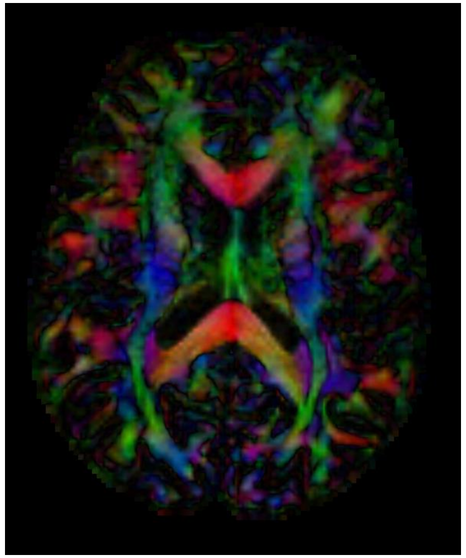
Sunnybrook Summer School: Diffusion MRI

Colleen Bailey, Liam Lawrence, Rachel Chan

August 18, 2025



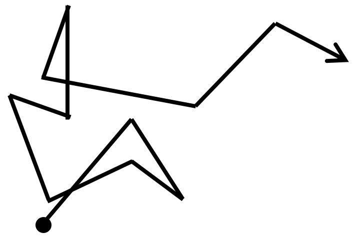
Outline

- Diffusion Basics
- Diffusion Sequences I (diffusion encoding)
  - Qualitative: diffusion weighted imaging
  - Quantitative: b-values and apparent diffusion coefficient
  - Anisotropy: diffusion tensor imaging
- Diffusion Sequences II (image encoding)
  - Tradeoffs
  - Artefacts
- Processing Pipeline and Analysis Tools

Water Motion: A relationship between space and time

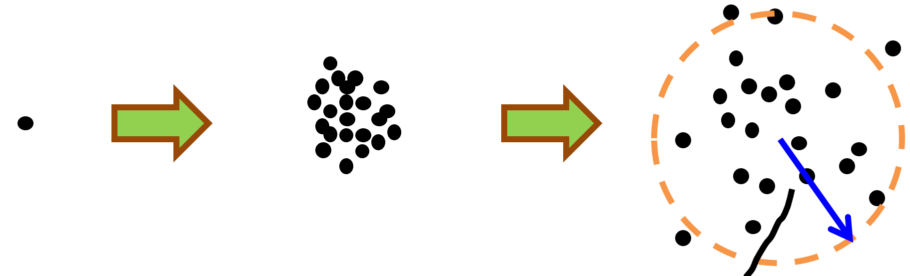
Brownian motion

“Random walk”



t=0

After time Δt



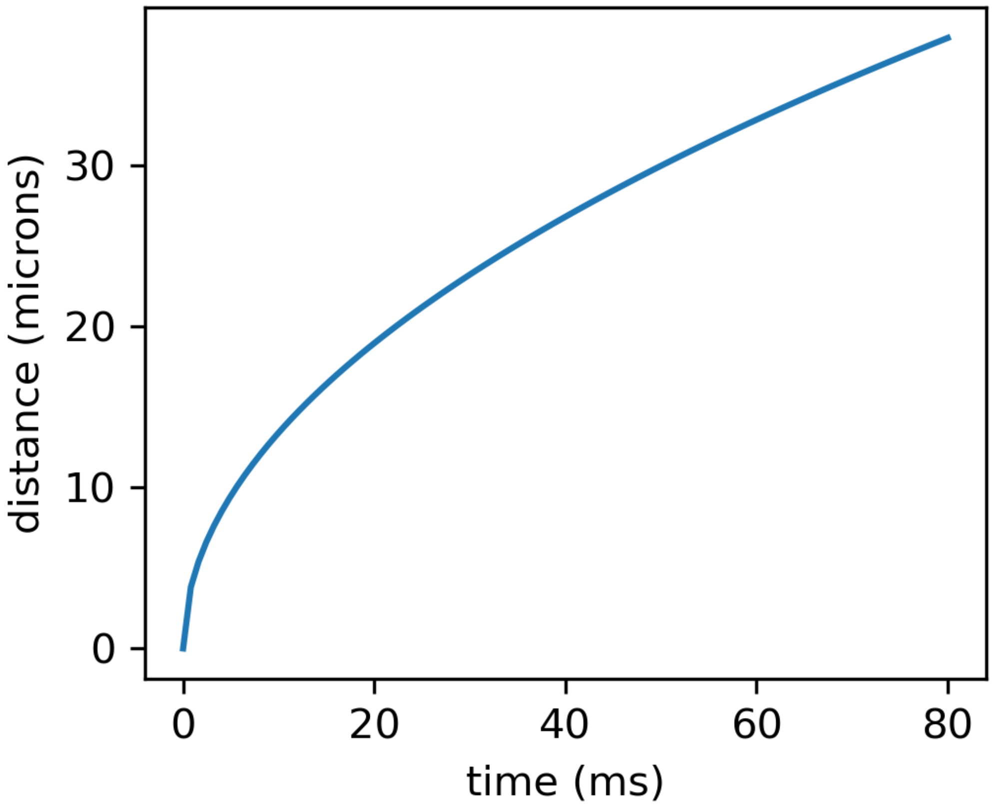
Diffusion coefficient
(temperature dependent)

Displacement $\Delta r = \sqrt{6D\Delta t}$

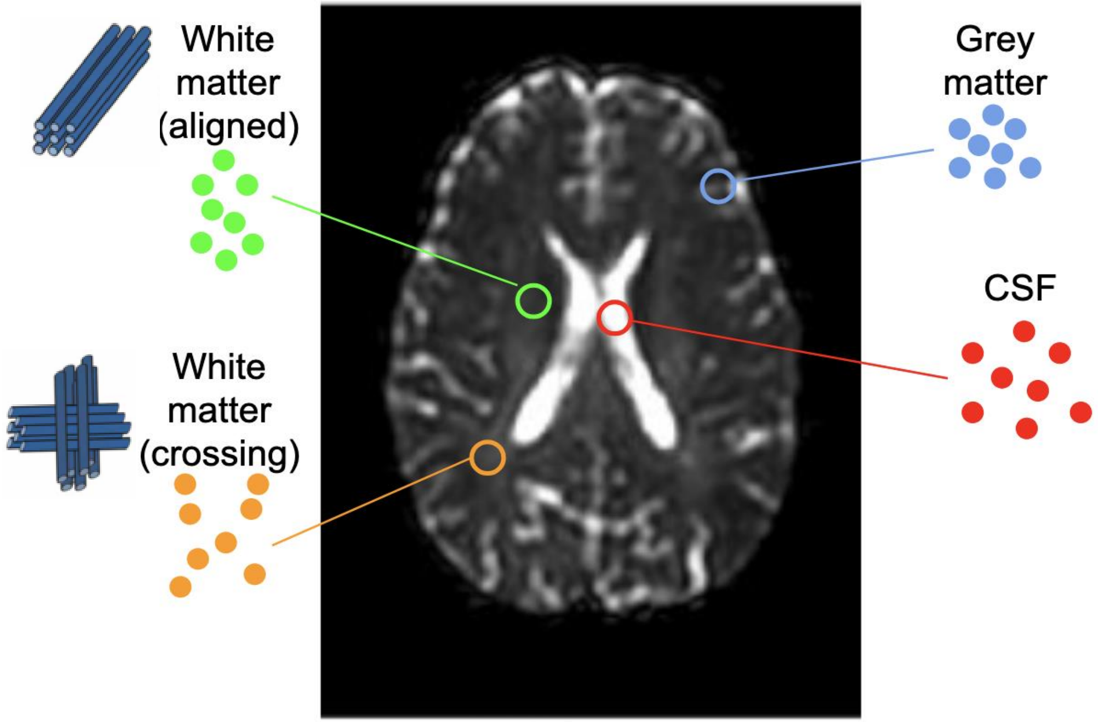
time

A rough idea of scale

Water at 37°C: D ~ 3 $\mu m^2/ms$



Example Probability Density Functions



MRI encodes spatial information using gradients

For Imaging

- The **phase and readout gradients** encode spatial location
- Acquire data in **k-space**

$$\mathbf{k} = \gamma \int g\left(\mathbf{r}, t\right) dt$$

$$S\left(\mathbf{k}\right) = \int \rho\left(\mathbf{r}\right) e^{-2\pi i \mathbf{k}\cdot\mathbf{r}} d\mathbf{r}$$

- Further out in k-space, higher spatial frequencies

For Diffusion

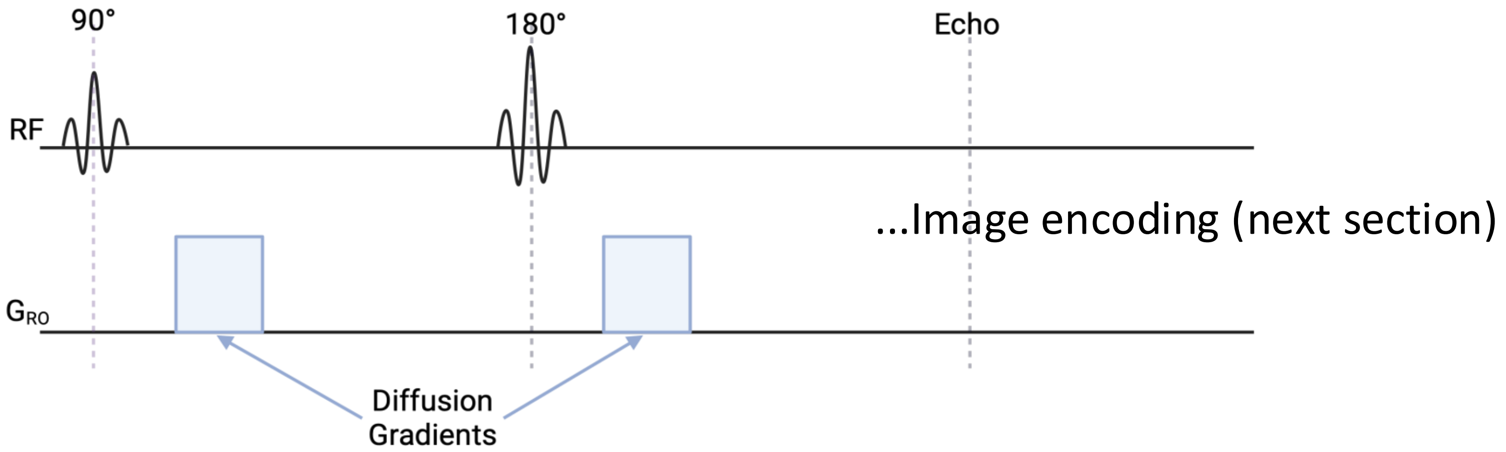
- The **diffusion gradients** encode spatial location
- Acquire data in **q-space**

$$\mathbf{q} = \gamma \int g\left(\mathbf{r}, t\right) dt$$

$$S\left(\mathbf{q}\right) = \int P\left(\mathbf{r}, t\right) e^{-2\pi i \mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$$

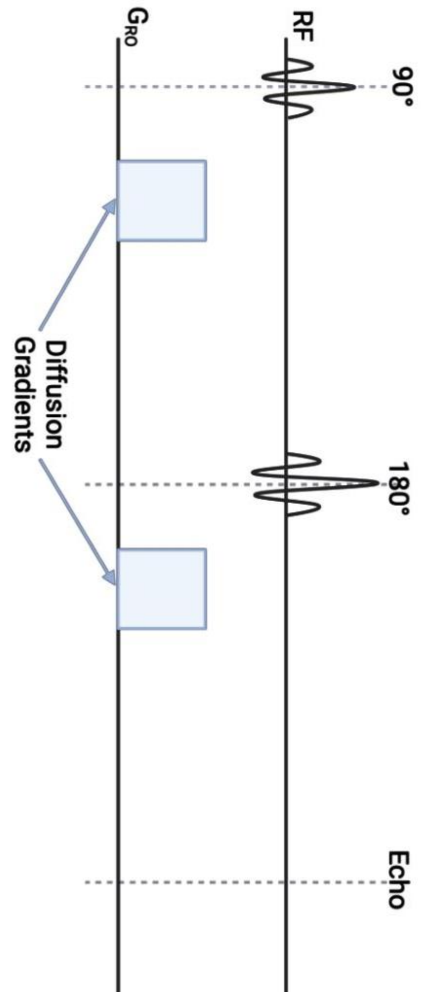
- Further out in q-space, more sensitive to small displacements
- But also dependence on diffusion time

A diffusion pulse sequence walkthrough

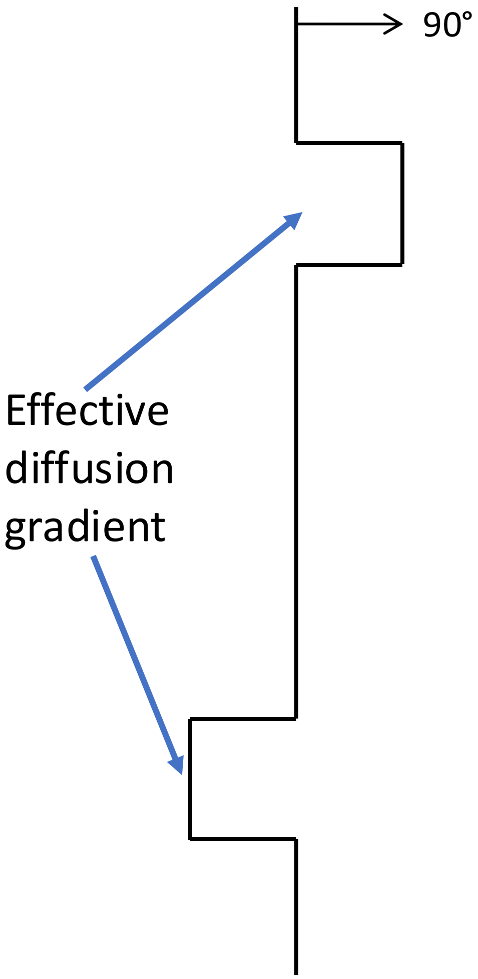


- Read left-to-right as a series of scanner instructions about rf pulses and gradients
- Most common diffusion sequence: Pulsed Gradient Spin Echo (PGSE) AKA Stejskal-Tanner diffusion sequence (1965)
- To explain how this sequence is sensitive to water motion, consider two situations: without and with motion.

A diffusion pulse sequence walkthrough

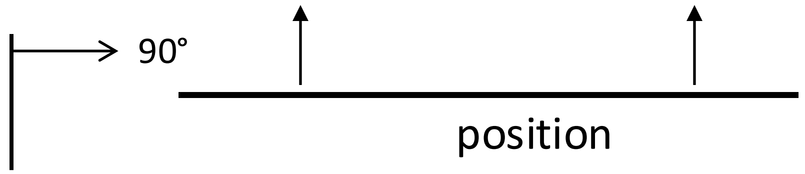


A diffusion pulse sequence walkthrough

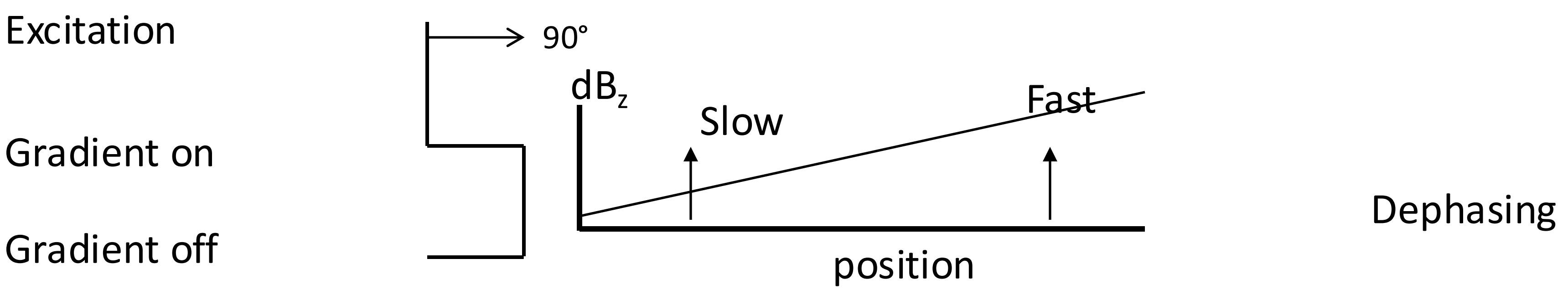


Consider two spins in different positions after a 90^o pulse

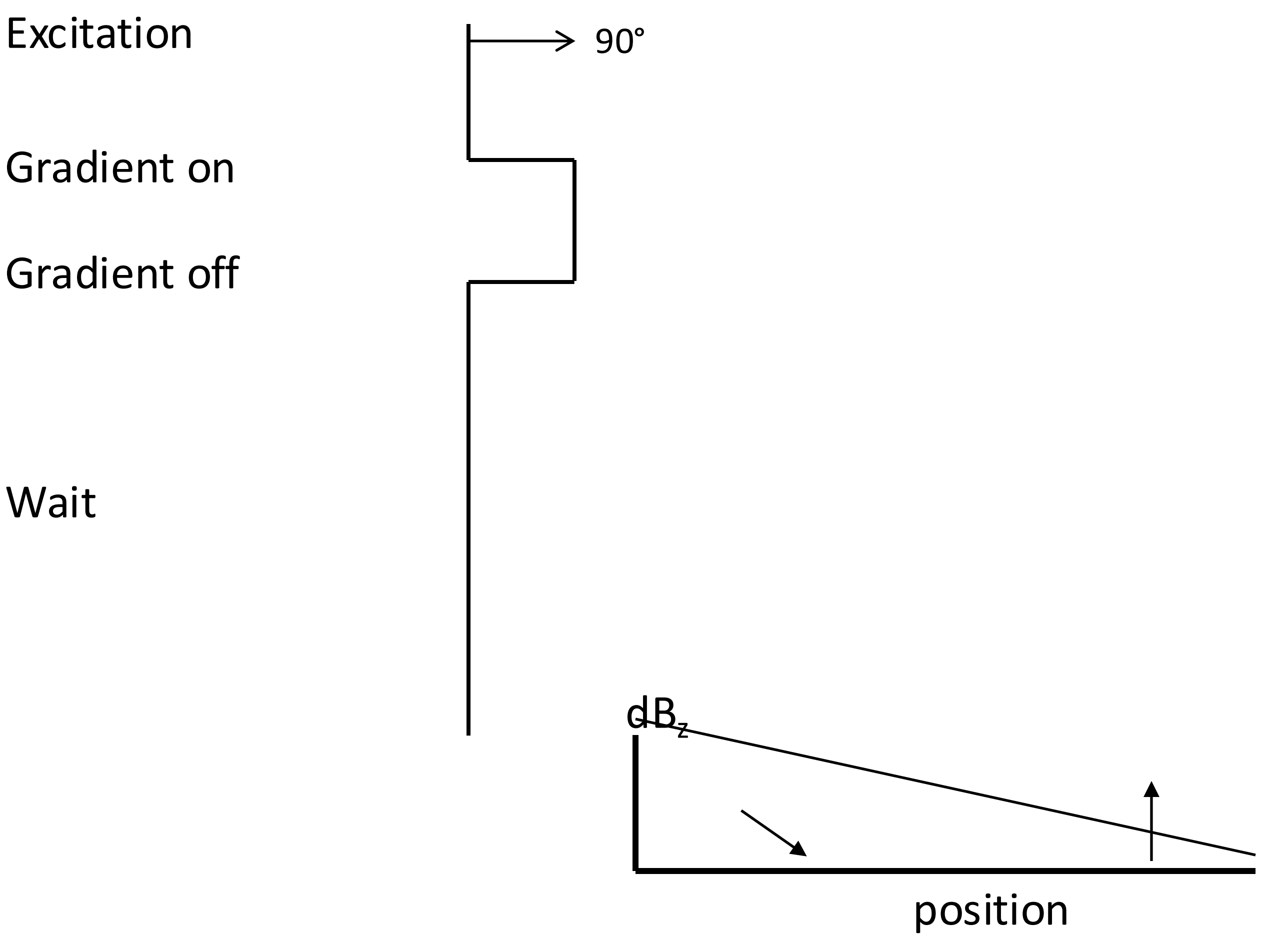
Excitation



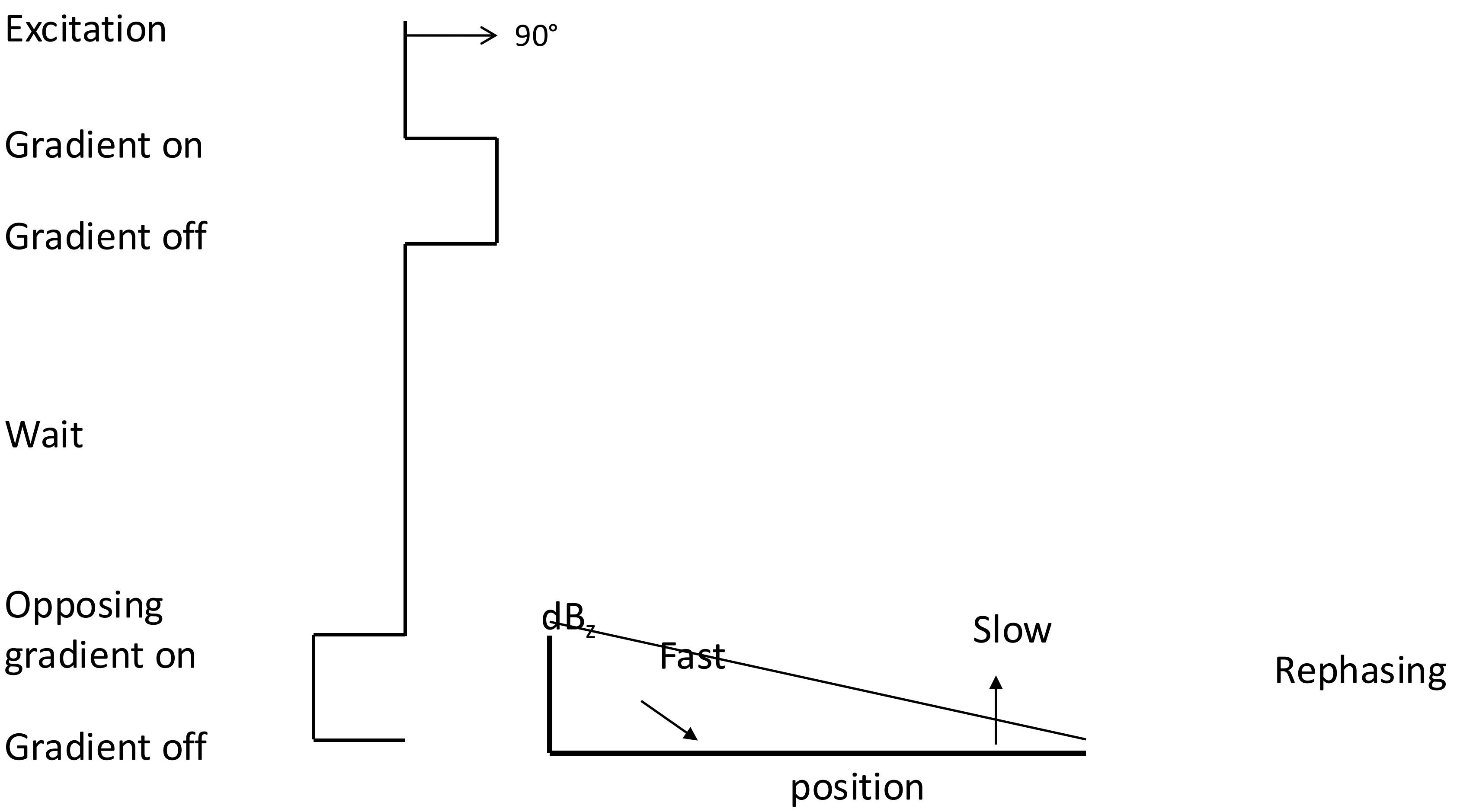
Turning on a magnetic field gradient means spins in different positions precess at different frequencies. When the gradient turns off, the spins are dephased.



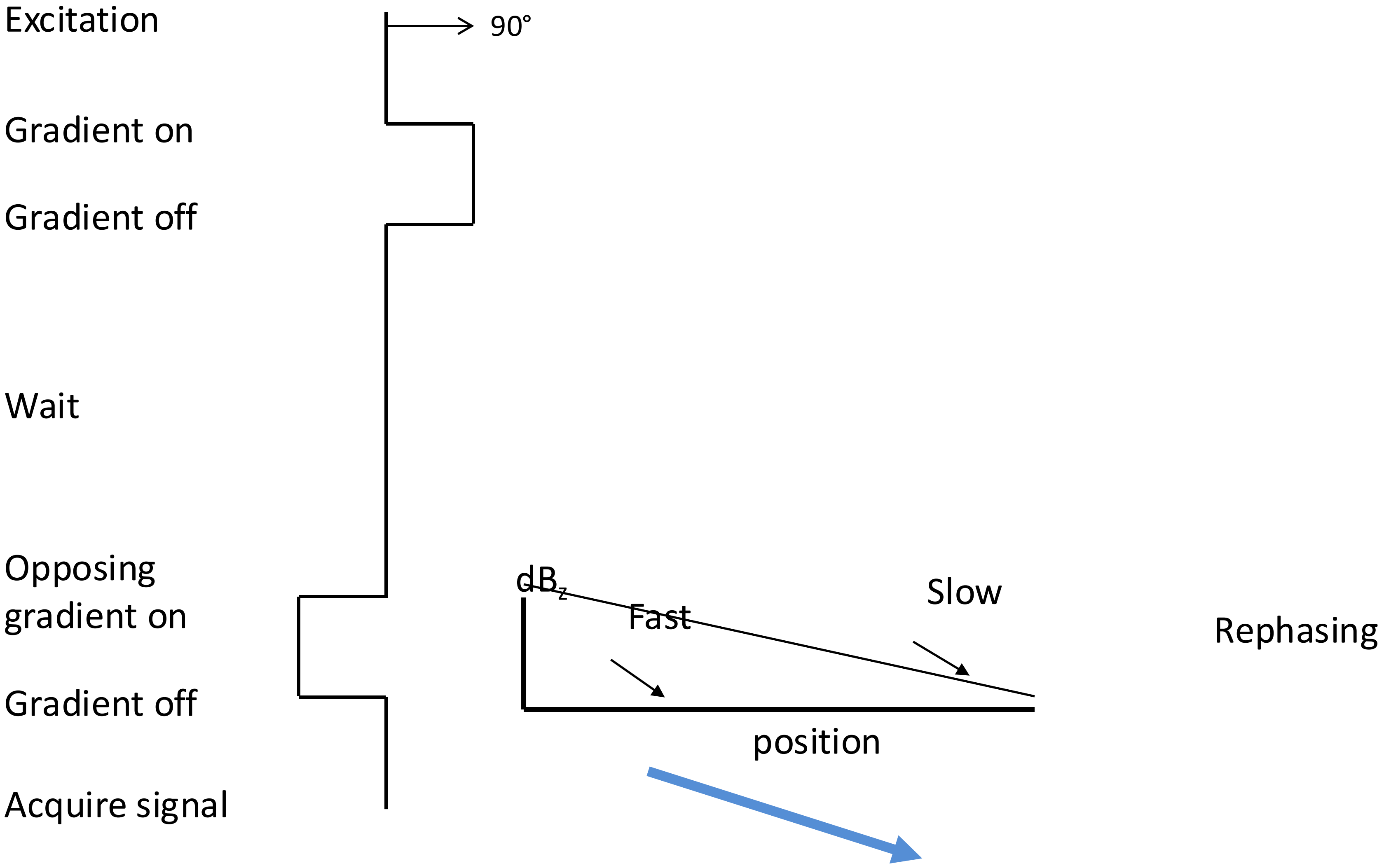
Then wait. During this diffusion time, the spins may move. First consider the case with no motion.



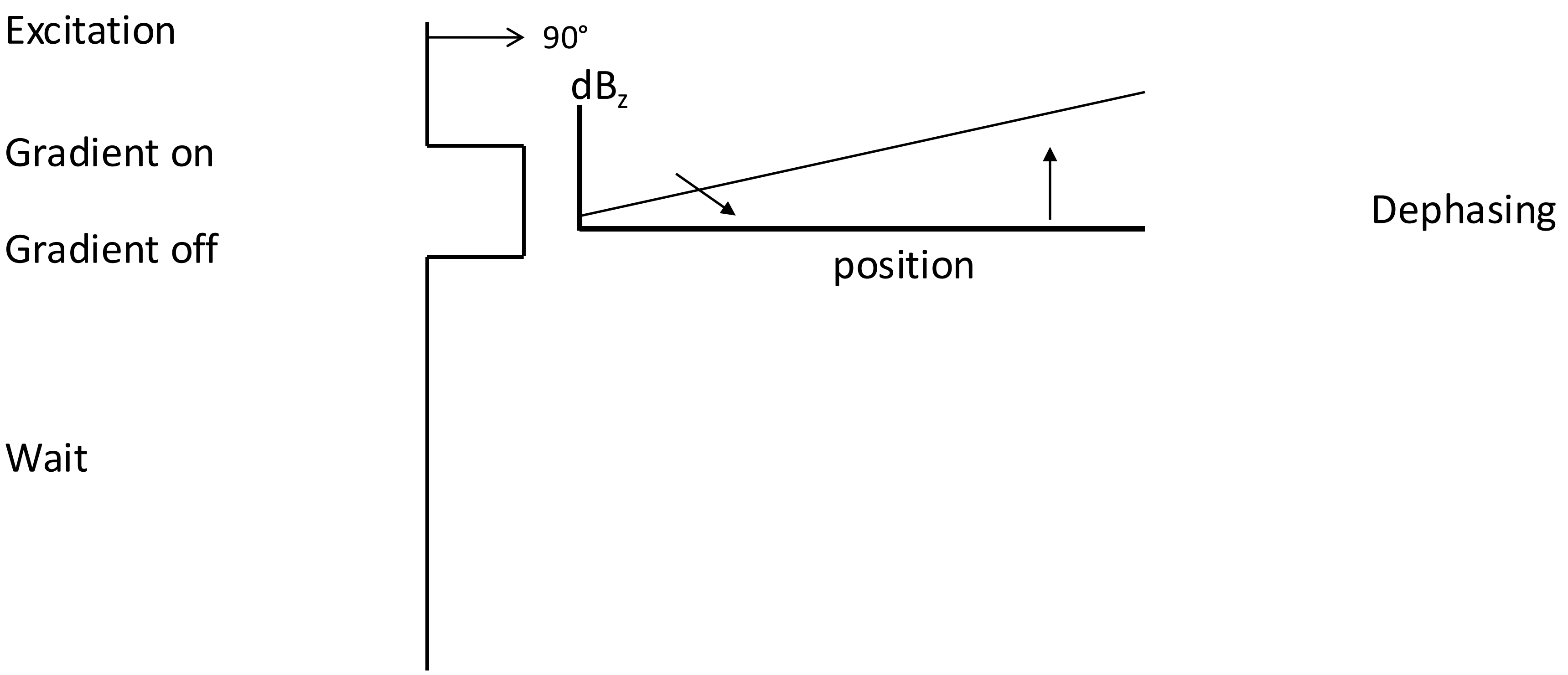
When gradients are applied in the opposite direction, the left spin now precesses faster than the right. If this gradient is on for the same time, the spins rephase.



If we read out (sum) the signal from all spins in the voxel, they will all be aligned and signal will be high



Now consider spins that move during the diffusion time



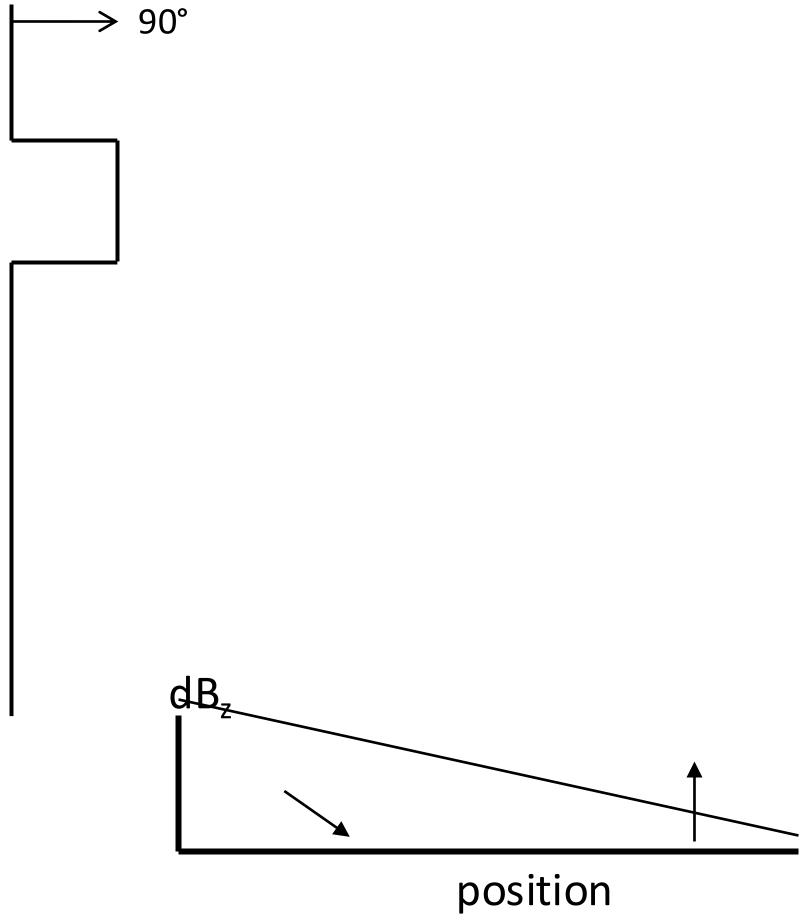
A dephased spin moves…

Excitation

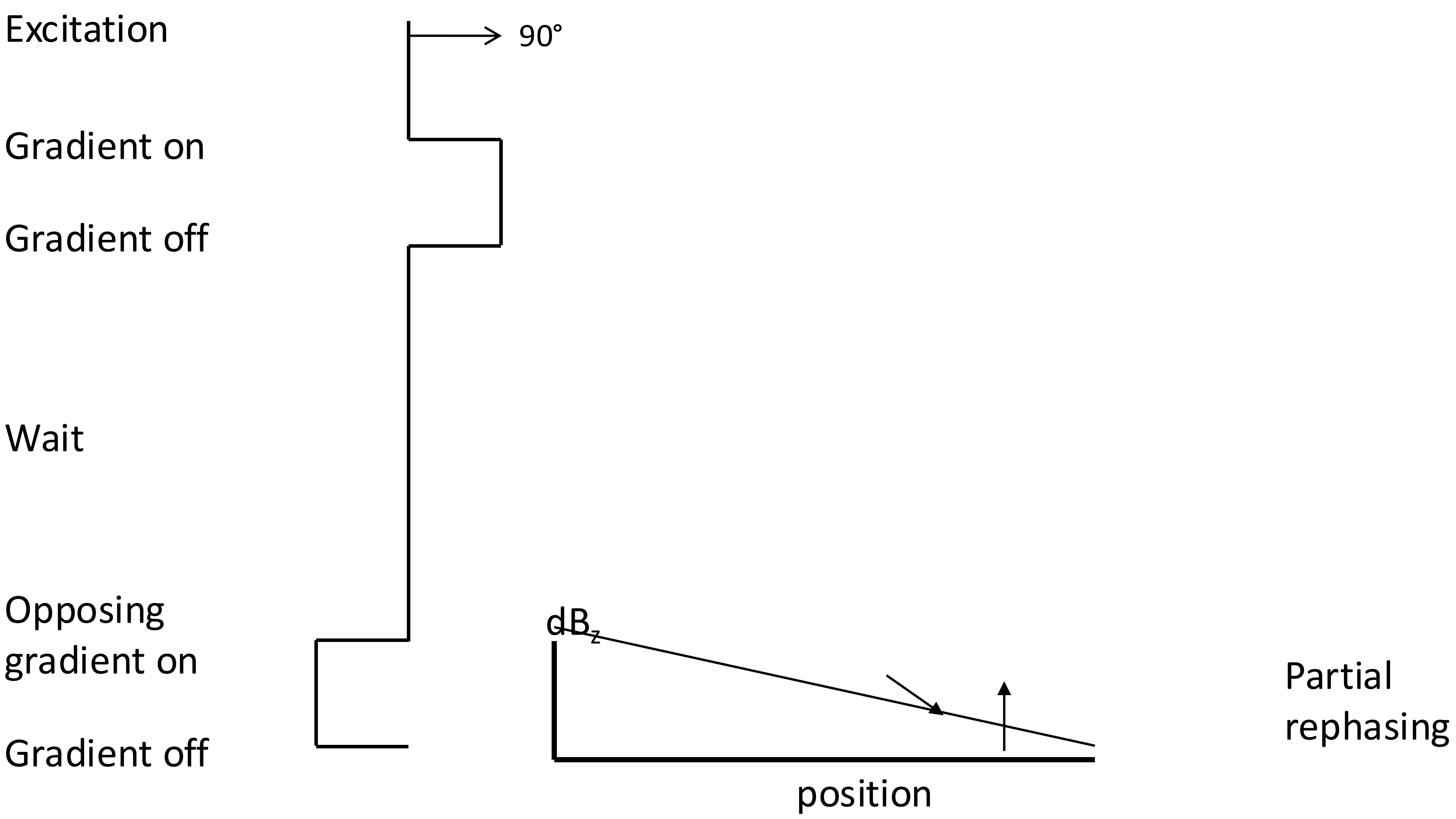
Gradient on

Gradient off

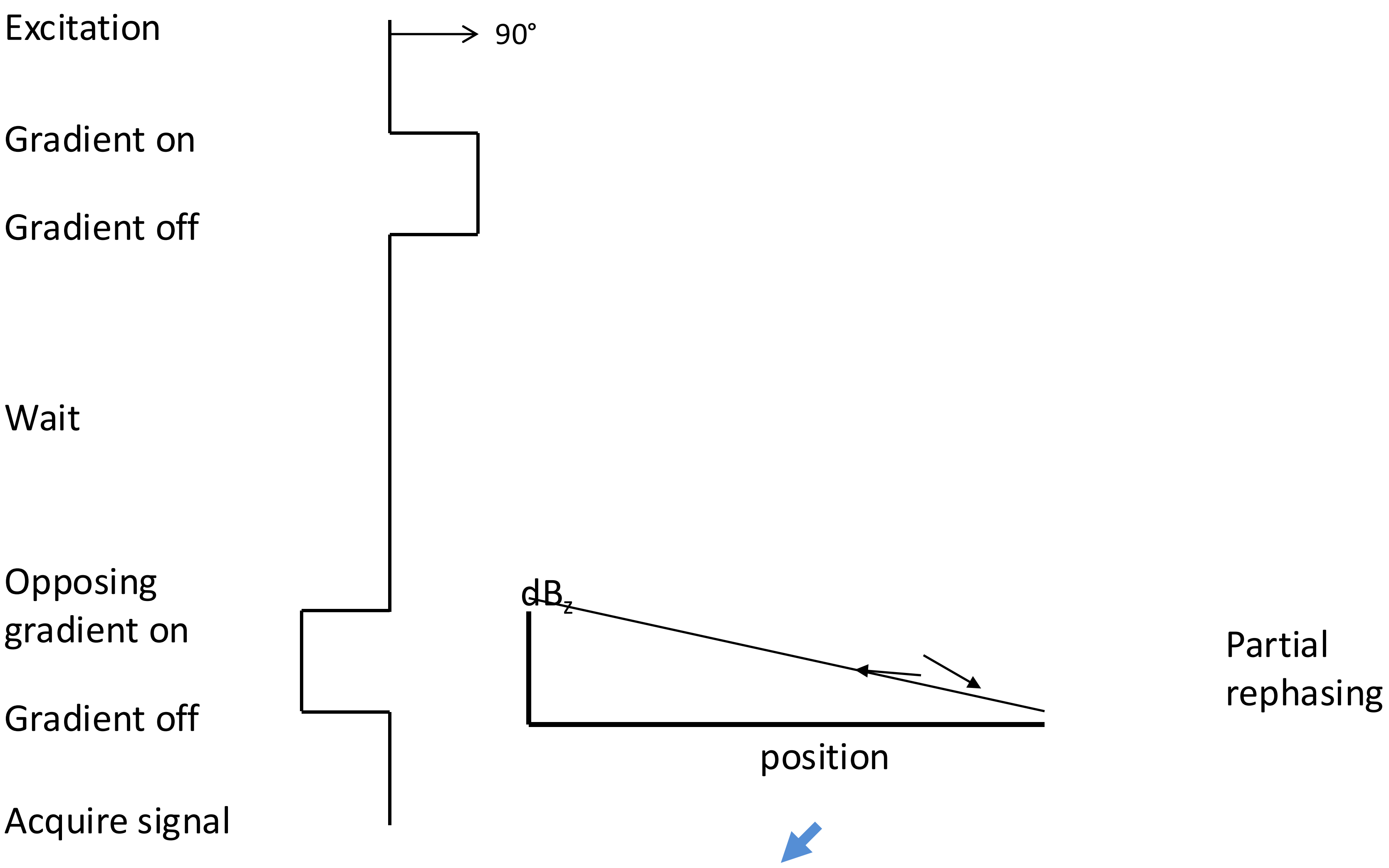
Wait



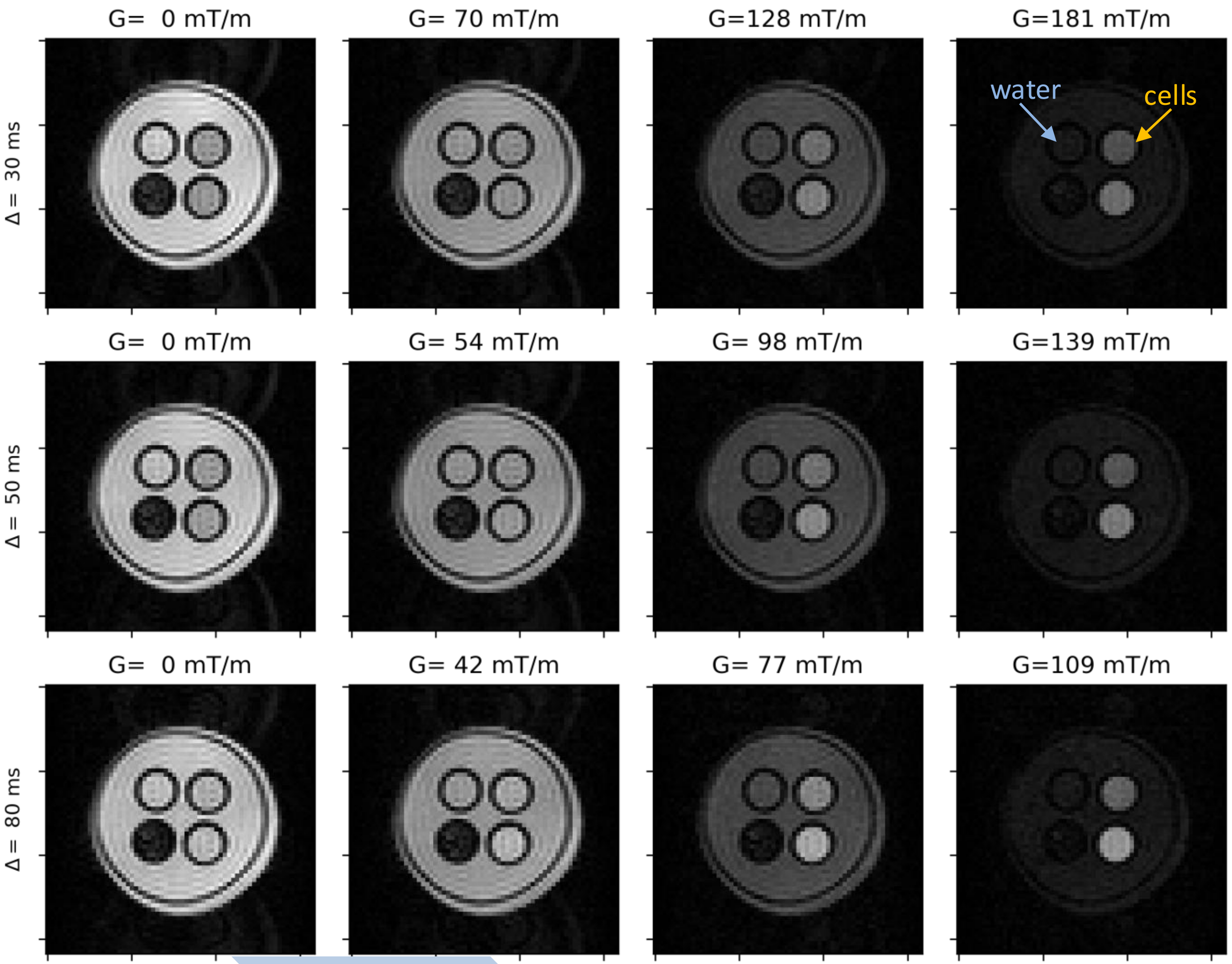
The gradient at the new position is not as large. The spins do not completely rephase.

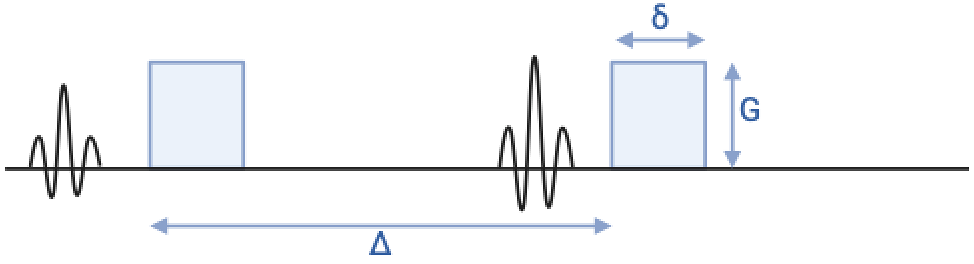


Summing all spins at all positions in the voxel, they are no longer in phase and we get a much smaller signal (the signal attenuates more when spins have moved more)



When water moves, DWI signal gets lower





How much lower depends on

- Scan parameters: gradient amplitude (G), gradient duration (δ), gradient separation (Δ)
- Tissue properties

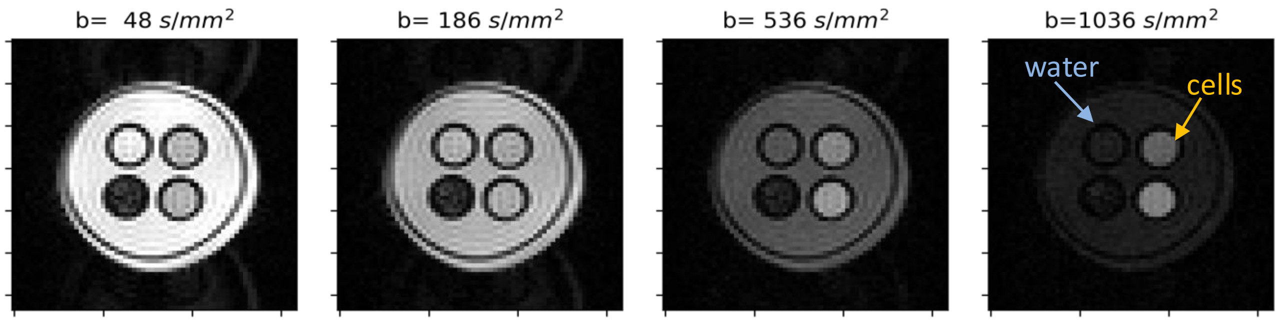
Can we go beyond DWI and quantify diffusion?
Where are these parameters on the scanner?
Which should you set and to what values?

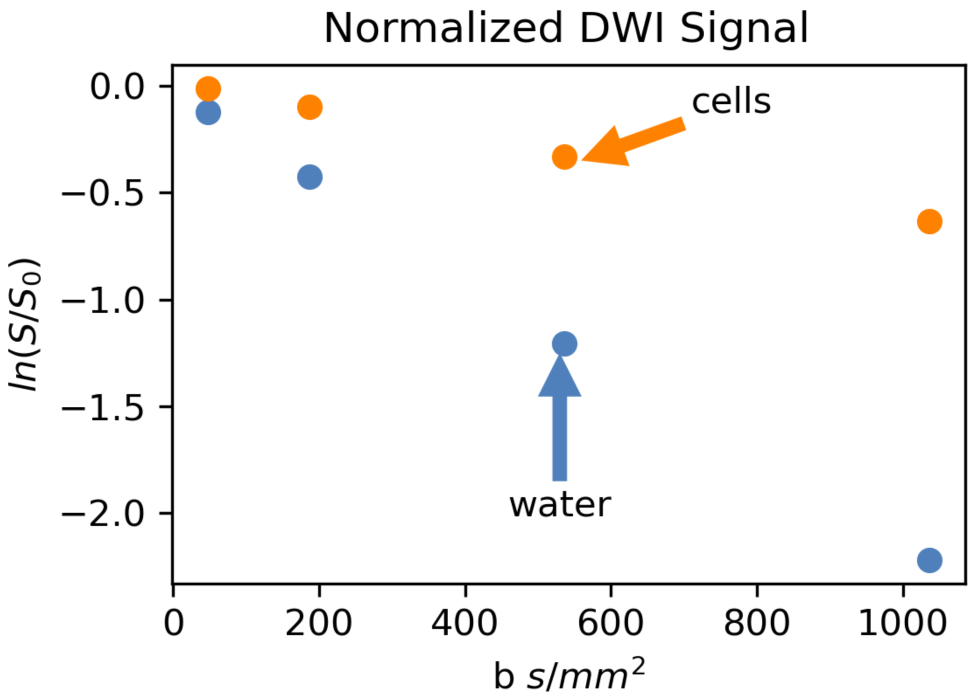
Quantification: b-value and Apparent Diffusion Coefficient (ADC)

- The b-value summarizes the scan parameters into one number: $b=(\gamma G\delta)^2(\Delta-\delta/3)$

$$\mathbf{q} = \gamma \int g(\mathbf{r}, t)\, dt \quad S(\mathbf{q}) = \int P(\mathbf{r}, t)\, e^{-2\pi i \mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$$

- For Gaussian diffusion: $\ln(S/S_0)=-b*ADC$



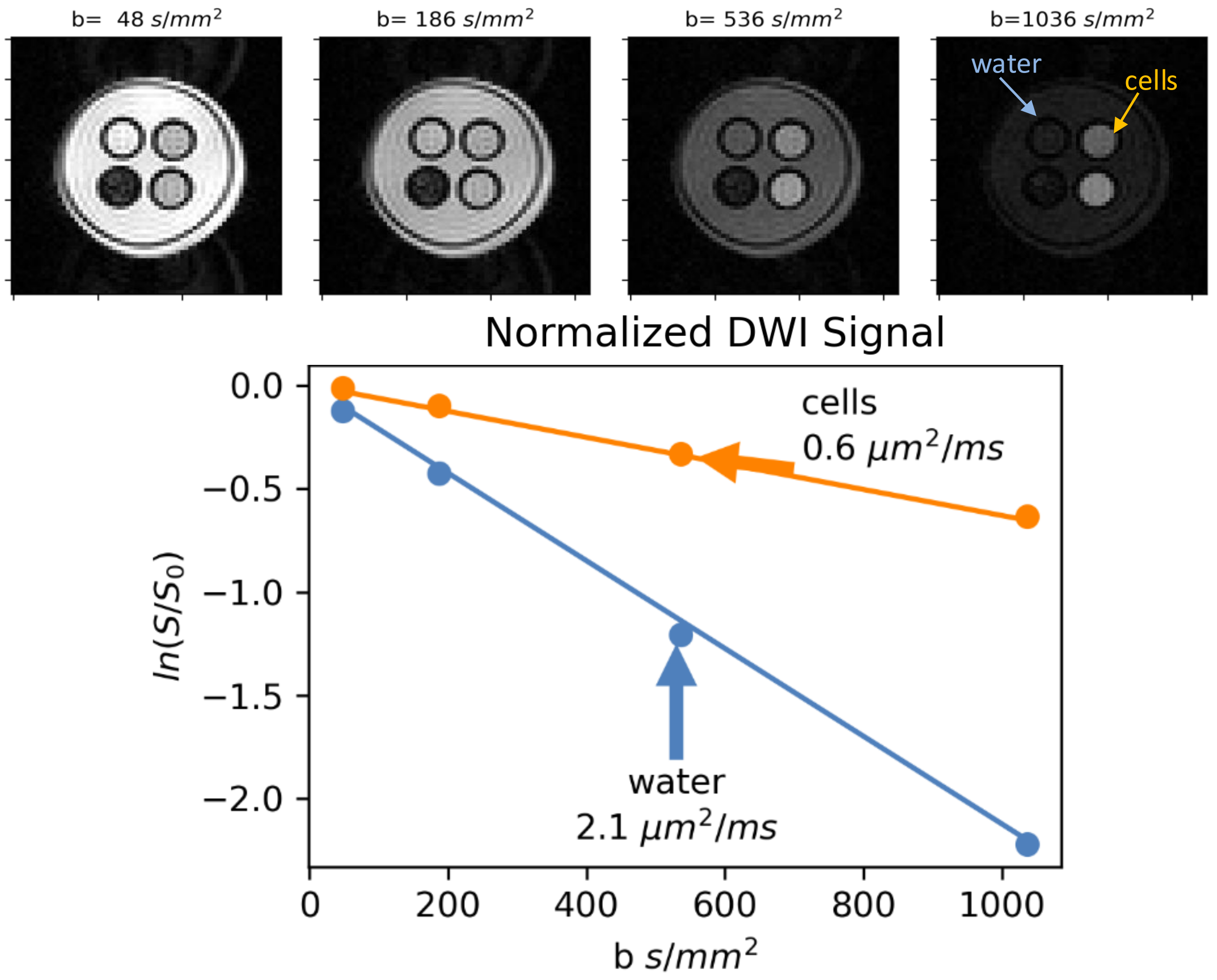


Quantification: b-value and Apparent Diffusion Coefficient (ADC)

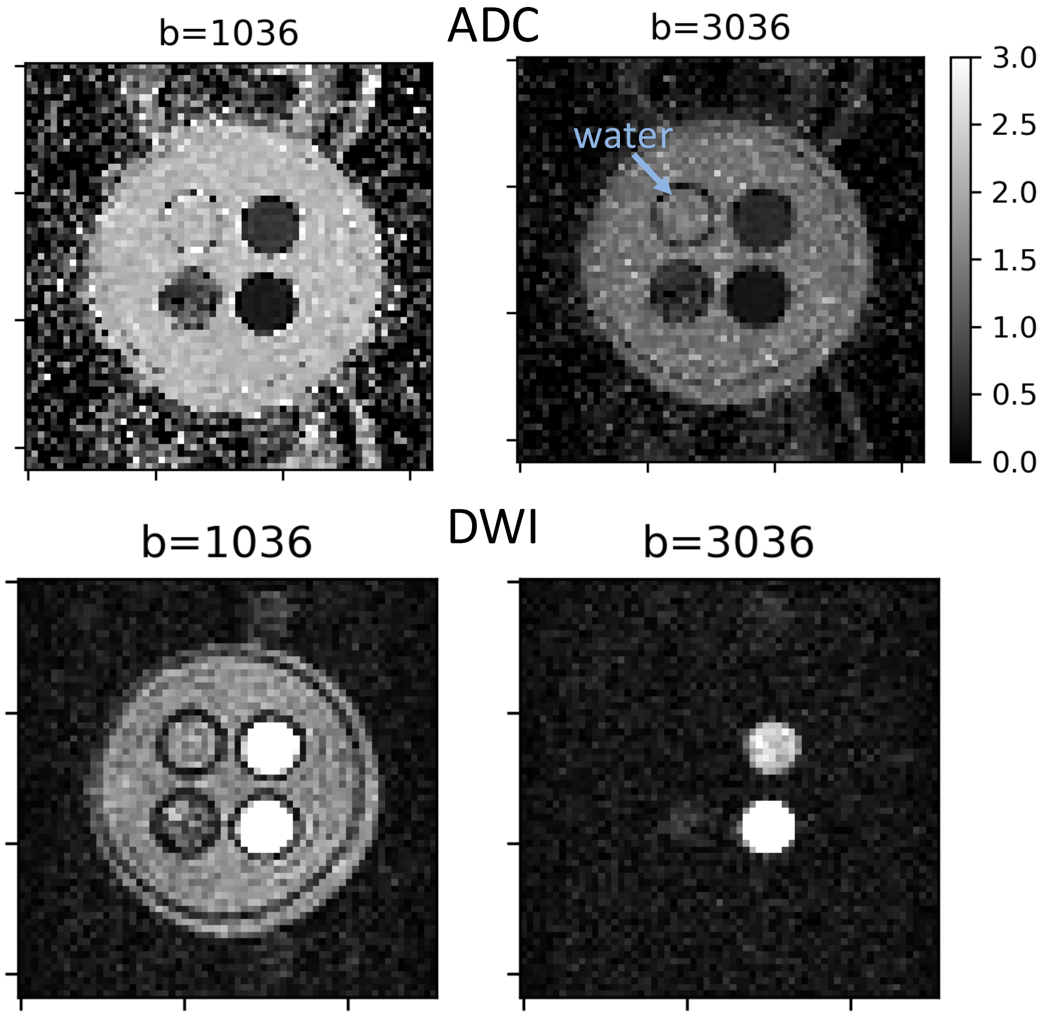
- The b-value summarizes the scan parameters into one number: $b=(\gamma G\delta)^2(\Delta-\delta/3)$

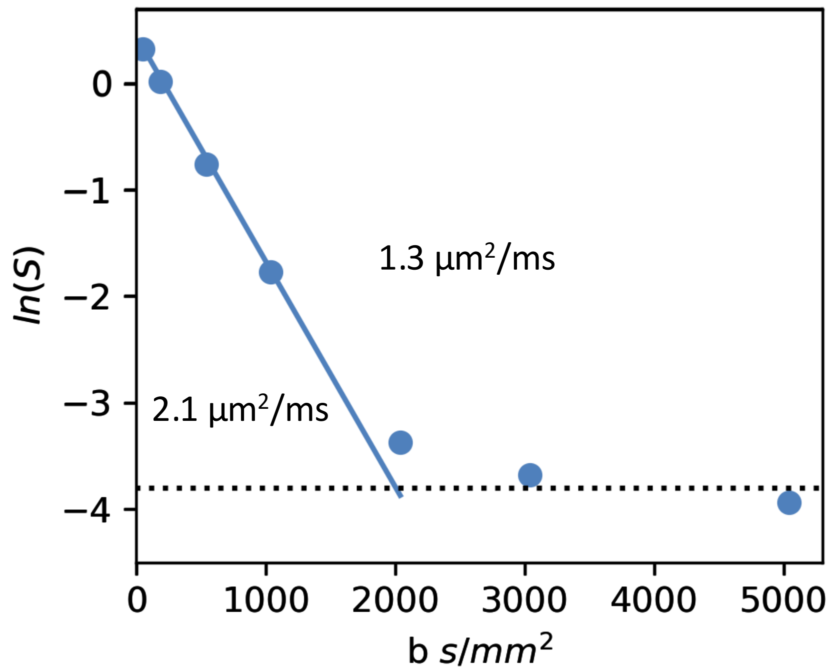
$$\mathbf{q} = \gamma \int g(\mathbf{r}, t)\, dt \quad S(\mathbf{q}) = \int P(\mathbf{r}, t)\, e^{-2\pi i \mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$$

- For Gaussian diffusion: $\ln(S/S_0)=-b*ADC$
- Can calculate ADC from 2 b-values



Why don't these ADCs for water agree?





It's noise!

Always check the DWI ***and*** the ADC map

SNR limits are one reason diffusion has larger voxels than structural

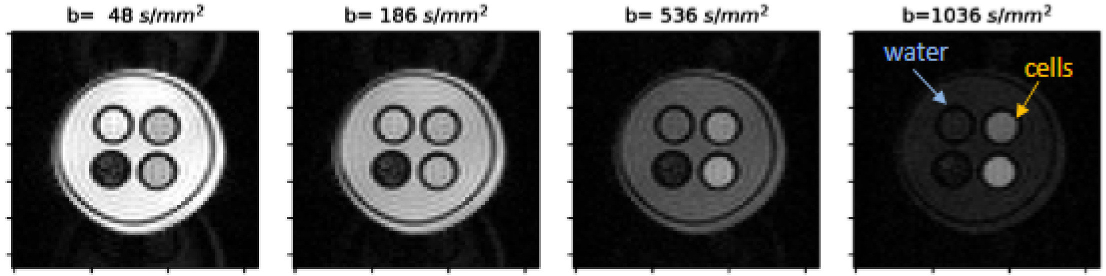
Why don't these ADCs for cells agree?

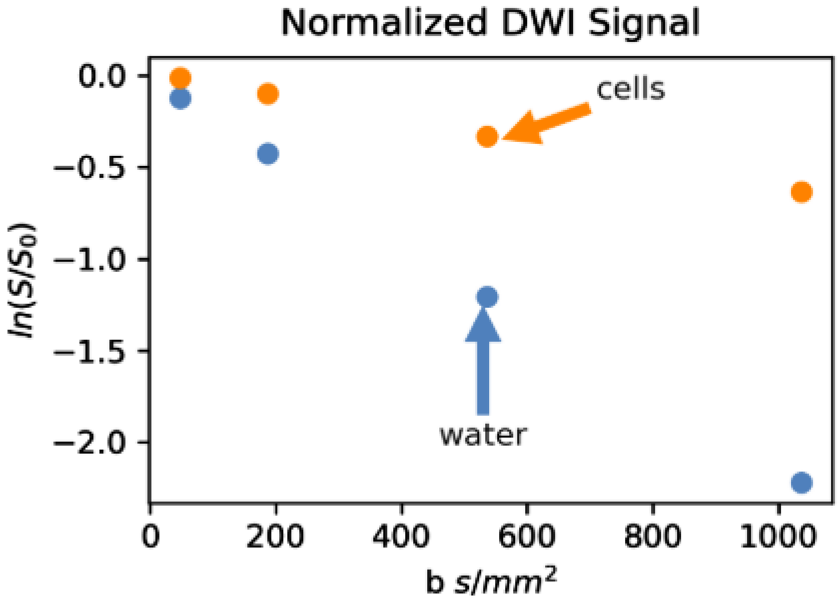
Quantification: b-value and Apparent Diffusion Coefficient (ADC)

- The b-value summarizes the scan parameters into one number: $b=(\gamma G\delta)^2(\Delta-\delta/3)$

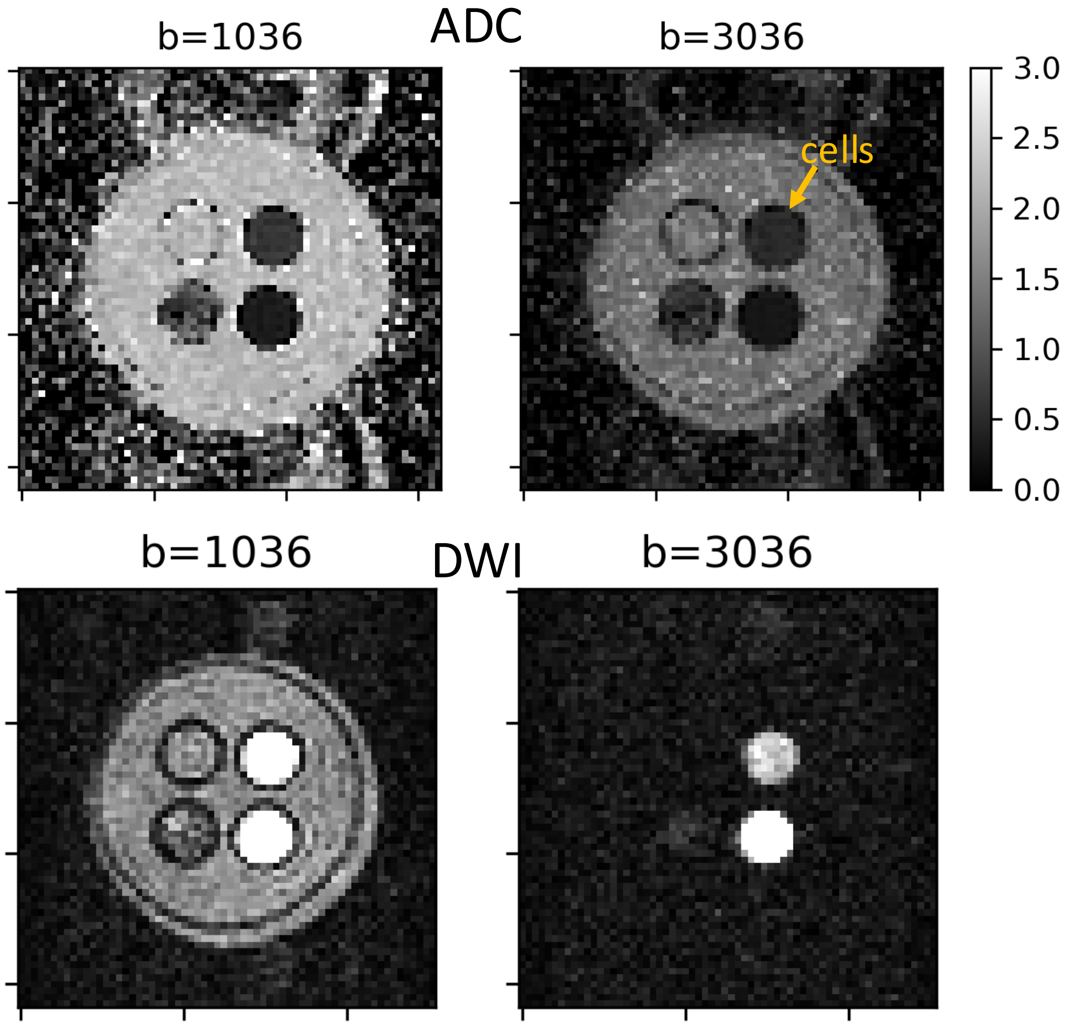
$$\mathbf{q} = \gamma \int g(\mathbf{r}, t)\, dt \quad S(\mathbf{q}) = \int P(\mathbf{r}, t)\, e^{-2\pi i \mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$$

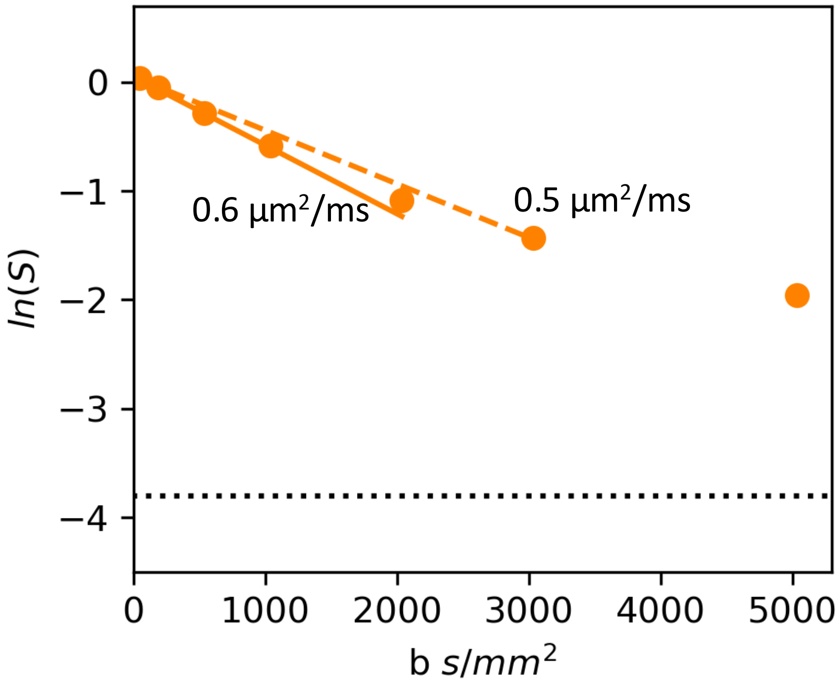
- For Gaussian diffusion: $\ln(S/S_0)=-b*ADC$





Why don't these ADCs for cells agree?





Water in cells does not undergo Gaussian diffusion at high b!

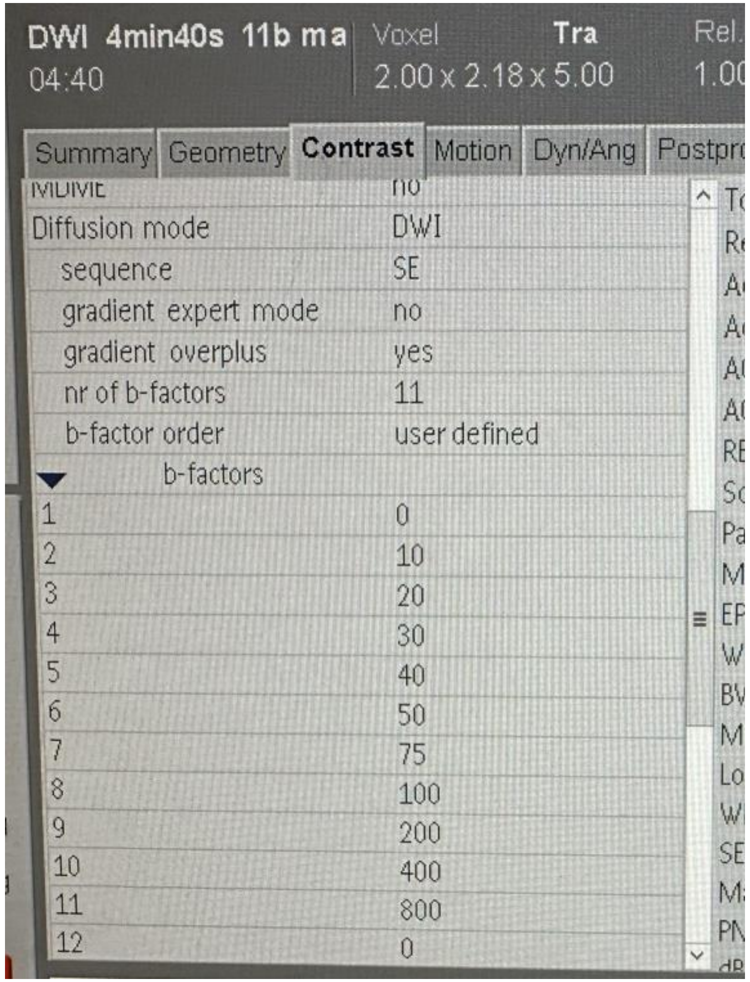
Even if you see signal in DWI, ADC may not be appropriate at high b (>1000 s/mm^2)

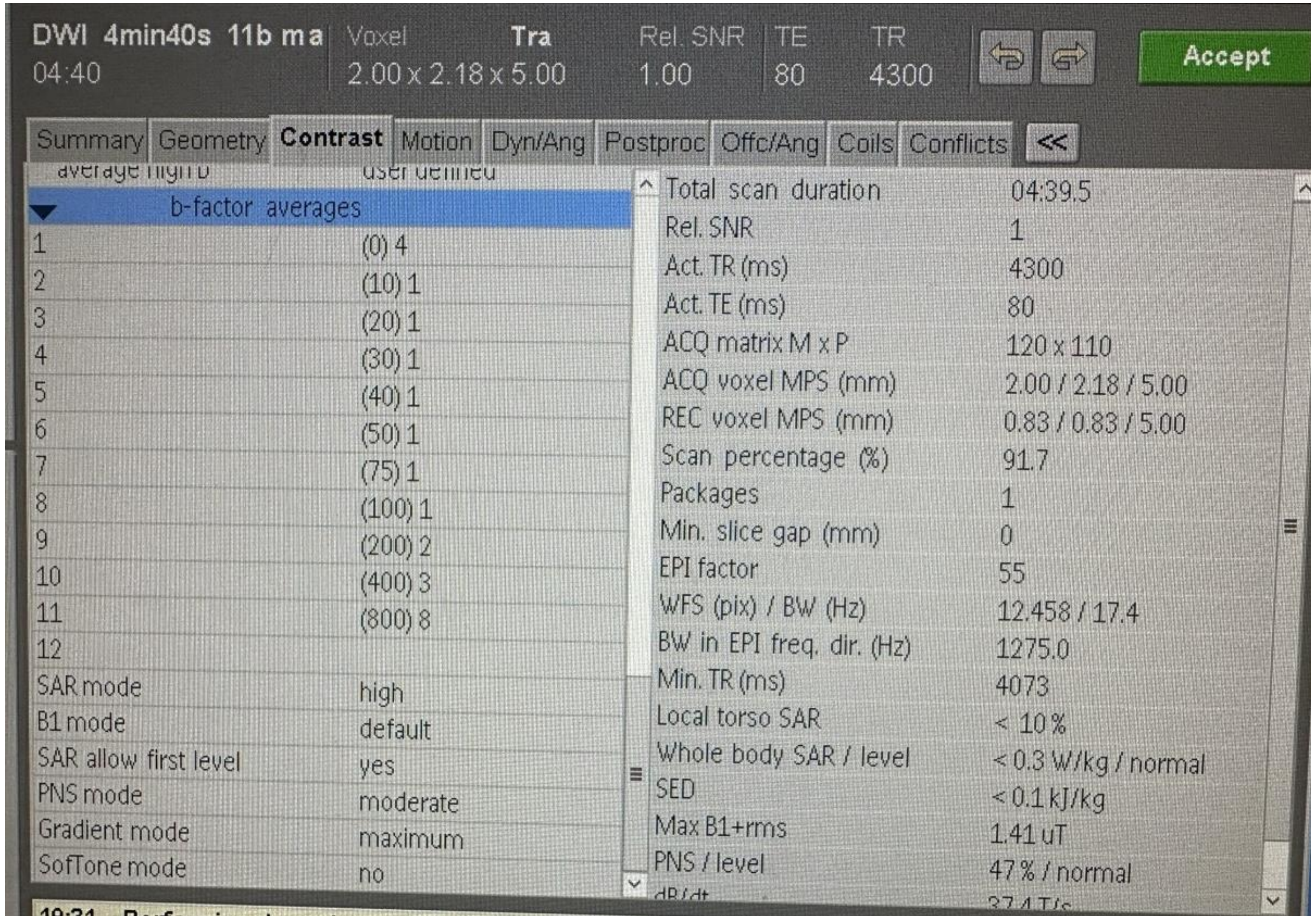
There are hardware and practical limitations

- The scanner has a maximum gradient amplitude
- TE limits gradient duration and separation
- Longer TEs have lower SNR (T2 decay)
- A "b-value" may not mean the same thing on different scanners.

Scanner (3 T)	Max Gradient (mT/m)
GE Signa Architect	44
GE Signa Premier	80
Philips Ingenia Elition	45
Philips MR 7700	65
Siemens Skyra	45
Siemens Prisma	80
Siemens Connectom	300

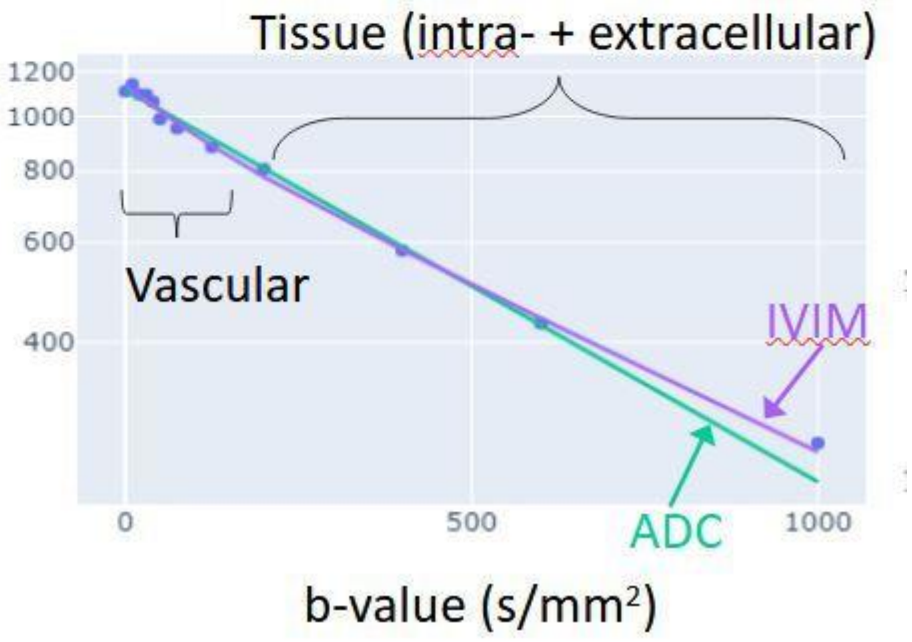
Example Scanner Interface





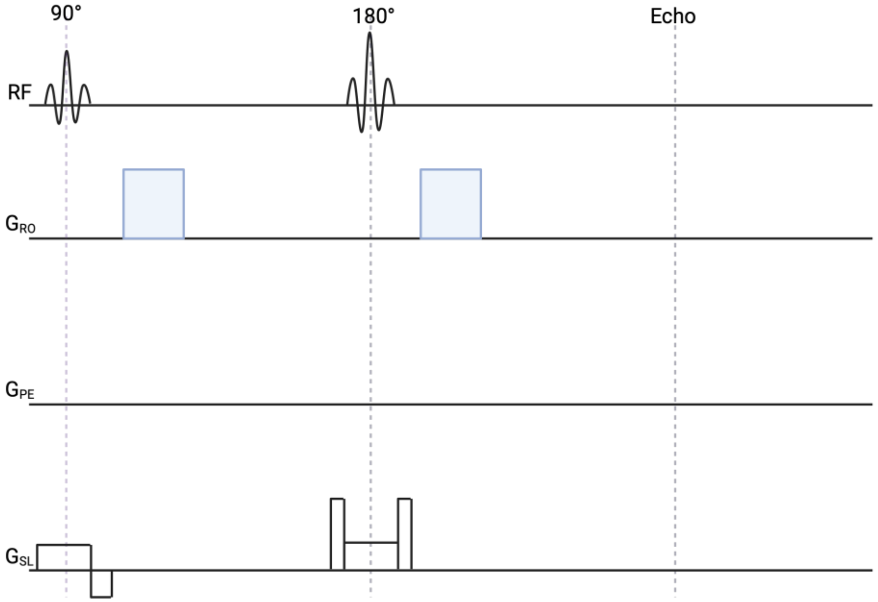
Other Considerations for DWI and ADC

- ADC is commonly used for quantification at low b (exact value is scanner-dependent)
- Kurtosis or axon diameter mapping can be used for quantification at high b

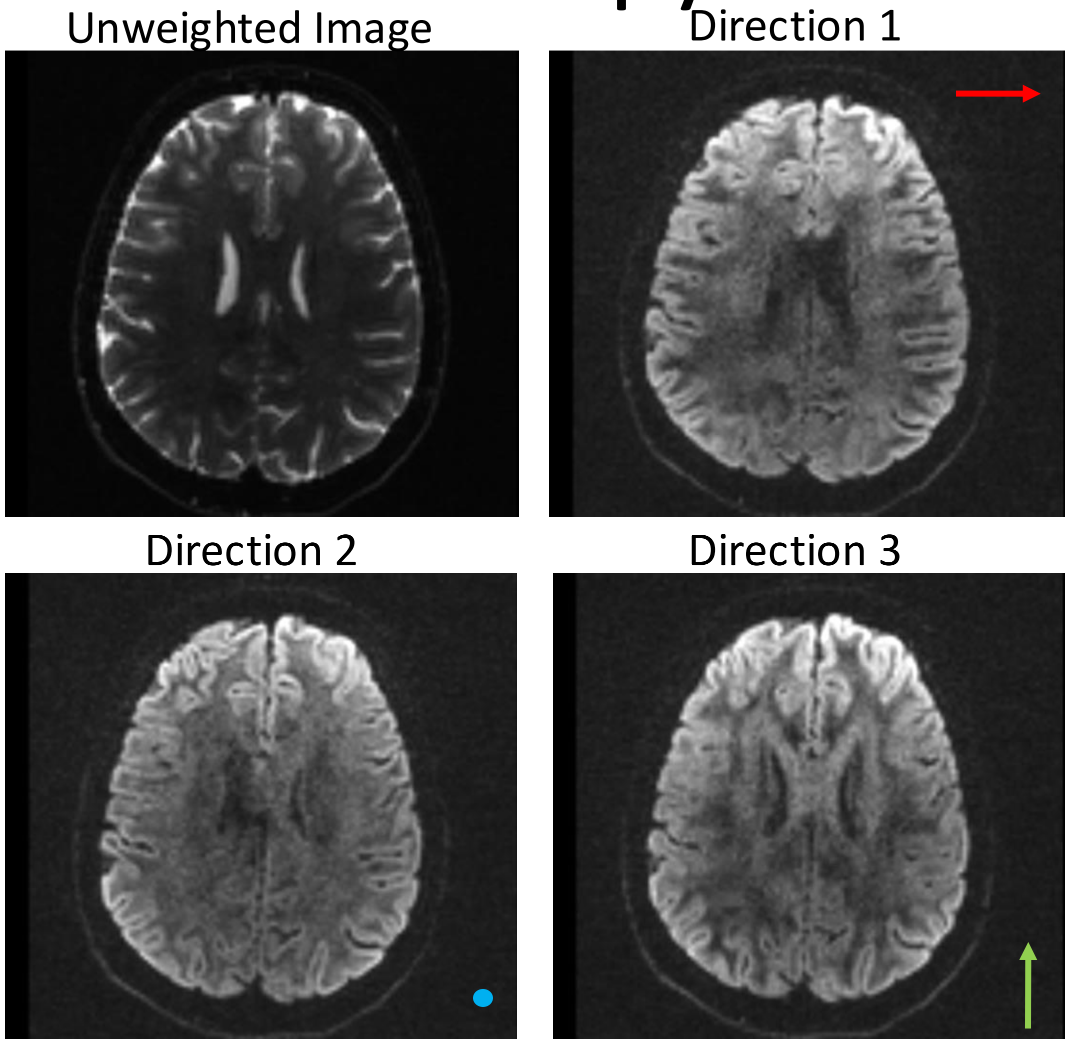


b (s/mm^2)

- Low b-values may be affected by blood flow in capillaries (modelled by IVIM)
- Slice selection gradients can contribute to diffusion weighting (mainly at low b).



Anisotropy and Diffusion Tensor Imaging (DTI)

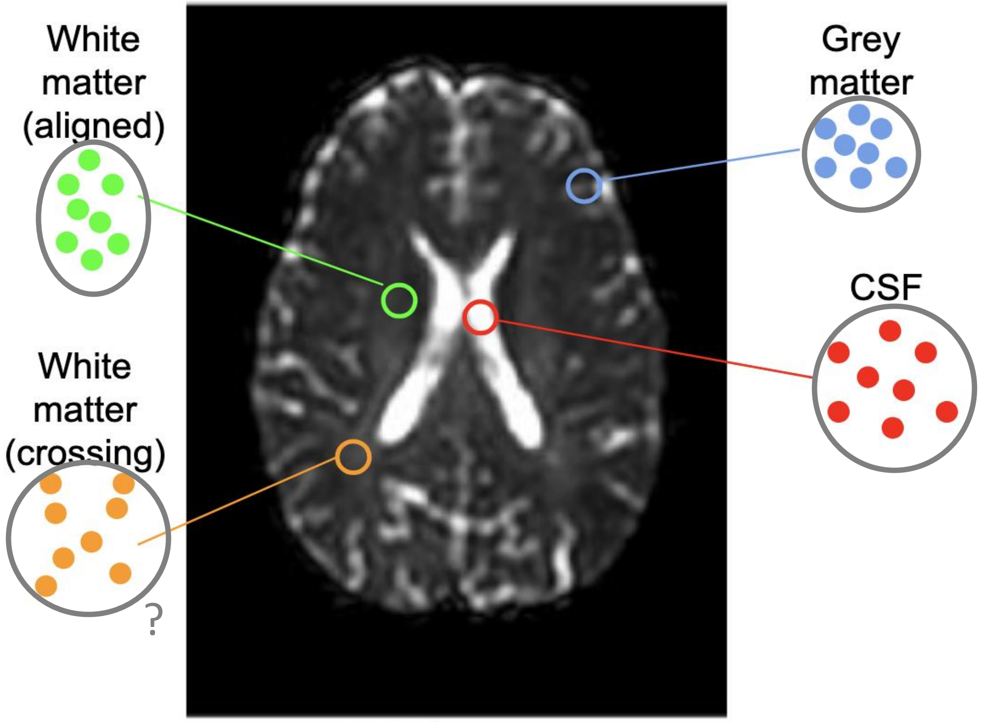


- Scanners typically have 3 gradient coils
- Turn on in various combinations to examine water motion in different directions
- Diffusion signal varies with gradient direction (eg. WM fibre orientation)

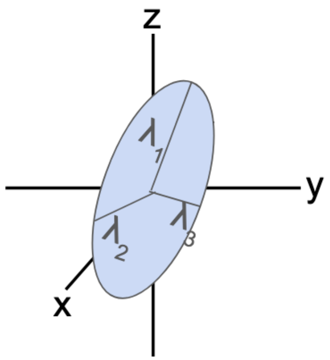
Data: https://dmri.slicer.org/docs/

What is a tensor and what can it tell us?

Example Probability Density Functions



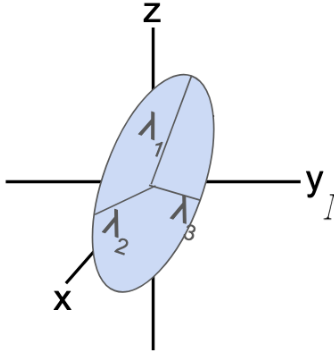
- Can be described by a tensor
- Requires 6 independent directions (plus unweighted image)
- This basic tensor shape doesn't describe all possibilities but works well for aligned axons vs grey matter



$$S(\mathbf{B}) = S_0 e^{-\mathbf{BD}}$$

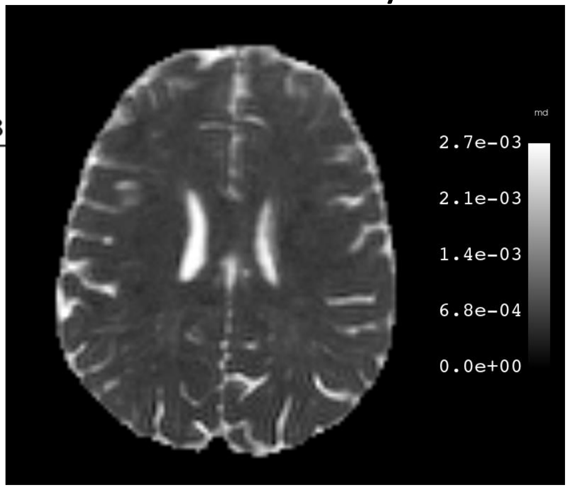
$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}$$

DTI – Common Summary Parameters

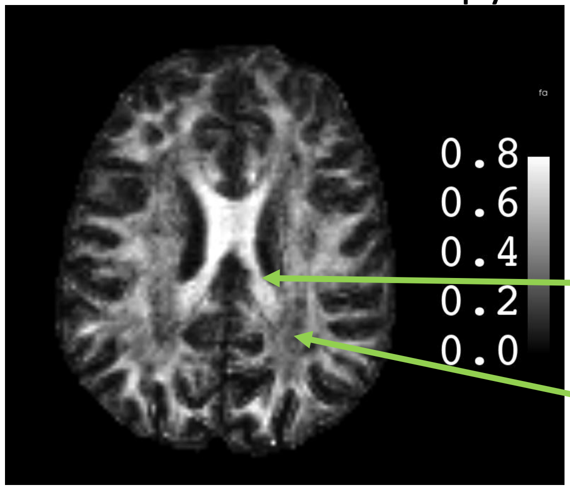


$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$$

Mean diffusivity

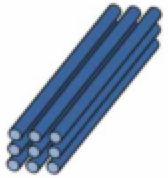


Fractional Anisotoropy

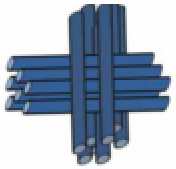


$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

Aligned fibres

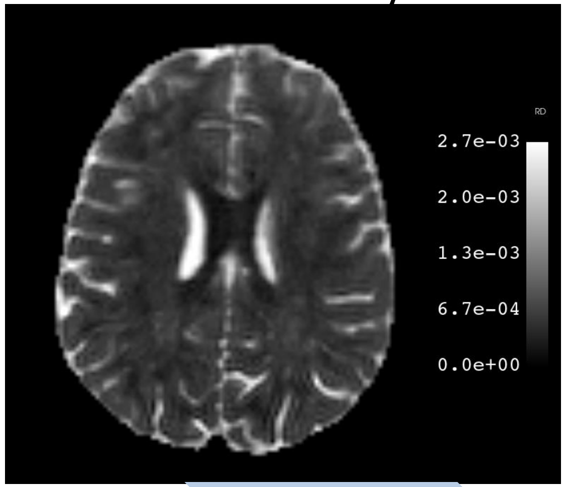


Crossing fibres

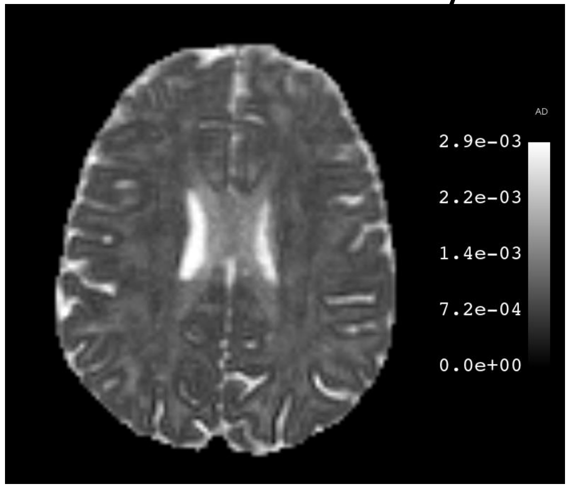


$$RD = \frac{\lambda_2 + \lambda_3}{2}$$

Radial diffusivity



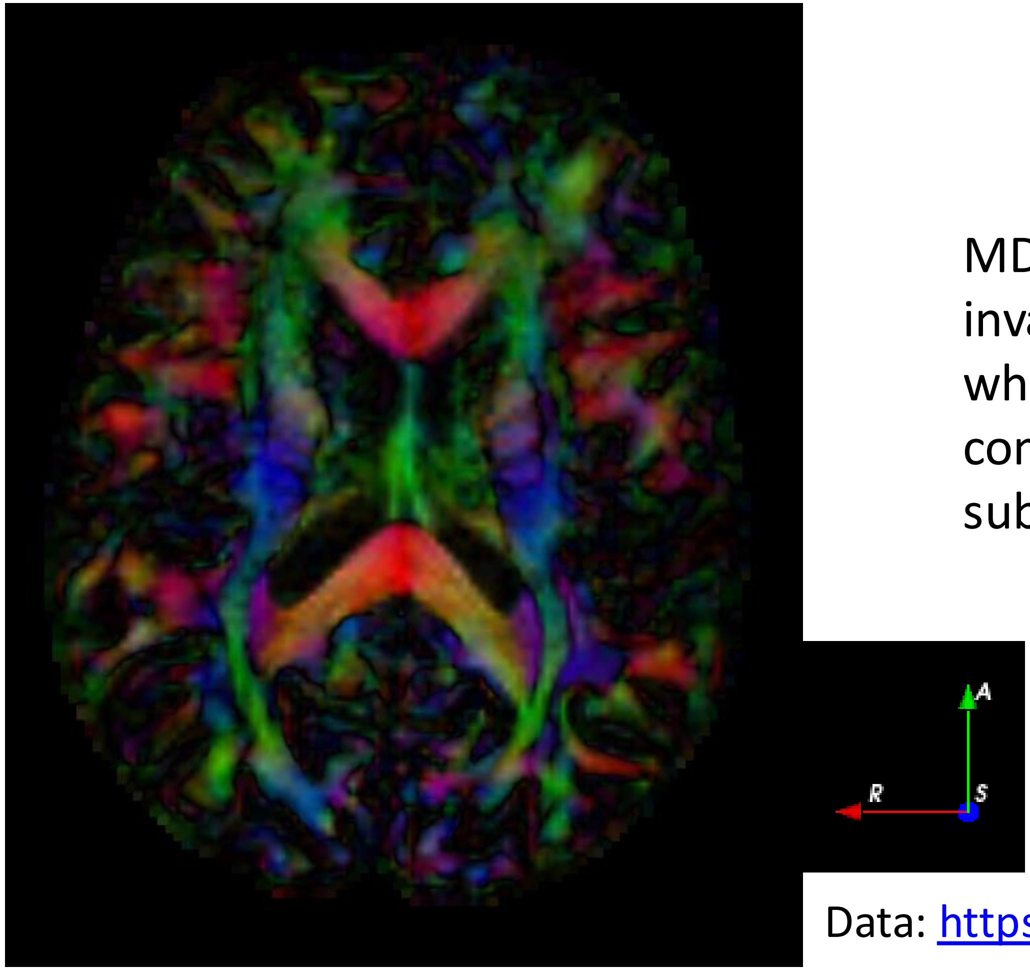
Axial diffusivity



$$AD = \lambda_1$$

Data: https://dmri.slicer.org/docs/

Colour FA: Summarizes FA and Fibre Direction

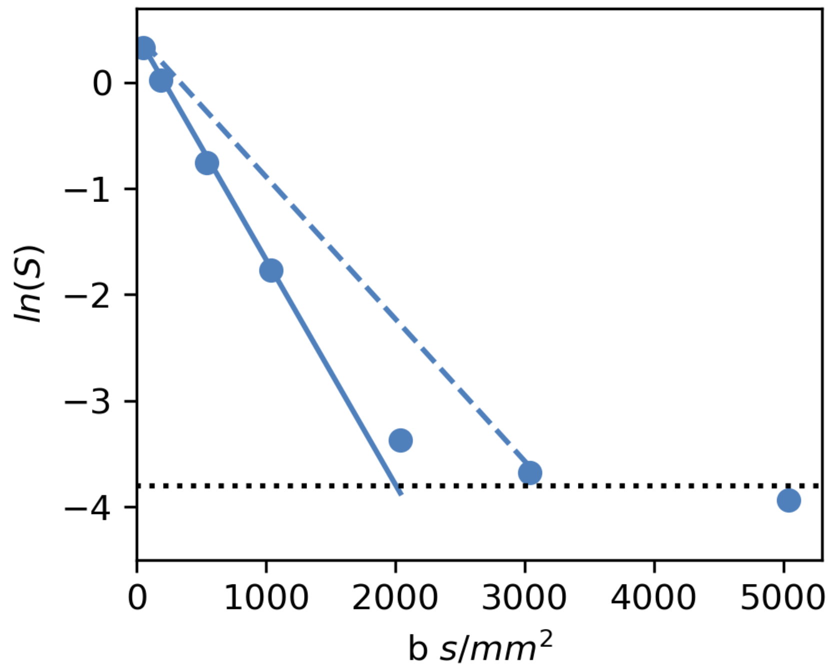


MD and FA are rotationally invariant scalar values, which make for easier comparisons between subjects.

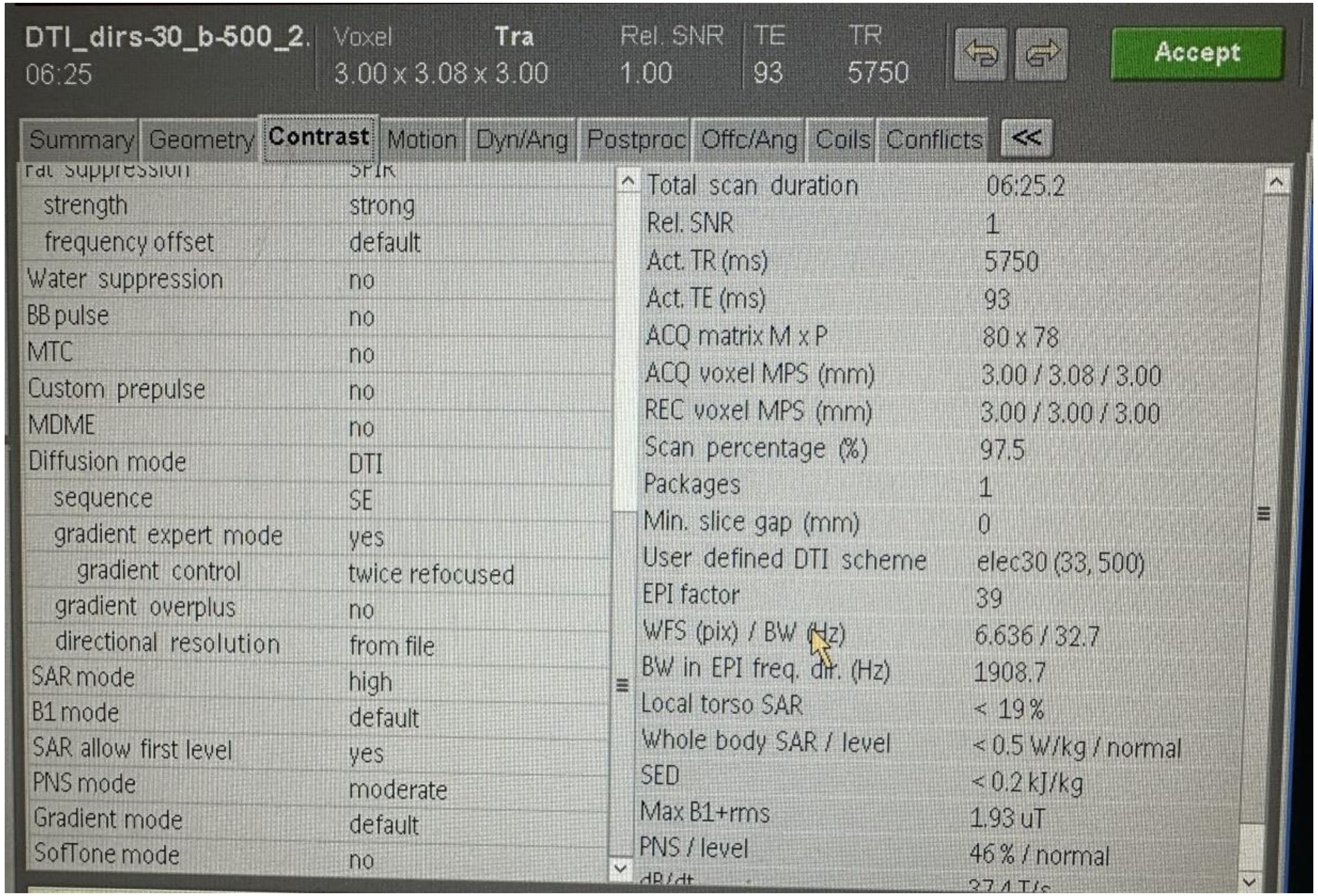
Data: https://dmri.slicer.org/docs/

Limitations of DTI

- FA is particularly sensitive to noise
- Cannot account for multiple fibre orientations within a voxel
    - Possible to explore shapes beyond tensor
    - Requires more measurements and/or assumptions about the distribution
- DTI assumes Gaussian diffusion
    - Usually done at one b-value.
    - Sets of directions at multiple b-values (shells) are possible but time-consuming
    - Quantification options include DKI and NODDI.

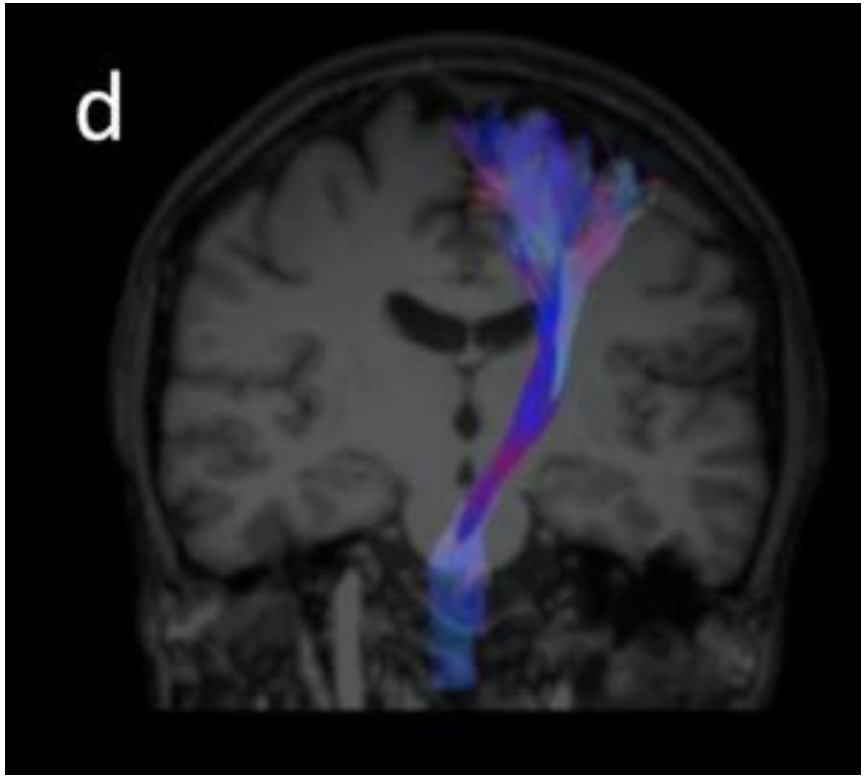


Example Scanner Interface



Tractography for Fibre Bundles

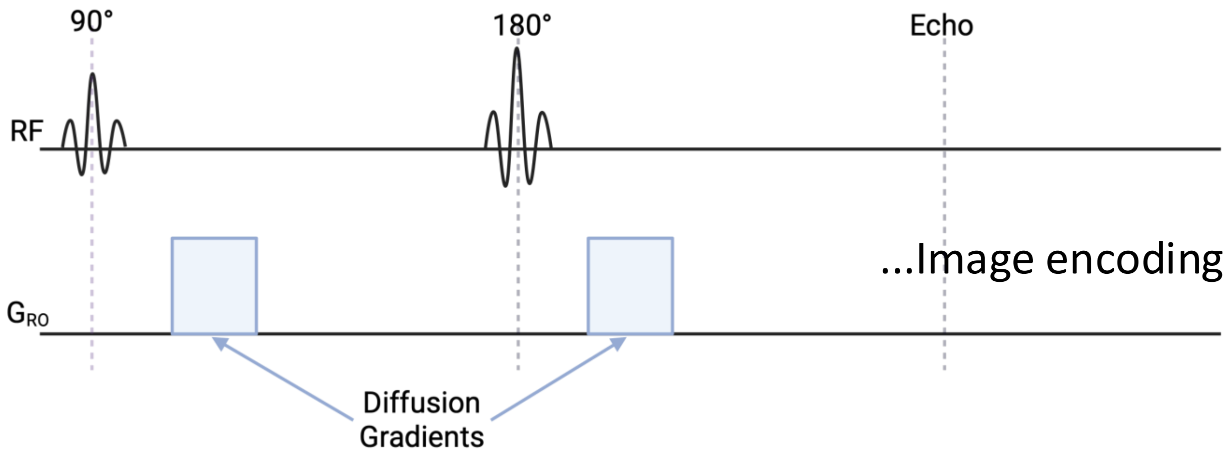
- Tensors in neighbouring voxels can be used to **model** the likely trajectory of white matter pathways
- Prone to false positives. Incorrect tracts need to be edited
- There are a variety of algorithms
    - Select a point in a seed region
    - Follow the likely local direction a small step size
    - Repeat until termination criterion is reached (eg. low FA)
    - These steps can be repeated for multiple points in the seed region to generate many streamlines

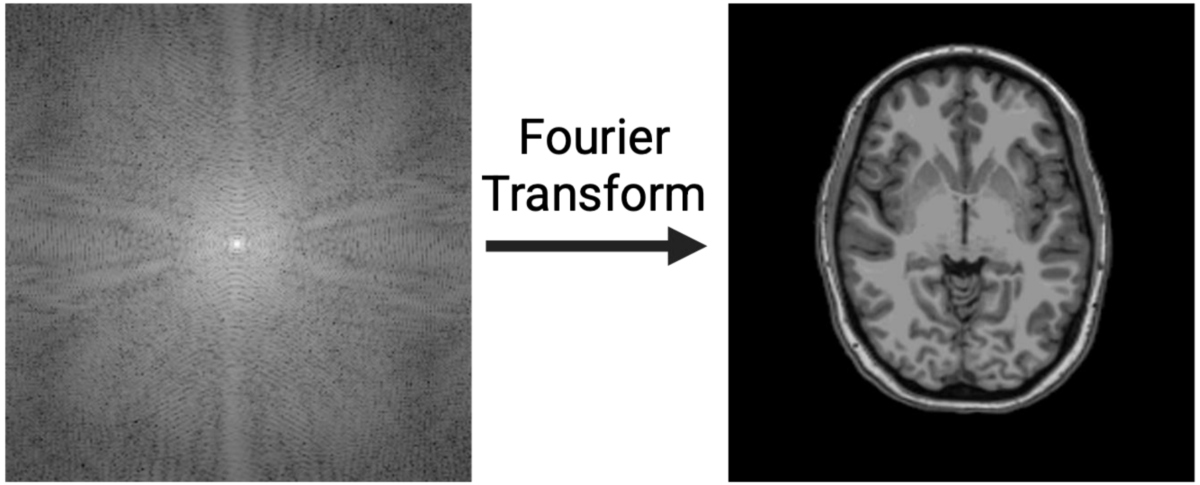


Christidi et al. 2022. Neurol. Int. 14: 841.
https://pmc.ncbi.nlm.nih.gov/articles/PMC9589952/

How do we acquire an image?

How do we acquire an image?



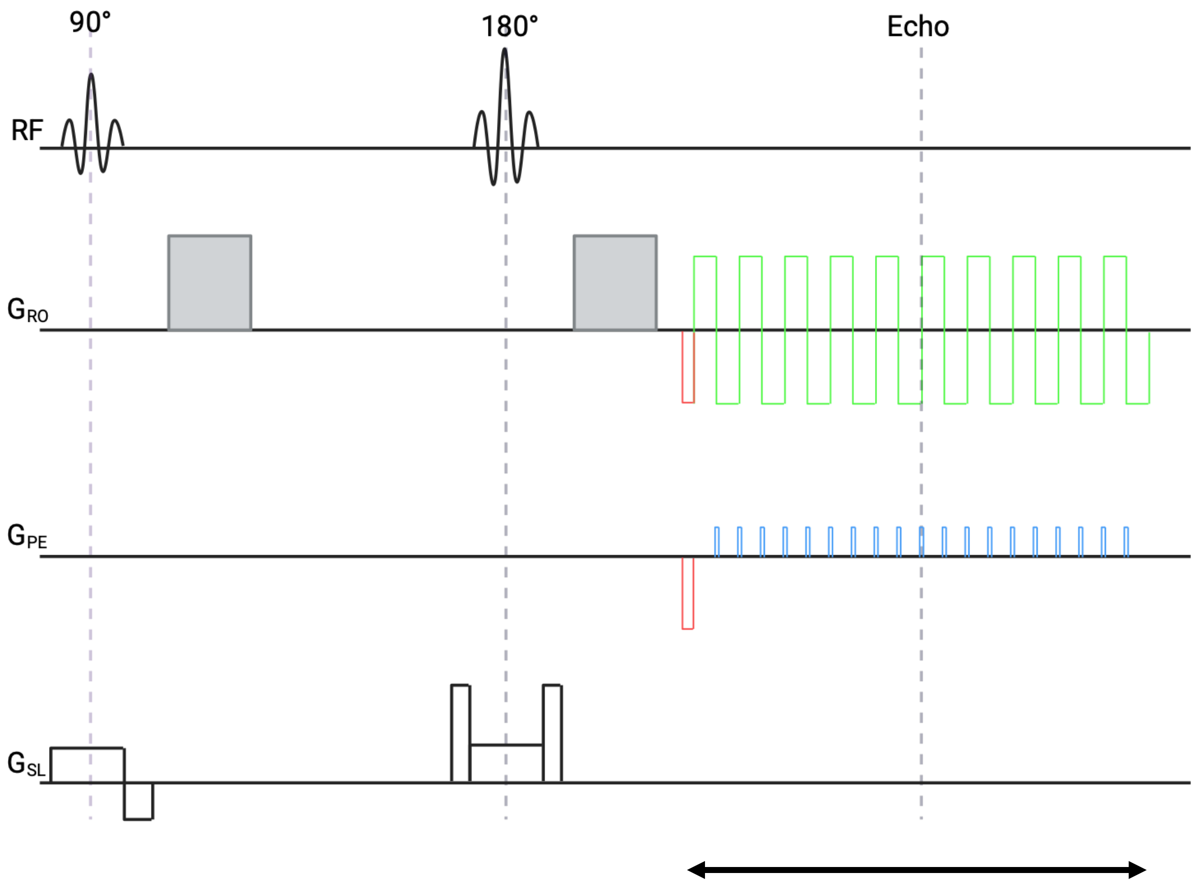


There are many different ways to traverse k-space

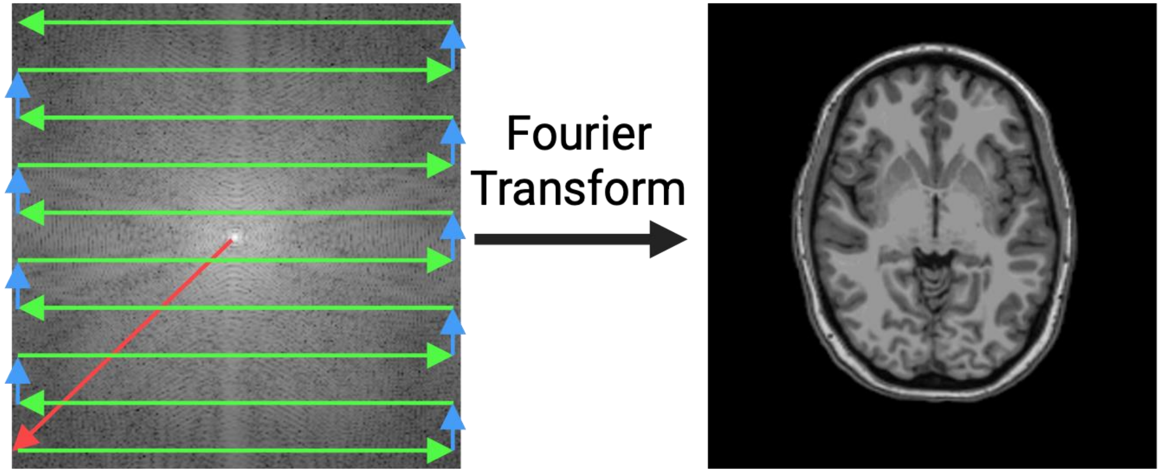
The most common method for clinical diffusion imaging uses echo planar imaging (EPI)

It is fast (for MRI) and less sensitive to motion but prone to certain artefacts

Echo Planar Imaging (EPI)

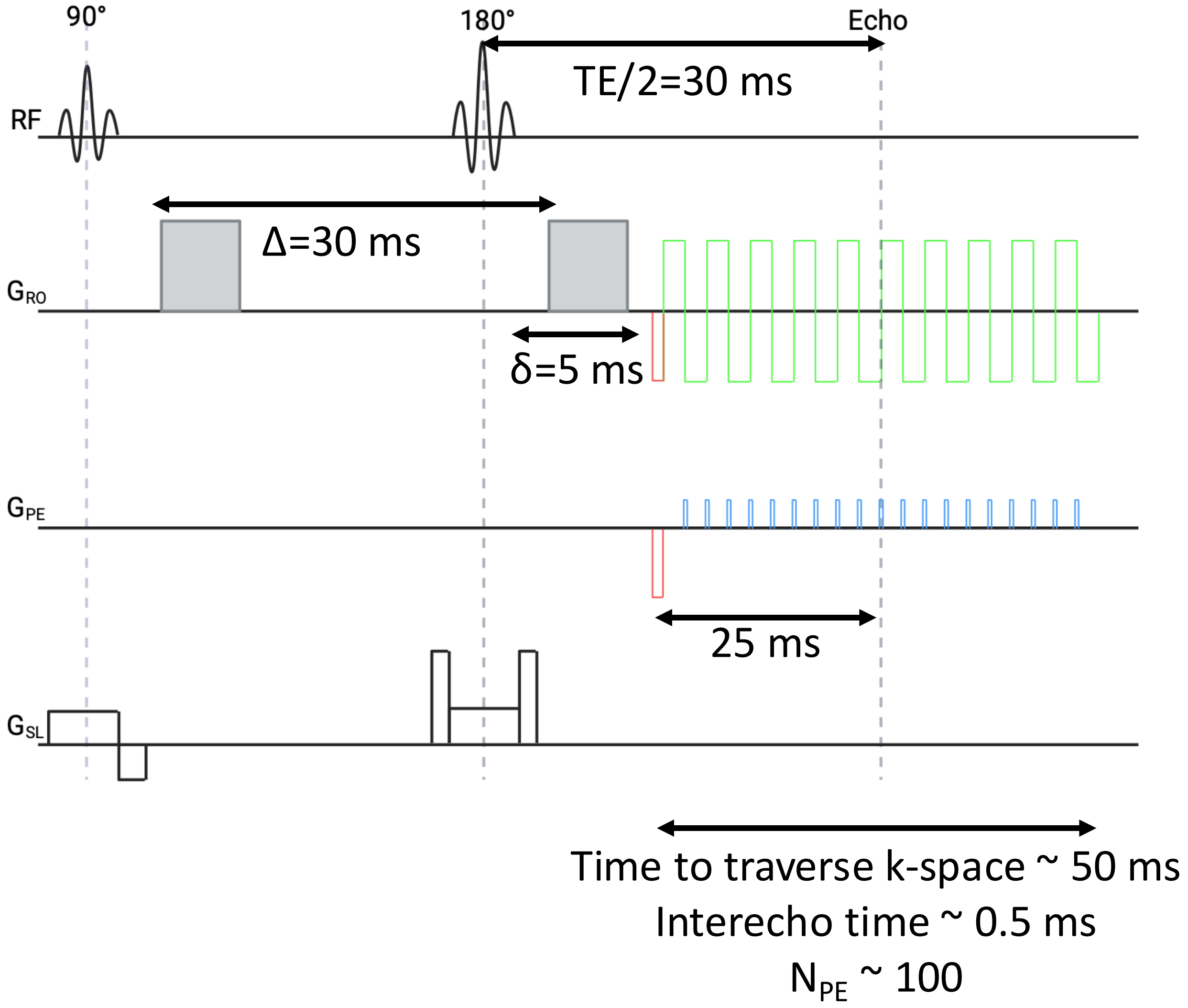


Time to traverse k-space ~ 50 ms
Interecho time ~ 0.5 ms
N_{PE} ~ 100



For a spin echo readout, where every phase encode line is done in a separate TR, the acquisition would take 100x as long!

Tradeoffs: Diffusion Weighting and Echo Times



Suppose:

- Maximum gradient 40 mT/m
- 100 phase encodes (20 cm FOV, 2 mm resolution)
- TE = 60 ms

What is the maximum b-value?

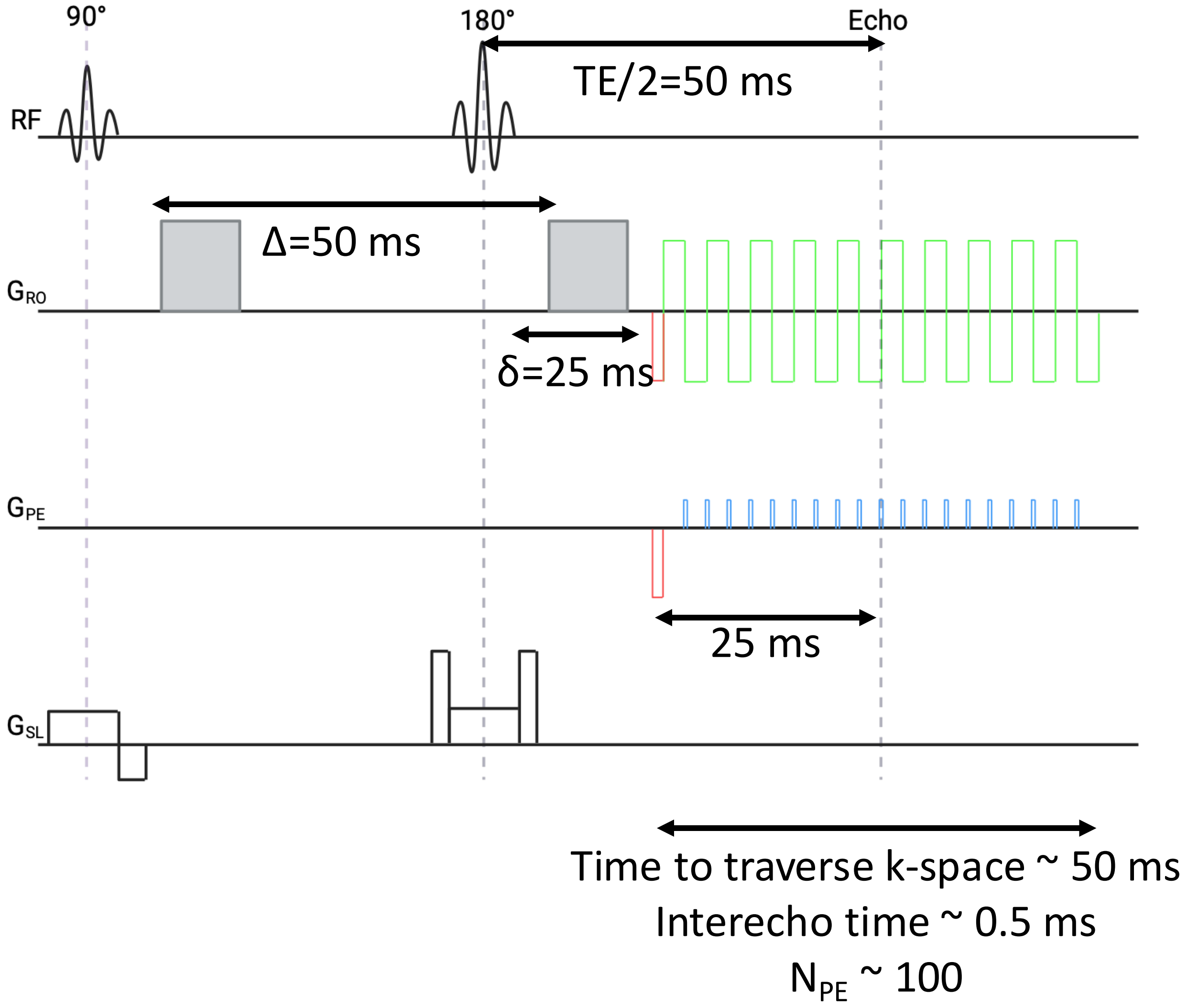
$b=(\gamma G\delta)^2(\Delta-\delta/3)$

$b=(2.68e^8*0.04*0.005)^2*(0.03-0.005/3)$

$b=8.11e^7 \text{ s/m}^2$

$b=81 \text{ s/mm}^2$

Exercise: Help me get more diffusion weighting



Option 1: Increase TE to 100 ms

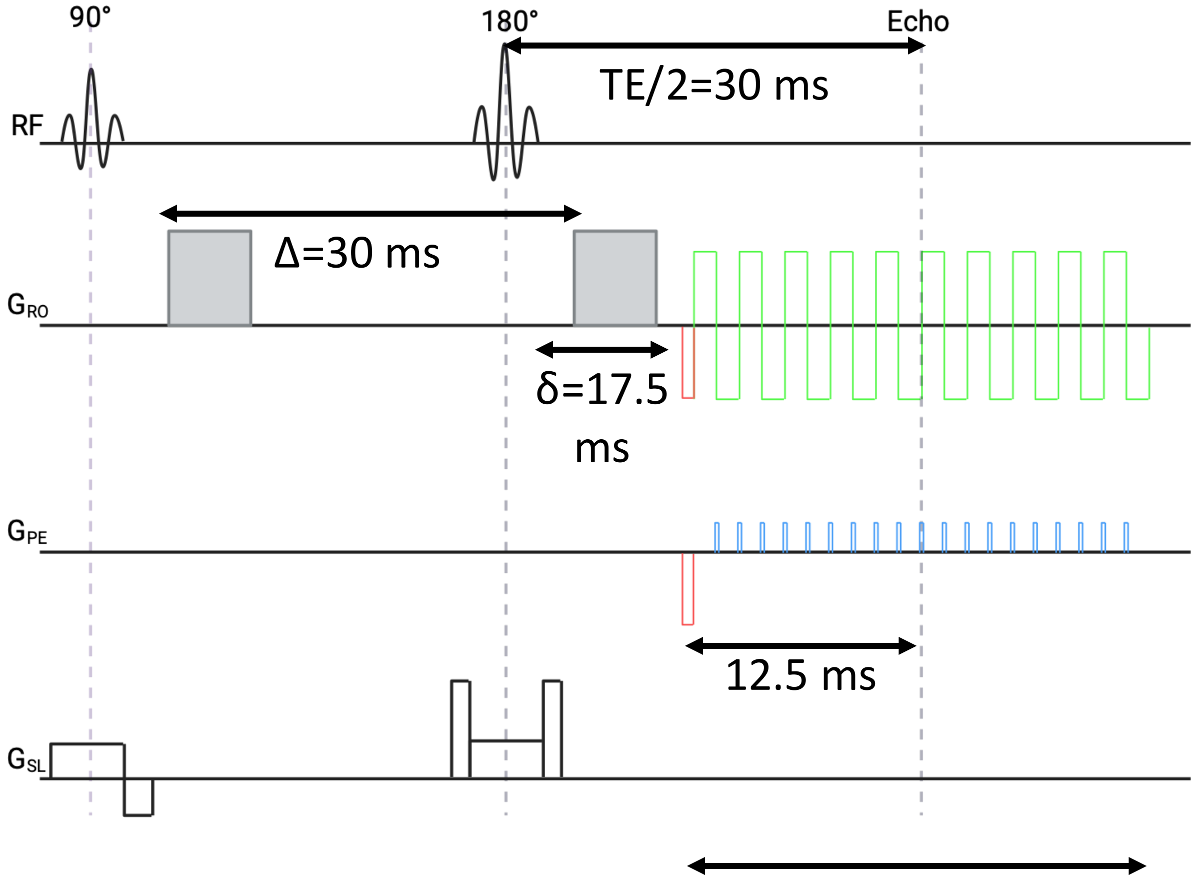
$b=(\gamma G\delta)^2(\Delta-\delta/3)$

$b=2982\ s/mm^2$

Why would we not want to do this?

T2 decay and lower SNR

Exercise: Help me get more diffusion weighting



Time to traverse k-space ~ 25 ms
Interecho time ~ 0.5 ms
N_{PE} ~ 50

Option 2: Lower resolution 4 mm

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

$b=847\ \mathrm{s/mm^2}$

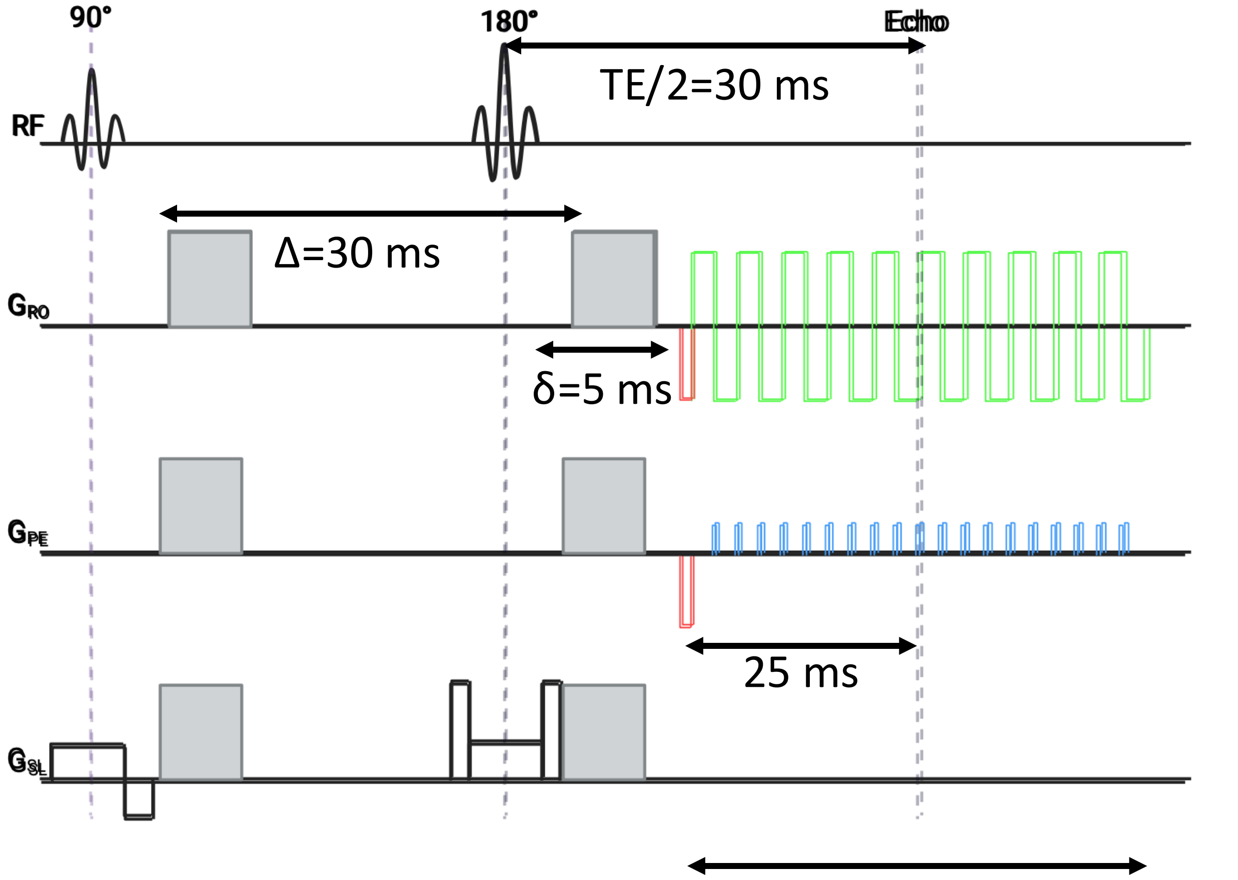
What are the advantages and disadvantages of this?

Lower spatial resolution

4x as much SNR

Would need to scan 16x a long for equivalent SNR gain

Exercise: Help me get more diffusion weighting



Time to traverse k-space ~ 50 ms
Interecho time ~ 0.5 ms
N_{PE} ~ 100

Option 3: Turn on all 3 gradients

G_{all}=sqrt(3)*0.04

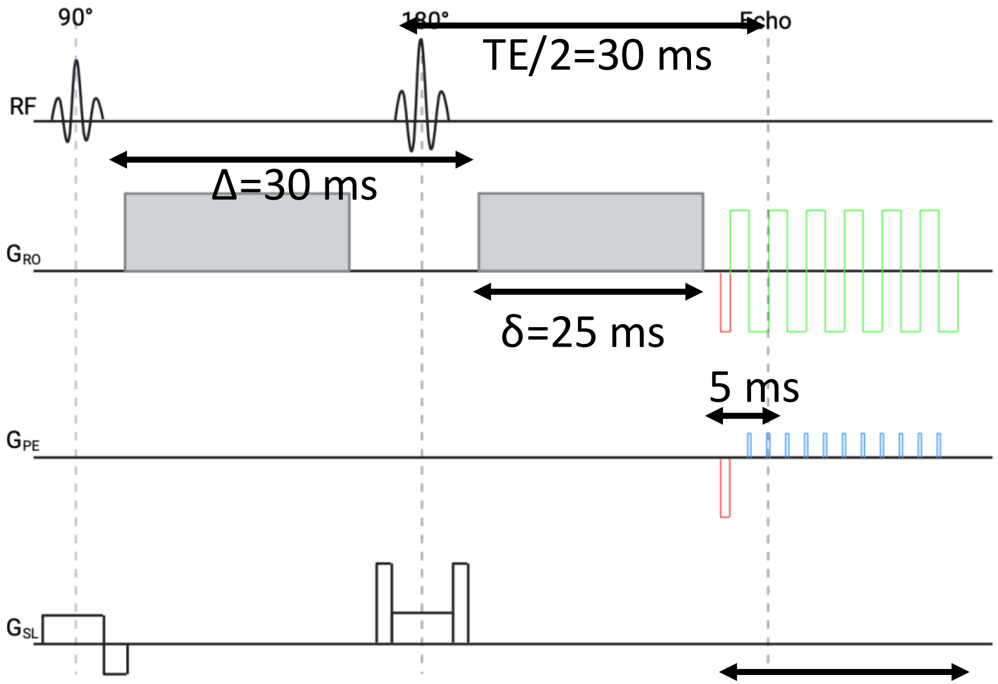
$b=(\gamma G\delta)^2(\Delta-\delta/3)$

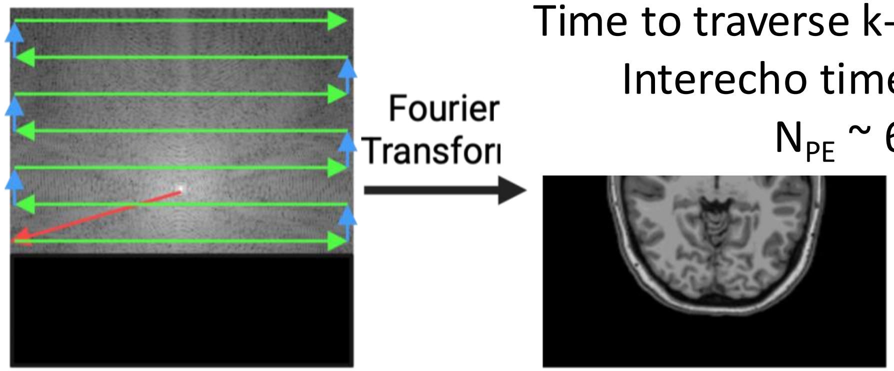
$b=243\ s/mm^2$

Why would we not want to do this?

This only works in 4 directions

Exercise: Help me get more diffusion weighting





Time to traverse k-space ~ 30 ms

Interecho time ~ 0.5 ms

N_{PE} ~ 60

Option 4: Partial Fourier Encoding of 0.6

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

$b=1550\ s/mm^2$

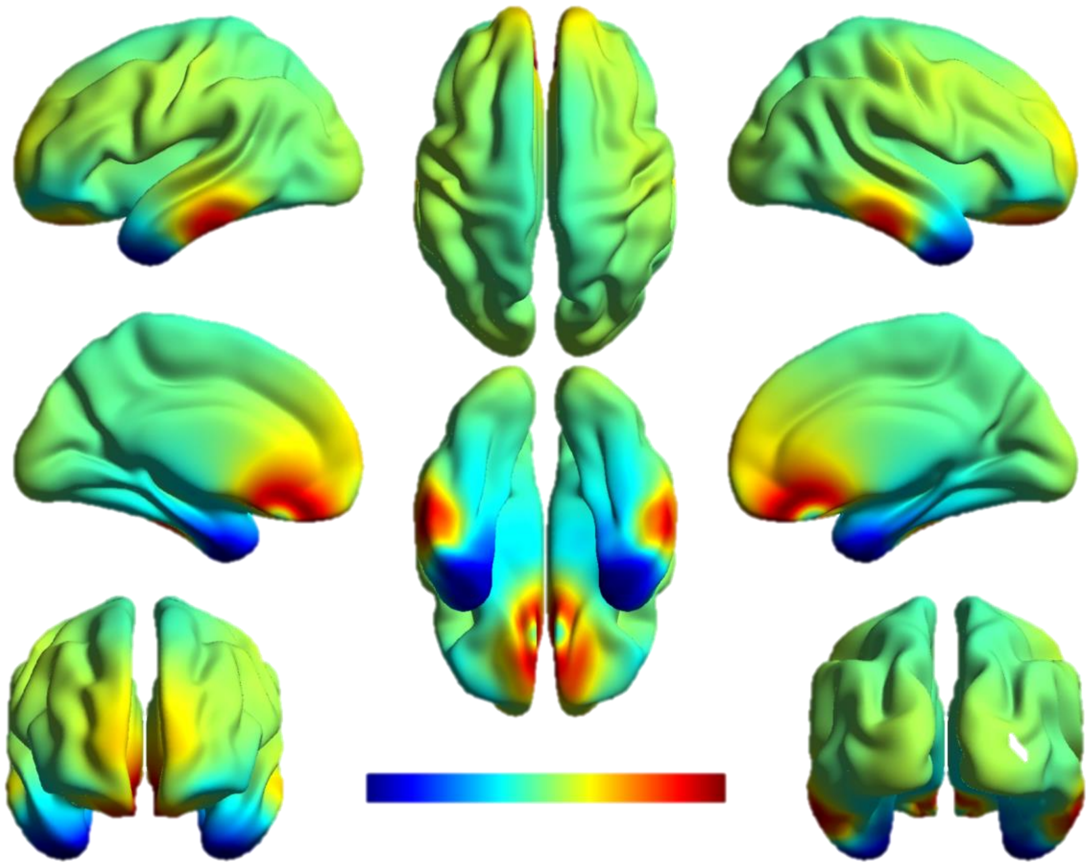
Parallel imaging is another option to reduce the number of phase encodes:

- Multiple receive coils
- SENSE or GRAPPA

A note about slices

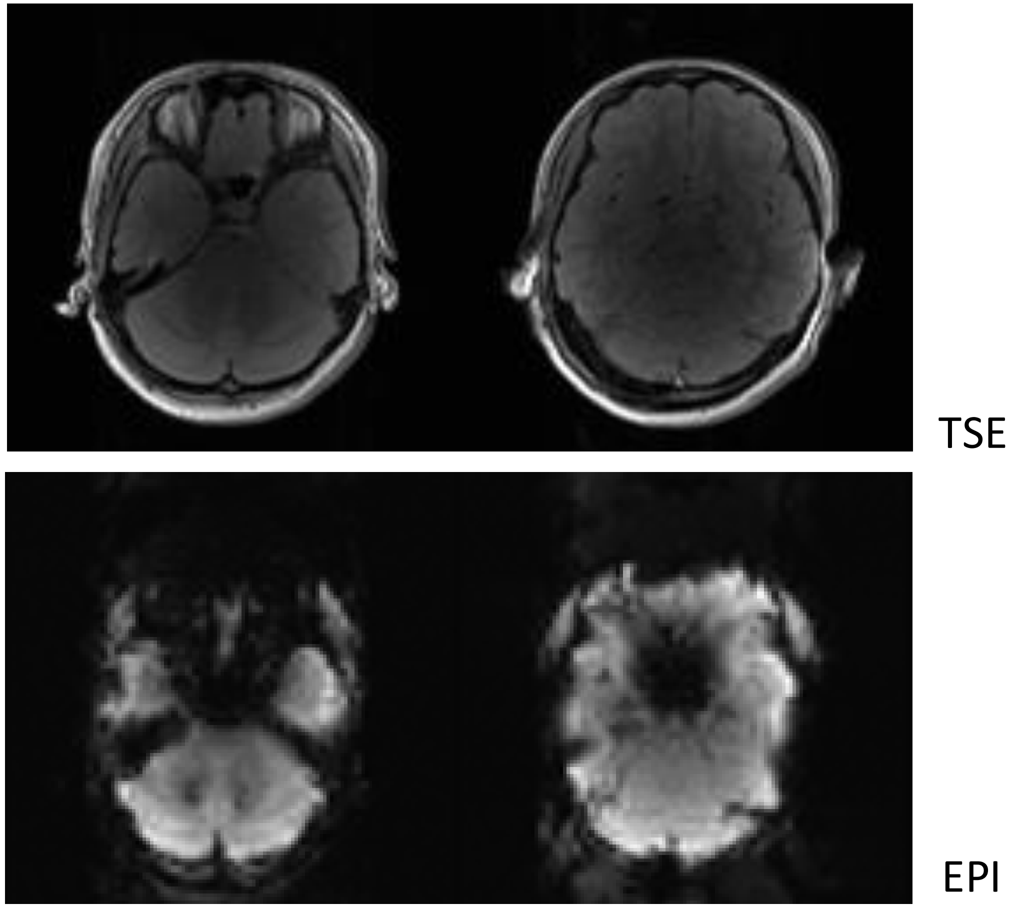
- Examples above are for one slice
- EPI data is often acquired "multi-slice" not "3D"
  - Slice selection + 2D k-space
  - Repeat the slice-selective excitation for different slices during TR
  - More slices require longer TR, so there is some increase in imaging time
- It is possible to acquire 3D k-space and do a 3D Fourier transform
  - Requires long echo times unless there are few phase encodes
  - Has advantages for resolution

EPI Artefacts: Susceptibility-based distortion



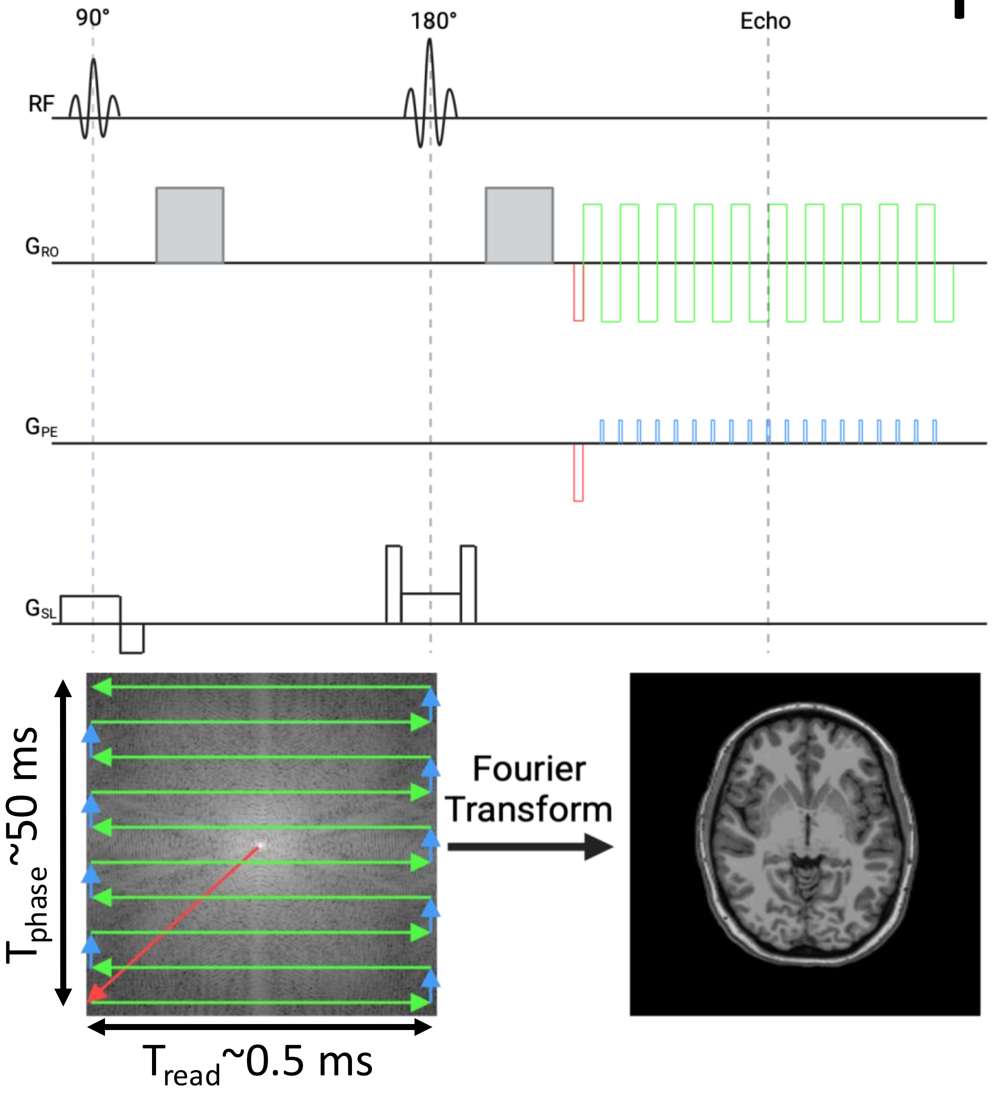
Off-resonance distribution at 3T

Images: Fa-Hsuan Lin



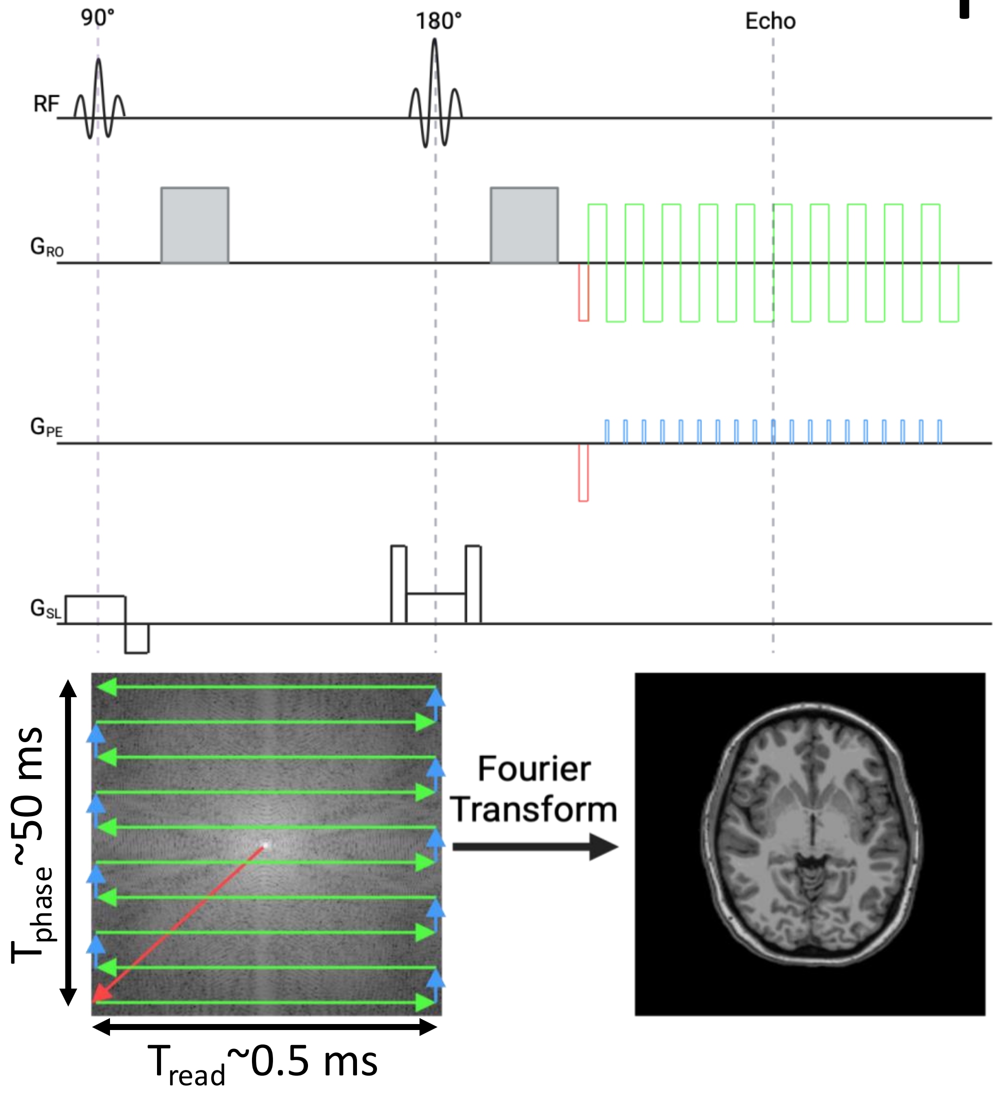
Diffusion Basics | Sequences I: Diffusion Encoding | b-value and ADC | Anisotropy and DTI | Sequences II: Image Encoding | Sequences II: Tradeoffs | Sequences II: Artefacts | Processing Pipeline

EPI Artefacts: Susceptibility-based distortion



- The bandwidth per pixel in the readout direction is:
  1/0.0005 = 2000 Hz/px
- In the phase direction, bandwidth per pixel is:
  1/0.05 = 20 Hz/px
- Susceptibility that varies the B0 field by 100 Hz creates a shift:
  - 100/2000 = 0.05 pixels in readout
  - 100/20 = 5 pixels in phase
- Appears as stretch or compression in phase encode direction

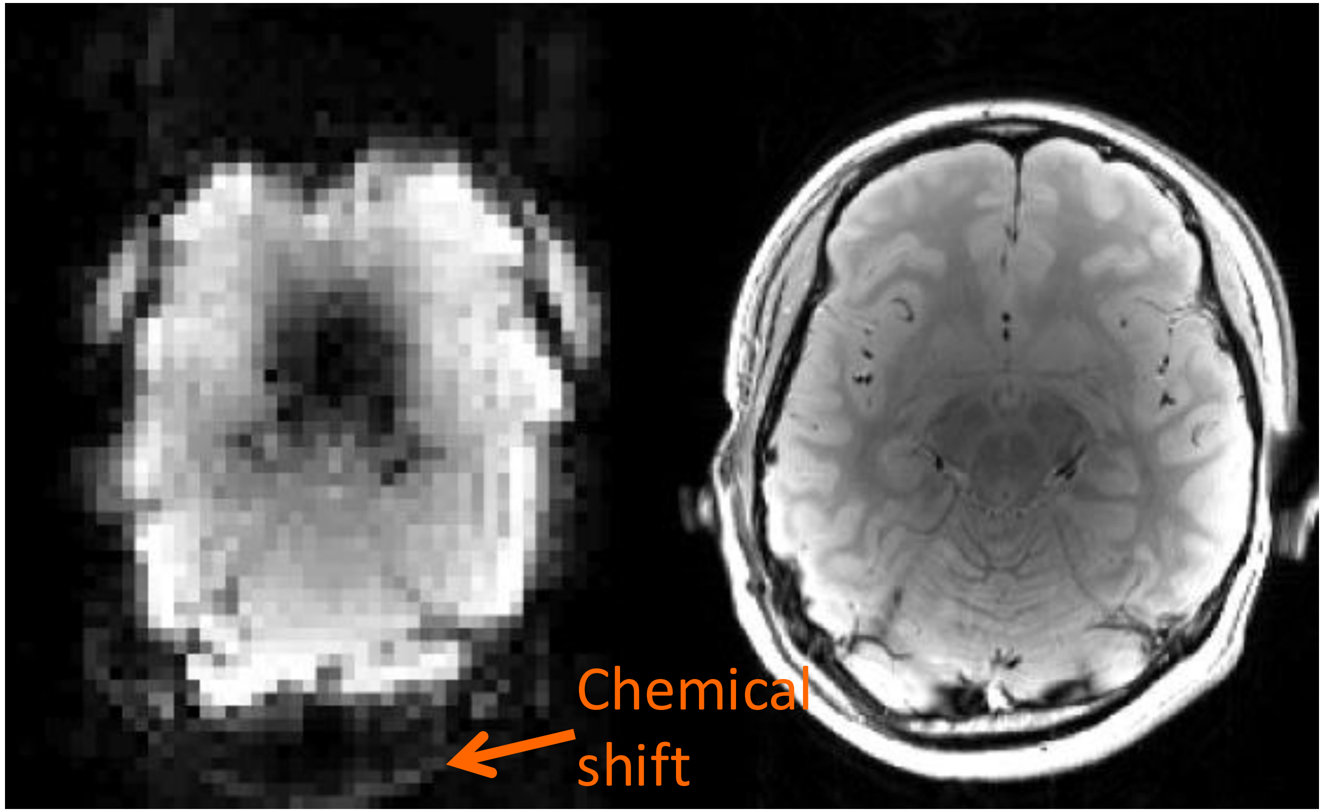
EPI Artefacts: Susceptibility-based distortion



Fixes

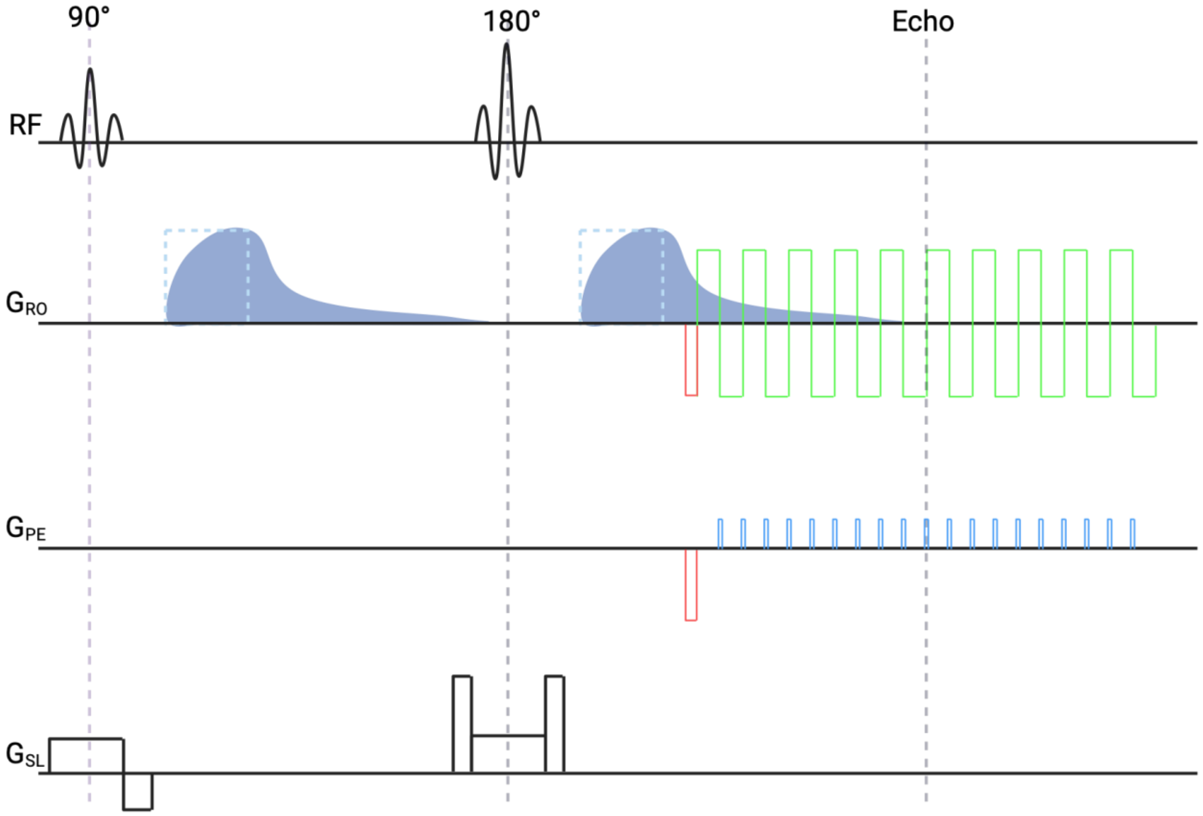
- Reduce susceptibility
  - Shimming
  - Thinner slices
- Correct for susceptibility
  - Acquire a B0 map and correct in post-processing or use to shim
  - Acquire an image with "reverse blip"
- Reduce # phase encodes
  - Partial Fourier
  - Parallel imaging
  - Choose the "right" phase encode direction

Chemical Shift Artefact



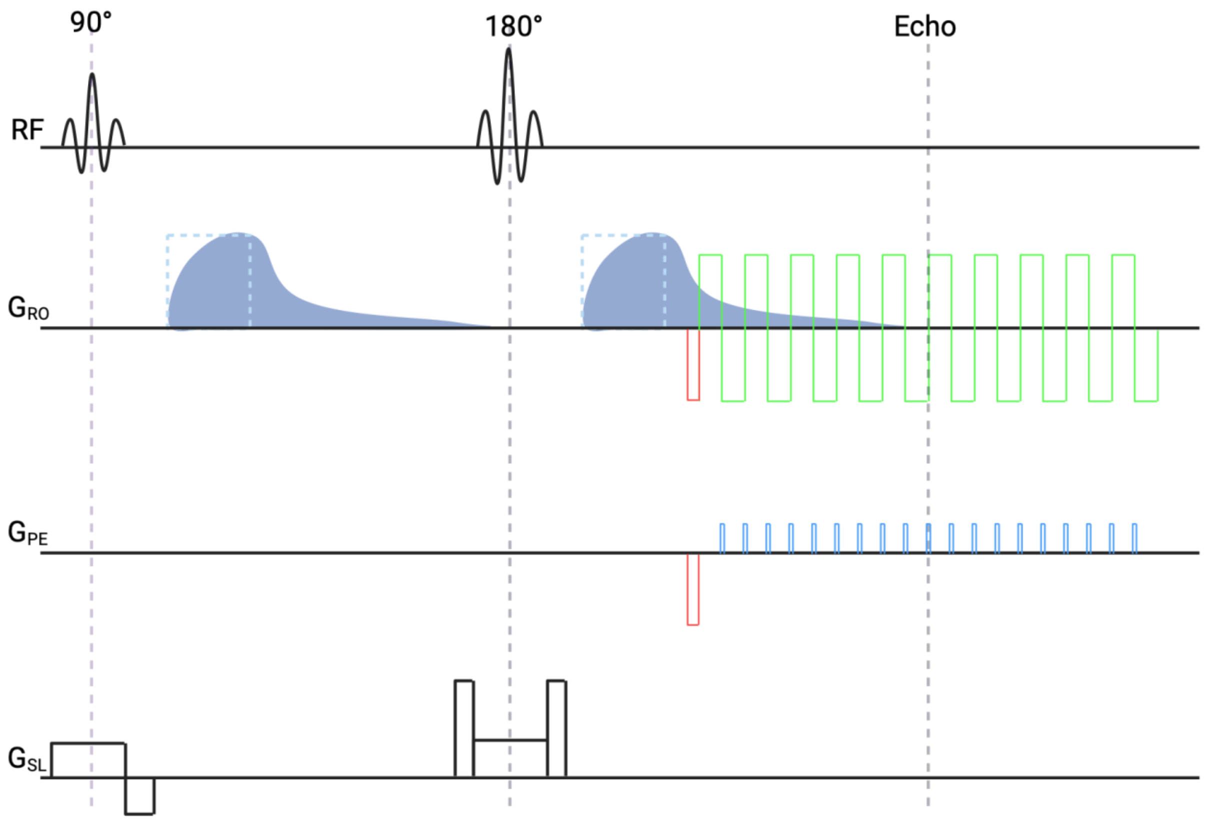
- Fat is shifted 3.5 ppm from water, about 450 Hz at 3 T
- For readout bandwidth of 2000 Hz/px, this is 0.2 pixels
- For phase bandwidth of 20 Hz/px, this is 22 pixels
- Fat signal attenuates less than water, so artefact may be more obvious in DWI than unweighted image.
- Use scanner fat saturation

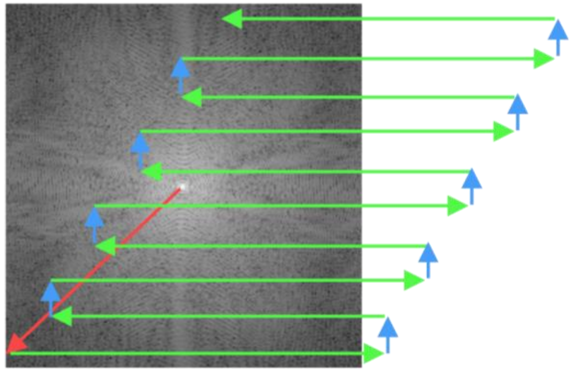
Eddy current distortions

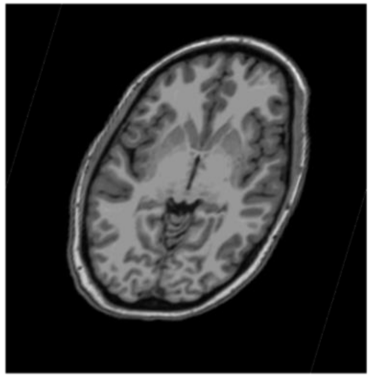


- Turning on and off gradients induces eddy currents, changes effective gradient
- Two problems:
  1. Diffusion gradients aren't what we plan (altered b-value)
  2. EPI readout will not be as expected, creating artefacts
- This effect will be worse for stronger diffusion gradients

Eddy Current Distortions - Readout

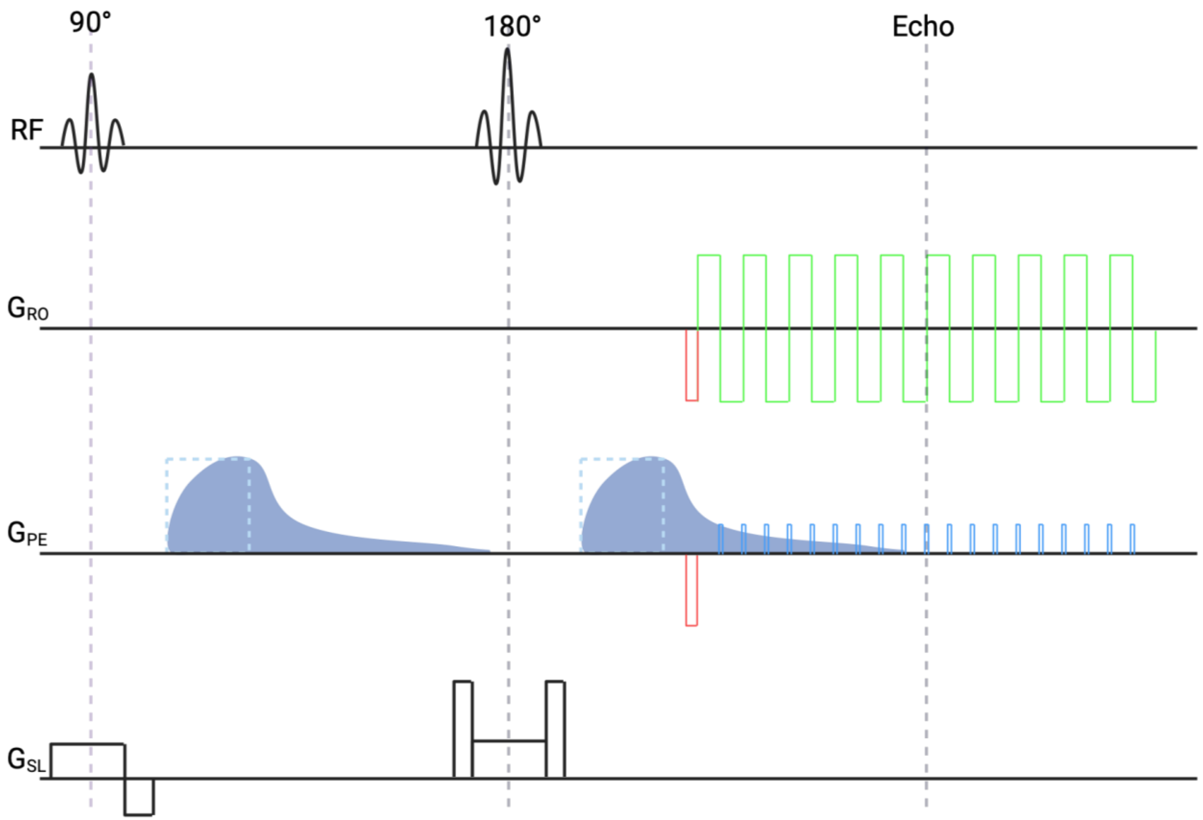


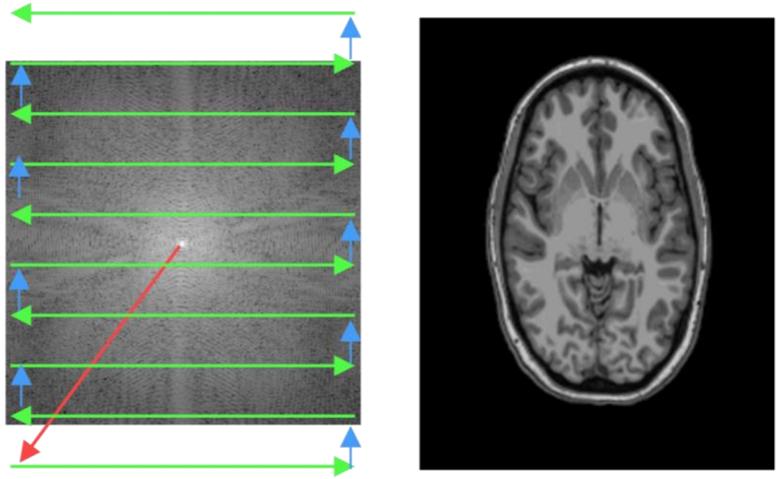




Shear

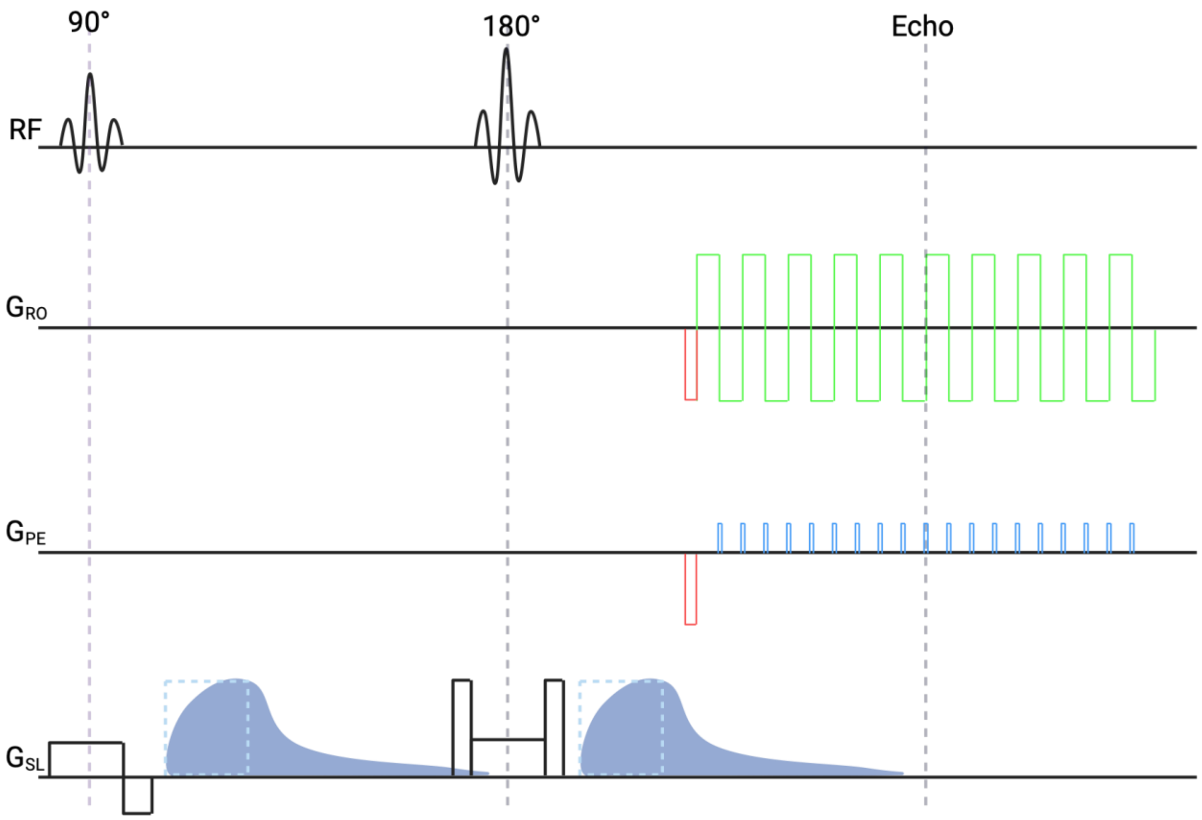
Eddy Current Distortions - Phase





Stretch

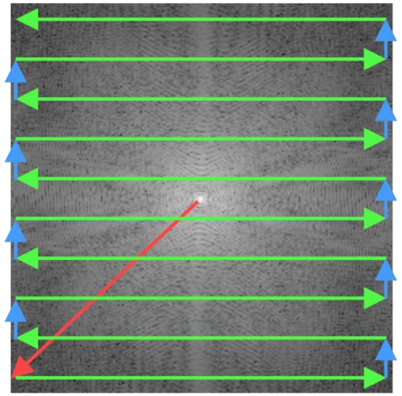
Eddy Current Distortions - Slice

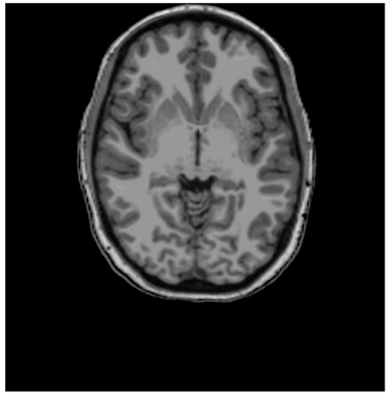


These are all approximately affine transformations.

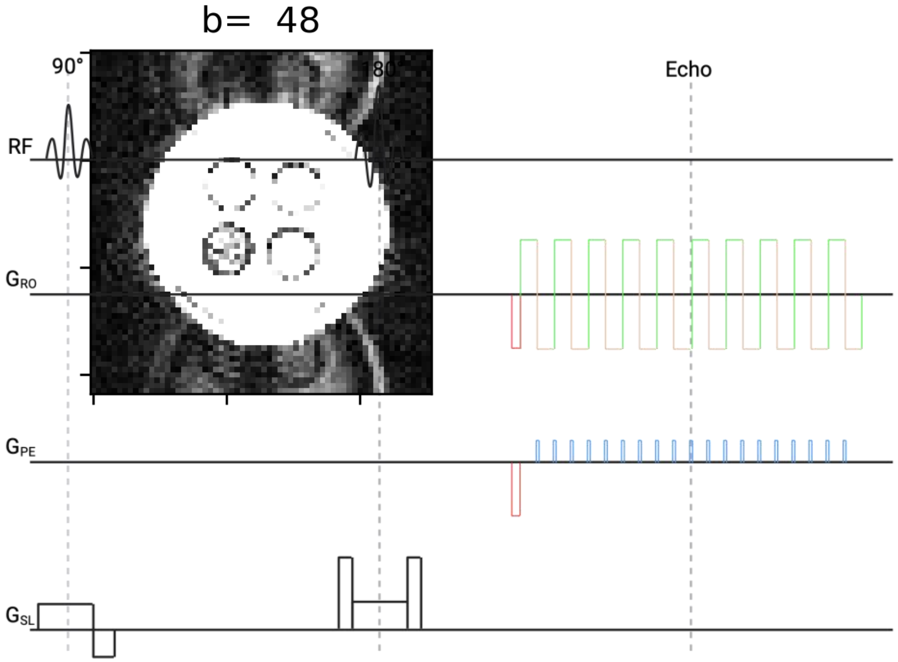
Can be corrected through affine registration.

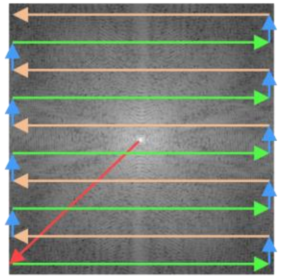
Translation

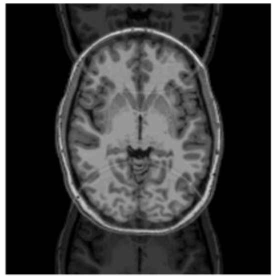




EPI ghosts



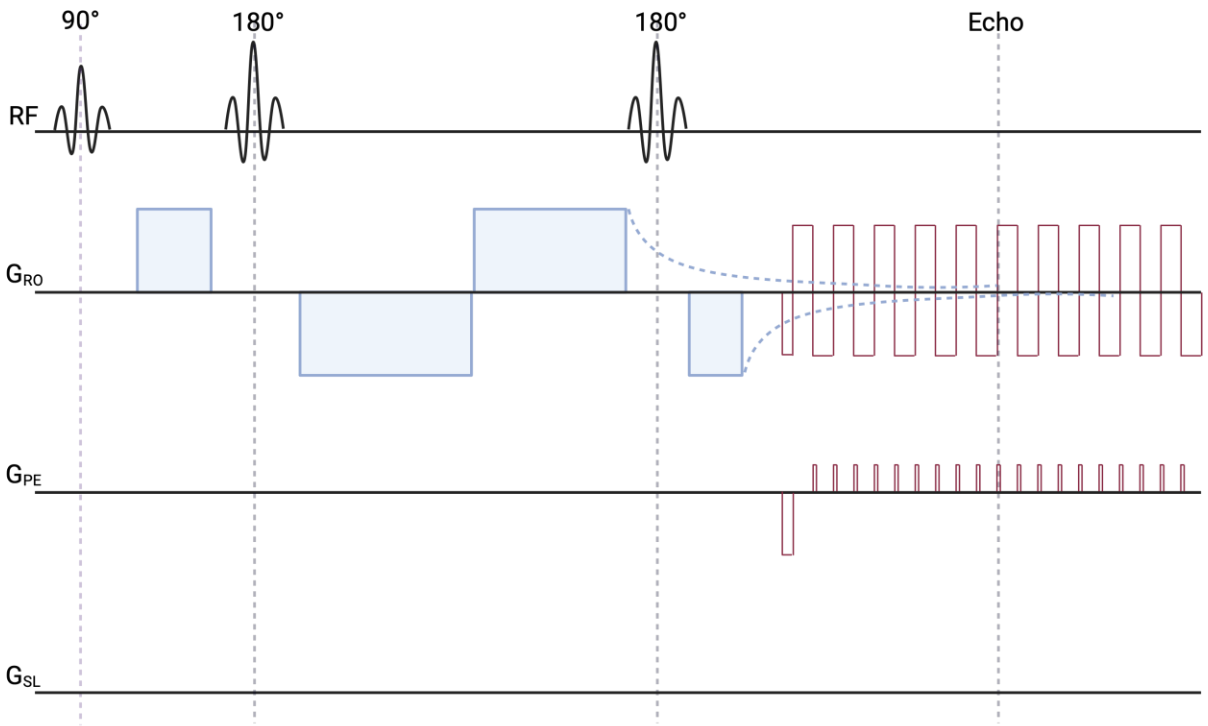




- k-space acquisition alternates direction for even and odd lines
- Small timing differences produce phase differences
- These appear as ghosts in image space
- Always in the phase direction, a distance of N/2 apart
- Many scanners have automatic correction based on pre-scan or navigators

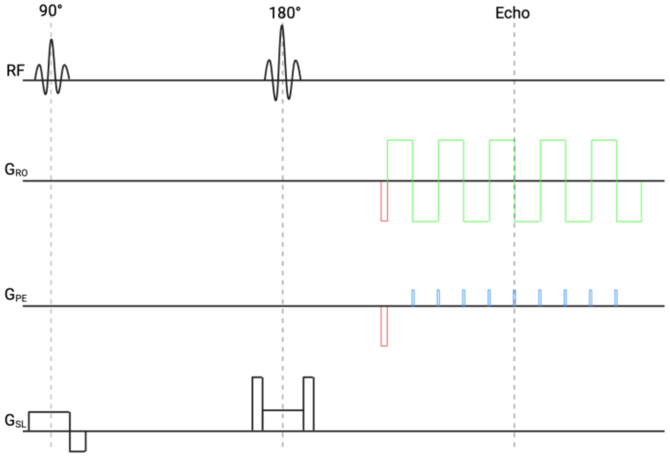
Potential Eddy Current Solutions

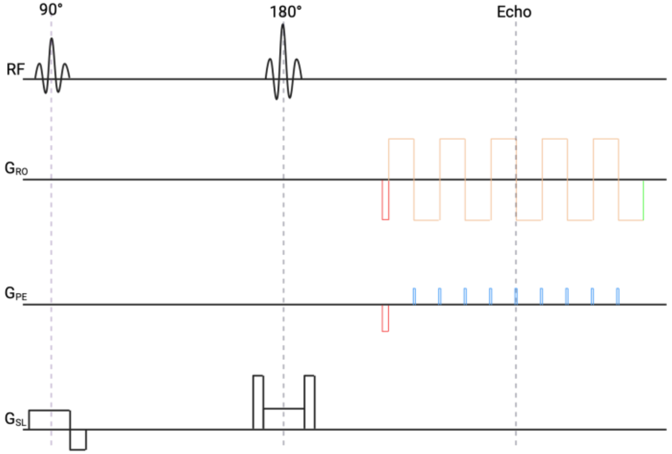
- Worse at high gradients
- Use a "self-compensating" sequence like TRSE
    - Longer TE for a given b-value
- Correction acquired during pre-scan
- Distortions can often be fixed with affine registration
    - Can be challenging as contrast changes

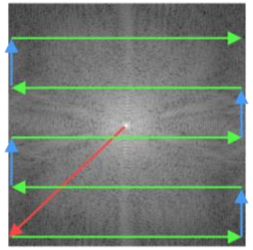


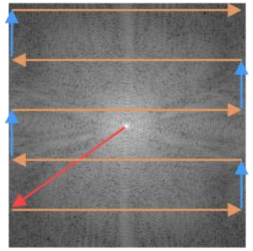
Multi-Shot Imaging and Motion

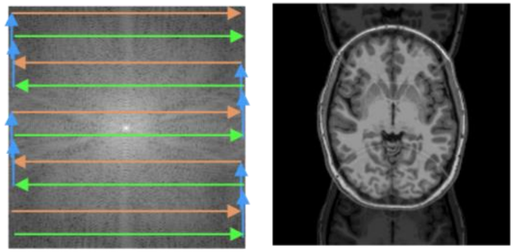
- EPI phase encodes can be split into separate TRs (multi-shot EPI)
- Shots may be seconds apart vs 50 ms for single shot
- Motion between shots causes phase errors
  - Brain pulsations from heart
  - Cardiac gating adds scan time
  - Acquire a navigator for phase correction











Multi-Shot Imaging and Motion

- Multi-shot EPI with interleaved phases may have gaps in k-space after correction
- Could leave residual ghosting
- Readout-segmented EPI is an alternative (eg. RESOLVE)
- Overlap segments so there are no gaps after correction

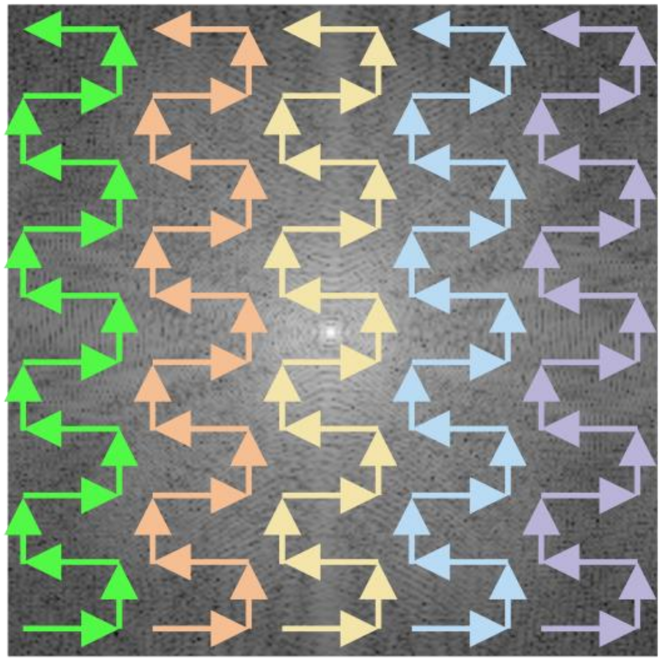


Image Processing Steps and Analysis Tools

- Look at the data (SNR, artefacts)
- Susceptibility correction with B0 map or reverse blip
- Motion correction
  - b-vector correction
- Eddy current correction
- Segment brain
- Apply model (eg. DTI)
- Calculate metrics (MD, FA)
- Compare subjects

- Many scanners automatically do eddy current correction, calculate ADC and FA
- Software often has its own tutorial
  - FSL (Oxford)
  - DTIStudio (Johns Hopkins)
  - Camino (UCL)
  - 3DSlicer with SlicerDMRI extension (NIH), Tutorial.