Brain Imaging and Beyond

Sofia Olhede[1,2,3]

[1]Department of Statistical Science, University College London, UK

[2]Department of Computer Science, University College London, UK

[3]Institute for Mathematical Sciences, Imperial College London, UK.

St Mary's College, Durham, UK, 14.00

Outline

- Why do Statistics?

Outline

- Why do Statistics?
- Some of my research interests.

Outline

- Why do Statistics?
- Some of my research interests.
- PhD in Statistics in the UK.

Outline

- Why do Statistics?
- Some of my research interests.
- PhD in Statistics in the UK.
- PhD Opportunities at UCL's Department of Statistical Science.

- Compared to most branches of Mathematics, Statistics is a relatively young discipline, few important results date back to much earlier than 1900.

- Compared to most branches of Mathematics, Statistics is a relatively young discipline, few important results date back to much earlier than 1900.
- Developments are data driven, with a rapid change of focus resulting from this fact.

- Compared to most branches of Mathematics, Statistics is a relatively young discipline, few important results date back to much earlier than 1900.
- Developments are data driven, with a rapid change of focus resulting from this fact.
- Strong current influences come from innovations in data acquisition in biochemistry/biology and neuroscience. Great volumes of data are generated.

- Compared to most branches of Mathematics, Statistics is a relatively young discipline, few important results date back to much earlier than 1900.
- Developments are data driven, with a rapid change of focus resulting from this fact.
- Strong current influences come from innovations in data acquisition in biochemistry/biology and neuroscience. Great volumes of data are generated.
- In 2007 281 billion *gigabytes* of data was generated. In terms of digital bits this corresponds to a number exceeding the number of stars in the universe or Avogadro's number.

- A great proportion of this data is unfortunately without value, or will never be used, due to poor experimental design.

- A great proportion of this data is unfortunately without value, or will never be used, due to poor experimental design.
- Soon storage of raw data will become unfeasible, necessitating on-line analysis and data compression *before* storage.

- A great proportion of this data is unfortunately without value, or will never be used, due to poor experimental design.
- Soon storage of raw data will become unfeasible, necessitating on-line analysis and data compression *before* storage.
- The need for careful data collection and analysis has **never** been greater.

- A great proportion of this data is unfortunately without value, or will never be used, due to poor experimental design.
- Soon storage of raw data will become unfeasible, necessitating on-line analysis and data compression *before* storage.
- The need for careful data collection and analysis has **never** been greater.
- John Tukey once said “The best thing about being a statistician is that you get to play in everyone’s backyard.”

Data

- Brain imaging has many characteristics of typical problems facing modern statistics.

Data

- Brain imaging has many characteristics of typical problems facing modern statistics.
- An enormous amount of data is collected, with a typical single MRI scan corresponding to $16 \times 60 \times 128^2 \times 30 \approx 400$ million data points.

Data

- Brain imaging has many characteristics of typical problems facing modern statistics.
- An enormous amount of data is collected, with a typical single MRI scan corresponding to $16 \times 60 \times 128^2 \times 30 \approx 400$ million data points.
- "Few of us expect to ever see a man who has analyzed, or even handled, a sequence of a million numerical values …"
  J. W. Tukey, (1959).

Data

- Brain imaging has many characteristics of typical problems facing modern statistics.
- An enormous amount of data is collected, with a typical single MRI scan corresponding to $16 \times 60 \times 128^2 \times 30 \approx 400$ million data points.
- "Few of us expect to ever see a man who has analyzed, or even handled, a sequence of a million numerical values …"
  J. W. Tukey, (1959).
- The hypotheses are sometimes difficult to formalize, and the temporal sampling very sparse.

MRI

- MRI technology is based on using a very powerful magnet (often 3 Tesla or more).

MRI

- MRI technology is based on using a very powerful magnet (often 3 Tesla or more).
- This magnet is used to align water molecules.

MRI

- MRI technology is based on using a very powerful magnet (often 3 Tesla or more).
- This magnet is used to align water molecules.
- A radio-wave makes the molecules resonate, and during this phase absorb energy.

MRI

- MRI technology is based on using a very powerful magnet (often 3 Tesla or more).
- This magnet is used to align water molecules.
- A radio-wave makes the molecules resonate, and during this phase absorb energy.
- Extra magnetic fields are created by gradient coils and these may be used to moderate the magnetic field.

MRI

- MRI technology is based on using a very powerful magnet (often 3 Tesla or more).
- This magnet is used to align water molecules.
- A radio-wave makes the molecules resonate, and during this phase absorb energy.
- Extra magnetic fields are created by gradient coils and these may be used to moderate the magnetic field.
- A scanner can then measure the energy released by the molecules.

Page 1 of

k-space

- The demodulated MR signal is equivalent to the Fourier Transform (FT) of the effective spin density (see [3, 6]) at each voxel.

k-space

- The demodulated MR signal is equivalent to the Fourier Transform (FT) of the effective spin density (see [3, 6]) at each voxel.
- Because of this, to obtain an estimate of the density, we can take the Inverse FT of the measurements in each orientation.

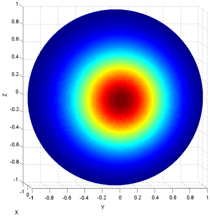
k-space

- The demodulated MR signal is equivalent to the Fourier Transform (FT) of the effective spin density (see [3, 6]) at each voxel.
- Because of this, to obtain an estimate of the density, we can take the Inverse FT of the measurements in each orientation.
- The inference question of interest is usually related to the “connectivity” of the brain, and how pathologies change connections.

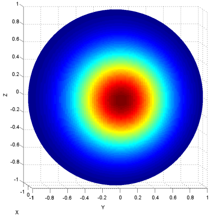
k-space

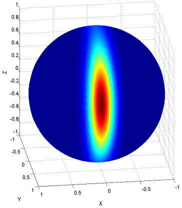
- The demodulated MR signal is equivalent to the Fourier Transform (FT) of the effective spin density (see [3, 6]) at each voxel.
- Because of this, to obtain an estimate of the density, we can take the Inverse FT of the measurements in each orientation.
- The inference question of interest is usually related to the “connectivity” of the brain, and how pathologies change connections.
- In neurodegenerative diseases the directionality of the density usually becomes less clear.

Diffusions in Space

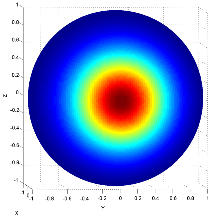


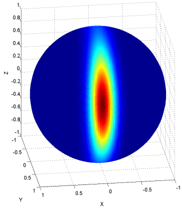
Diffusions in Space

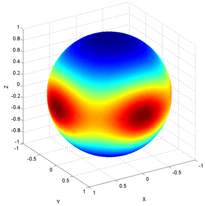




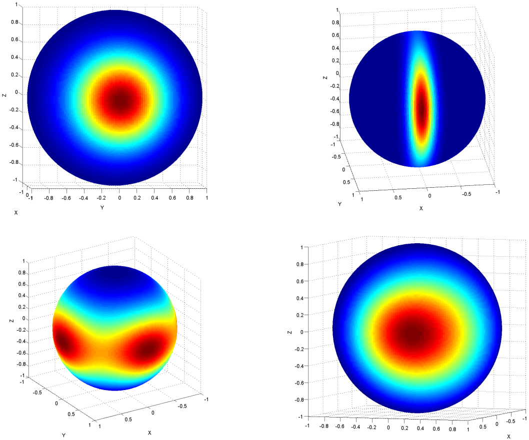
Diffusions in Space



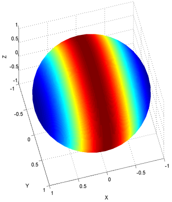




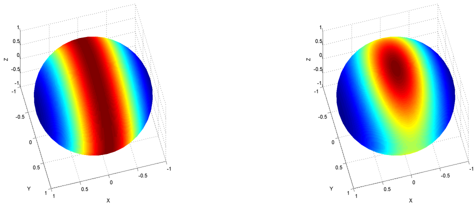
Diffusions in Space



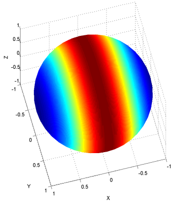
Diffusions in Frequency

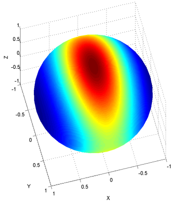


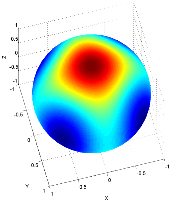
Diffusions in Frequency



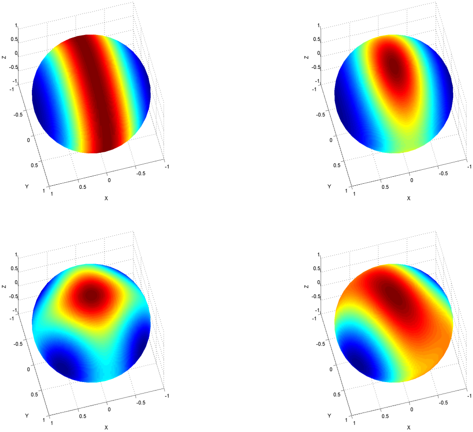
Diffusions in Frequency







Diffusions in Frequency



Modelling

- About 6–60 directions are used at each frequency (if you can leave something dead in the scanner you can get substantially more).

Modelling

- About 6–60 directions are used at each frequency (if you can leave something dead in the scanner you can get substantially more).
- If we wanted to numerically invert the density, substantially more measurements are required. We need to model the observations [1].

Modelling

- About 6–60 directions are used at each frequency (if you can leave something dead in the scanner you can get substantially more).
- If we wanted to numerically invert the density, substantially more measurements are required. We need to model the observations [1].
- Usually a Gaussian density is assumed in 3-D with zero mean (no net diffusion), based on truncating the three-dimensional Bolzmann equation.

Modelling

- About 6–60 directions are used at each frequency (if you can leave something dead in the scanner you can get substantially more).
- If we wanted to numerically invert the density, substantially more measurements are required. We need to model the observations [1].
- Usually a Gaussian density is assumed in 3-D with zero mean (no net diffusion), based on truncating the three-dimensional Bolzmann equation.
- This is known not to fit well (excess kurtosis), even in the simplest case of no orientational heterogeneity [5, 2].

Modelling

- About 6–60 directions are used at each frequency (if you can leave something dead in the scanner you can get substantially more).
- If we wanted to numerically invert the density, substantially more measurements are required. We need to model the observations [1].
- Usually a Gaussian density is assumed in 3-D with zero mean (no net diffusion), based on truncating the three-dimensional Bolzmann equation.
- This is known not to fit well (excess kurtosis), even in the simplest case of no orientational heterogeneity [5, 2].
- From all individually fitted PDFs, a connectivity map is obtained from “connect the dots”, or “tracking”.

What Can We Say?

- A realistic model is to take the PDF of a single fibre as ellipsoid, i.e. with decay of elliptical contours.

What Can We Say?

- A realistic model is to take the PDF of a single fibre as ellipsoid, i.e. with decay of elliptical contours.
- From such mild assumptions it is possible to determine if there is one orientational preference, and if the density appears Gaussian.

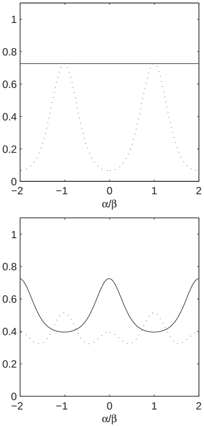
What Can We Say?

- A realistic model is to take the PDF of a single fibre as ellipsoid, i.e. with decay of elliptical contours.
- From such mild assumptions it is possible to determine if there is one orientational preference, and if the density appears Gaussian.
- Various summaries of non-Gaussianity [4] can then be feed into tracking algorithms. Especially important is determining fanning/forking. Distributional theory can be calculated for the summaries.

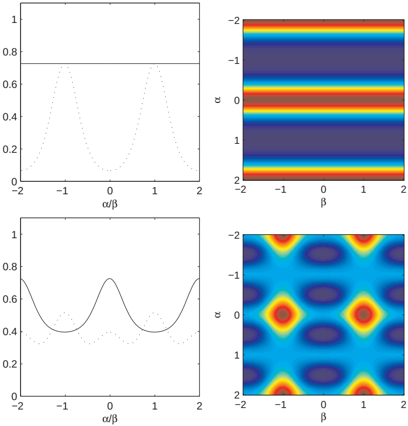
What Can We Say?

- A realistic model is to take the PDF of a single fibre as ellipsoid, i.e. with decay of elliptical contours.
- From such mild assumptions it is possible to determine if there is one orientational preference, and if the density appears Gaussian.
- Various summaries of non-Gaussianity [4] can then be feed into tracking algorithms. Especially important is determining fanning/forking. Distributional theory can be calculated for the summaries.
- “Far better an approximate answer to the right question, than the exact answer to the wrong question, which can always be made precise.”

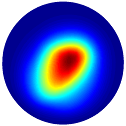
What Can We Say?



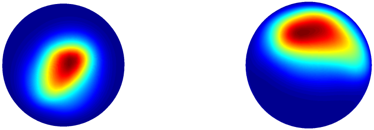
What Can We Say?



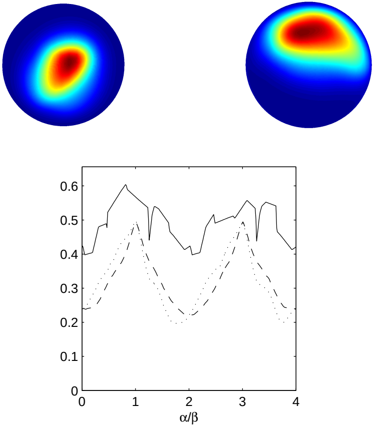
What Do We Find?



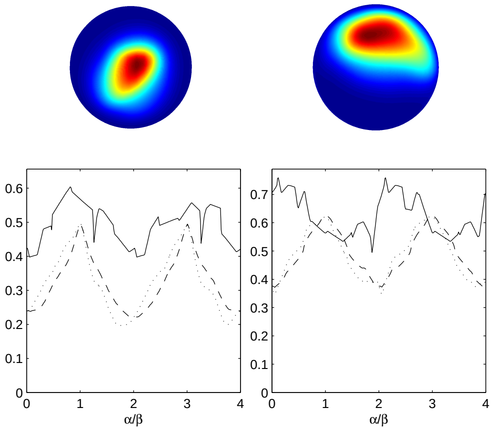
What Do We Find?

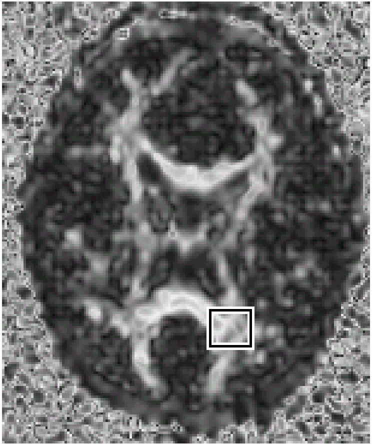


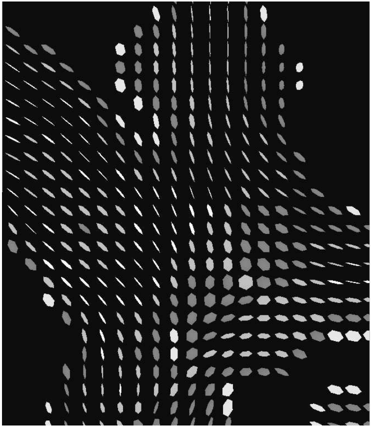
What Do We Find?

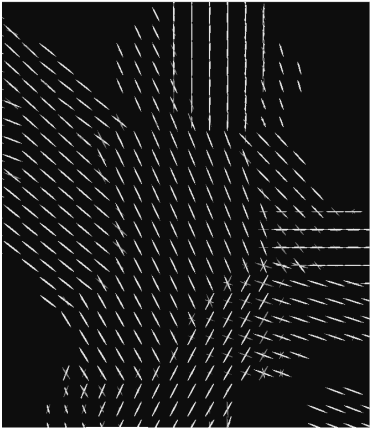


What Do We Find?









Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.
- Ocean circulation.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.
- Ocean circulation.
- Inhomogeneous Random Fields.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.
- Ocean circulation.
- Inhomogeneous Random Fields.
- High-Frequency Finance data.

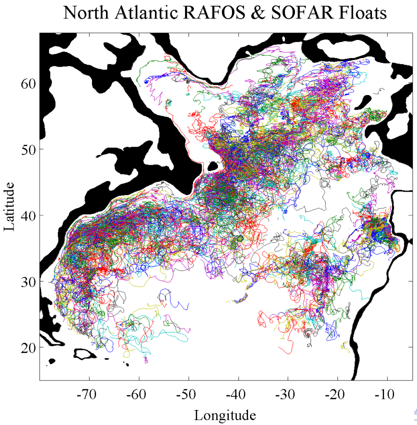
Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.
- Ocean circulation.
- Inhomogeneous Random Fields.
- High-Frequency Finance data.
- Shape analysis.

Beauty of maths is Generality - Some of My Active Projects

- Theory and Methods for Non-Stationary Multivariate Time Series.
- Gravity anomalies.
- Pain perception in babies.
- Ocean circulation.
- Inhomogeneous Random Fields.
- High-Frequency Finance data.
- Shape analysis.
- Ocean Ecology.

Observed Series



Before

UK

- The UK has a number of strong Statistics departments.

UK

- The UK has a number of strong Statistics departments.
- Less recent strong figures include Karl and Egon Pearson, Fisher, David Cox etc. The oldest department of Statistics can be found in the UK.

UK

- The UK has a number of strong Statistics departments.
- Less recent strong figures include Karl and Egon Pearson, Fisher, David Cox etc. The oldest department of Statistics can be found in the UK.
- To improve the training of PhD students there are a number of graduate training centres, notably Academy for PhD Training in Statistics, the London Taught Course Centre and the Scottish Mathematical Sciences Training Centre.

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.
- *large p small n*. What happens if you have more parameters than observations? Traditionally such problems are not solvable: recent theory begs to differ (Isaac Newton Institute).

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.
- *large p small n*. What happens if you have more parameters than observations? Traditionally such problems are not solvable: recent theory begs to differ (Isaac Newton Institute).
- Functional Data Analysis.

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.
- *large p small n*. What happens if you have more parameters than observations? Traditionally such problems are not solvable: recent theory begs to differ (Isaac Newton Institute).
- Functional Data Analysis.
- Continuous time observations governed by an SDE or a set of ODEs/PDEs.

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.
- *large p small n*. What happens if you have more parameters than observations? Traditionally such problems are not solvable: recent theory begs to differ (Isaac Newton Institute).
- Functional Data Analysis.
- Continuous time observations governed by an SDE or a set of ODEs/PDEs.
- Problems motivated by biology/genetics.

Important Developments:

- Shape Analysis: much of existing theory and methods have been developed in the UK. Shapes don't add.
- *large p small n*. What happens if you have more parameters than observations? Traditionally such problems are not solvable: recent theory begs to differ (Isaac Newton Institute).
- Functional Data Analysis.
- Continuous time observations governed by an SDE or a set of ODEs/PDEs.
- Problems motivated by biology/genetics.
- Problems motivated by imaging (i.e. fMRI).

UK

- Areas include: *Biostatistics* (Wilkinson (Newcastle), Green (Bristol), Richardson (ICL), Balding (ICL)), *Causality*, (Dawid (Cambridge), Lauritzen (Oxford), Smith (Warwick)), *MCMC* (Andrieu (Bristol), Green (Bristol), Richardson (ICL), Roberts (Warwick)...), *time series and statistical signal processing* (Holmes (Oxford), Nason (Bristol), Olhede (ICL & UCL), Walden (ICL)), Machine learning (Cristianini (Bristol), Shawe-Taylor (UCL), Titterington (Glasgow)), *Shape Analysis*, (Bowman (Glasgow), Kent, Mardia (Leeds)), *large p small n*, (Meinshausen (Oxford), Cheng (UCL), Yao (LSE), Samworth (Cambridge)), *functional data analysis*, (Farraway (Bath), Guillas (UCL)), *Algebraic Statistics* (Wynn (LSE)), *Model Choice*, (Jon Foster (Southampton), Brown (Kent)), *Bayesian Theory* Walker (Kent), *Ecological Statistics* (Morgan (Kent)) *etc*.

UCL

- **Computational Statistics** Trevor Sweeting, Jing-Hao Xue, Serge Guillas, Ricardo Silva, Alex Beskos.
- **Medical Statistics** Rumana Omar, Julie Barber and Gareth Ambler.
- **Multivariate and High Dimensional Data** Tom Fearn, Christian Hennig, Jing-Hao Xue, Ming-Yen Cheng, Ricardo Silva, Sofia Olhede.
- **Non-Parametric Methods** Ming-Yen Cheng , Christian Hennig, Sofia Olhede, Serge Guillas.
- **Stochastic Modelling** Valerie Isham, Paul Nothrup, Richard Chandler, Hilde Herdeboot.
- **Time Series** Richard Chandler, Sofia Olhede, Serge Guillas.

UCL

- The Department of Statistical Science was founded in 1911 with Karl Pearson as its head. It has 20 members of staff, covering most areas of statistics.

UCL

- The Department of Statistical Science was founded in 1911 with Karl Pearson as its head. It has 20 members of staff, covering most areas of statistics.
- UCL also has a **Centre for Statistical Machine Learning**, run by John Shawe-Taylor.

UCL

- The Department of Statistical Science was founded in 1911 with Karl Pearson as its head. It has 20 members of staff, covering most areas of statistics.
- UCL also has a **Centre for Statistical Machine Learning**, run by John Shawe-Taylor.
- General inquiries should be directed to Afzal Siddiqui (afzal at stats.ucl.ac.uk), the PG admission tutor.

UCL

- The Department of Statistical Science was founded in 1911 with Karl Pearson as its head. It has 20 members of staff, covering most areas of statistics.
- UCL also has a **Centre for Statistical Machine Learning**, run by John Shawe-Taylor.
- General inquiries should be directed to Afzal Siddiqui (afzal at stats.ucl.ac.uk), the PG admission tutor.
- All staff can be e-mailed using firstname at stats.ucl.ac.uk .

UCL

- The Department of Statistical Science was founded in 1911 with Karl Pearson as its head. It has 20 members of staff, covering most areas of statistics.
- UCL also has a **Centre for Statistical Machine Learning**, run by John Shawe-Taylor.
- General inquiries should be directed to Afzal Siddiqui (afzal at stats.ucl.ac.uk), the PG admission tutor.
- All staff can be e-mailed using firstname at stats.ucl.ac.uk .
- UCL also has an interdisciplinary PhD training centre CoMPLEX of which Statistical Science is a member.

References I

- D. C. Alexander,
  Multiple-fibre reconstruction algorithms for diffusion MRI,
  Annals of the New York Academy of Sciences, **1046**, 113–133, 2005.

- K. M. Jansons and D. C. Alexander,
  Persistent angular structure: New insights from diffusion magnetic resonance imaging data,
  Inverse Problems, **19**, pp. 1031–1046, 2003.

References II

- S. Ljunggren,
  A Simple Graphical Representation of Fourier Based Imaging Methods,
  J. Magnetic Resonance, **54**, pp. 338–343, 1983.
- S. C. Olhede and B. Whitcher,
  A Statistical Framework to characterise microstructure in high angular resolution diffusion imaging,
  5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2008, pp. 899 - 902, 2008.

References III

D. S. Tuch,
Q-ball Imaging,
MRM, **52**, 1358–1372, 2004.

D. Tweig,
The k-trajectory of the NMR imaging process with applications in analysis and synthesis of imaging methods,
Med. Phys., **10**, pp. 610–21, 1983.

References IV

V. Wedeen, P. Hagmann, W.-Y. Tseng, T. G. Reese and R. M. Weisskoff,
Mapping Complex Tissue Architecture with Diffusion Spectrum Imaging,
Magnetic Resonance in Medicine, **54**, pp. 1377–1386, 2005.