Neurodata MRI to Graphs (NDMG)

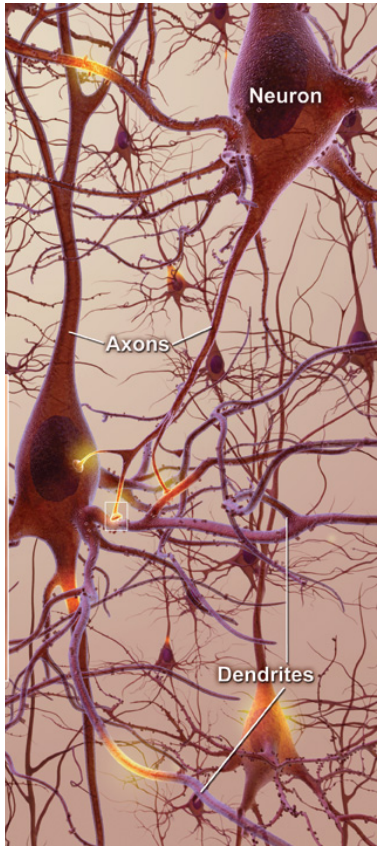
A low resource multimodal pipeline to democratize connectome estimation and analysis

Derek Pisner, Alex Loftus, and Ross Lawrence

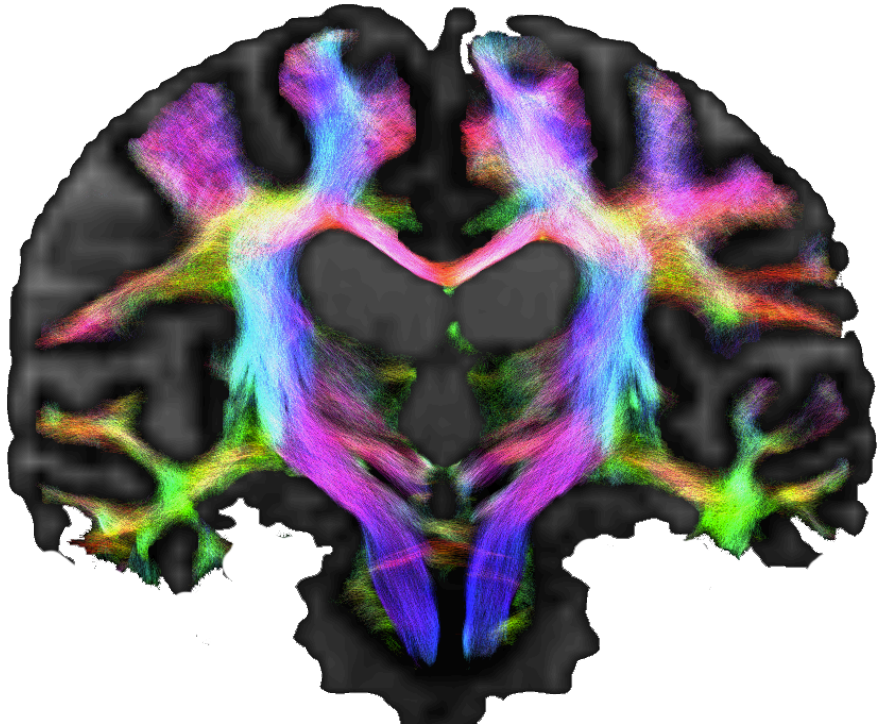
08/19/2019

A brief overview of diffusion MRI (dMRI)

White-matter imaging as a study of microstructural *connectivity*



- Axons measure ~μm in width
- Group together in bundles that traverse the white matter
- Can't image individual axons but we can image *bundles* (i.e. and more broadly *fascicles*) with dMRI
- Useful for studying neurodegenerative diseases, stroke, aging, development…



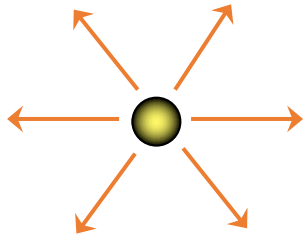
≈ ?

Diffusion in brain tissue

- Differentiate tissues based on the diffusion (random motion) of water molecules within them

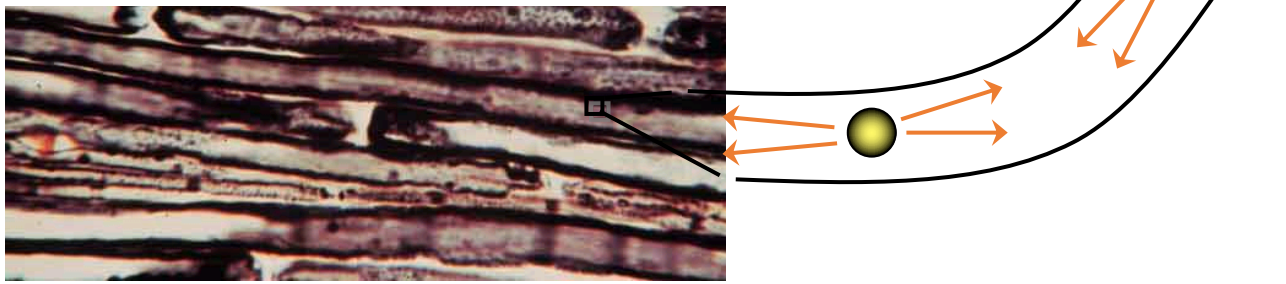
- Gray matter: Diffusion is uniformly restricted

  $\Rightarrow$ isotropic



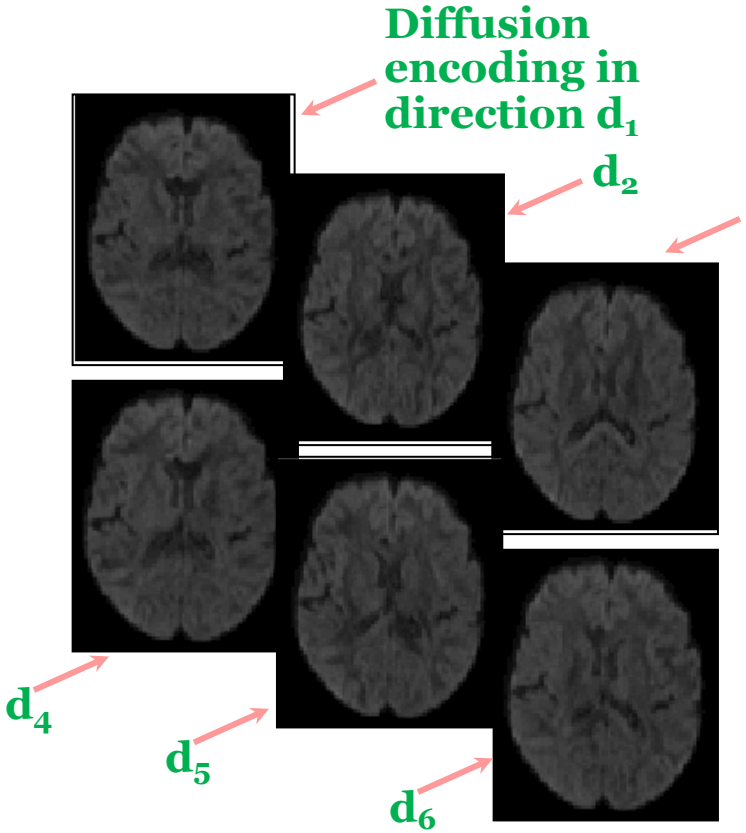
- White matter: Diffusion is preferentially restricted

  $\Rightarrow$ anisotropic



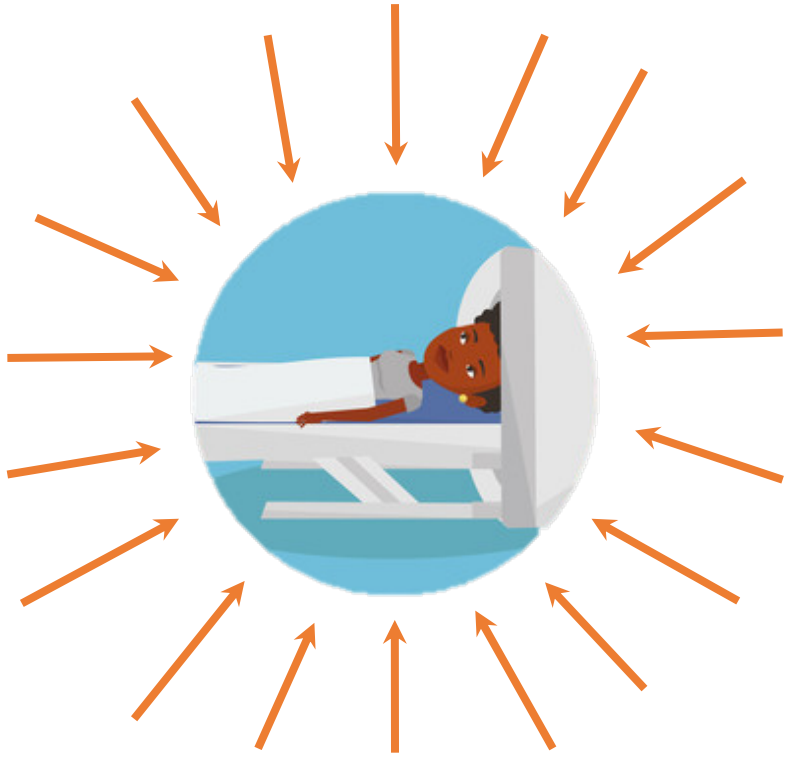
Diffusion MRI

- Magnetic resonance imaging can provide "diffusion encoding"
- Magnetic field strength is varied by gradients in different directions
- Image intensity is attenuated depending on *water diffusion* in each direction
- Compare with reference image(s) to infer diffusion relative to non-diffusion signal



Diffusion MRI =
Gradient Magnitudes + Directions

Gradient directions



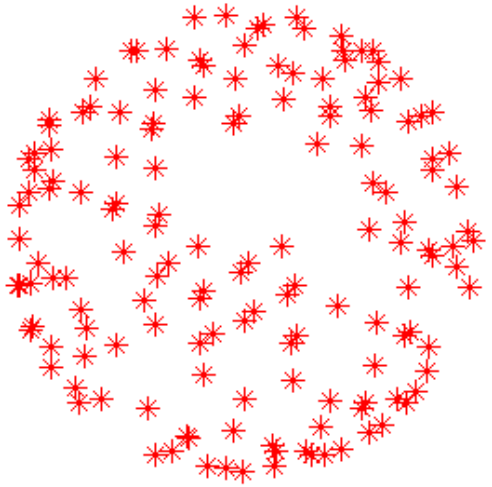
B-vectors are the x, y, z coordinates whose polarity is set by the scanner gradients and are described with respect to the person in the scanner

Number of gradient directions matters

- In principle, ↑ directions = ↑ reliable estimation of diffusion.
  - Caveat: increased imaging time = ↑ cumulative noise (Jones, 2004)

- DTI:
  - >6, but usually 10's of directions

- DSI:
  - Usually few 100's of directions

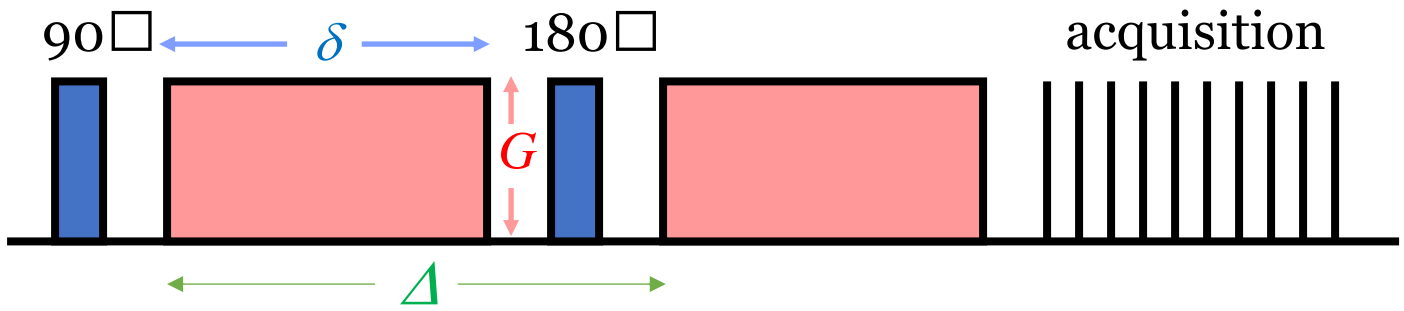


Gradient magnitude

- **B-values** depend on acquisition parameters:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$$

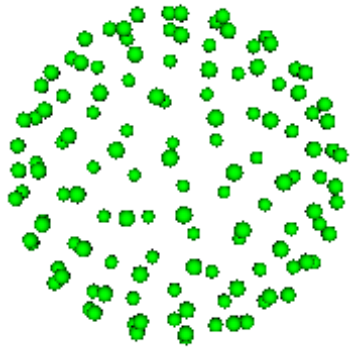
  - γ gyromagnetic ratio
  - G strength of diffusion-encoding gradient
  - δ duration of diffusion-encoding pulse
  - Δ interval of diffusion-encoding pulses

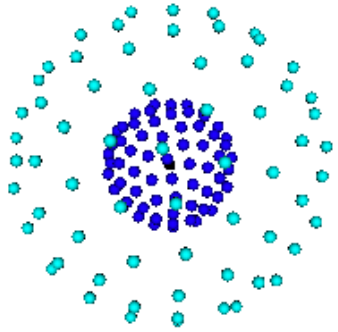


Strength of gradients matters

- Increasing b-value leads to:
  - Increased tissue *contrast* with areas of higher / lower diffusivity
  - More noise relative to signal (↓ SNR) ⇒ ↓ reliable estimation of diffusion

- DTI(*uniform b-values*): b ~ 1000 sec/mm^2
- DSI (*many directions*): b ~ 10,000 sec/mm^2
- Multi-shell (*multiple b-values*): b ~ 1,000; 2,500; 5,000 sec/mm^2

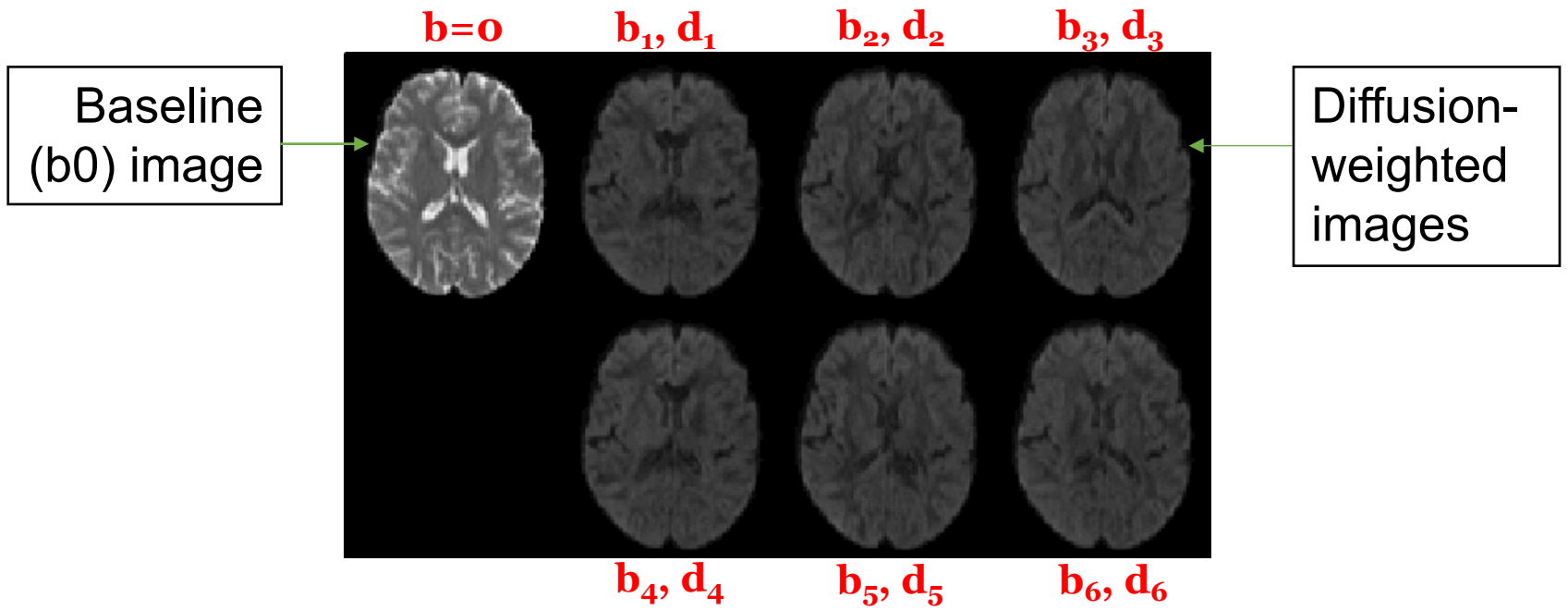




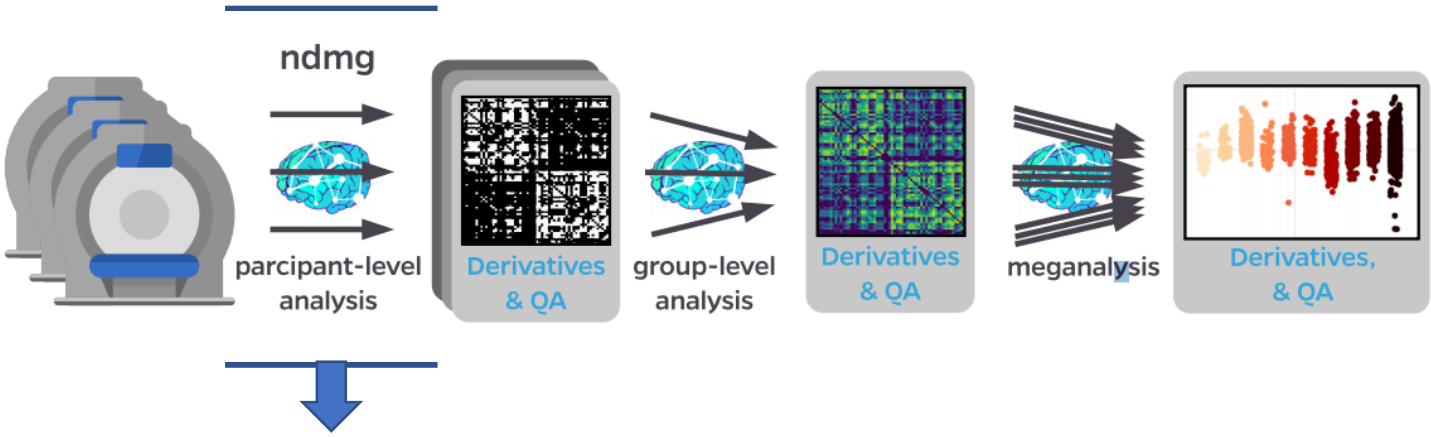
Looking at the data

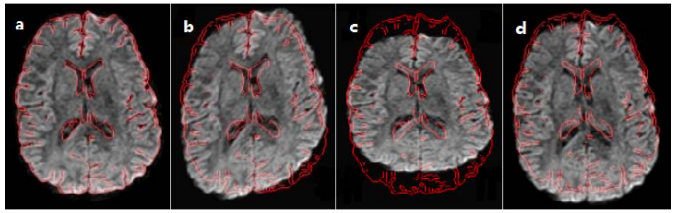
A diffusion dataset consists of:

- Non-diffusion-weighted a.k.a “low-b” reference images (b-value = 0)
- Diffusion-weighted images (DWI) with different gradient directions d_1, d_2, and b-value >0.
  - bvec: B-vectors
  - bval: B-values
- DWI have lower intensity values

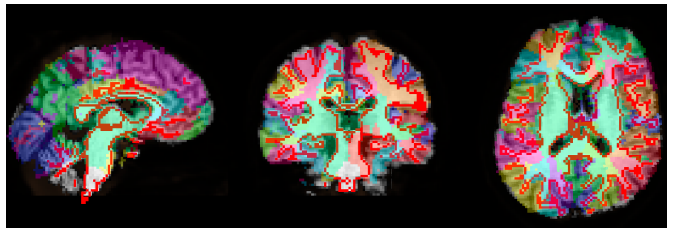


Data analysis steps

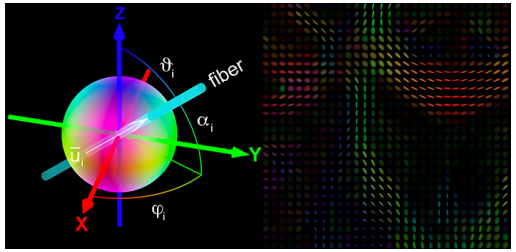




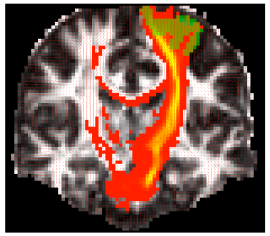
1. Pre-process images to reduce distortions



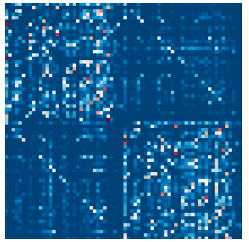
2. Registration



3. Reconstruction



4. Tractography



5. Connectome

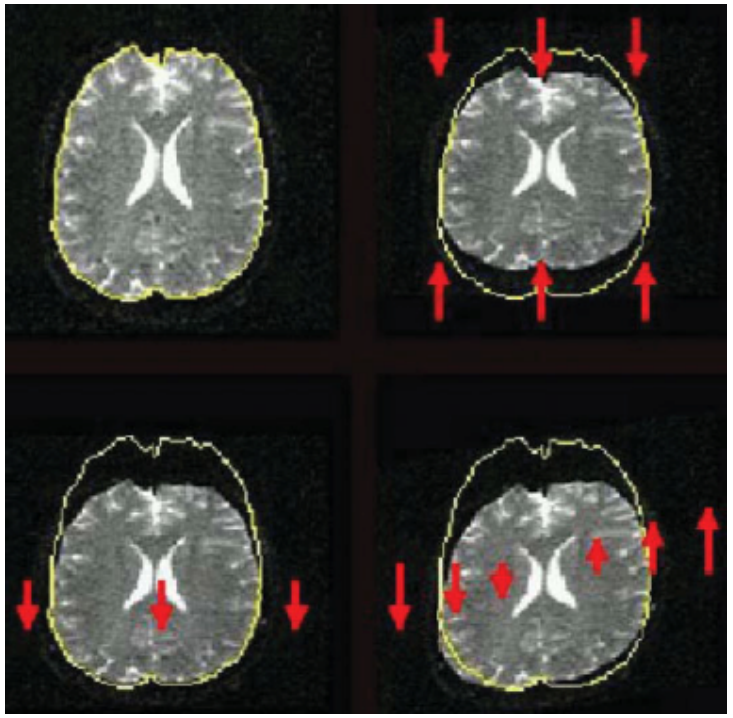
NDMG

- An end-end tool for efficiently generating high quality structural connectomes.
- Scalable to large datasets, intended for mega-analysis
- Begin with raw NIFTI1 data (BIDS spec), end with graph(s)
- Total runtime 20 min – 1 hour
- Tractography in different spaces:
  - native-space, native-space with MNI-normalization, and MNI-space
- Different types of tractography:
  - Deterministic, probabilistic, particle
- Use different diffusion models:
  - CSA, CSD

Preprocessing

Distortions: Eddy currents

- Cause: Fast switching of diffusion-encoding gradients induces eddy currents in conducting components
- Eddy currents lead to residual gradients that shift the diffusion gradients
- The shifts are direction-dependent, *i.e.*, different for each DW image
- Result: Geometric distortions



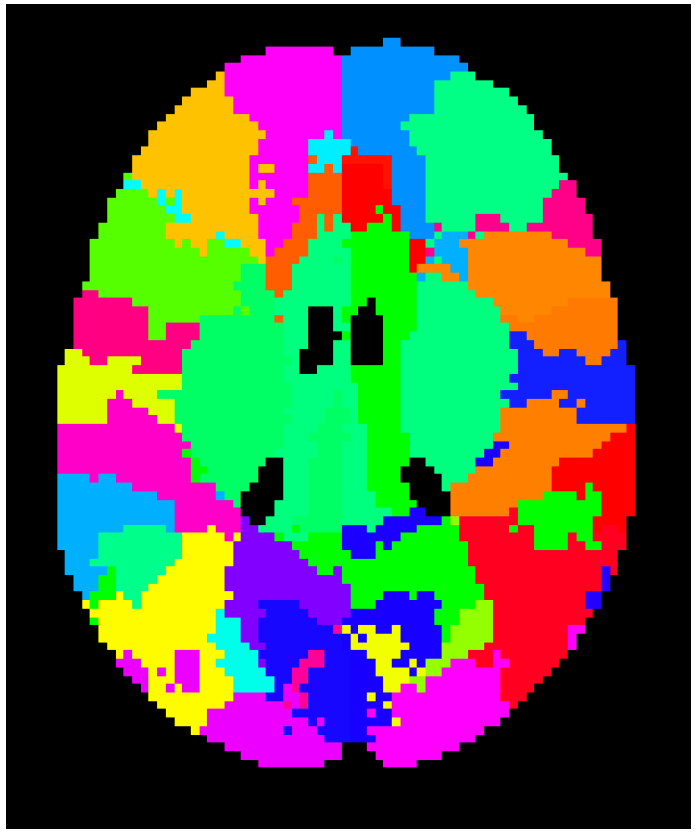
From Le Bihan *et al.*, Artifacts and pitfalls in diffusion MRI, JMRI 2006

NDMG Preprocessing

- Perform eddy correction on the dMRI series.
- Reorient all neuroimages to RAS+ and reslice to either 1 mm or 2 mm voxel resolution
- Create a mean B0 brain mask of all B0's found in the dMRI series

Registration

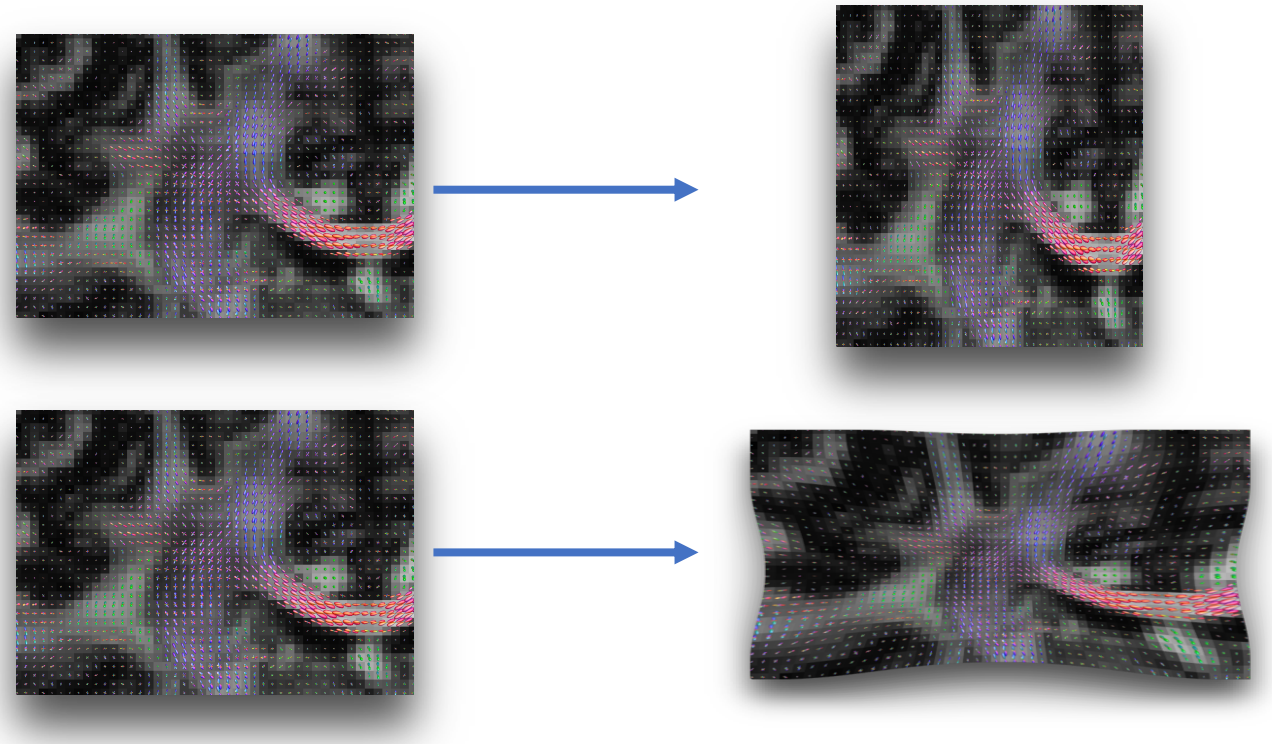
GOAL: Get atlas (what's that?) to native diffusion space (what's that?)



Parcellation	Type of Parcellation	Number of vertices
AAL [8]	anatomical	116
Desikan [7]	anatomical	70
Harvard-Oxford combined [10]	anatomical	111
Harvard-Oxford Cortical [10]	anatomical	48
Harvard-Oxford Subcortical [10]	anatomical	21
Talairach [11]	anatomical	1105
Brodmann	anatomical	41
JHU [9]	anatomical	48
Glasser [15]	semi-automated	180
slab907 [12]	algorithmic	907
slab1068 [13]	algorithmic	1068
CPAC200	algorithmic	200
DS00071 [18]	algorithmic	70
DS00108 [18]	algorithmic	107
DS00140 [18]	algorithmic	139
DS00195 [18]	algorithmic	194
DS00278 [18]	algorithmic	277
DS00350 [18]	algorithmic	349
DS00446 [18]	algorithmic	445
DS00583 [18]	algorithmic	582
DS00833 [18]	algorithmic	832
DS01216 [18]	algorithmic	1215
DS01876 [18]	algorithmic	1875
DS03231 [18]	algorithmic	3231
DS06481 [18]	algorithmic	6481
DS16784 [18]	algorithmic	16784
DS72784 [18]	algorithmic	72784

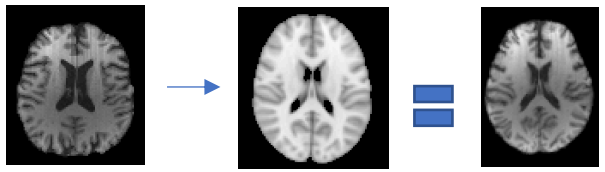
Why not just register a dMRI directly to MNI-space like we do with fMRI?

This requires *inter*modal registrations, which presents a unique challenge for dMRI.

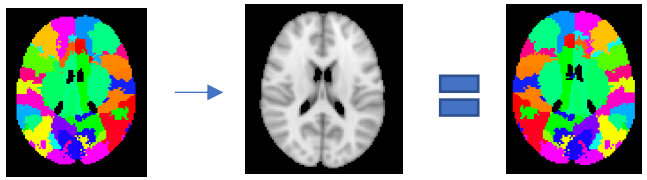


NDMG registration

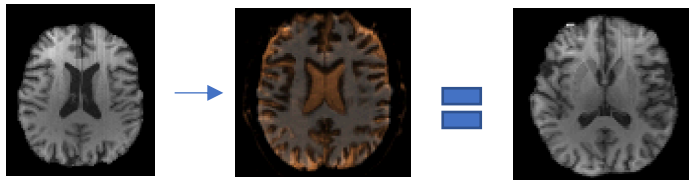
1. Learn non-linear T1w → MNI



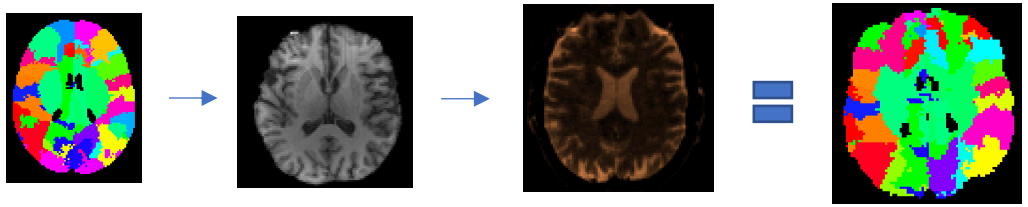
2. Learn non-linear warp from Atlas → T1w in MNI space



3. Learn T1w → native dMRI (linear with 'boundary-based' tissue priors from T1w white-matter segmentation)

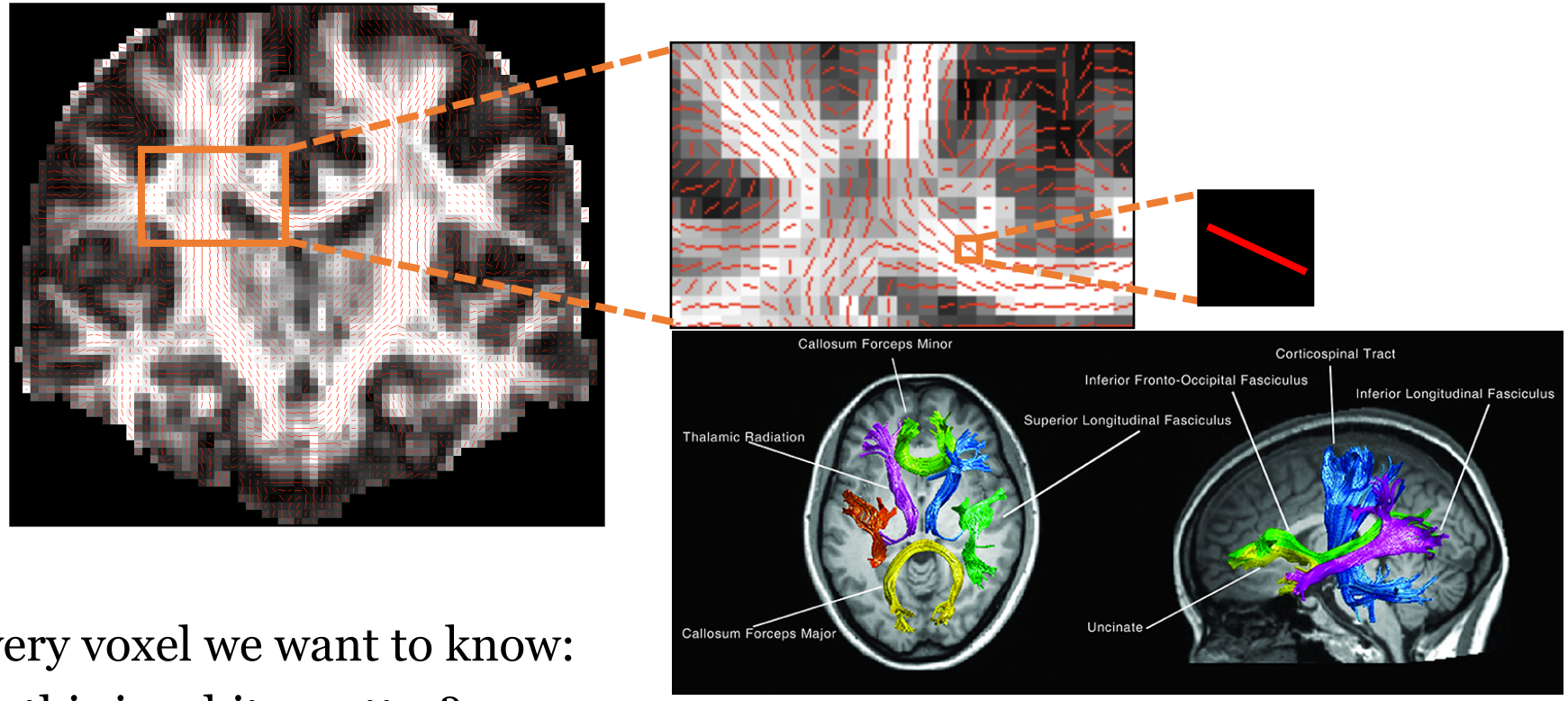


4. Apply T1w MNI → T1w (apply inverse of non-linear warp from (1)) → native dMRI (transform from (3))



Estimate diffusion directions at each voxel (i.e. ‘Reconstruction’)

How to represent diffusion



- At every voxel we want to know:
  - Is this in white matter?
  - If yes, what pathway(s) is it part of?
    - What is the orientation of diffusion?
    - What is the magnitude of diffusion?
- A standard T1-weighted image cannot capture all this!

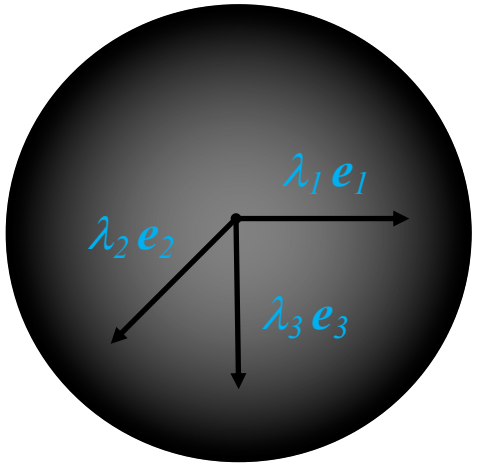
Simplest model: Tensor

- A tensor is a 3x3 symmetric, positive-definite matrix:
- Eigenvectors → diffusion direction
- Eigenvalues → diffusion magnitude

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{12} & d_{22} & d_{23} \\ d_{13} & d_{23} & d_{33} \end{bmatrix}$$

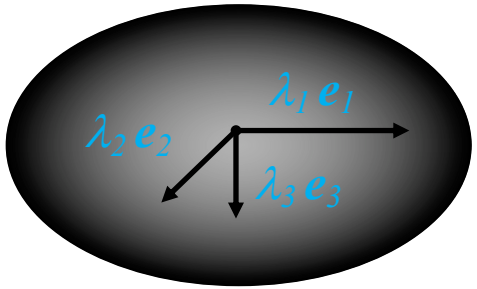
Isotropic diffusion:

$\lambda_1 \approx \lambda_2 \approx \lambda_3$



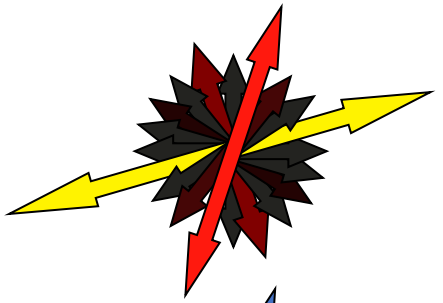
Anisotropic diffusion:

$\lambda_1 >> \lambda_2 \approx \lambda_3$



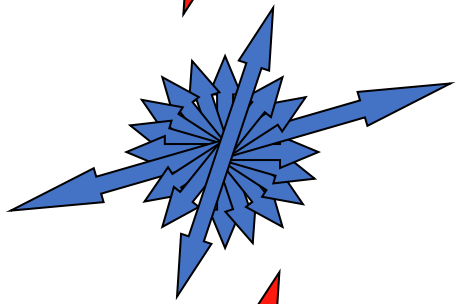
- Tensor → Ellipsoid = Likelihood of water molecule displacements at that voxel

The many models of diffusion



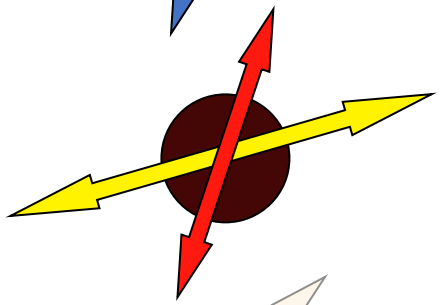
Variable number of fiber populations:

Probability distribution of fiber populations at different diffusion orientations *and* magnitude (FOD, diffusion spectrum)



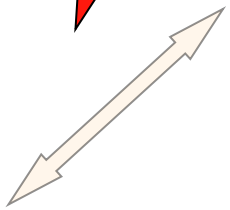
Orientation distribution function (ODF):

Probability distribution of spin diffusion (orientation only)



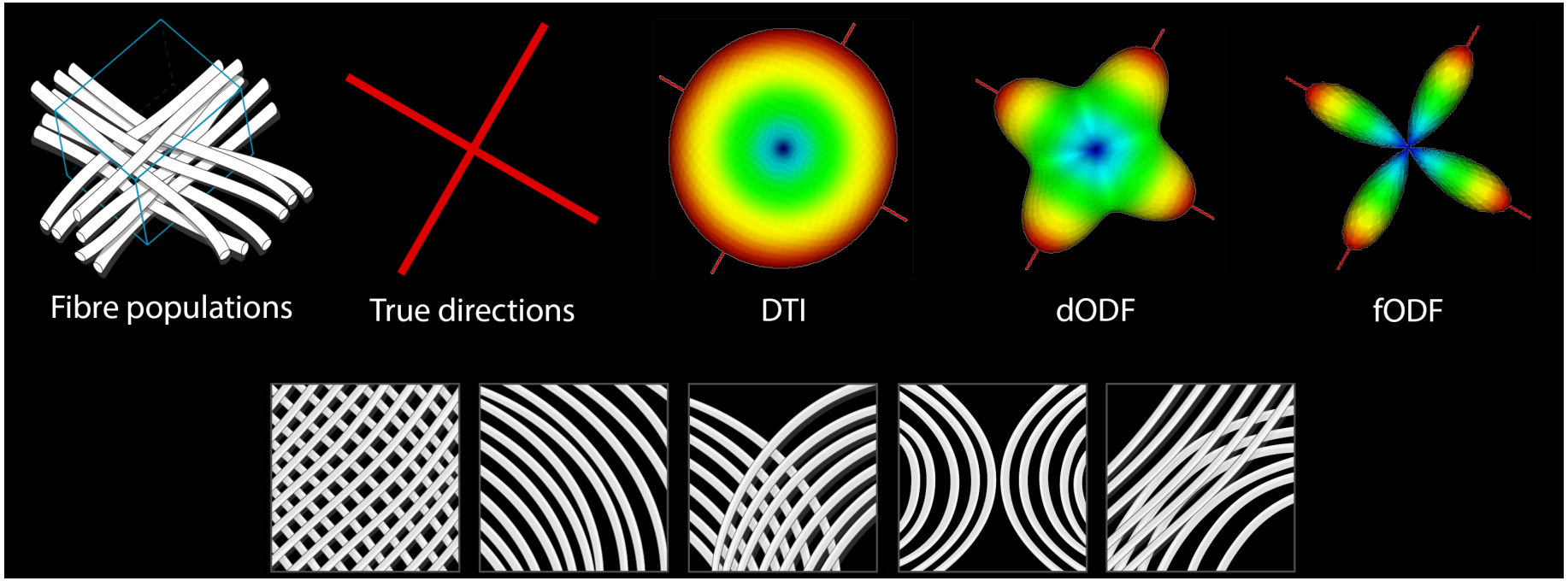
Fixed number of fiber populations:

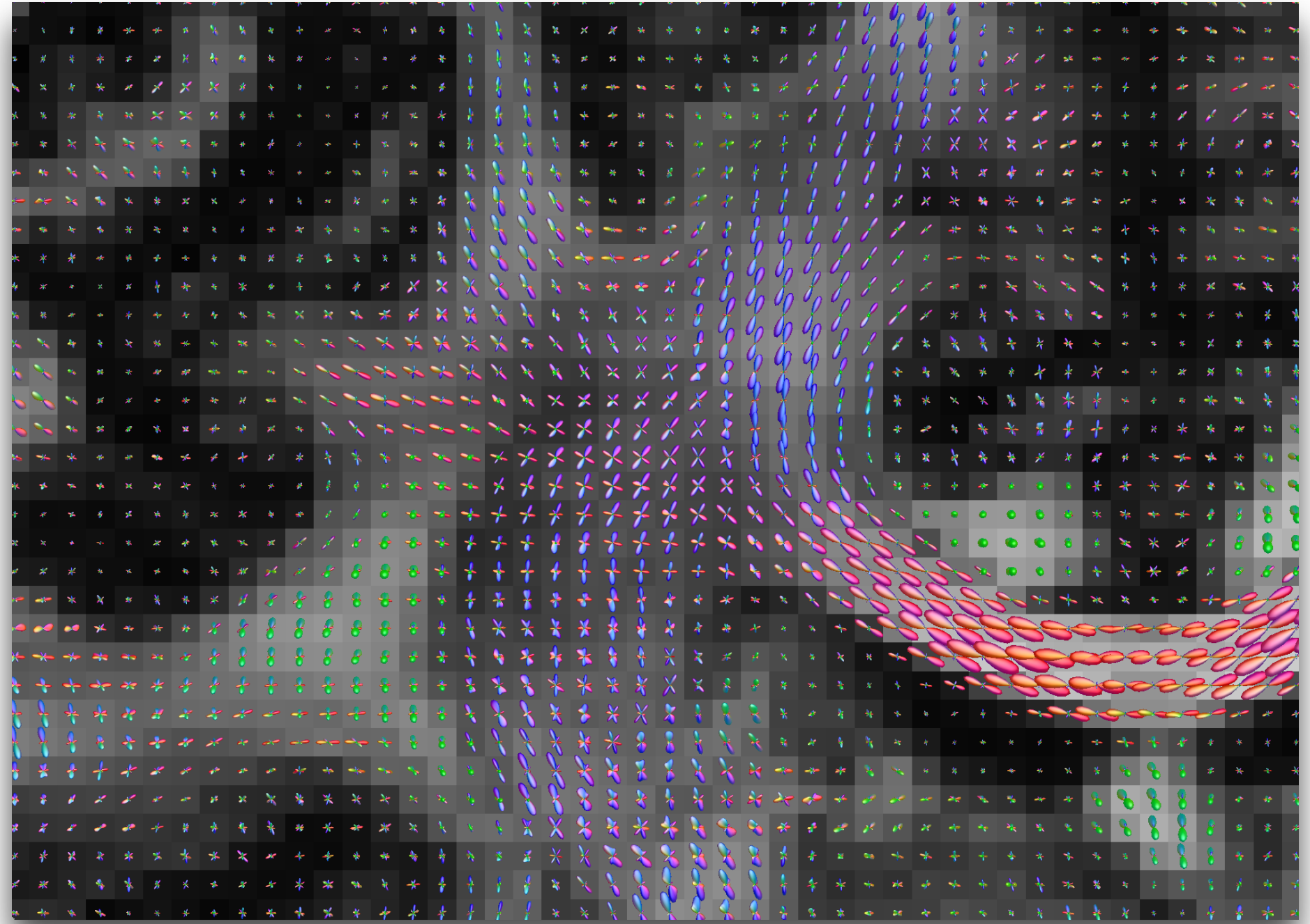
Diffusion orientation and magnitude for N compartments (mixture of tensors, ball-and-stick, *etc.*)



Tensor:

Single diffusion orientation and magnitude

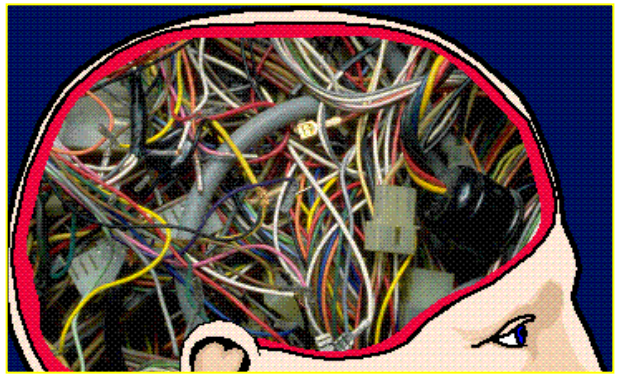

Fibre populations
True directions
DTI
dODF
fODF



NDMG Reconstruction

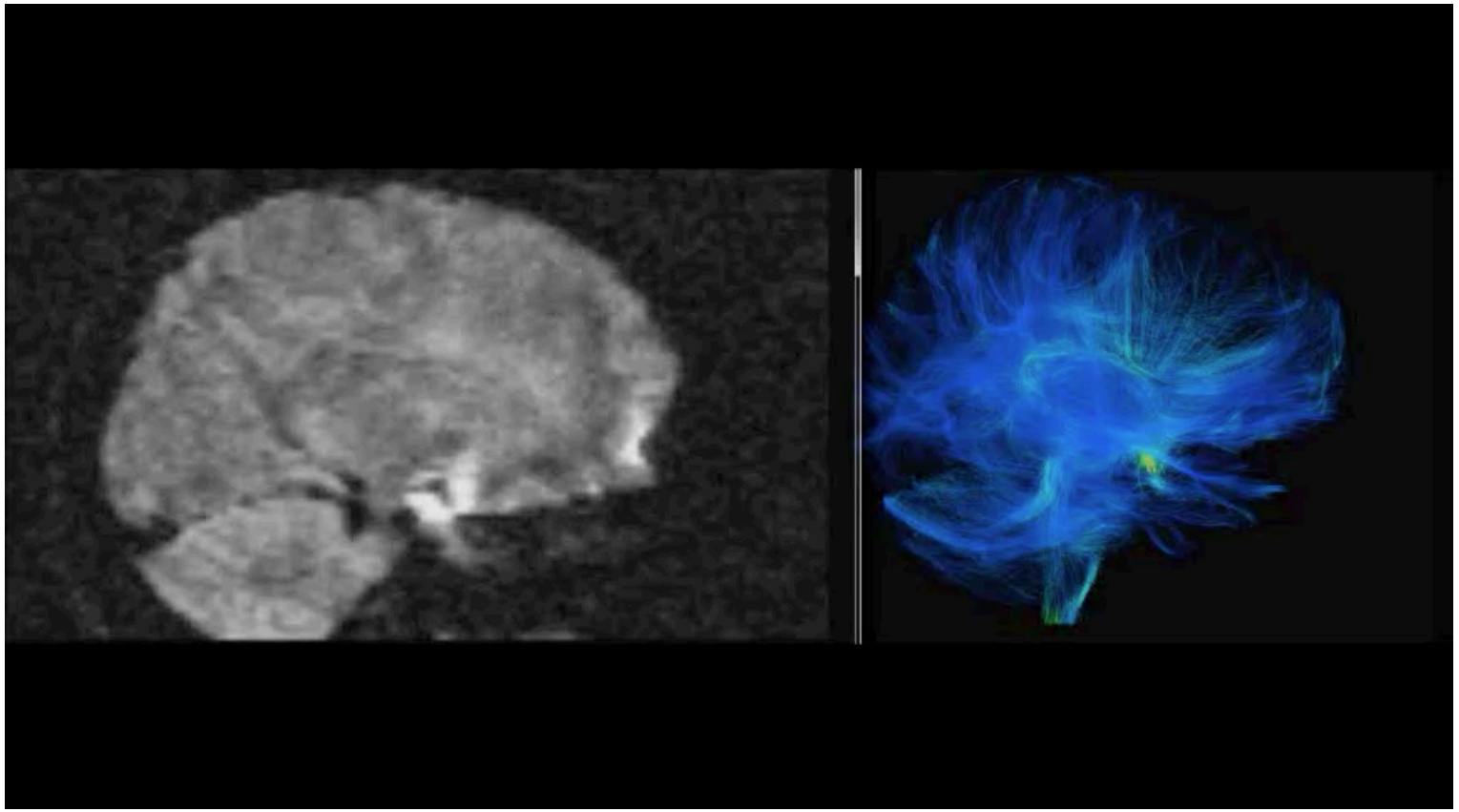
- Multiple options available (dODF or fODF):
  - dODF : Constant Solid Angle (CSA)
  - fODF : Constrained Spherical Deconvolution (CSD)

Tractography

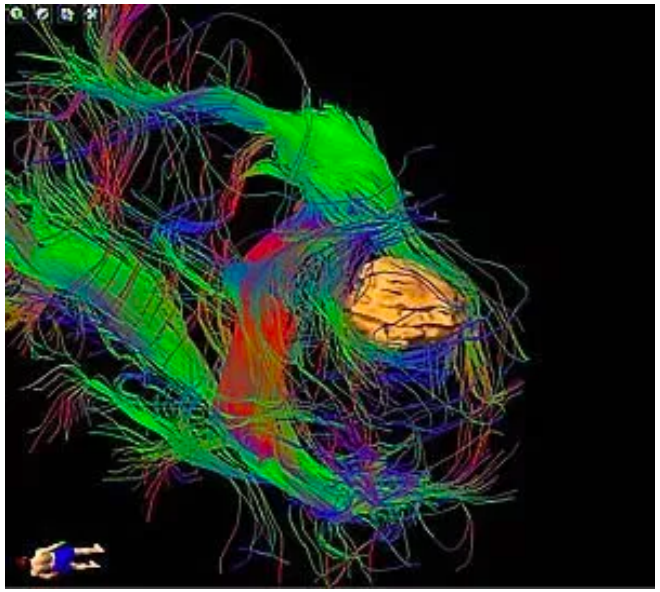


Single Subject
Diffusion Weighted MRI

Single Subject
White Matter Reconstruction



Connectomics is a more recent type of tractography application



NEUROSURGICAL PLANNING

(E.G. TUMOR REMOVAL)

EXPLORING INDIVIDUAL DIFFERENCES IN MICROSTRUCTURE OF KNOWN FASCICLES/ BUNDLES OF WHITE-MATTER

Tracking

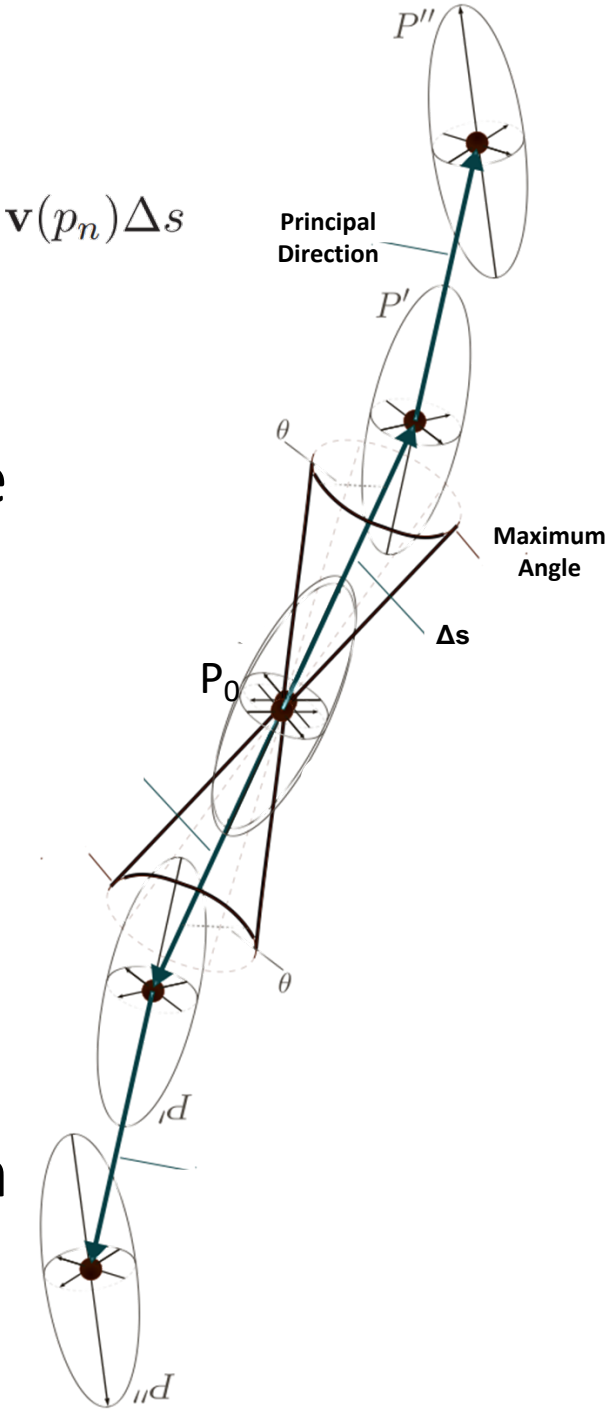
$$p_{n+1} = p_n + \mathbf{v}(p_n)\Delta s$$

1. For each seed position (initial position)

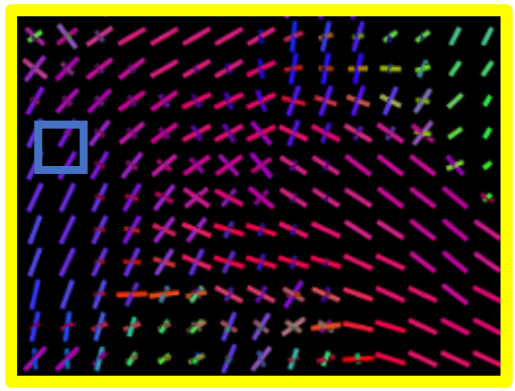
2. Follow the path along sequential voxels based on the underlying directional estimate.

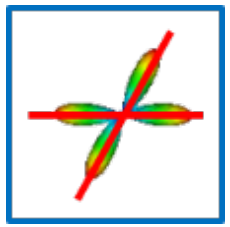
3. Stop if:
   - The position p_n is outside of the mask
   - The angle between the vector v_n and v_{n-1} is greater than the fixed maximum (e.g. opening angle $\theta=60^o$)

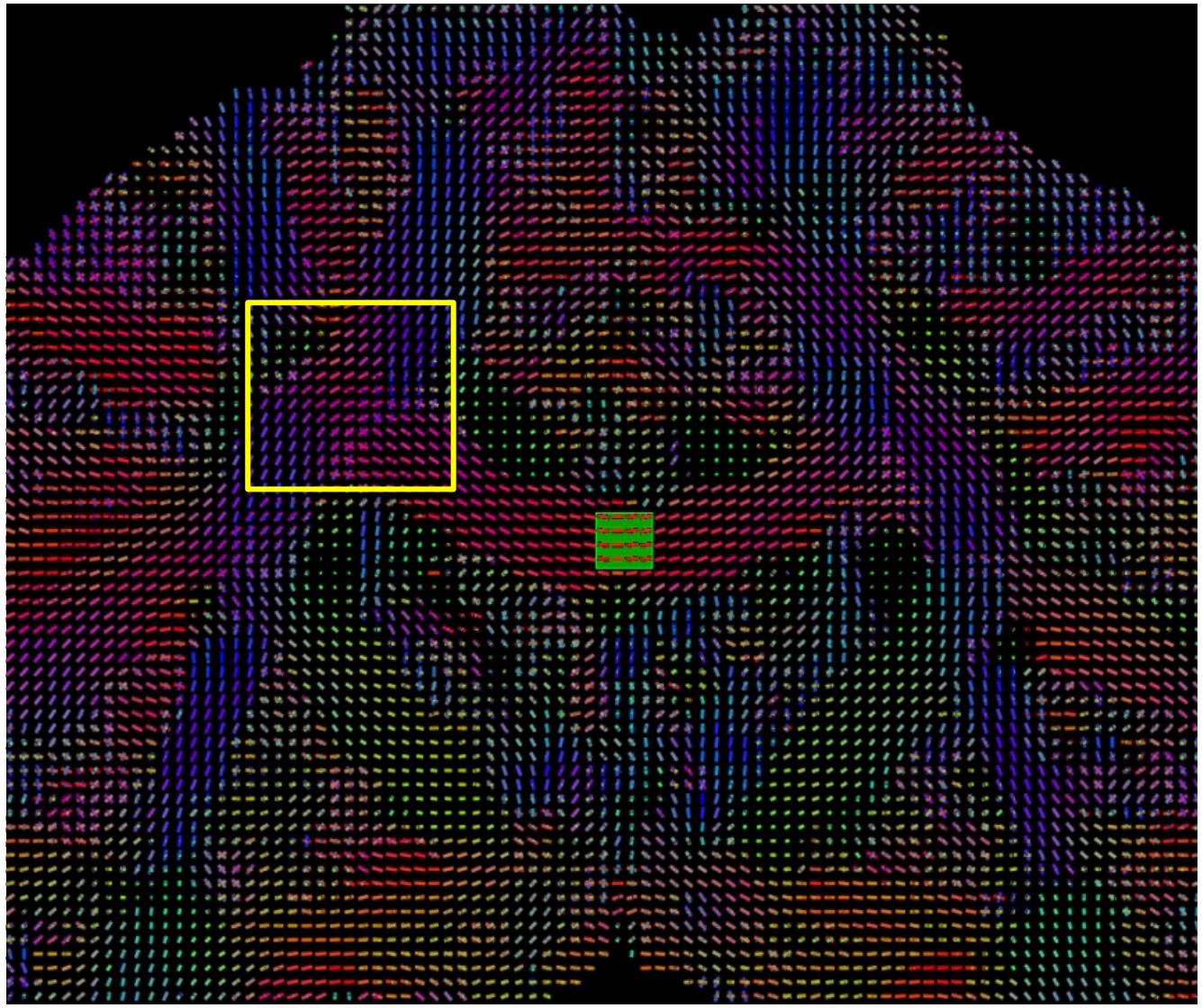
4. Restart from the seed in the opposite initial direction



Fiber ODF Tractography

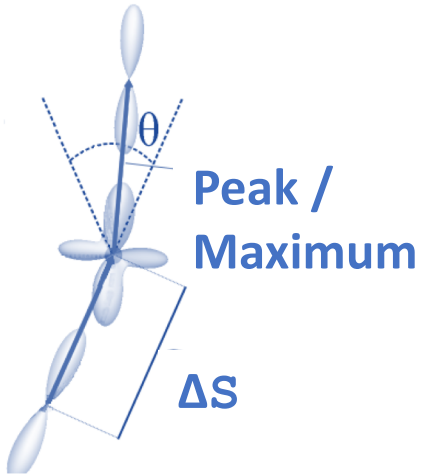




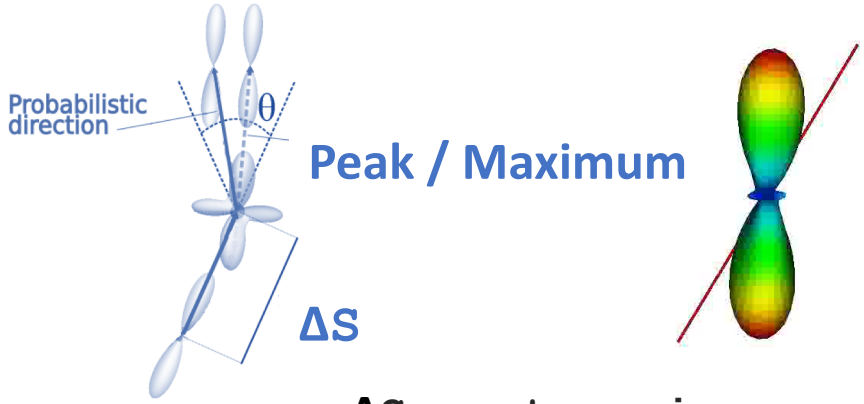


Deterministic vs. Probabilistic Tractography

- Deterministic:
  - Follow the peak direction of the fiber ODF in the (CSA / CSD case)
  - If multiple peaks are available in the distribution, follow the closest to the incoming direction



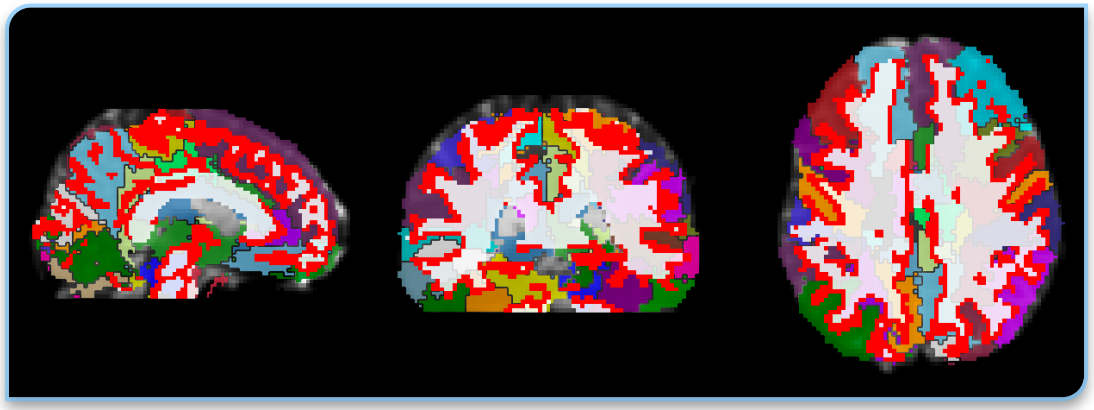
- Probabilistic:
  - Sample a direction from the model at each voxel to probabilistically select a direction to follow

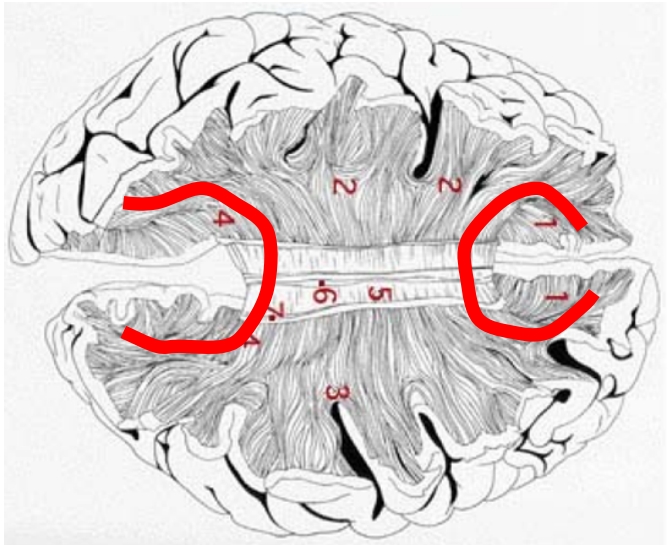


ΔS : step size
θ : opening angle

NDMG Tractography

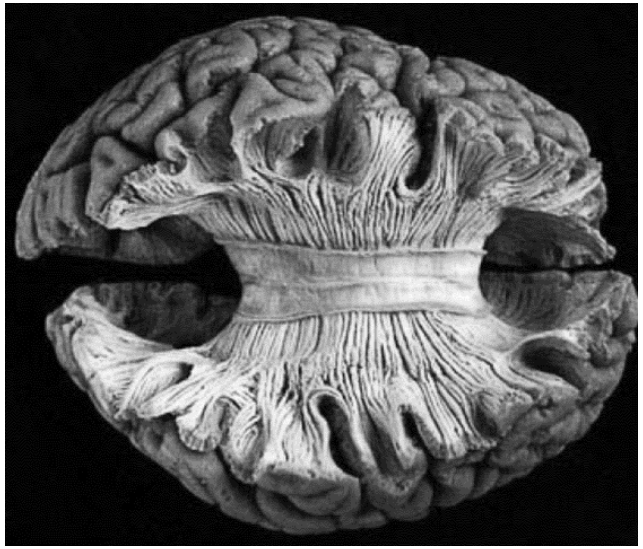
- Probabilistic, deterministic, and particle tracking methods available.
- Denser seeding (configurable with --seed flag) yields more streamlines.
- 500,000 -1 million streamlines recommended (need --seed 20/25) (Roine et al. 2019)
  - Reproducibility increases with streamlines, but it's compute hungry.
- Seeding occurs in:
  - (1) GM-WM interface
  - (2) Corpus Callosum

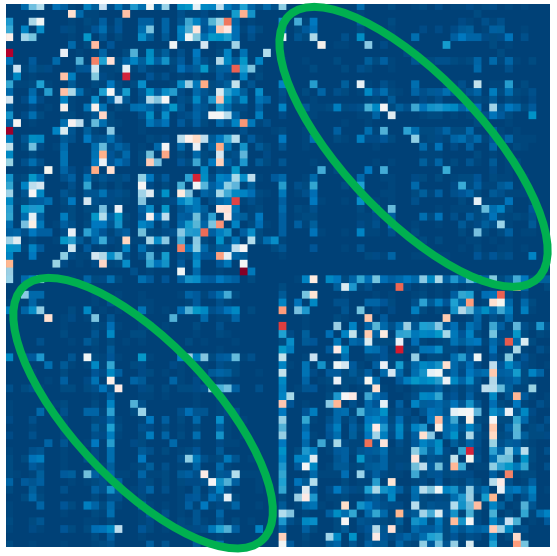




Challenge:
Recovering contralateral connections

- Especially populous (150-200 million) commissural (interhemispheric) fibers need to be tracked somehow!
  - Sufficiently large step-size in tractography can help. NDMG uses 0.2 by default.
  - Denser seeding from Corpus Callosum can also help, so in NDMG we do this.





Increasing biological validity

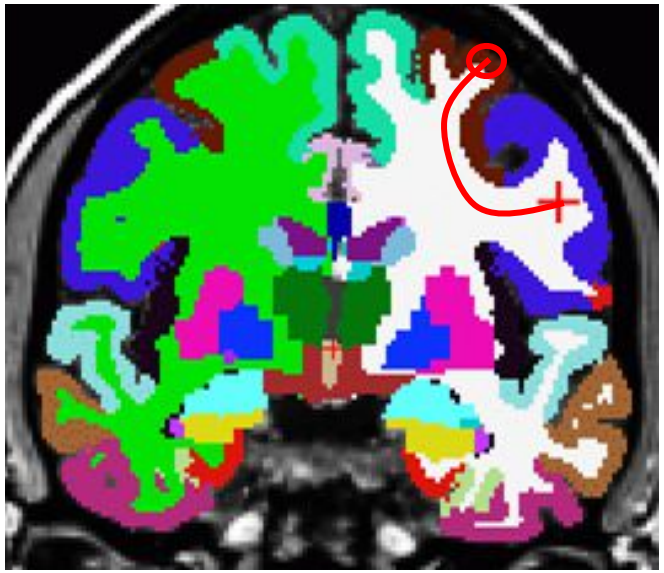
- Step 1: Use underlying tissue information (from the T1w) in the tracking:
  - Avoid ventricular CSF
  - Only track within tissue that has >0.05 probability of being WM.

- Step 2: Use streamline filtering (Malcolm et al., 2010)
  - By minimum fiber length (>10-60mm)
  - **May reduce false positives by 80-90%!**

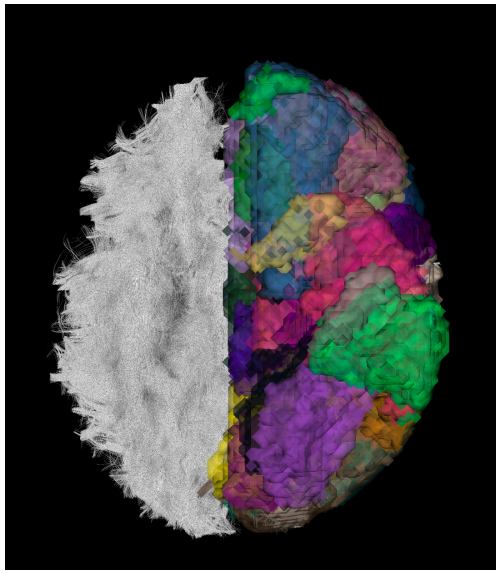
Graph Generation

Nodes = Grey matter regions of interest (from an atlas) containing streamline endpoints

Edges = Properties of white matter revealed by tractography streamlines (i.e. their number and integrity)

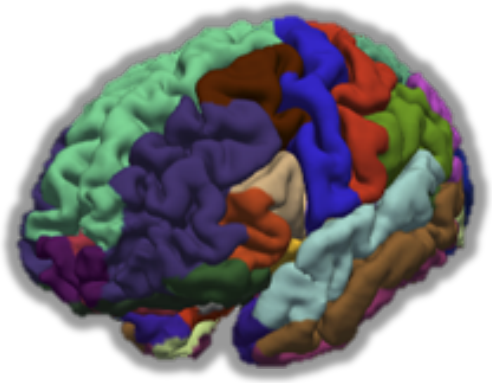


	TL	ML	MR	TR	th	G	Br
TL							
ML			cc		tML		cstL
MR		cc			tMR		cstR
TR							
th		tML	tMR				
G							
Br		cstL	cstR				

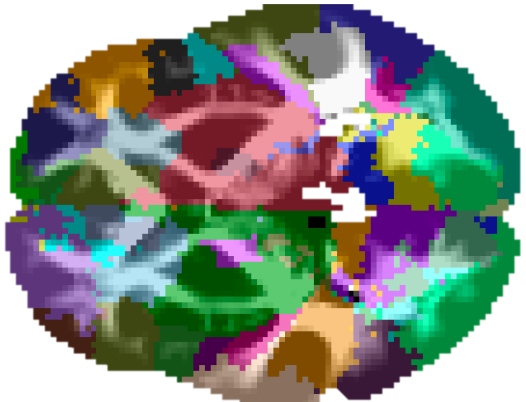


What counts as nodes? (Surfaces vs. Volumes)

Streamlines points intersecting with surfaces

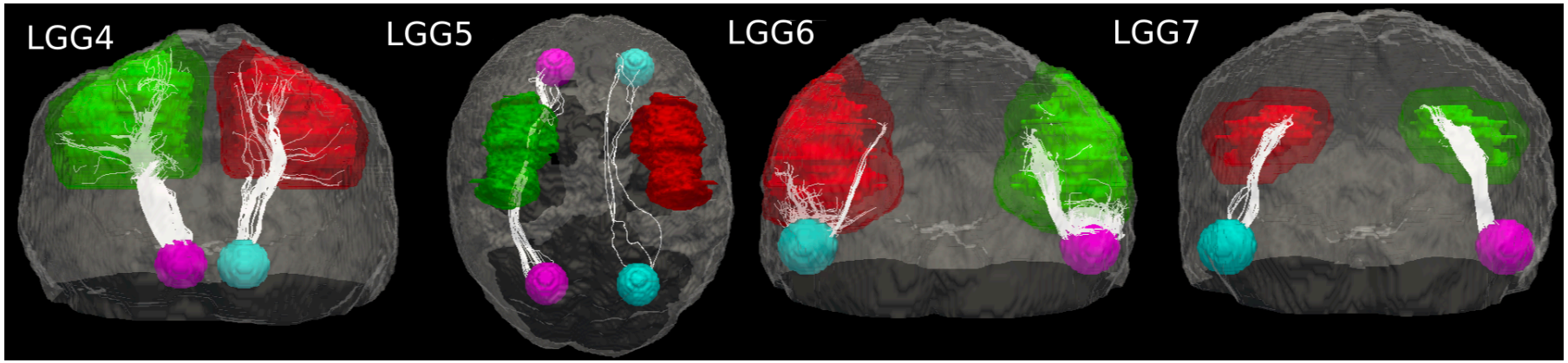


Streamline points intersecting with volumes



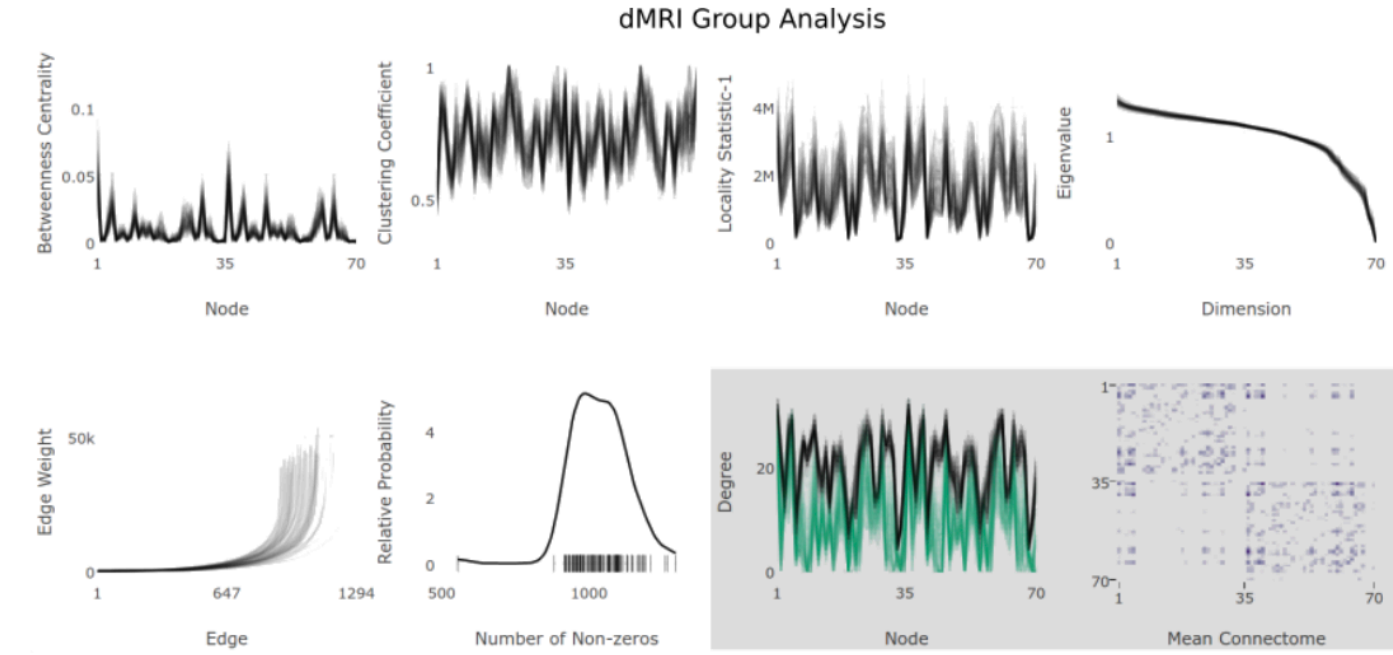
What counts as an edge?

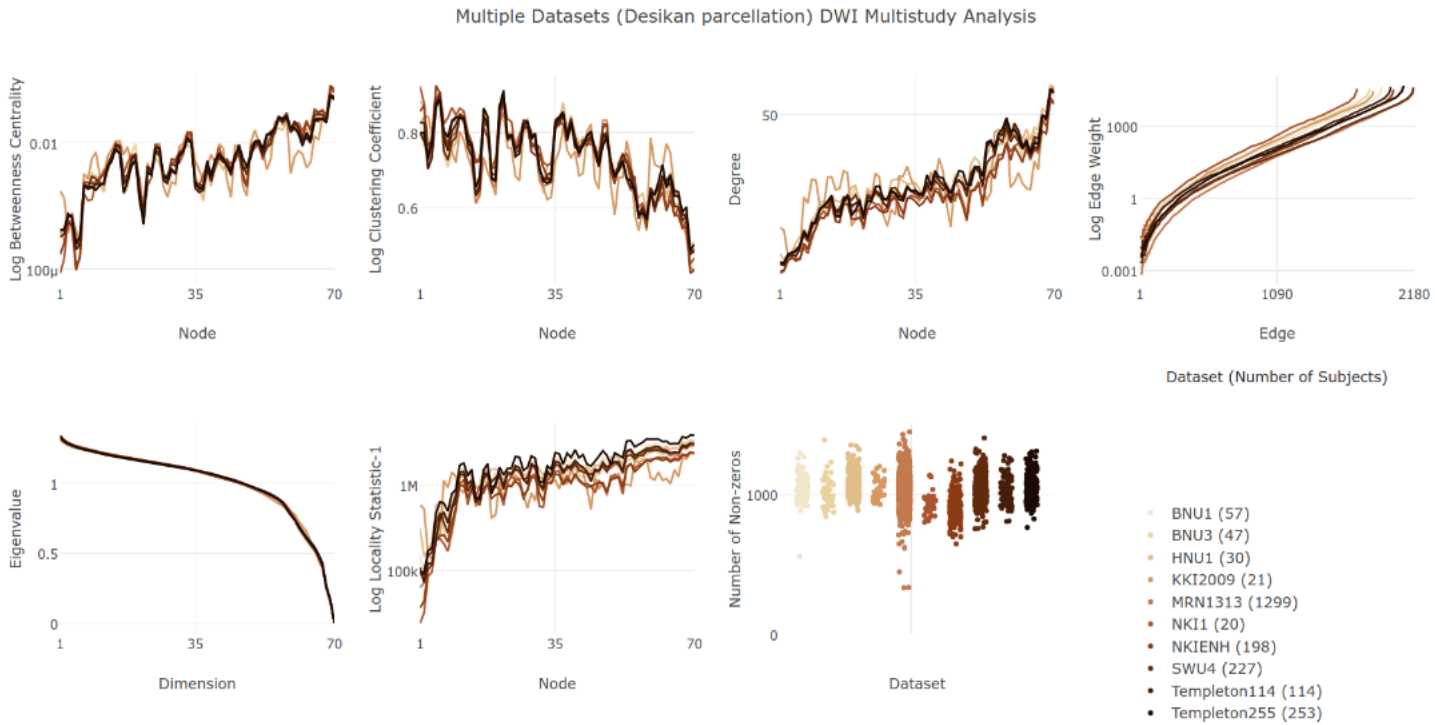
- Must volumetrically pass through the voxels of an atlas label (i.e. node).
  - What degree of overlap constitutes a “connection”?
  - Limitation of volumetric nodes (vs. more compute-demanding surface-style nodes)



Discriminability

Dataset	Scanner	# Dirs	Age (yrs)	% Male	# Subj's	Rep's	Total Scans	Discr
BNU1 [2]	Siemens	30	23.0 ± 2.3	53	57	2	114	0.999
SWU4 [2]	-	93	20.0 ± 1.3	51	235	2	454	0.815
NKI1 [2]	Siemens	137	34.4 ± 12.8	0	24	2	40	0.918
HNU1 [2]	GE	33	24.4 ± 2.3	50	30	10	300	0.989


dMRI Group Analysis
Betweenness Centrality
Node
Clustering Coefficient
Node
Locality Statistic-1
Node
Eigenvalue
Dimension
Edge Weight
Edge
Relative Probability
Number of Non-zeros
Degree
Node
Mean Connectome


Multiple Datasets (Desikan parcellation) DWI Multistudy Analysis
Log Betweenness Centrality
Node
Log Clustering Coefficient
Node
Degree
Node
Log Edge Weight
Edge
Dataset (Number of Subjects)
Eigenvalue
Dimension
Log Locality Statistic-1
Node
Number of Non-zeros
Dataset
BNU1 (57)
BNU3 (47)
HNU1 (30)
KKI2009 (21)
MRN1313 (1299)
NKI1 (20)
NKIENH (198)
SWU4 (227)
Templeton114 (114)
Templeton255 (253)

Ways to run NDMG:

- Locally on your laptop/desktop
- In a Docker or singularity container
- On Amazon Web Services (AWS)
- On HPC (e.g. SLURM scheduler)
- In Gigantum

dpisner@utexas.edu
https://github.com/dPys