a result of human error. As the tractograms were scored in sequence, with the no-noise tractogram scored first, followed by tractograms with increasing increments of added WGN, confirmation bias or the observer-expectancy effect may have influenced the tractogram scores in part, since a deterioration of these structures is expected with increasing noise. Therefore, the relative scores for each of the algorithms, and to a certain extent, each of the structures, is potentially more of a reflection of the specific features the scorer was looking for rather than the actual presence or absence of structures in the tracotgram. Confirmation bias could be eliminated if the scans were provided to the scorer in a random order, without any information about the level of added noise that was in the input dMRI.

On average, iFOD2 and Tensor_Prob consistently reproduce a larger fraction of the structures of interest than SD_STREAM at the same SNR. This is evident from the complete loss of major structures when applying SD_STREAM to dMRI scans with added WGN of power 35 and above, compared with between 5% and 40% of structures being retained for iFOD2, and between 10% and 35% for Tensor_Prob with added WGN 40.

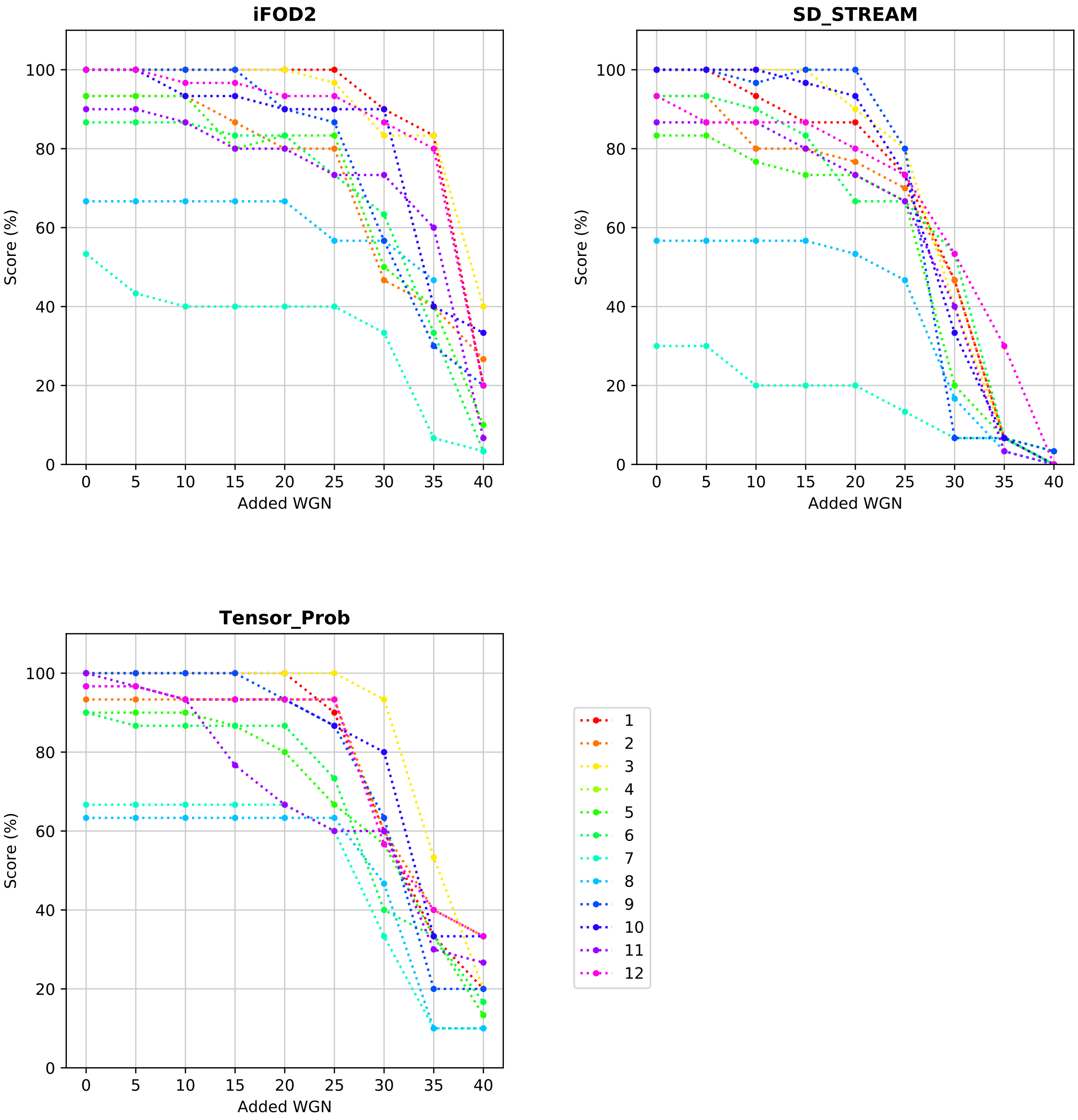


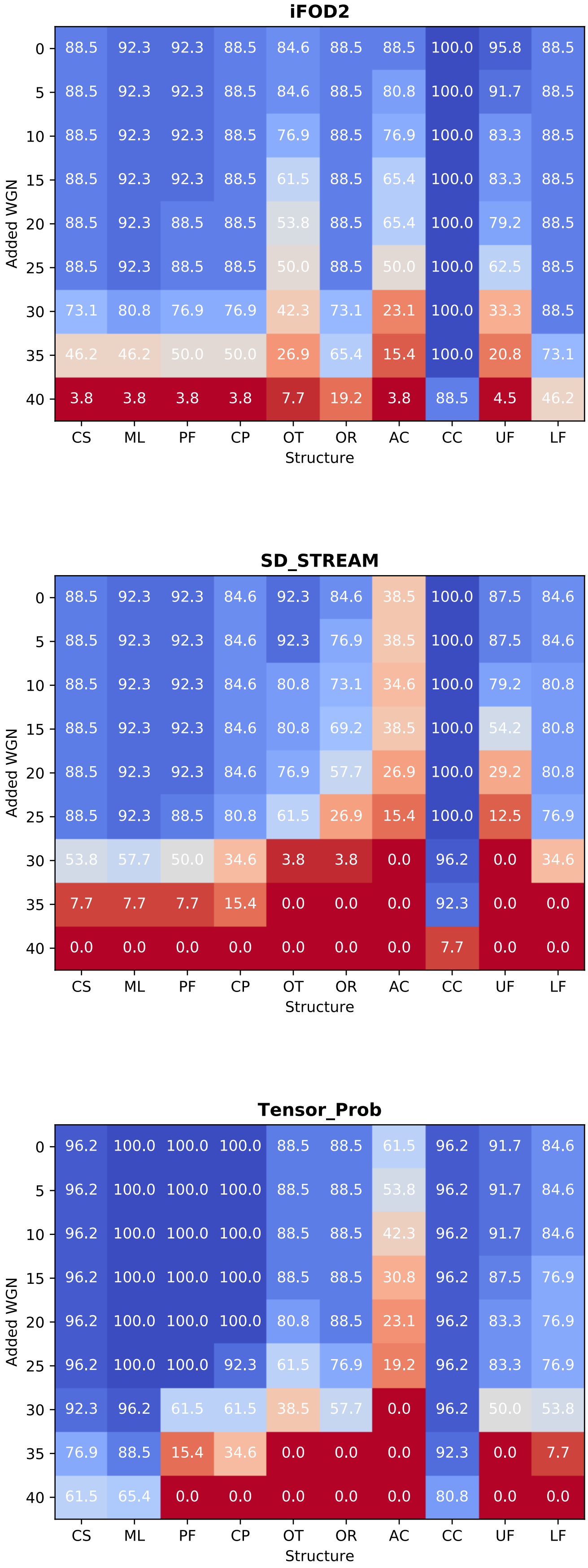
Figure 3.3: Plots showing the change in the qualitative tractogram score for each patient as WGN is added using iFOD2 (top), SD_STREAM (middle) and Tensor_Prob (bottom).

The effect of added WGN on the fidelity of the reconstruction of each structure of interest is shown for each algorithm in Figure 3.4. The first row represents the score assigned to each structure averaged over all patients for tractograms with no added WGN. For iFOD2, the CC is the most commonly observed structure, being present in 100% of no-noise tractograms for the 12 patients. This is followed by the UF, ML and PF, which are demonstrated in more than 90% of patients. The least commonly observed major structures are the OT, present in just under 85% of patient tractograms. As WGN is added, the AC, OT and are the first structures to deteriorate, respectively. These structures are depleted in about 50% of patient tractograms with 25 added WGN, and in more than 92% at 40 added WGN. This is followed by the CS, ML, PF, CP, which are depleted to a score of about 50% at added WGN 35, then the OR and LF, which have deteriorated to a score of between 65% and 75% at the same noise level. The CC is retained even at added WGN 40 for 88% of patients.

For SD_STREAM, the CC is the most commonly demonstrated major tract in the tractograms with no added WGN, followed by the OT, ML, and PF, which are observed in more than 90% of patient tractograms without added noise. The least frequently recontructed tract is the AC, which is demonstrated in less than 40% of patients prior to noise addition. The OT, OR, AC, and UF are absent in most patients at added WGN 30, with scores between 0 and 4%. All other structures, excluding the CC, deteriorate in most patients by added WGN 35. The majority of patients continue to demonstrate a CC until the level of added noise reaches a power of 40.

For Tensor_Prob, the ML, PF and CP are the most commonly demonstrated structures in the no added noise tractograms, present for all patients. This is followed by the CC and CS, which are present in about 96% of patient tractograms. The AC is demonstrated about 60% of the patients and is the first structure to deteriorate from the tractograms with added noise, followed by the OT, OR, UF, LF, PF and CP. The latter structures are present in less than 60% of patients at added WGN 30. The CC is retained for most patients at all noise levels, along with the CS and ML. With added noise of a power of 40, the percentages of patients with discernible CS and ML are 61.5% and 65.4%, respectively.

The qualitative scores of each of the 12 patient tractograms at each noise level show that the scores degrade with noise irrespective of the tractography algorithm used. However, it is evident that particular algorithms are better able to reproduce certain structures than others. This was particularly evident for the CS and ML, with Tensor_Prob reconstructing these structures with the highest fidelity of the three algorithms at all noise levels. iFOD2 produces the highest qualitative scores above WGN 25, followed closely by Tensor_Prob. SD_STREAM significantly underperforms compared to the other two tractography algorithms, with almost all structures unable to be discerned with added WGN above a power of 30 for most patients.



iFOD2

Added WGN \ Structure	CS	ML	PF	CP	OT	OR	AC	CC	UF	LF
0	88.5	92.3	92.3	88.5	84.6	88.5	88.5	100.0	95.8	88.5
5	88.5	92.3	92.3	88.5	84.6	88.5	80.8	100.0	91.7	88.5
10	88.5	92.3	92.3	88.5	76.9	88.5	76.9	100.0	83.3	88.5
15	88.5	92.3	92.3	88.5	61.5	88.5	65.4	100.0	83.3	88.5
20	88.5	92.3	88.5	88.5	53.8	88.5	65.4	100.0	79.2	88.5
25	88.5	92.3	88.5	88.5	50.0	88.5	50.0	100.0	62.5	88.5
30	73.1	80.8	76.9	76.9	42.3	73.1	23.1	100.0	33.3	88.5
35	46.2	46.2	50.0	50.0	26.9	65.4	15.4	100.0	20.8	73.1
40	3.8	3.8	3.8	3.8	7.7	19.2	3.8	88.5	4.5	46.2

SD_STREAM

Added WGN \ Structure	CS	ML	PF	CP	OT	OR	AC	CC	UF	LF
0	88.5	92.3	92.3	84.6	92.3	84.6	38.5	100.0	87.5	84.6
5	88.5	92.3	92.3	84.6	92.3	76.9	38.5	100.0	87.5	84.6
10	88.5	92.3	92.3	84.6	80.8	73.1	34.6	100.0	79.2	80.8
15	88.5	92.3	92.3	84.6	80.8	69.2	38.5	100.0	54.2	80.8
20	88.5	92.3	92.3	84.6	76.9	57.7	26.9	100.0	29.2	80.8
25	88.5	92.3	88.5	80.8	61.5	26.9	15.4	100.0	12.5	76.9
30	53.8	57.7	50.0	34.6	3.8	3.8	0.0	96.2	0.0	34.6
35	7.7	7.7	7.7	15.4	0.0	0.0	0.0	92.3	0.0	0.0
40	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.7	0.0	0.0

Tensor_Prob

Added WGN \ Structure	CS	ML	PF	CP	OT	OR	AC	CC	UF	LF
0	96.2	100.0	100.0	100.0	88.5	88.5	61.5	96.2	91.7	84.6
5	96.2	100.0	100.0	100.0	88.5	88.5	53.8	96.2	91.7	84.6
10	96.2	100.0	100.0	100.0	88.5	88.5	42.3	96.2	91.7	84.6
15	96.2	100.0	100.0	100.0	88.5	88.5	30.8	96.2	87.5	76.9
20	96.2	100.0	100.0	100.0	80.8	88.5	23.1	96.2	83.3	76.9
25	96.2	100.0	100.0	92.3	61.5	76.9	19.2	96.2	83.3	76.9
30	92.3	96.2	61.5	61.5	38.5	57.7	0.0	96.2	50.0	53.8
35	76.9	88.5	15.4	34.6	0.0	0.0	0.0	92.3	0.0	7.7
40	61.5	65.4	0.0	0.0	0.0	0.0	0.0	80.8	0.0	0.0

Figure 3.4: Plots showing the rating per structure averaged across all patients against the level of added WGN for iFOD2 (top), SD_STREAM (middle) and Tensor_Prob (bottom).Each value on the colourmap indicates the percentage of patient tractograms that exhibited the structure in that column with the specified level of added WGN in that row.

3.6 Selection of white matter structures for analysis

Initially, major tracts were selected for analysis via the following process:

- Whole brain tractography was performed using MRtrix's iFOD2 algorithm.
- Individual streamlines were selected for analysis based on whether they traversed a manually specified volume at the base of the brain stem.
- The selection was refined by retaining only those streamlines which passed through a second volume which did not intersect the first.

The chosen subset of streamlines was intended to isolate a single white matter tract, with the quantity of streamlines in the final selection providing a measure of the strength of the white matter connectivity in the tract. However, this technique was flawed in that the manual determination of the starting and terminating regions for the tract significantly influenced the number and form of the streamlines selected, and hence the inferred strength of the connection. This was due to the fact that this selection process essentially eliminates streamlines which deviate from the path defined by the initial and final volumes, yielding a sample that potentially poorly represents the 'true' shape and distribution of streamlines constituting the desired tract.

An alternative method of isolating a major tract was therefore devised. Rather than performing tractography on the whole brain and then extracting individual streamlines, tractography was performed in such a way that it yielded only one tract or structure that could be compared between patients. This was achieved by defining a spherical point at the base of the brain stem and generating a tractogram of possible streamlines emanating from that point. This effectively produced a tractogram consisting of a single tract or white matter structure with a form that could be expected to be roughly conserved across patients.

The location of the seed point for the tract was determined visually by locating the base of the brain in the original .mif image and utilising the cursor in MRtrix to obtain the position of the corresponding voxel. The voxel was then used as an input to the -seed_sphere option for each tractography algorithm, to specify a seeding volume, with the default radius of the seed volume being 1 pixel.

The change in the way the streamlines evolved from the chosen point as noise or motion were added could therefore be examined. This technique provided two key measures of the software's performance; namely the length and distribution of reconstructed streamlines. Using streamline density to quantify tractography performance was also considered. However, a constant number of streamlines was generated for each tractogram, independent of the length or complexity of the tract represented in the tractogram. Consequently, while the density in a particular region of a tractogram may demonstrate some correlation with added noise or motion, it is not necessarily informative as different paths may be traced out by chance which do not fall in the region of interest.

3.7 Added noise vs SNR

The SNR in each input dMRI was calculated using the single image method described in section 3.3. The SNR for each scan was obtained by applying this method to only the diffusion-weighted scan images. An alternative estimate of the SNR was computed by applying the difference method to the first two volumes of each scan. In figures 3.5 and 3.6, each version of the computed SNR is plotted against the power of added WGN for each patient.

Applying the single image method directly to the diffusion-weighted images results in a continuum of SNR versus added WGN curves for the twelve patients, with initial SNRs ranging from about 6 to 16. As increased powers of WGN are added, the SNRs for all patients start to converge, approaching an SNR of about 3 at added WGN 40. There appears to be two clusters of patients with initial SNRs around 6 and 11, which may indicate that the initial SNRs of these dMRI scans were similar, perhaps due to similar scan parameters or conditions.

Using the difference method on both b=0 and diffusion images, there is only evidence of two distinct subsets of patients for levels of added WGN above 20, where a small bulge occurs for all patients excluding 6, 9 and 11.

As the SNR curves appear to be more stable in form using the single image method, the values obtained with this technique are used in the analysis which follows.

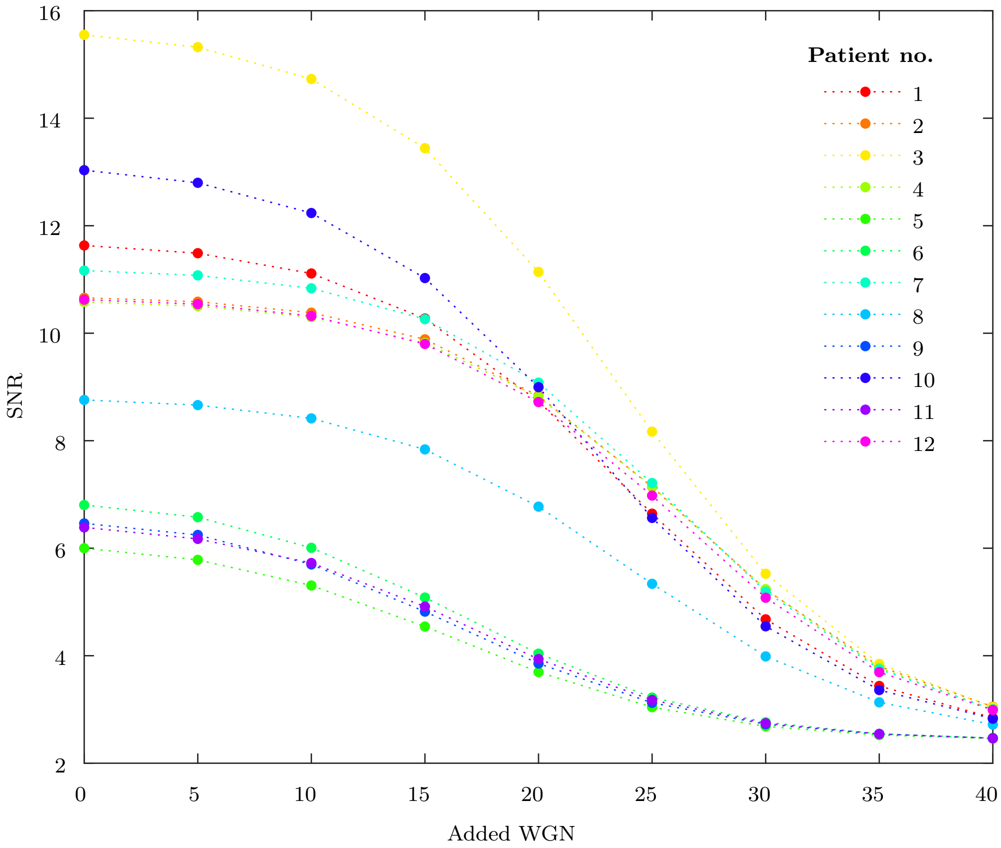


Figure 3.5: Plot of the SNR calculated using the single image method, applied to the diffusion-weighted images for each patient.

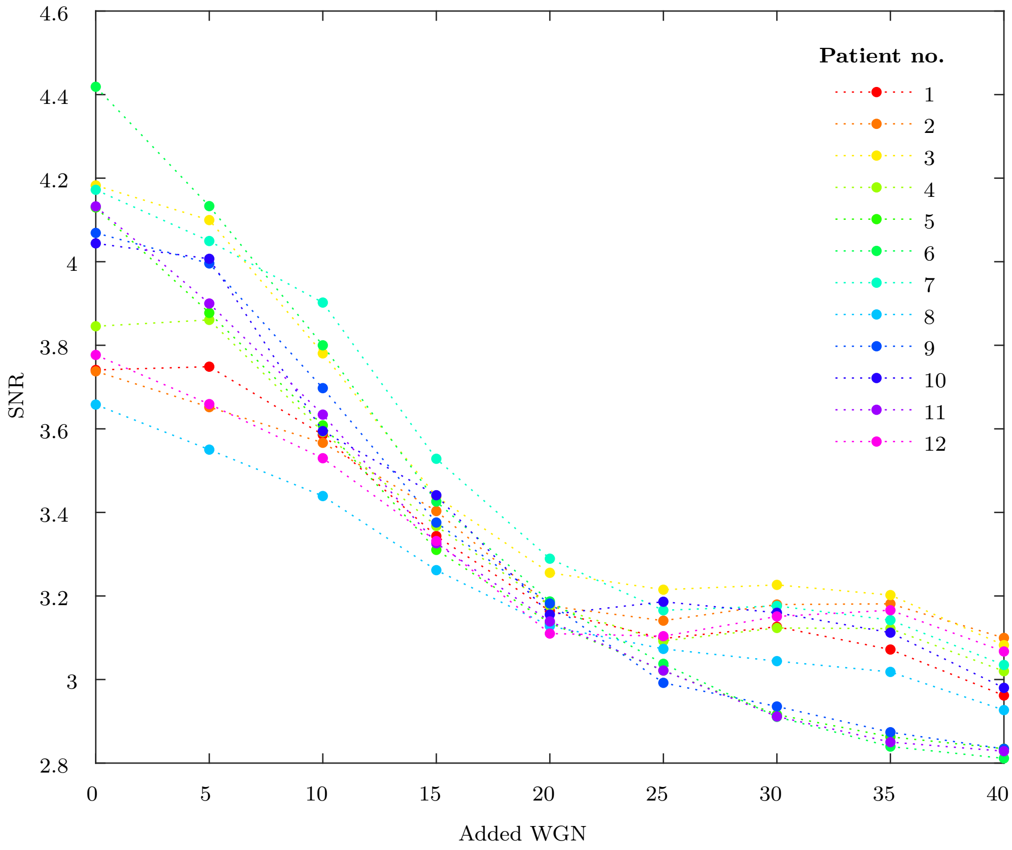


Figure 3.6: Plot of the computed SNR using the difference of two volumes method versus the amount of added WGN, using the diffusion-weighted images for each patient.

3.8 Noise versus streamline length

3.8.1 iFOD2

iFOD2 is a probabilistic tractography algorithm which models the uncertainty at the whole fibre orientation distribution (FOD) level, since dMRI is not capable of providing sufficiently high angular resolution to distinguish different levels of streamline dispersion. The presence of fibre dispersion in the brain is now well-established, meaning that fibres cannot be assumed to be coherently oriented. This justifies the use of iFOD2 over Tensor_Prob, as the tensor tracking methods implemented in the Tensor_Prob algorithm assume that fibres are coherently oriented. For probabilistic approaches, the measure of interest is the density of the streamlines, which provides an indication of the confidence in the results.

For each tract generated using iFOD2 from the dMRI scans with 9 variations of added WGN for a particular patient, a histogram was generated of the lengths of individual streamlines. For a given patient, the low-noise (i.e. added WGN < 30) spectra share the same characteristic shape, which is related to the underlying white matter structure represented in the unmodified tractogram. Figure 3.7 depicts the change in the distribution of streamline lengths with the addition of WGN for patient 1, for tractograms generated using iFOD2.

In general, the iFOD2 streamline length distributions progressively lose their characteristic peaks and become skewed to the left as the noise is increased. This reflects the tendency of iFOD2 to produce premature and/or late terminated streamlines as additional noise is added. The blurring of the peaks in the streamline length histogram with added noise likely relates to loss or alteration of the finer structures in the tractogram. This effect is only dominant at low levels of added WGN (i.e. added WGN < 30). With WGN beyond a power of 30, the effect of early streamline termination begins to dominate, where not only are finer structures lost, but more significant structures also start to deteriorate.

The change in the mean streamline length and SD with noise provides an insight into how the overall shape of the distribution is affected by increased noise in the input dMRI. However, as the histogram distributions are highly skewed and not normally distributed, changes in the mean streamline length and standard deviation are emphasised by the increase in skew of the histograms. Figure 3.8 shows the mean streamline lengths and SDs superimposed over individual streamline lengths for patient 1, versus the level of added WGN. It is evident that the maximum possible streamline length is 253 mm (100 times the voxel size of 2.53 mm). More than 95% of the streamline length distribution falls within 200 mm for all noise variations. Marginal differences in the mean and SD of the length distribution are evident for added WGN up to WGN 20. Beyond this power, a progressive shortening of the streamlines is observed, with the mean and standard deviation steadily decreasing to about half their original values at WGN 40.

One possible interpretation of this trend is that adding a small degree of uncertainty to the input dMRI (i.e. adding WGN of powers 5 through 20) will only affect signals of a similar order of magnitude. Accordingly, intricate structures are first to go, and significantly more noise is required ($>$ power 20 WGN) to impact the signal of regions of higher connectivity. The algorithm hence consistently reproduces the majority of the tract (and preserves most of the original length of each streamline comprising it) for the first few levels of added WGN.

The longest lengths in each tract are composed of a minority of the streamlines. Additionally, it appears that a certain level of noise prohibits streamlines with lengths beyond a certain threshold from being delineated. This provides evidence that increased noise leads to early termination of streamlines by iFOD2. This is due to the fact that the algorithm terminates the streamline if it does not deem the next step to be of a sufficiently high probability. In this sense it takes a cautious approach, ensuring that those tracts reproduced in the final tractogram are present in reality and are thus reliable. However, as tracts may not be fully delineated as a result of part or all of the tract being obscured by noise, motion, artefacts, or insufficient resolution in the input image, the tractograms the software yields may not provide a complete representation of the physical structure of the patient's brain. Therefore, the absence of tracts in a tractogram is not sufficient information to conclude that they do not exist in reality, while tracts that are present are likely to be reliable.

Figure 3.9 shows the normalised mean streamline length (i.e. the mean streamline length for the tractogram derived from the dMRI with added noise, divided by the mean streamline length for the whole-brain and single-tract tractograms derived from the unmodified dMRI) versus the power of added WGN for the iFOD2 tractograms for all patients. The mean streamline length remains within 5% of the original value for input dMRI scans with added WGN up to a power of about 15. For higher noise levels, a rapid decline in the mean streamline length is observed for the majority of patients. For a small subset (patients 5, 6, 9 and 11), the mean streamline length remains within 10% for all whole brain tractograms. This may be due to the original dMRI scans for these patients having lower initial SNR, such that the addition of further noise didn't significantly increase the uncertainty in streamline delineation.

Plotting the normalised mean streamline lengths versus the SNR calculated using the single image method, shows that the streamline lengths remain constant on average above an SNR of about 10. Below this, the mean streamline length progressively drops (figure 3.10). Patients 5, 6, 9 and 11 follow a separate curve than the other patients, with mean streamline length only starting to decrease below an SNR of about 5.

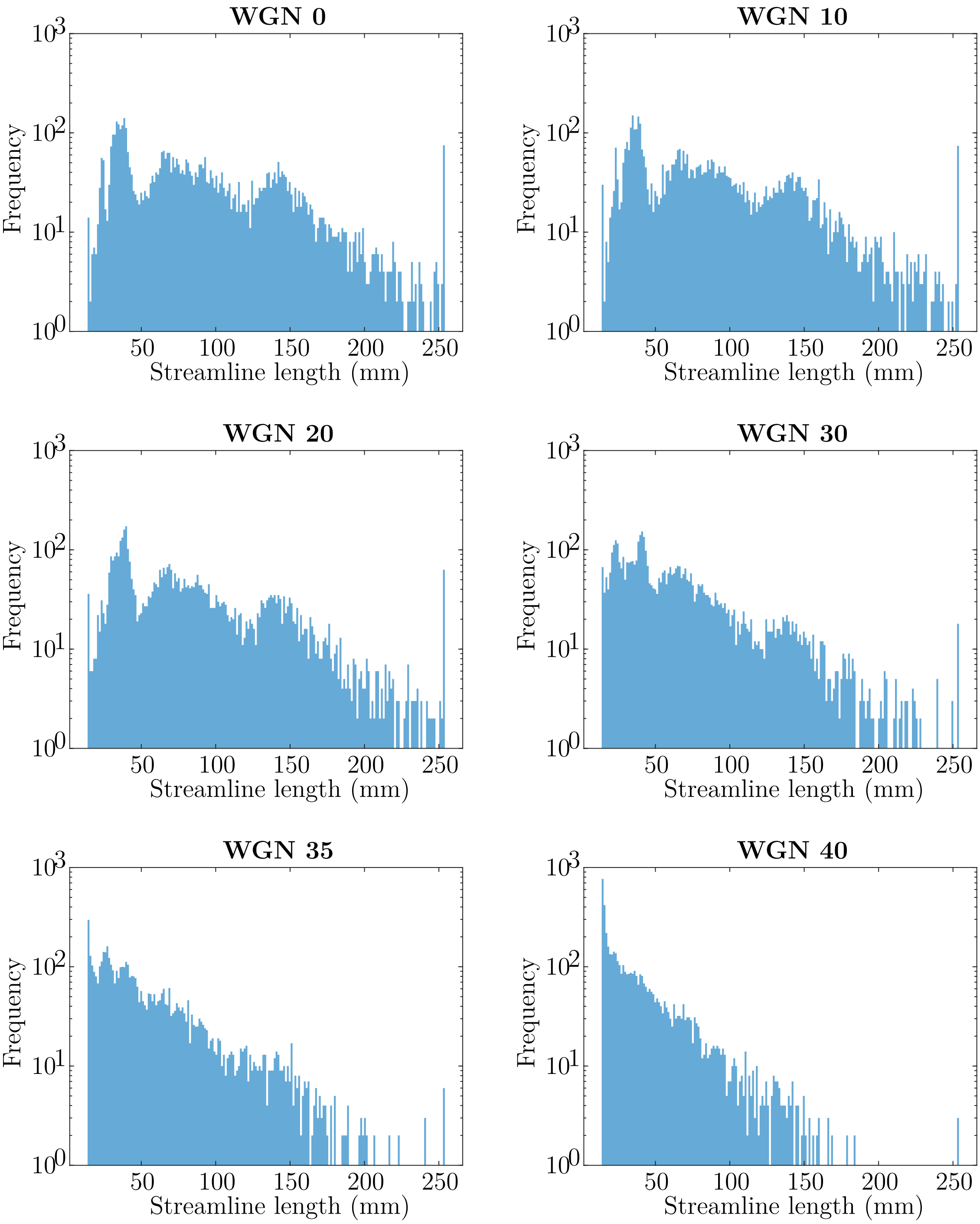


Figure 3.7: Histograms of streamline length for tractograms generated using MRtrix's iFOD2 algorithm with varied levels of added WGN. These results are derived from the dMRI scan taken of patient 1. Each tractogram consists of 5000 streamlines seeded at a point at the base of the brain stem, such that the integral of each histogram is equivalent. The histogram bin size was set to 1, equal to the step size used by the algorithm. The bins therefore represent all possible streamline lengths that the algorithm can reproduce. As additional noise is added, the histogram becomes more skewed to the left, representing the tendency of the iFOD2 algorithm to prematurely terminate streamlines when the input dMRI is noisier.

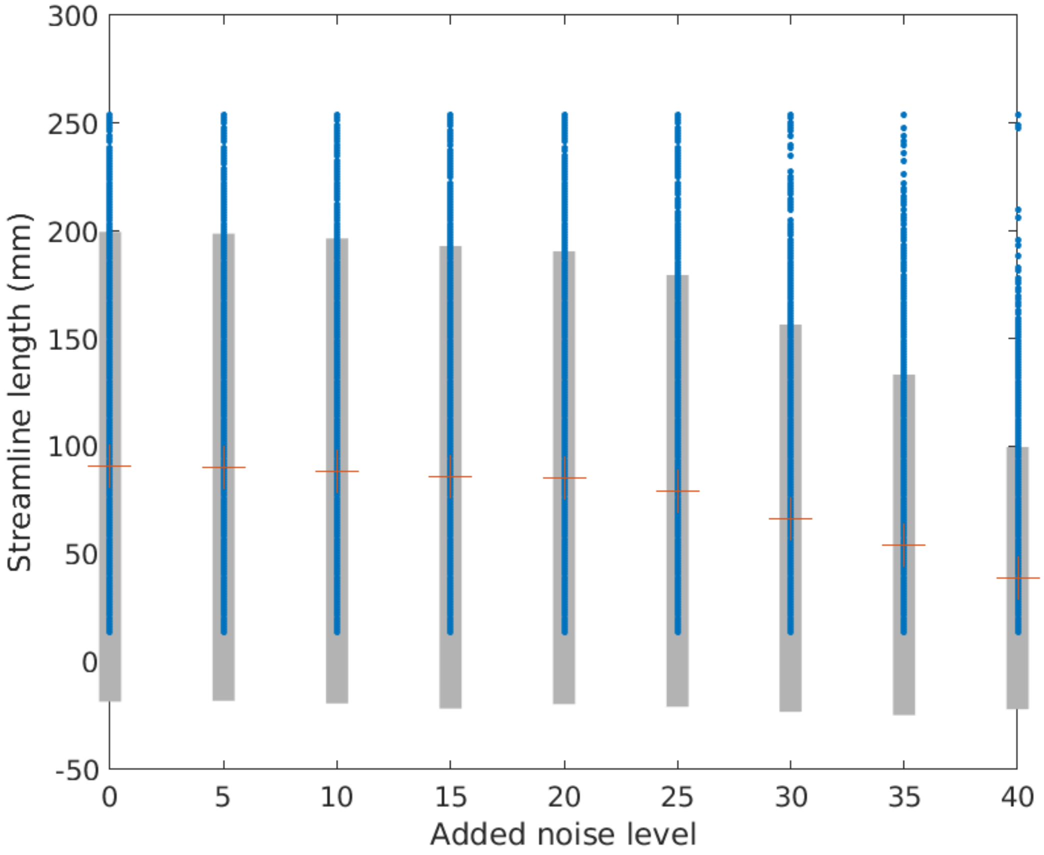


Figure 3.8: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with iFOD2 as WGN is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

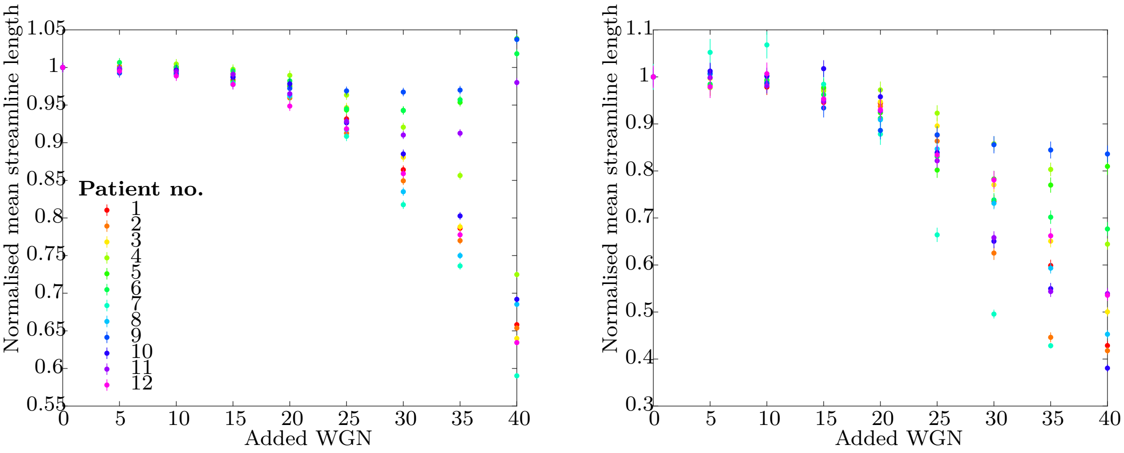


Figure 3.9: Plot of the normalised mean streamline length versus added noise for all whole-brain (left) and single tract tractograms (right) generated using iFOD2.

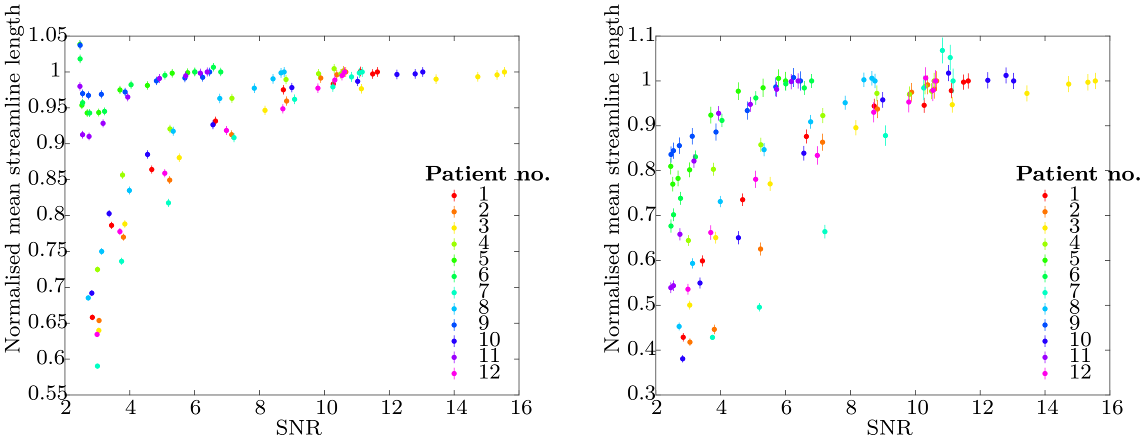


Figure 3.10: Plot of the normalised mean streamline length versus SNR (calculated using the single image method) for all whole-brain (left) and single tract tractograms (right) generated using iFOD2.

3.8.2 SD_STREAM

Since the SD_STREAM approach is deterministic, it produces cleaner-looking results than probabilistic approaches. This is evident in both the tractograms produced by the algorithm, as well as the distributions of streamline lengths derived from these tractograms.

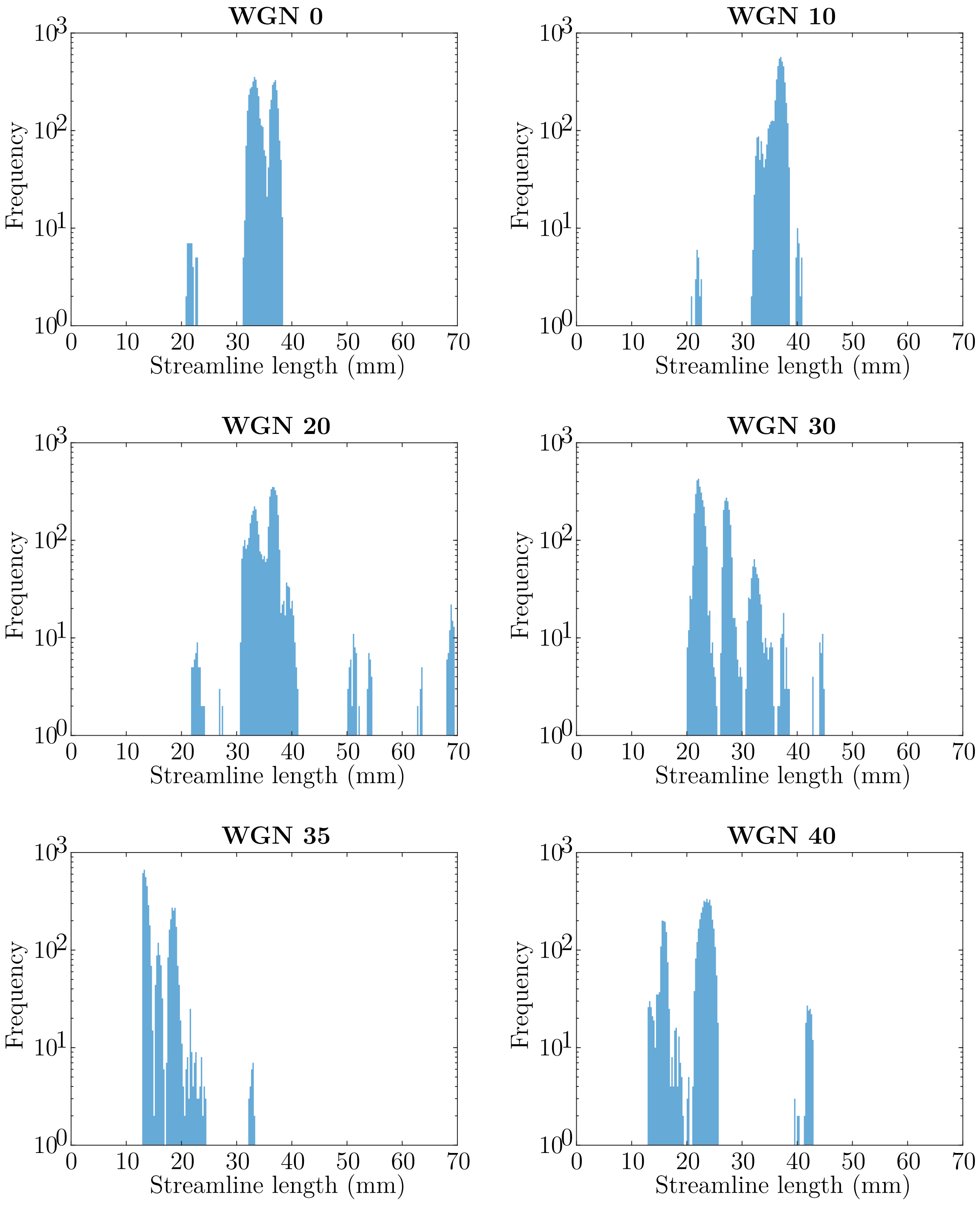
Similarly to iFOD2, SD_STREAM produces streamline length distributions that share a characteristic shape up to an added WGN power of about 20. However, the shapes of the initial distributions of streamline lengths for the unmodified dMRI scan differs significantly for the two algorithms, which reflects the fact that SD_STREAM only delineates streamlines with low uncertainty, while iFOD2 produces both high and low probability streamlines. This also explains why streamline lengths are typically significantly shorter in tractograms generated using SD_STREAM than those generated using iFOD2. In the case of patient 1, represented in Figure 3.11, the majority of streamlines for the first few tractograms (i.e. WGN 0 to 20) have lengths between 30 and 40 mm, with a small portion having lengths just above 20 mm. At WGN 20, the algorithm appears to generate some false positives, with clusters of streamlines having lengths between 50 mm and 70 mm. This behaviour demonstrates that a certain level of added noise can by chance produce a signal that the SD_STREAM algorithm interprets as an elongation of a continuous diffusion signal pathway (or a warping), which causes some streamlines to be terminated later than they would be for an input dMRI with less noise.

For higher noise levels, these false positives largely disappear, and the distribution becomes skewed to the left, as seen for iFOD2. Again, this shows the early termination of streamlines at high noise, where increased uncertainty causes the algorithm to trace out fewer steps.

An unusual feature of the SD_STREAM streamline length distributions in comparison with those from iFOD2, is the presence of well-defined peaks even at very high levels of added noise. A smoothing out of the distribution is not observed. Rather, it appears that with each step in noise, the distribution simply takes on a new form.

Figure 3.12 shows the change in the mean, SD, and range of the streamline lengths with added WGN. The average streamline length is constant up to about a power of 25 added WGN, at which point it begins to decrease. The presence of false positives is immediately evident in the WGN 20 to 25 tractograms (and to some extent in the WGN 15, 30 and 40 tractograms) from the range of individual streamline lengths, with the lengths of the longest streamlines far exceeding those in the original tractogram.

The normalised mean streamline length for the whole brain tractograms decreases progressively with noise, as shown in Figure 3.13. Similarly to iFOD2, two sets of trends emerge. Given that the same patients occur in the two sets for both algorithms, this suggests, as previously stated, that this may be an artefact of the original dMRI scans falling within two ranges of initial SNRs, or being primarily composed of shorter/longer streamlines. Plotting the normalised mean streamline lengths versus the calculated SNR gives a threshold very close to that for iFOD2 (i.e. around 10 for the majority of patients, and about 6 for the minority group), below which the mean streamline length rapidly drops off.



WGN 0

WGN 10

WGN 20

WGN 30

WGN 35

WGN 40

Figure 3.11: Histograms of streamline length for tractograms generated using MRtrix's SD_STREAM algorithm with varied levels of added WGN. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered again consists of 5000 streamlines seeded from the base of the brain stem. The SD_STREAM algorithm uses a finer step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are therefore shown on a logarithmic scale to better visualise regions of low counts in the distribution.

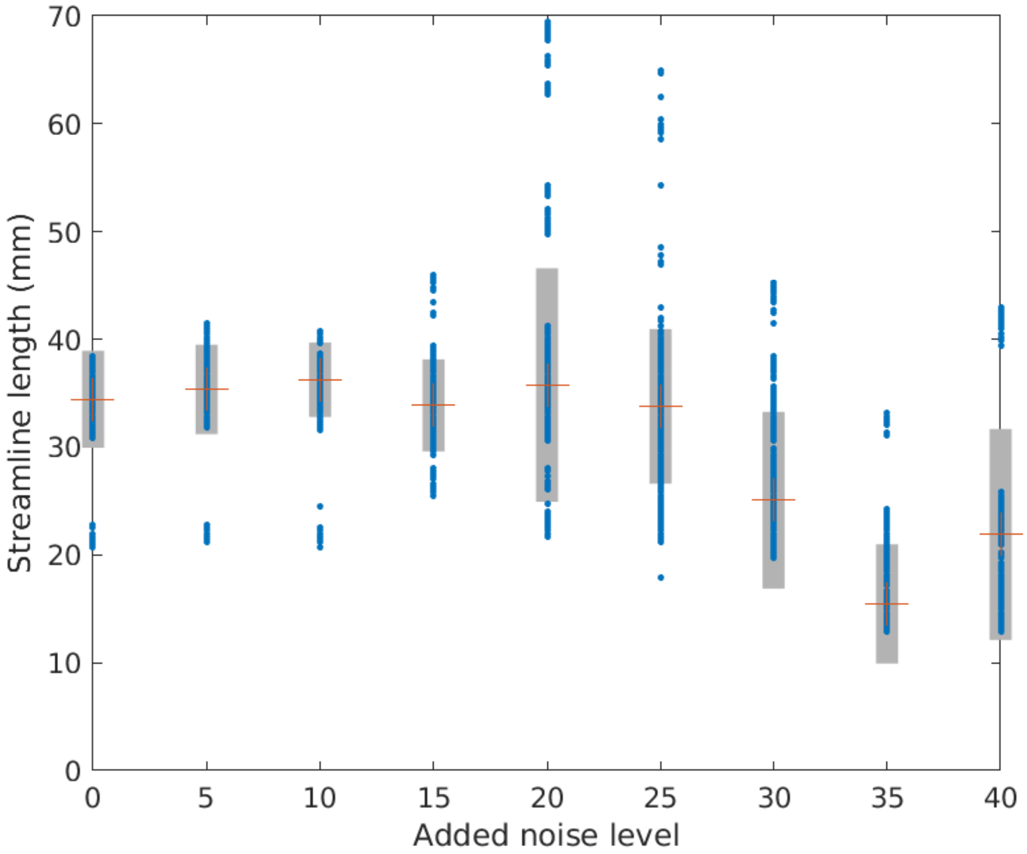


Figure 3.12: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with SD_STREAM as WGN is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

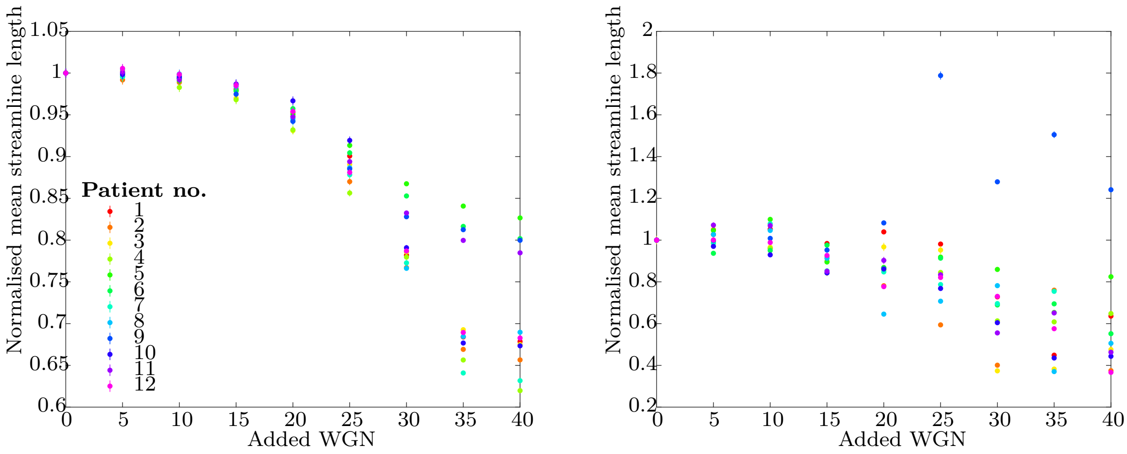


Figure 3.13: Plot of the normalised mean streamline length versus added noise for all whole-brain (left) and single tract tractograms (right) generated using SD_STREAM.

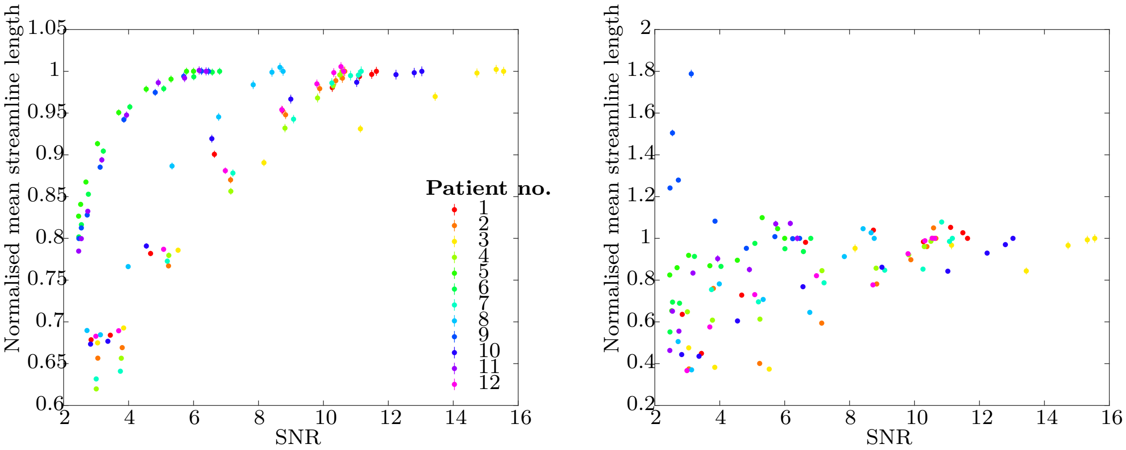


Figure 3.14: Plot of the normalised mean streamline length versus SNR (calculated using the single image method) for all whole-brain (left) and single tract tractograms (right) generated using SD_STREAM.

3.8.3 Tensor_Prob

The Tensor_Prob algorithm, similarly to the iFOD2 algorithm, is a probabilistic approach. However, this algorithm differs in that it employs a wild-bootstrap method to estimate the uncertainty in fibre orientations. Typical bootstrap methods used in tractography perform resampling using data with the same diffusion gradients, which necessitates multiple acquisitions at each diffusion gradient. The wild-bootstrap technique, by contrast, can be used even for dMRI scans with a single acquisition at each diffusion gradient. This method estimates the uncertainty in fibre orientations through model-based resampling. It is therefore only capable of modelling the imaging noise. As a result, the input data must be of very high quality (i.e. must contain minimal noise) to obtain streamlines with minimal dispersion.

The streamline length histograms are shown for patient 1 in Figure 3.15. Similar behaviour is observed as for the other tractography algorithms. The first few noise levels show a similar characteristic structure, with main peaks centred at around 30 mm and 60 mm. Smaller peaks representing streamlines with longer lengths disappear beyond this point. This reflects a loss of finer detail. The distribution is then pushed to the left due to early termination of streamlines.

Figure 3.16 shows that the difference in the mean streamline length is insignificant for tractograms with added WGN noise powers between 0 and 25. Beyond this, only a slight decrease in the mean streamline length is observed. However, the spread of values drops significantly from added WGN 25 to WGN 30. The range of lengths in each distribution also shows that streamlines are terminated prematurely with added WGN above a power of about 20.

The normalised mean streamline length falls with added WGN with a power of beyond 10 for Tensor_Prob, as depicted in Figure 3.17. Similarly to the iFOD2 and SD_STREAM algorithms, the normalised mean streamline lengths for tractograms generated using Tensor_Prob follow a separate trend for patients 5, 6, 9 and 11 compared to the other patients. Again, the mean streamline lengths remain consistent for SNRs down to about 10 for the majority of patients. Patients 5, 6, 9 and 11 show consistent mean streamline lengths down to an SNR of about 6.

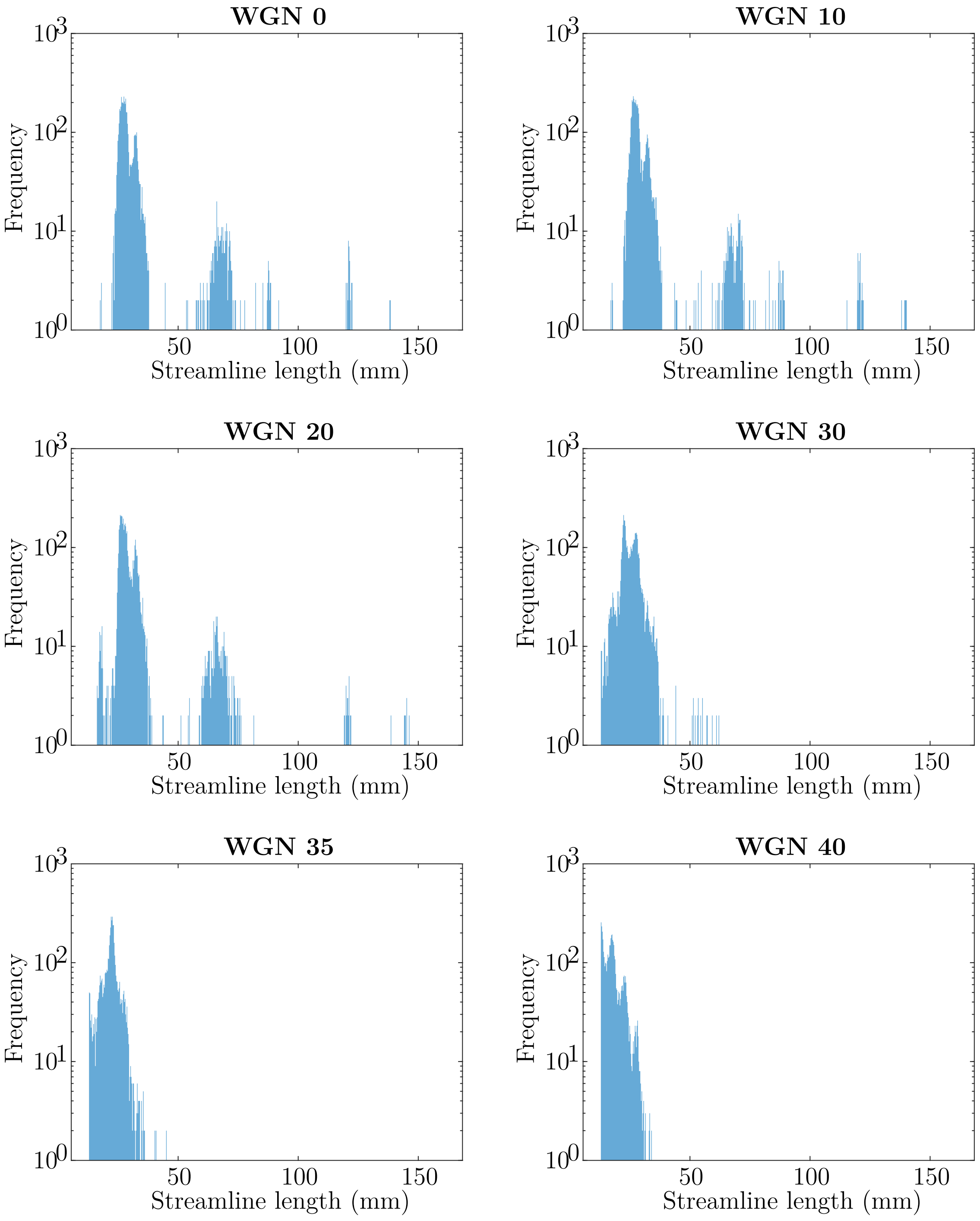


Figure 3.15: Histograms of streamline length for tractograms generated using MRtrix's Tensor_Prob algorithm with varied levels of added WGN. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered again consists of 5000 streamlines seeded from the base of the brain stem. The Tensor_Prob algorithm uses a step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are therefore shown on a logarithmic scale to better visualise regions of low counts in the distribution.

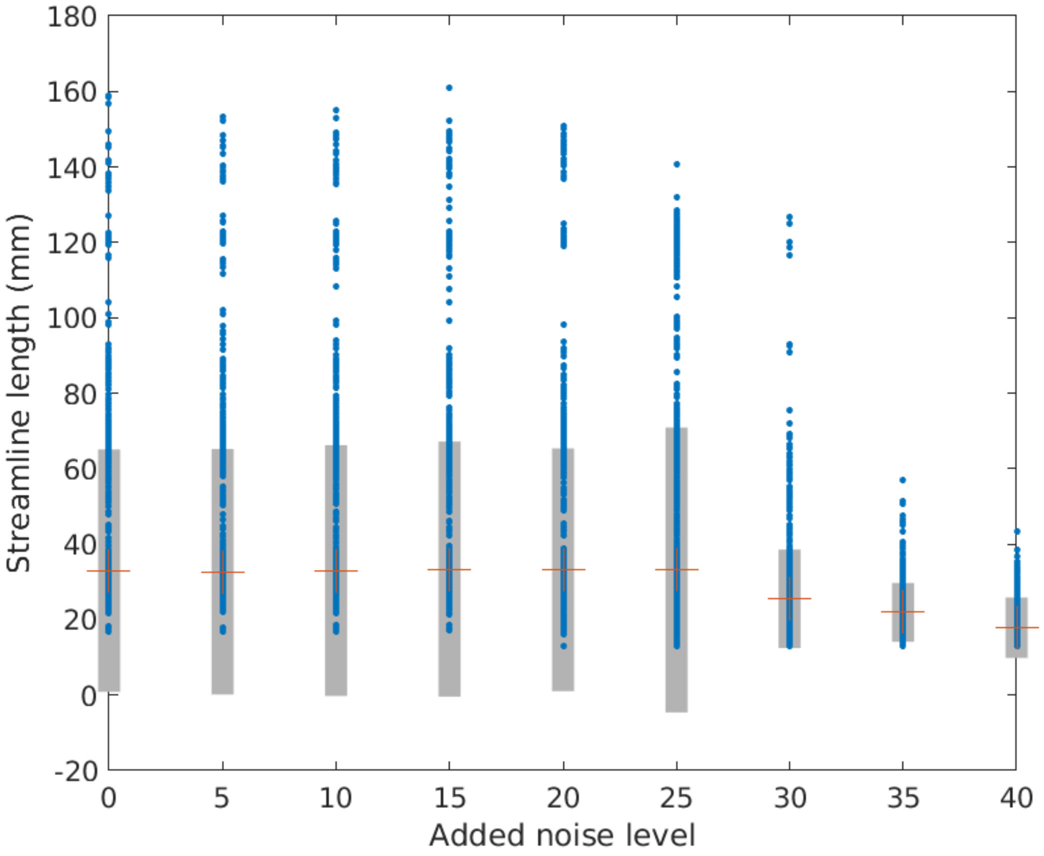


Figure 3.16: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with Tensor_Prob as WGN is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red, and grey, respectively.

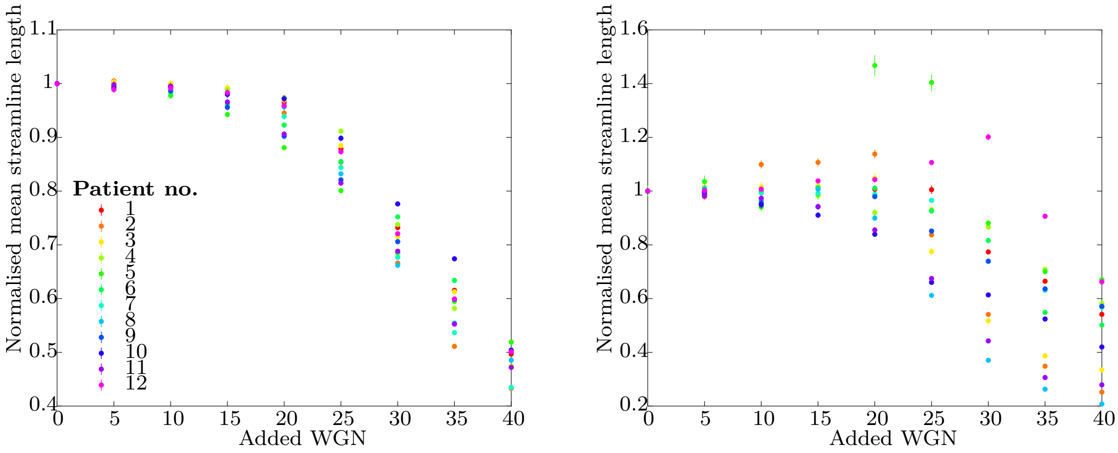


Figure 3.17: Plot of the normalised mean streamline length versus added noise for all whole-brain (left) and single tract tractograms (right) generated using Tensor_Prob.

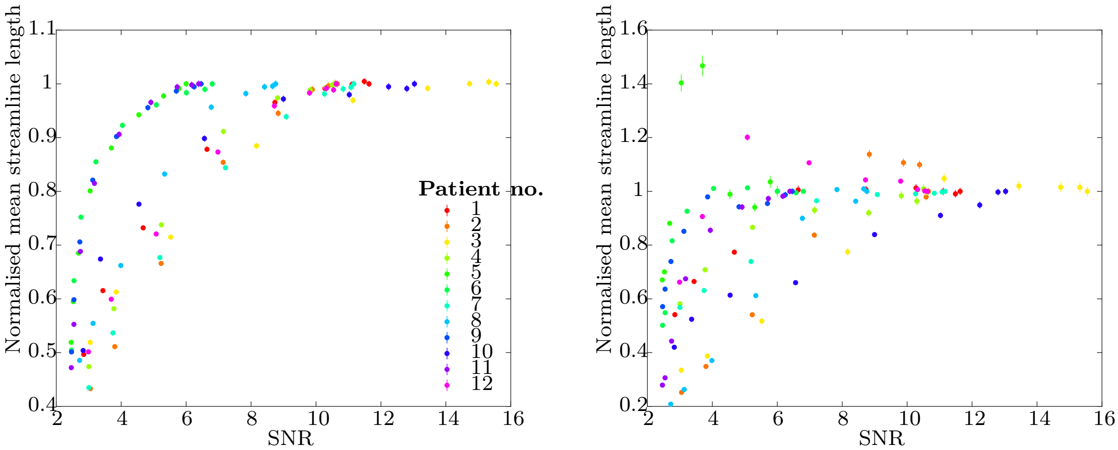


Figure 3.18: Plot of the normalised mean streamline length versus SNR (calculated using the single image method) for all whole-brain (left) and single tract tractograms (right) generated using Tensor_Prob.

3.9 Polar plots

For a selection of streamlines for each patient and increment of noise, a polar plot was produced, visualising the change in the angle of the streamline with respect to the initial angle of the tract (i.e. the zenith angle), θ, as the distance from the base of the tract, r, increases. The change in the angle of the streamline with respect to the x-axis (the azimuthal angle), ϕ was also plotted against r. 3D histograms were generated to illustrate the density of the streamlines in these plots.

Each streamline which comprises a tractogram is stored as a series of points in 3D space. A polar plot illustrates the relationship between the distance from the origin, r, and the direction of the streamline with respect to its initial direction. The direction at some point along the streamline is given by the vector connecting the point and the previous point in the streamline:

$$P_i - P_{i-1} = \frac{(x_i, y_i, z_i) - (x_{i-1}, y_{i-1}, z_{i-1})}{\sqrt{x_i^2 + y_i^2 + z_i^2} \cdot \sqrt{x_{i-1}^2 + y_{i-1}^2 + z_{i-1}^2}} \tag{3.8}$$

Directionality was imparted to the streamlines by specifying an origin or seed point and tracking each streamline in one direction (i.e. by running each tractography algorithm with the option -unidirectional). With these specifications, θ varies between $-\pi$ and π and ϕ varies between 0 and 2π. The default behaviour of MRtrix is to track streamlines bidirectionally, which would result in loss of directional information.

3.9.1 iFOD2

For each patient and level of added noise, a polar plot was generated from the tractogram produced by iFOD2. For the majority of patients the characteristic tract structure was retained for noise levels up to about 20 added WGN. Figure 3.19 shows the change in the θ versus r polar plot as noise is added to the input dMRI scan for patient 1.

A progressive deterioration of the white matter structures is observed with increasing noise, with only the base of the tract remaining at the highest noise level. Parts of the tract with strong connectivity are retained, since the SNR is sufficiently high in these regions in the initial image that the noise does not overwhelm the diffusion signal. However, more intricate pathways are altered or lost as the SNR decreases, since lower diffusion signals become increasingly difficult to discern.

The deterioration of the tract with noise is shown from a different perspective in Figure 3.20. These polar plots illustrate the change in the distribution of ϕ vs r with added noise. Since the streamlines superior to the brain stem spread out in all directions, the distribution of ϕ appears largely uniform around r = 60. This uniformity is compromised for noise levels of about 30 WGN and above. This is likely due to early termination of the streamlines at these noise levels, which reduces the densities of regions with higher r values.

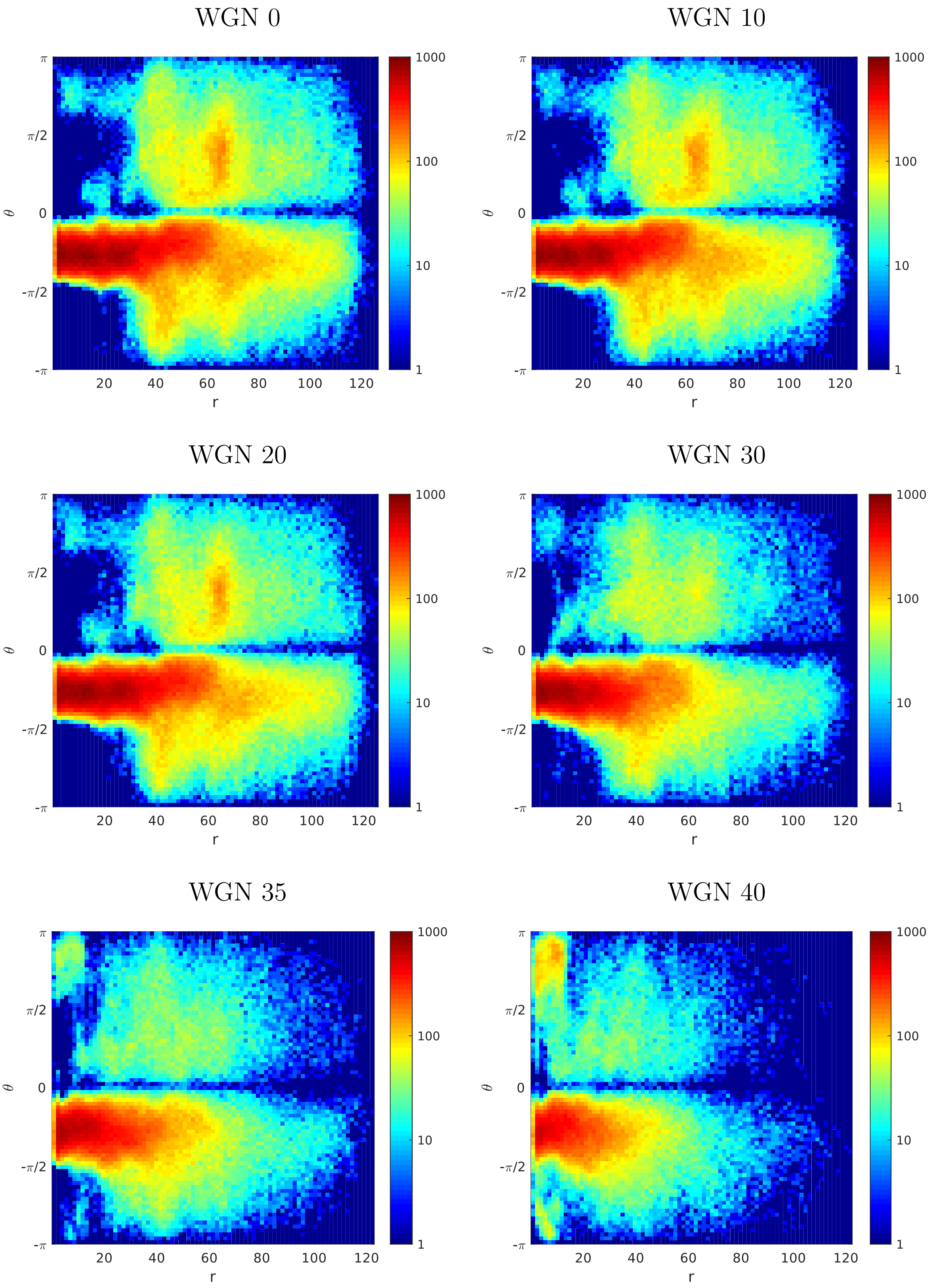
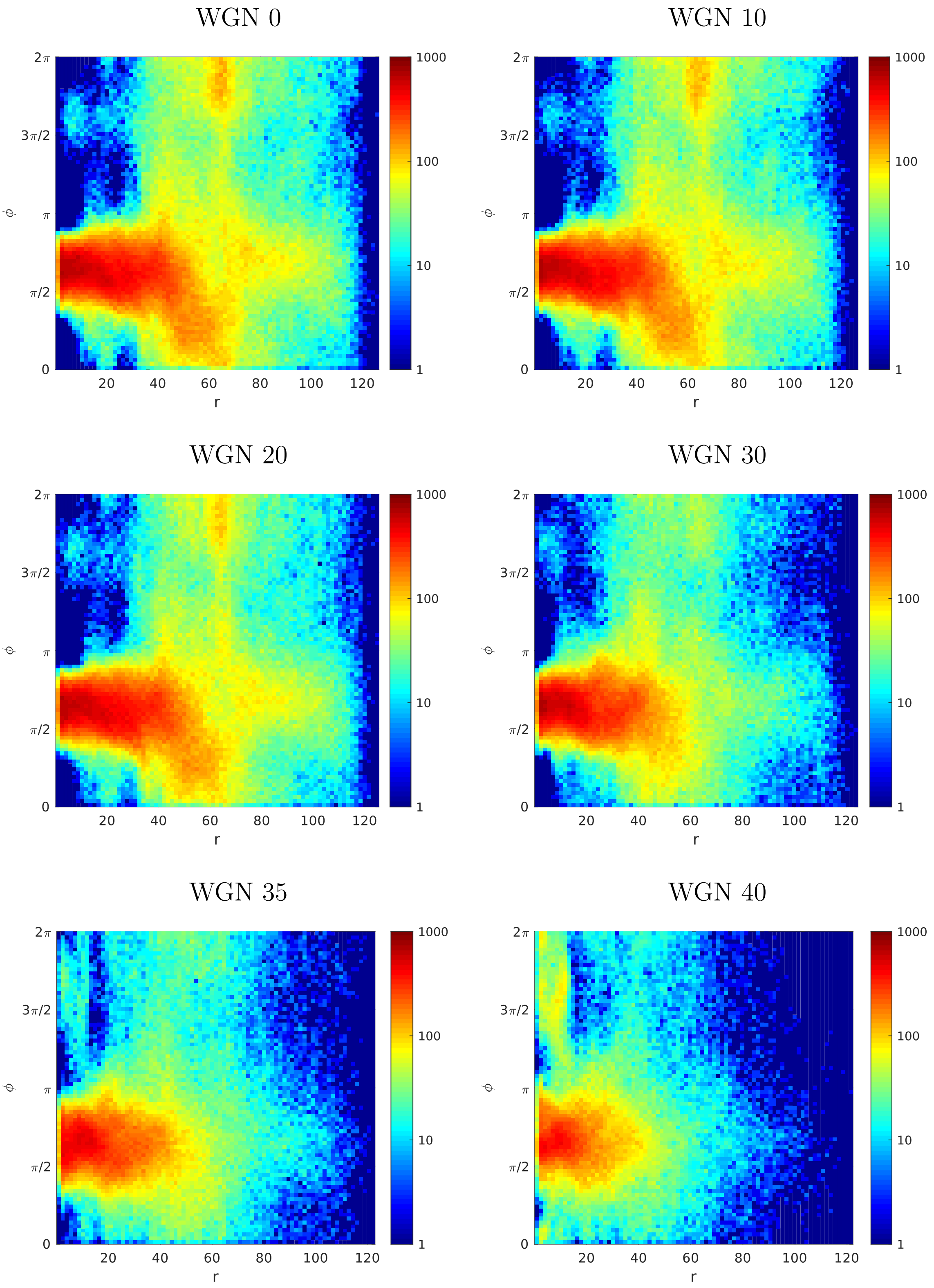


Figure 3.19: Polar plots of θ vs r for tractograms with added noise generated using iFOD2 for patient 1.The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.



WGN 0

WGN 10

WGN 20

WGN 30

WGN 35

WGN 40

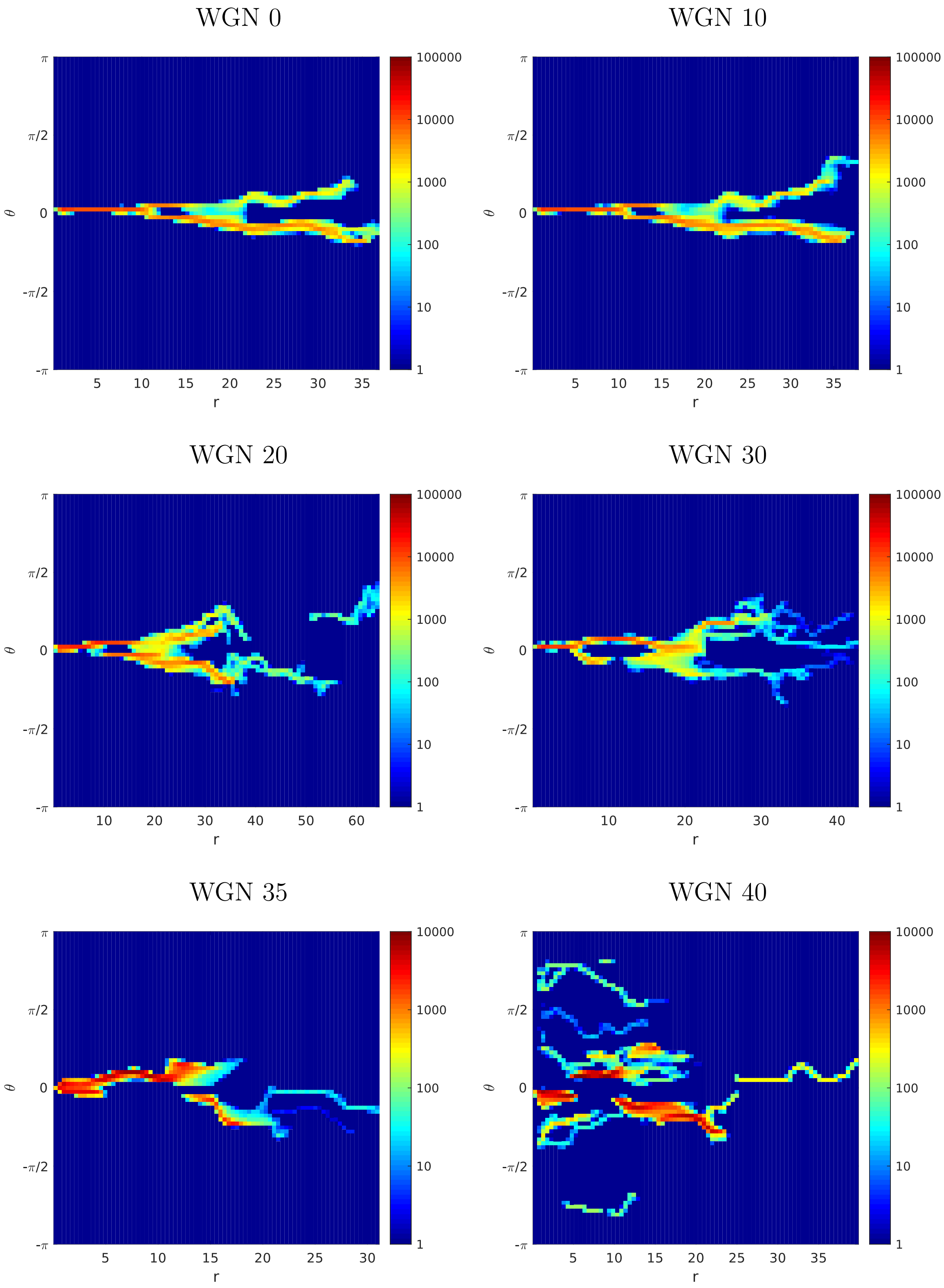
Figure 3.20: Polar plots of ϕ vs r for tractograms with added noise generated using iFOD2 for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

3.9.2 SD_STREAM

The polar plot analysis was repeated with the SD_STREAM tractography algorithm, with similar results. Tracts were observed to retain their characteristic structure up to an added WGN level of about 20. Beyond this noise level, the lower intensity features in the initial polar plot began to deteriorate or become warped out of shape.

In the case of Patient 1, shown in Figure 3.21, a well-defined forked tract is present in the initial θ versus r polar plot. With added WGN 10, most of the structure is retained. However, a portion of the streamlines following the path with a positive θ appear to have been falsely elongated. The lower streamlines also seem to have lost some structure near the end of their path. At noise levels 20 and above, the loss of the tract structure becomes increasingly severe until, at WGN 40, there is no apparent structure left.

This progression from a defined to a chaotic tract is also reflected in the ϕ versus r polar plots for the patient (see Figure 3.22). The polar plots for added WGN 0 and 10 closely resemble one another. However, some loss of density is evident in the WGN 10 plot, where the upper branch forks off from the tract base. The elongation of the upper streamlines and the change in structure of the lower streamlines is also visualised more clearly by the differences in these plots. Note the change in the r scale from WGN 10 to WGN 20. The elongation of a selection of the streamlines becomes more pronounced, and the fork structure more diffuse, since the added noise has obscured the original strong diffusion signals. While the original forked tract possessed paths with specific directions in the x-y plane, the tractogram produced for the dMRI with added 30 WGN has streamlines which branch out roughly uniformly in the x-y plane at r = 20. This demonstrates the inability of the algorithm to decide on a set path or paths for the streamlines to follow. Unusual behaviour occurs at WGN 40, where the streamlines appear to intertwine together into an intense band.


WGN 0

WGN 10

WGN 20

WGN 30

WGN 35

WGN 40

Figure 3.21: Polar plots of θ vs r for tractograms with added noise generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

ϕ vs r

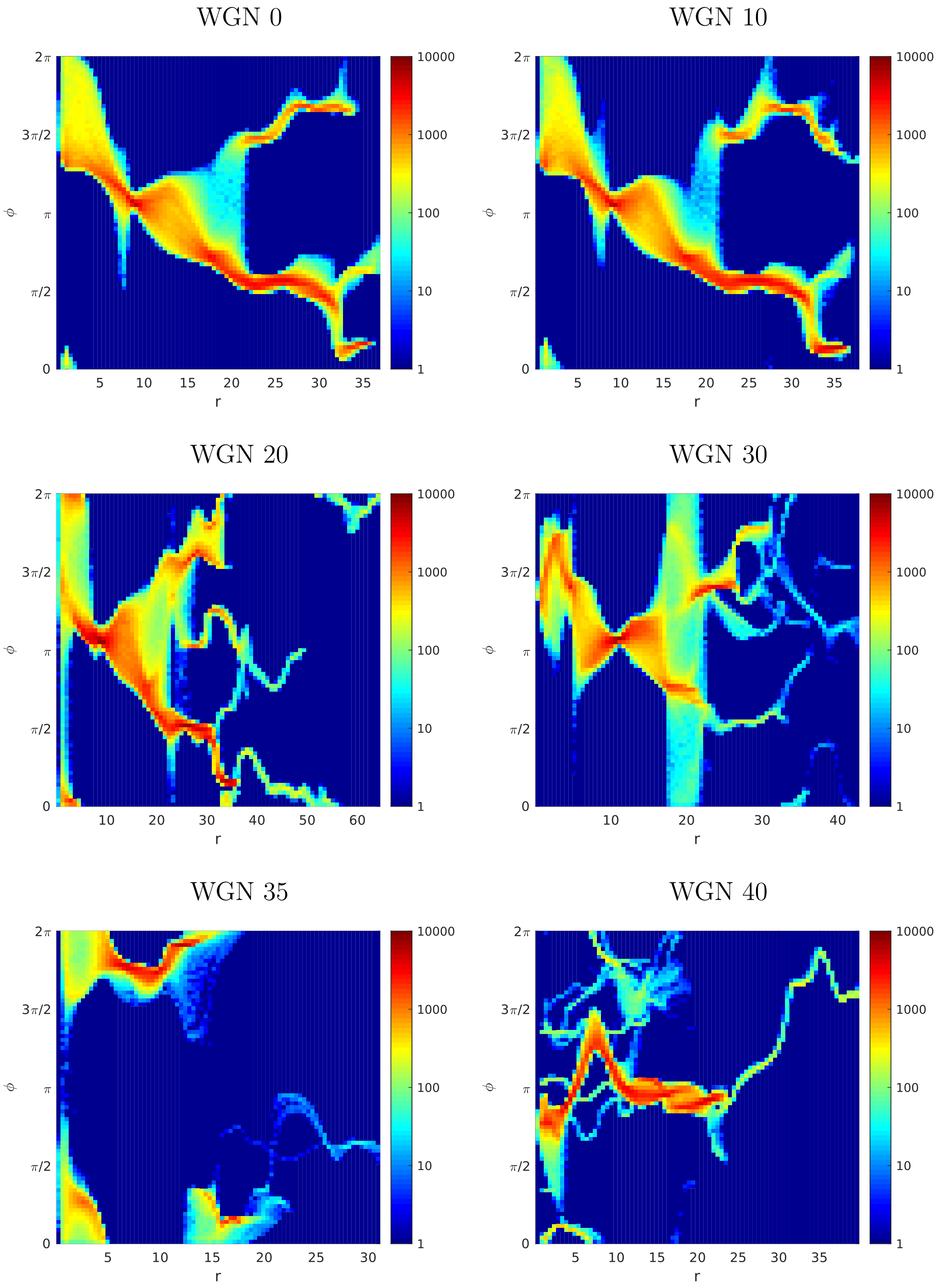


Figure 3.22: Polar plots of ϕ vs r for tractograms with added noise generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

3.9.3 Tensor_Prob

Polar plots were also generated following application of the Tensor_Prob algorithm to each of the input dMRI scans. Observations for these polar plots were similar, with the key point beyond which the tract loses its structure being WGN 20.

Again, intricate structures are the first to go. This is evident in the transition of the θ vs r polar plot for Patient 1 from WGN 10 to WGN 20 (see Figure 3.23). In the area surrounding r = 10, $\theta = 0$, the original polar plot contains a highly detailed structure, which is reconstructed with high accuracy for WGN 10. However, in the WGN 20 polar plot, a significant drop in intensity is observed in this region, which reflects the fact that fewer streamlines were delineated with that length and shape. The portion of the original polar plot which appears as red (i.e. the region with the highest streamline density) remains even when WGN 40 is added, since the diffusion signal was sufficiently strong to remain above the noise.

The ϕ vs r plots in figure 3.24 show the expected deterioration with noise. Initially, the streamlines have some characteristic structure in the x-y plane as the streamline is traced from the origin point. Detail is lost around WGN 20, and the density distribution starts to rapidly diffuse beyond this point. At the highest level of added noise (i.e. WGN 40), the probability of a streamline tracing a direction in the x-y plane is essentially equal in all directions. This implies that the streamlines trace random paths through the x-y plane.

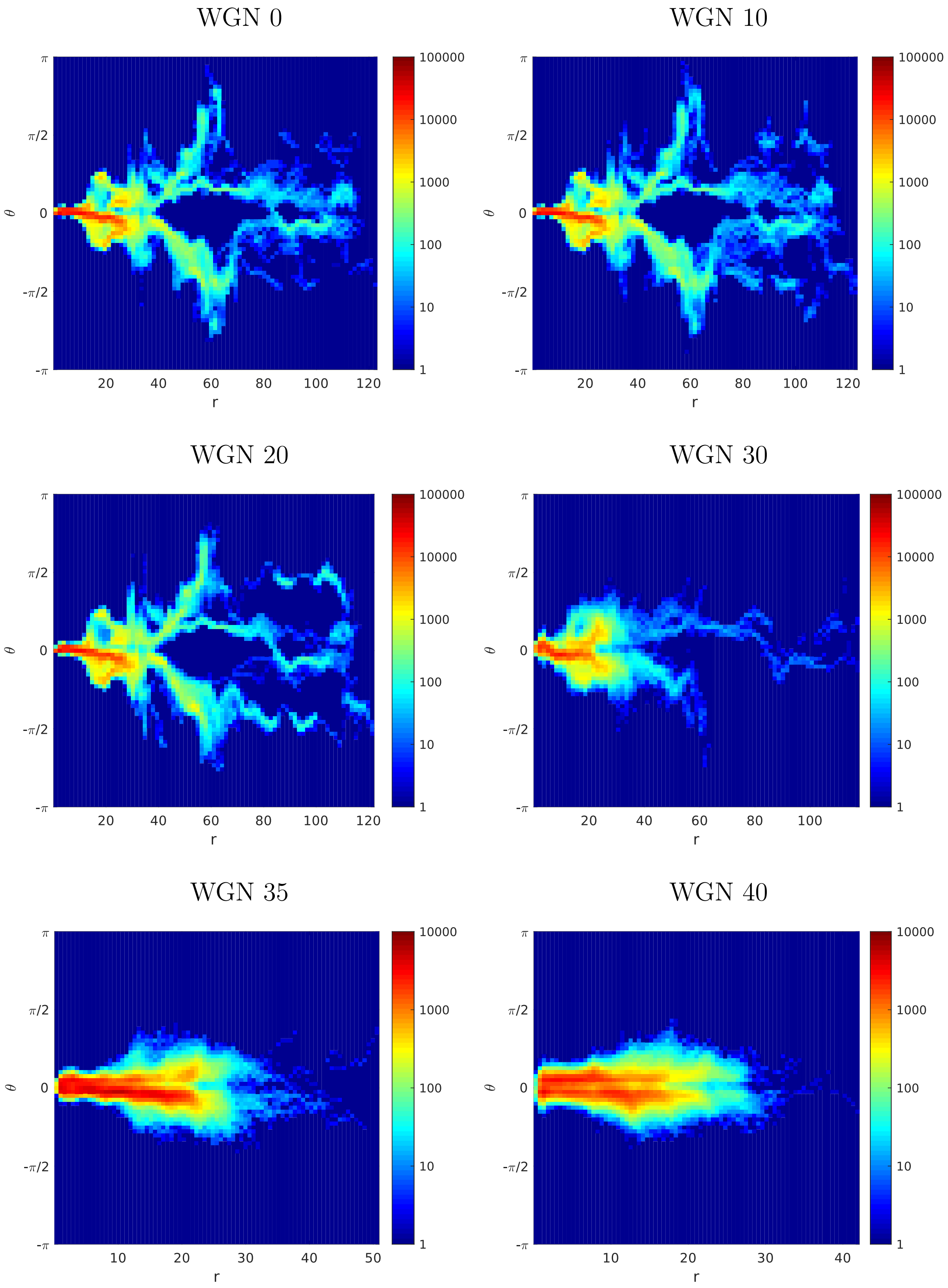


Figure 3.23: Polar plots of θ vs r for tractograms with added noise generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

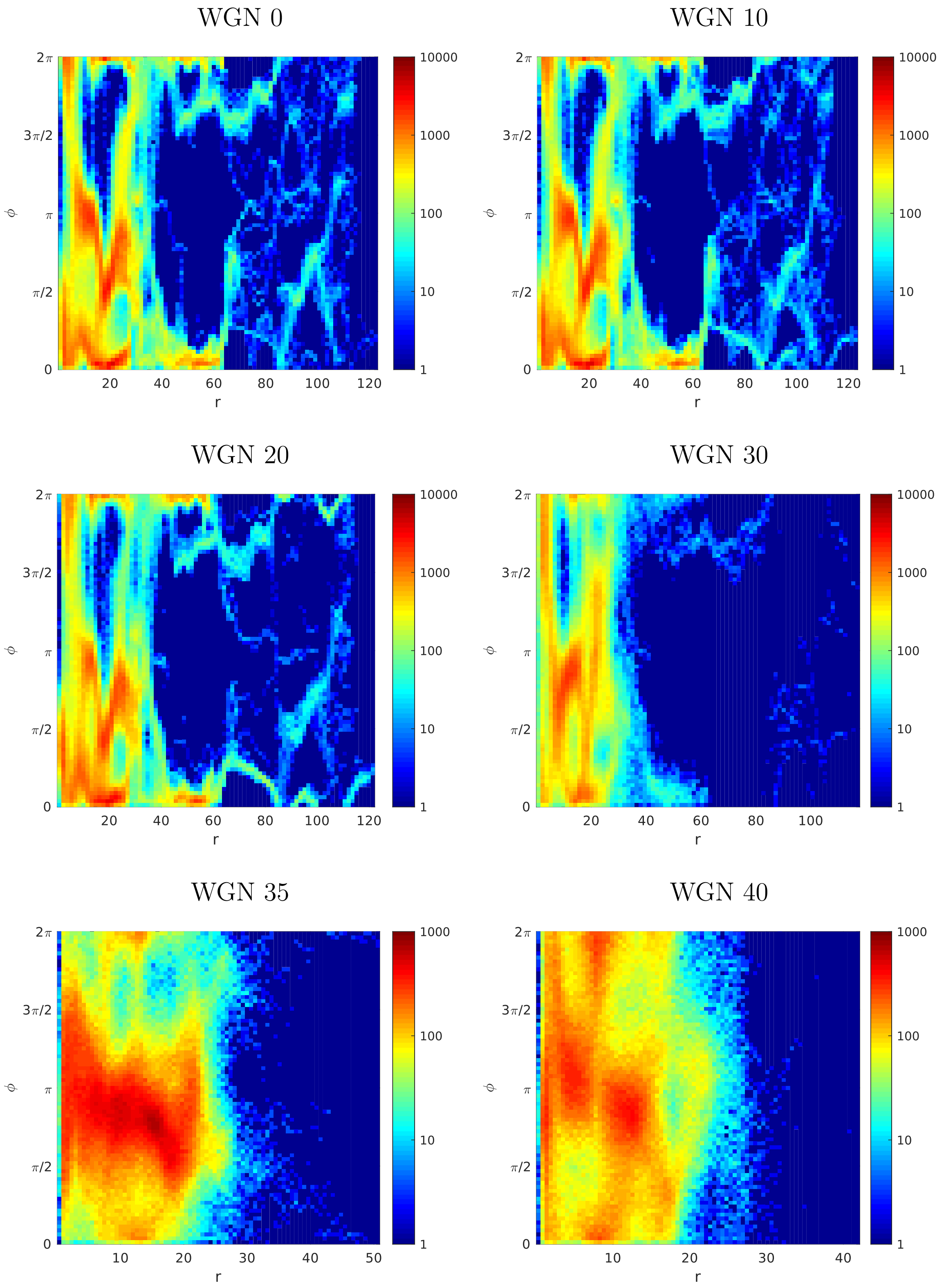


Figure 3.24: Polar plots of ϕ vs r for tractograms with added noise generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

3.10 Resemblance plots

A metric was devised to measure the resemblance of each modified tractogram to the original tractogram without added WGN. For each patient and increment of noise, a tract map was generated from a 50k streamline tractogram of the whole brain, with the value in each voxel representing the density of streamlines in that region.

The resemblance value was obtained via the following method. Repeats of the original tractogram and tract-weighted map were generated from the same dMRI scan. Due to the inherent randomness of seeding streamlines, these repeat tract-weighted maps differed to some extent. For each voxel in the unmodified tract-weighted map, a threshold was set, equal to 3 times the SD of the values belonging to that voxel from the repeat tract-weighted maps. For voxels where the repeat original tract-weighted maps had the same value, and the SD is therefore zero, the threshold was replaced with a default value.

If a voxel in the noise-affected tract-weighted map was within +/- the threshold of the corresponding voxel in the original tract-weighted map, then that voxel of the modified tract-weighted map was classified as "like" the corresponding voxel of the original tract-weighted map. If it lay outside of these bounds it was classified as "unlike". Any voxels for which both the modified and original tract-weighted maps contained no streamlines were excluded from the calculation. The resemblance was then calculated as the sum of the "like" voxels for the noisy tract-weighted map, divided by the number of "like" voxels for the no WGN tract-weighted map.

i.e. The threshold for voxel $vox_{WGN=0,rpt=1}$ is:

$$T = 3 \times \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(vox_{WGN=0,rpt=i} - \overline{vox_{WGN=0}}\right)^2} \tag{3.9}$$

where N is the number of repeat unmodified tract-weighted maps.

A voxel in the modified tract-weighted map, $vox_{WGN=n,rpt=1}$, is therefore accepted if it satisfies the following condition:

$$vox_{WGN=0,rpt=1} - T < vox_{WGN=n} < vox_{WGN=0,rpt=1} + T \tag{3.10}$$

The number of "like" voxels for the tract-weighted map with no added WGN was determined by comparing it with a second tract-weighted map generated from the same input dMRI. Since the repeat tract-weighted map was also used to derive the threshold values, the resemblance value for the unmodified tract-weighted map is 1 by definition. The resemblance is therefore the fraction of voxels with values within 3 SD of the original tract-weighted map. This provides a measure of how well the white matter structure is retained as noise is added.

Figures 3.25, 3.26 and 3.27 show the relationship between the computed resemblance value and the level of added WGN or the SNR calculated using the difference method for each tractography algorithm.

For the probabilistic algorithms, iFOD2 and Tensor_Prob, the resemblance value remains constant at low levels of added WGN, with a spread of about 10%. For iFOD2, the resemblance drops from 1 at no noise to about 0.7 (or 70%) for WGN of powers 5 through 15. For Tensor_Prob, the resemblance for WGN powers 5 and 10 is approximately 0.8 (or 80%). For iFOD2 this drop in resemblances coincides with an SNR of approximately 4. A progressive decrease in resemblance is observed as additional noise is added beyond this point. For Tensor_Prob the drop in resemblance is more gradual but appears to start at around an SNR of 8. A sharper drop in resemblance at an SNR of about 4 is evident for patients 5, 6, 9 and 11.

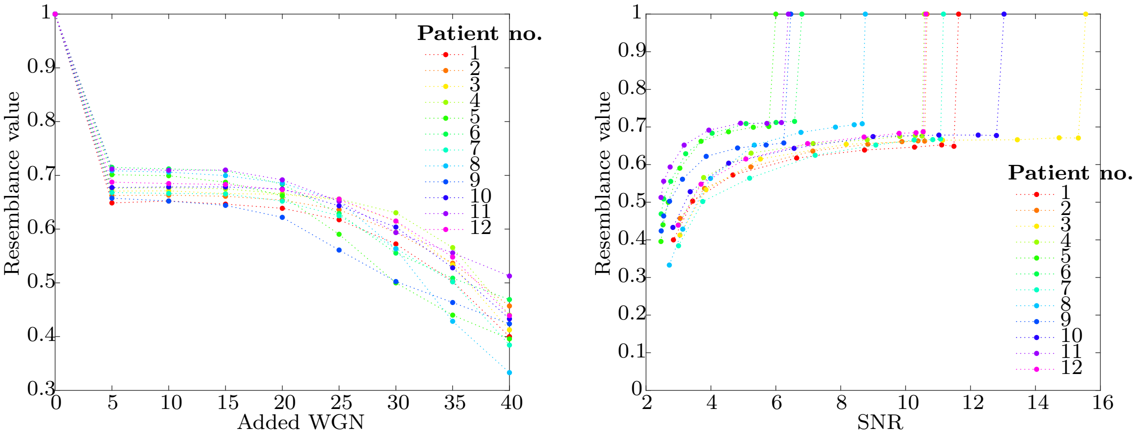


Figure 3.25: Plot of the calculated resemblance value versus added WGN (left) and SNR (right) for tract-weighted maps generated using iFOD2.

For the deterministic SD_STREAM algorithm, the resemblance value decreases linearly with added WGN. The SNR also progressively drops off, in contrast to the sharp drop observed for iFOD2.

A significant drop in tract-weighted map similarity is evident between WGN 0 and 5. To explain this effect, consider the tree-like structure of the whole-brain tractogram. Regions where there is a strong diffusion signal will appear as regions of high streamline density in the resulting tract-weighted map. More intricate white matter pathways produce less intense diffusion signals. Structures which facilitate the least diffusion (i.e. fine white matter structures) will therefore appear as regions with the lowest streamline density.

In theory, repeat tract-weighted maps generated from the same dMRI should be identical, as they are derived from the same diffusion signals. However, since streamlines are seeded at random points in the brain for each tractogram, there may be slight variations in output with each repetition. These fluctuations primarily affect the ability of the algorithm to accurately reproduce regions of the brain where the diffusion signal is weak (i.e. of the order of the image noise).

Adding WGN of power 5 (i.e. Gaussian noise with mean 0 and variance 3.16) magnifies the uncertainty in the dMRI, such that early termination of streamlines occurs within regions of the weakest diffusion signal (i.e. the finest white matter structures). The main white matter structures are retained, as diffusion signals are sufficiently high that they are not overwhelmed by the added noise. Based on the plot, the proportion of the brain comprised of these fine white matter structures is somewhere between 20% and 30%.

With the addition of WGN of power 10 (i.e. Gaussian noise with mean 0 and variance 10), the tractogram similarly loses fine white matter structures. However, many of the same fine structures are lost in this tractogram as for the WGN 5 tractogram. Again, the majority of the main white matter structures are unaffected. This explains why a significant drop occurs between WGN 0 and WGN 5, but not between WGN 5 and WGN 10. A significant amount of noise (i.e. WGN 20, with mean 0 and variance 100) must be added before the core, high diffusion-signal white matter structures begin to deteriorate. This hypothesis is supported by observation of the tractograms and polar plot histograms, where only slight differences are evident between WGN 0 and WGN 20, and more severe loss of information occurs beyond WGN 20 for the probabilistic algorithms.

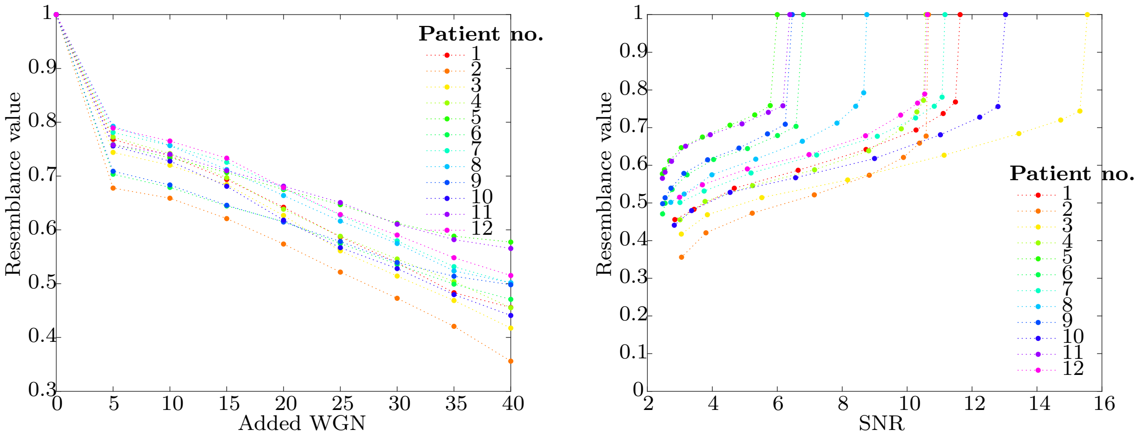


Figure 3.26: Plot of the calculated resemblance value versus added WGN (left) and SNR (right) for tract-weighted maps generated using SD_STREAM.

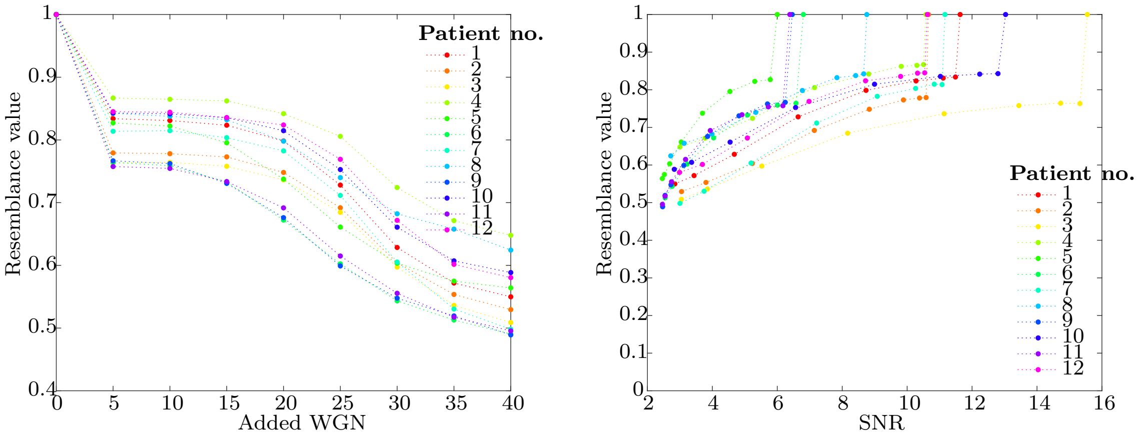


Figure 3.27: Plot of the calculated resemblance value versus added WGN (left) and SNR (right) for tract-weighted maps generated using Tensor_Prob.

The resemblance value can also be used to compare the polar plot histograms for each pair of original and modified tracts. Figures 3.28 through 3.30 show the relationship between the resemblance value calculated for the θ (left) and ϕ (right) polar plots and added WGN and the calculated SNR. Very similar results are observed as for the whole-brain tract-weighted maps for iFOD2, with the fall in resemblance occurring at an SNR of around 8 for the majority of patients, and at about an SNR of 4-5 for patients 5, 6, 9 and 11. For SD_STREAM and Tensor_Prob, there is no obvious dip in resemblance for a particular SNR value. However, the resemblance is still inversely related to the level of noise in the input dMRI. For both algorithms the resemblance appears to sharply decrease with the initial few increments of added WGN, irrespective of the initial SNR of the scan.

The spread in resemblance values for the 12 patients is significantly larger than that calculated from the whole-brain tract-weighted maps. One contributing factor may be a simple reduction in statistics, since the polar plots were derived from a 5k tractogram, whereas the tract-weighted maps were derived from a 50k tractogram. Since only a single tract is represented by each polar plot, the presence of false positives in some tractograms may also be a source of divergence in resemblance values, whereas the effect of premature streamline termination would likely predominate over false elongation of streamlines in the full brain tract-weighted map.

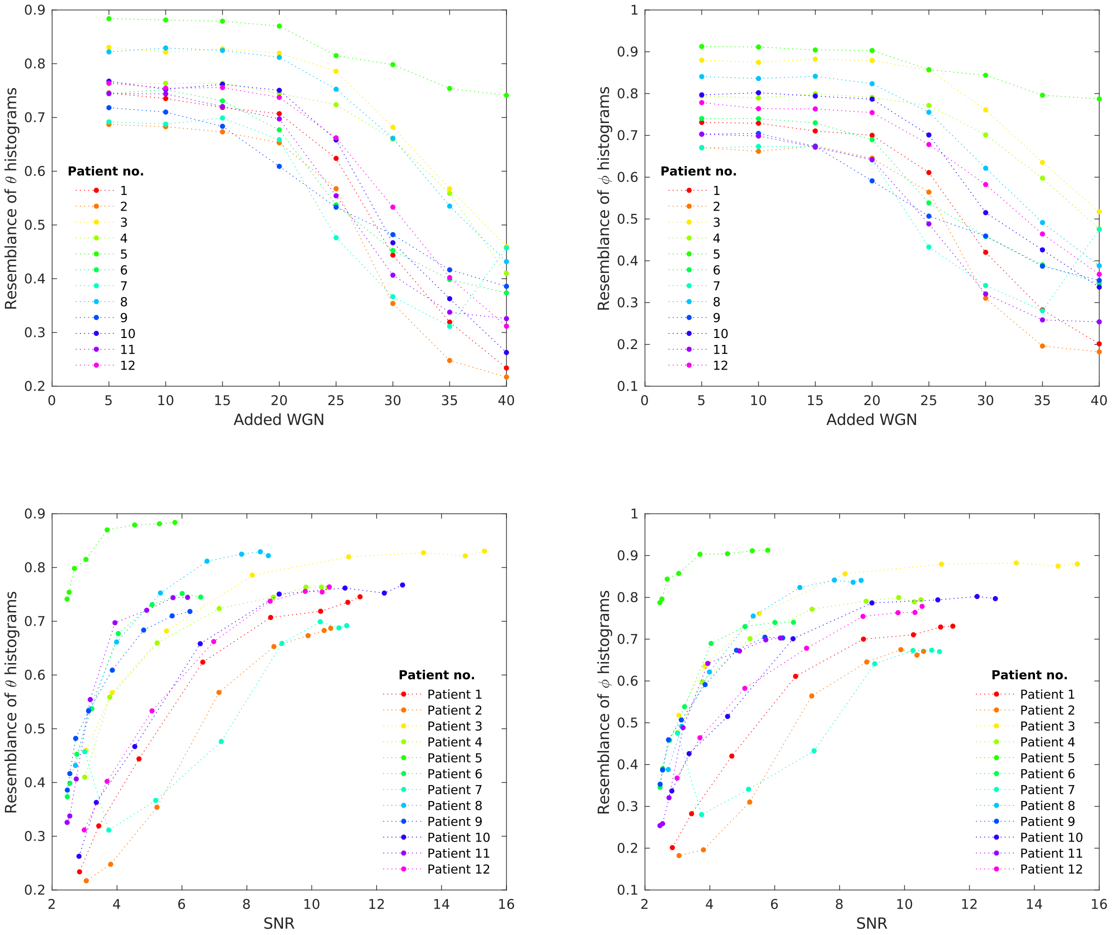


Figure 3.28: Plot of resemblance values versus added WGN for polar plot histograms of θ (top left) and ϕ (top right) generated from iFOD2 tractograms. Also shown are the resemblance values plotted against the calculated SNR using the single image method.

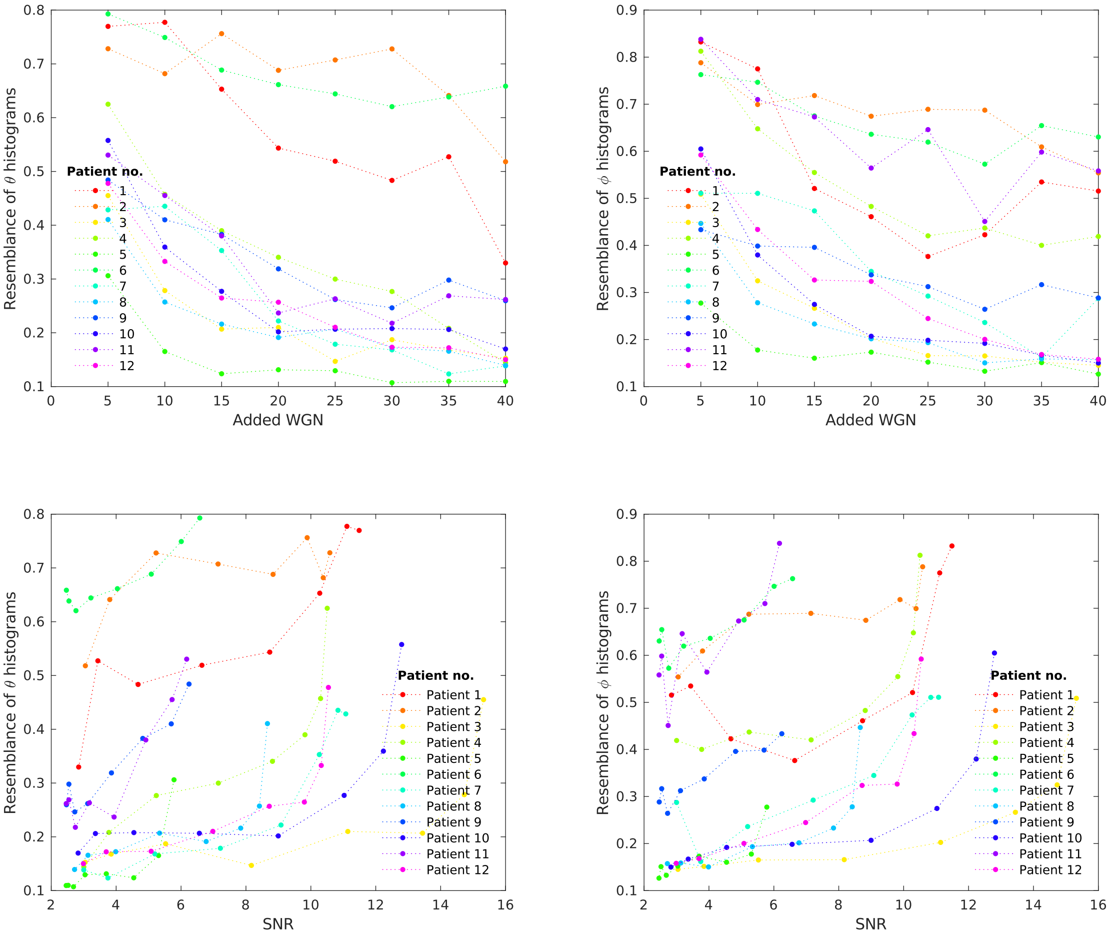


Figure 3.29: Plot of resemblance values versus added WGN for polar plot histograms of θ (top) and ϕ (bottom) generated from SD_STREAM tractograms. Also shown are the resemblance values plotted against the calculated SNR using the single image method.

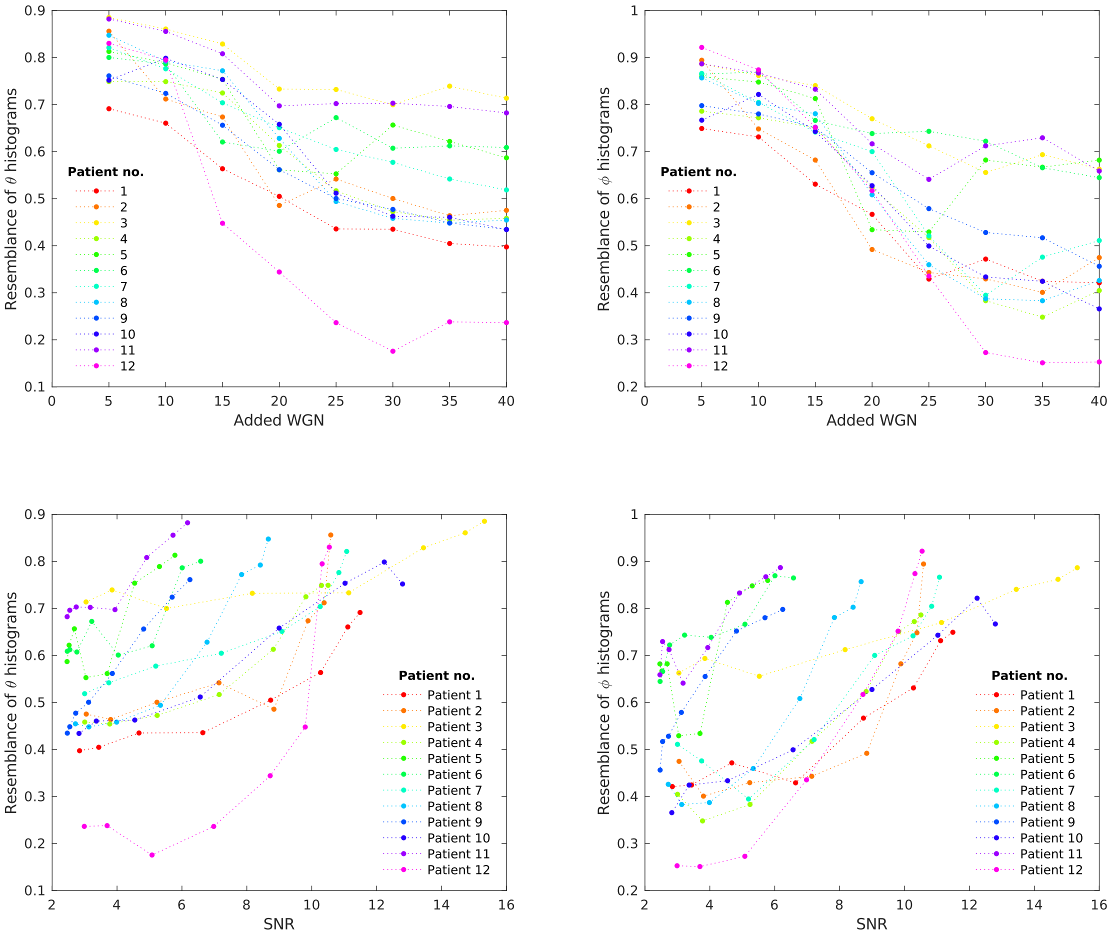


Figure 3.30: Plot of resemblance values versus added WGN for polar plot histograms of θ (top) and ϕ (bottom) generated from Tensor_Prob tractograms. Also shown are the resemblance values plotted against the calculated SNR using the single image method.

3.11 RMS difference in tract-weighted maps

Tract weighted maps generated for different noise levels were compared using Matlab for each of the tractography algorithms of interest. The RMS difference between each unmodified tract-weighted map and noisy tract-weighted map pair was calculated and plotted against the level of added WGN and the calculated SNR using the difference method.

Figure 3.31 demonstrates that the RMS difference between the tract-weighted maps is consistent up to about WGN 15 for iFOD2. For higher noise levels, the tract-weighted maps start to diverge more significantly. As in the resemblance plots, a turning point is evident at a SNR of about 4 for patients 5, 6, 9 and 11, and about 8 for the remaining patients.

Figure 3.32 also depicts results that are in line with observations from the resemblance plots for SD_STREAM. The RMS difference in the tract-weighted maps is a continuously increasing function of the added WGN. Again, the tract-weighted map progressively worsens with increasing noise. For patients 5, 6, 9 and 11 the difference metric rises more sharply below about an SNR of 3, indicating some significant loss of information below this.

For Tensor_Prob, Figure 3.33 shows that the tract-weighted maps are consistent (within

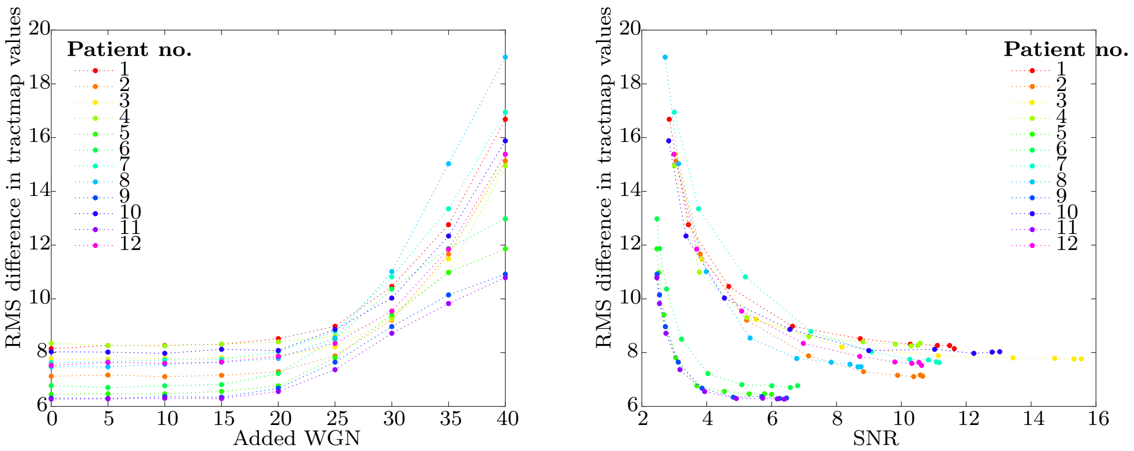


Figure 3.31: Plots of the RMS difference between the unmodified and noisy tract-weighted maps versus added WGN (left) and SNR (right) for iFOD2.

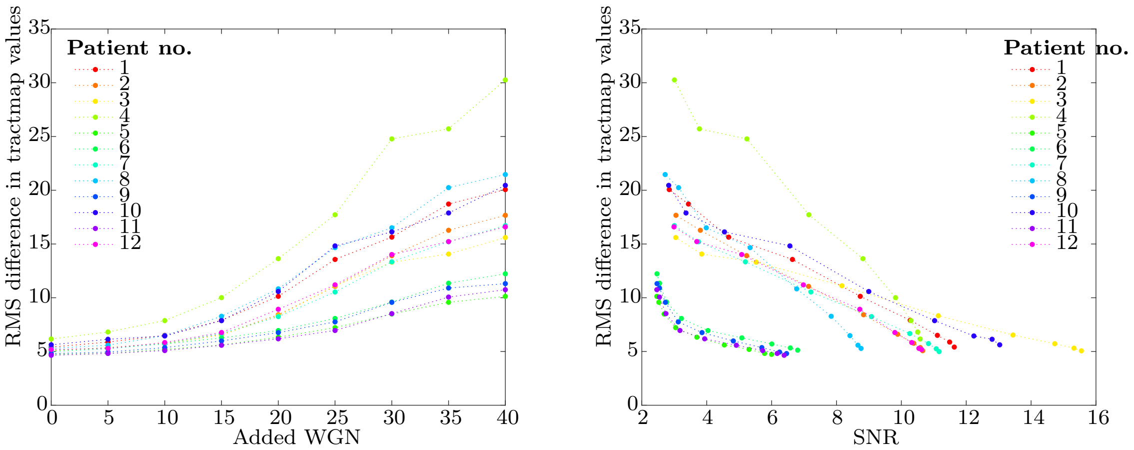


Figure 3.32: Plots of the RMS difference between the unmodified and noisy tract-weighted maps versus added WGN (left) and SNR (right) for SD_STREAM.

statistical uncertainty) up to added WGN 15. Beyond this, the RMS difference in the tract-weighted maps increases approximately linearly with added WGN. The SNR below which this decrease in performance seems to occur is about 4 for patients 5, 6, 9 and 11, and about 8 for the other patients.

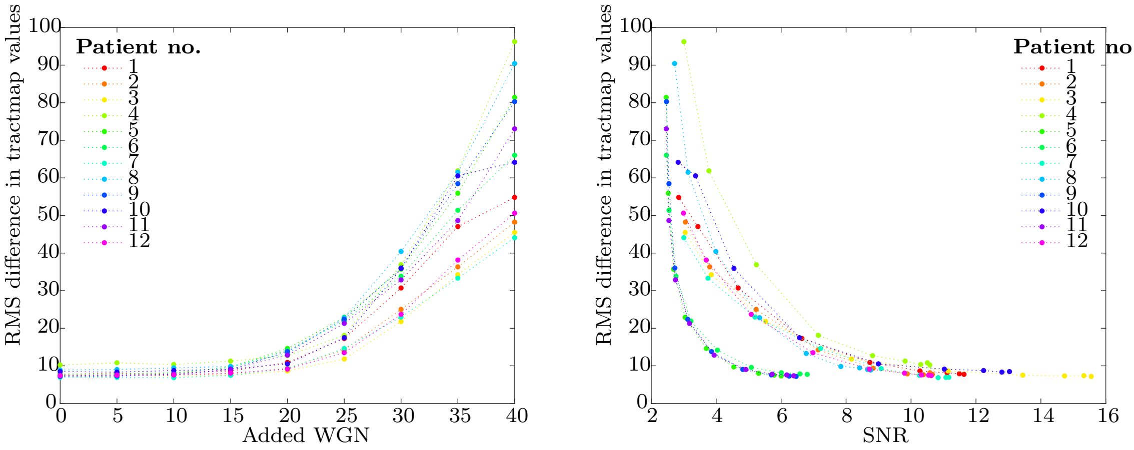


Figure 3.33: Plots of the RMS difference between the unmodified and noisy tract-weighted maps versus added WGN (left) and SNR (right) for Tensor_Prob.

3.12 Relationship between noise and the rate of false positives and false negatives

False positive and false negative rates were approximated by a nearest point search method. Repeat zero-added WGN tractograms (tractograms X and Y) were input into Matlab's k-nearest neighbours search method (knnsearch) to obtain the distance to the nearest point in tractogram X for each point in tractogram Y. The resulting distances were sorted and a threshold distance D was defined as the distance for which 99.7% (i.e. 3 σ) of points in tractogram X were within D of a point from tractogram Y.

For each tractogram derived from an artificial noisy dMRI Z, the k-nearest neighbours method was used to obtain the distances of each point in Z from a point in the unmodified tractogram X. The number of distances within D were counted and the true positive rate then calculated by taking the ratio of the total distances less than or equal to D and the total number of points in the modified tractogram. A value of 1 (or 100%) therefore implies that the original tractogram was reconstructed with 100% fidelity (i.e. all points in Z are within the threshold distance D of a point in X). The false positive rate was therefore defined as 1 minus this value (i.e. the number of points in Z that do not fall within the threshold distance D of a point in X). This measure gives the fraction of reconstructed points in the modified tractogram that were outside of the unmodified tractogram (with a certainty of 3 σ).

The false negative rate was determined by inverting the inputs for the k-nearest neighbours search and repeating this process. That is, the distances of each point in X to the nearest point in Z were obtained then the distances within D were counted and divided by the number of points in X, then subtracted from 1, giving the fraction of points in the original tractogram that were not reconstructed in the modified tractogram (with a certainty of 3 σ).

The false positive and negative rates were recalculated using the repeat unmodified tractogram (Y), which provided an estimate of their respective errors. The error in the false positive rate was then determined as the standard deviation of the number of distances less than or equal to D for each unmodified tractogram. Similarly, the error in the false negative rate was calculated as the standard deviation of the number of distances greater than D for each unmodified tractogram. A significantly larger number of repeat tractograms would need to be generated to obtain a reliable estimate of the error in the calculated false positive and false negative rates.

As the level of added WGN is increased, false negative and positive rates generally increase. The shape of the trend which describes how false positive and false negatives increase with noise closely resembles the trends which relate the resemblance value and the RMS tractogram difference to the level of added WGN. The percentage of false positives and negatives is relatively constant for each of the probabilistic algorithms for added WGN up to about 15. The rates of both false positive and false negatives increase in general beyond this point. The increase in false positive and negative rates for SD_STREAM is a continuously increasing function of the added WGN.

What differs substantially between the algorithms is the fraction of false positives and negatives produced. For iFOD2, the percentage of false positives goes as roughly a quarter of the percentage of false negatives, with the maximum false positive and negative rates being 4% and 16% respectively at added WGN 40.

For SD_STREAM, the percentage of false positives is directly proportional to the percentage of false negatives, both reaching a maximum of around 90% at added WGN 40.

The Tensor_Prob algorithm also produces false positives and false negatives at a similar rate, with both reaching approximately 70% at added WGN 40.

It is clear that iFOD2 produces the least false positive and false negative results compared to SD_STREAM and Tensor_Prob. This algorithm therefore appears to be the most robust to noise.

The superiority of iFOD2 over Tensor_Prob can be attributed to several factors. Firstly, Tensor_Prob is a first-order algorithm where each step is sampled once which has a tendency to overshoot curved bundles. In contrast, iFOD2 is a second order algorithm where each "step" is actually an arc through space where the FOD is sampled at multiple spatial points. In iFOD2, the amplitudes of the FOD at a tangent to all points along the arc are calculated and multiplied together to give the probability of following that arc. This process reduces overshoot in curved bundles.

Additionally, the Tensor_Prob algorithm only models imaging noise because it uses the wild bootstrap approach, which produces minimal streamline dispersion when there is little noise and deteriorates as further noise is added. In contrast, iFOD2 models uncertainty at the level of the entire FOD, since dMRI does not have sufficiently high angular resolution to distinguish between mild and no dispersion. It has also been established that fibre dispersion exists in the brain (Mollink et al., 2017), so the assumption that fibres are perfectly coherently oriented, which is intrinsic to the Tensor_Prob algorithm, is false.

SD_STREAM appears to perform the poorest with respect to the production of false positives and negatives. This is likely due to the fact that this algorithm is deterministic rather than probabilistic like the iFOD2 and Tensor_Prob algorithms. SD_STREAM does not take uncertainties in the estimated fibre orientations into account, so provided a given pathway does not trigger its termination criteria, it will delineate a low probability pathway with equal weighting to a high probability pathway, while the probabilistic algorithms weight paths by their probability.

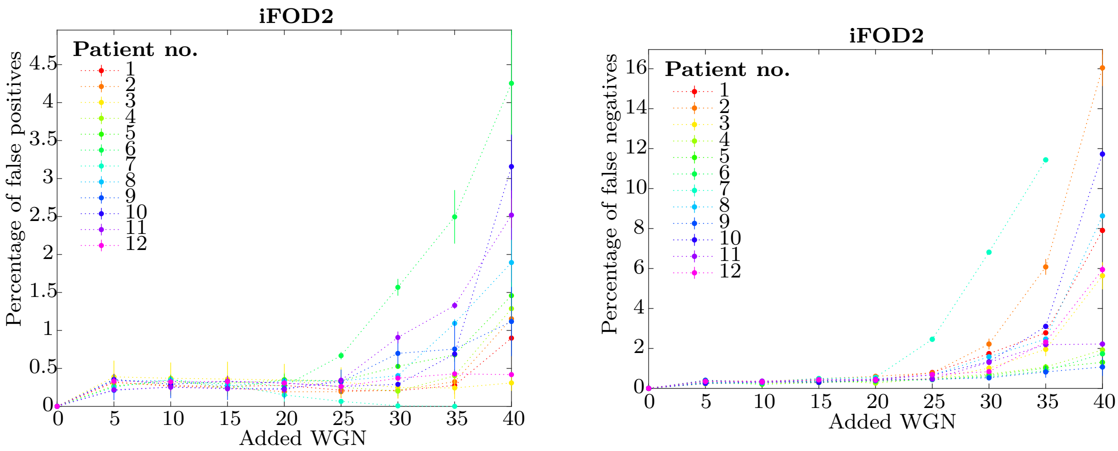


Figure 3.34: False positives (top) and false negatives (bottom) in iFOD2 tractograms versus noise calculated using a nearest point search method with a threshold of 3 SD.

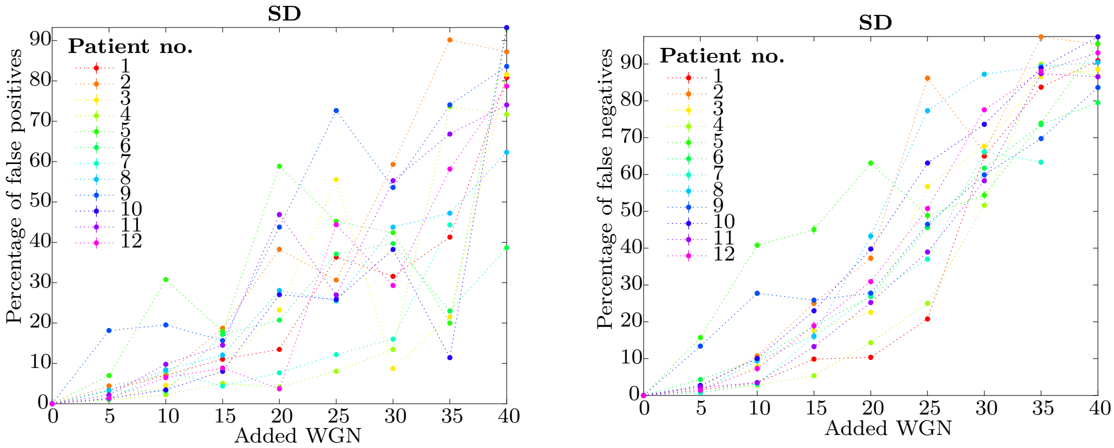


Figure 3.35: False positives (top) and false negatives (bottom) in SD_STREAM tractograms versus noise calculated using a nearest point search method with a threshold of 3 SD.

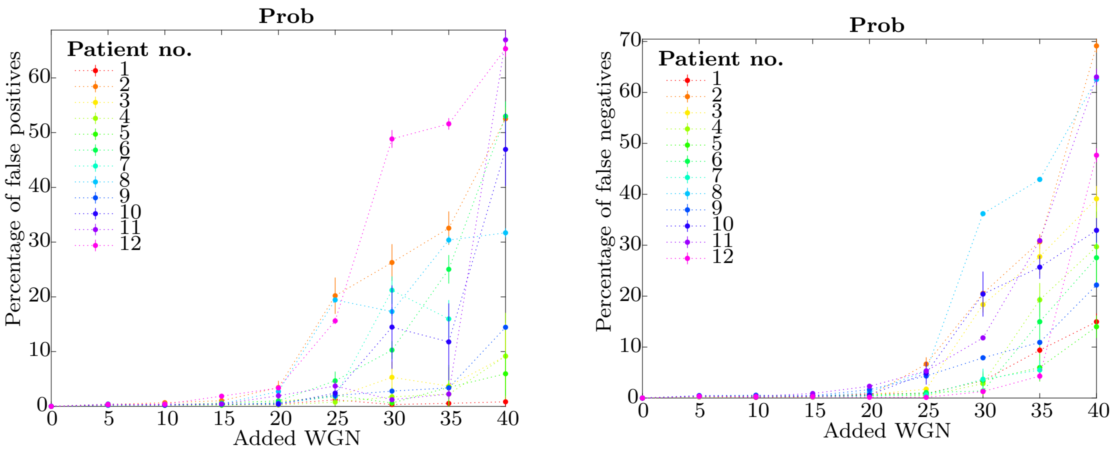


Figure 3.36: False positives (top) and false negatives (bottom) in Tensor_Prob tractograms versus noise calculated using a nearest point search method with a threshold of 3 SD.

3.13 Chapter Summary

The results presented in this chapter support the expected relationship between noise and tractogram integrity, namely, that increased noise results in progressive deterioration of reconstructed tracts.

Qualitative analysis of the patient tractograms with added noise revealed a progressive deterioration of major white matter structures for added WGN levels beyond a power of 20 for the deterministic SD_STREAM algorithm, and 25 for the probabilistic algorithms iFOD2 and Tensor_Prob. Additionally, the corpus callosum was reconstructed with the highest fidelity with added noise for all three tractography algorithms, followed by the medial lemnisci and pontine fibres.

Analysis of the streamline length distributions for tractograms with added noise demonstrated that premature tract termination was the primary effect induced by added noise in all three algorithms.

A method of assessing the level of resemblance between an unmodified tractogram and tractogram with added WGN was devised. A constant resemblance value was obtained for tractograms with added noise up to WGN 20 for the probabilistic algorithms, while a continuous decrease in resemblance was observed using SD_STREAM. The RMS difference in tract-weighted maps derived from the tractograms revealed a similar trend.

The differing results for each of the three tractography algorithms were expounded upon, with iFOD2 standing out as the optimal algorithm for repeatably reproducing a tractogram with the most intricacy, and the least false positive or false negative streamlines.

Plotting each metric used to compare the accuracy of the noisy tractograms with respect to the unmodified tractogram against the computed SNR (using the difference method), revealed that for all three algorithms, dMRI scans with an SNR of less than about 4 should always be considered unreliable. An SNR above 8 is ideal for the tractography to produce reliable results.

CHAPTER 4

The effect of motion on tractography performance

4.1 Addition of artificial patient motion

A number of forms of patient motion may be present in a patient diffusion magnetic resonance imaging (dMRI) scan, including rigid body motion, elastic motion and flow (Zaitsev et al., 2015). Rigid body motion encompasses translations in one or multiple dimensions, and rotation about some axis. Involuntary head motion requires six degrees of freedom, to account for translations and rotations in x, y and z. Elastic motion includes stretching, compression and shearing along three axes, which, with rigid motion, can be represented by 12 degrees of freedom. As patient motion can occur at any point in the scanning process, this can lead to misregistration of diffusion signals between slices or volumes. Forms of flow include blood and cerebrospinal fluid (CSF) flow. In their 2017 study, Rydhog et al. (2017) demonstrated that diffusion MRI is sensitive to perfusion, particularly at low b-values. While b-values of the order of 1000 s/mm^2, which are not strongly impacted by perfusion, are typical for diffusion MRI sequences, at least one b=0 volume is typically used, which is susceptible to perfusion.

Here we consider two basic forms of rigid motion which may be present in a dMRI scan: sudden translations or rotations between volumes. Increments of each of these forms of motion were artificially added to each patient's original dMRI scan. Other forms of motion that may be present in a scan include translations between slices, which may represent random or smooth movements.

The effects of these two forms of motion are considered in terms of change relative to the unmodified tractogram. There is therefore an inherent assumption that any motion artefacts present in the unmanipulated data have no effect on the resulting analysis. Additionally, the effects of added motion on the tractogram should be considered independently for each patient, as variations in scan paramters such as the number of slices, number of diffusion directions, diffusion gradient order, etc. may influence the results.

To investigate the effect of translation of the patient between volumes, the second half of each scan was translated in one direction in the axial plane by increments of 1 mm up to a maximum of 5 mm using the Matlab function imtranslate. For scans which had 2 volumes, artificial translation was added to the second volume. For scans which had 4 volumes, artificial translation was added to the third and fourth volumes. The averaging of these volumes prior to tractography resulted in half of the scan being affected by motion.

To assess the impact of rotation of the patient in a clockwise direction between volumes, the second half of each scan was rotated by increments of 1 degree up to a maximum of 10 degrees using the Matlab function imrotate. The rotations were applied to the second volume for 2-volume scans, and the third and fourth volumes for 4-volume scans.

The maximum level of added motion in each case was chosen to be representative of the expected scale of movement that would occur in a typical dMRI scan, limited by the amount of added motion that resulted in complete deterioration of the tractogram.

4.2 Measuring patient movement

An automatic three-dimensional (3D) registration technique available in Matlab was used to quantify the amount of translation and rotation present in each dMRI. For each dMRI scan with added translation or rotation, the command dwi2mask was used to generate a mask of the brain. For each increment of added translation or rotation, the mask of the corresponding dMRI was registered to the unmodified dMRI for that patient. This was achieved using the Matlab function imregtform, which estimates the geometric transformation that aligns a moving with a fixed image. An affine transformation matrix was estimated for each unmodified-modified mask pair, of the form:

$$T = \begin{bmatrix} x_{ii} & x_{ij} & x_{ik} & 0 \\ x_{ji} & x_{jj} & x_{jk} & 0 \\ x_{ki} & x_{kj} & x_{kk} & 0 \\ x_{li} & x_{lj} & x_{lk} & 1 \end{bmatrix} \tag{4.1}$$

The function imregtform with an affine transform type can estimate differences in the masks of the form of translation, rotation, scaling and/or shear. The amount of translation required to align the modified mask with the unmodified mask can be extracted by taking the magnitude of the vector defined by the first three elements of the last row of the transformation matrix (i.e. $[x_{li}, x_{lj}, x_{lk}]$). The resulting translation estimates are shown in Figure 4.1 against the known added translations for each scan.

The scaling factor can be extracted from the transformation matrix by taking the length of the first three column vectors:

$$s_x = ||\langle x_{ii}, x_{ij}, x_{ik} \rangle|| \tag{4.2}$$

$$s_y = ||\langle x_{ji}, x_{jj}, x_{jk} \rangle|| \tag{4.3}$$

$$x_z = ||\langle x_{ki}, x_{kj}, x_{kk} \rangle|| \tag{4.4}$$

$$\vec{s} = \langle s_x, s_y, s_z \rangle \tag{4.5}$$

The matrix defined by elements in the first three rows and columns can then be used to extract the amount of rotation and shear. Given that there is no expected contribution from shear, the amount of rotation between the two masks can be straightforwardly extracted from the rotation matrix, R.

$$R = \begin{bmatrix} x_{ii}/s_x & x_{ij}/s_x & x_{ik}/s_x \\ x_{ji}/s_y & x_{jj}/s_y & x_{jk}/s_y \\ x_{ki}/s_z & x_{kj}/s_z & x_{kk}/s_z \end{bmatrix} \tag{4.6}$$

The following equations can then be used to obtain the Euler angles.

$$\theta_x = atan2(x_{kj}/s_y, x_{kk}/s_z) \tag{4.7}$$

$$\theta_y = atan2(-x_{ki}/s_z, \sqrt{(x_{kj}/s_z)^2 + (x_{kk}/s_z)^2}) \tag{4.8}$$

$$\theta_z = atan2(x_{ji}/s_y, x_{ii}/s_x) \tag{4.9}$$

These equations are further simplified by assuming that there is no contribution from scaling, given that each scan was only either translated or rotated. Setting all s terms to 1 and computing the amount of rotation about the z axis provides an estimate of the added rotation. The estimated rotation is shown in Figure 4.2 against the known added rotation.

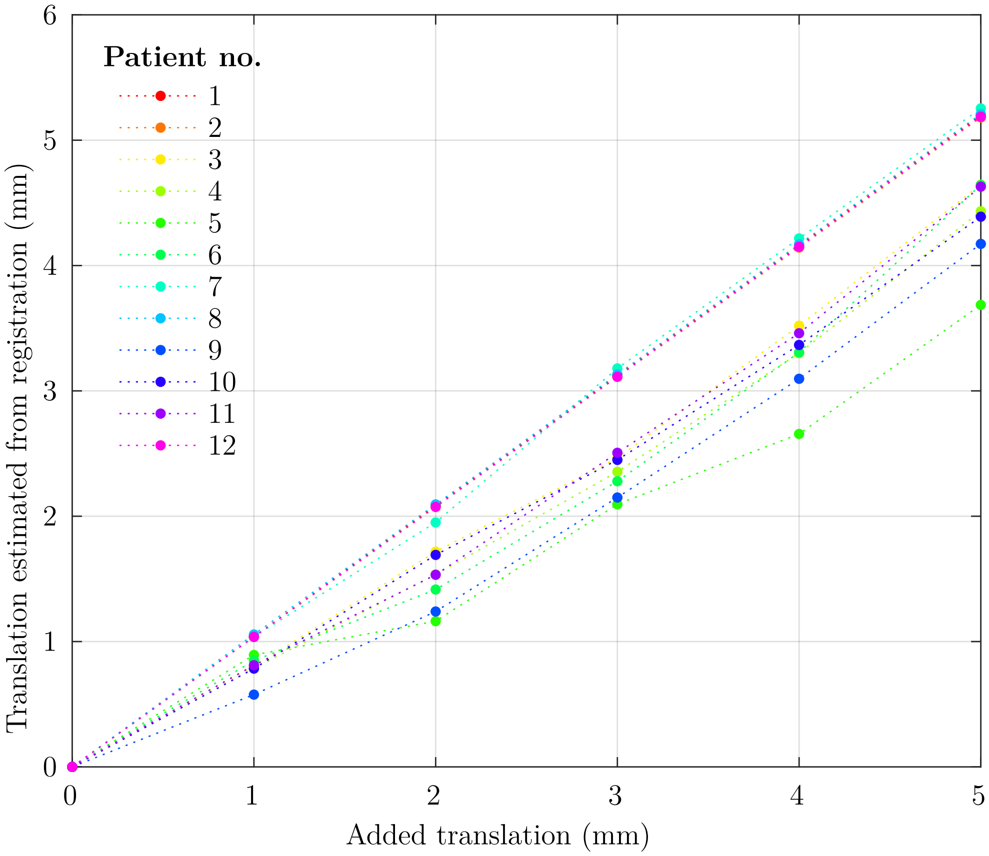


Figure 4.1: Plot of the translation obtained through registration of the unmodified and translated volumes of the dMRI scan for each patient versus the known added translation.

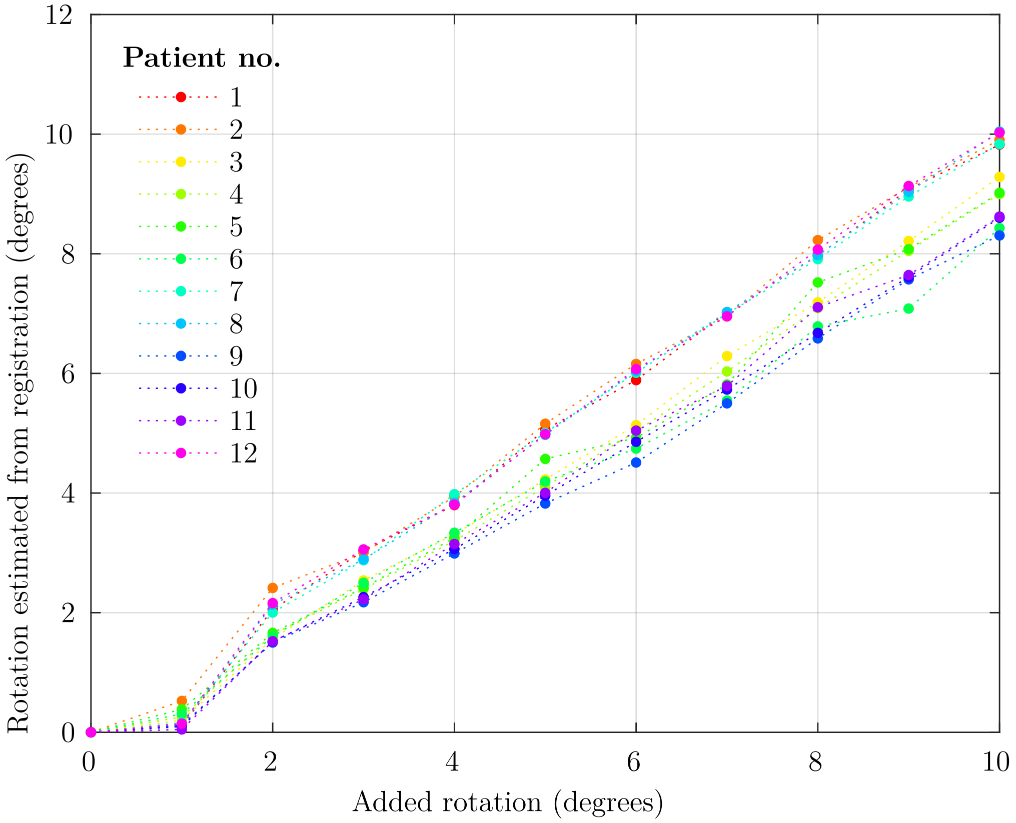


Figure 4.2: Plot of the rotation obtained through registration of the unmodified and translated volumes of the dMRI scan for each patient versus the known added rotation.

The plots of measured versus true added translation and rotation demonstrate that this technique is imperfect, even in the idealised case where there are no other parameters or external factors that differ between the two scans and the motion is in only one direction. From the plot of added translation, it is evident that the amount of translation is underestimated for the majority (10/12) of patients, and slightly overestimated for the others (patients 8 and 12). This reflects the limitations of the registration technique for determining the correct transformation that will convert the original into the modified scan. Similar observations can be made for the plot of added rotations. In this case there is also a noticeable dip at 1 degree of added rotation, suggesting that the registration technique is not sensitive to such small scale rotations, and may falsely attribute it to another form of motion.

4.3 Qualitative results

The scoring system described in section 3.5 was used to assess the quality of each of the tractograms without and with added translation and rotation. Figure 4.3 shows the change in the tractogram score (as a percentage) with added translation in mm for each patient using the iFOD2 algorithm. For the majority of patients analysed, the addition of translation up to 5 mm has no effect on the tractogram scores. For a small subset of patients (3, 5, 7 and 12), the structures begin to visibly deteriorate within 5 millimetres of added translation. However, it is not clear whether this indicates true deterioration of the tractograms or an artefact of expectation bias.

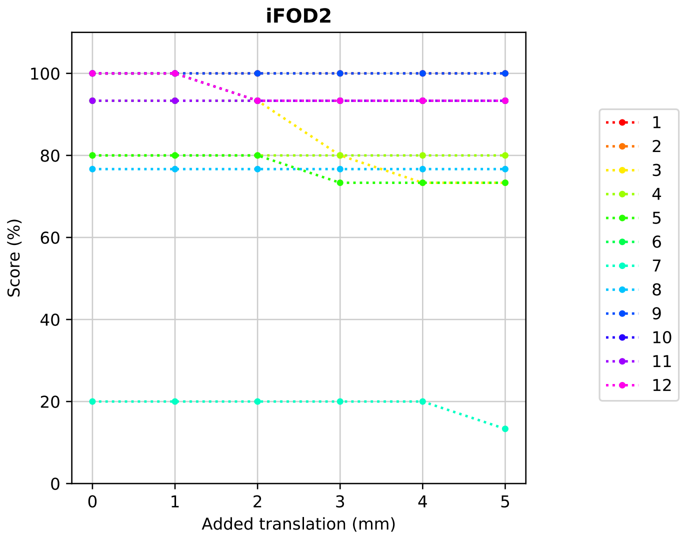


Figure 4.3: Plots showing the change in the qualitative tractogram score as translation in millimetres is added for each patient using iFOD2.

For tractograms with added rotation, the tractogram scores were similarly unaffected by the variable under consideration. A plot of the tractogram score versus the degrees of added rotation has been omitted, as no significant changes were observed between tractograms for any of the patients.

While all major structures initially present in the tractograms are retained even up to the maximum translations or rotations considered, it is clear that the physical locations of these structures have shifted. The approach taken to artificially add translation and rotation does not provide a fully realistic simulation of the effects of these forms of motion on a dMRI scan. By modifying one half of the scan volumes, the diffusion signal on average is just displaced, which evidently has very little to no effect on the ability of the tractography algorithm to delineate pathways, but does have an effect on the positions of the structures in the resulting tractogram. This is exemplified by the tractograms shown in figures 4.4 and 4.5, which demonstrate no observable change in the integrity of the tractogram, but show that a shift in position has occurred as a result of adding translation or rotation, respectively.

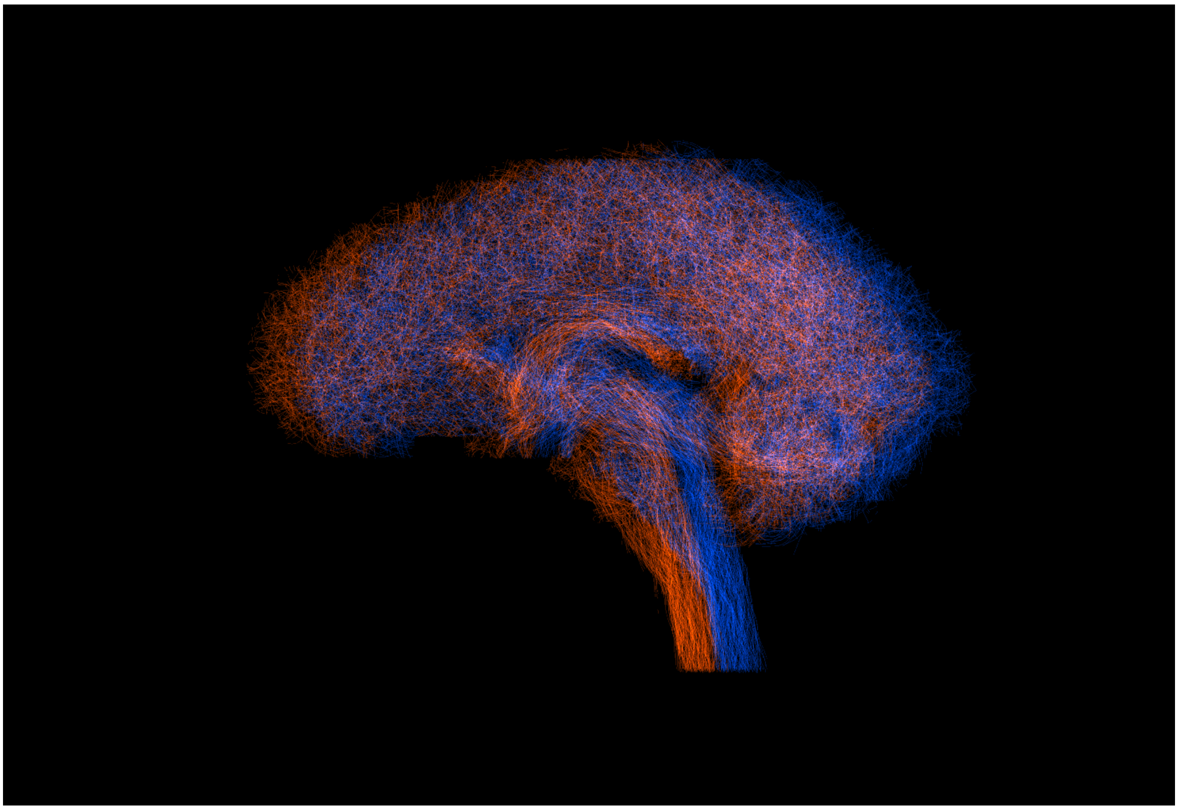


Figure 4.4: Screen capture of two overlaid iFOD2 tractograms produced from the patient 1 dMRI with and without added translation. The sagittal view of the initial tractogram is shown in blue, and the same view of the tractogram with 5 mm of added translation is shown in red.

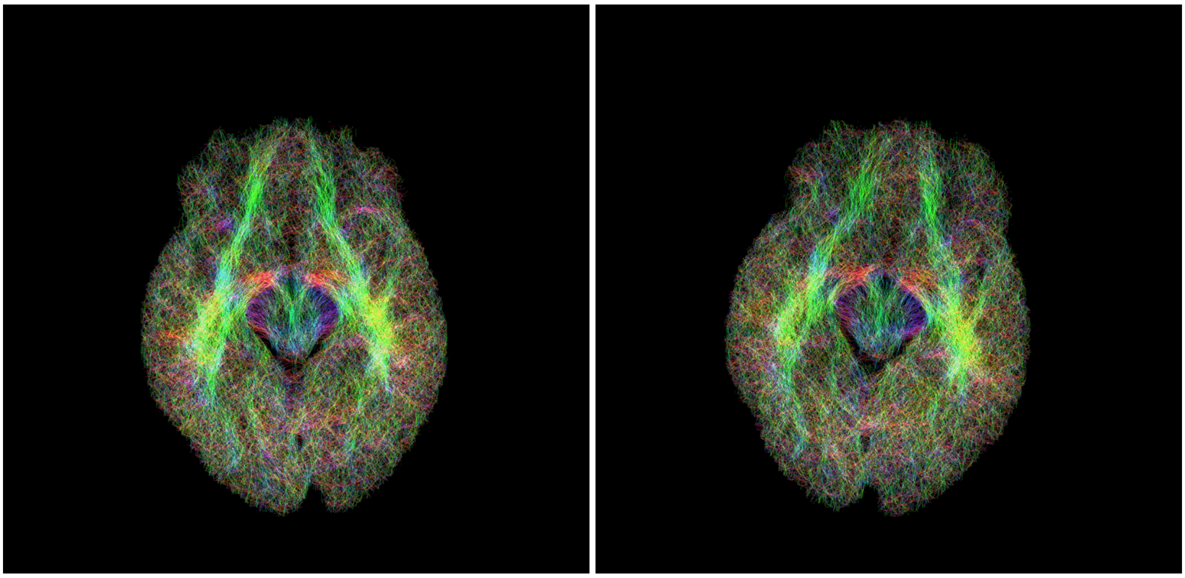


Figure 4.5: Screen captures of iFOD2 tractograms produced from the patient 1 dMRI with and without added rotation. The image on the left shows the axial view of the patient 1 tractogram with no added rotation. The same view of the tractogram is shown on the right, with 10 degrees of added rotation.

4.4 Translation versus streamline length

4.4.1 iFOD2

For each dMRI scan with added translation, a 5k tractogram seeded at the base of the brain stem was generated using iFOD2. For each tractogram, a histogram of the streamline lengths comprising the tract was produced. The change in these histograms with added translation provides an insight into the effect of translation in one direction on the tractography performance. An example of this is shown in Figure 4.6, which depicts the change in the distribution of streamline lengths with the addition of translation for patient 1, for tractograms generated using iFOD2. A progressive decrease in streamline lengths is observed as translation is added for Patient 1. The distribution of streamline lengths in the unmodified tractogram contains a distinct peak centred at around 40 mm. The counts in this peak are largely preserved across all tractograms. Two additional peaks with broader distributions appear to exist in the initial tractogram. The distribution centred around 60 mm is largely retained, but the counts at higher lengths become redistributed to lower length bins, causing a peak to appear below 20 mm. This suggests that premature termination of streamlines primarily affects the longest streamlines in the distribution.

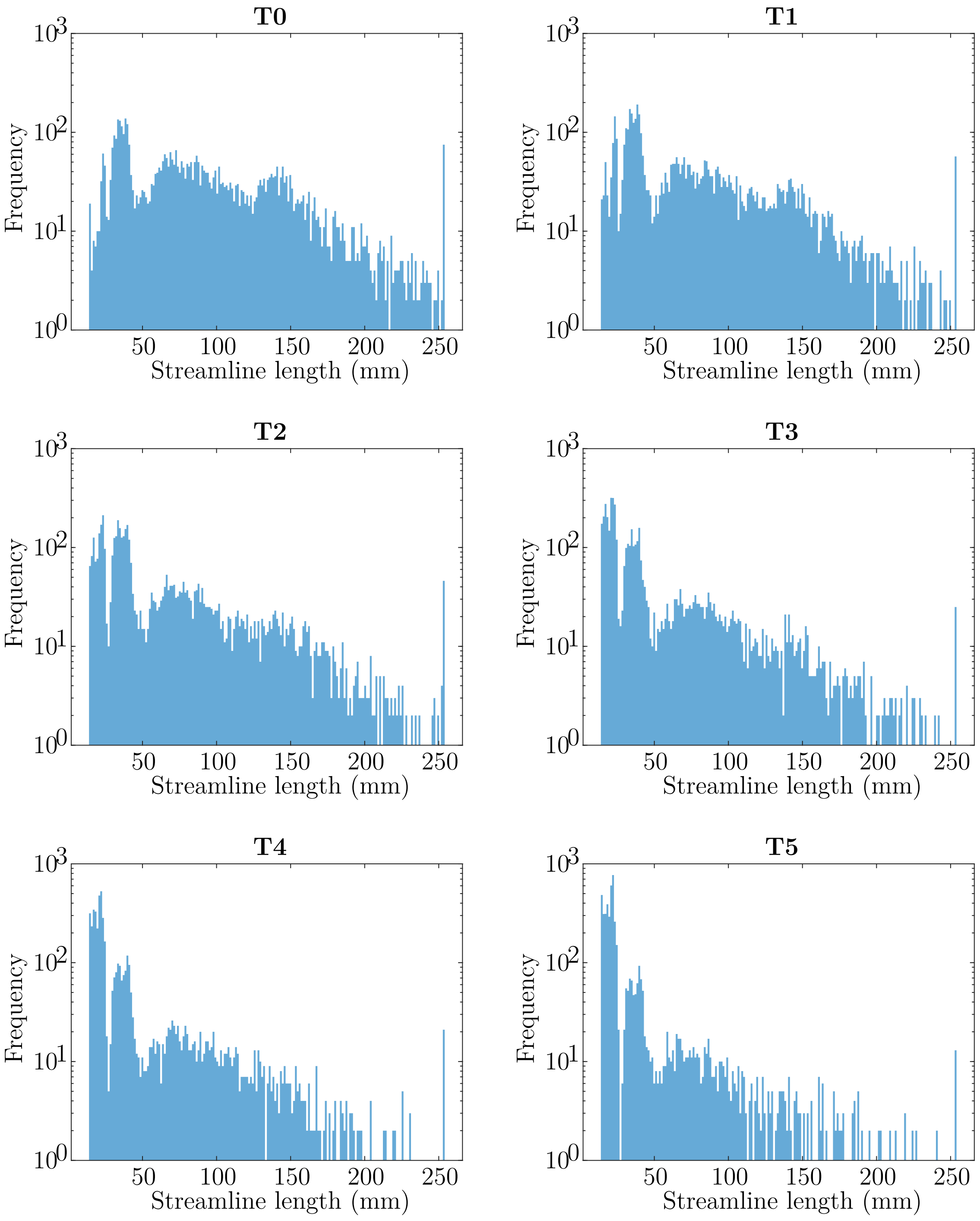


Figure 4.6: Histograms of streamline length for tractograms generated using MRtrix's iFOD2 algorithm with varied levels of added translation. These results are derived from the dMRI scan taken of patient 1. Each tractogram consists of 5000 streamlines seeded at a point at the base of the brain stem, such that the integral of each histogram is equivalent. The histogram bin size was set to 1, equal to the step size used by the algorithm. The bins therefore represent all possible streamline lengths that the algorithm can reproduce. The labels T0 to T5 indicate the amount of translation added in millimetres (i.e. Tx => added translation of x mm). As additional translation is added, the histogram becomes more skewed to the left, reflecting the tendency of the iFOD2 algorithm to prematurely terminate streamlines when the input dMRI contains a larger amount of translation.

Figure 4.7 shows the mean streamline length and standard deviation for each 5k point-seeded tractogram generated for patient 1. A decline in mean streamline length and reduction in spread is observed with added translation.

Figure 4.8 depicts the normalised mean streamline length for the whole-brain tractogram generated for each patient for each increment of added translation. The error bars represent the standard error in the mean streamline length for each tractogram. Different classes of patients appear to emerge from the data. For one subset of patients (patients 5, 6, 9, 10, and 11) the addition of translation results in an increase in false positives, with the mean streamline length increasing by up to 12% with 5 mm added translation. The mean streamline length also follows the same increasing trend for added translation up to 2 mm for patients 3 and 4. Beyond this, the mean streamline length remains roughly level. For the remaining patients (patients 1, 2, 7, 8 and 12), the mean streamline length is consistent to within 2% for added translations up to 5 mm.

The results obtained using the tractograms of the full brain differ from those for the seeded tractograms. The streamline lengths in the whole-brain tractograms appear to become elongated or remain the same with added translation, while the seeded tractograms show a steady decline in streamline length with increased translation. This may indicate that the response of different white matter structures to added motion may vary. Alternatively, it may indicate that the algorithm fails to delineate intricate streamlines more frequently with added motion (i.e. shorter streamlines are lost entirely), so only structures that are prominent in the scan are reconstructed, causing the mean streamline length to increase. In the seeded tractogram this would not be the case, as the algorithm is forced to repeatedly trace the required number of paths from the same starting point, so while it may take more attempts to delineate the more intricate paths (i.e. the algorithm has a low probability of delineating the path successfully), it will still do so.

4.4.2 SD_STREAM

For each tractogram generated using SD_STREAM for patient 1 with added translation, a histogram of streamline lengths was produced, displayed in Figure 4.9. The histogram of the unmodified tractogram contains a band of counts between 32 mm and 38 mm peaked at 33 mm and 37 mm. This band is retained with 1 mm of added translation. However, the addition of 1 mm translation results in the loss of a smaller peak centred at about 22 mm. As further noise is added the peaks begin to shift towards lower lengths, demonstrating the tendency towards early streamline termination. The production of false positives is also evident with an added translation of 3 mm, with a significant number of streamlines having lengths of up to 45 mm.

The mean streamline length and standard deviation for 5k seeded tractograms generated for patient 1 are shown in figure 4.10. The mean and spread are consistent with 1 mm of added translation. Beyond this, the spread increases significantly and a large portion of the streamlines delineated are shorter. With an added translation of 5 mm, all reproduced streamlines have lengths less than 25 mm.

The normalised mean streamline length and error in the mean for each tractogram generated with added translation are plotted in figure 4.11. In general the mean streamline length is conserved for added translations up to 5 mm. However, for a small subset of patients (patients 3, 10, and 11), the mean increases by about 3% for added translations of 2 mm through 5 mm.

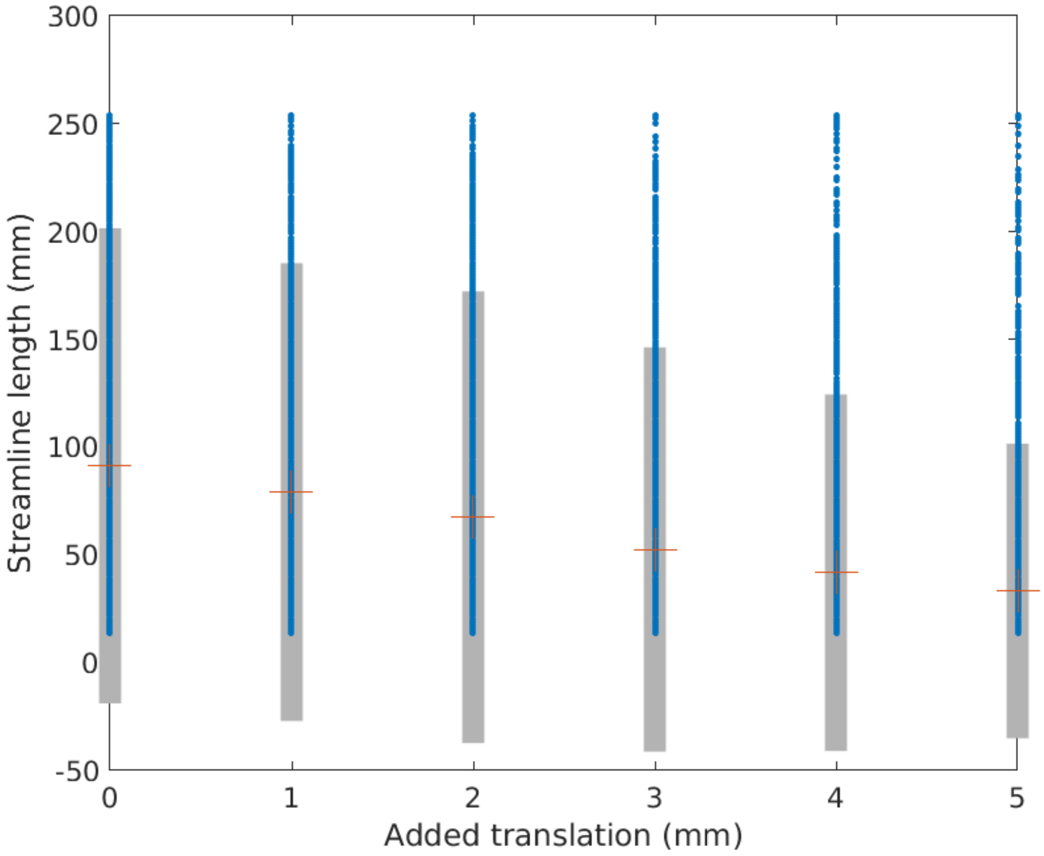


Figure 4.7: Plot illustrating the change in the overall distribution of streamline lengths for 5k point-seeded tractograms generated with iFOD2 as translation is added to the input dMRI. The individual streamline lengths, and the mean and 2 standard deviation (SD) of each distribution are shown in blue, red and grey, respectively.

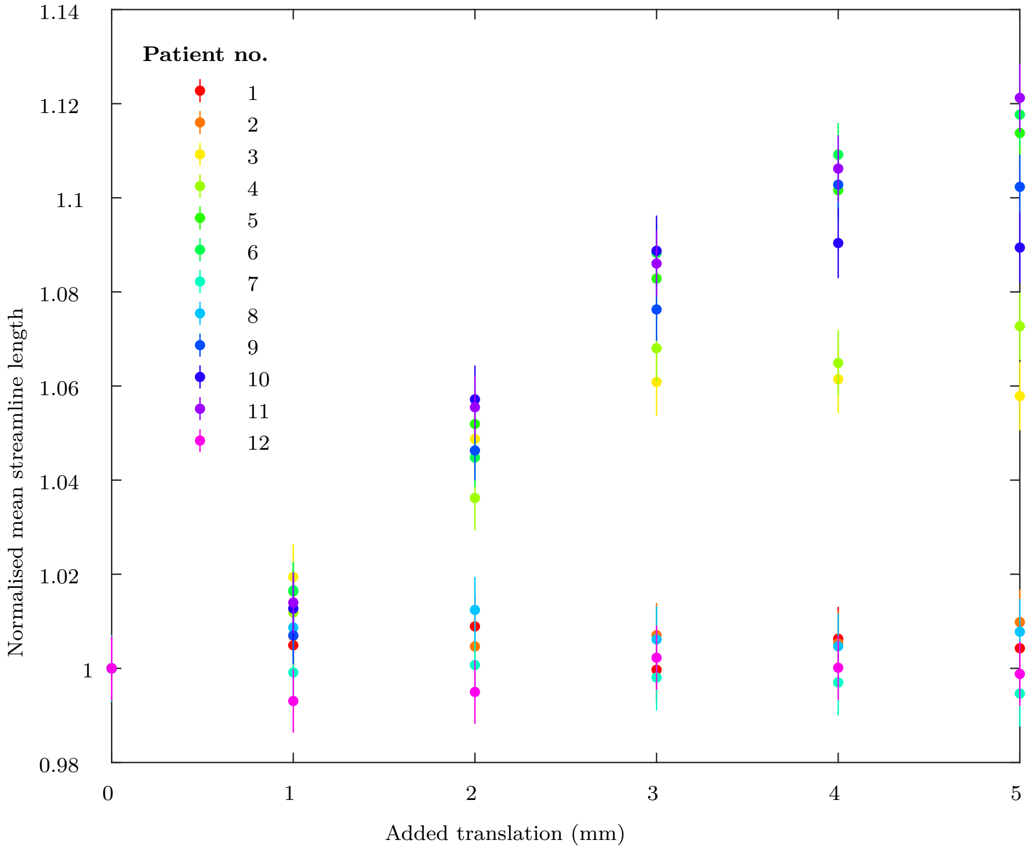


Figure 4.8: Plot of the normalised mean streamline length versus added translation for all whole-brain tractograms generated using iFOD2.

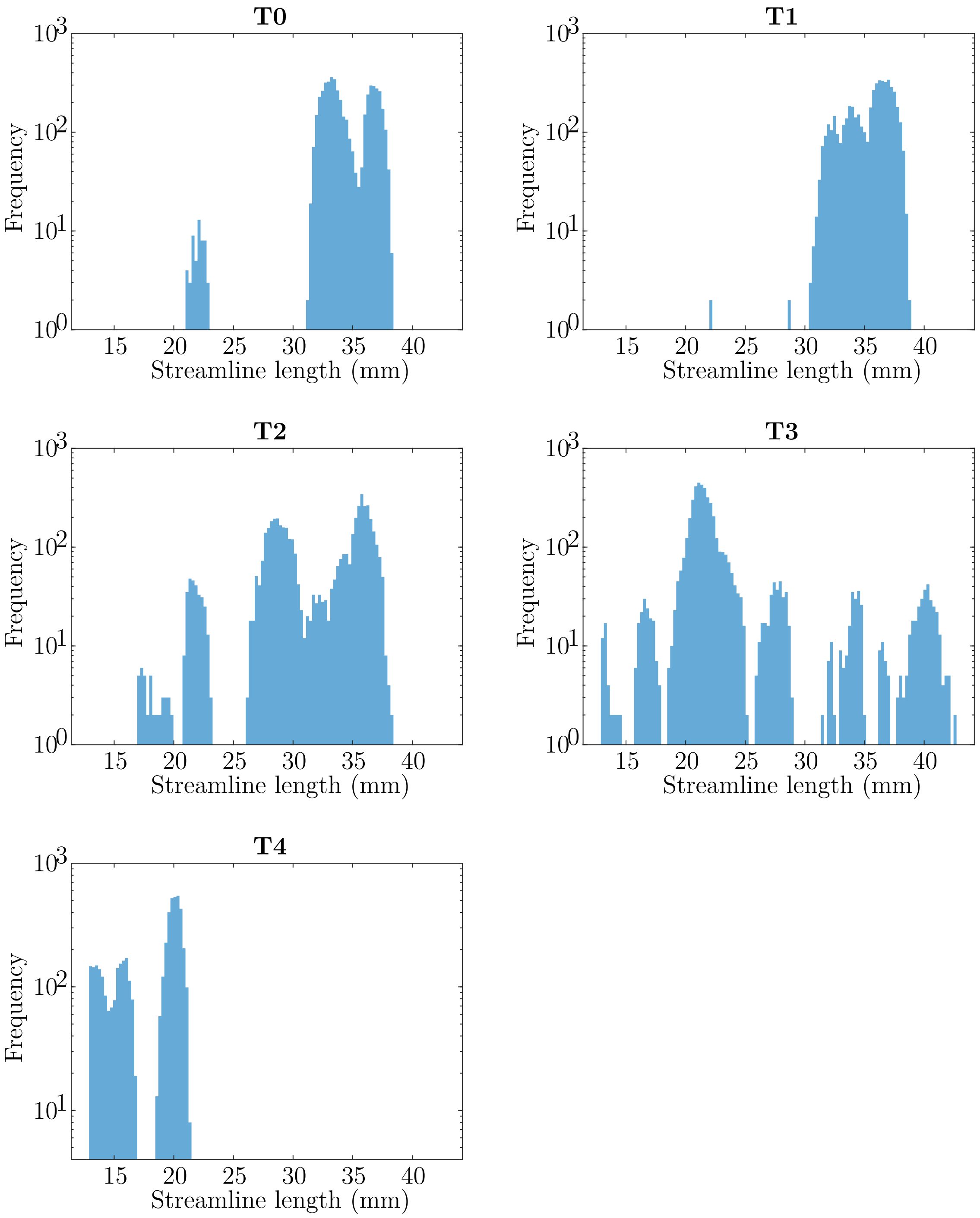


Figure 4.9: Histograms of streamline length for tractograms generated using MRtrix's SD_STREAM algorithm with varied levels of added translation. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered again consists of 5000 streamlines seeded at a point at the base of the brain stem. The SD_STREAM algorithm uses a finer step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are shown on a logarithmic scale to better visualise regions of low counts in the distribution. The labels T0 to T4 represent the amount of added translation in millimetres. T5 is omitted as the tractography failed with 5 mm of added translation.

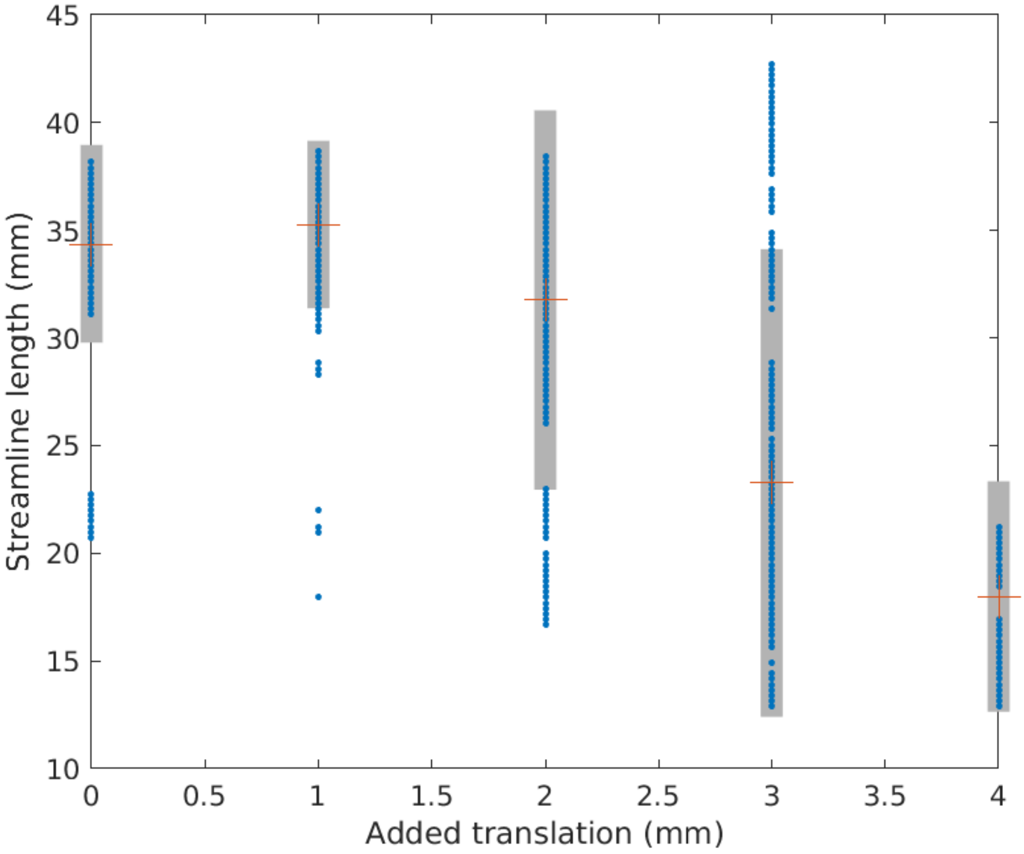


Figure 4.10: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with SD_STREAM as translation is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

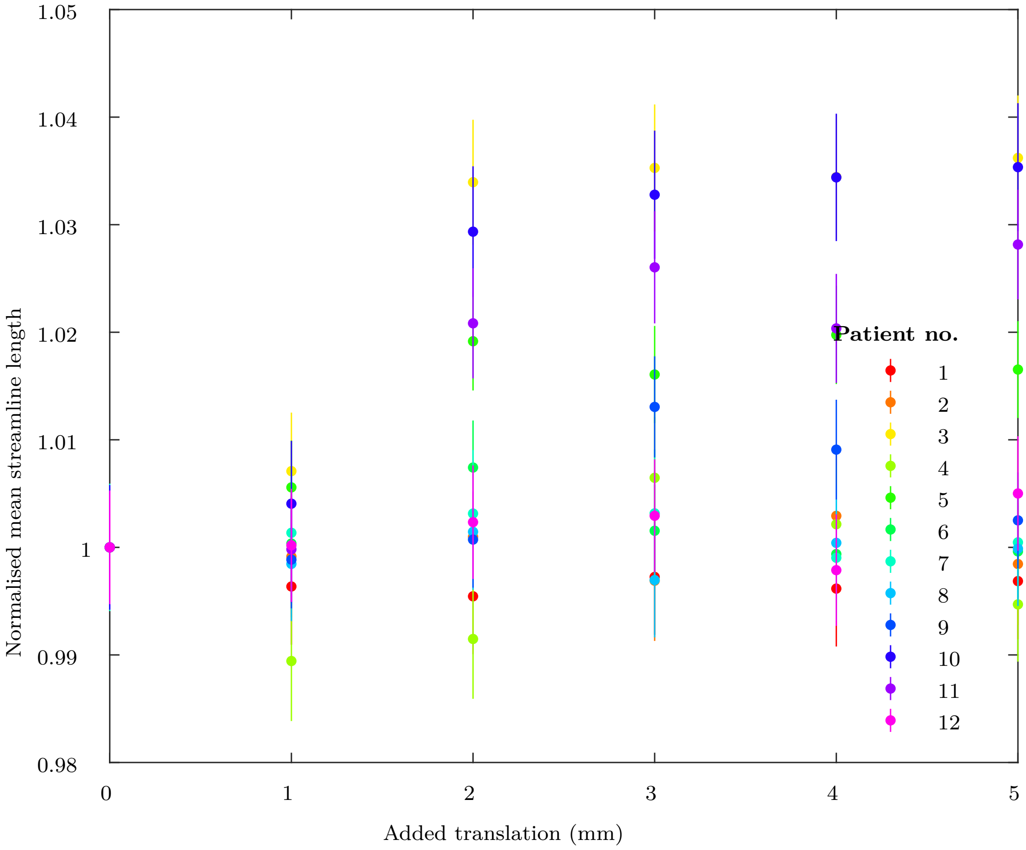


Figure 4.11: Plot of the normalised mean streamline length versus added translation for all whole-brain tractograms generated using SD_STREAM.

4.4.3 Tensor_Prob

Histograms of the streamline lengths for 5k tractograms seeded at the base of the brain stem with added translation generated using Tensor_Prob for patient 1 were generated. The results for patient 1 are shown in figure 4.12. The initial distribution is comprised of a large peak centred at about 25 mm, and progressive lower peaks at greater lengths. The two most prominent peaks in the distribution are mostly conserved with an added translation of 1 mm. However, additional streamlines are delineated with higher lengths than in the original distribution. At high levels of added translation the distribution collapses to lengths under 50 mm. A tractogram seeded at the base of the brain stem could not be produced from the dMRI scans with an added translation of 4 mm and 5 mm. This is likely due to the displacement of the location of the base of the brain stem through averaging of the shifted and un-shifted volumes of the scan.

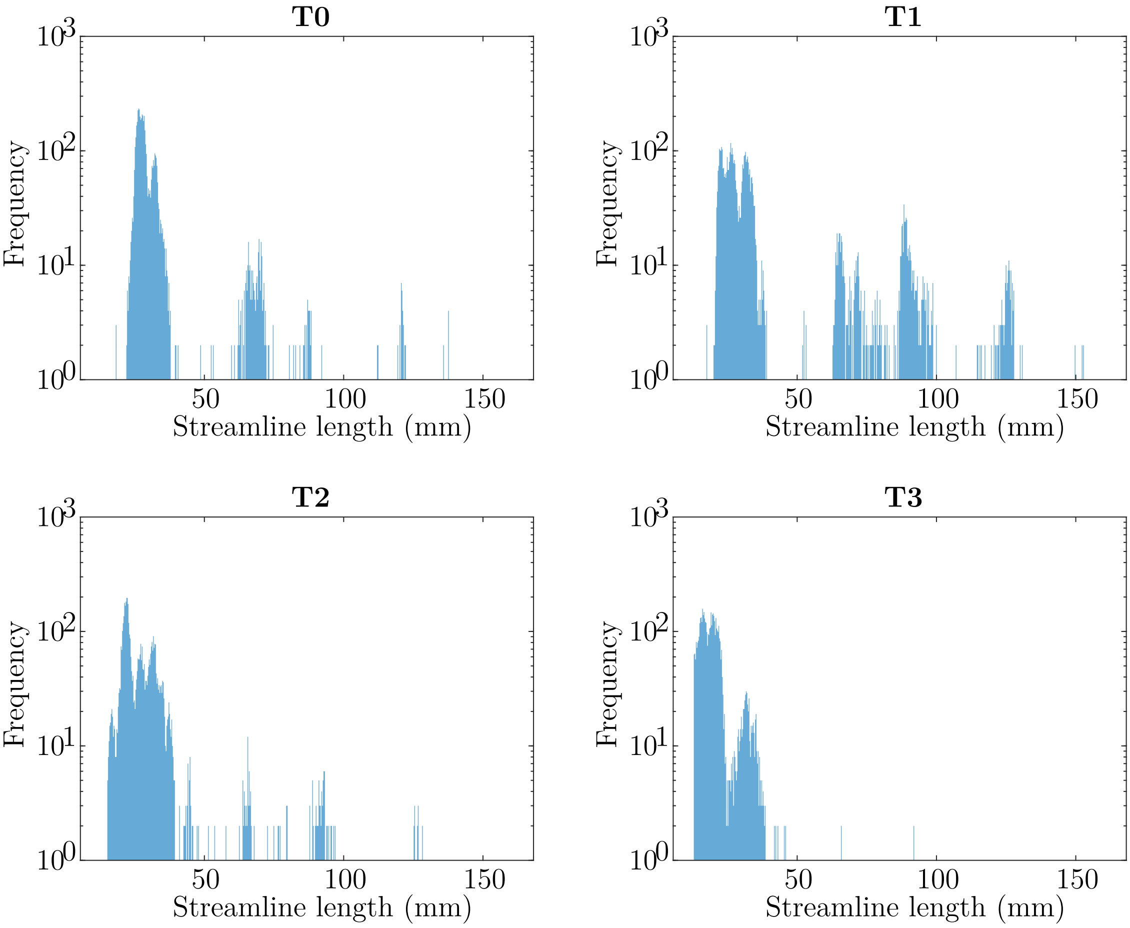


Figure 4.12: Histograms of streamline length for tractograms generated using MRtrix's Tensor_Prob algorithm with varied levels of added translation. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered consists of 5000 streamlines seeded at a point at the base of the brain stem. The Tensor_Prob algorithm uses a step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are shown on a logarithmic scale to better visualise regions of low counts in the distribution. The labels T0 to T3 represent the amount of translation added in millimetres. The tractography was unsuccessful for input scans with more than 3 mm of added translation.

The mean streamline length remains consistent to within 1 standard deviation for trac-

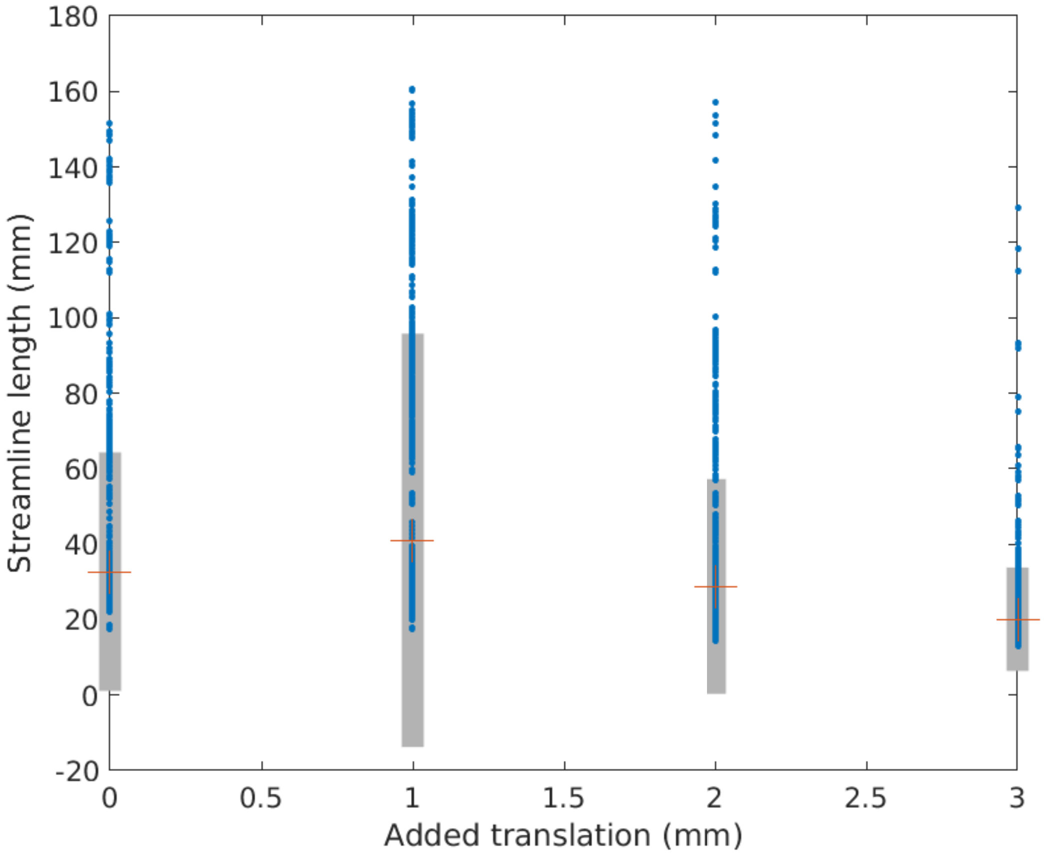


Figure 4.13: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with Tensor_Prob as translation is added to the input dMRI for patient 1. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

tograms with added translation up to 3 mm for patient 1. This is evident from figure 4.13, which shows the mean sitting between about 20 mm and 45 mm. The spread of the streamline lengths initially increases, then declines as more streamlines are prematurely terminated with an added translation of 3 mm.

The consistency in the mean streamline length is also evident in figure 4.14 for patients 1, 2, 7, 8, 9, 11 and 12. However, for the remaining patients a progressive decline in the mean streamline length is observed. For patients 4 and 10, the magnitude of this decrease is about 14% for an added translation of 5 mm. For patients 3, 5, 6 and 11, the mean streamline length decreases by about 5% over the same range.

4.4.4 Comparison of translation effects on iFOD2, SD_STREAM and Tensor_Prob

The response of tractography to the addition of translation differs for each algorithm considered. For iFOD2, the mean streamline length for a whole-brain tractogram either remains consistent for added translation up to 5 mm, or increases roughly linearly then begins to

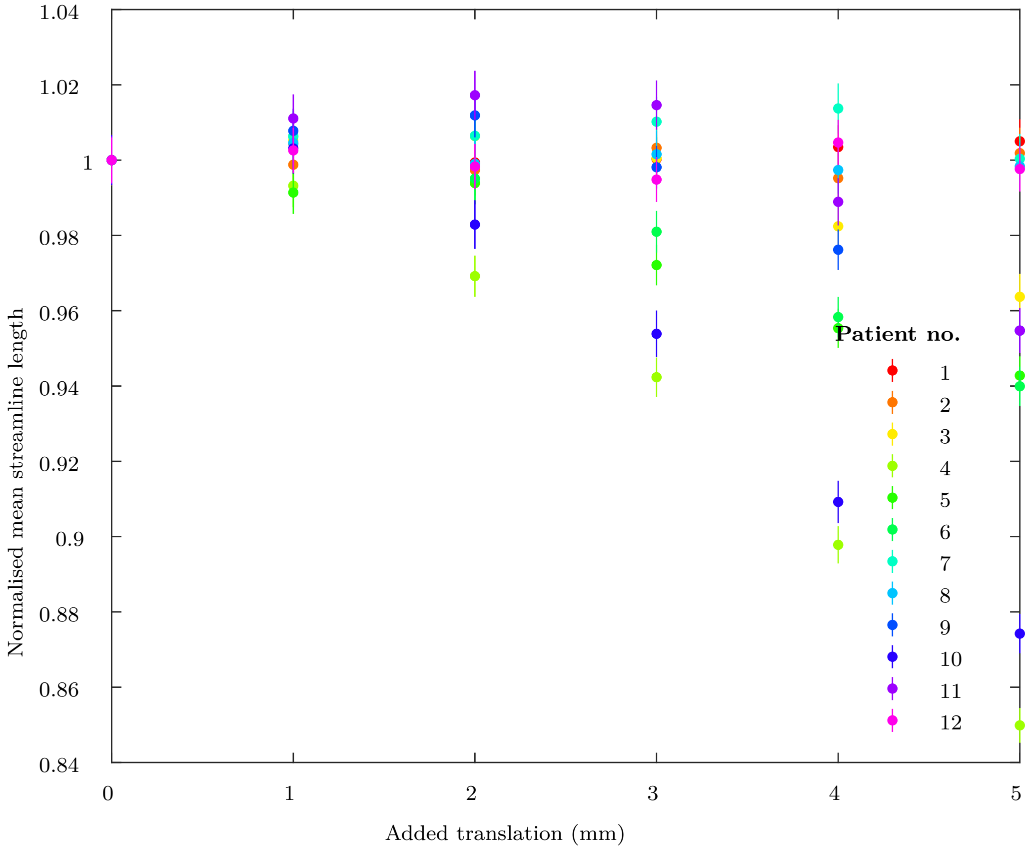


Figure 4.14: Plot of the normalised mean streamline length versus added translation for all whole-brain tractograms generated using Tensor_Prob.

plateau, reaching 10% with 5 mm of added translation. For SD_STREAM a similar trend is observed. However, rather than a gradual increase in the mean streamline length for a subset of the patients, a step change is observed at an added translation of 2 mm. Additionally, the magnitude of the increase in mean streamline length remains under 4% for added translations up to 5 mm. For Tensor_Prob, again a portion of the patients have tractograms with consistent mean streamline lengths (within about 2%). However, there are two additional subsets of patients which show a reduction in mean streamline length with added translation, of up to about 5% for one group, and 14% for the other. In all cases, the trend towards higher or lower mean streamline lengths does not appear to be correlated with the mean streamline length of the initial (unmodified) tractogram, which implies this initial condition has no causal effect on the direction of the trend.

These results suggest that the algorithms iFOD2 and SD_STREAM tend to favour the production of false positives when there is translatory motion present in the input scan, whereas Tensor_Prob responds to increased translation by producing more false negatives.

Different trends which arise for different patient groups may be due to a gross difference in scan parameters between those groups. For instance, patients 1, 2, 7, 8 and 12 which show consistency in mean streamline length up to 5 mm added translation with iFOD2, have a repetition time of 4200, an echo time of 105 and 2 B_0 images. Of these, the echo time is least likely to have led to two diverging trends in the data. Despite the scans for patients 1, 2, 7, 8 and 12 having equal echo times, the absolute value (105) falls between that of patients 3, 4, 5, 10 and 11 (83-93), and patients 6 and 9 (114), so one would expect to see an additional two trend groups represented in the plot, which is not the case. The two trends evident in the plot may, however, correspond to two patient groups with differing repetition times and number of B_0 images. For patients 3, 4, 5, 6, 9, 10 and 11, the repetition time falls within the range 3000 to 3300, and the number of B_0 images is 4.

In practice, the repetition time is likely to have a significant impact on the amount of patient movement present in a dMRI scan. If there is greater time between successive pulse sequences applied to the same slice (i.e. a higher TR), there is more time for involuntary movements to occur between acquisition of signals. However, in this case motion was artificially added to one volume of the scan, so each input dMRI should contain the same amount of movement between slices as the original scan.

Similarly, a larger number of B_0 images could increase sensitivity to patient motion by increasing the scan acquisition time. However, the amount of inter-slice patient movement should remain constant in each scan. The addition of artificial motion affects half of the B_0 images, and should not depend on whether there were 2 or 4 B_0 images in the original scan.

The scan parameters do not appear to align with trends observed for SD_STREAM. However, Tensor_Prob shows similar groupings of patients. Scan parameters which may have an effect include the number of signals averaged, the repetition and echo times, and the number of B_0 images.

Averaging of the first volume with different translations of the second volume may result in different diffusion signals aligning. Therefore, translation of a volume could result in certain streamlines being lost and others (false positives) produced. It is therefore conceivable that added translation could result in an increase or decrease in mean streamline length.

It is unclear why iFOD2 and SD_STREAM appear to favour the production of false positives as translation is added to one volume of the scan, whereas Tensor_Prob favours false negatives. This behaviour does not appear to agree with the observed trend towards false negatives for point-seeded tractograms.

4.5 The effect of rotation on streamline length

4.5.1 iFOD2

For each 5k point-seeded tractogram generated for each increment of added clockwise rotation in the axial plane, a histogram was produced of the streamline lengths. The histograms for patient 1 iFOD2 tractograms with added rotation are shown in Figure 4.15. The general form of the distribution remains consistent for all added rotations, up to 10 degrees. However, features in the distribution become gradually blurred as rotation is added.

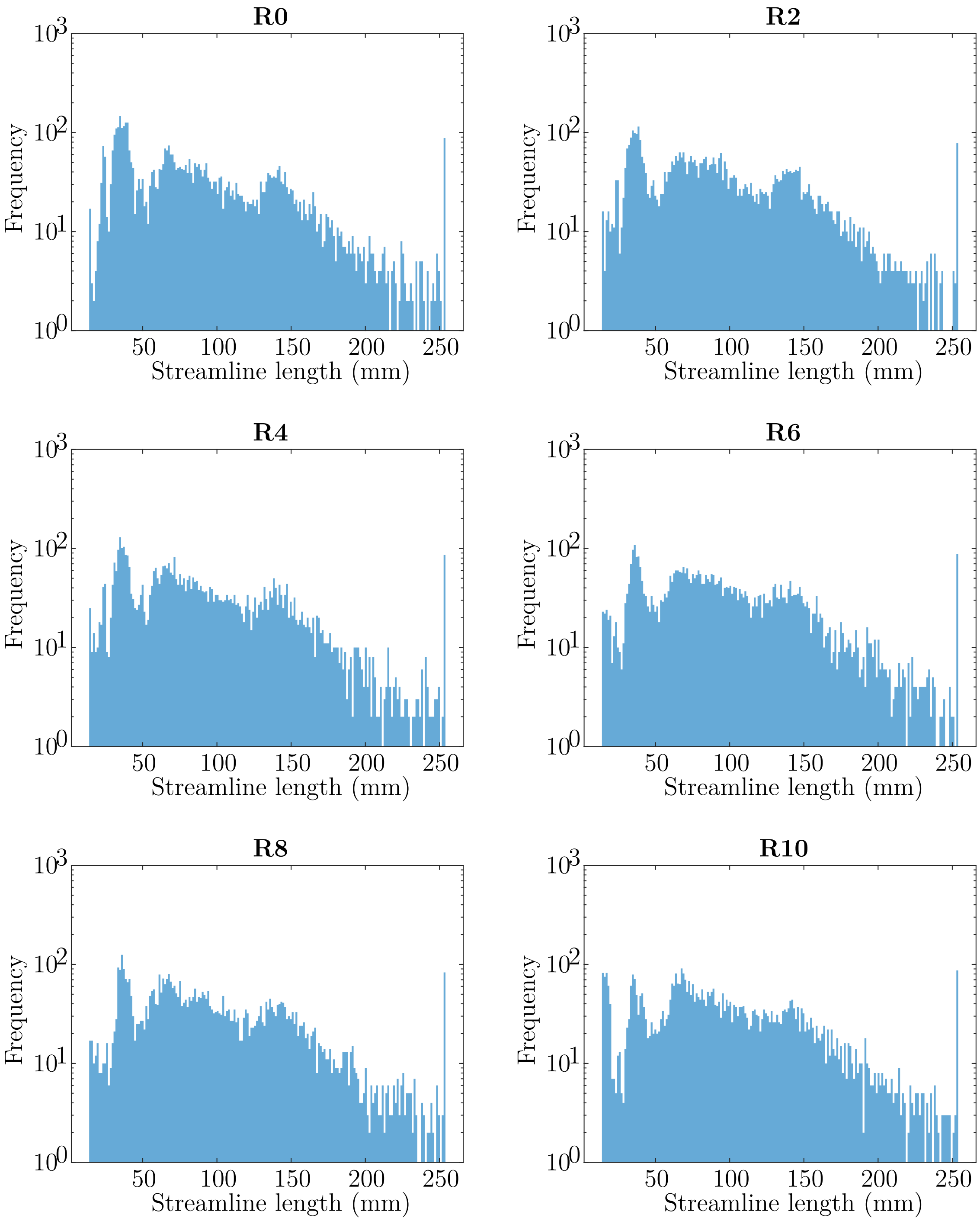


Figure 4.15: Histograms of streamline length for tractograms generated using MRtrix's iFOD2 algorithm with varied levels of added rotation. These results are derived from the dMRI scan taken of patient 1. Each tractogram consists of 5000 streamlines seeded at a point at the base of the brain stem, such that the integral of each histogram is equivalent. The histogram bin size was set to 1, equal to the step size used by the algorithm. The bins therefore represent all possible streamline lengths that the algorithm can reproduce. As additional rotation is added, the histogram becomes more skewed to the left, reflecting the tendency of the iFOD2 algorithm to prematurely terminate streamlines when the input dMRI contains a larger amount of rotation.

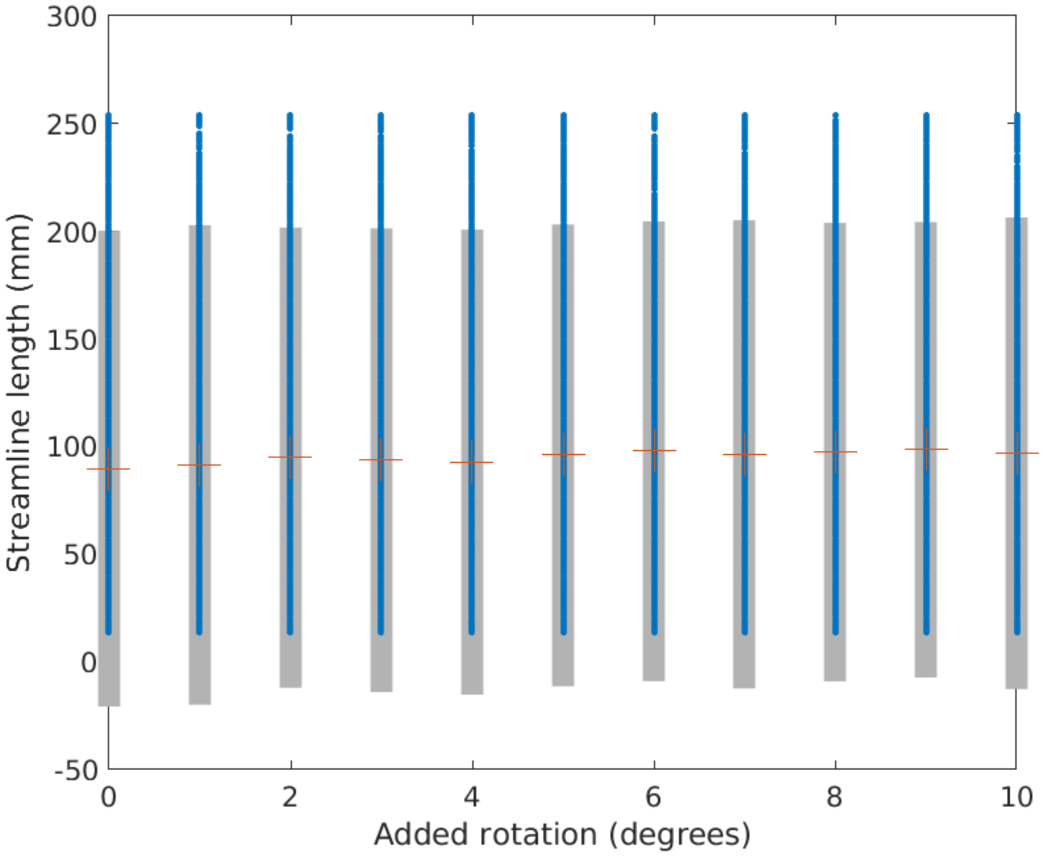


Figure 4.16: Plot illustrating the change in the overall distribution of streamline lengths for 5k point-seeded tractograms generated with iFOD2 as rotation is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

Plotting the mean streamline length and standard deviation for each of the rotation-affected tractograms for patient 1 (see Figure 4.16) reveals that the mean and standard deviation remain consistent for rotations of up to 10 degrees.

Plotting the mean streamline length relative to the mean streamline length with no added rotation against the level of added rotation for the iFOD2 whole-brain tractograms reveals two subsets of patients with upward and downward trending mean streamline lengths. The mean streamline length for patients 3, 4, 5, 6, 9, 10 and 11 increases approximately proportionally with added rotation, reaching about 10% with 10 degrees of added rotation. The mean streamline length for patients 1, 2, 7, 8 and 12 decreases as rotation is added, with these patients having a mean streamline length on average 3% shorter than the initial tractogram.

As for translation, a clear difference is observed between the change in mean streamline length for seeded tractograms versus whole brain tractograms. In this case, the mean streamline length for patient 1 remained relatively consistent for all degrees of added rotation for the seeded tractograms. Conversely, for the whole-brain tractograms, two trends are observed: either streamlines tend to be terminated prematurely, or become falsely elongated with increased rotation in the input scan. The level of consistency in the mean streamline length for the patient 1 seeded tractograms is likely due to the fact that the algorithm is repeatedly tracing out the same pathways (or near-identical pathways), as it is being forced

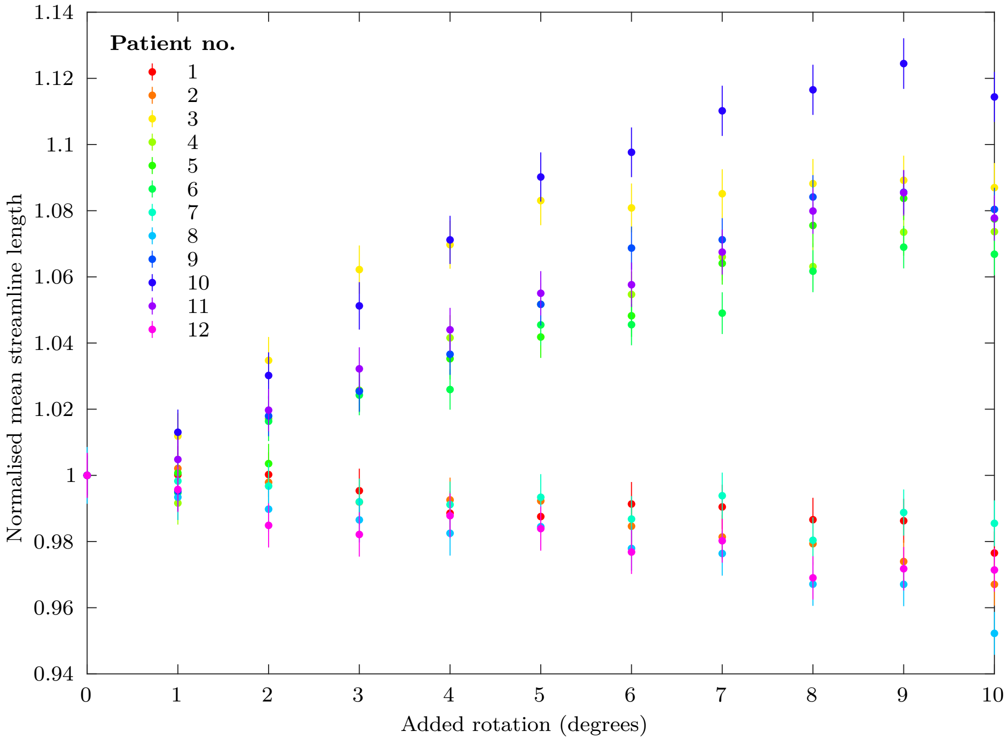


Figure 4.17: The normalised mean streamline length was calculated as the ratio of the mean streamline length to the mean streamline length with no added rotation. Plotting the normalised mean streamline length against the added rotation for all whole-brain tractograms generated using iFOD2 reveals two diverging trends, with the streamline lengths increasing on average for one subset of patients, and decreasing for the others.

to reconstruct streamlines starting from the same point in the brain. On the other hand, for the whole brain tractograms there are more possibilities for the algorithm to either trace out a similar path as in the unmodified tractogram but terminate when approaching an intricate structure (yielding a shorter streamline than expected), or to generate longer streamlines by artificially connecting shorter paths (yielding a longer streamline than expected).

4.5.2 SD_STREAM

A histogram of streamline lengths was generated for each SD_STREAM 5k point-seeded tractogram with added rotation (see Figure 4.18). The initial distribution contains two sharp peaks which sit between 30 mm and 40 mm, and a short, narrow band of counts at around 22 mm. With progressive addition of rotation, the distribution shows evidence of false positives, with additional peaks appearing at higher lengths. With an added rotation of 10 degrees, the distribution becomes restricted to a band between 37 mm and 48 mm. A possible explanation for this behaviour is that rotation of the gross white matter structure by this amount has resulted in most structures becoming too far unaligned between volumes to be delineated. Most streamlines will be rejected as the algorithm has greater difficulty traversing the path of the white matter, so only one or a few structures which happen to align in the unmodified and rotated volumes remain. It is important to note that rotation of half of the volumes by a significant amount may cause structures that would otherwise not be directly connected being traced out as a single structure.

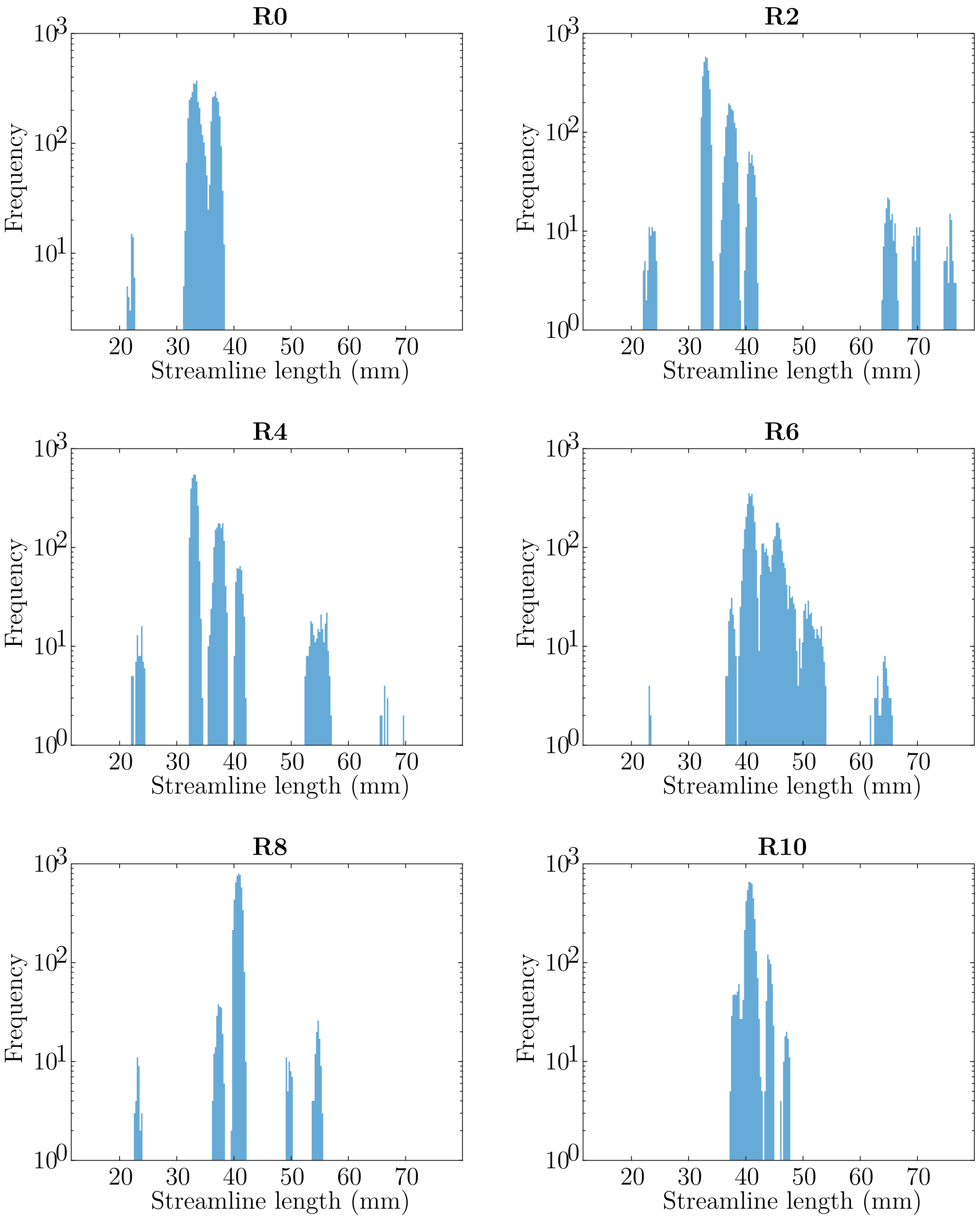


Figure 4.18: Histograms of streamline length for tractograms generated using MRtrix's SD_STREAM algorithm with varied levels of added rotation. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered again consists of 5000 streamlines seeded at a point at the base of the brain stem. The SD_STREAM algorithm uses a finer step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are shown on a logarithmic scale to better visualise regions of low counts in the distribution.

Plotting the mean streamline length and standard deviation for each point-seeded tractogram against the degree of added rotation (Figure 4.19) shows erratic changes in the stream-

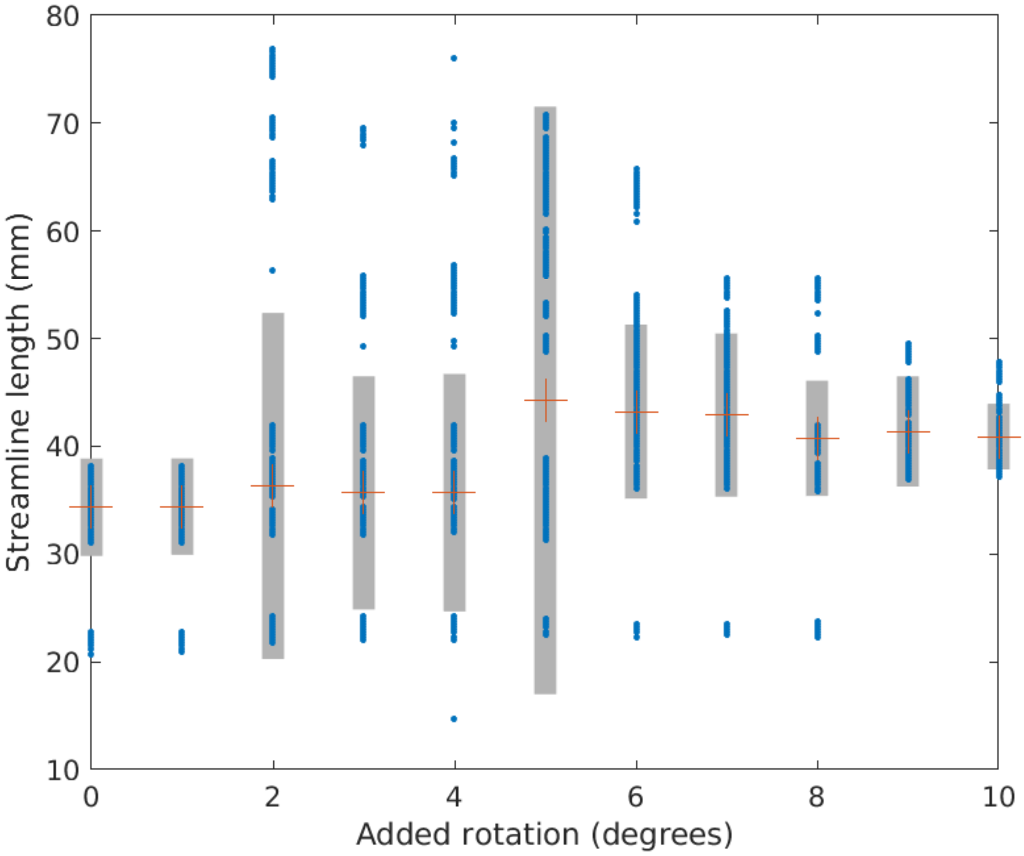


Figure 4.19: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with SD_STREAM as rotation is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

line length distribution. A step change is observed in the mean streamline length at 5 degrees of added rotation, from about 35 mm to roughly 40 mm. The standard deviation initially increases up to 5 mm added rotation, then reduces, presumably as the algorithm has increased difficulty delineating most structures and instead traces out very similar paths repeatedly to make up the 5,000 streamlines which comprise the tractogram.

Figure 4.20 shows the mean streamline length normalised by the initial mean streamline length for whole-brain tractograms against the degree of added rotation. A progressive decline in mean streamline length is observed with added rotation for all patients. The magnitude of this effect varies between patients, with the mean streamline length decreasing by between 1% and 14%. This suggests that SD_STREAM favours the production of false negatives as opposed to false positives with increased rotation between scan volumes.

4.5.3 Tensor_Prob

Histograms of the streamline lengths comprising the Tensor_Prob 5k point-seeded tractograms with added rotation are shown in Figure 4.21. The unmodified tractogram contains a large peak centred at about 30 mm, and peaks of decreasing magnitude spaced at regular

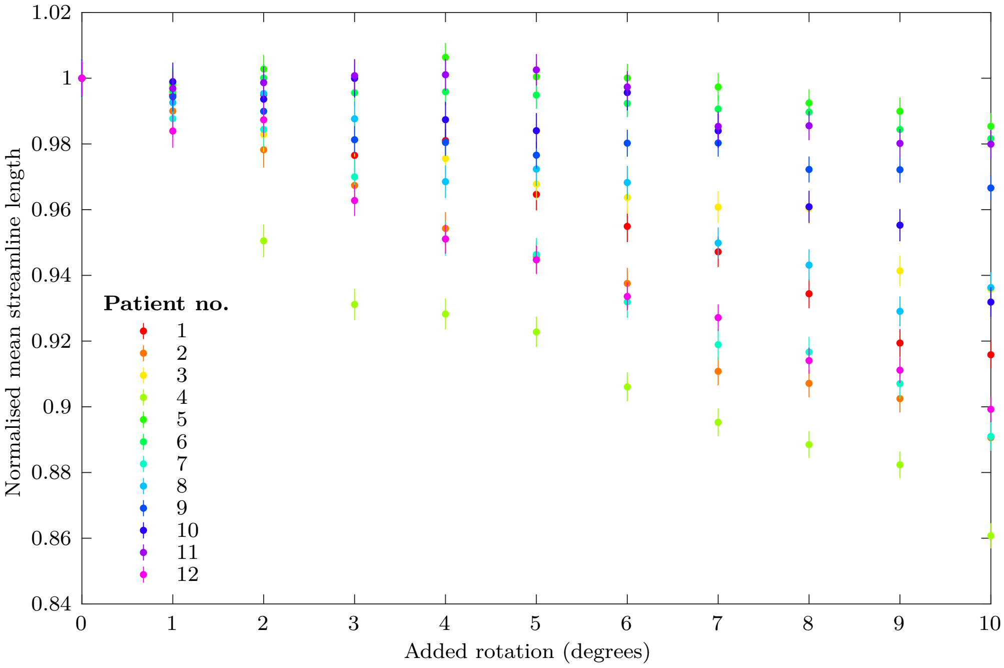


Figure 4.20: Plot of the normalised mean streamline length versus added rotation for all whole-brain tractograms generated using SD_STREAM.

length intervals up to about 120 mm. The distribution with 2 degrees of added rotation is similar. However, an increased number of counts are present in the peaks which fall between 50 mm and 100 mm. Similar changes occur between 2 and 6 degrees of added rotation. Beyond this, the distribution begins to blur out, and false positives are produced with added rotation 10 degrees.

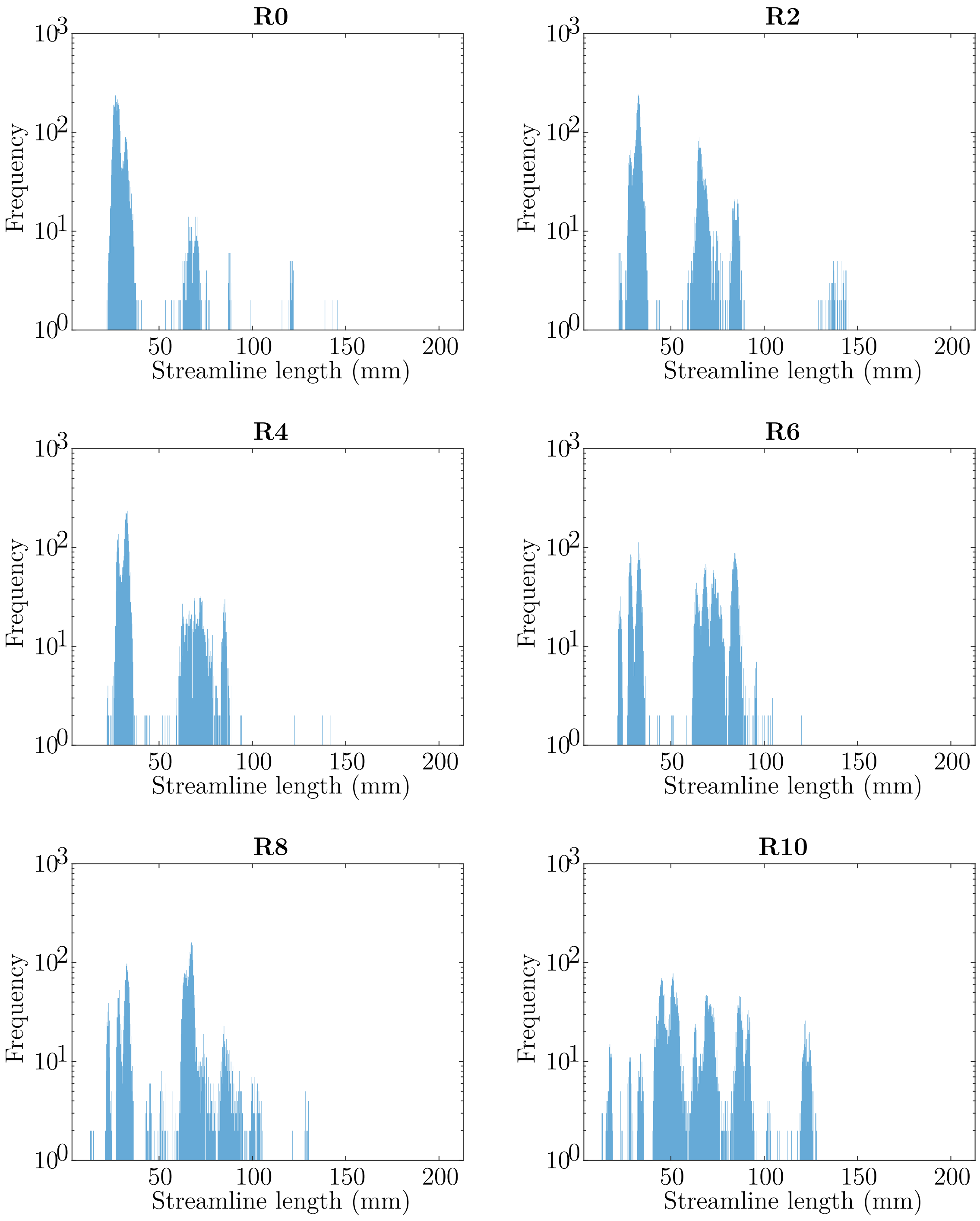


Figure 4.21: Histograms of streamline length for tractograms generated using MRtrix's Tensor_Prob algorithm with varied levels of added rotation. These results are derived from the dMRI scan taken of patient 1. Each tractogram considered consists of 5000 streamlines seeded at a point at the base of the brainstem. The Tensor_Prob algorithm uses a step size of 0.2 by default, which was inputted as the histogram bin size. The histograms are shown on a logarithmic scale to better visualise regions of low counts in the distribution.

The change in mean streamline length and standard deviation for point-seeded tractograms with added rotation is shown in Figure 4.22. The mean streamline length increases between 0 and 10 added degrees of rotation. However, the spread in the distribution remains

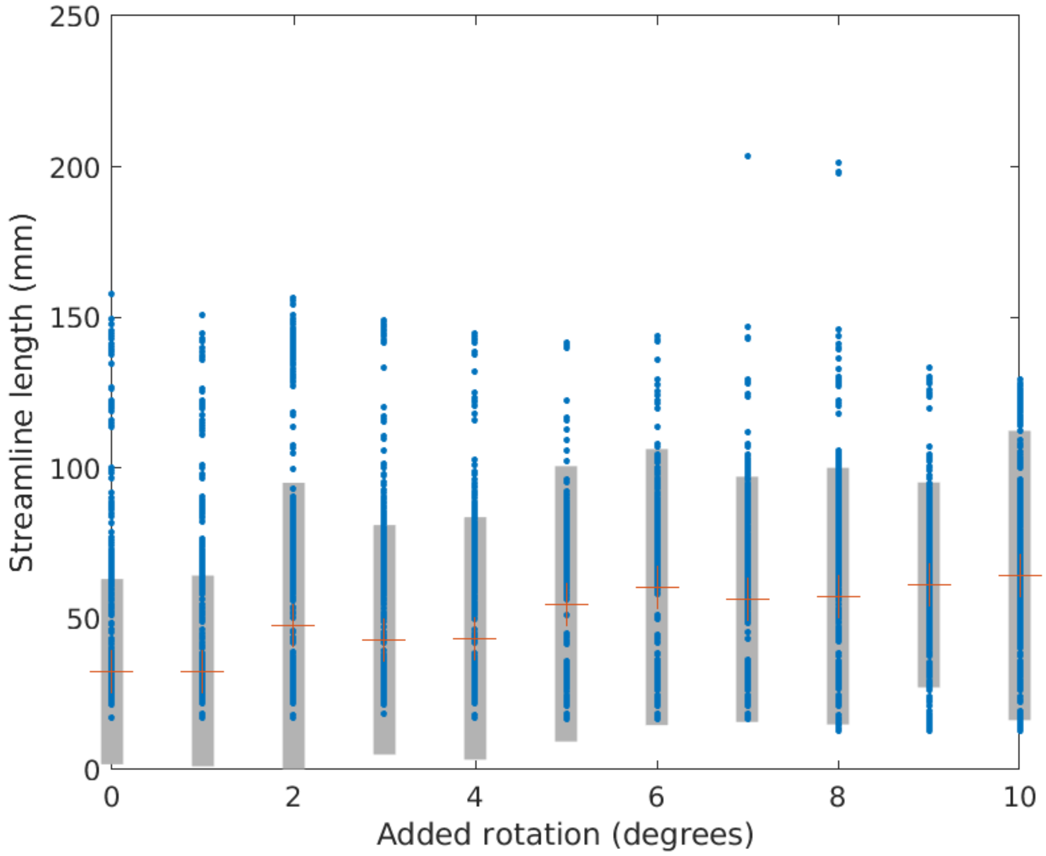


Figure 4.22: Plot illustrating the change in the overall distribution of streamline lengths for tractograms generated with Tensor_Prob as rotation is added to the input dMRI. The individual streamline lengths, and the mean and 2 SD of each distribution are shown in blue, red and grey, respectively.

fairly consistent across this range.

Figure 4.23 depicts the change in normalised mean streamline length for whole-brain tractograms with added rotation generated using Tensor_Prob. A smooth decrease in the mean streamline length is observed for all patients, with the mean streamline length at 10 degrees having decreased by between 11% and 34%.

4.5.4 Comparison of rotation effects on iFOD2, SD_STREAM and Tensor_Prob

For both SD_STREAM and Tensor_Prob, the mean streamline length decreases with added rotation for the whole-brain tractograms, whereas for iFOD2, a subset of patients show a decline in mean streamline length, while another subset show an increase of larger magnitude in mean streamline length. This suggests that SD_STREAM and Tensor_Prob are more likely to terminate streamlines early with added rotation (i.e. produce a greater proportion of false negatives), whereas iFOD2 may respond to rotation addition through generating false positives or false negatives.

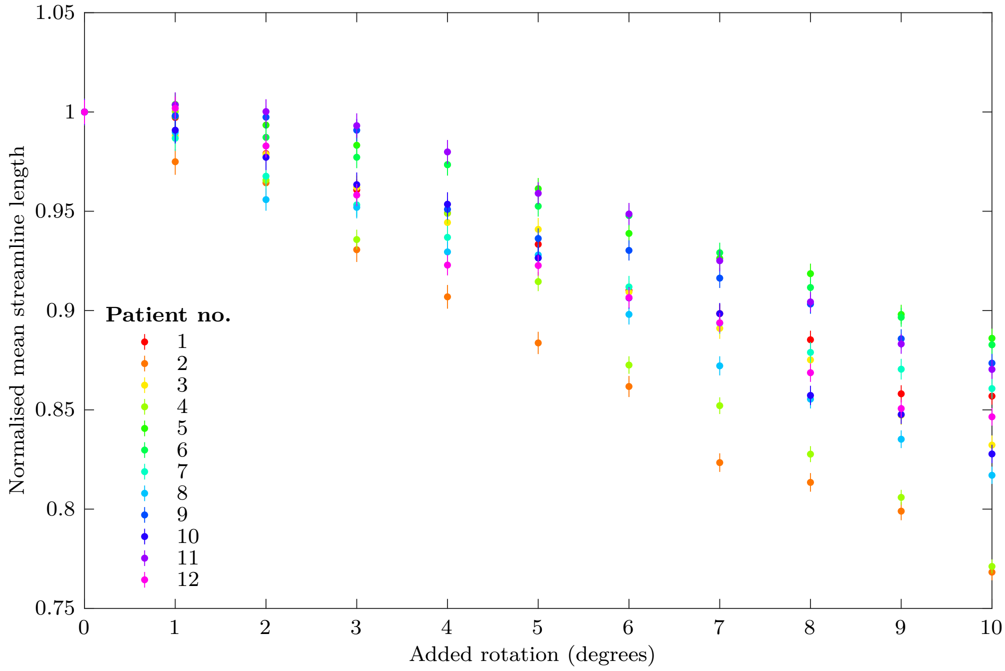


Figure 4.23: Plot of the normalised mean streamline length versus added rotation for all whole-brain tractograms generated using Tensor_Prob.

A different trend is evident when the mean streamline length for patient 1's point-seeded tractograms are plotted against rotation. For iFOD2, mean streamline length remains consistent up to added 10 degrees rotation, while SD_STREAM and Tensor_Prob show slight increases in mean streamline length over the same range. The size of the difference in mean streamline length is less than 10 mm for iFOD2 and SD_STREAM, and up to about 50 mm for Tensor_Prob, which suggests delination of individual tracts using Tensor_Prob is more prone to changes induced by rotation of the subject in the axial plane.

4.6 Polar plots with added translation

A 5000 streamline tractogram seeded from the base of the brain stem was generated for each input scan with added pure translation using the algorithms iFOD2, SD_STREAM and Tensor_Prob. The seed point was set as the centre point of the tract shown in the most caudal slice of the unmodified scan.

From each tractogram generated, polar plots were produced of the change in the angle of the streamline with respect to the initial tract angle, θ, vs the distance from the base of the tract, r; and the change in the angle of the streamline with respect to the x-axis, ϕ against r. 3D histograms were generated to illustrate the density of the streamlines in these plots.

A 5k tractogram seeded from a point at the base of the brain stem could not be generated for the scan taken of patient 1 with an added translation of 5 mm using SD_STREAM, or with added translations of 4 mm and 5 mm using Tensor_Prob.

The algorithm was potentially unable to generate a tractogram in these cases as the scan volumes were displaced to such an extent that averaging across them removed most of the diffusion information. Consequently, the algorithm was unable to trace out streamlines of minimum length with sufficient confidence.

For the seeded tractograms generated using iFOD2, shown in Figure 4.24, a progressive deterioration of the tract is observed with added translation. Major structures are retained for an added translation of 1 mm. Beyond this, the density of streamlines declines significantly, and longer streamlines are lost, until only the tract base remains with 5 mm added translation. This is evident in both polar plots showing the density of streamline points for θ versus r, and ϕ versus r.

Interestingly, when a whole brain tractogram is produced from a combination of unmodified volume/s and displaced volume/s, the resulting tractogram retains most of the structures as in the unmodified whole brain tractogram, but they become displaced from their original positions.

4.6.1 iFOD2

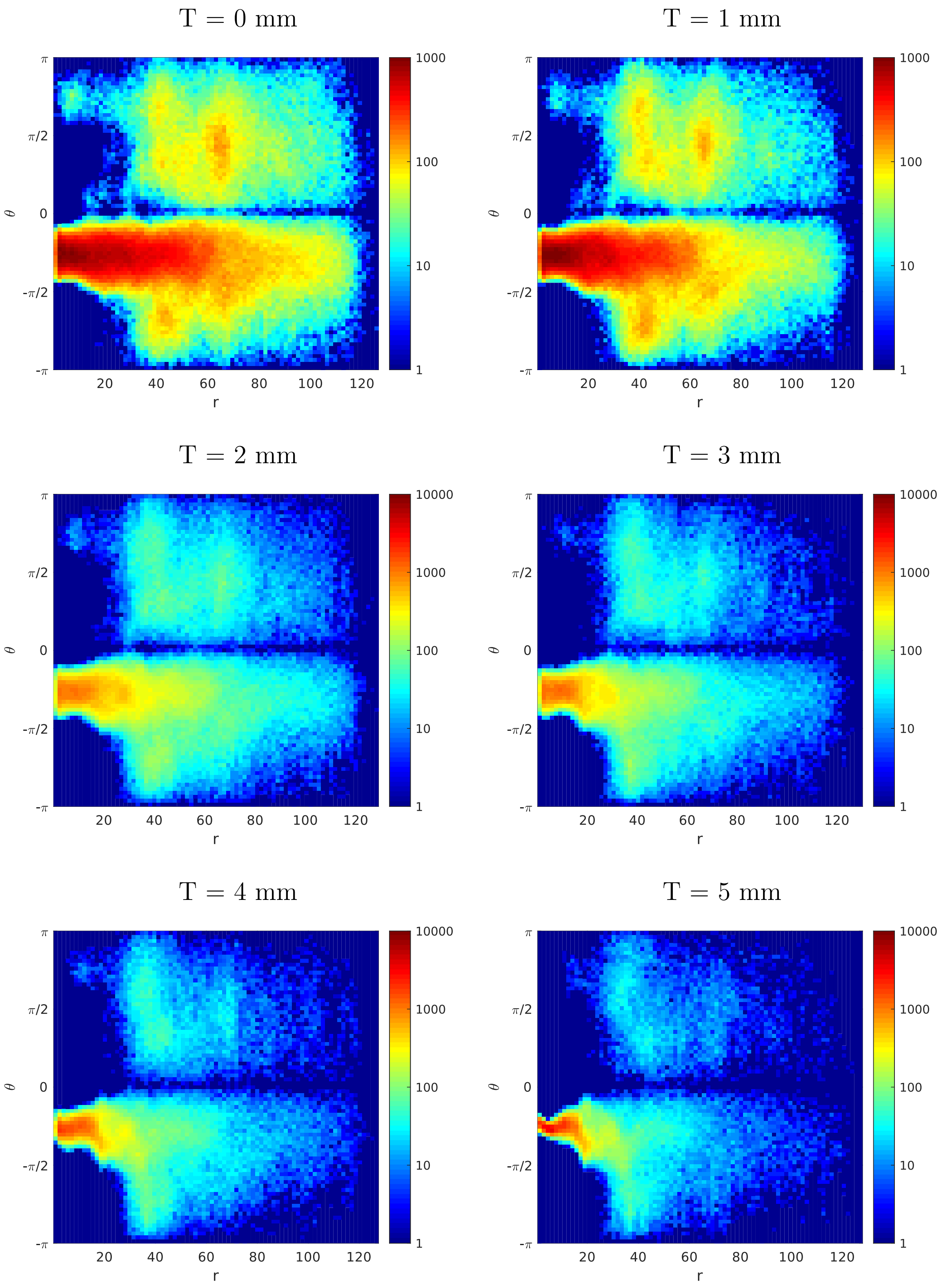


Figure 4.24: Polar plots of θ vs r for tractograms with added translation generated using iFOD2 for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

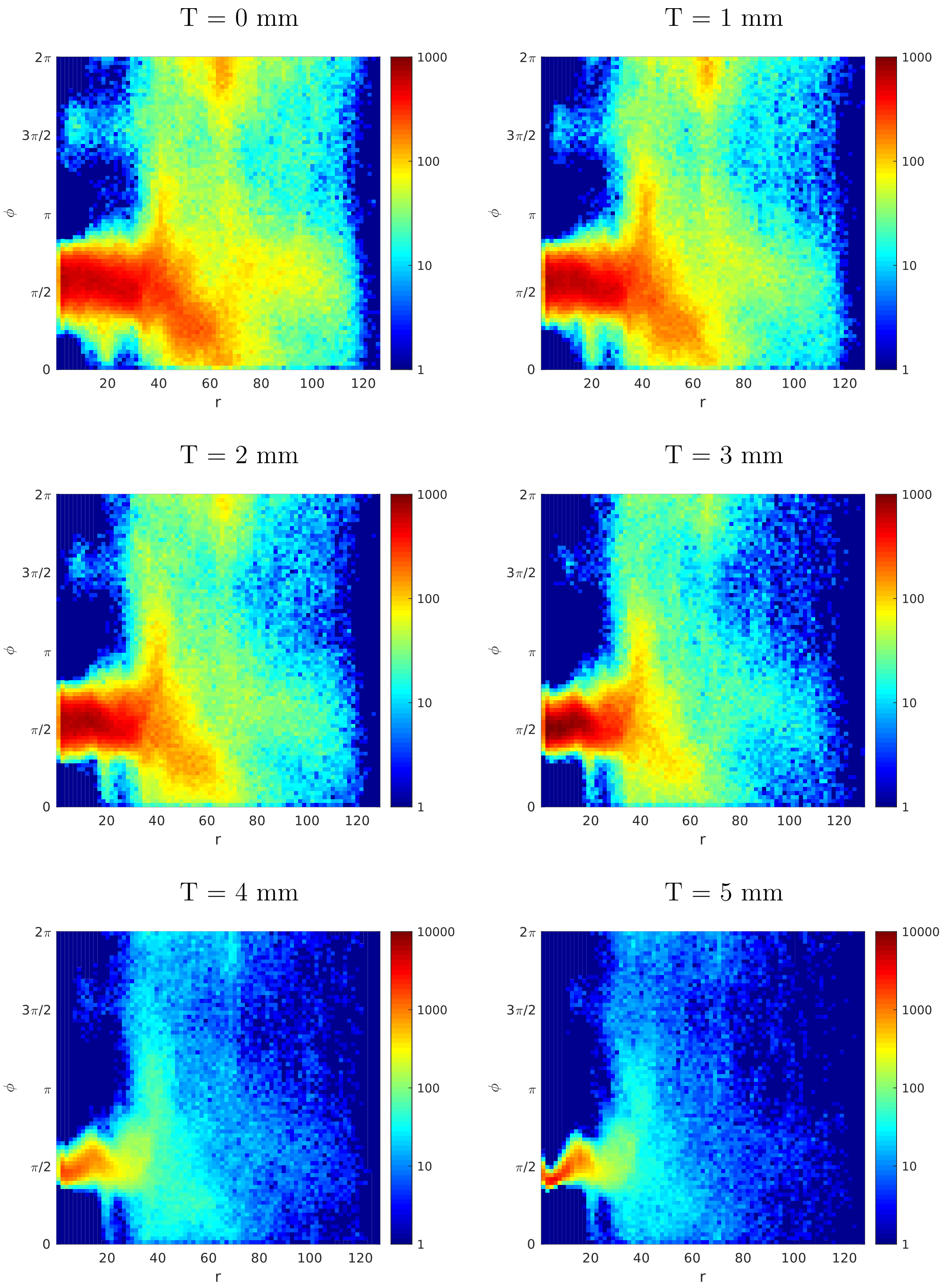


Figure 4.25: Polar plots of ϕ vs r for tractograms with added translation generated using iFOD2 for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

4.6.2 SD_STREAM

The polar plots with increments of added translation generated using SD_STREAM also show a deterioration in structure as translation is added. The initial polar plots demonstrate that streamlines follow two distinct paths branching off from the base of the brainstem. With an added translation of 1 mm in the second scan volume, the θ vs r polar plot is minimally affected, with some divergence evident in the ϕ vs r polar plot. Further increasing the amount of added translation causes the observed relationship between each angle and r to progressively deteriorate. Both the shape and length of streamlines are affected. The tractograms with 4 mm added translation terminate nearer to the seed point, and streamline paths are generally more erratic with greater translation.

The loss of streamline length results in reduced density for higher r values. This is particularly evident for streamline points with ϕ approximately equal to π and r between 10 and 15. With minimal added translation, this region is fairly dense. For an added translation of 3 mm, a significant decline in the density in this region of the polar plot is observed, demonstrating that the corresponding streamline paths are significantly less likely to be delineated.

Regions of the tract which are generally retained, even with significant translation, are features that are most prominent in the initial scan, and therefore the highest density of streamlines. This is evidenced by consistency in the shape of red regions in the plots.

As observed for noise, the SD_STREAM algorithm typically produces more refined structures than the iFOD2 algorithm, and appears to be less strongly impacted by modifications of the input scan. This is likely due to the fact that SD_STREAM primarily delineates prominent features in the tractogram, omitting more intricate structures to produce a tractogram with less false positive, but substantially more false negative streamlines.

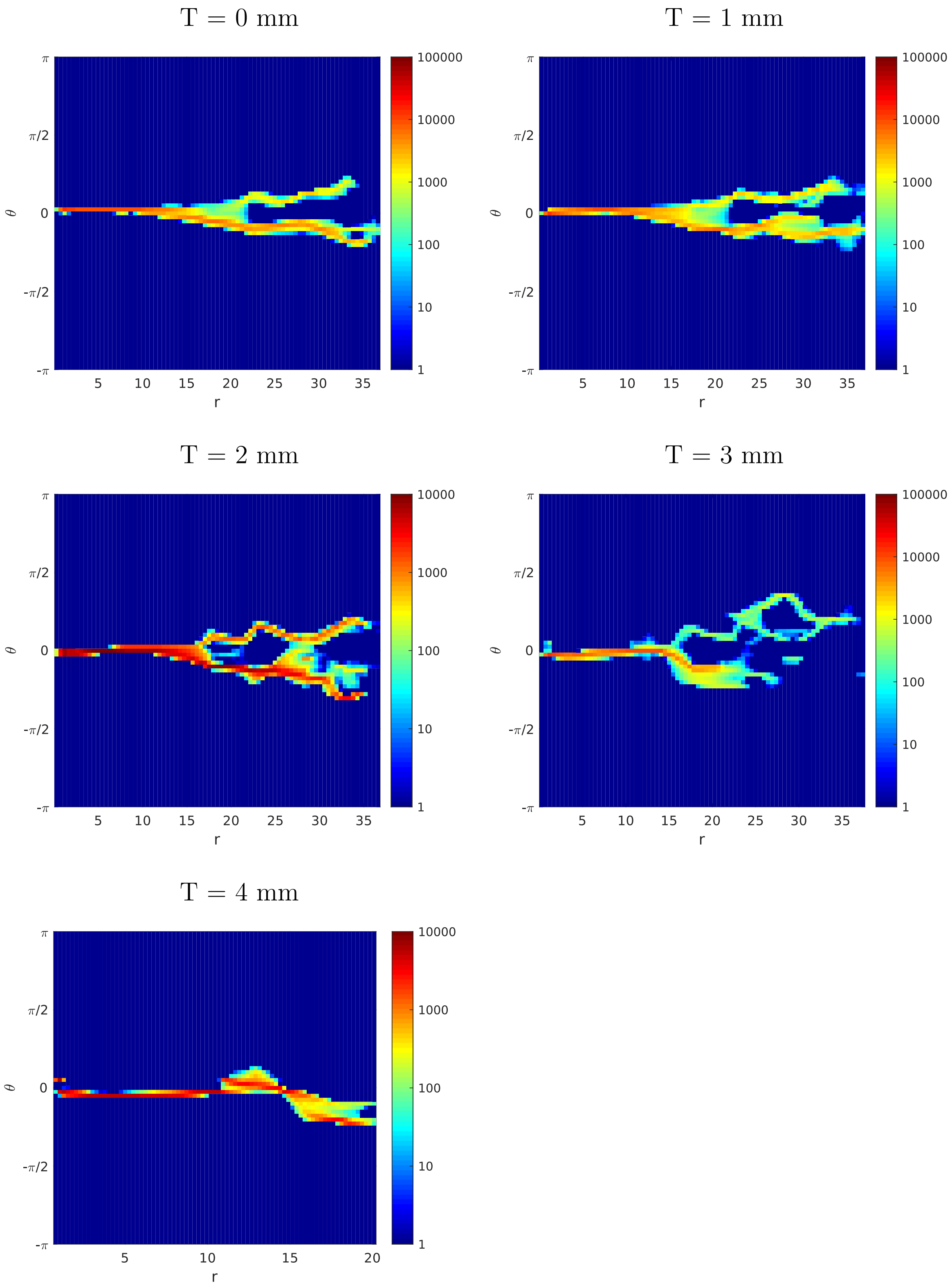
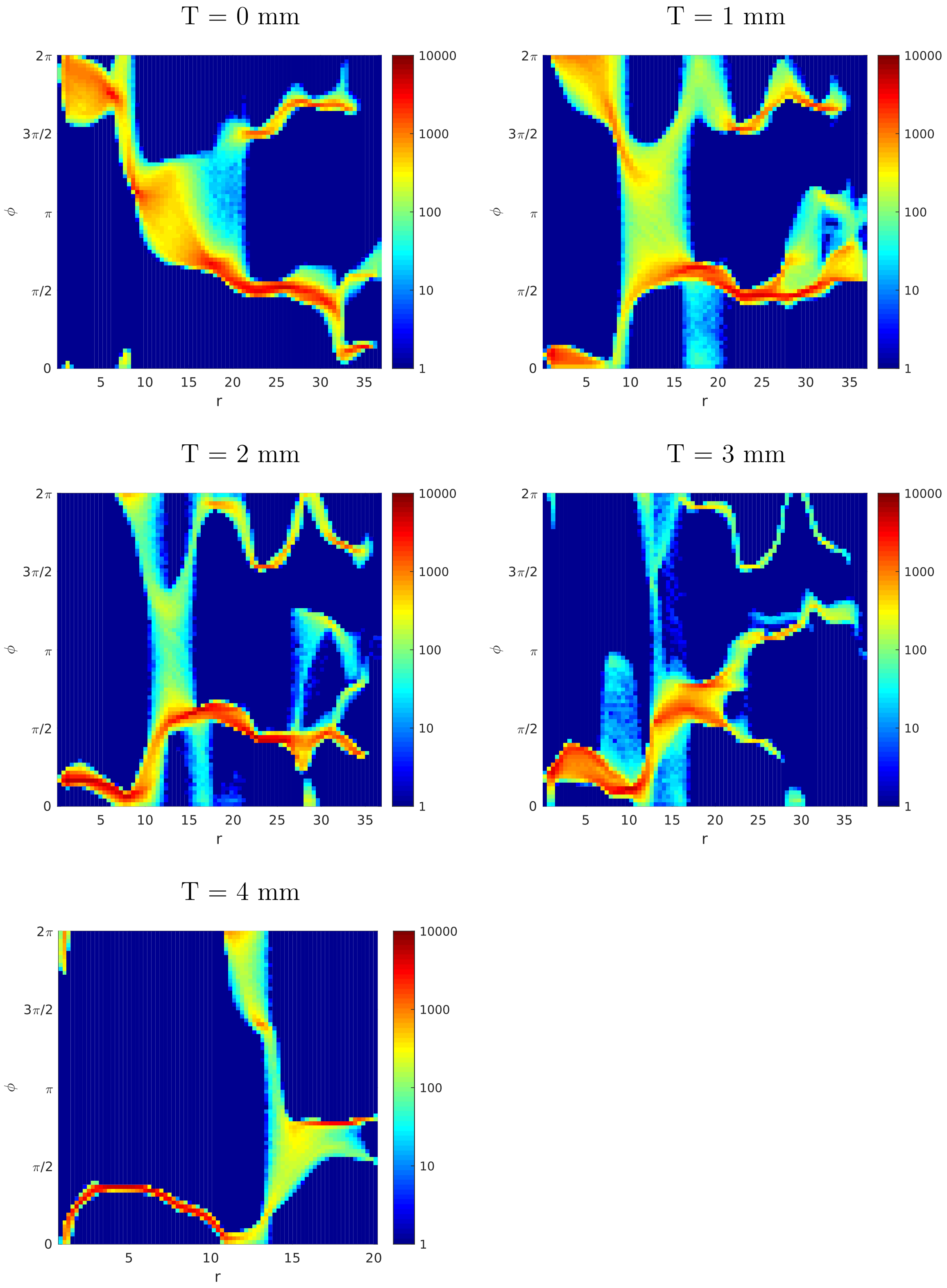


Figure 4.26: Polar plots of θ vs r for tractograms with added translation generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.


T = 0 mm

T = 1 mm

T = 2 mm

T = 3 mm

T = 4 mm

Figure 4.27: Polar plots of ϕ vs r for tractograms with added translation generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

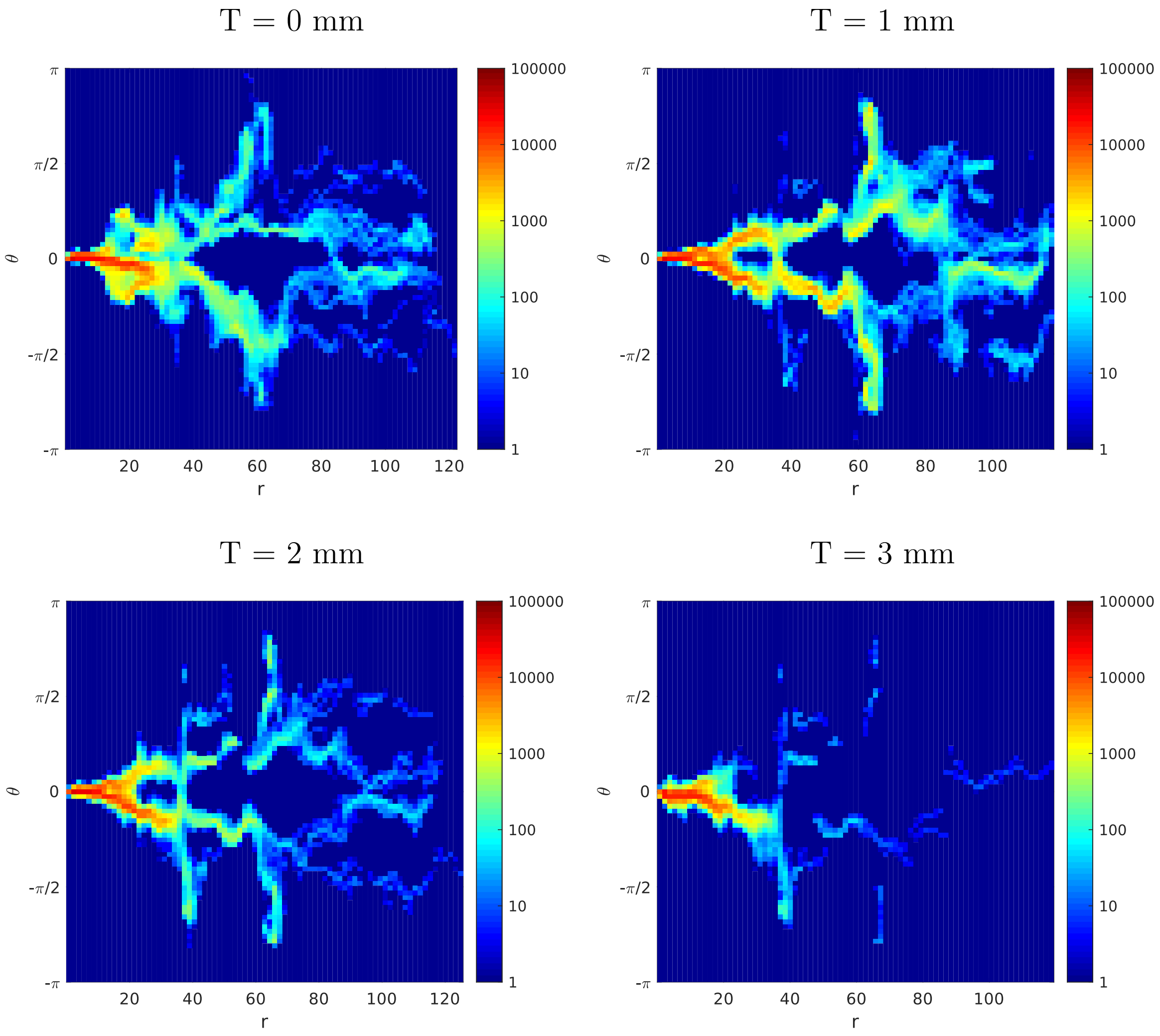
4.6.3 Tensor_Prob

The Tensor_Prob algorithm also produces more structured tractograms than iFOD2, with mainly high diffusion signals represented by the reconstructed streamlines. However, a higher level of intricacy is present in the tractogram structure than for SD_STREAM.

The high density streamline regions are conserved for added translations of up to 3 mm, which in this case mainly consists of the brain stem. Detailed structures are increasingly impacted by added translation, with a translation of 3 mm compromising the diffusion signals to such an extent that most of the detailed tract structure beyond about r = 30 is gone.

The deterioration of the tract observed with added translation for Tensor_Prob occurs much more rapidly than for iFOD2 or SD_STREAM. This may reflect that this algorithm relies on a higher degree of certainty (i.e. strength of the diffusion signal) to delineate a given streamline path.

With the addition of translation, the polar plot converges to a similar structure as for the highest translation plotted for SD_STREAM, corresponding to regions in the scan of highest diffusion signal intensity.



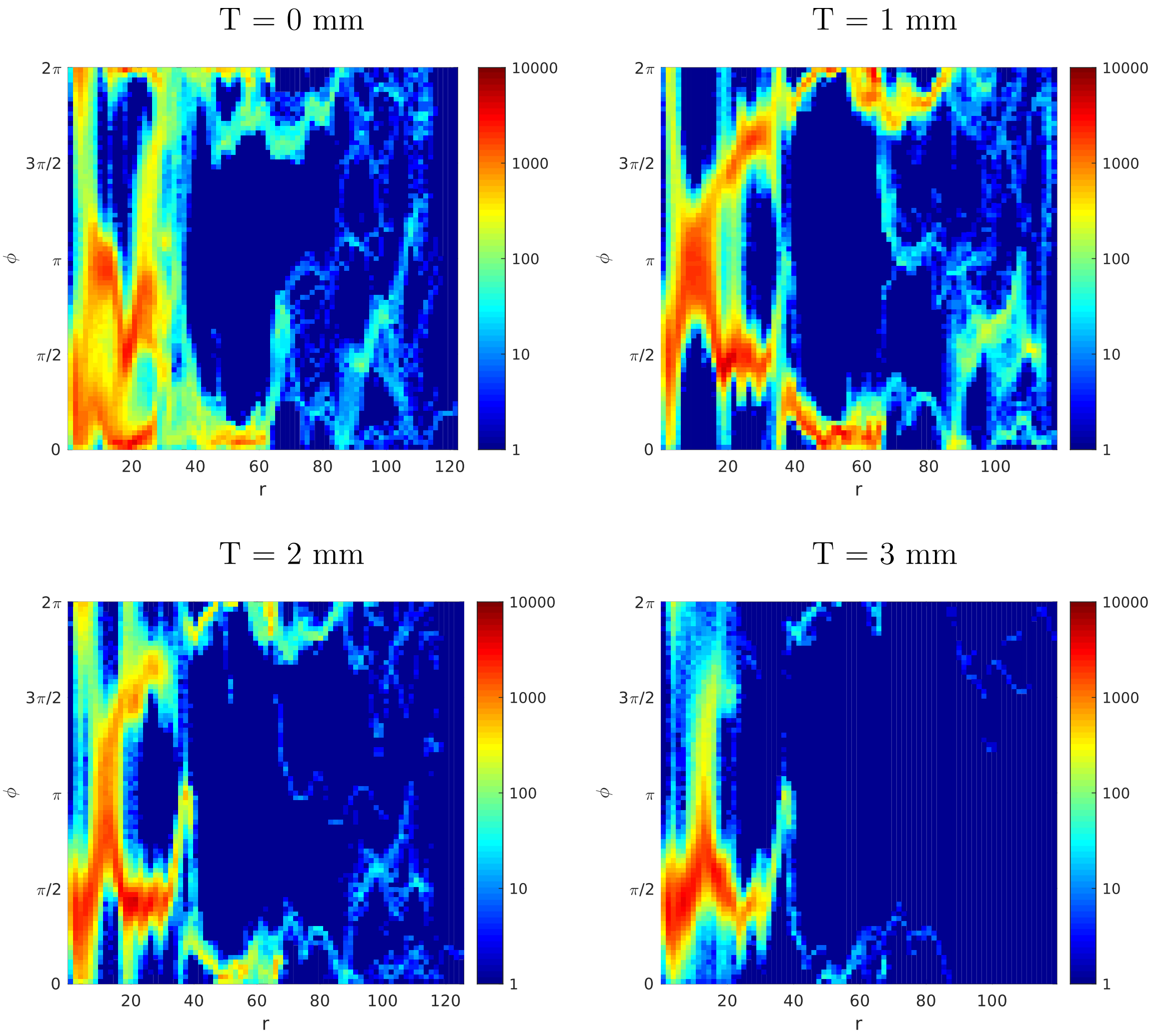
T = 0 mm

T = 1 mm

T = 2 mm

T = 3 mm

Figure 4.28: Polar plots of θ vs r for tractograms with added translation generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.


T = 0 mm

T = 1 mm

T = 2 mm

T = 3 mm

Figure 4.29: Polar plots of ϕ vs r for tractograms with added translation generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

4.7 Polar plots with added rotation

A 5000 streamline tractogram seeded from the base of the brain stem was generated for each input scan with added pure rotation. The seed point was set as the centre point of the tract shown in the most caudal slice of the unmodified scan. Polar plots were produced for each seeded tractogram with added rotation.

Incremental addition of rotation to the input dMRI scan results in slight deterioration of the corresponding polar plots for iFOD2. The loss of detail is not significant for added rotations of up to 10 degrees.

4.7.1 iFOD2

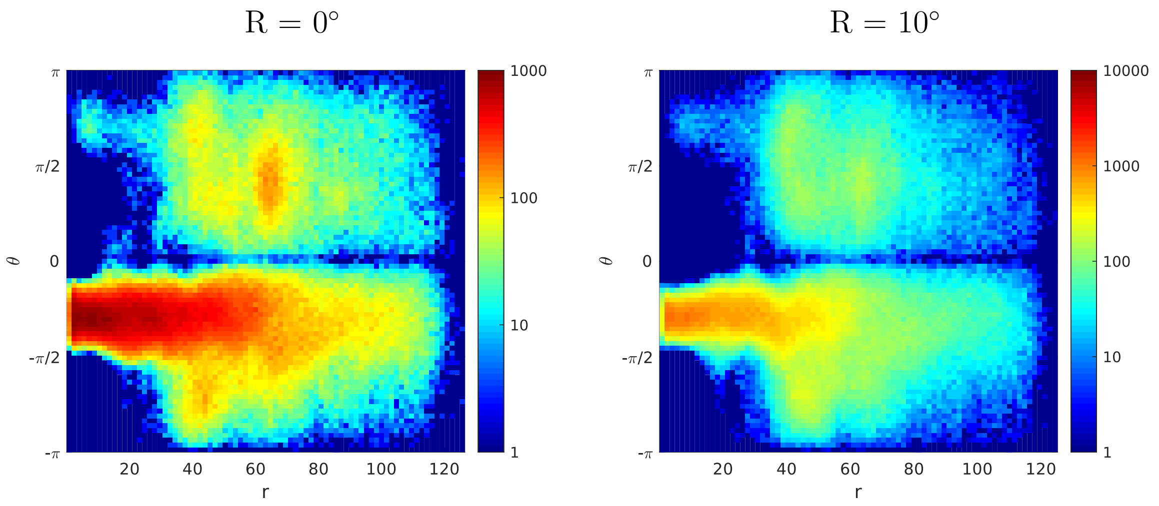


Figure 4.30: Polar plots of θ vs r for tractograms with added rotation generated using iFOD2 for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

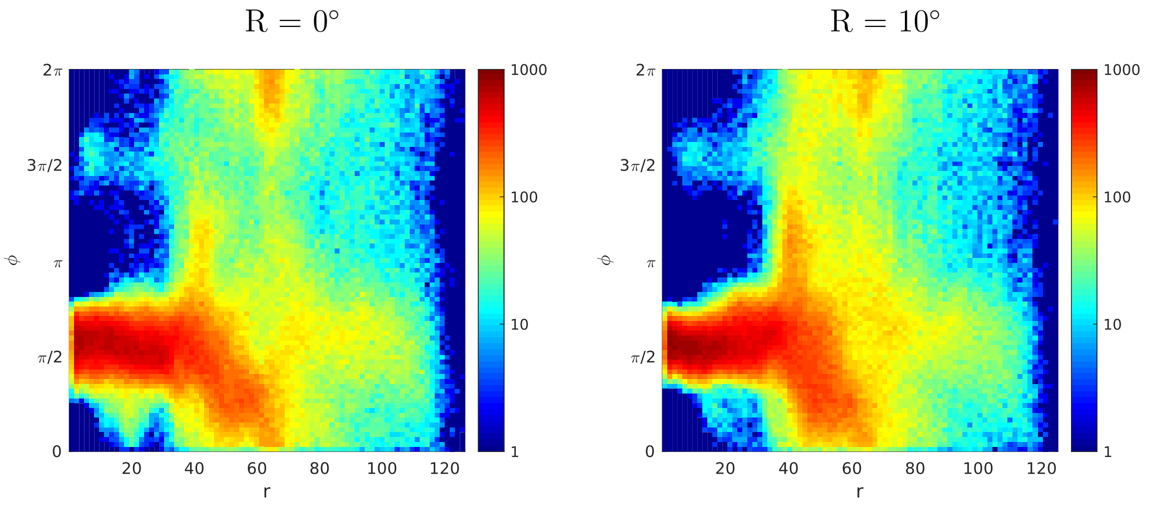


Figure 4.31: Polar plots of ϕ vs r for tractograms with added rotation generated using iFOD2 for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

4.7.2 SD_STREAM

For SD_STREAM, more significant changes are observed with small amounts of added rotation. Initially, with no added rotation, the θ vs r polar plot shows a tract structure with two forked paths. With only 2 degrees of added rotation, one of these forked paths is lost. The second is retained up to an added rotation of about 8 degrees, beyond which it appears only the tract base remains.

The ϕ vs r polar plots illustrate similar changes with added rotation. In the first 2 degree increment, one of the major tract paths is largely lost. The formation progressively changes as further rotation is added.

This example illustrates that the output of SD_STREAM is highly susceptible to rotational differences in scan volumes, relative to iFOD2.

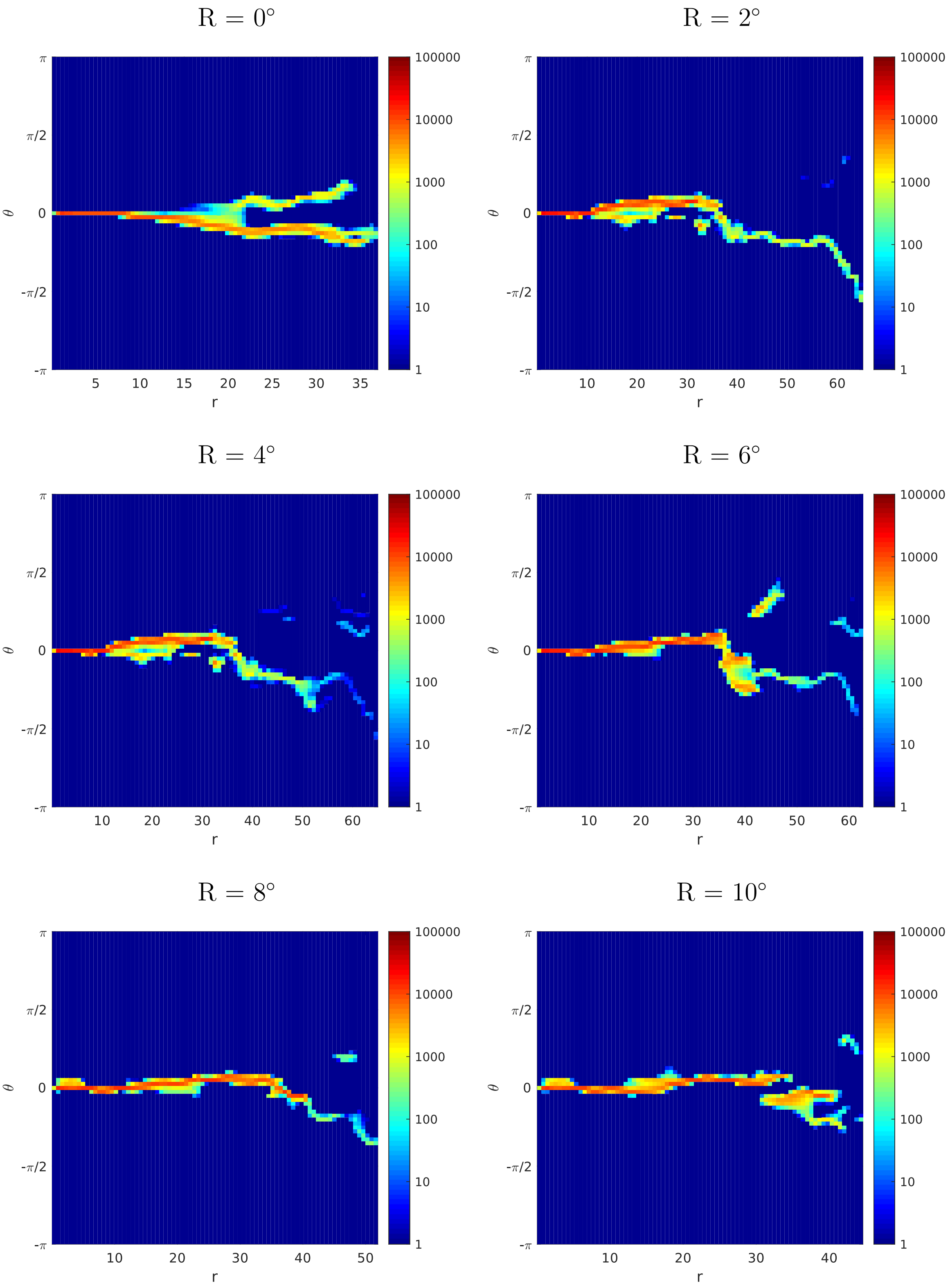
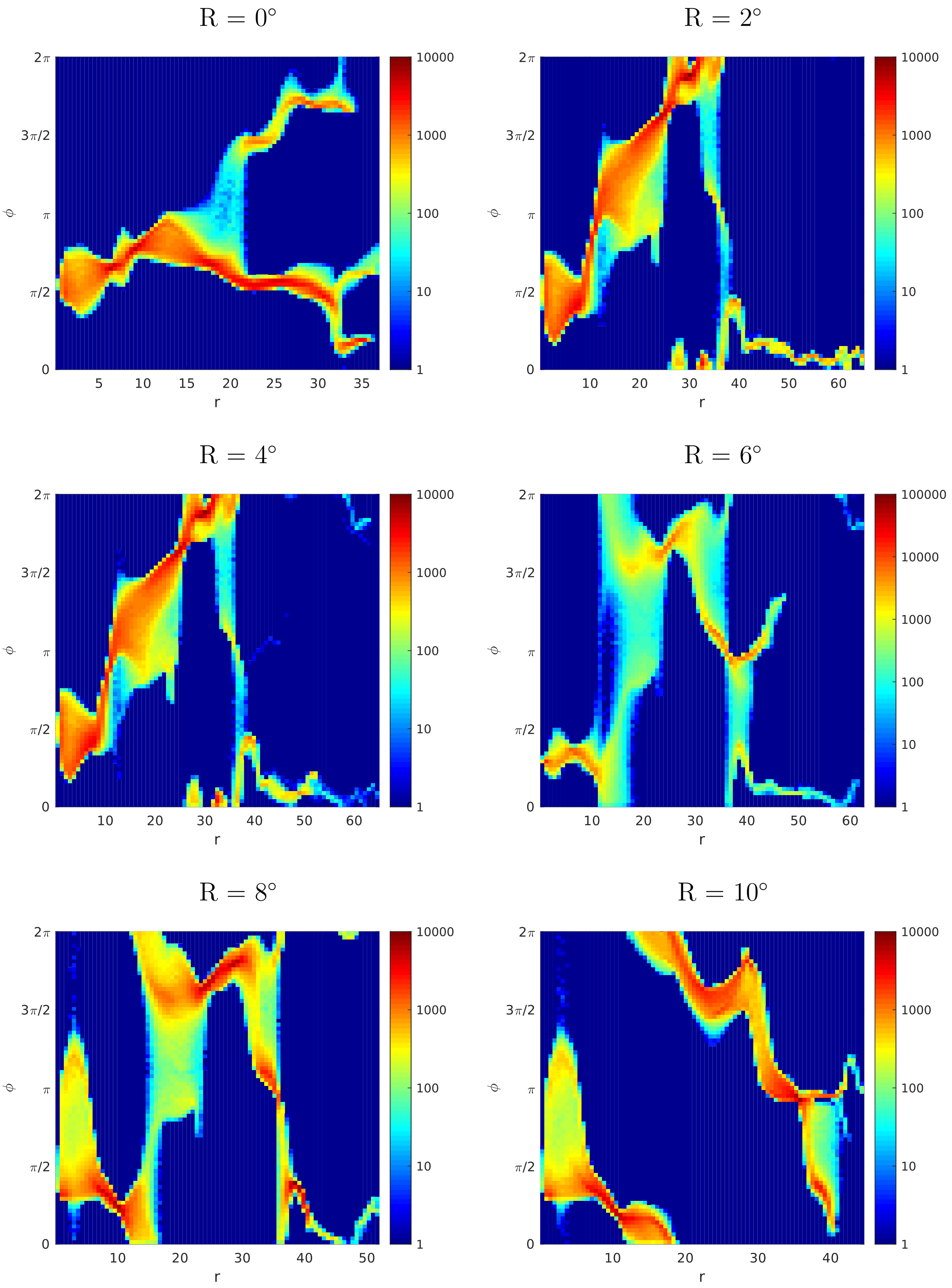


Figure 4.32: Polar plots of θ vs r for tractograms with added rotation generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.



R = 0°

R = 2°

R = 4°

R = 6°

R = 8°

R = 10°

Figure 4.33: Polar plots of ϕ vs r for tractograms with added rotation generated using SD_STREAM for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

4.7.3 Tensor_Prob

Tensor_Prob appears to be less strongly impacted by added rotation than SD_STREAM. The addition of rotation to half of the scan volumes introduces false positive and negatives. However, with a small amount of rotation (i.e. about 2 degrees), the effect is not significant. In the θ vs r polar plots with no added rotation and 2 degrees of added rotation, an increase in the intensity of the lower tract path, and a decrease of intensity in the upper tract path (particularly more intricate structures further along the tract) is observed. More intricate structures are lost with added translation up to about 6 degrees. Beyond this, the rotation seems to cause different diffusion paths to become aligned, resulting in higher intensity forked paths which resemble but are not wholly consistent with those for the 0 degree tract.

Similar observations can be made from the polar plots showing the change in the ϕ vs r with added rotation. The structure towards the tract base remains consistent up to a rotation of about 6 degrees. Beyond this, false positives appear in the data.

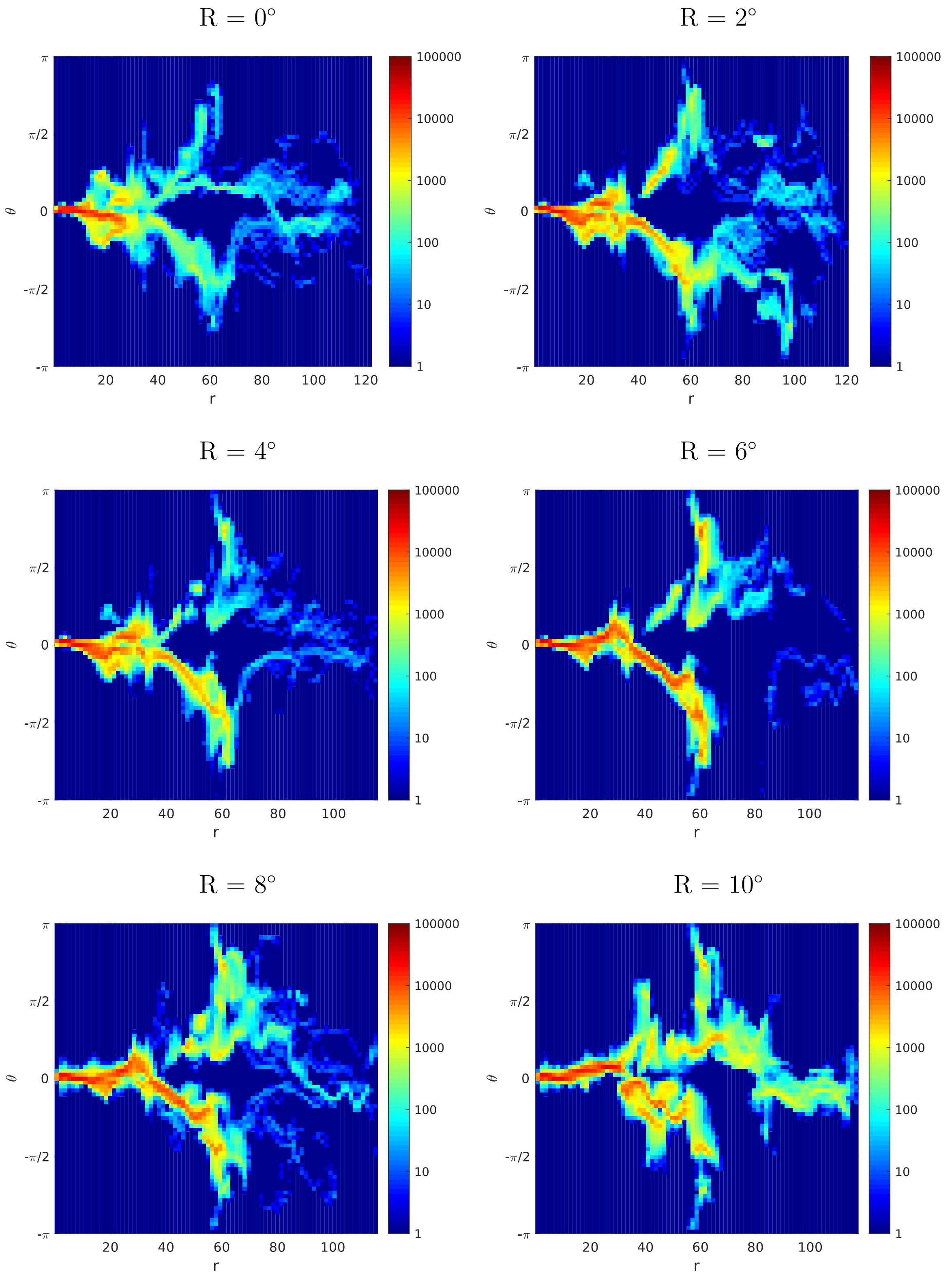


Figure 4.34: Polar plots of θ vs r for tractograms with added rotation generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

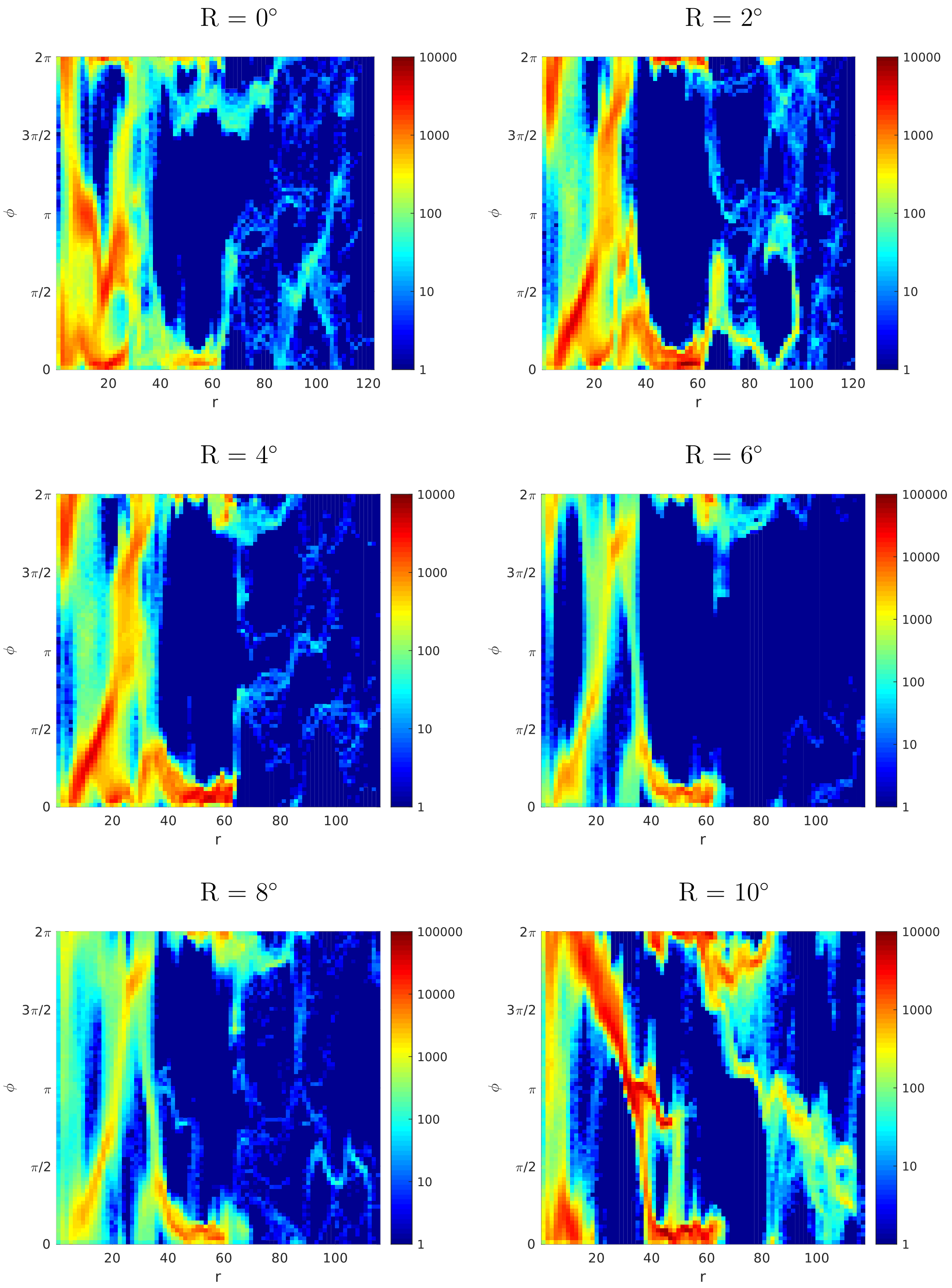


Figure 4.35: Polar plots of ϕ vs r for tractograms with added rotation generated using Tensor_Prob for patient 1. The colour scale indicates the number of streamline points with a given displacement (i.e. angle θ and distance r) from the origin.

4.8 Resemblance plots

The resemblance metric, described in section 3.10, was used to assess the loss of consistency between tract-weighted maps with added motion and the unmodified tract-weighted maps.

For each patient and increment of added translation and rotation, a tract map was generated from a 50k streamline tractogram of the whole brain. For each of these modified tract-weighted maps the resemblance value was computed.

4.8.1 Translation

Figure 4.36 shows the change in resemblance metric with added translation up to 5 mm for tract-weighted maps generated using iFOD2. A large drop in the resemblance value occurs with 1 mm added translation. This decrease in resemblance of about 35% may reflect the loss or modification of fine diffusion structures in the tractogram with even small translational differences between scan volumes.

Further addition of translation then causes the resemblance to decrease at a progressively slower rate, suggesting that only high intensity diffusion structures are retained with 5 mm added translation. If significant further translation is added, the iFOD2 algorithm becomes incapable of delineating streamlines. This is due to diffusion paths becoming displaced to such an extent in the two scan volumes that averaging across them removes the diffusion information, such that the algorithm cannot confidently delineate any white matter structures.

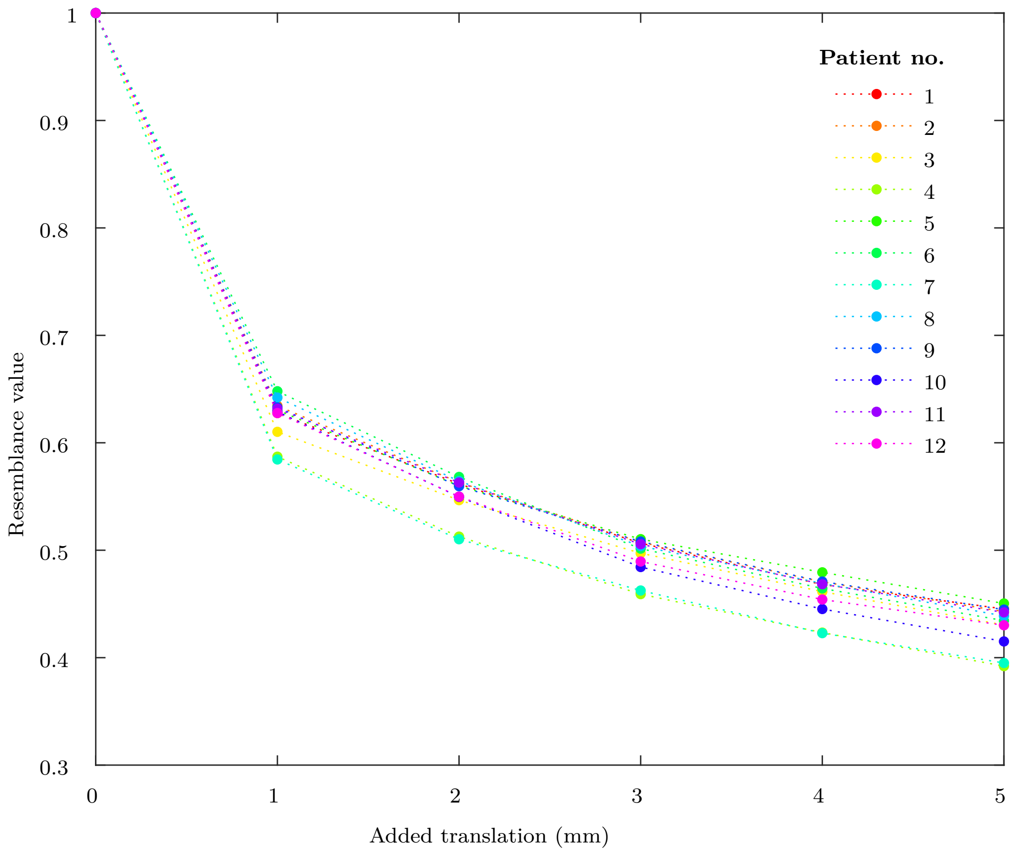


Figure 4.36: Plot of the calculated resemblance value versus added translation for whole-brain tract-weighted maps generated using iFOD2.

For the SD_STREAM algorithm, the resemblance value similarly shows a significant drop with the first increment of added translation. The decrease in the resemblance value with further added translation again has a plateau which begins to become obvious at around 5 mm added translation.

While the relationship between the resemblance value and the amount of added translation appears to be fairly consistent between patients when using iFOD2, there is significantly more spread in resemblance values when using SD_STREAM. Patient 4 also shows a much larger drop in resemblance of about 50% with 1 mm added translation, compared with a drop of between 30% and 40% for the other patients.

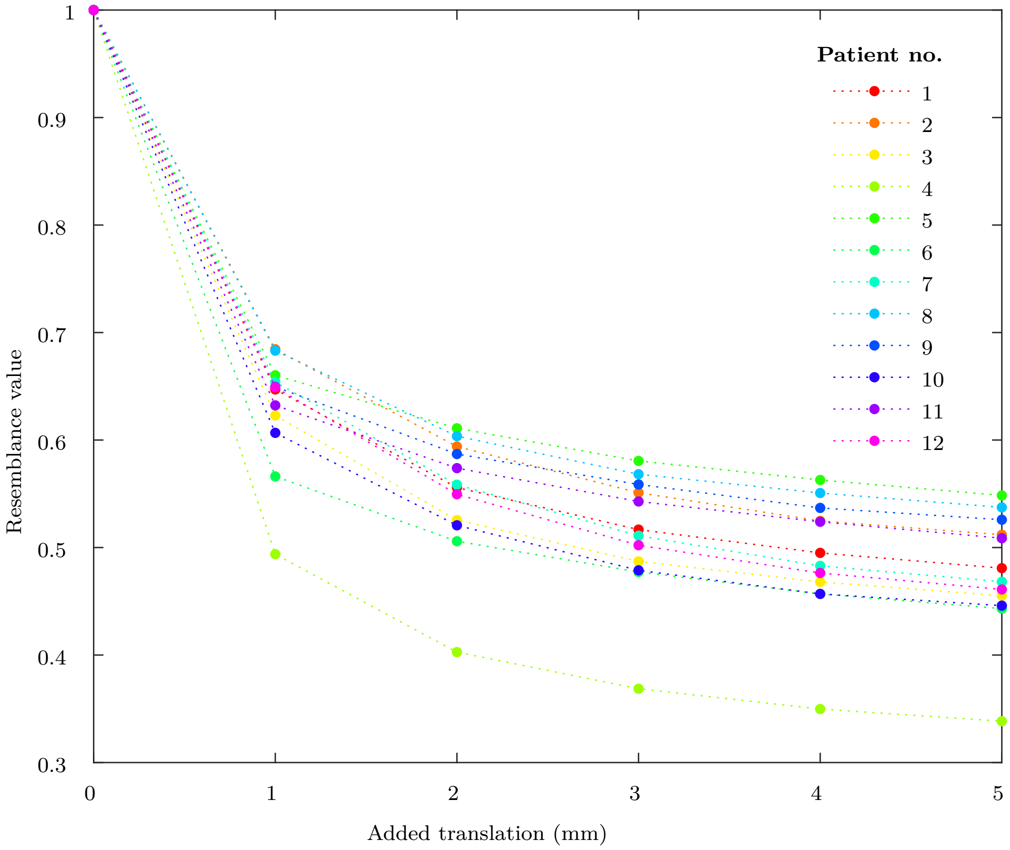


Figure 4.37: Plot of the calculated resemblance value versus added translation for tract-weighted maps generated using SD_STREAM.

Figure 4.38 depicts the change in resemblance value with added translation for the Tensor_Prob algorithm. The drop in resemblance with the first increment of translation is less significant than for iFOD2 or SD_STREAM. Again, resemblance drops by smaller amounts with each additional increment. The spread in resemblance values is substantial, compared with iFOD2 and SD_STREAM.

There does not appear to be any consistency in which patients' tractograms are most strongly affected by added translation between algorithms. For instance patient 4 has the lowest resemblance values at each increment of added translation when SD_STREAM is used, and the highest resemblance values when Tensor_Prob is used.

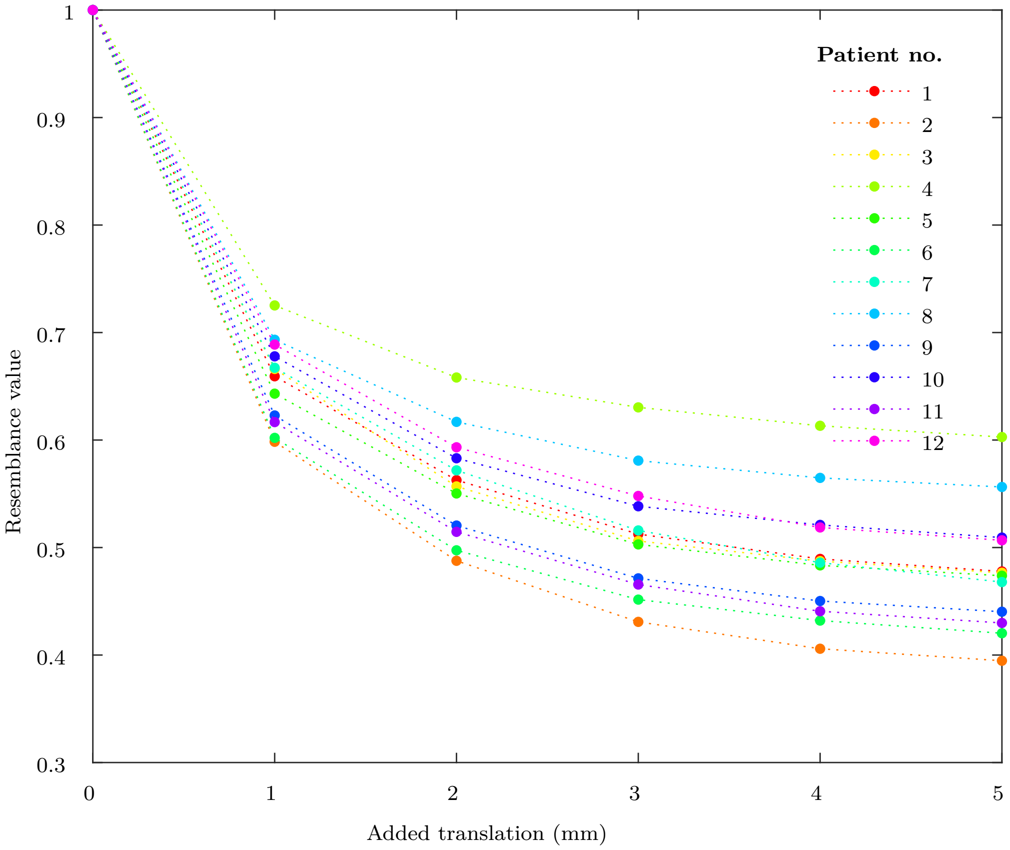


Figure 4.38: Plot of the calculated resemblance value versus added translation for tract-weighted maps generated using Tensor_Prob.

4.8.2 Rotation

Figure 4.39 shows the change in resemblance value with added rotation up to 10 degrees for iFOD2. As observed for the translation resemblance plots, a significant drop, of about 30%, occurs after the first increment of added rotation. The resemblance then decreases progressively with additional rotation, flattening towards 10 degrees of added rotation.

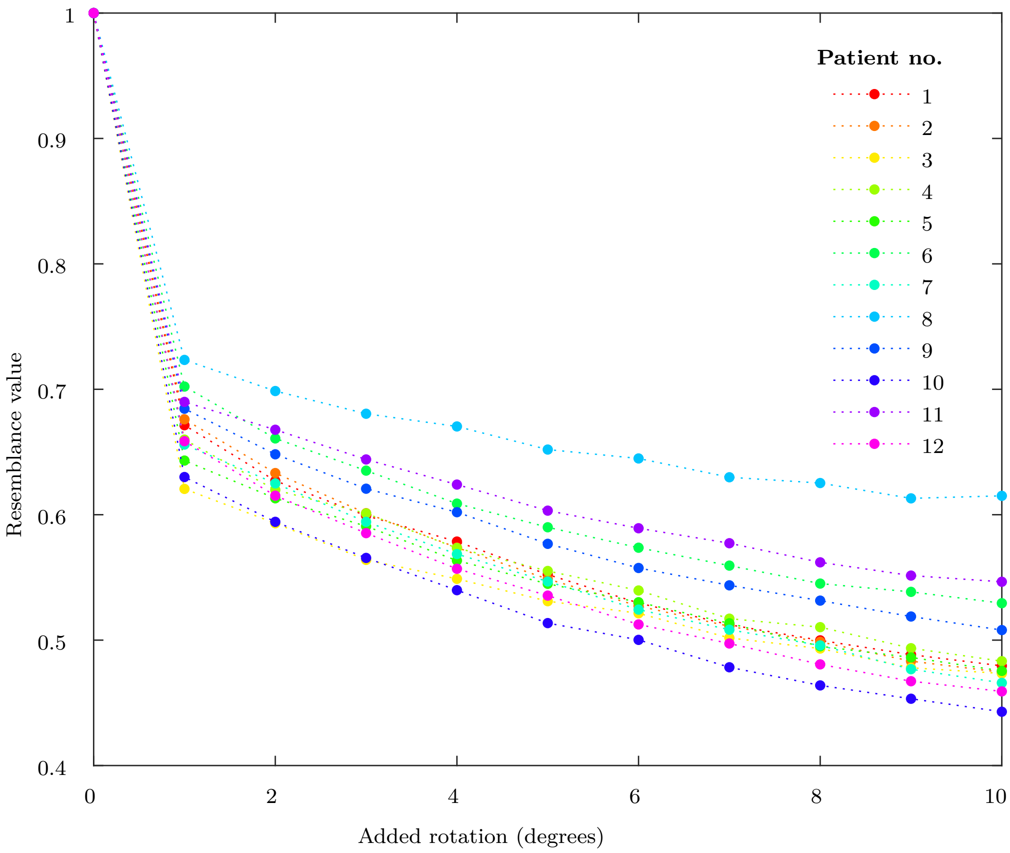


Figure 4.39: Plot of the calculated resemblance value versus added rotation for tract-weighted maps generated using iFOD2.

The change in resemblance is more gradual for SD_STREAM. However, the spread of values is similar to that for iFOD2.

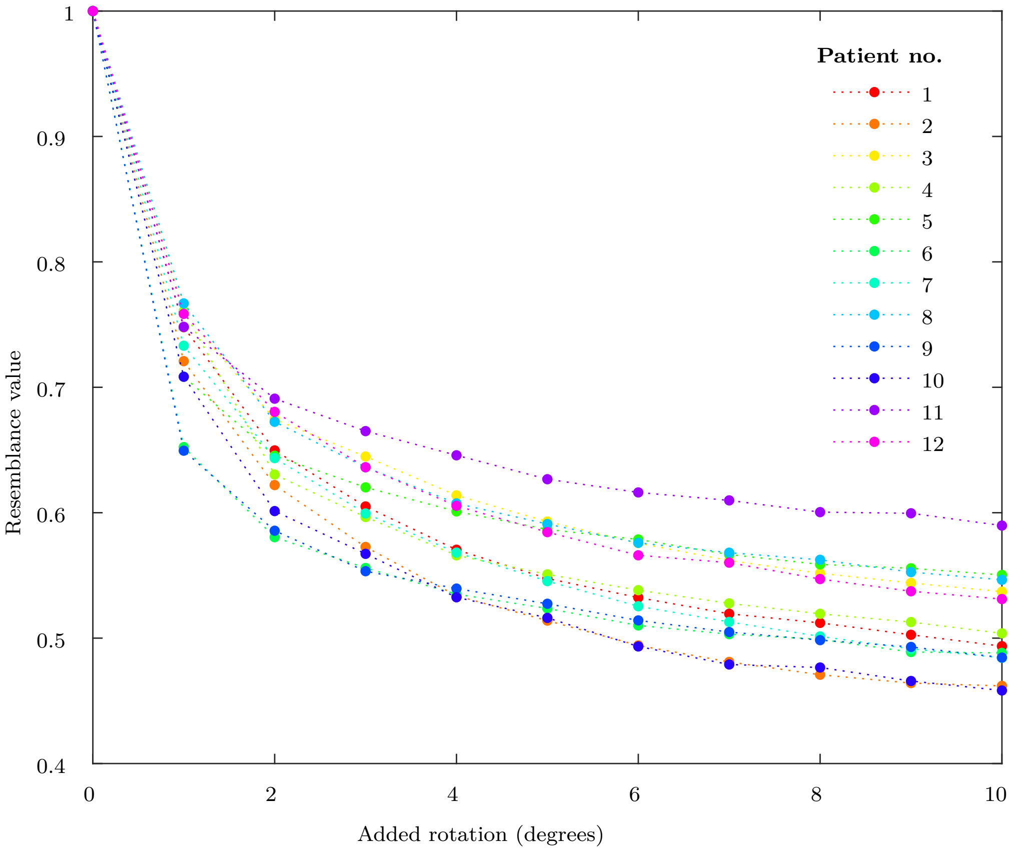


Figure 4.40: Plot of the calculated resemblance value versus added rotation for tract-weighted maps generated using SD_STREAM.

For Tensor_Prob, an even more gradual decline in resemblance occurs with added rotation. The spread in resemblance values for different patients is slightly greater than the other two algorithms.

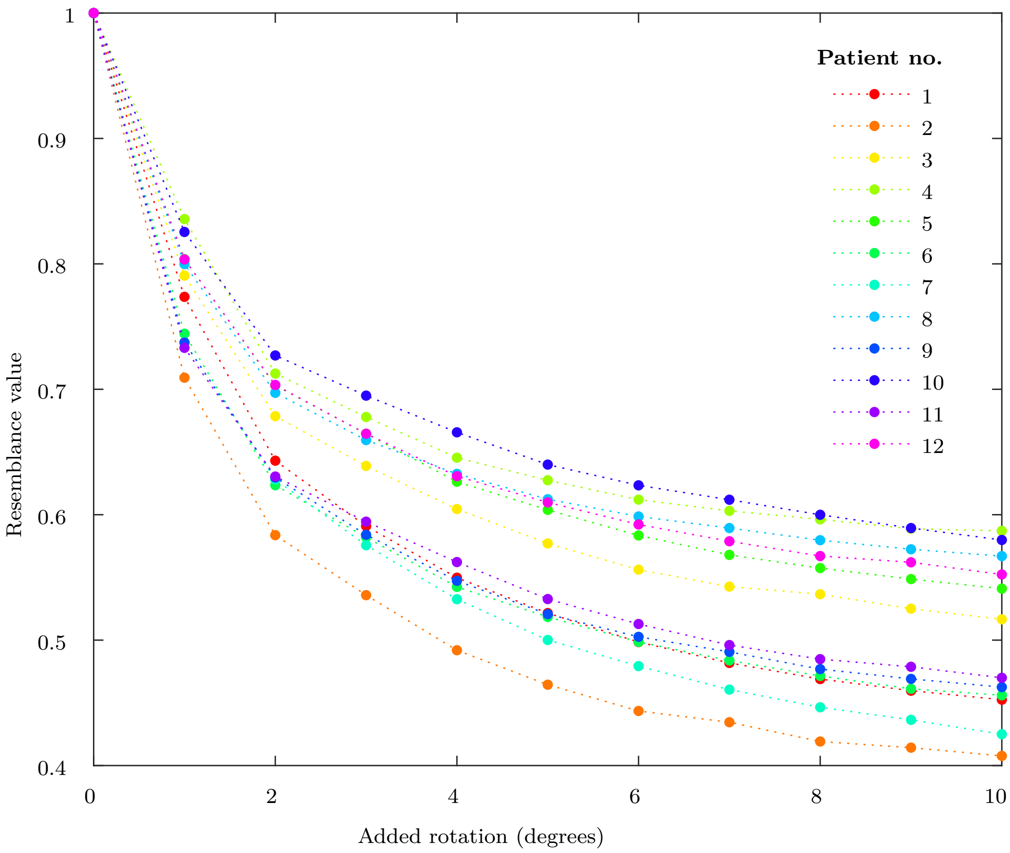


Figure 4.41: Plot of the calculated resemblance value versus added rotation for tract-weighted maps generated using Tensor_Prob.

4.9 Root-mean-square (RMS) difference in tract-weighted maps

An alternative measure of the change in a tract-weighted map with added motion is the RMS difference between the original and modified tract-weighted maps, described in section 3.11.

4.9.1 Translation

Figure 4.42 shows the RMS difference between the modified and original tract-weighted maps generated using iFOD2 for each patient with added translation. With intrinsic motion only, the RMS difference in tract-weighted map voxel values is up to 8. This provides a baseline level with which to compare the observed RMS difference for tract-weighted maps with added translation. With 1 mm added translation, a large proportion of patients have RMS difference in tract-weighted map voxels under the maximum baseline RMS difference, which suggests that for most patients, the majority of tract structures are retained with minimal change.

For translations beyond 1 mm, the RMS difference increases as translation is increased to 5 mm. The spread in RMS difference values also increases proportionally with translation.

An increase in the RMS difference does not necessarily indicate complete loss of structures in the corresponding tract-weighted map. The value instead reflects any difference in streamline density at any position in the scan. Therefore, if displacement of scan volumes causes part of the diffusion signal to be lost, the resultant decrease in streamline density in affected regions will produce an increase in the RMS difference. By the same reasoning, displacements which cause areas of one scan volume with low diffusion signals to become overlapped with higher diffusion signals from an adjacent region will result in an increase in streamline density which also increases the RMS difference.

The increase of the RMS difference therefore cannot discriminate between false positive and negatives in the tract-weighted maps, but indicates the amount of change in reconstructed connectivity.

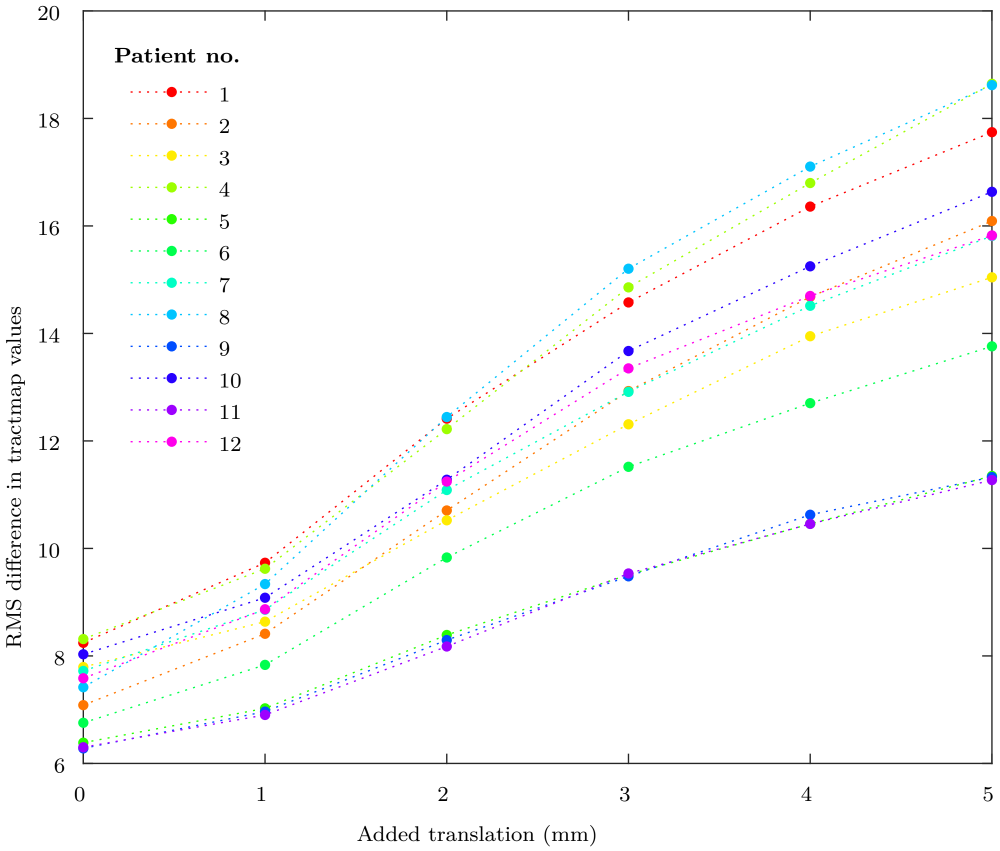


Figure 4.42: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added translation versus added translation for iFOD2.

Figure 4.43 shows the variation in the RMS difference between the modified and unmodified tract-weighted map voxels as a function of added translation for each patient using SD_STREAM. The base level RMS difference between repeat unmodified tract-weighted maps is slightly lower for SD_STREAM than iFOD2, with the majority of patients having an initial RMS difference of about 5. While this is expected, since SD_STREAM reconstructs only structures with high diffusion signals, the difference is not substantial, and not a justified reason for favouring SD_STREAM.

With 1 mm of added translation, the RMS difference between tract-weighted maps for most patients exceeds the baseline level.

The range of RMS difference values also varies significantly between patients. There appear to be two subsets of patients, with one subset (patients 5, 6, 9 and 11) being less susceptible to changes in streamline density caused by added translation, and the remainder of the patients being more susceptible. The magnitude of the RMS differences are similar for the latter group as for iFOD2, indicating that SD_STREAM is similarly susceptible to translation between scan volumes for these patients. For the patients with lower RMS differences, SD_STREAM shows a lower sensitivity to added translation than iFOD2.

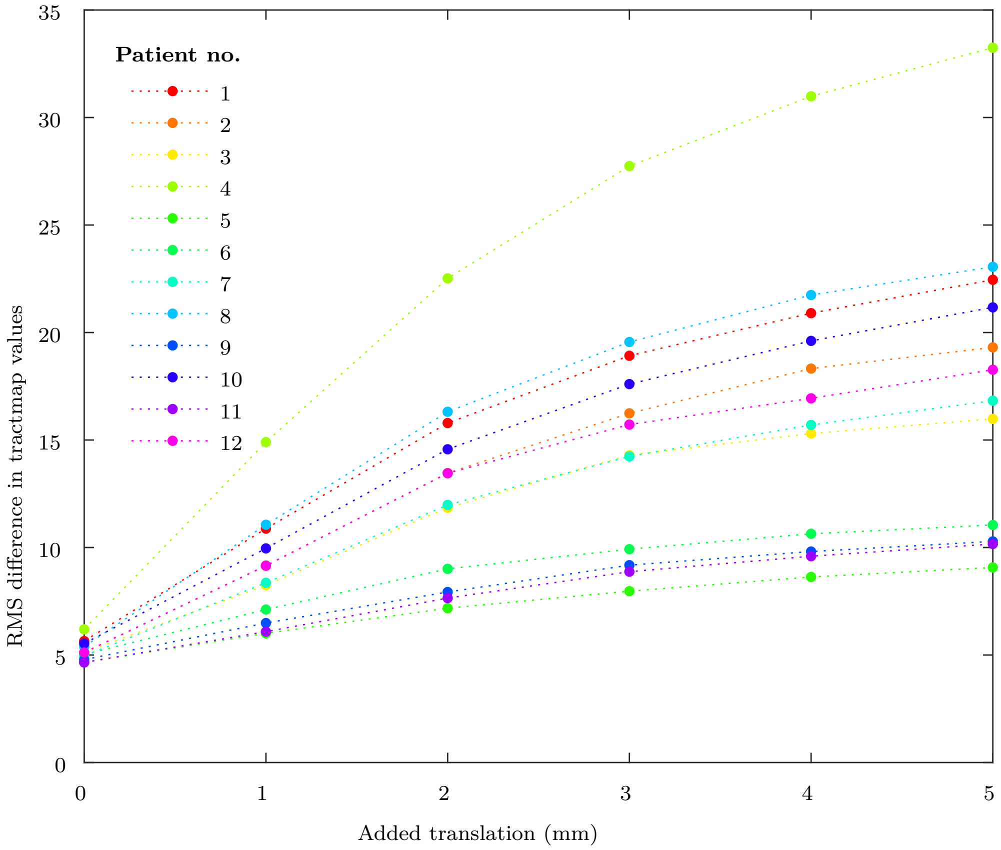


Figure 4.43: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added translation versus added translation for SD_STREAM.

For Tensor_Prob, the baseline RMS difference is higher than both iFOD2 and SD_STREAM. This indicates that the uncertainty in repeat tractography of a given scan is higher for this algorithm. Additionally, with 1 mm added translation, all patients show a RMS difference higher than the baseline, indicating modifications to streamline density that cannot be accounted for by random error alone.

With further translation, the RMS difference progressively increases, flattening off gradually towards 5 mm. The magnitude of the RMS difference is more than a factor of 2 higher than for iFOD2, indicating that this algorithm is most susceptible to false positives and negatives due to translational differences between scan volumes.

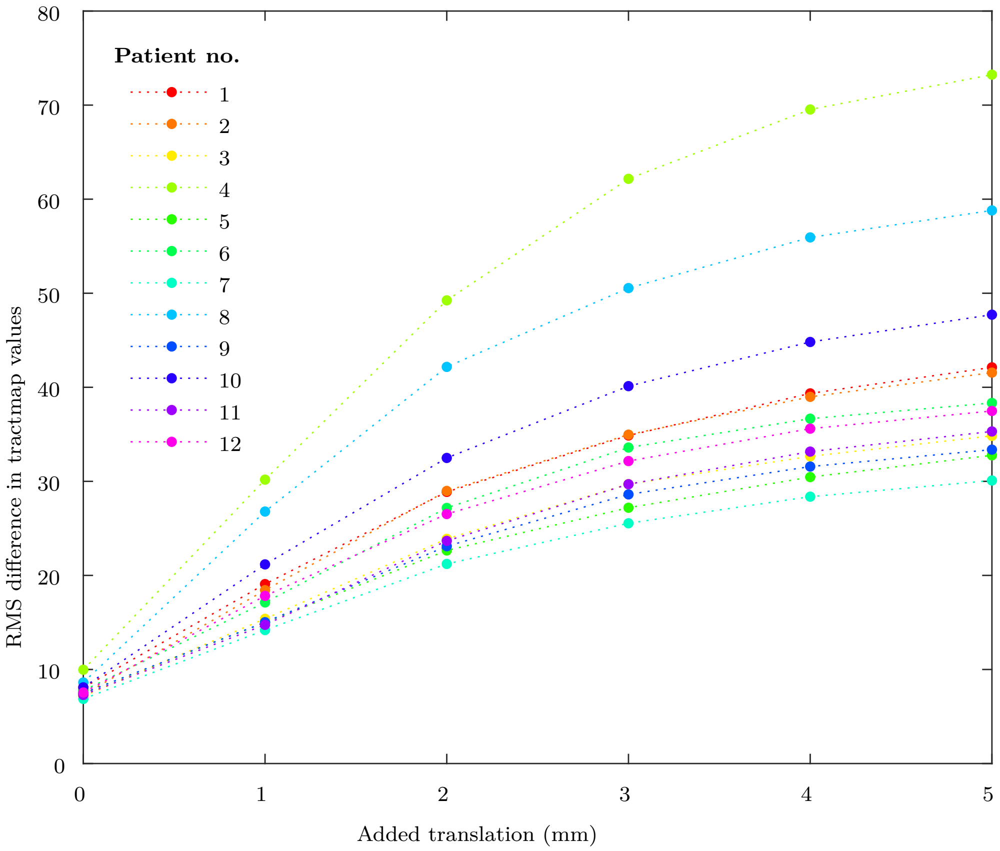


Figure 4.44: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added translation versus added translation for Tensor_Prob.

4.9.2 Rotation

The same approach was applied to the tract-weighted maps generated with added rotation up to 10 degrees. For iFOD2, shown in figure 4.45, the baseline RMS difference is between 6 and 8.5 for all but one patient, which aligns well with that observed for the repeat tract-weighted maps generated for the translation RMS differences. Patient 8 shows an unusually high baseline RMS difference compared to the other patients.

With 1 degree of added rotation, almost all patients have RMS differences below 8. About half of the patients remain below this baseline level with added rotations up to about 3 degrees. Beyond this, the RMS difference progressively increases. The spread in RMS difference values increases with rotation.

Patients 5, 6, 9, and 11 have lower RMS differences than the other patients at equivalent rotations. This suggests these scans were less susceptible to being degraded by added rotation.

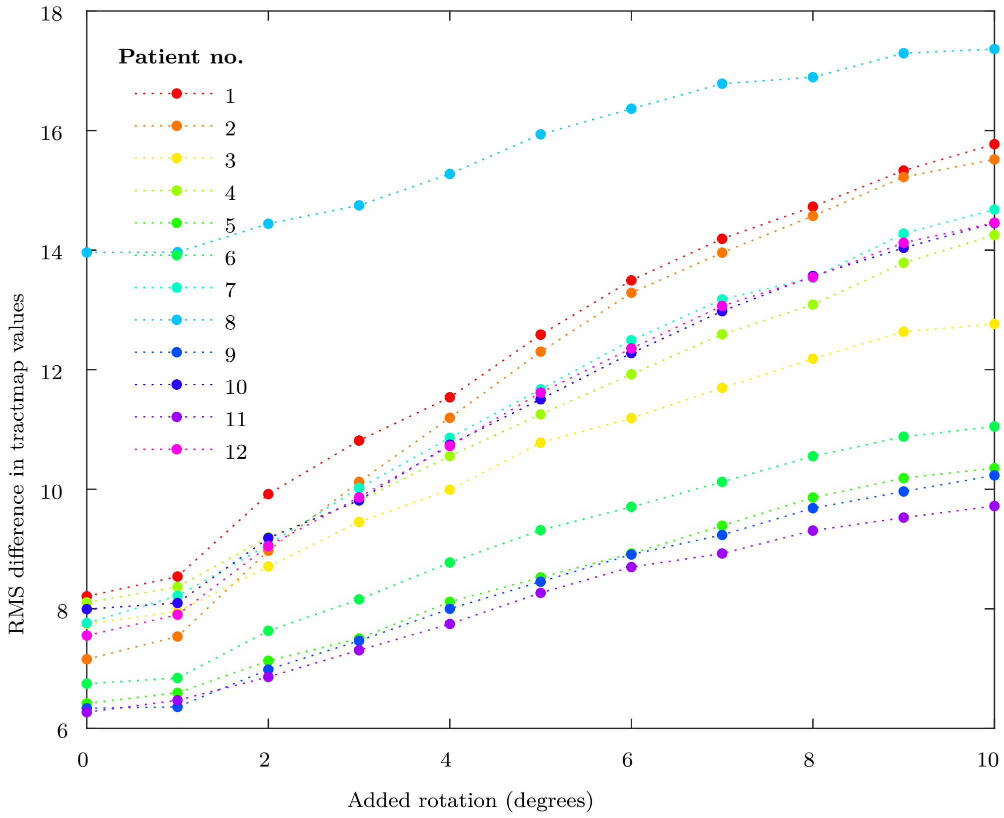


Figure 4.45: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added rotation versus added rotation for iFOD2.

For SD_STREAM, the baseline level RMS difference is consistent with that obtained from the repeat no translation tract-weighted maps. With added rotation, the RMS difference rapidly rises then increases by smaller amounts towards 10 degrees. A small subset of patients (patients 5, 6, 9 and 11) show lower RMS differences than the other patients, as was observed for added translation using SD_STREAM, and for added rotation using iFOD2. Again, the spread widens with larger added rotations.

The magnitude of the RMS difference is in general greater for SD_STREAM than iFOD2 (see Figures 4.46 and 4.45), suggesting that this algorithm is more strongly affected by rotation between scan volumes.

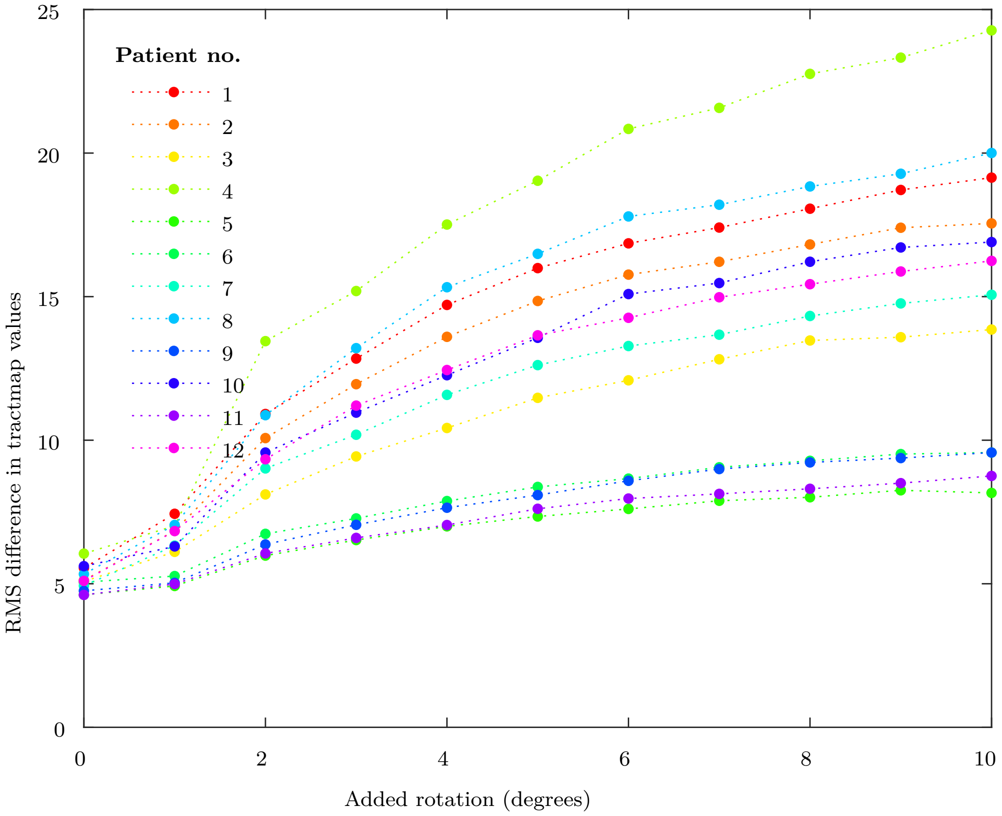


Figure 4.46: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added rotation versus added rotation for SD_STREAM.

For Tensor_Prob a similar trend is seen with increasing tract-weighted map differences for greater rotations (see Figure 4.47). The four patients (patients 5, 6, 9 and 11) which demonstrated lower RMS differences for the iFOD2 and SD_STREAM algorithms are also among the patients with lowest RMS differences for Tensor_Prob.

The baseline RMS difference for Tensor_Prob agrees well with that found by generating repeat tractograms for the translation analysis. With increased rotation, a sharp incline in the RMS difference occurs, between 1 and 2 degrees of added rotation. At higher rotations, the increase in RMS difference is gradual, tapering off towards 10 degrees.

Patient 4 shows particularly high RMS difference values with added rotation for both SD_STREAM and Tensor_Prob.

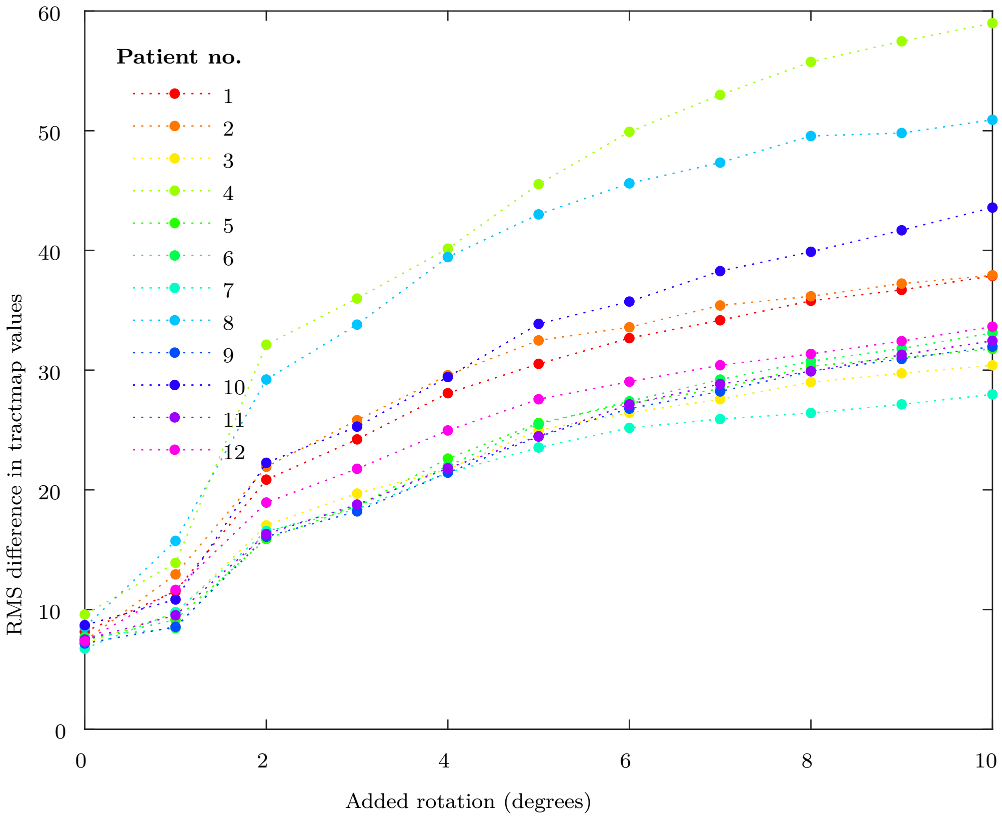


Figure 4.47: Plots of the RMS difference between the unmodified tract-weighted maps and tract-weighted maps with added rotation versus added rotation for Tensor_Prob.

4.10 Change in false positives and false negative rates with added motion

The rates of false positives and false negatives were calculated using the same method as described in section 3.12 with the 50k streamline whole brain tractograms generated for each increment of added translation or rotation.

4.10.1 Translation

For iFOD2, the rates of false positives and negatives are similar with added translation. The rate of increase is initially shallower, then increases approximately linearly beyond 2 mm added translation. At 5 mm of added translation, the percentage of false positives is between 4% and 6% and percentage of false negatives between 3% and 6%.

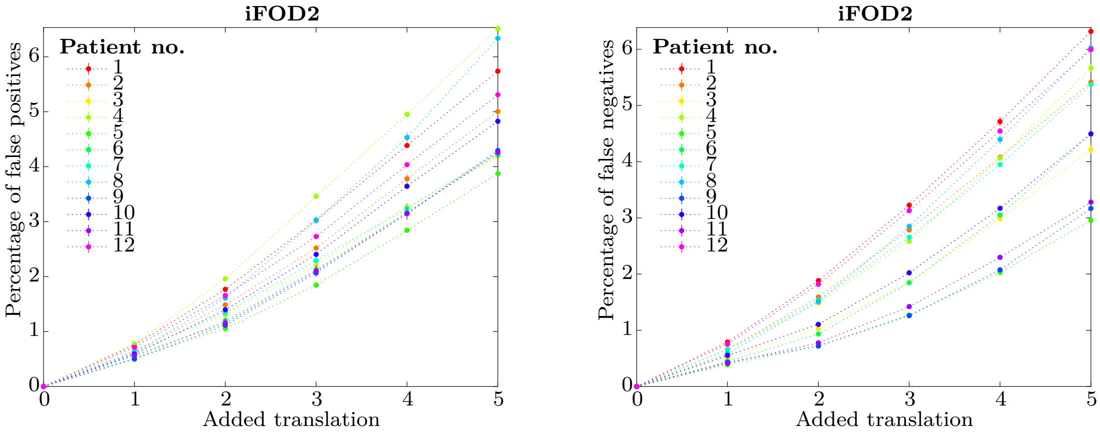


Figure 4.48: False positives (top) and false negatives (bottom) in iFOD2 tractograms versus added translation calculated using a nearest point search method with a threshold of 3 SD.

For SD_STREAM, the false positive rate is higher for some patients than the false negative rate. However, for the majority of patients, the rates are similar to those shown for iFOD2. Despite this, the spread in false positive and negative rates increases more significantly with added translation than for iFOD2.

Two subsets of patients are evident from the plots, with patients 5, 6, 9 and 11 having lower false positive and negative rates than the remainder of patients, which is consistent with the observed lower RMS difference for these patients using SD_STREAM to generate tract-weighted maps.

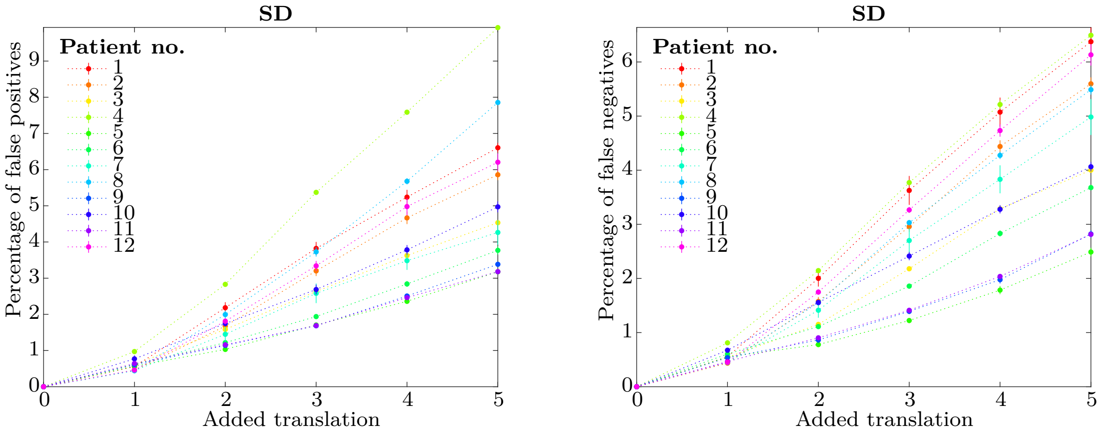


Figure 4.49: False positives (top) and false negatives (bottom) in SD_STREAM tractograms versus added translation calculated using a nearest point search method with a threshold of 3 SD.

The increase in false positive and negative rates for Tensor_Prob also increases gradually for small amount of added translation, then more rapidly for translations higher than 2 mm. Again, the patients 5, 6, 9 and 11 exhibit lower false positive and negative rates than the remainder of patients.

The amount of false positives and negatives is significantly greater for Tensor_Prob than for either iFOD2 or SD_STREAM. A preference towards false positives or negatives is not obvious for Tensor_Prob.

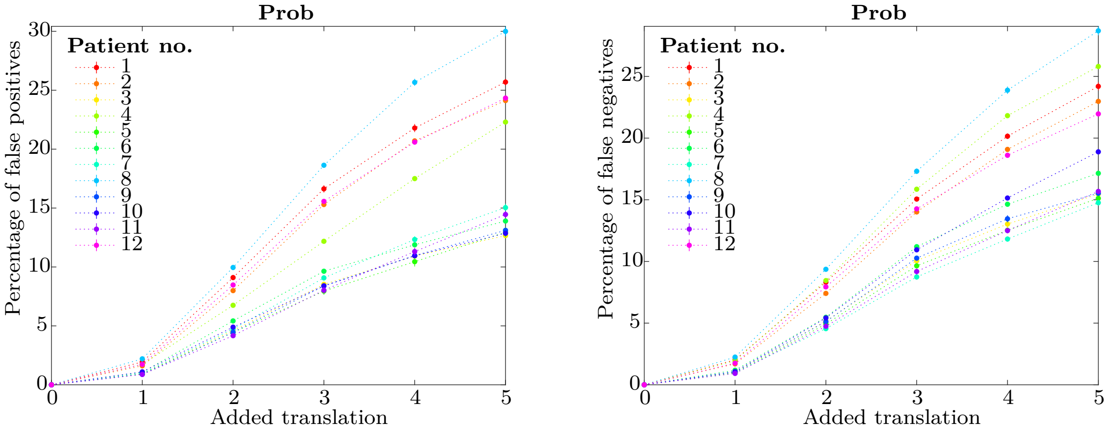


Figure 4.50: False positives (top) and false negatives (bottom) in Tensor_Prob tractograms versus added translation calculated using a nearest point search method with a threshold of 3 SD.

4.11 Rotation

The false positive and negative rates were also computed for each tractogram generated with added clockwise rotation. For iFOD2, a slightly higher rate of false positives is seen than false negatives. Both false positive and negative rates show a steady increase with added rotation, with increasing spread.

Additionally, one subset of patients has false negative rates approximately a factor of 2 higher than the remainder of patients. One possible explanation for this is that the tractograms for those patients with low false negative rates may have consisted primarily of large white matter structures with high diffusion signals, such that small displacements between scan volumes did not significantly reduce the overall signal in the tracts.

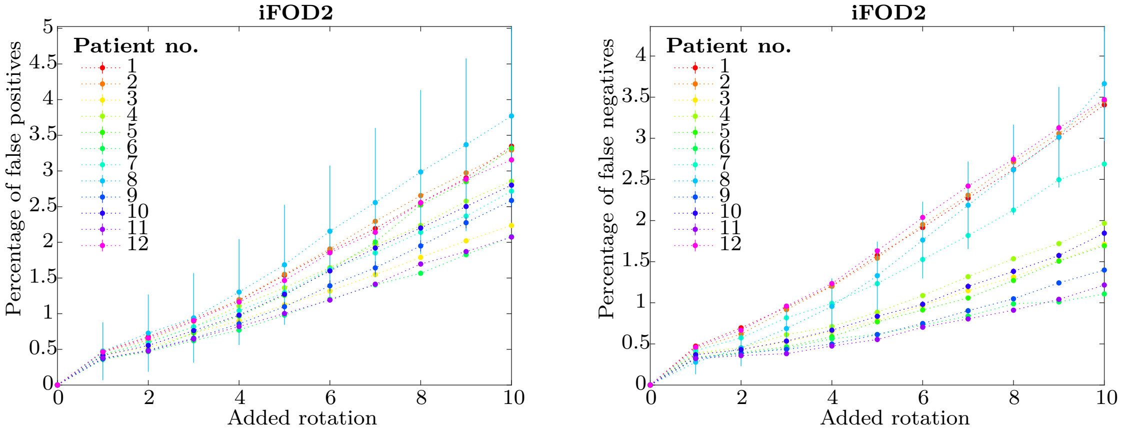


Figure 4.51: False positives (top) and false negatives (bottom) in iFOD2 tractograms versus added rotation calculated using a nearest point search method with a threshold of 3 SD.

For SD_STREAM, the false positive and negative rates are similar in magnitude, and have substantially greater spread at lower rotations than observed for iFOD2, despite the range of maxima (false positive and negative rates at 5 mm) being similar. As has been noted previously, patients 5, 6, 9 and 11 show the least false positives and negatives.

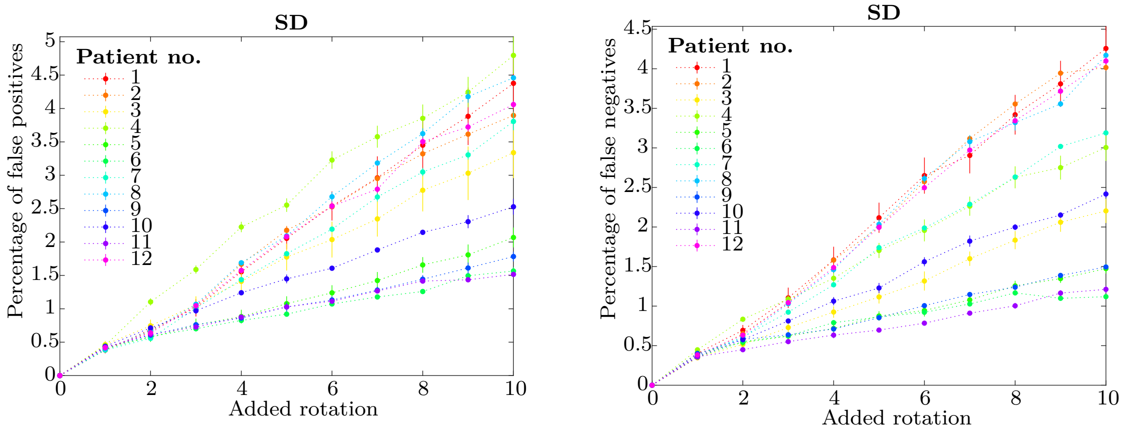


Figure 4.52: False positives (top) and false negatives (bottom) in SD_STREAM tractograms versus added rotation calculated using a nearest point search method with a threshold of 3 SD.

Tensor_Prob produces significantly higher false positive and negative rates than iFOD2 or SD_STREAM. The same division of patients with lower and higher rates are evident as for iFOD2, further supporting the suggestion that there is some feature about the underlying structure of these tracts which makes them less sensitive to added rotation.

The rate of increase of false positives and negatives is slower initially, then increases beyond about 2 mm. Similarly to both iFOD2 and SD_STREAM, the spread in values increases with greater rotation.

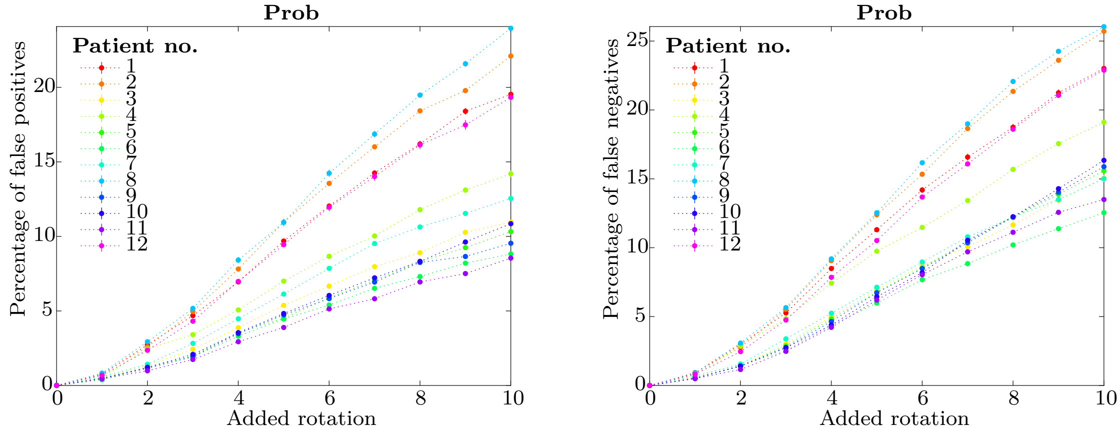


Figure 4.53: False positives (top) and false negatives (bottom) in Tensor_Prob tractograms versus added rotation calculated using a nearest point search method with a threshold of 3 SD.

Added translation or rotation appears to promote false positive and negative rates, through the degree of this dependence varies between patients. In general, through visual inspection of whole brain tractograms, major structures are retained, but become displaced from their true positions by added movement. This suggests that tractography can provide reliable evidence of the presence and strength of white matter connections in the brain, but an alternative imaging modality should be used to provide the relevant structural and spatial information for locating particular features in the brain.

4.12 Summary of the effects of added motion

The effect of the different forms of motion on the quality of the tractogram varies, with deterioration of individual tracts being evident for added translations within a few millimetres, or rotations within a few degrees. Major structures are largely retained for whole brain tractograms with added translations exceeding 5 mm, and rotations beyond 10 degrees. However, structures become warped by the averaging of displaced scan volumes, such that spatial information within the scan is compromised. If multi-modal imaging is used to obtain both diffusion and spatial information it is possible that registration could reinstate the integrity of the tractogram.

The RMS difference between modified and unmodified tract-weighted maps, and the rate of false positives and negatives were found to be the most instructive measures of tractogram changes caused by added translation or rotation. In both cases, the changes to tractogram quality are gradual.

In general, iFOD2 appears to have the most optimal performance with varying degrees of motion present in the scan. It produces fewer false positives and negatives than SD_STREAM or Tensor_Prob, and reports the lowest magnitude differences between modified and unmodified whole brain tract-weighted maps.

This may be a result of the iFOD2 algorithm producing less overshoot in curved paths, as well as its improved handling of streamline dispersion and use of probability in weighting streamline paths, as discussed in section 3.12.

CHAPTER 5

Conclusion

The effect of adding noise or movement to the diffusion magnetic resonance imaging (dMRI) scans on the resulting tractograms was investigated through qualitative and quantitative means. The performance of each of the three MRtrix algorithms considered provides an indication of their relative viability for clinical use. The analysis of the effects of adding varied levels of noise or motion to each input scan is also instructive for quality assurance of patient data in a clinical setting.

Due to the lack of a standardized model for white matter tracts with which to compare patient tractograms, the performance of the MRtrix software was assessed based on its robustness to artificially added noise and patient movement. Increments of artificial noise and motion representing sudden translations and rotations were added to 12 patient dMRI scans. The MRtrix algorithms iFOD2, SD_STREAM and Tensor_Prob were utilised to produce tractograms from the resulting set of input images. Maps of the density of streamlines comprising the tractograms were then generated and individual tracts extracted for analysis.

5.1 Effect of decreased SNR on MRtrix tractography

The deterioration of tractograms with added noise was assessed initially qualitatively using a scoring system devised by McRobbie and Agzerian (2016). A progressive decline in the integrity of 10 main structures was observed in the tractograms as white Gaussian noise (WGN) was added. For the SD_STREAM algorithm, the structures maintained their form with added WGN up to about a power of 20 for the majority of patients. For the probabilistic algorithms iFOD2 and Tensor_Prob, the decline in structure was only noticable beyond about a power of 25 WGN for the majority of tractograms, suggesting that these algorithms are more robust to increased levels of noise. However, the onset of the dip in structure scores varied between patients, as expected due to the pre-existing variations in the signal to noise ratio (SNR) between the dMRI scans. The structure reconstructed with the highest fidelity by all three algorithms was the corpus callosum, followed by the medial lemnisci and pontine fibres.

Tractograms seeded from a single point in the brain stem were read into Matlab and for each patient histograms of streamline length for all noise levels were plotted. The addition of noise promoted premature termination of streamlines in all considered algorithms, due to the increased uncertainty in the diffusion signal triggering the stopping criteria of the tractography algorithms. Estimates of the percentage of false negatives for scans with added noise corroborated this finding. In addition, an increase in the percentage of false positive streamlines was also observed with increasing SNR. It is hypothesised that in these cases, the additional noise in the input scan led to different diffusion paths being connected, causing one tractography path to merge into another. Observing the changes in the tractograms, the added noise also contributed to changes in the distribution and density of streamlines, ultimately inhibiting the algorithms' ability to delineate more intricate diffusion structures.

The correlation between noise and tract deterioration was assessed by applying a range of different metrics to ascertain a particular SNR at which tractograms become unreliable for each algorithm and patient. By combining the results of polar plots, false positive and negative rates, resemblance measures and qualitative scoring, input diffusion tensor magnetic resonance imaging (dtMRI) scans with a SNR of greater than 8 (calculated using the method described in section 3.3) were found to be more likely to produce reliable tractograms (i.e. tractograms that most closely resembled the tractogram derived from the unmodified scan), provided the effects of patient movement or artefacts are negligible. This value was ascertained by using the relationship between added WGN and SNR depicted in Figure 3.6 to map the metrics calculated at each WGN level to SNR. The value of 8 was found to be relatively consistent across patients 1-4, 7, 8, 10 and 12 and between the two probabilistic tractography algorithms. For patients 5, 6, 9 and 11 the metrics only demonstrated worsening performance once the SNR dropped below about 4.

The resemblance metric and RMS difference between tract-weighted maps indicated that the probabilistic algorithms perform consistently with added noise up to about WGN 20, while the deterministic SD_STREAM demonstrated a continuous decrease in resemblance and increase in RMS difference with added noise. Therefore, it was not possible to discern a clear threshold SNR that would indicate whether a scan was clinically viable for use with SD_STREAM. SD_STREAM appeared to consistently underperform with respect to iFOD2 and Tensor_Prob, suggesting it is highly susceptible to changes in noise. Based on the results obtained, iFOD2 stood out as the optimal algorithm for repeatably reproducing tractograms with the highest level of detail, and the fewest false positives or false negatives with added noise.

5.2 Effect of motion on MRtrix tractography

The addition of motion, in the form of sudden translations and rotations in the slice plane between scan volumes, was also anticipated to lead to a deterioration of tractogram structures. This was observed for individual tracts with added translations within a few millimetres, or rotations within a few degrees. However, for whole brain tractograms, major structures were largely retained with added translations exceeding 5 mm, and rotations beyond 10 degrees. However, the structures became warped by the averaging of the displaced scan volumes. For clinical applications, multi-model imaging may be able to correct for the spatial displacement of structures by rotation.

The addition of motion resulted in a significant rise in the percentage of false positives compared with an unmodified tractogram for each patient. This is suspected to be in part due to the warping of the tractogram by added motion. That is, if a structural scan was available with which the tract density map could be registered prior to computing the false positive and false negative rates, it is expected a significant drop in false positive rates would be observed.

The RMS difference between the original tract-weighted maps and the tract-weighted maps produced from scans with added motion was found to be the most informative measure of changes caused by added translation or rotation. The shapes of the responses to motion for the three tractography algorithms were similar. However, iFOD2 was found to be superior to SD_STREAM and Tensor_Prob, as it produced fewer false positives and negatives, and yielded whole brain tract-weighted maps which deviated the least from the originals (i.e.

had the lowest magnitude RMS differences between modified and unmodified tract-weighted maps).

5.3 Limitations and further work

This study encompasses a number of limitations, the most notable of which are the limitations in the data itself. All subjects considered in this study underwent a dMRI scan for assessment of a particular neurological condition. Consequently, all unmodified tractograms that were used to define the baseline tractography performance in this study were compromised to some extent. Most of the metrics used in this study attempted to measure a change in the initial state of the scan, so the absence of features in the original scans was not of strong importance. However, this particularly affected the results obtained from qualitative scoring.

As this is a retrospective study of existing patient scans, the acquisition parameters used to obtain each scan also differed. In particular, a range of b-values were used for different subjects, which could impact the tractography performance irrespective of noise or motion effects. However, McRobbie and Agzarian (McRobbie and Agzarian, 2016) demonstrated that b-values as low as 1000 were sufficient to perform constrained spherical deconvolution (CSD). The dependence of the SNR on acquisition parameters such as number of signal averages, diffusion directions, and geometric scan parameters was also not considered in this study. It is therefore advised that caution is taken in interpreting the SNR results presented here.

Additionally, the effects of increased noise, rotation or translation were considered in isolation. In the case of motion, artifical changes were also only considered along one axis (translation) or one direction (rotation) in the plane of the dMRI slices. In reality, it is reasonable to expect patient movement to consist of a combination of these and other forms of motion in any direction, with some perhaps being more probable than others, as well as a level of noise that will depend on a variety of factors. To assess how these factors affect tractography in combination, experimental data representative of the possible range of noise levels and degrees of patient movement anticipated to occur in a clinical setting should be obtained. This is essential to analyse the effects of particular forms of motion, such as motion tangential to the slice plane, as it is impractical to artificially simulate this.

The form of noise added in this study was WGN. However, the true noise present in the signal obtained in a dMRI is affected by a number of different factors. Echo-planar-imaging based acquisitions are particularly susceptible to distortions from eddy currents or inhomogeneities in the magnetic field (Eichner et al., 2020). Therefore, while the results of this study suggest a clear cutoff for the amount of noise that may be present in a scan and the image still be viable, this value may differ for experimental data.

The actual patient movement during each of the dMRI acquisitions will also impact the diffusion signals. The signal intensities from these original scans are intact in this retrospective analysis, however, the positional information has changed. This is a limitation of the study.

A possible source of error in the method of specifying a seed point for reconstruction of a single tract is the choice of the seed point for the tract for each patient, particularly given that the radius of the seed volume was significantly smaller than the radius of the brain stem generated in a randomly seeded full brain tractogram. Although this ensured that only streamlines belonging to the brain stem were selected, if the distribution of streamlines at the base of the brain stem was non-uniform, this may have produced artefacts such as early terminated streamlines in the final reconstruction. A method of correcting for this would be to sample a number of points at the base of the brainstem and perform a repeated analysis for each seed point, or alternatively to widen the radius of the seed volume to encompass the entirety of the brain stem base.

Factors other than noise and patient motion may influence how well the MRtrix software performs tractography by compromising the input dMRI. For instance, artefacts in the dMRI could create distortions which would likely lead to the tractography algorithms delineating tracts in displaced locations. MRtrix has built-in tools for correcting geometric distortions, eddy currents and inhomogeneities present in diffusion images. A tool for denoising input dMRI scans has also been implemented into MRtrix. A comprehensive analysis of MRtrix's clinical validity should therefore include an assessment of how well the available tools are capable of correcting for noise and/or distortions. Failing this, an important step in validating this software would be to develop a method of identifying any features of a dMRI scan that could compromise the accuracy of the resulting tractogram. Therefore, the use of significantly compromised dMRI scans could be avoided.

Another aspect of tractography that has been largely unexplored is the computational detection of abnormalities in white matter. With the collection of sufficient data, it is conceivable that this could be achieved through machine learning, using dMRI scans and either qualitative scores or some other metric to inform the algorithm whether a feature is typical or atypical.

This study has demonstrated the robustness of MRtrix tractography, in particular that of the algorithm iFOD2, to artificial variations in SNR and patient movement. This information supports the use of MRtrix in a clinical setting, with the high level of detail that is achievable using this software potentially providing a means of significantly improving patient outcomes.

References

Aboitiz, F., Scheibel, A., Fisher, R. and Zaidel, E. (1992), 'Fiber composition of the human corpus callosum', *Brain Research* **598**, 143–153.

Alexander, A., Hassan, K., Kindlmann, G., Parker, D. and Tsuruda, J. (2000), 'A geometric analysis of diffusion tensor measurements of the human brain', *Magnetic Resonance in Medicine* .

Anderson, A. (2001), 'Theoretical analysis of the effects of noise on diffusion tensor imaging', *Magnetic Resonance in Medicine* **46**, 1174–1188.

Assaf, Y., Johansen-Berg, H. and Thiebaut de Schotten, M. (2017), 'The role of diffusion mri in neuroscience', *NMR in Biomedicine* **32**.

Assaf, Y. and P.J., B. (2005), 'Composite hindered and restricted model of diffusion (charmed) mr imaging of the human brain', *Neuroimage* **27**, 48–58.

Basser, P., Mattiello, J. and Le Bihan, D. (1994), 'Mr diffusion tensor spectroscopy and imaging', *Biophysical Journal* **66**, 259–267.

Basser, P. and Pierpaoli, C. (1996), 'Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri', *Journal of Magnetic Resonance* **213**, 560–570.

Beaulieu, C. (2002), 'The basis of anisotropic water diffusion in the nervous system – a technical review', *NMR in Biomedicine* **15**, 435–455.

Behrens, T., Berg, H., Jbabdi, S., Rushworth, M. and Woolrich, M. (2007), 'Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?', *NeuroImage* **34**(1), 144–155.

Budde, M. and Annese, J. (2013), 'Quantification fo anisotropy and fiber orientation in human brain histological sections', *Frontiers in Integrative Neuroscience* **7**.

Burdette, J., Durden, D., Elster, A. and Yen, Y. (2001), 'High b-value diffusion-weighted mri of normal brain', *Journal of Computer Assisted Tomography* **25**, 515–519.

Bushberg, J., Seibert, J., Leidholdt Jr, E. and Boone, J. (2012), *The essential phyics of medical imaging*, 3 edn, Lippincott Williams and Wilkins, Philadelphia.

Chenevert, T., Brunberg, J. and Pipe, J. (1990), 'Anisotropic diffusion in human white matter: demonstration with mr techniques in vivo', *Radiology* **177**(2).

Chilla, G., Tan, C., Xu, C. and Poh, C. (2015), 'Diffusion weighted magnetic resonance imaging and its recent trend – a survey', *Quantitative Imaging in Medicine and Surgery* **5**, 407–422.

Conturo, T., N.F., L., T.S., C., E., A., Snyder, A., Shimony, J., McKinstry, R., Burton, H. and Raichle, M. (1999), 'Tracking neuronal fiber pathways in the living human brain', *Proceedings of the National Academy of Sciences of the United States of America* **96**, 10422–10427.

DellAcqua, F., Simmons, A., Williams, S. and Catani, M. (2013), 'Can spherical deconvolution provide more information than fiber orientations? hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion', *Human Brain*

Mapping **34**, 2464–2483.

DellAcqua, F. and Tournier, J. (2018), ‘Modelling white matter with spherical deconvolution: How and why?’, *NMR in Biomedicine* **32**.

Descoteaux, M., Deriche, R., Knosche, T. and Anwander, A. (2009), ‘Deterministic and probabilistic tractography based on complex fibre orientation distributions’, *IEEE Transactions on Medical Imaging* **28**, 269–286.

Dietrich, O., Raya, J., Reeder, S., Reiser, M. and Schoenberg, S. (2007), ‘Measurement of signal-to-noise ratios in mr images: Influence of multichannel coils, parallel imaging, and reconstruction filters’, *Journal of Magnetic Resonance Imaging* **26**, 375–385.

Douaud, G., Jbabdi, S., Behrens, T., Menke, R., Gass, A., Monsch, A., Rao, A., Whitcher, B., Kindlmann, G., Matthews, P. and S., S. (2011), ‘Dti measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in mci and mild alzheimer’s disease’, *NeuroImage* **55**, 880–890.

Eichner, C., Paquette, M., Mildner, T., Schlumm, T., Pleh, K., Samuni, L., Crockford, C., Wittig, R., Jager, C., Moller, H., Friederici, A. and Anwander, A. (2020), ‘Increased sensitivity and signal-to-noise ratio in diffusion-weighted mri using multi-echo acquisitions’, *NeuroImage* **221**.

Einstein, A. (1956), ‘Investigations on the theory of the brownian movement’, *Courier Corportion* .

Frank, L. (2002), ‘Characterization of anisotropy in high angular resolution diffusion-weighted mri’, *Magnetic Resonance in Medicine* **47**, 1083–99.

Glenn, G., Kuo, L., Chao, Y., Lee, C., Helpern, J. and Jensen, J. (2016), ‘Mapping the orientation of white matter fiber bundles: A comparative study of diffusion tensor imaging, diffusional kurtosis imaging, and diffusion spectrum imaging’, *American Journal of Neuroradiology* **37**, 1216–1222.

Godenschweger, F., Kagebein, U., Stucht, D., Yarach, U., Sciarra, A., Yakupov, R., Lusebrink, F., Schulze, P. and Speck, O. (2016), ‘Motion correction in mri of the brain’, *Physics in Medicine and Biology* **61**.

Gonzalez, R. (2012), ‘Clinical mri of acute ischemic stroke’, *Journal of Magnetic Resonance Imaging* **36**, 259–271.

Gudbjartsson, H. and Patz, S. (1995), ‘The rician distribution of noisy mri data’, *Magnetic Resonance in Medicine* **34**, 910–914.

Havsteen, I., Ohlhues, A., Madsen, K., Nybing, J., Christensen, H. and Christensen, A. (2017), ‘Are movement artifacts in magnetic resonance imaging a real problem? - a narrative review’, *Frontiers in Neurology* **8**.

Hess, C., Mukherjee, P., Han, E., D., X. and D.B., V. (2006), ‘Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis’, *Magnetic Resonance in Medicine* **56**, 104–117.

Hosey, T., Williams, G. and Ansorge, R. (2005), ‘Inference of multiple fiber orientation in high angular resolution diffusion imaging’, *Magnetic Resonance in Medicine* **54**, 1480–1489.

Huang, H., Zhang, J., Zijl, P. and Mori, S. (2004), ‘Analysis of noise effects on dti-based tractography using the brute-force and multi-roi approach’, *Magnetic Resonance in Medicine* **52**.

Hubbard, P. and Parker, G. (2009), ‘Validation of tractography’, *Diffusion MRI* pp. 353–375.

Hutchinson, E., Avram, A., Irfanoglu, M., Koay, C., Barnett, A., Komlosh, M., Ozarslan, E., Schwerin, S., Juliano, S. and Pierpaoli, C. (2017), 'Analysis of the effects of noise, dwi sampling and value of assumed parameters in diffusion mri models', *Magnetic Resonance in Medicine* **28**.

Jansons, K. and Alexander, D. (2003), 'Persistent angular structure: new insights from diffusion magnetic resonance imaging data', *Information Processing in Medical Imaging* **18**, 672–683.

Jeurissen, B., Descoteau, M., Mori, S. and Leemans, A. (2017), 'Diffusion mri fiber tractography of the brain', *NMR in Biomedicine* **32**.

Jeurissen, B., Leemans, A., Tournier, J., Jones, D. and Sijbers, J. (2012), 'Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging', *Human Brain Mapping* **34**, 2747–2766.

Jeurissen, B., Leemans, A., Tournier, J., Jones, D. and Sijbers, J. (2013), 'Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging', *Human Brain Mapping* **34**, 2747–2766.

Jones, D. (2008), 'Tractography gone wild: Probabilistic tracking using the wild bootstrap', *IEEE Transactions on Medical Imaging* pp. 1268 – 1274.

Jones, D., Knosche, T. and Turner, R. (2013), 'White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion mri', *NeuroImage* pp. 239–254.

Lampinen, B., Szczepankiewicz, F., Martensson, J., van Westen, D., Sundgren, P. and Nilsson, M. (2016), 'Neurite density imaging versus imaging of microscopic anisotropy in diffusion mri: A model comparison using spherical tensor encoding', *Neuroimage* **147**, 517–531.

Le Bihan, D. (2014), 'Diffusion mri: what water tells us about the brain', *EMBO Molecular Medicine* **6**(5), 569–573.

Leiva-Salinas, C. and Wintermark, M. (2010), 'Imaging of ischemic stroke', *Neuroimaging Clinics of North America* **20**, 455–468.

McClymont, D., Teh, I. and Schneider, J. (2017), 'The impact of signal-to-noise ratio, diffusion-weighted directions and image resolution in cardiac diffusion tensor imaging – insights from the ex-vivo rat heart', *Journal of Cardiovascular Magnetic Resonance* **2017**, 19–90.

McRobbie, D. and Agzarian, M. (2016), Quality control measures for constrained spherical deconvolution mr diffusion tractography in clinical use, *in* 'ISMRM 24th Annual Meeting and Exhibition', 1205, Singapore.

McRobbie, D., Moore, E., Graves, M. and Prince, M. (2006), *MRI from picture to proton*, 2 edn, Cambridge University Press.

Minati, L. and Weglarz, W. (2007), 'Physical foundations, models and methods of diffusion magnetic resonance imaging of the brain: A review', *Concepts in Magnetic Resonance Part A* pp. 278–307.

Moldrich, R., Pannek, K., Hoch, R., Rubenstein, J., Kurniwan, N. and Richards, L. (2010), 'Comparative mouse brain tractography of diffusion magnetic resonance imaging', *NeuroImage* **51**, 1027–1036.

Mollink, J., Kleinnijenhuis, M., van Cappellen van Walsum, A., Sotiropoulos, S., Cottaar, M., Mirfin, C., Heinrich, M., Jenkinson, M., Pallebage-Gamarallage, M., Ansorge, O., Jbabdi, S. and Millera, K. (2017), 'Evaluating fibre orientation dispersion in white matter: Comparison of diffusion mri, histology and polarized light imaging', *Neuroimage* **157**, 561–

574.

Mori, S. and J.D., T. (2014), *Introduction to Diffusion Tensor Imaging (Second Edition)*, Academic Press, chapter Chapter 8 - Moving Beyond DTI: High Angular Resolution Diffusion Imaging (HARDI).

Moseley, M., Kucharczyk, J., Mintorovitch, J., Cohen, Y., Kurhanewicz, J., Derugin, N., Asgari, H. and Norman, D. (1990), 'Diffusion-weighted mr imaging of acute stroke: correlation with t2-weighted and magnetic susceptibility-enhanced mr imaging in cats', *American Journal of Neuroradiology* **11**(3), 423–429.

NEMA, R. (2008), 'Nema standard publication ms 1-2008 (r2014): Determination of signal-to-noise ratio (snr) in diagnostic magnetic resonance imaging', *National Electrical Manufacturers Association* .

Pierpaoli, C. and Basser, P. (1996), 'Toward a quantitative assessment of diffusion anisotropy', *Magnetic Resonance in Medicine* **36**(6), 893–906.

Reese, T., Wedeen, V. and Weisskoff, R. (1996), 'Measuring diffusion in the presence of material strain', *Journal of Magnetic Resonance series B* **112**(3), 258–258.

Roine, T., Jeurissen, B., D., P., Aelterman, J., Leemans, A., W., P. and Sijbers, J. (2014), 'Isotropic non-white matter partial volume effects in constrained spherical deconvolution', *Frontiers in Neuroinformatics* **8**.

Roine, T., Jeurissen, B., D., P., Aelterman, J., Philips, W., Leemans, A. and Sijbers, J. (n.d.), 'Informed constrained spherical deconvolution (icsd)', *Medical Image Analysis* **24**, 269–281.

Rydhog, A., Szczepankiewicz, F., Wirestam, R., Ahlgren, A., Westin, C., Knutsson, L. and Pasternak, O. (2017), 'Separating blood and water: Perfusion and free water elimination from diffusion mri in the human brain', *Neuroimage* **156**, 423–434.

Schilling, K., Janve, V., Gao, Y., Stepniewska, I., Landman, B. and Anderson, A. (2016), 'Comparison of 3d orientation distribution functions measured with confocal microscopy and diffusion mri', *Neuroimage* **129**, 185–197.

Smith, R., Tournier, J., Calamante, F. and Connelly, A. (2012), 'Anatomically-constrained tractography: Improved diffusion mri streamlines tractography through effective use of anatomical information', *NeuroImage* **62**, 1924–1938.

Toselli, B., Tortora, D., Severino, M., Arnulfo, G., Canessa, A., Morana, G., Rossi, A. and Fato, M. (2017), 'Improvement in white matter tract reconstruction with constrained spherical deconvolution and track density mapping in low angular resolution data: A pediatric study and literature review', *Frontiers in Pediatrics* **5**.

Tournier, J., Calamante, F. and Connelly, A. (2007), 'Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution', *NeuroImage* **35**, 1459–1472.

Tournier, J., Calamante, F. and Connelly, A. (2012), 'Mrtrix: Diffusion tractography in crossing fiber regions', *International Journal of Imaging Systems and Technology* **22**, 55–66.

Tournier, J., Calamante, F., Gadian, D. and Connelly, A. (2004), 'Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution', *NeuroImage* **23**, 1176–1185.

Tournier, J., Mori, S. and Leemans, A. (2011), 'Diffusion tensor imaging and beyond', *Magnetic Resonance in Medicine* **65**, 1532 – 1556.

Tournier, J., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C. and Connelly, A. (2019), ‘Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation’, *NeuroImage* **202**, 116–137.

Tournier, J., Yeh, C., Calamante, F., Cho, K., Connelly, A. and Lin, C. (2008), ‘Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data’, *Neuroimage* **42**, 617–625.

Tuch, D. (2004), ‘Q-ball imaging’, *Magnetic Resonance in Medicine* **52**, 1358–1372.

Tuch, D., Reese, T., Wiegell, M., Makris, N., Belliveau, J. and Wedeen, V. (2002), ‘High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity’, *Magnetic Resonance in Medicine* **48**.

Wakana, S., Caprihan, A., Panzenboeck, M., Fallon, J., Perry, M., Gollub, R., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P. and Mori, S. (2007), ‘Reproducibility of quantitative tractography methods applied to cerebral white matter’, *Neuroimage* **36**, 630–644.

Wheeler-Kingshott, C. and Cercignani, M. (2009), ‘About “axial” and “radial” diffusivities’, *Magnetic Resonance in Medicine* **61**.

Yeh, C., Tournier, J., Cho, K., Lin, C., Calamante, F. and Connelly, A. (2010), ‘The effect of finite diffusion gradient pulse duration on fibre orientation estimation in diffusion mri’, *Neuroimage* **51**, 743–751.

Zaitsev, M., Maclaren, J. and Herbst, M. (2015), ‘Motion artefacts in mri: A complex problem with many partial solutions’, *Journal of Magnetic Resonance Imaging* **42**, 887–901.

Zhan, W. and Yang, Y. (2006), ‘How accurately can the diffusion profiles indicate multiple fiber orientations? a study on general fiber crossings in diffusion mri’, *Journal of Magnetic Resonance* **183**, 193–202.