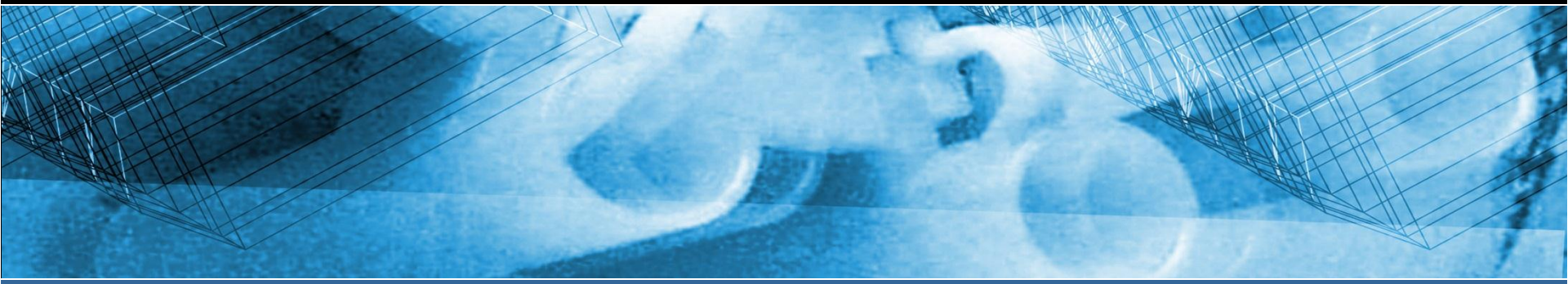


Reconstruction Algorithms for MRI

Berkin Bilgic

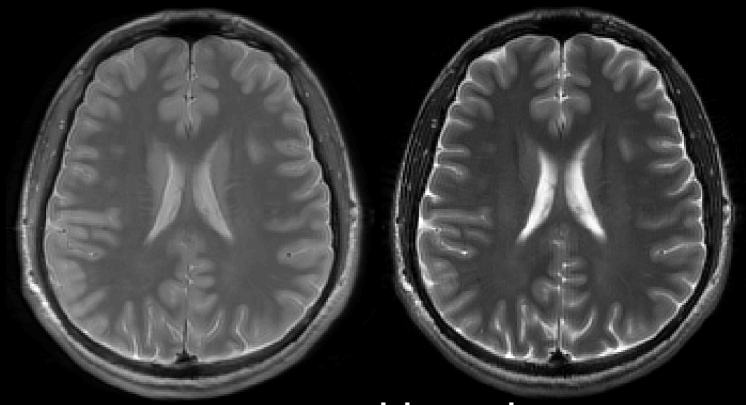
17 December 2012

Outline

- **Magnetic Resonance Imaging (MRI)**

Outline

- **Magnetic Resonance Imaging (MRI)**
  - Non-invasive imaging, great versatility



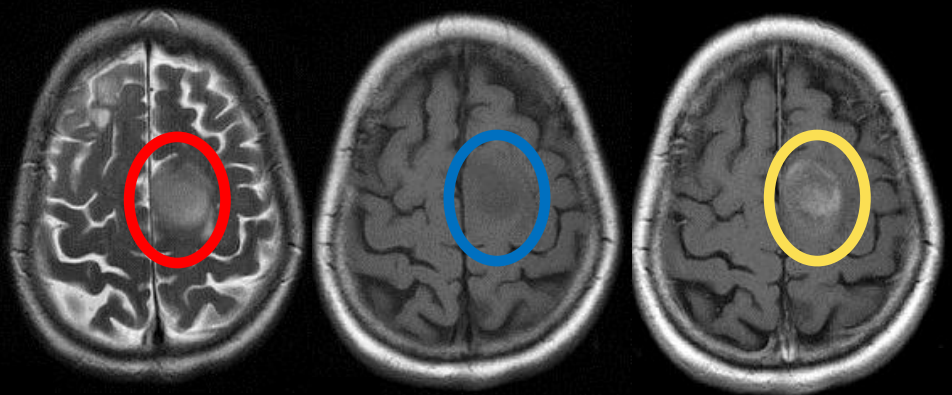
structural imaging

Outline

❑ Magnetic Resonance Imaging (MRI)

- Non-invasive imaging, great versatility

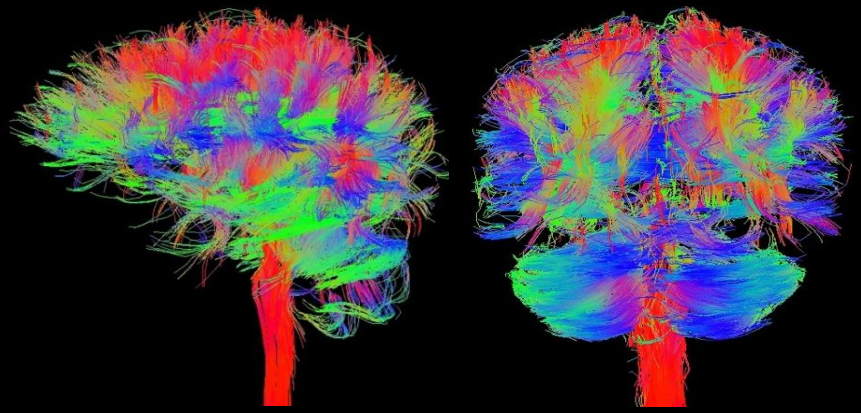
Patient with Glioma[1]



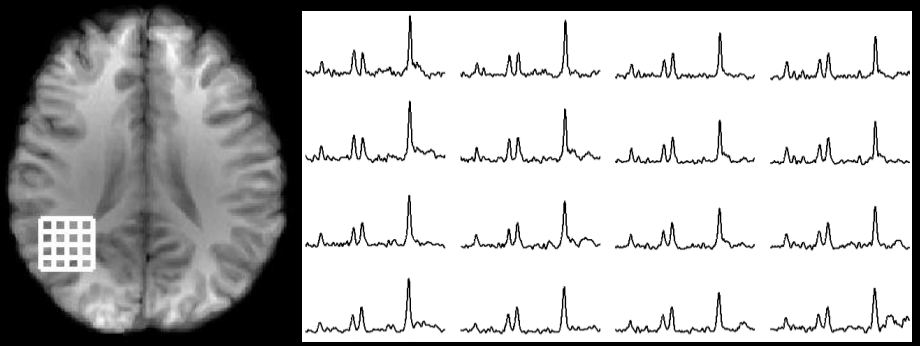
T2-Weighted

T1-Weighted

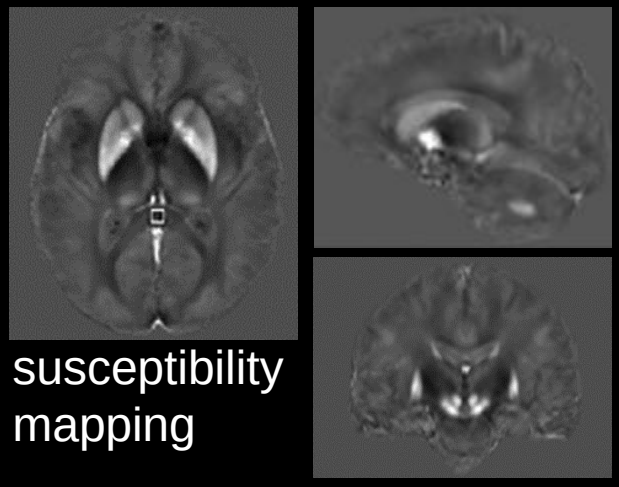
T1-Weighted
+ Gadolinium



diffusion imaging



spectroscopic imaging



susceptibility
mapping

[1] Burnet et al. Radiotherapy & Oncology'07

Outline

- **Magnetic Resonance Imaging (MRI)**
  - Non-invasive imaging, great versatility
  - Inherently slow, protocol takes ≥ 30 min

Outline

- **Magnetic Resonance Imaging (MRI)**
  - Non-invasive imaging, great versatility
  - Inherently slow, protocol takes ≥ 30 min
  - This limits the quality and resolution of the images

Outline

- **Magnetic Resonance Imaging (MRI)**
  - Non-invasive imaging, great versatility
  - Inherently slow, protocol takes ≥ 30 min
  - This limits the quality and resolution of the images
- **This thesis:** use prior knowledge about MR signals to
  - Reduce imaging time without sacrificing image quality
  - Mitigate image artifacts and provide quantitative imaging

Contributions

- Joint reconstruction

Contributions

- **Joint reconstruction**
  - Images with multiple contrasts are clinically routine

Contributions

❑ Joint reconstruction

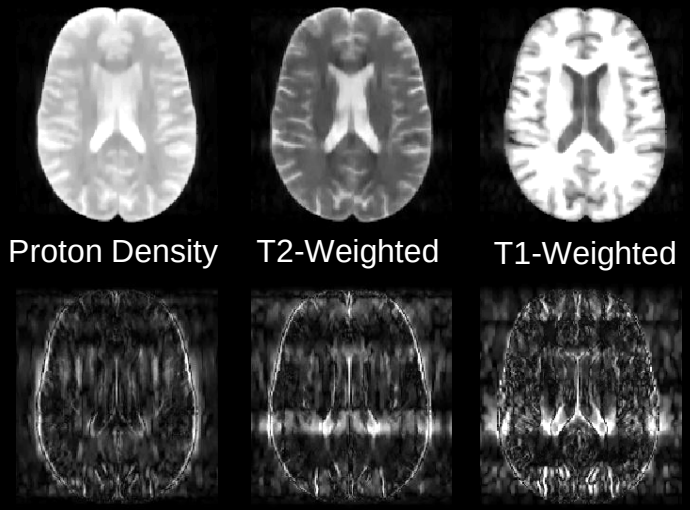
- Images with multiple contrasts are clinically routine

- Using 4-times less data than conventional (4x speed up):

SparseMRI  9.4 % error

State of the art: Sparse MRI

Lustig et al. MRM'07



Difference to fully-sampled: 9.4%

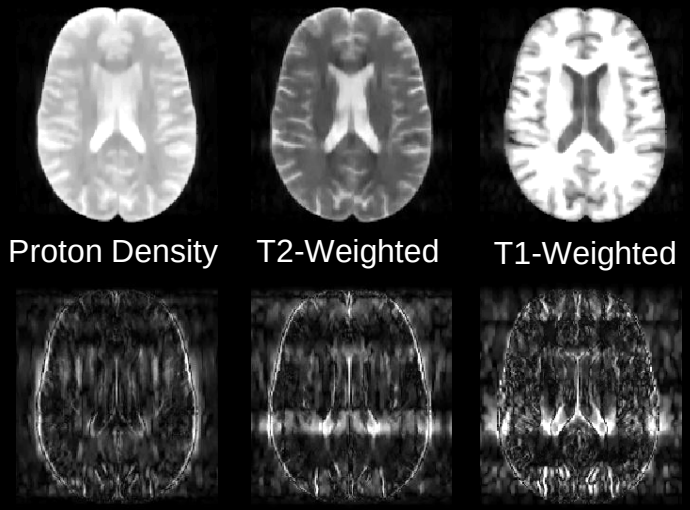
Contributions

Joint reconstruction

- Images with multiple contrasts are clinically routine
- Exploit their similarity for accelerated imaging
- Using 4-times less data than conventional (4x speed up):

SparseMRI 9.4 % error
Proposed 2.3 % error

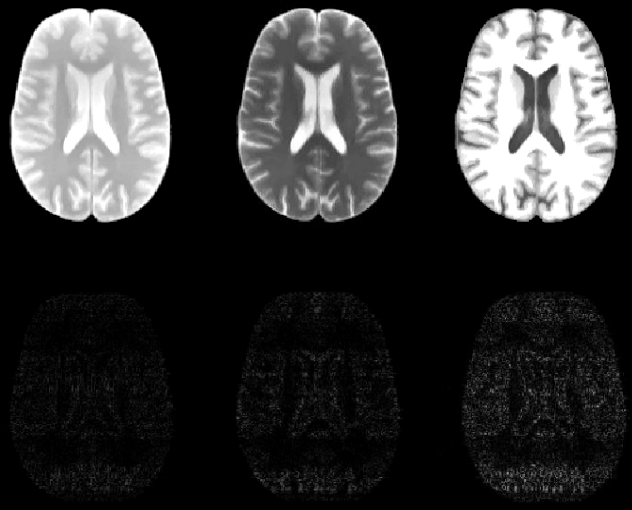
State of the art: Sparse MRI
Lustig et al. MRM'07



Proton Density T2-Weighted T1-Weighted

Difference to fully-sampled: 9.4%

Proposed: Joint Reconstruction
Bilgic et al. MRM'11



Difference to fully-sampled: 2.3%

Contributions

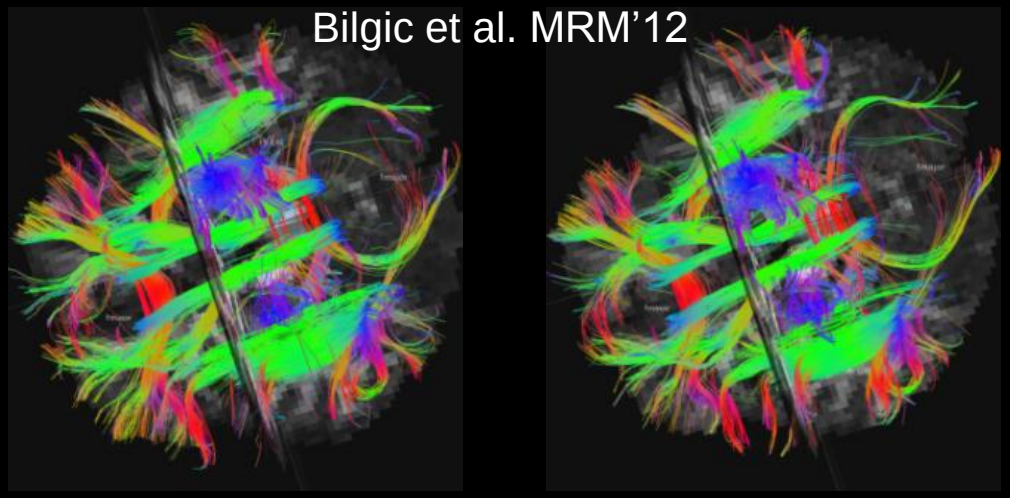
- Joint reconstruction
- **Diffusion Spectrum Imaging (DSI)**
  - DSI allows investigation of white matter connectivity of the brain
  - But suffers from very long scan times (~50 min)

Contributions

- Joint reconstruction
- **Diffusion Spectrum Imaging (DSI)**
  - DSI allows investigation of white matter connectivity of the brain
  - But suffers from very long scan times (~50 min)
  - 3-times less data than conventional → 17 min

Fully-sampled 50 min
Proposed 17 min

White matter fiber tracts
Bilgic et al. MRM'12



Fully-sampled data: 50-min scan time

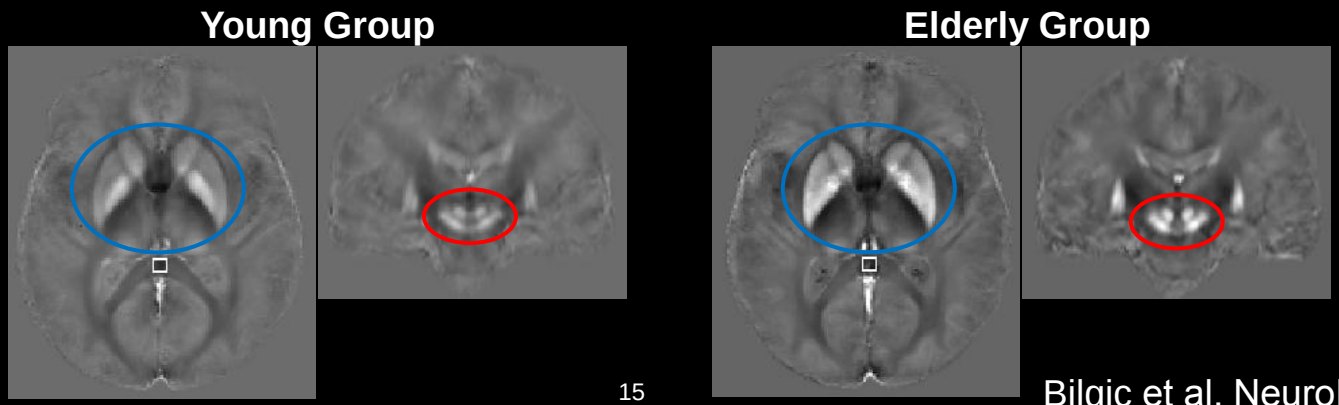
Proposed: 17-min scan time

Contributions

- Joint reconstruction
- Diffusion Spectrum Imaging (DSI)
- **Quantitative Susceptibility Mapping (QSM)**
  - QSM quantifies tissue iron concentration and vessel oxygenation
  - Susceptibility cannot be observed directly, needs to be inferred from MR signal phase

Contributions

- **Joint reconstruction**
- **Diffusion Spectrum Imaging (DSI)**
- **Quantitative Susceptibility Mapping (QSM)**
  - QSM quantifies tissue iron concentration and vessel oxygenation
  - Susceptibility cannot be observed directly, needs to be inferred from MR signal phase
  - QSM reveals increased iron during aging in striatal and brain stem regions



Bilgic et al. NeuroImage'12

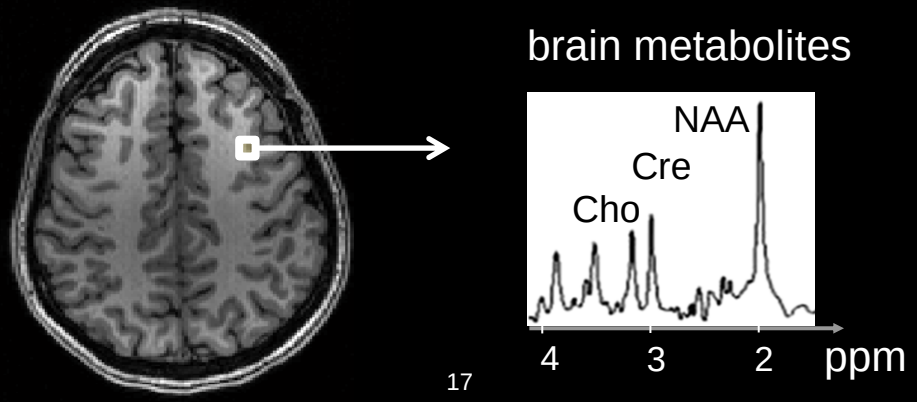
Contributions

- Joint reconstruction
- Diffusion Spectrum Imaging (DSI)
- Quantitative Susceptibility Mapping (QSM)
- **MR Spectroscopic Imaging (MRSI)**
  - In addition to spatial mapping, MRSI also provides encoding in resonance frequency

Contributions

- Joint reconstruction
- Diffusion Spectrum Imaging (DSI)
- Quantitative Susceptibility Mapping (QSM)
- **MR Spectroscopic Imaging (MRSI)**

- In addition to spatial mapping, MRSI also provides encoding in resonance frequency
- At each voxel, this yields a 1-d spectrum of relative biochemical metabolite concentrations



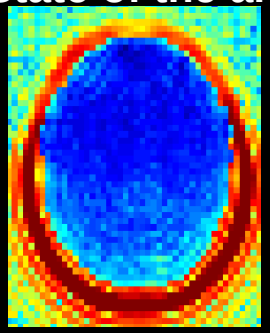
Contributions

- Joint reconstruction
- Diffusion Spectrum Imaging (DSI)
- Quantitative Susceptibility Mapping (QSM)
- **MR Spectroscopic Imaging (MRSI)**

  - Due to limited spatial resolution, strong lipid signals outside the brain contaminate the metabolite spectra inside the brain

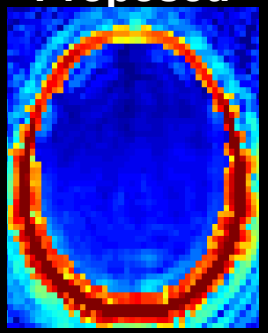
Sum over Lipid Frequencies

State of the art



Lee et al.
ISMRM'10

Proposed



Bilgic et al.
MRM'12

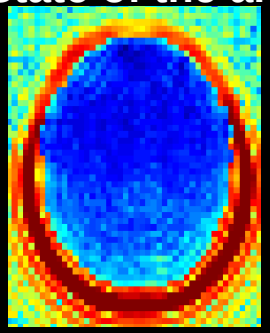
Contributions

- Joint reconstruction
- Diffusion Spectrum Imaging (DSI)
- Quantitative Susceptibility Mapping (QSM)
- **MR Spectroscopic Imaging (MRSI)**

- Due to limited spatial resolution, strong lipid signals outside the brain contaminate the metabolite spectra inside the brain

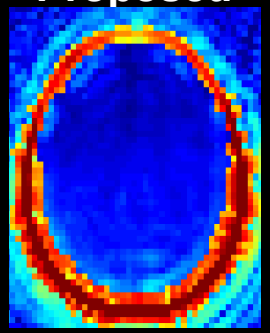
Sum over Lipid Frequencies

State of the art

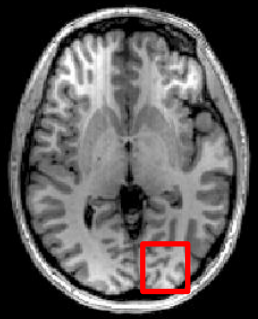


Lee et al.
ISMRM'10

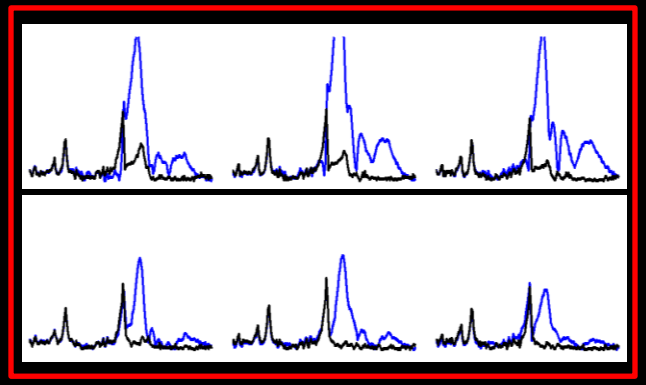
Proposed



Bilgic et al.
MRM'12



Structural image



Black: proposed
Blue: Lee et al.

Outline

- Problems that were addressed, why they are worth solving
- Contribution to the field

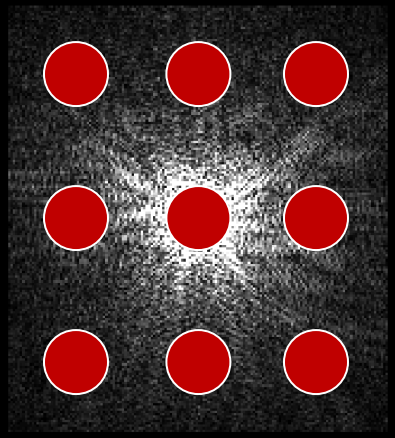
Outline

- Problems that were addressed, why they are worth solving
- Contribution to the field
- In particular,
  - Joint reconstruction of similar images
  - Accelerated Diffusion Spectrum Imaging
  - Quantifying tissue iron concentration
  - Lipid artifact suppression for Spectroscopic Imaging
    - Postpone to closed session

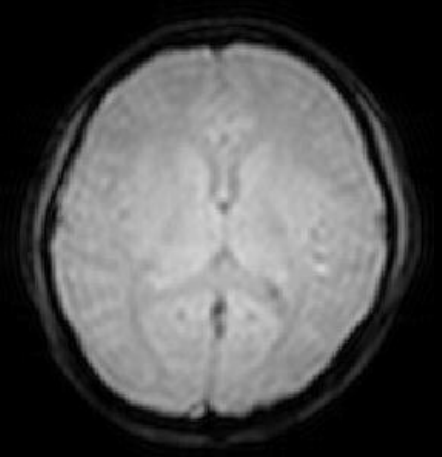
Outline

- Problems that were addressed, why they are worth solving
- Contribution to the field
- In particular,
  - Joint reconstruction of similar images
  - Accelerated Diffusion Spectrum Imaging
  - Quantifying tissue iron concentration
  - Lipid artifact suppression for Spectroscopic Imaging
    - Postpone to closed session

MRI Image Reconstruction

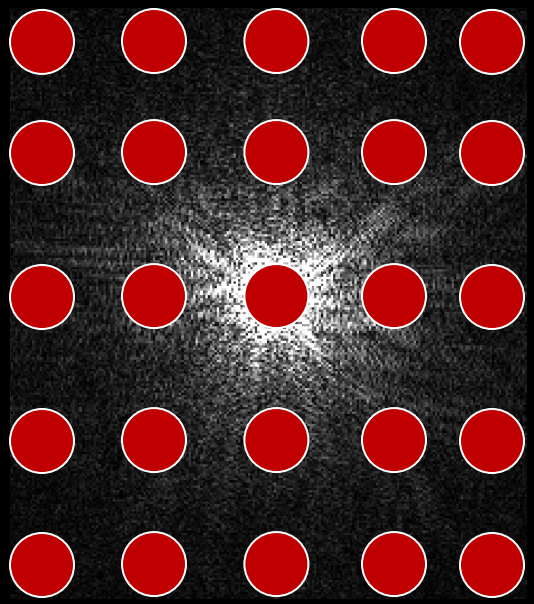


IDFT

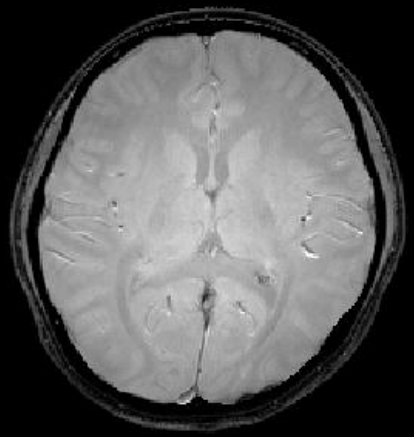


- In MRI, the data acquired are the Discrete Fourier Transform (DFT) samples of the object being imaged.
- Given sufficiently many samples (i.e. at Nyquist rate), taking the inverse DFT gives the spatial image.

MRI Image Reconstruction



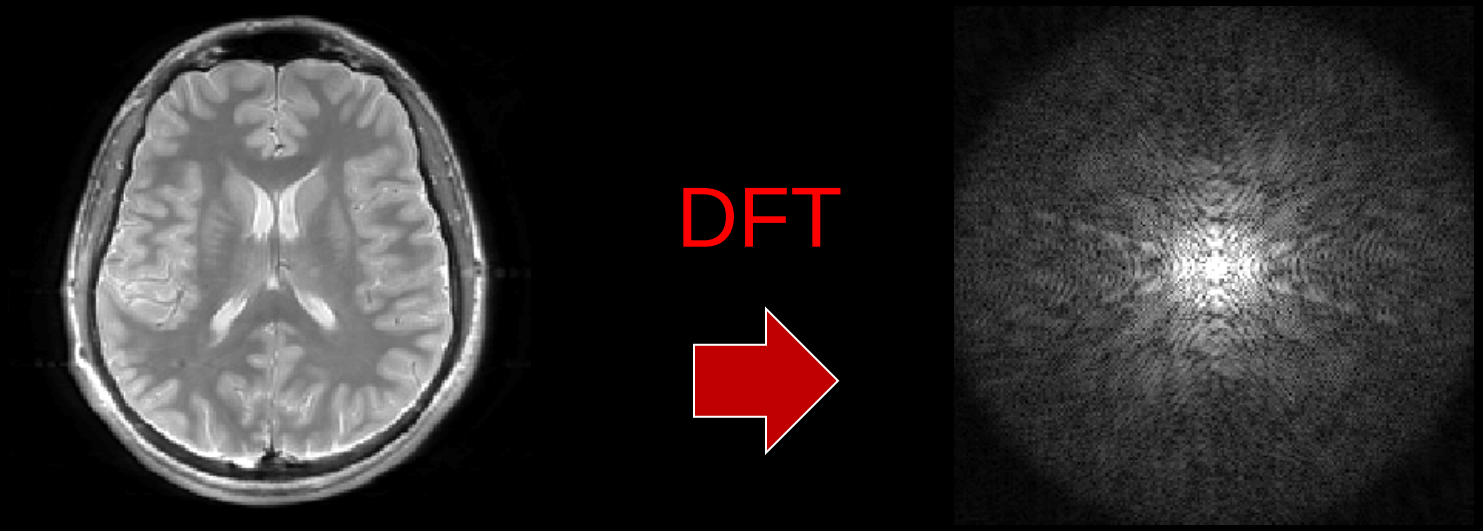
IDFT



- If we sample more of k-space, scan time increases
- For higher resolution images, we need to go further out in k-space => increased scan time

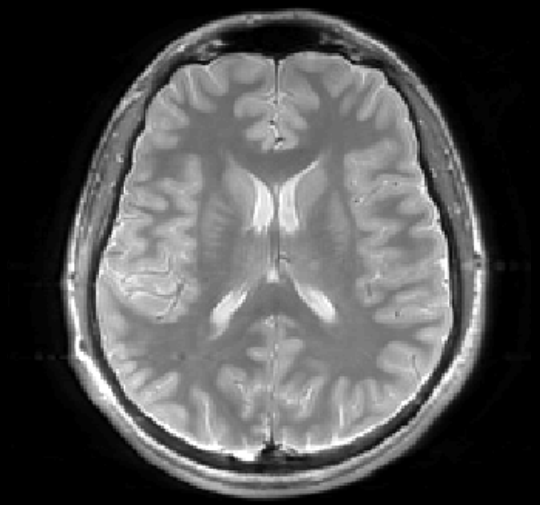
MRI Image Reconstruction

- For faster imaging, we can acquire less data (below Nyquist rate) but this incurs aliasing.



MRI Image Reconstruction

- For faster imaging, we can acquire less data (below Nyquist rate) but this incurs aliasing.

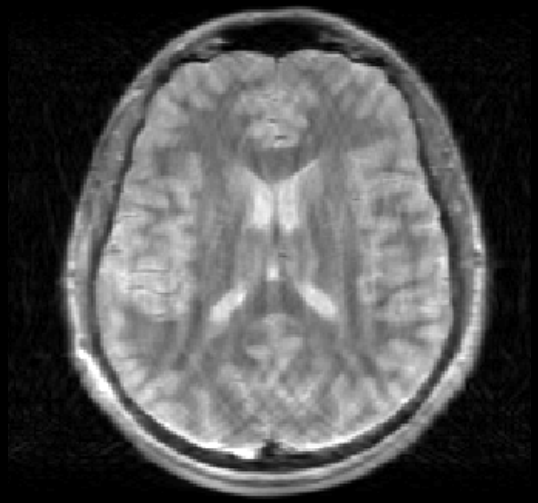


Undersample
remove 60% of data

60% reduction in scan time

MRI Image Reconstruction

- For faster imaging, we can acquire less data (below Nyquist rate) but this incurs aliasing.



IDFT

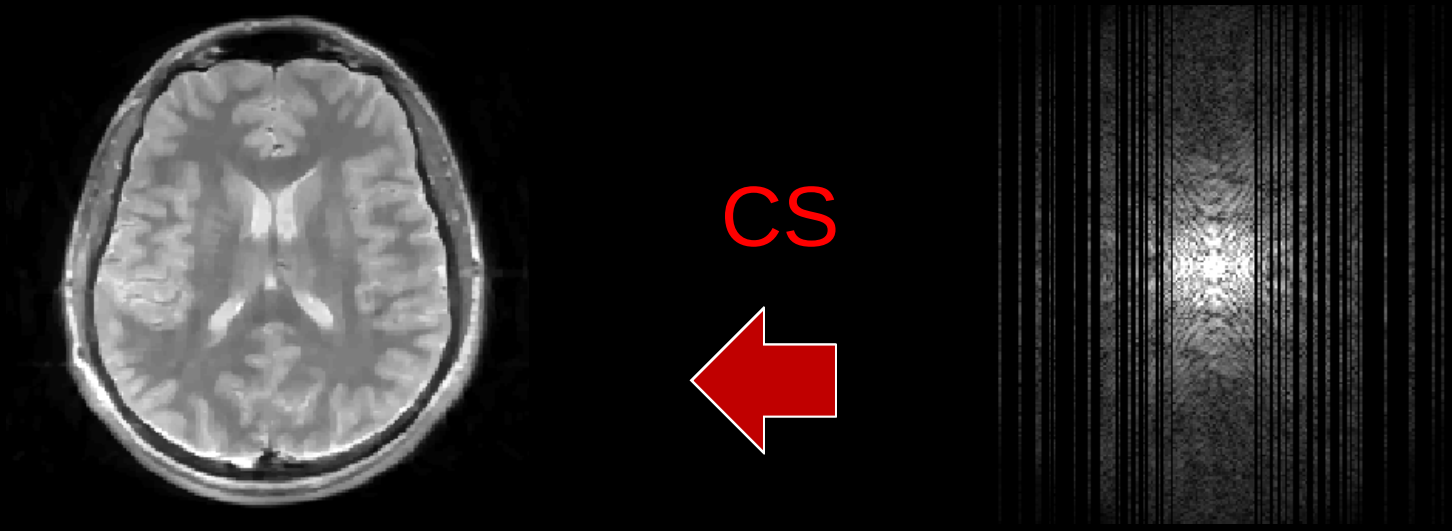
Undersample
remove 60% of data

RMSE = 11.7 %

60% reduction in scan time

Compressed Sensing (CS) reconstruction

- Reduce aliasing artifacts by imposing prior knowledge in reconstruction[1]
- CS prior: image is sparse under a transform

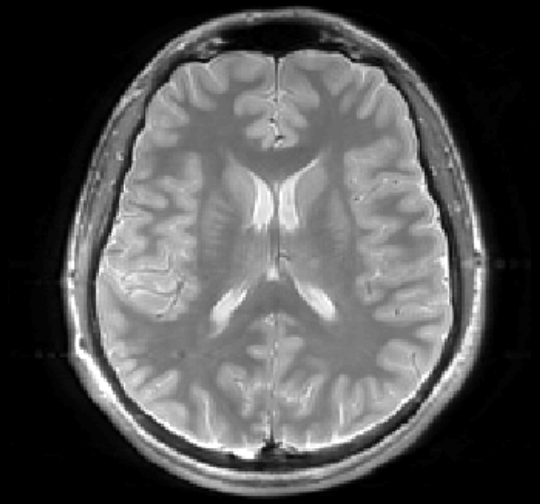


RMSE = 5.9 %

[1] Lustig *et al*. MRM 2007

Total Variation prior

- Total Variation (TV): Most popular transform for CS recon
- Prior: spatial gradient of the image is sparse



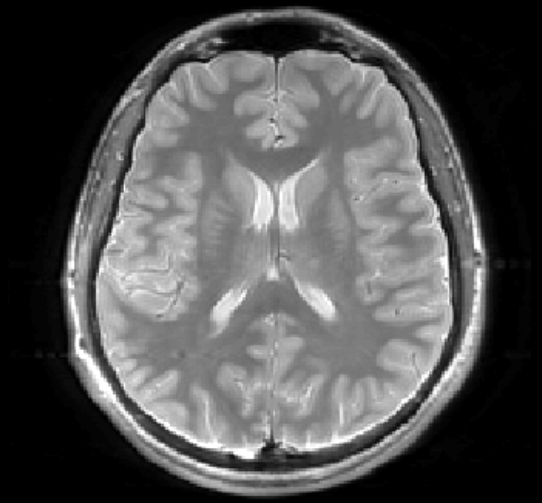
Gradient

Total Variation prior

- Total Variation (TV): Most popular transform for CS recon
- Prior: spatial gradient of the image is sparse

$$min_{img} \|\mathbf{F}_\Omega \cdot img - data\|_2{}^2 + \lambda \cdot \|\mathbf{G} \cdot img\|_1$$

- $\mathbf{F}_\Omega$: undersampled DFT
- img: image
- $data$: k-space samples
- $\mathbf{G}$: gradient operator



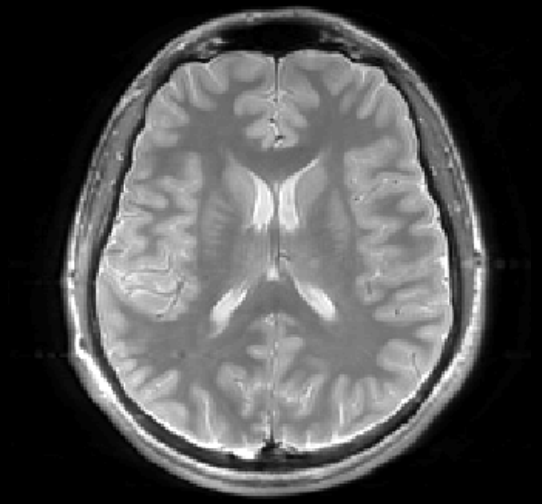
Gradient

Total Variation prior

- Total Variation (TV): Most popular transform for CS recon
- Prior: spatial gradient of the image is sparse

$$min_{img} \|\mathbf{F}_\Omega \cdot img - data\|_2^{\,2} + \lambda \cdot \underbrace{\|\mathbf{G} \cdot img\|_1}_{\text{Total Variation}}$$

$\mathbf{F}_\Omega$: undersampled DFT; img: image; $data$: k-space samples



Gradient

Multi-contrast data acquisition

- In clinical MRI, it is common to image the same region of interest under multiple contrast settings
- This aims to increase the diagnostic power of MRI as tissues exhibit different characteristics under different contrasts
- For instance, SRI24 atlas[1] contains such multi-contrast data,

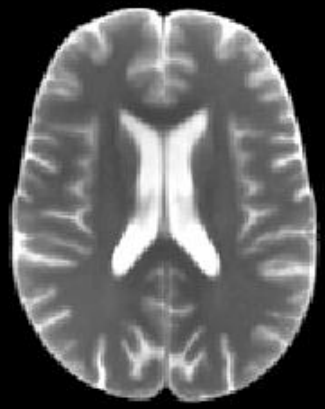
proton density

[1] Rohlfing *et al*. Hum Brain Map, 2010

Multi-contrast data acquisition

- In clinical MRI, it is common to image the same region of interest under multiple contrast settings
- This aims to increase the diagnostic power of MRI as tissues exhibit different characteristics under different contrasts
- For instance, SRI24 atlas[1] contains such multi-contrast data,

T2 weighted



[1] Rohlfing *et al*. Hum Brain Map, 2010

Multi-contrast data acquisition

- In clinical MRI, it is common to image the same region of interest under multiple contrast settings
- This aims to increase the diagnostic power of MRI as tissues exhibit different characteristics under different contrasts
- For instance, SRI24 atlas[1] contains such multi-contrast data,

T1 weighted

[1] Rohlfing *et al*. Hum Brain Map, 2010

Multi-contrast data acquisition

- In clinical MRI, it is common to image the same region of interest under multiple contrast settings
- This aims to increase the diagnostic power of MRI as tissues exhibit different characteristics under different contrasts
- For instance, SRI24 atlas[1] contains such multi-contrast data,

T2 weighted

T1 weighted

proton density

Joint reconstruction with ℓ_1-ℓ_2 regularization

- To couple multi-contrast signals,
  - take the ℓ_2 norm across the contrast dimension,
  - then apply ℓ_1 regularization to the combination,

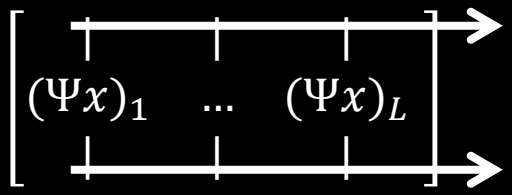
$$\sum_{i=1}^{L} \|\mathbf{F}_\Omega x_i - y_i\|_2^{\,2} + \lambda \cdot \sum_{j=1}^{N} \left(\sum_{i=1}^{L} (\Psi x)_{i,j}^{\,2} \right)^{1/2}$$

data consistency for L images

ℓ_2 across contrasts in transform domain

ℓ_1 over combination

- Prior: few non-zero rows



ℓ_2 across contrasts

Joint reconstruction with ℓ_1-ℓ_2 regularization

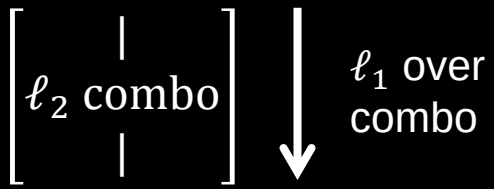
- To couple multi-contrast signals,
  - take the ℓ_2 norm across the contrast dimension,
  - then apply ℓ_1 regularization to the combination,

$$\sum_{i=1}^{L} \|\mathbf{F}_\Omega x_i - y_i\|_2^{\,2} + \lambda \cdot \sum_{j=1}^{N} \left(\sum_{i=1}^{L} (\Psi x)_{i,j}^{\,2} \right)^{1/2}$$

data consistency for L images

ℓ_2 across contrasts in transform domain

ℓ_1 over combination



Joint reconstruction with ℓ_1-ℓ_2 regularization

- To couple multi-contrast signals,
  - take the ℓ_2 norm across the contrast dimension,
  - then apply ℓ_1 regularization to the combination,

$$\sum_{i=1}^{L} \|\mathbf{F}_\Omega x_i - y_i\|_2^{\,2} + \lambda \cdot \sum_{j=1}^{N} \left(\sum_{i=1}^{L} (\Psi x)_{i,j}^{\,2} \right)^{1/2}$$

data consistency for L images

ℓ_2 across contrasts in transform domain

ℓ_1 over combination

- **M-FOCUSS**[1] is an iteratively reweighted ℓ_2 regularization algorithm that solves this optimization problem

[1] Cotter *et al*. T Signal Proces, 2005

Joint reconstruction with Bayesian CS

- Alternative approach: model the transform coefficients across contrasts for a single voxel as random variables with common variance

Joint reconstruction with Bayesian CS

- Alternative approach: model the transform coefficients across contrasts for a single voxel as random variables with common variance

- The most likely variance at each voxel is estimated using Bayesian inference given the observed *k*-space data

Joint reconstruction with Bayesian CS

- Alternative approach: model the transform coefficients across contrasts for a single voxel as random variables with common variance
- The most likely variance at each voxel is estimated using Bayesian inference given the observed *k*-space data
- This model is more flexible than L1-L2 regularization, as there is no common sparsity support assumption across contrasts

BCS Theory: Observation model

$$\mathbf{F}_\Omega \, \boldsymbol{x} \;=\; \boldsymbol{y}$$

$\mathbf{F}_\Omega$: partial Fourier transform

$\boldsymbol{x}$: image to be estimated

$\boldsymbol{y}$: undersampled k-space data

Observation model – sparse representation

$$\mathbf{V}\,\mathbf{F}_{\Omega}\,\boldsymbol{x} = \mathbf{V}\,\boldsymbol{y}$$

$$\mathbf{V} = \left(\mathbf{1} - \boldsymbol{e}^{-2\pi j\omega/n}\right)$$

k-space representation of differencing: $x_i - x_{i-1}$

Observation model – sparse representation

$$\mathbf{F}_\Omega \, \boldsymbol{\delta} = \tilde{\boldsymbol{y}}$$

$\boldsymbol{\delta}$: image gradient to be estimated

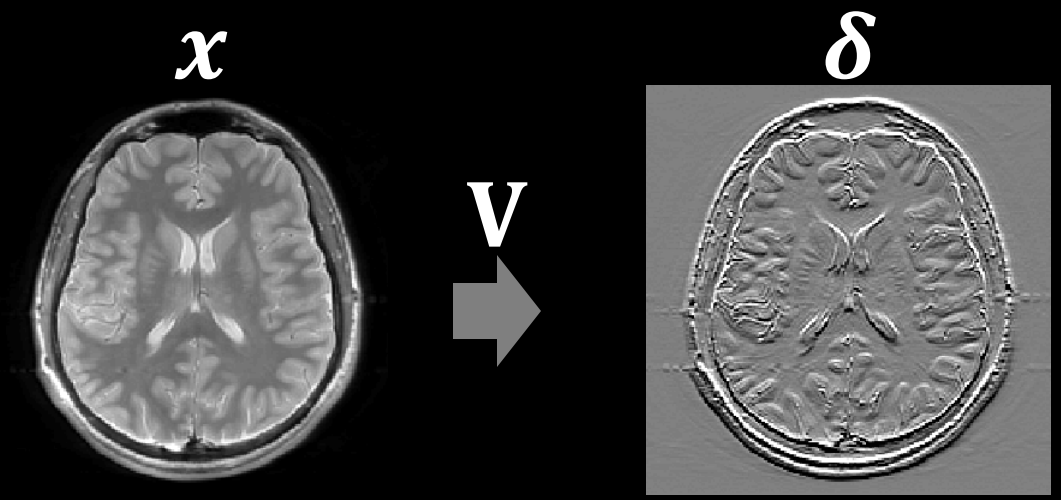
$\tilde{\boldsymbol{y}}$: modified k-space data

Observation model – sparse representation

$$\mathbf{F}_\Omega \, \boldsymbol{\delta} = \tilde{\boldsymbol{y}}$$

$\boldsymbol{\delta}$: image gradient to be estimated

$\tilde{\boldsymbol{y}}$: modified k-space data



Data likelihood

- Assuming that the k-space data are corrupted by complex-valued Gaussian noise with σ^2 variance,

$$\underbrace{\mathrm{p}(\tilde{\boldsymbol{y}} \mid \boldsymbol{\delta}, \sigma^2)}_{\text{Gaussian likelihood}} \sim \mathcal{N}(\mathbf{F}_\Omega \boldsymbol{\delta} - \tilde{\boldsymbol{y}}, \sigma^2)$$

Prior distribution on gradient coefficients

- Bayesian CS places hyperparameters γ on each pixel,

$$\underbrace{p(\delta_i \mid \gamma_i)}_{\text{Gaussian prior}} \sim \mathcal{N}(0, \gamma_i)$$

- So that i[th] pixel is a zero-mean Gaussian with variance γ_i

Prior distribution on gradient coefficients

- Bayesian CS places hyperparameters γ on each pixel,

$$\underbrace{p(\delta_i \mid \gamma_i)}_{\text{Gaussian prior}} \sim \mathcal{N}(0, \gamma_i)$$

- So that i[th] pixel is a zero-mean Gaussian with variance γ_i

- Multiplicative combination of all pixels give the full prior distribution,

$$p(\boldsymbol{\delta} \mid \boldsymbol{\gamma}) \sim \prod_i \mathcal{N}(0, \gamma_i)$$

Posterior distribution for gradient coefficients

- Using the likelihood and the prior, we invoke Bayes' Rule to arrive at the posterior,

$$\mathrm{p}(\boldsymbol{\delta} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\gamma}) \;\propto\; \mathrm{p}(\boldsymbol{\delta} \mid \boldsymbol{\gamma}) \cdot \mathrm{p}(\tilde{\boldsymbol{y}} \mid \boldsymbol{\delta})$$

Posterior distribution for gradient coefficients

- Using the likelihood and the prior, we invoke Bayes' Rule to arrive at the posterior,

$$\underbrace{\mathrm{p}(\boldsymbol{\delta} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\gamma})}_{\text{Gaussian posterior}} \propto \underbrace{\mathrm{p}(\boldsymbol{\delta} \mid \boldsymbol{\gamma})}_{\text{Gaussian prior}} \cdot \underbrace{\mathrm{p}(\tilde{\boldsymbol{y}} \mid \boldsymbol{\delta})}_{\text{Gaussian likelihood}}$$

Posterior distribution for gradient coefficients

- Using the likelihood and the prior, we invoke Bayes' Rule to arrive at the posterior,

$$p(\boldsymbol{\delta} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\gamma}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Posterior distribution for gradient coefficients

- Using the likelihood and the prior, we invoke Bayes' Rule to arrive at the posterior,

$$p(\boldsymbol{\delta} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\gamma}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

$$\boldsymbol{\Gamma} = diag(\boldsymbol{\gamma})$$

$$\mathbf{A}^{-1} = (\sigma^2\mathbf{I} + \mathbf{F}_{\Omega}\boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H})^{-1} \quad \rightarrow 10^4 \times 10^4 \text{ matrix inversion}$$

Posterior distribution for gradient coefficients

- Using the likelihood and the prior, we invoke Bayes' Rule to arrive at the posterior,

$$p(\boldsymbol{\delta} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\gamma}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

$$\boldsymbol{\Gamma} = diag(\boldsymbol{\gamma})$$

$$\mathbf{A}^{-1} = (\sigma^2\mathbf{I} + \mathbf{F}_{\Omega}\boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H})^{-1}$$

Inversion using Lanczos algorithm[1]

[1] Seeger et al. MRM, 2010

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Maximization step:

$$\gamma_i = \frac{|\mu_i|^2}{1 - \Sigma_{ii}/\gamma_i}$$

[1] Wipf et al. IEEE Trans Signal Process, 2007

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boldsymbol{\mu} = \boxed{\boldsymbol{\Gamma}}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boxed{\boldsymbol{\Gamma}} - \boxed{\boldsymbol{\Gamma}}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boxed{\boldsymbol{\Gamma}}$$

Maximization step:

$$\boxed{\gamma_i} = \frac{|\mu_i|^2}{1 - \Sigma_{ii}/\gamma_i}$$

[1] Wipf et al. IEEE Trans Signal Process, 2007

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boxed{\boldsymbol{\mu}} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boxed{\boldsymbol{\Sigma}} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Maximization step:

$$\gamma_i = \frac{\left|\boxed{\mu_i}\right|^2}{1 - \boxed{\Sigma_{ii}}/\gamma_i}$$

[1] Wipf et al. IEEE Trans Signal Process, 2007

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Maximization step:

$$\gamma_i = \frac{|\mu_i|^2}{1 - \Sigma_{ii}/\gamma_i}$$

→ for a single image

[1] Wipf et al. IEEE Trans Signal Process, 2007

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Maximization step:

$$\gamma_i = \frac{\|\mu_1, \ldots, \mu_L\|^2}{L - L \cdot \Sigma_{ii}/\gamma_i}$$

$\rightarrow$ for L images jointly

[1] Wipf et al. IEEE Trans Signal Process, 2007

EM algorithm for optimization

- Expectation-maximization algorithm[1] is used to estimate the hyperparameters and the posterior iteratively,

Expectation step:

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\tilde{\boldsymbol{y}}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{F}_{\Omega}{}^{H}\mathbf{A}^{-1}\mathbf{F}_{\Omega}\boldsymbol{\Gamma}$$

Maximization step:

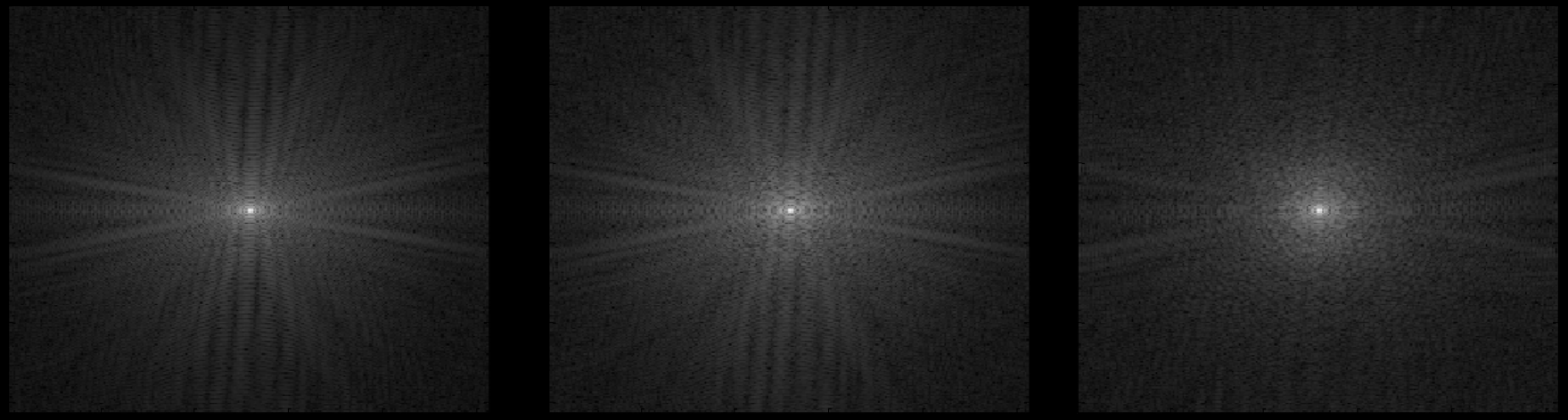
$$\gamma_i = \frac{\|\mu_1, \ldots, \mu_L\|^2}{L - L \cdot \Sigma_{ii}/\gamma_i}$$

All images are used to estimate the variance:

Contrasts are coupled

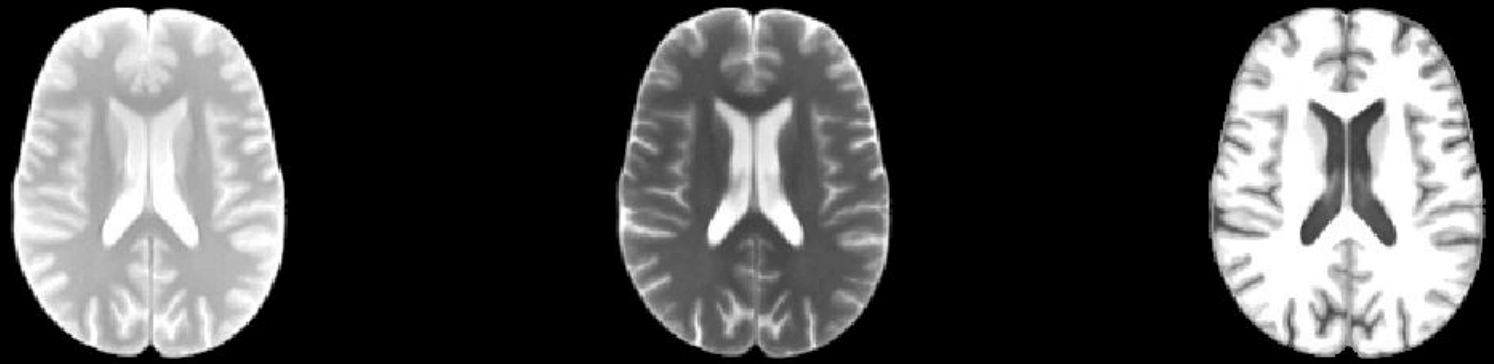
[1] Wipf et al. IEEE Trans Signal Process, 2007

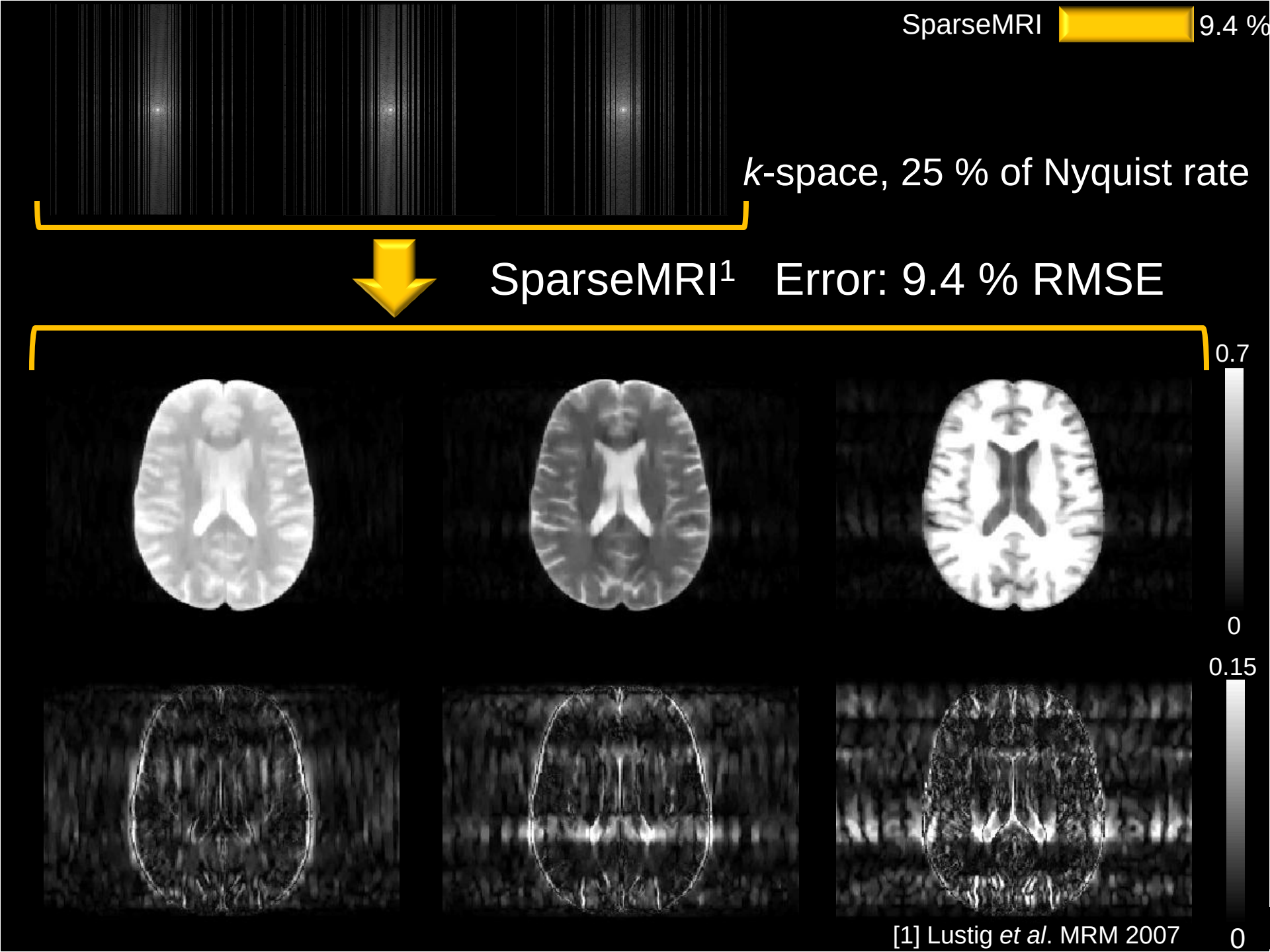
SRI24 Atlas

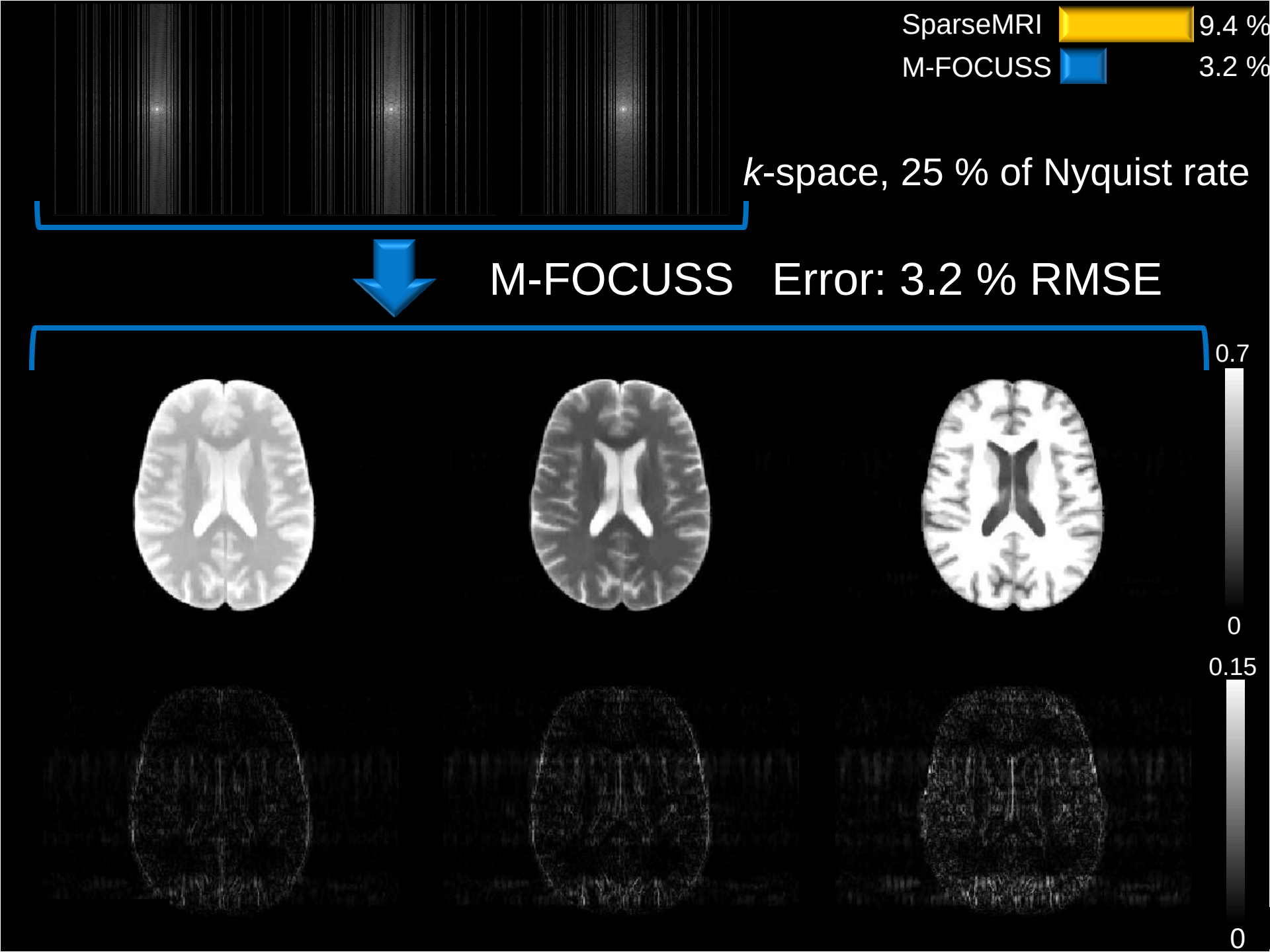


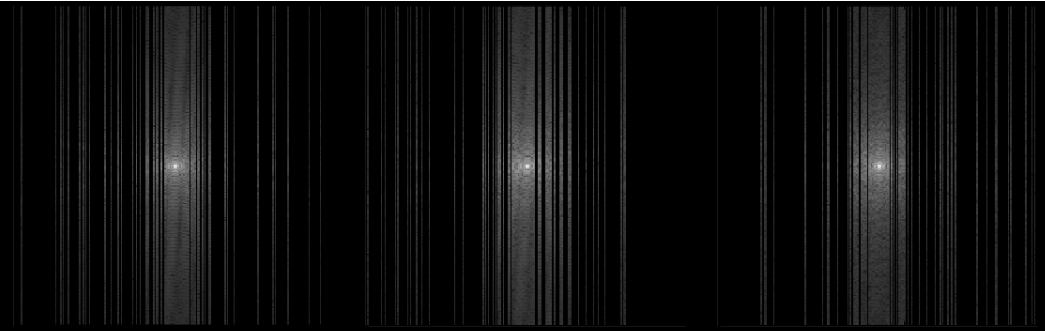
k-space, 100 % of Nyquist rate

Inverse FFT Error: 0 % RMSE




SparseMRI
9.4 %
k-space, 25 % of Nyquist rate
SparseMRI[1] Error: 9.4 % RMSE
0.7
0
0.15
0
[1] Lustig et al. MRM 2007


SparseMRI
9.4 %
M-FOCUSS
3.2 %
k-space, 25 % of Nyquist rate
M-FOCUSS Error: 3.2 % RMSE
0.7
0
0.15
0



SparseMRI 9.4 %
M-FOCUSS 3.2 %
Joint Bayes 2.3 %

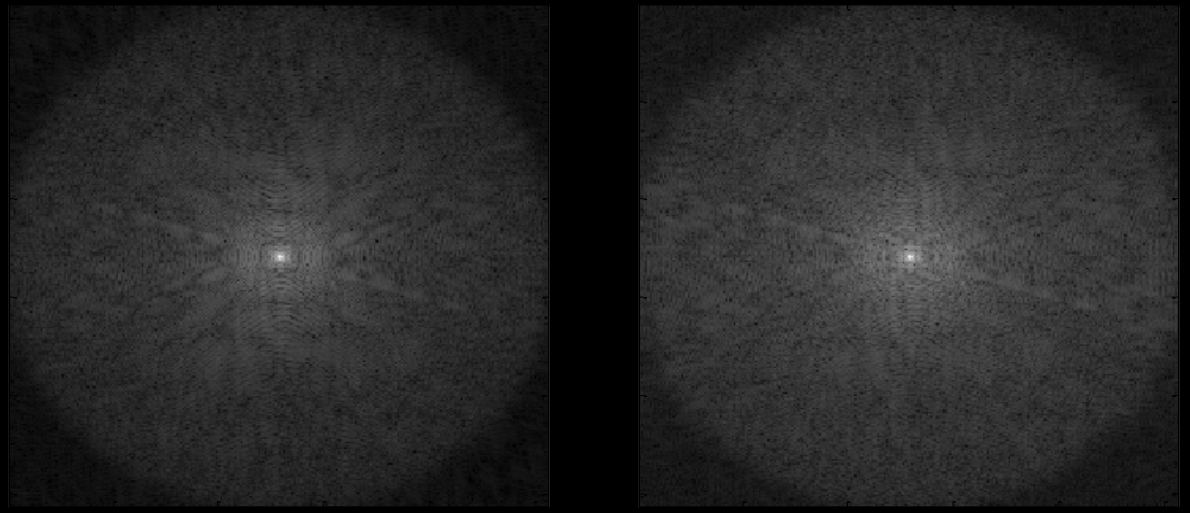
k-space, 25 % of Nyquist rate

Joint BCS Error: 2.3 % RMSE

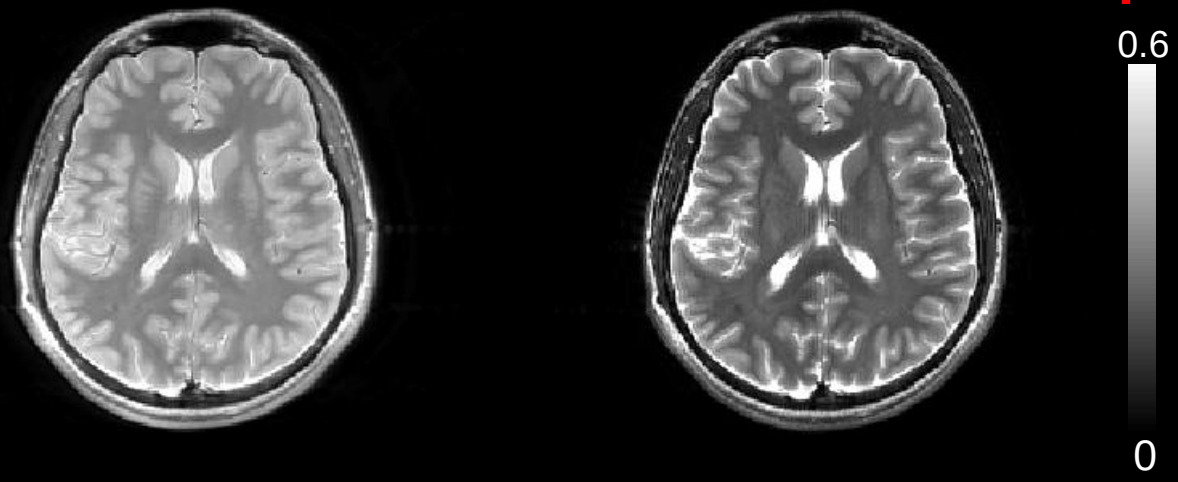
0.7
0
0.15
0

TSE Scans : *in vivo* acquisition

k-space
100 % of Nyquist rate



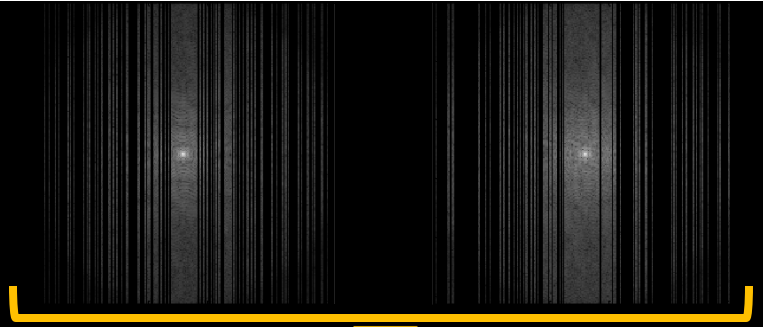
Inverse FFT



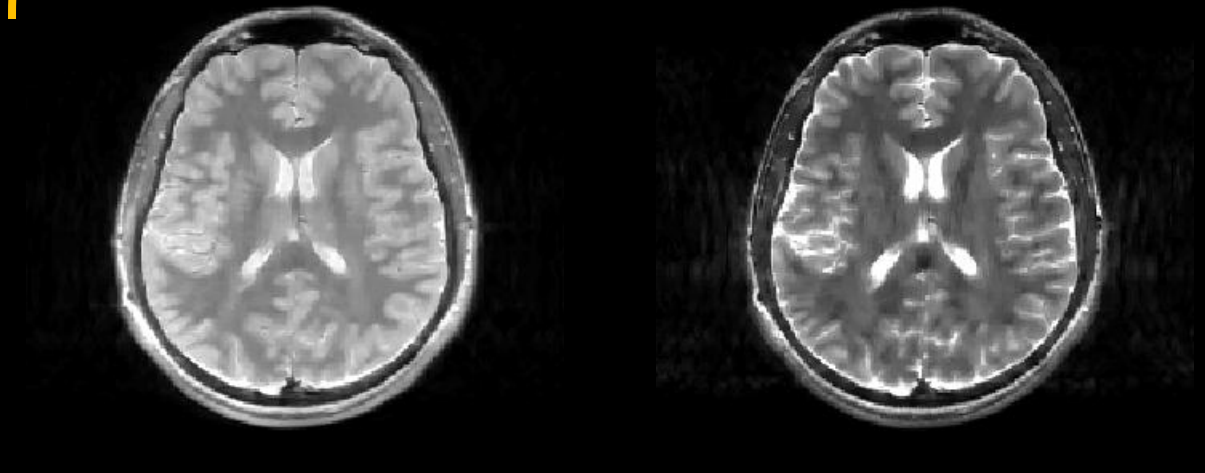
Error: 0 % RMSE

SparseMRI 9.4 %

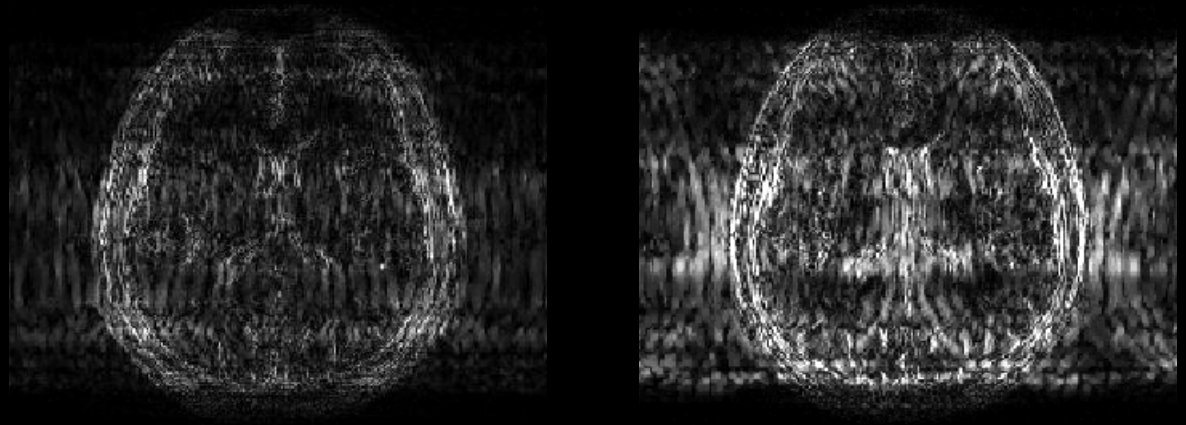
k-space, 40 % of Nyquist rate



SparseMRI[1]



Error: 9.4 % RMSE

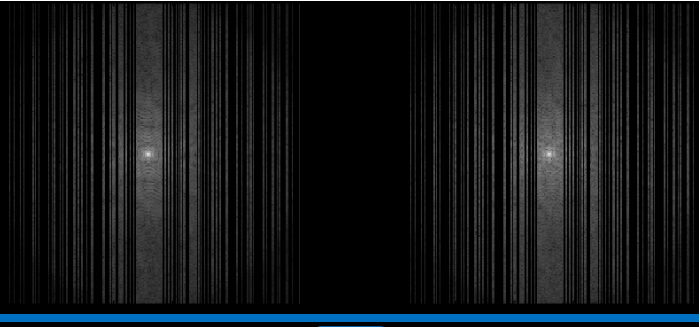


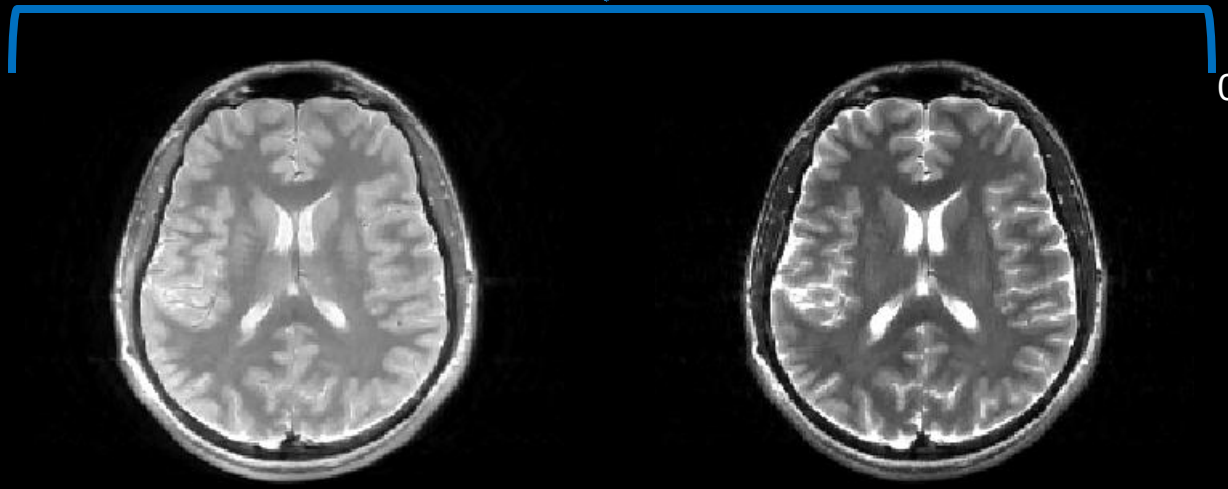
[1] Lustig *et al*. MRM 2007

SparseMRI 9.4 %

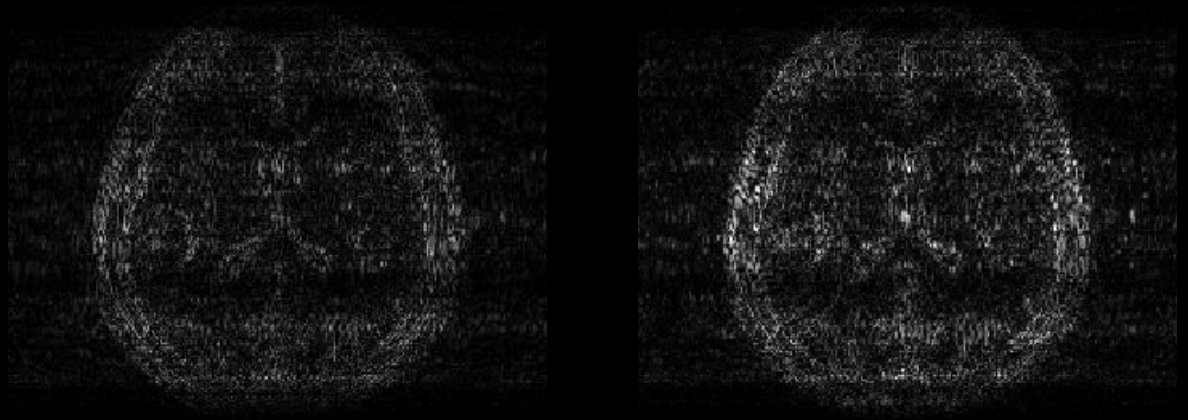
M-FOCUSS 5.1 %

k-space, 40 % of Nyquist rate

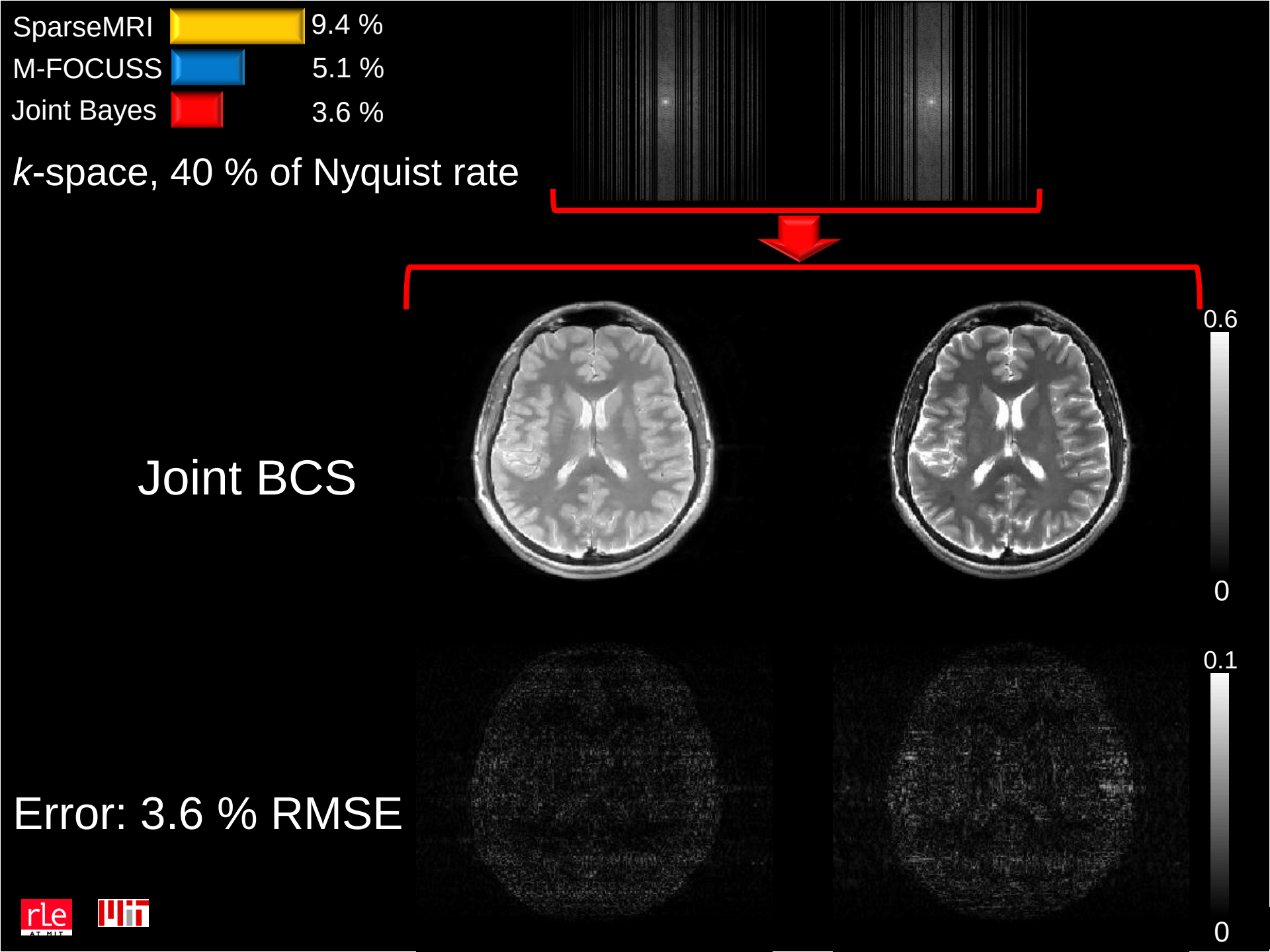




M-FOCUSS



Error: 5.1 % RMSE


SparseMRI
9.4 %
M-FOCUSS
5.1 %
Joint Bayes
3.6 %
k-space, 40 % of Nyquist rate
Joint BCS
0.6
0
Error: 3.6 % RMSE
0.1
0

Joint Reconstruction Conclusion

- Demonstrated improved reconstruction quality for multi-contrast imaging by exploiting similarity across contrasts

- Proposed to use two methods for joint reconstruction:
  - M-FOCUSS: ℓ_1-ℓ_2 regularization
  - Bayesian CS: common variance

Joint Reconstruction Conclusion

	method	speed	quality
SparseMRI[1]	1 by 1	~ minutes	good
M-FOCUSS	joint	~ minutes	better
Bayesian CS	joint	~ hours	best

[1] Lustig *et al*. MRM 2007

Future Directions in Joint Reconstruction

- **Bayesian CS computation speed**
  - Current implementation: several hours / slice
  - Bottleneck: matrix inversion for covariance estimation
  - Initial results with sparse matrix inversion: several minutes[1]

[1] Cauley *et al*. ISMRM 2013 *submitted*

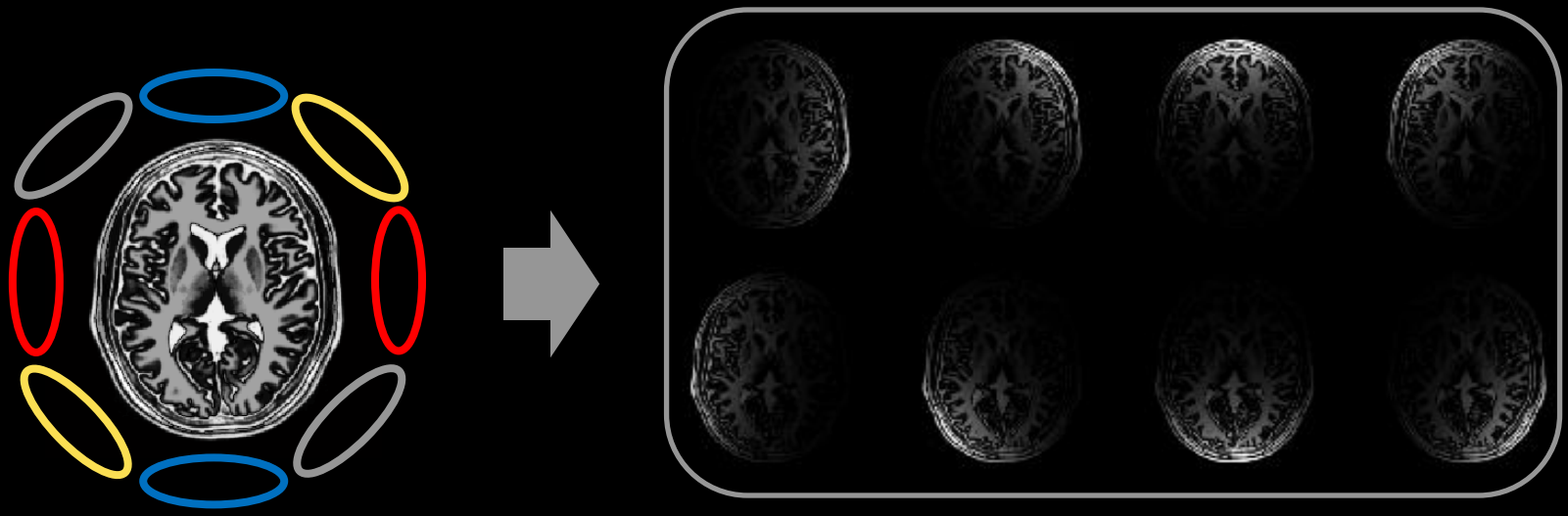
Future Directions in Joint Reconstruction

Bayesian CS computation speed

- Initial results with sparse matrix inversion: several minutes

Extension to Parallel Imaging

- Information from multiple receivers facilitate reconstruction from undersampled data



8-receivers

each receiver has different spatial sensitivity

Future Directions in Joint Reconstruction

- **Bayesian CS computation speed**
  - Initial results with sparse matrix inversion: several minutes
- **Extension to Parallel Imaging**
  - Information from multiple receivers facilitate reconstruction from undersampled data
  - Matrix inversion becomes ~ $10^5 \times 10^5$, ongoing research

Future Directions in Joint Reconstruction

Bayesian CS computation speed

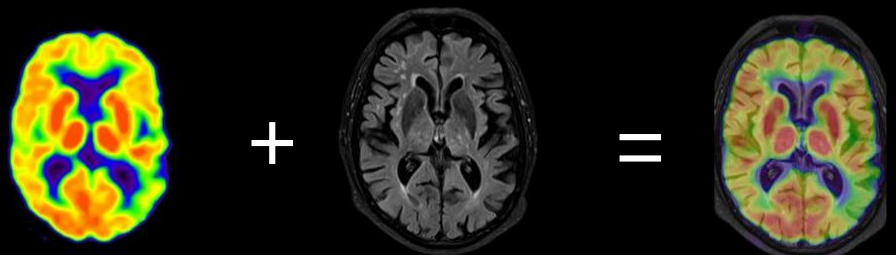
- Initial results with sparse matrix inversion: several minutes

Extension to Parallel Imaging

- Information from multiple receivers facilitate reconstruction from undersampled data
- Matrix inversion becomes ~ $10^5 \times 10^5$, ongoing research

Multi-modal Imaging

- Extend joint reconstruction to PET / MRI[1] etc.



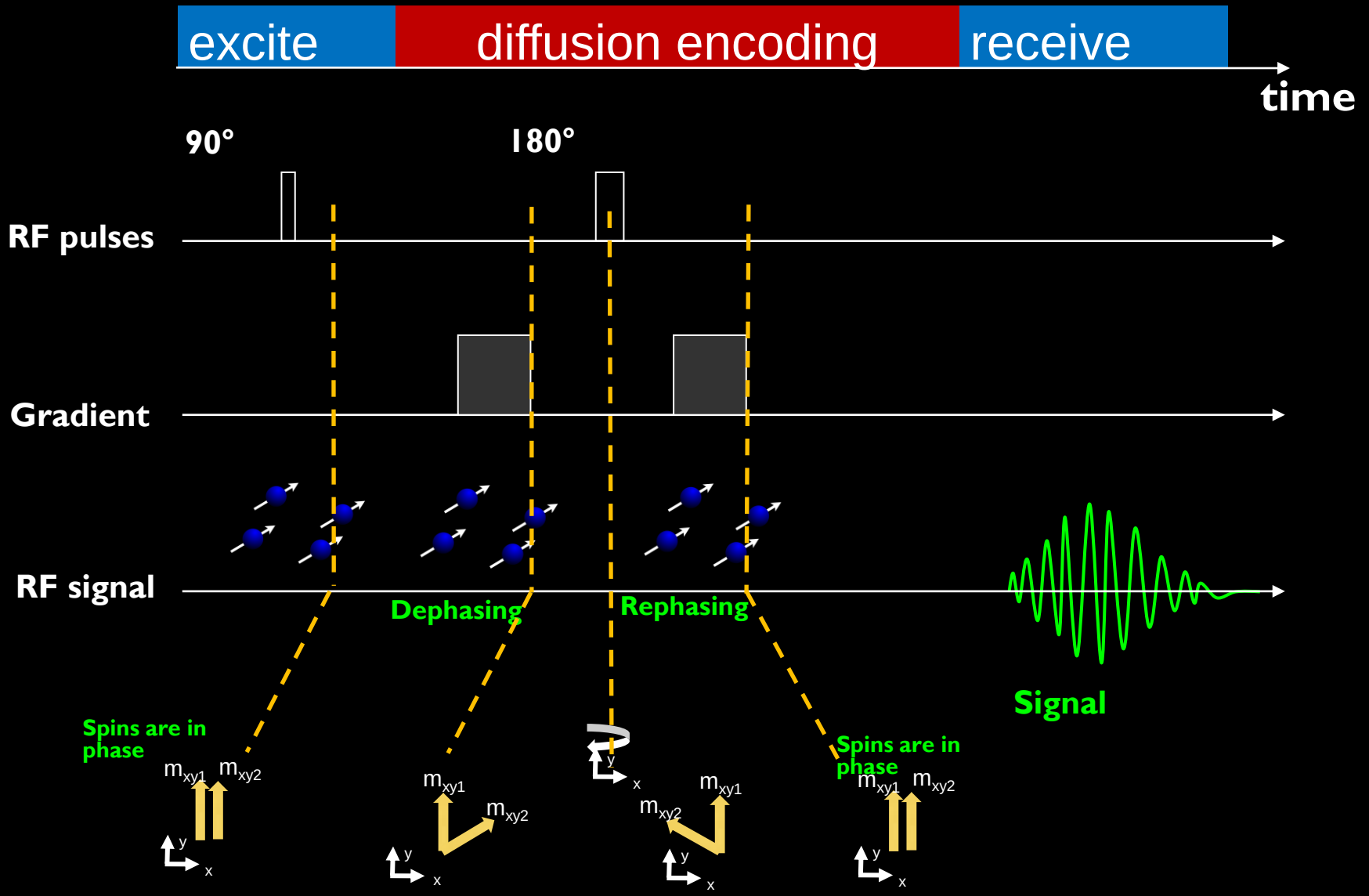
PET MRI [73] PET + MRI

[1] Siemens Biograph mMR

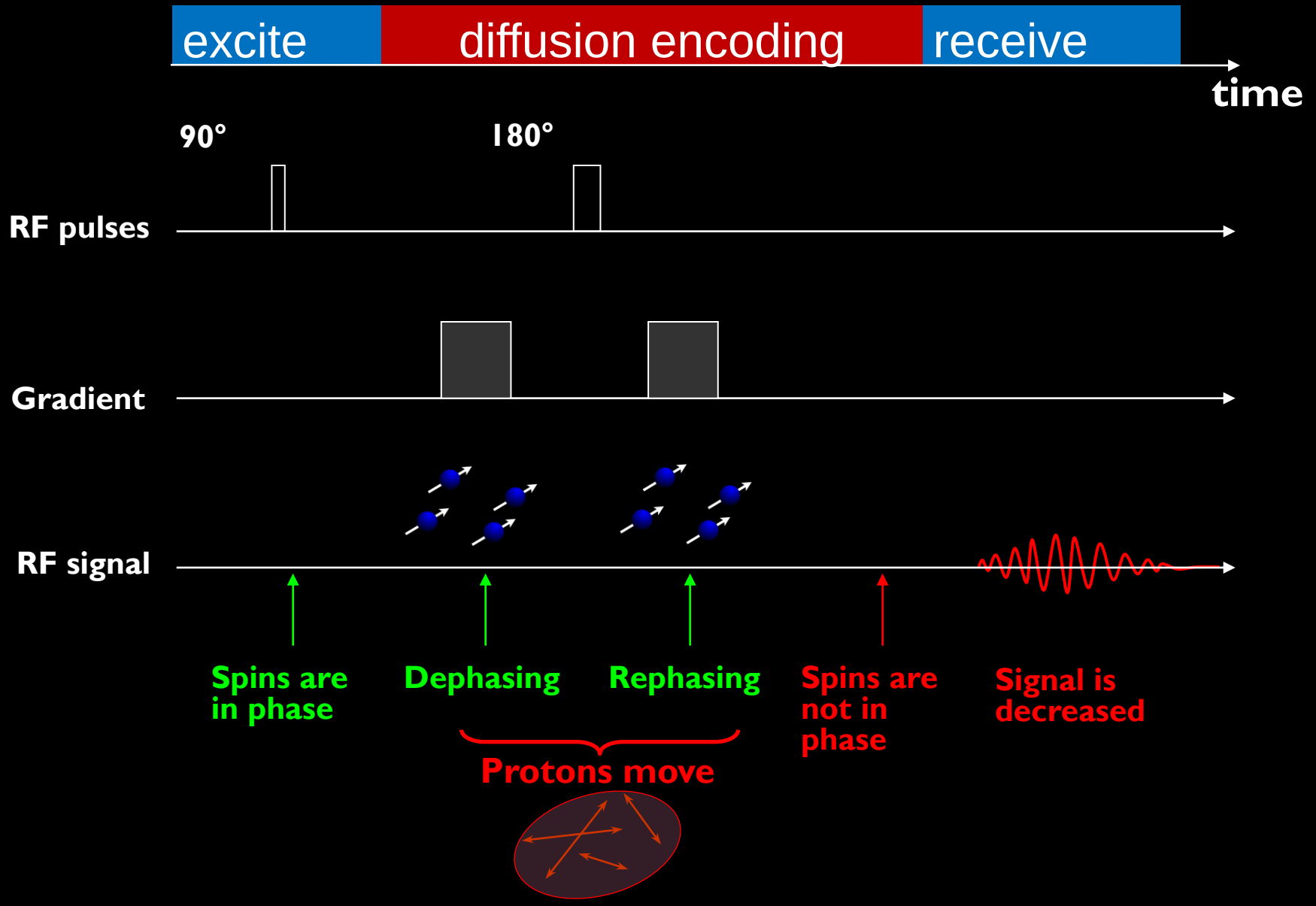
Outline

- Problems that were addressed, why they are worth solving
- Contribution to the field
- In particular,
  - Joint reconstruction of similar images
  - Accelerated Diffusion Spectrum Imaging
  - Quantifying tissue iron concentration
  - Lipid artifact suppression for Spectroscopic Imaging
    - Postpone to closed session

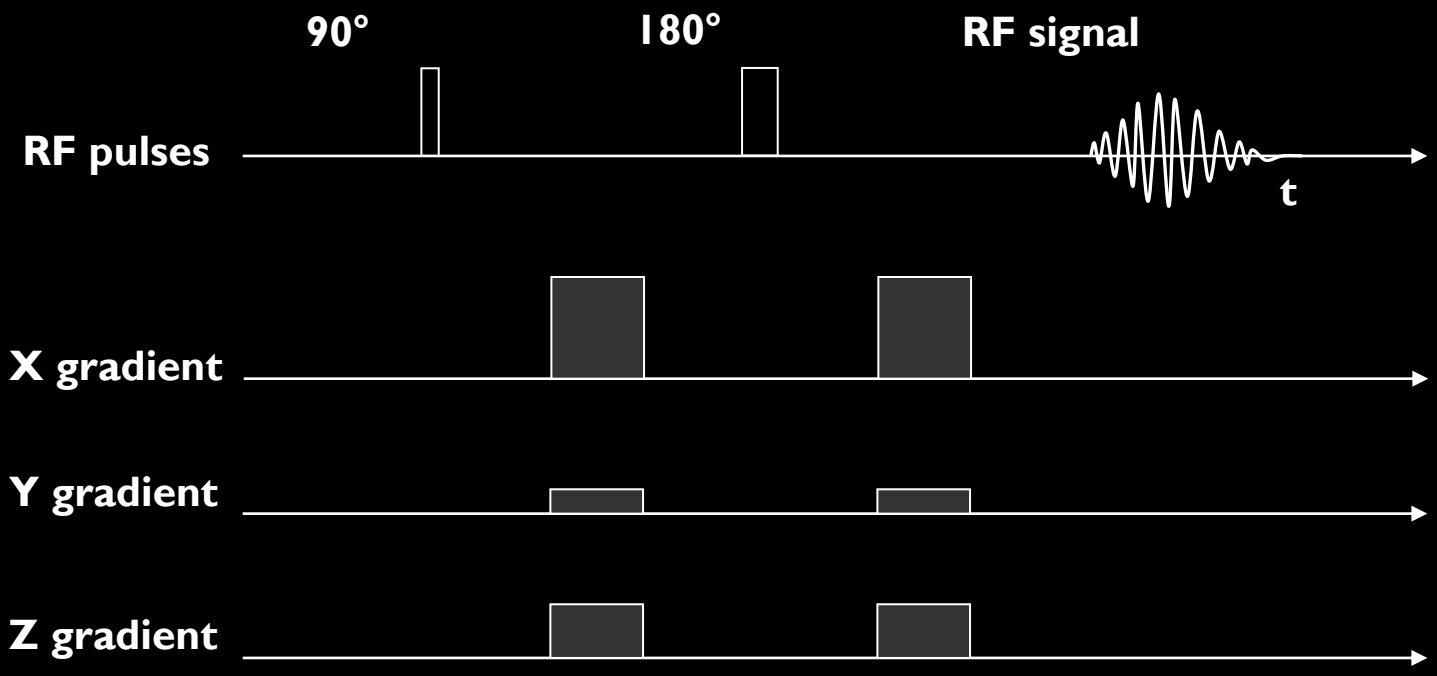
Diffusion imaging



Diffusion imaging – moving water molucules

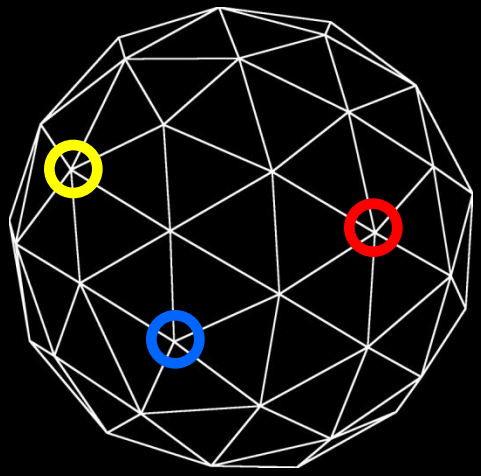


Direction of diffusion

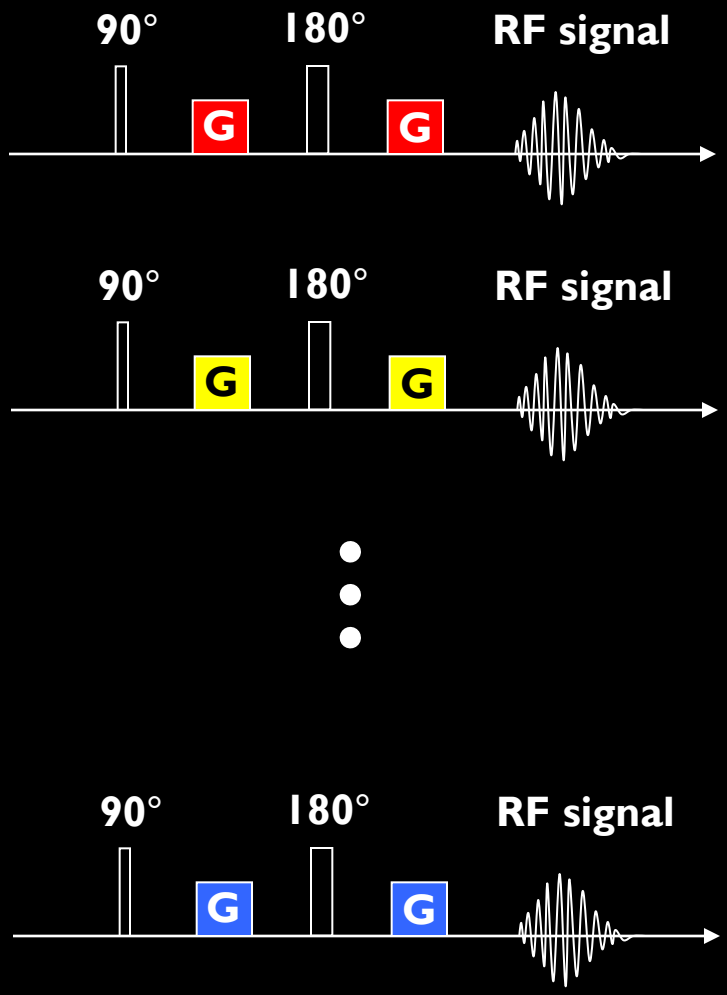


- Weight the diffusion in the desired direction of space using magnetic gradients in 3-D

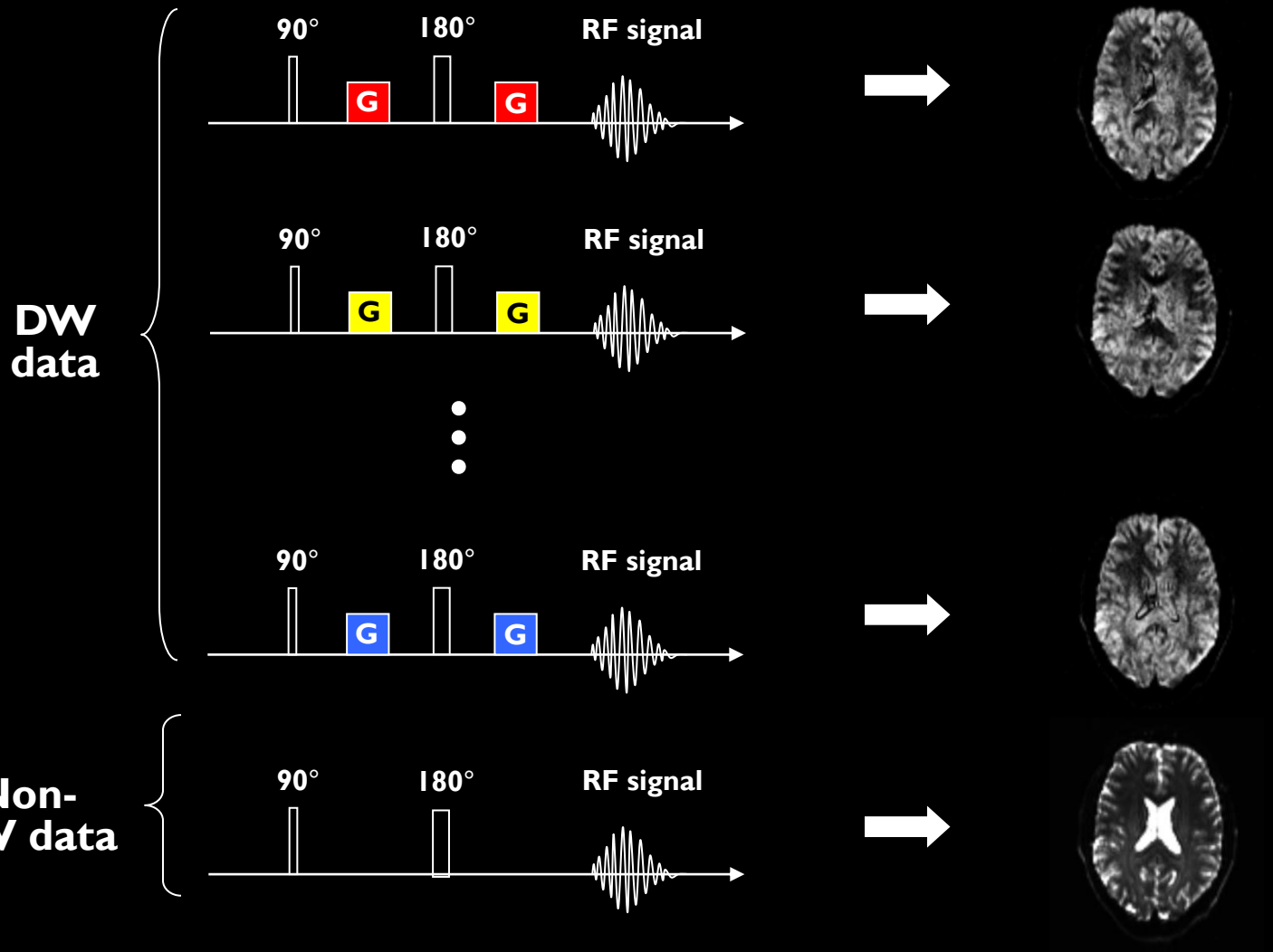
Diffusion imaging



q-space



Diffusion imaging

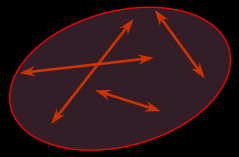


- Image intensity attenuation is dependent on water diffusion in each direction

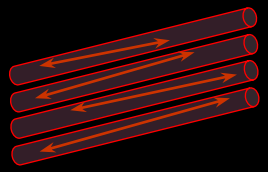
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

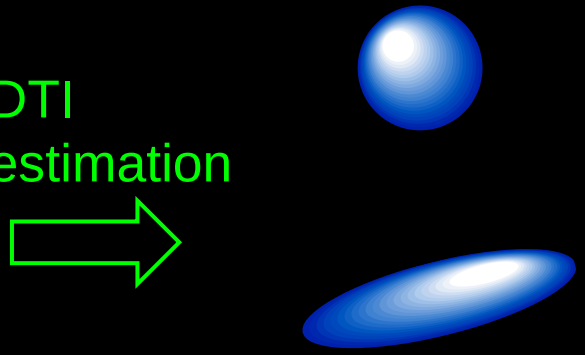
Isotropic tissue



Fibrous tissue



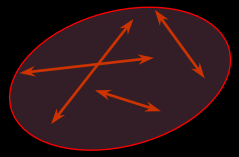
DTI estimation



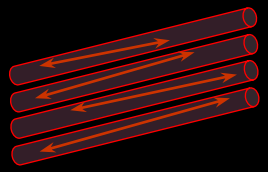
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

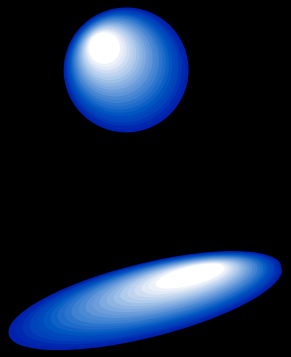
Isotropic tissue



Fibrous tissue



DTI estimation



- Tensor representation:

$$prob(\text{move to } \boldsymbol{r} \text{ in time } \boldsymbol{\Delta}) \propto \exp\left(-\frac{r^T \boldsymbol{D}^{-1} r}{4\boldsymbol{\Delta}}\right)$$

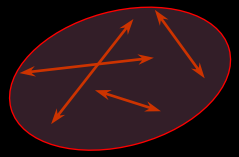
$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}$$

- r ~ 10 μm << 1 mm (voxel size)

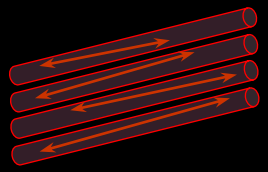
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

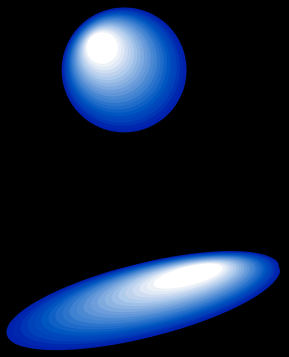
Isotropic tissue



Fibrous tissue



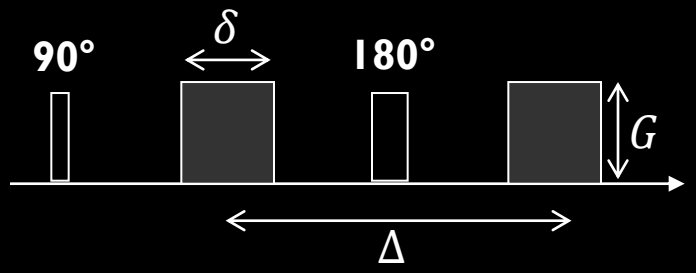
DTI estimation



- Tensor representation:

$$prob(\boldsymbol{r}, \Delta) \propto \exp\left(-\frac{\boldsymbol{r}^T \boldsymbol{D}^{-1} \boldsymbol{r}}{4\Delta}\right)$$

- r ~ 10 μm



MR signal detected:

$$\mathbf{S}(\boldsymbol{g}) = \mathbf{S}(0) \cdot \exp(-b \cdot \hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}})$$

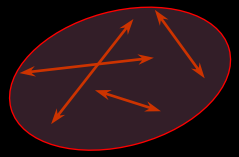
$b \propto G^2 \delta^2 (\Delta - \delta/3)$

$\hat{\boldsymbol{g}}$: unit vector along $\boldsymbol{g}$

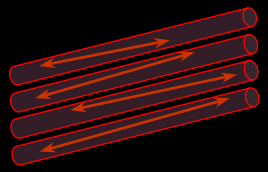
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

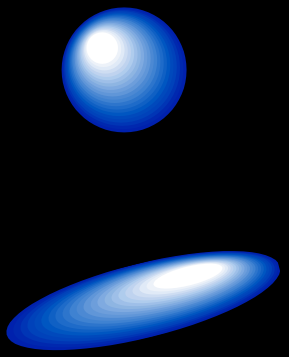
Isotropic tissue



Fibrous tissue



DTI estimation



- Tensor estimation:

$$\mathbf{S}(\boldsymbol{g}) = \mathbf{S}(0) \cdot \exp(-b \cdot \hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}})$$

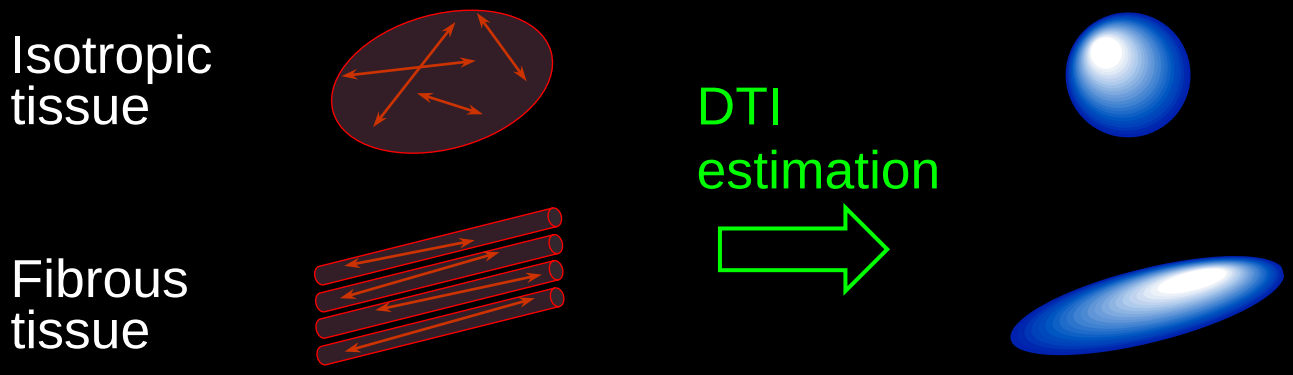
$$\hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}} = \frac{1}{b} \cdot \ln\left(\frac{\mathbf{S}(0)}{\mathbf{S}(\boldsymbol{g})}\right)$$

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

6 unknowns

Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:



- Tensor estimation:

$$\mathbf{S}(\boldsymbol{g}) = \mathbf{S}(0) \cdot \exp(-b \cdot \hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}})$$

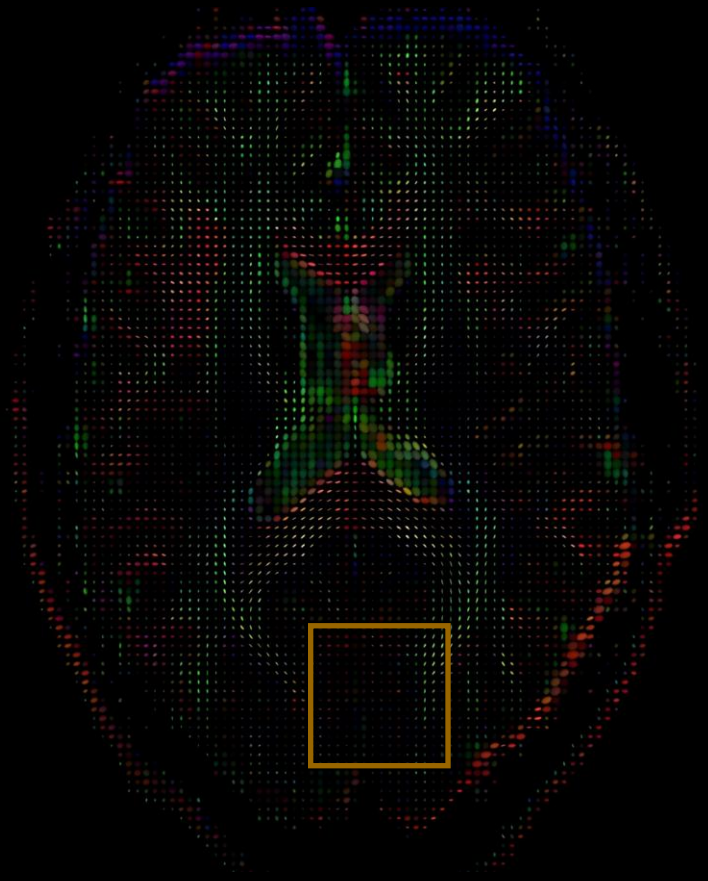
$$\hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}} = \frac{1}{b} \cdot \ln\left(\frac{\mathbf{S}(0)}{\mathbf{S}(\boldsymbol{g})}\right)$$

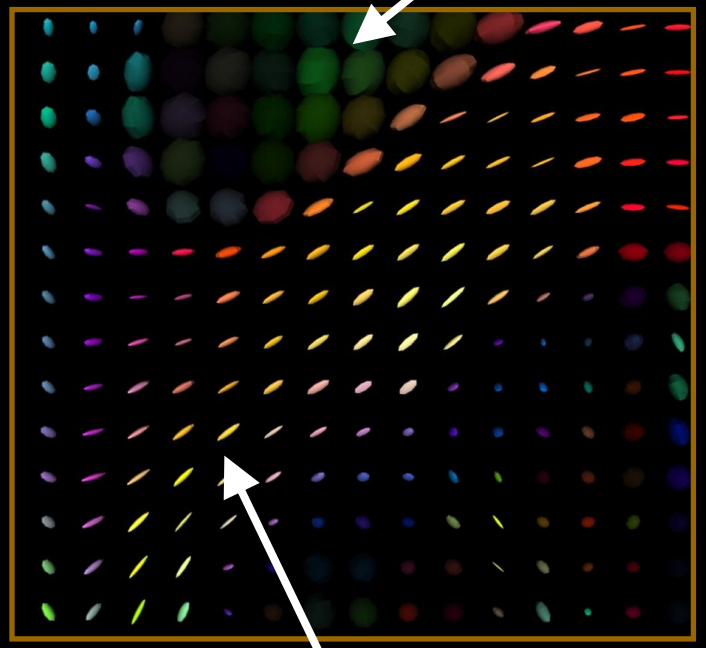
$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}$$

6 unknowns

- At least 6 DWI + 1 non-DWI acquisitions are needed for DTI

Tensor visualization

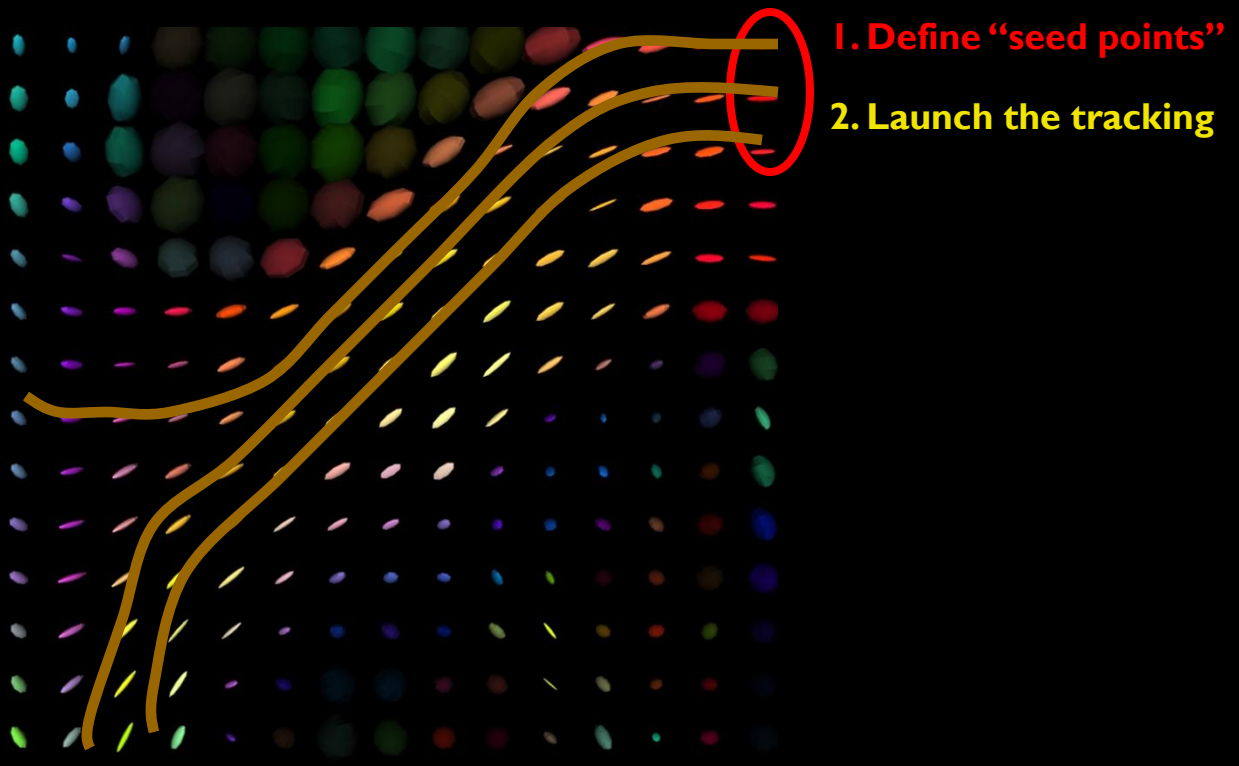




CSF: isotropic

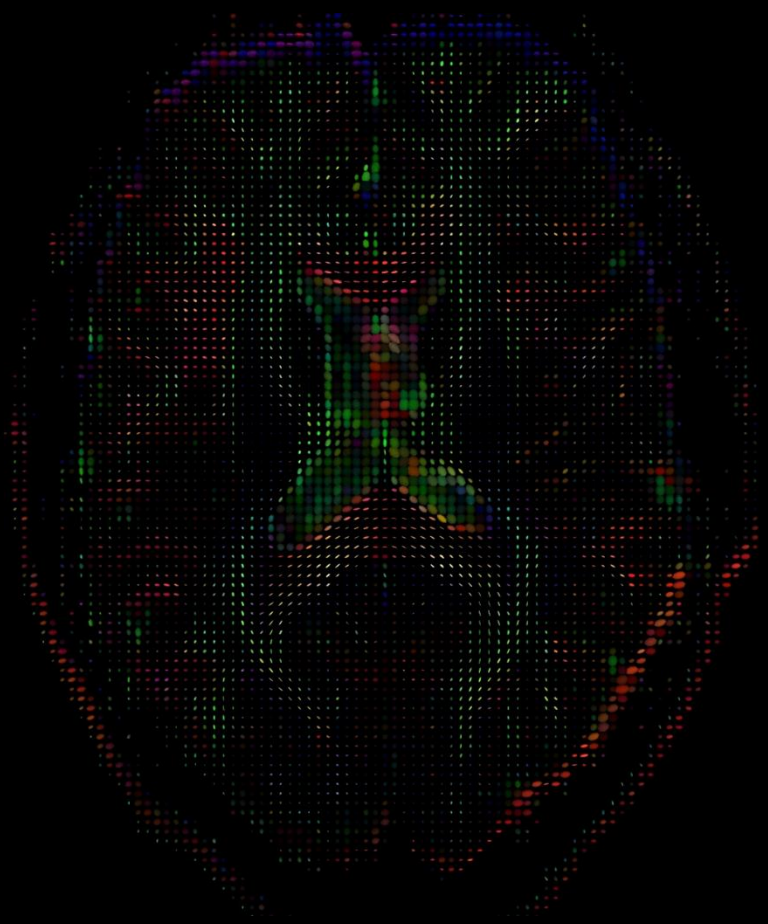
White matter: anisotropic

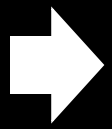
Fiber tracking

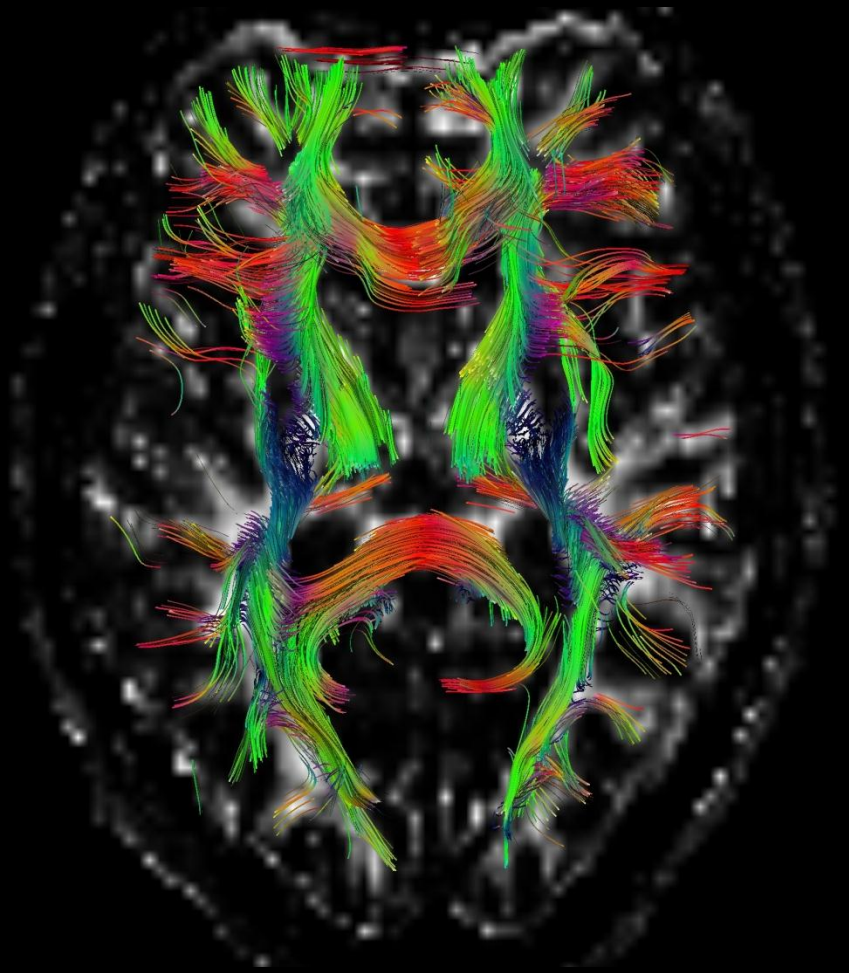


- Connect similar directions
- Variety of software is available

Fiber Tractography







Tensors

Tracts

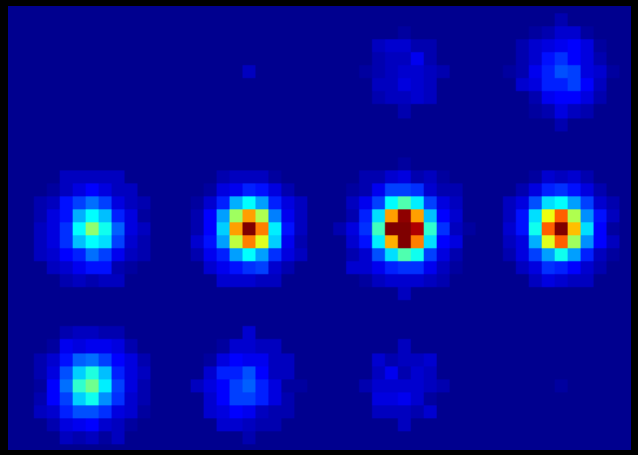
Diffusion Spectrum Imaging (DSI)

- Unlike tensor modeling, DSI offers a complete description of water diffusion
- And reveals complex distributions of fiber orientations
- DSI requires full sampling of q-space (DTI needs ≥7 points)

Diffusion Spectrum Imaging (DSI)

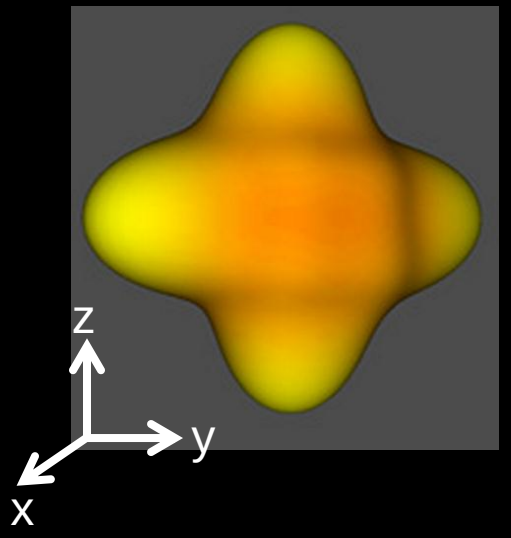
- Unlike tensor modeling, DSI offers a complete description of water diffusion
- And reveals complex distributions of fiber orientations
- DSI requires full sampling of q-space (DTI needs ≥7 points)

Q-space of a single voxel
515 directions



DFT

Probability Density Function (pdf)
of a single voxel



Sampling full q-space takes ~1 hour

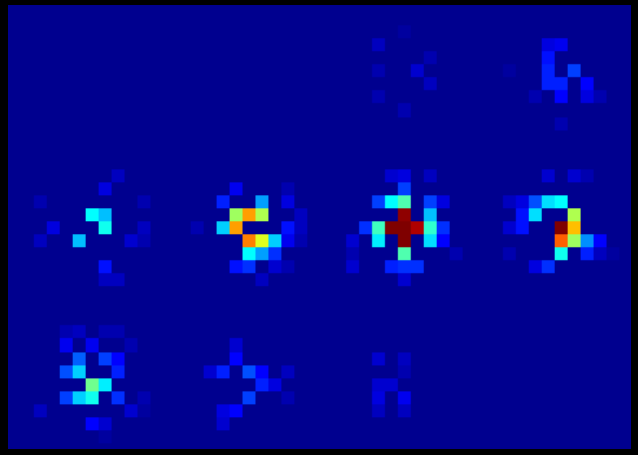
Undersampled DSI

- To reduce scan time, undersample q-space
- Use sparsity prior to reconstruct the pdfs [1]

$$min_{\boldsymbol{p}} \|\mathbf{F}_{\Omega}\boldsymbol{p} - \boldsymbol{q}\|_2{}^2 + \alpha \cdot \|\boldsymbol{\Psi}\boldsymbol{p}\|_1 + \beta \cdot \mathrm{TV}(\boldsymbol{p})$$

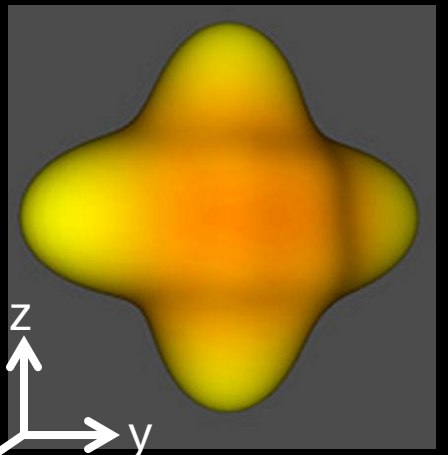
undersampled DFT — pdf — q-samples — wavelet — total variation

Undersampled q-space of a single voxel



CS

Probability Density Function (pdf) of a single voxel



1. Menzel MI et al MRM 2011

K-SVD algorithm for DSI

- Is pdf sparse in TV and wavelet?
- Use a transform tailored for sparse representation of pdfs

Step1: Create dictionary from a training pdf dataset [P]

$$min_{\mathbf{P},\mathbf{D}} \sum_i \|\boldsymbol{x}_i\|_0 \quad \text{subject to} \quad \|\mathbf{P} - \mathbf{D}\mathbf{X}\|_F{}^2 \leq \epsilon$$

K-SVD[1] iterative algorithm was used to obtain [D]

Step2: Use dictionary to impose sparsity constraint

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

FOCUSS[2] was used to provide parameter free recon

1. Aharon M, et al IEEE Trans Signal Processing 2006
2. Gorodnitsky IF, et al IEEE Trans Signal processing 1997

Methods

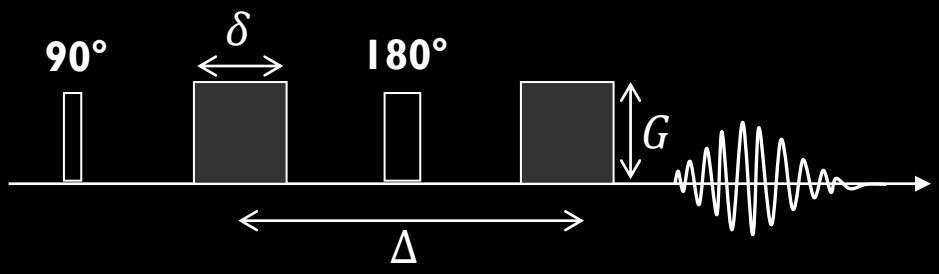
- 3 healthy volunteers, 3T Siemens Skyra

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients†, 64-chan head coil [1]

Gmax = 300 mT / m

Conventional = 45 mT / m



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

† MAGNETOM Skyra CONNECTOM system (Siemens Healthcare)

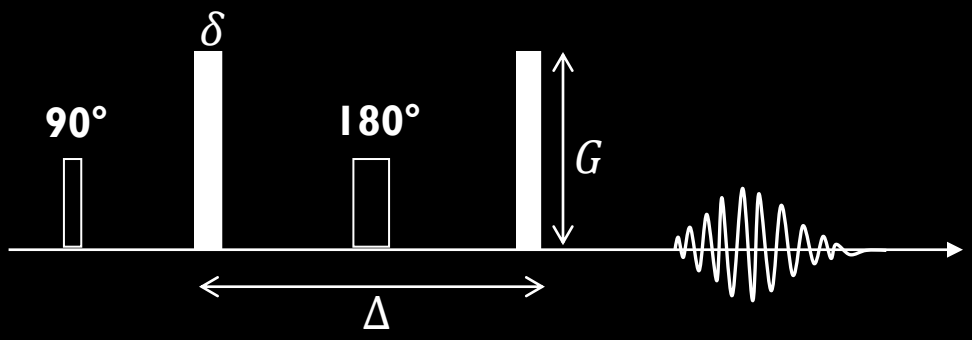
1. Keil B, et al MRM 2012

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]

Gmax = 300 mT / m

Conventional = 45 mT / m



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

At fixed b, larger $G \rightarrow$ shorter δ

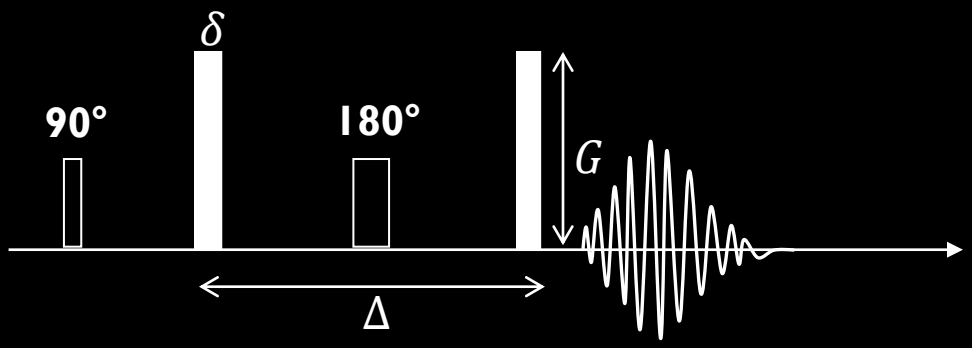
1. Keil B, et al MRM 2012

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]

Gmax = 300 mT / m

Conventional = 45 mT / m



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

At fixed b, larger $G \rightarrow$ shorter δ

Shorter echo time, higher signal

1. Keil B, et al MRM 2012

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2

1. Keil B, et al MRM 2012

Methods

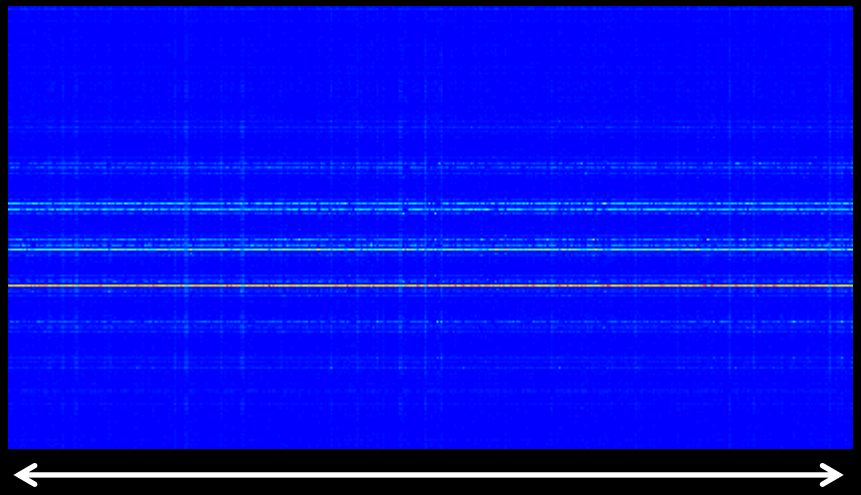
- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

1. Keil B, et al MRM 2012

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- One dictionary trained with data from each subject



12×12×12 = pdf grid size = 1728 rows

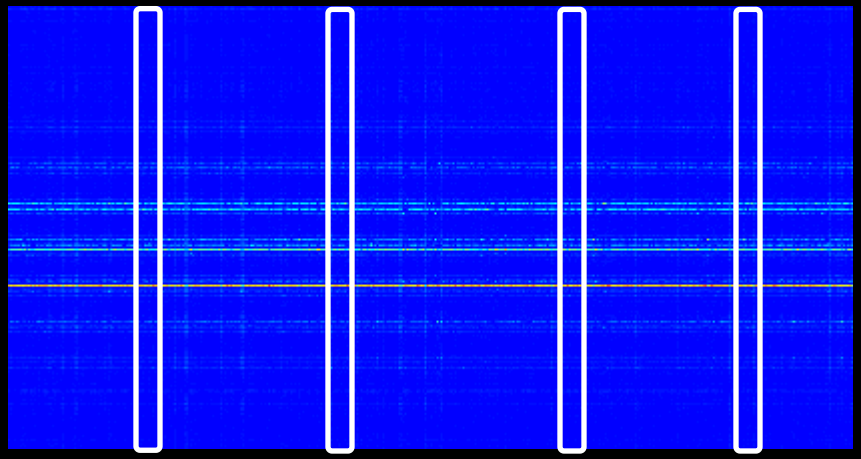
3172 columns

1. Keil B, et al MRM 2012

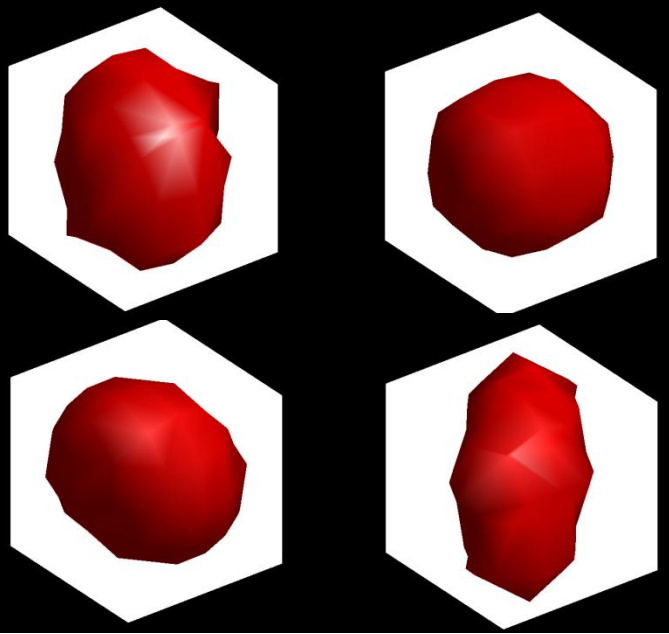
Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- One dictionary trained with data from each subject



particular columns →



Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- One dictionary trained with data from each subject
- Recon experiments at accelerations R = 3, 5 and 9

1. Keil B, et al MRM 2012

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- One dictionary trained with data from each subject
- Recon experiments at accelerations R = 3, 5 and 9

- <u>Comparison of methods:</u>
  i. Wavelet + TV (Menzel et al [2])
  ii. L1-FOCUSS (apply L1 penalty on pdfs)
  iii. Dictionary-FOCUSS (proposed)

1. Keil B, et al MRM 2012
2. Menzel MI et al MRM 2011

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- 10 average collected at 5 q-space points

  Low-noise data, serve as ground truth

1. Keil B, et al MRM 2012
2. Menzel MI et al MRM 2011

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm^2
- 515 q-space points, 50 min scan time
- Number of voxels = 96×96×57 ≈ **500.000**

- 10 average collected at 5 q-space points
  Low-noise data, serve as ground truth

- <u>Tractography comparison:</u>
  - Fully-sampled vs. R = 3 Dictionary-FOCUSS
  - Fractional Anisotropy compared for 18 major fiber bundles

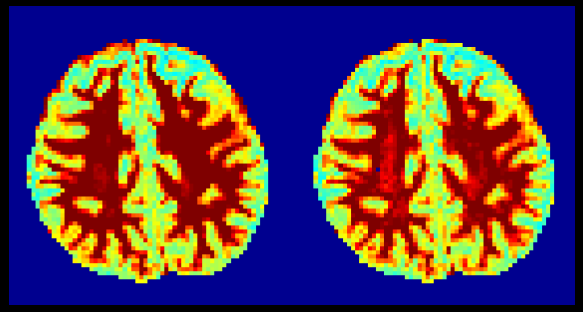
1. Keil B, et al MRM 2012
2. Menzel MI et al MRM 2011

Subject A, pdf reconstruction error

Slice 40

Wavelet+TV

ℓ_1-FOCUSS



Acceleration
R = 3

15.8% RMSE

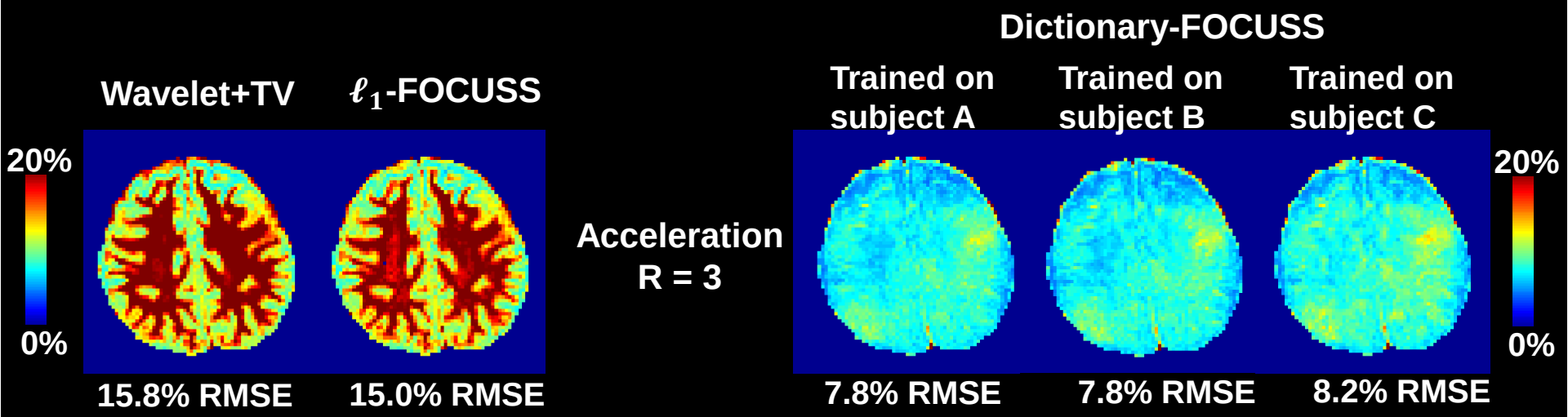
15.0% RMSE

Wav+TV @ R=3 **15.8% error**

ℓ_1-FOCUSS @ R=3 **15.0% error**

Subject A, pdf reconstruction error

Slice 40



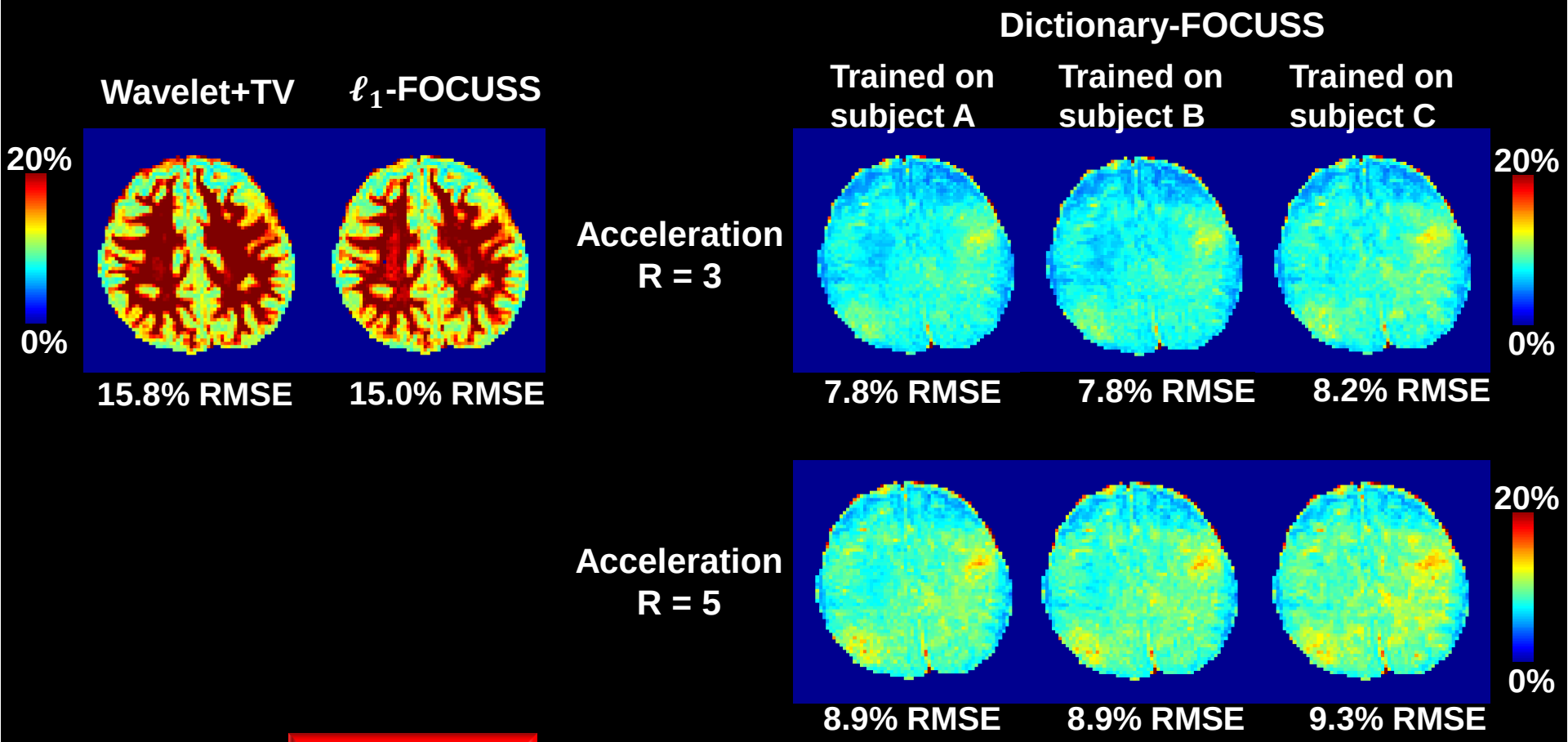
Wav+TV @ R=3 — 15.8% error

ℓ_1-FOCUSS @ R=3 — 15.0% error

Dictionary @ R=3 — 7.8% error

Subject A, pdf reconstruction error

Slice 40



Wav+TV @ R=3 — 15.8% error

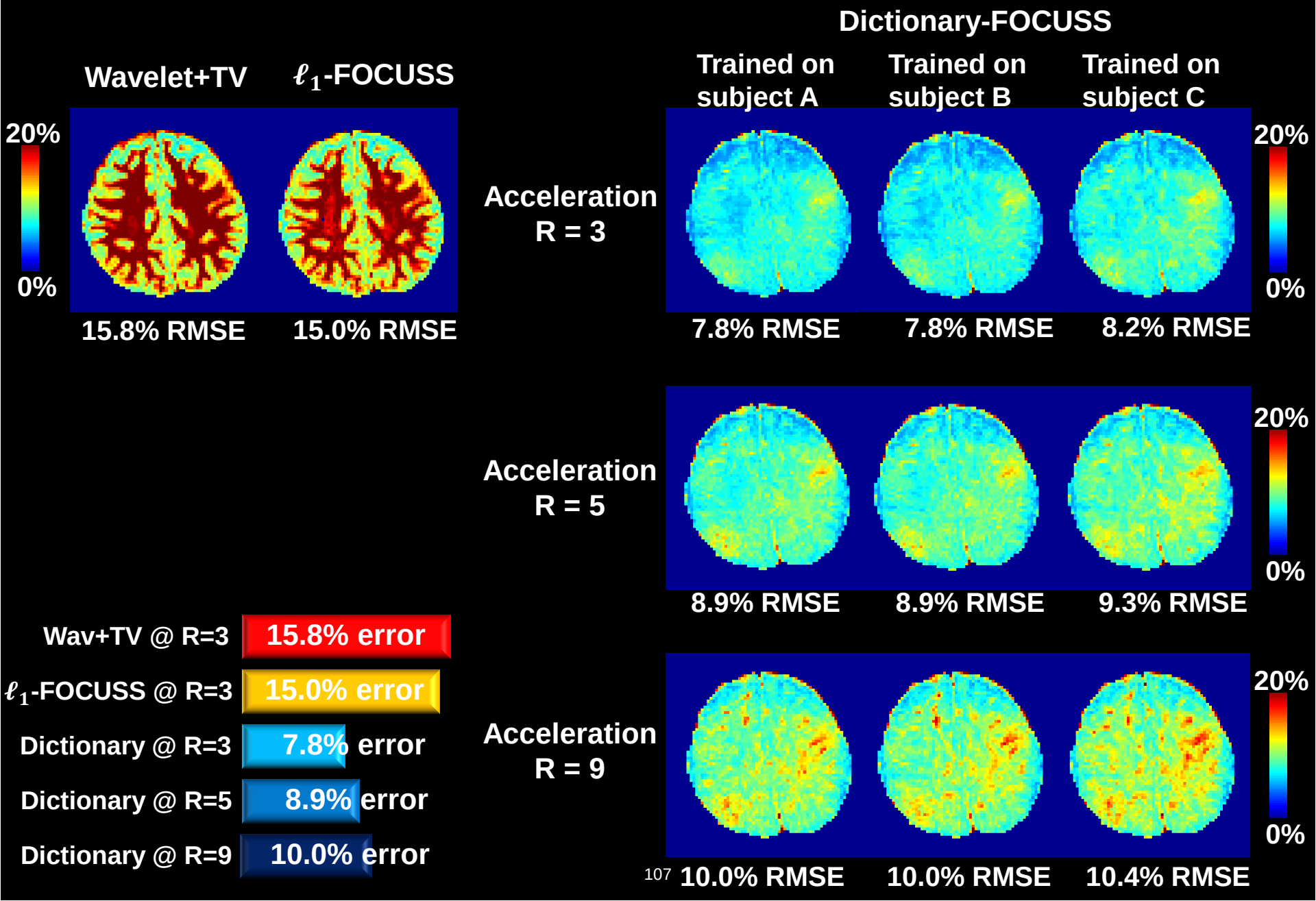
ℓ_1-FOCUSS @ R=3 — 15.0% error

Dictionary @ R=3 — 7.8% error

Dictionary @ R=5 — 8.9% error

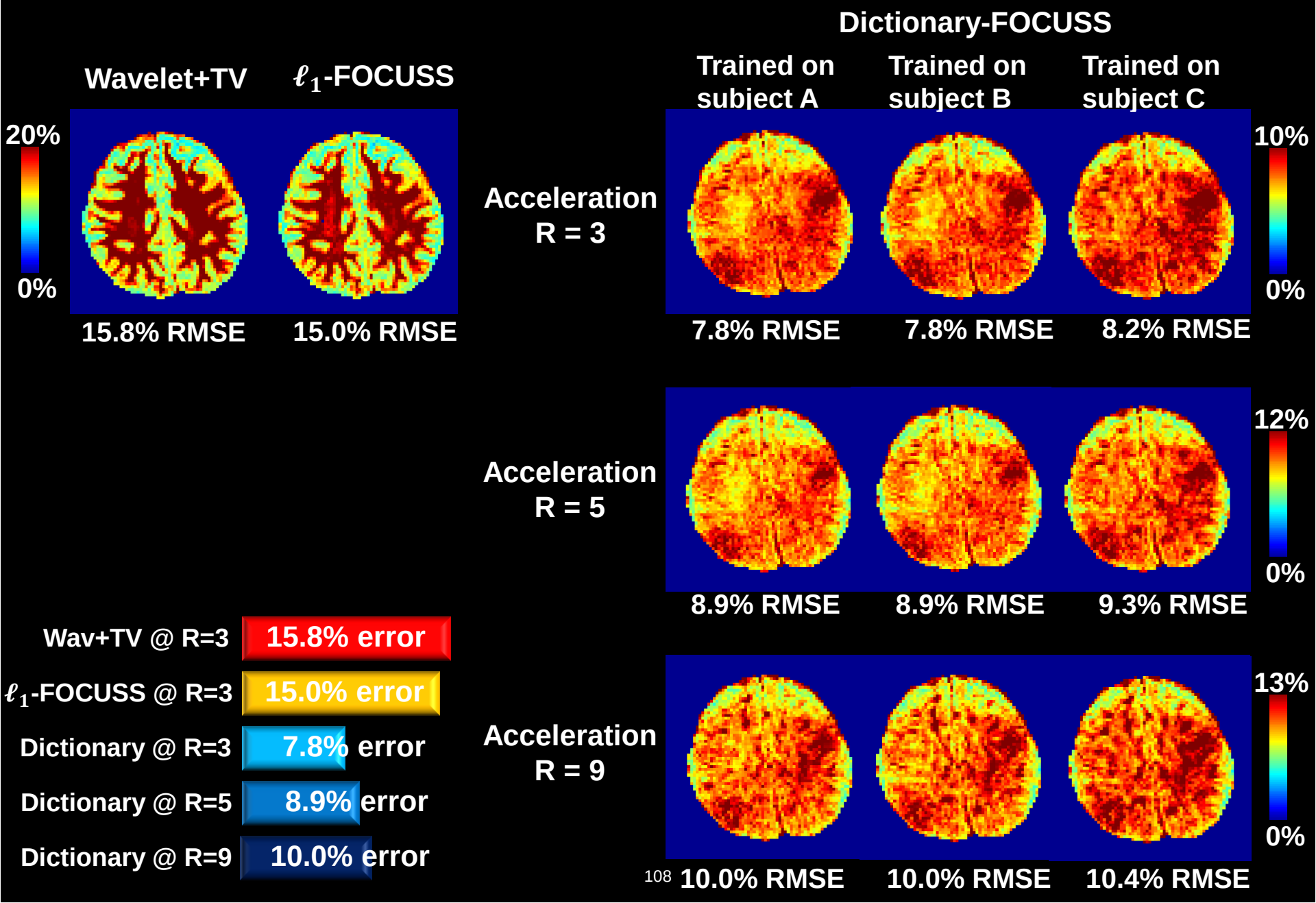
Subject A, pdf reconstruction error

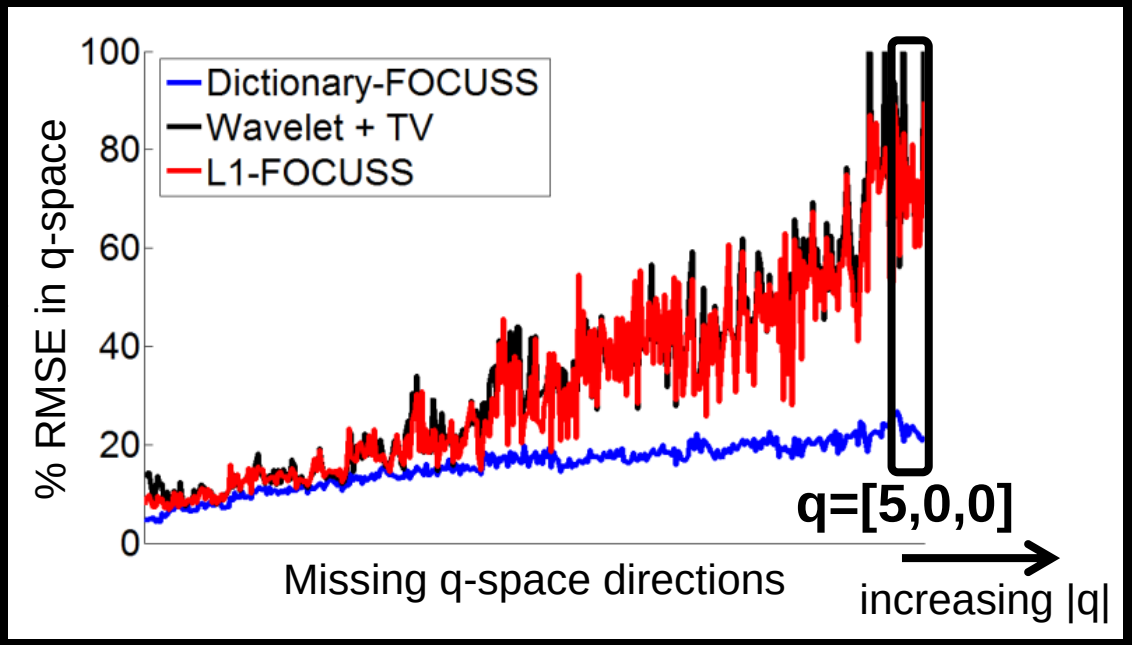
Slice 40



Subject A, pdf reconstruction error

Slice 40





q-space reconstructions at q=[5,0,0]

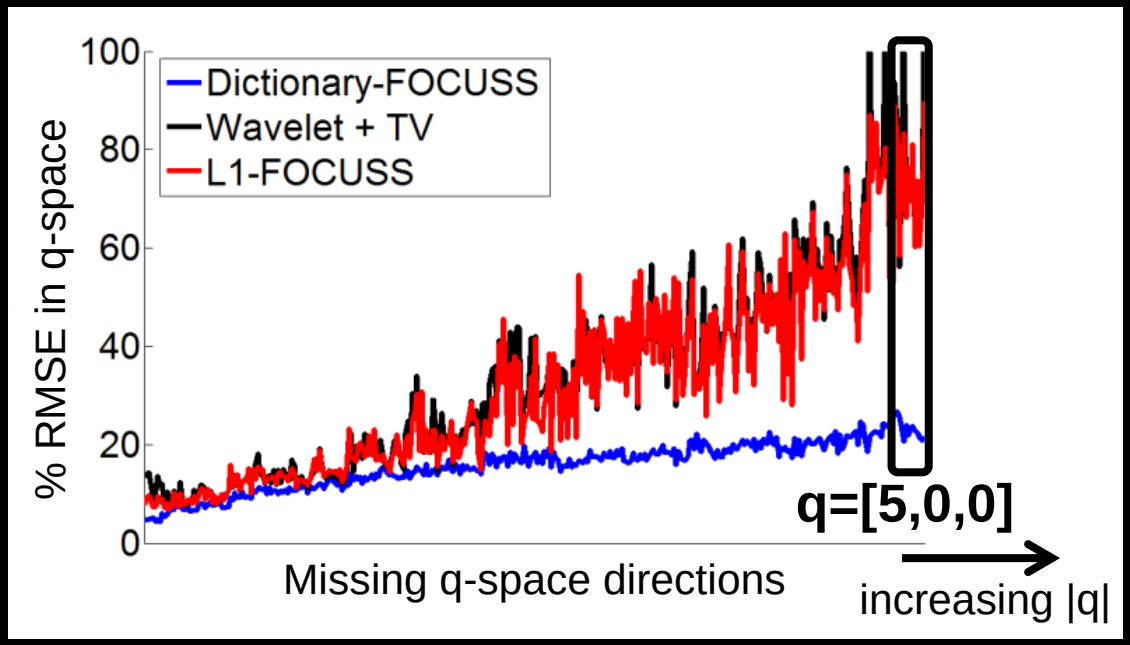
Wavelet+TV

ℓ_1-FOCUSS

Dict-FOCUSS

Fully-sampled

1 average



q-space reconstructions at q=[5,0,0]

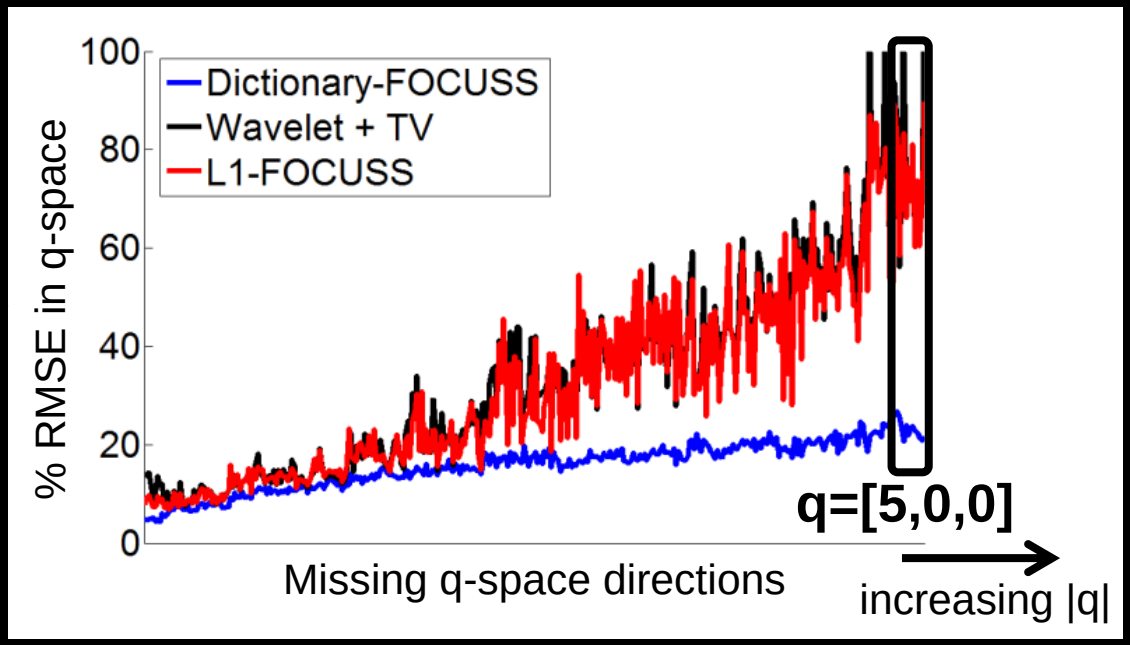
Wavelet+TV

ℓ_1-FOCUSS

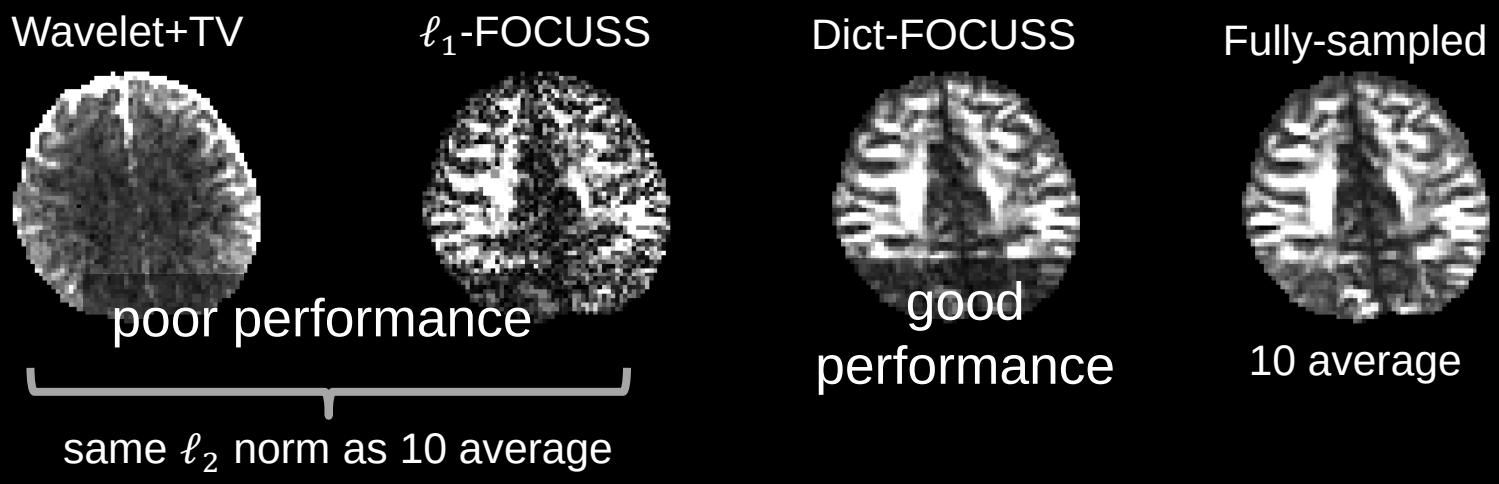
Dict-FOCUSS

Fully-sampled

10 average



q-space reconstructions at q=[5,0,0]



- SNR drops substantially at the outer q-space
- RMSE computed relative to 1 average fully-sampled data includes noise and recon error

- To isolate recon error, collected 10 avg on 5 q-space points

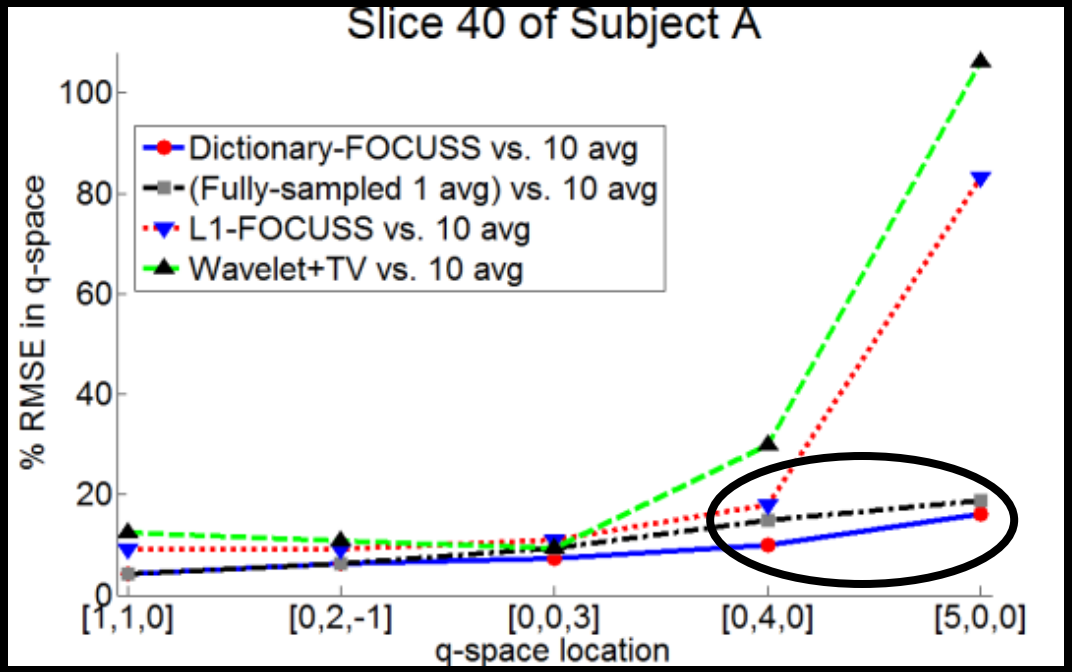
- SNR drops substantially at the outer q-space
- RMSE computed relative to 1 average fully-sampled data includes noise and recon error

1 avg fully-sampled

10 avg fully-sampled

q = [5,0,0]

- SNR drops substantially at the outer q-space
- RMSE computed relative to 1 average fully-sampled data includes noise and recon error

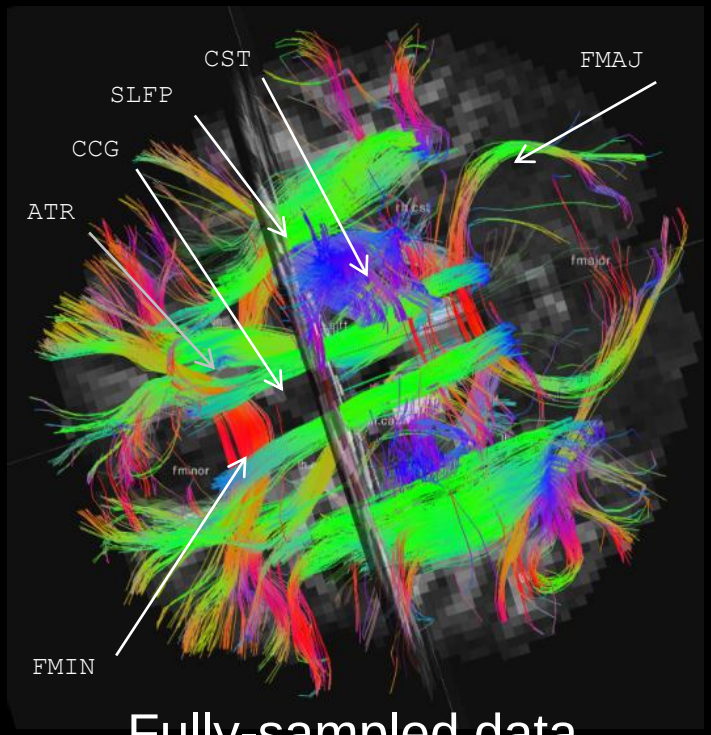


Lower RMSE than acquired data

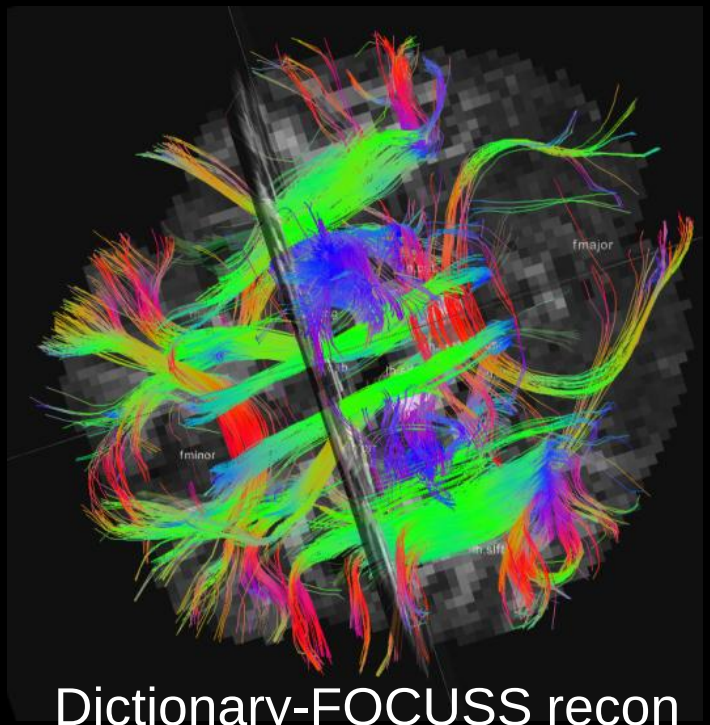
Denoising effect [1]

1. Patel V, et al ISBI 2011, p1805

Tractography solutions for subject A

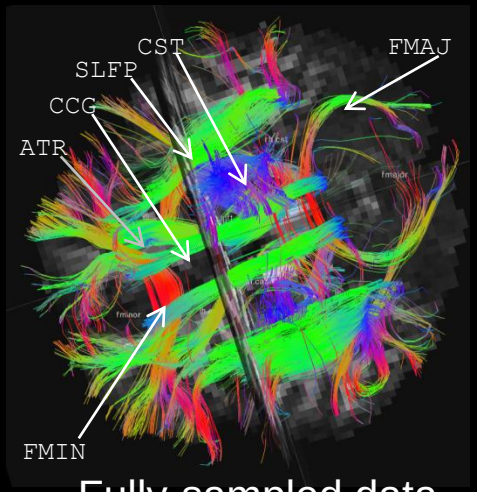


Fully-sampled data

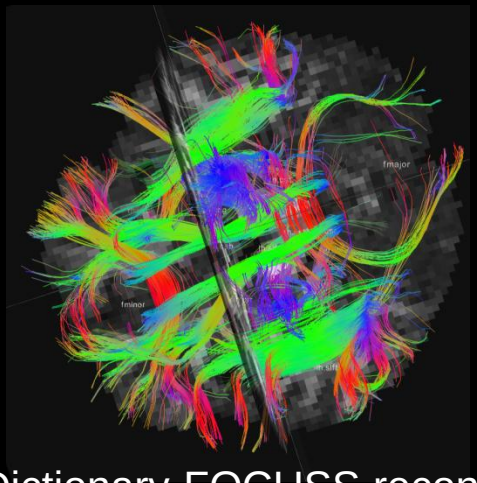


Dictionary-FOCUSS recon with 3-fold acceleration

Tractography solutions for subject A

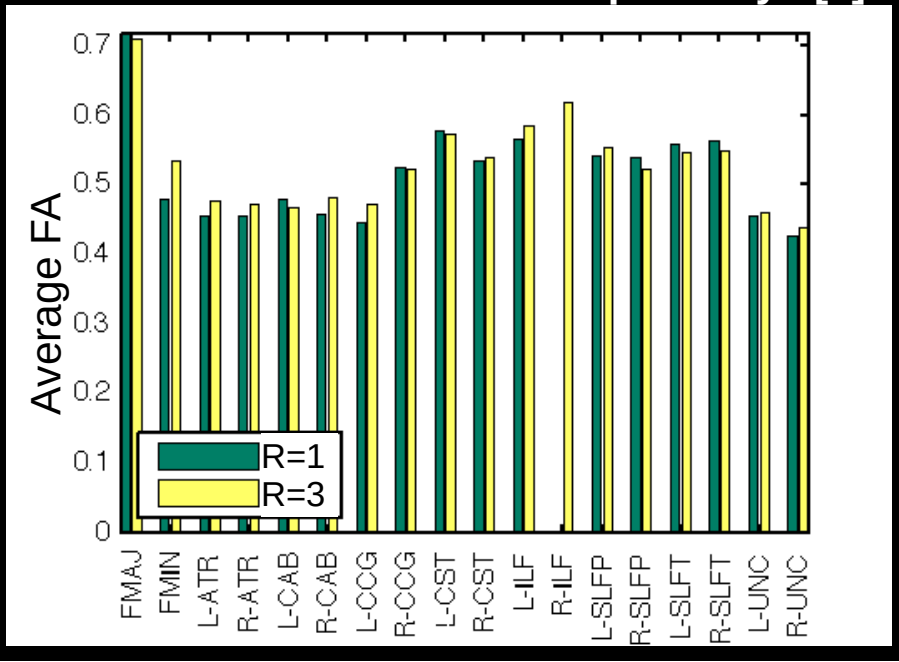


Fully-sampled data



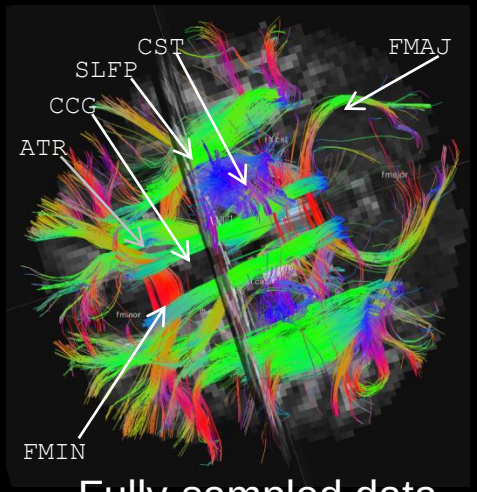
Dictionary-FOCUSS recon
with 3-fold acceleration

Average Fractional Anisotropy for 18 labeled white-matter pathways [1]

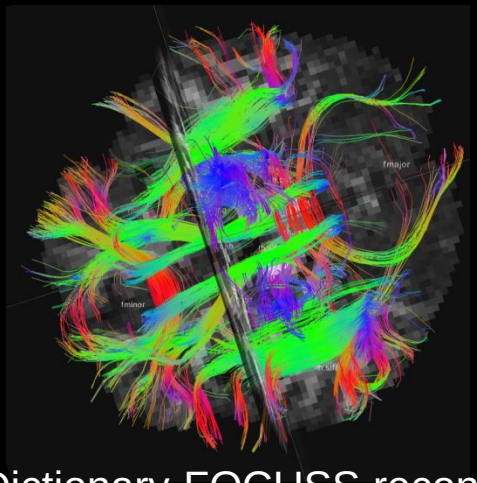


1. Yendiki A et al
Front Neuroinform 2011

Tractography solutions for subject A

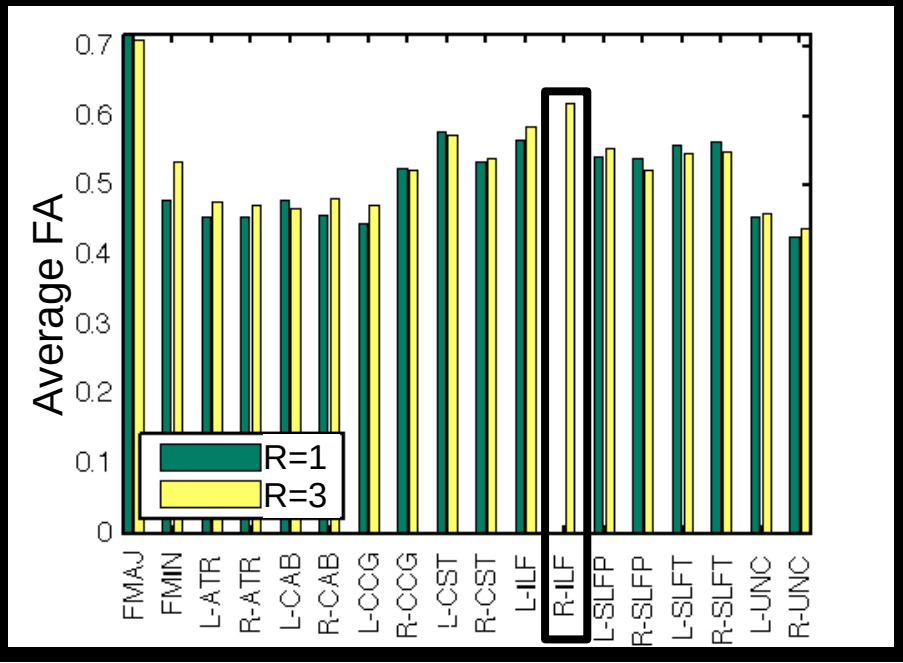


Fully-sampled data



Dictionary-FOCUSS recon
with 3-fold acceleration

Mean FA error = 3%



1. Yendiki A et al
Front Neuroinform 2011

Concluding Remarks

- Up to 2-times RMSE reduction in pdf domain
  - Dictionary-FOCUSS (proposed) vs. Wavelet+TV [1]

1. Menzel MI et al MRM 2011

Concluding Remarks

- Up to 2-times RMSE reduction in pdf domain
  - Dictionary-FOCUSS (proposed) vs. Wavelet+TV [1]
- 3-fold accelerated Dict-FOCUSS ≈ Fully-sampled data
  - Low-noise 10 average data validation
  - Tractography comparison

1. Menzel MI et al MRM 2011

Concluding Remarks

- Up to 2-times RMSE reduction in pdf domain
  - Dictionary-FOCUSS (proposed) vs. Wavelet+TV [1]
- 3-fold accelerated Dict-FOCUSS ≈ Fully-sampled data
- Dictionary from single slice seems to generalizes to other slices and to other subjects

1. Menzel MI et al MRM 2011

Concluding Remarks

- **Voxel-by-voxel recon**
  - Dictionary-FOCUSS: 12 sec / voxel
  - Wavelet+TV: 27 sec / voxel in Matlab

Concluding Remarks

- **Voxel-by-voxel recon**
  - Dictionary-FOCUSS: 12 sec / voxel
  - Wavelet+TV: 27 sec / voxel in Matlab

- Full-brain processing for 10^5 voxels: DAYS of computation

Addressed next

Concluding Remarks

- **Voxel-by-voxel recon**
  - Dictionary-FOCUSS: 12 sec / voxel
  - Wavelet+TV: 27 sec / voxel in Matlab
- Full-brain processing for 10^5 voxels: DAYS of computation
- Do dictionaries generalize across healthy vs. patient populations? across different age groups?

Fast DSI Reconstruction

- Two proposals that are computationally 1000-fold faster with image quality similar to Dictionary-FOCUSS:

Fast DSI Reconstruction

- Two proposals that are computationally 1000-fold faster with image quality similar to Dictionary-FOCUSS:

i. **PINV:**
   - Uses a dictionary trained with K-SVD
   - Rather than ℓ_1, applies ℓ_2 regularization to dictionary coefficients
   - Admits closed-form solution (Regularized Pseudoinverse (PINV))

Fast DSI Reconstruction

- Two proposals that are computationally 1000-fold faster with image quality similar to Dictionary-FOCUSS:

i. **PINV:**
   - Uses a dictionary trained with K-SVD
   - Rather than ℓ_1, applies ℓ_2 regularization to dictionary coefficients
   - Admits closed-form solution (Regularized Pseudoinverse (PINV))

ii. **PCA:**
   - Obtain a low-dimensional representation using training data
   - Retain maximum variance using Principal Component Analysis (PCA)
   - Admits closed-form solution, no need for K-SVD

PINV: ℓ_2 regularization

- Dictionary-FOCUSS iteratively solves

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

PINV: ℓ_2 regularization

- Dictionary-FOCUSS iteratively solves

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

- Instead, consider

$$min\ \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \lambda \cdot \|\boldsymbol{x}\|_2{}^2$$

PINV: ℓ_2 regularization

- Dictionary-FOCUSS iteratively solves

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

- Instead, consider

$$min\ \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \lambda \cdot \|\boldsymbol{x}\|_2{}^2$$

- Solution: $\tilde{\boldsymbol{x}} = ((\mathbf{F}_\Omega \mathbf{D})^H \mathbf{F}_\Omega \mathbf{D} + \lambda \mathbf{I})^{-1} (\mathbf{F}_\Omega \mathbf{D})^H q$

PINV: ℓ_2 regularization

- Dictionary-FOCUSS iteratively solves

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

- Instead, consider

$$min\ \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \lambda \cdot \|\boldsymbol{x}\|_2{}^2$$

- Solution: $\tilde{\boldsymbol{x}} = ((\mathbf{F}_\Omega \mathbf{D})^H \mathbf{F}_\Omega \mathbf{D} + \lambda \mathbf{I})^{-1} (\mathbf{F}_\Omega \mathbf{D})^H q$

Singular Value Decomposition: $\mathbf{F}_\Omega \mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$

PINV: ℓ_2 regularization

- Dictionary-FOCUSS iteratively solves

$$min\|\boldsymbol{x}\|_1 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} = \boldsymbol{q}$$

- Instead, consider

$$min\ \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \lambda \cdot \|\boldsymbol{x}\|_2{}^2$$

- Solution: $\tilde{\boldsymbol{x}} = ((\mathbf{F}_\Omega \mathbf{D})^H \mathbf{F}_\Omega \mathbf{D} + \lambda \mathbf{I})^{-1} (\mathbf{F}_\Omega \mathbf{D})^H q$

$$\tilde{\boldsymbol{x}} = \underbrace{\mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H}\, q$$

compute once

$$\mathbf{F}_\Omega \mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$$

$$\boldsymbol{\Sigma}^+ = (\boldsymbol{\Sigma}^H\boldsymbol{\Sigma} + \lambda \mathbf{I})^{-1}\boldsymbol{\Sigma}^H$$

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset

- Start with a training set of pdfs $\mathbf{P}$
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - \boldsymbol{p}_{mean}$$

$$\mathbf{Z}\mathbf{Z}^{H} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{H}$$

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset

- Start with a training set of pdfs $\mathbf{P}$
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - \boldsymbol{p}_{mean}$$

$$\mathbf{Z}\mathbf{Z}^{H} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{H}$$

- Pick the first T columns of $\mathbf{Q}$ corresponding to largest eigvals: $\mathbf{Q}_T$

$$\boldsymbol{pca} = \mathbf{Q}_T^{H}(\boldsymbol{p} - \boldsymbol{p}_{mean})$$

$\downarrow$

T - dimensional pca coefficients

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset

- Start with a training set of pdfs $\mathbf{P}$
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - \boldsymbol{p}_{mean}$$

$$\mathbf{Z}\mathbf{Z}^{H} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{H}$$

- Pick the first T columns of $\mathbf{Q}$ corresponding to largest eigvals: $\mathbf{Q}_T$

$$\boldsymbol{pca} = \mathbf{Q}_T^{H}(\boldsymbol{p} - \boldsymbol{p}_{mean})$$

- The location of $\boldsymbol{pca}$ in the pdf space,

$$\boldsymbol{p}_T = \mathbf{Q}_T\boldsymbol{pca} + \boldsymbol{p}_{mean}$$

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset
- Least-squares approximation in T - dimensions,

$$min \left\| \mathbf{F}_{\Omega} \boldsymbol{p}_{\boldsymbol{T}} - \boldsymbol{q} \right\|_2^{\,2}$$

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset
- Least-squares approximation in T - dimensions,

$$min \, \|\mathbf{F}_{\Omega}\boldsymbol{p}_{\boldsymbol{T}} - \boldsymbol{q}\|_2{}^2$$

- In PCA coordinates,

$$min_{\boldsymbol{pca}} \, \|\mathbf{F}_{\Omega}\mathbf{Q}_{T}\boldsymbol{pca} - (\boldsymbol{q} - \mathbf{F}_{\Omega}\boldsymbol{p}_{mean})\|_2^2$$

PCA Reconstruction

- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset
- Least-squares approximation in T - dimensions,

$$min \, \|\mathbf{F}_\Omega \boldsymbol{p_T} - \boldsymbol{q}\|_2{}^2$$

- In PCA coordinates,

$$min_{\boldsymbol{pca}} \, \|\mathbf{F}_\Omega \mathbf{Q}_T \boldsymbol{pca} - (\boldsymbol{q} - \mathbf{F}_\Omega \boldsymbol{p}_{mean})\|_2^2$$

- Closed-form solution:

$$\widetilde{\boldsymbol{pca}} = \underbrace{\text{pinv}(\mathbf{F}_\Omega \mathbf{Q}_T)}_{\text{compute once}}(\boldsymbol{q} - \mathbf{F}_\Omega \boldsymbol{p}_{mean})$$

Selection of regularization parameters

- PINV: selection of λ

$$min \; \|\mathbf{F}_\Omega \mathbf{D} \boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \boxed{\lambda} \cdot \|\boldsymbol{x}\|_2{}^2$$

Selection of regularization parameters

- PINV: selection of λ

$$min \, \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \boxed{\lambda} \cdot \|\boldsymbol{x}\|_2{}^2$$

- PCA: selection of PCA dimension T in $\mathbf{Q}_T$

$$min_{\boldsymbol{pca}} \, \|\mathbf{F}_\Omega \boxed{\mathbf{Q}_T} \boldsymbol{pca} - (\boldsymbol{q} - \mathbf{F}_\Omega \boldsymbol{p}_{mean})\|_2^2$$

Selection of regularization parameters

- PINV: selection of λ

$$min \, \|\mathbf{F}_\Omega \mathbf{D}\boldsymbol{x} - \boldsymbol{q}\|_2{}^2 + \boxed{\lambda} \cdot \|\boldsymbol{x}\|_2{}^2$$

- PCA: selection of PCA dimension T in $\mathbf{Q}_T$

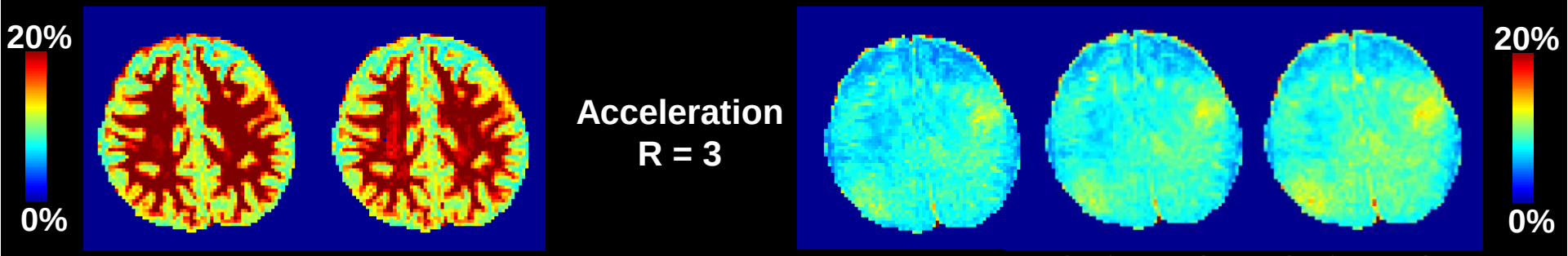
$$min_{\boldsymbol{pca}} \, \|\mathbf{F}_\Omega \boxed{\mathbf{Q}_T} \boldsymbol{pca} - (\boldsymbol{q} - \mathbf{F}_\Omega \boldsymbol{p}_{mean})\|_2^2$$

- Fully-sampled pdf training dataset $\mathbf{P}$ was used to generate the dictionary $\mathbf{D}$ and the eigenvectors $\mathbf{Q}$

- Find λ and T that yields the lowest reconstruction error on $\mathbf{P}$

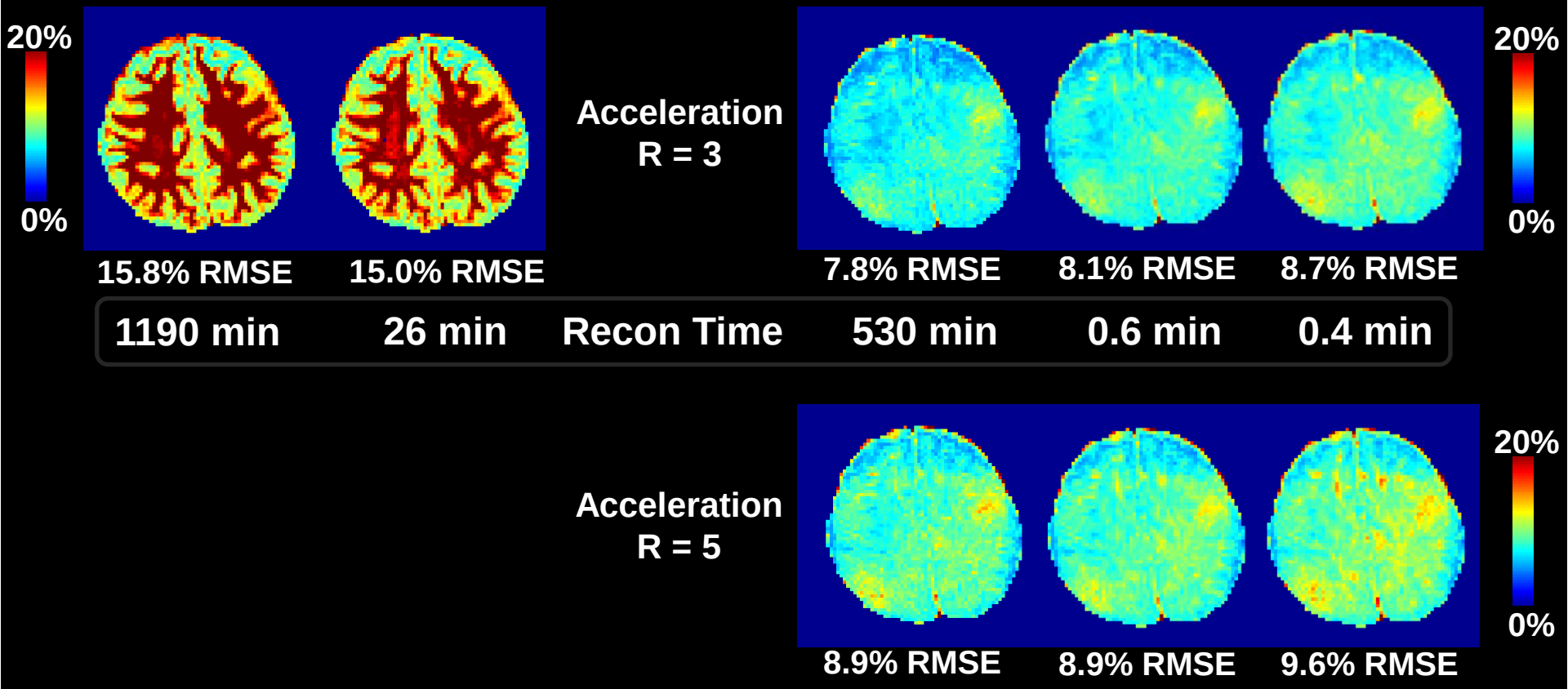
Subject A, pdf reconstruction error

Slice 40



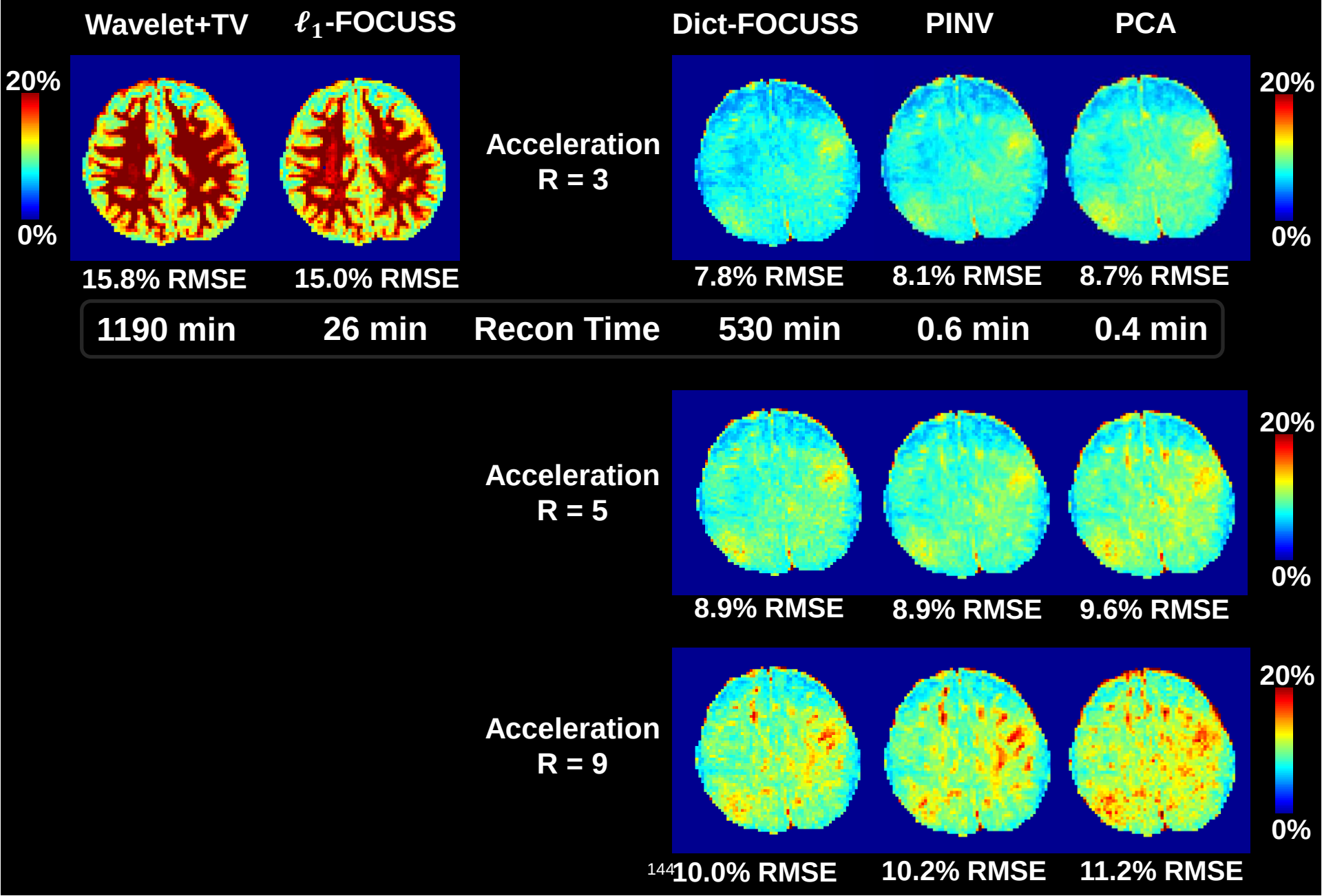
Subject A, pdf reconstruction error

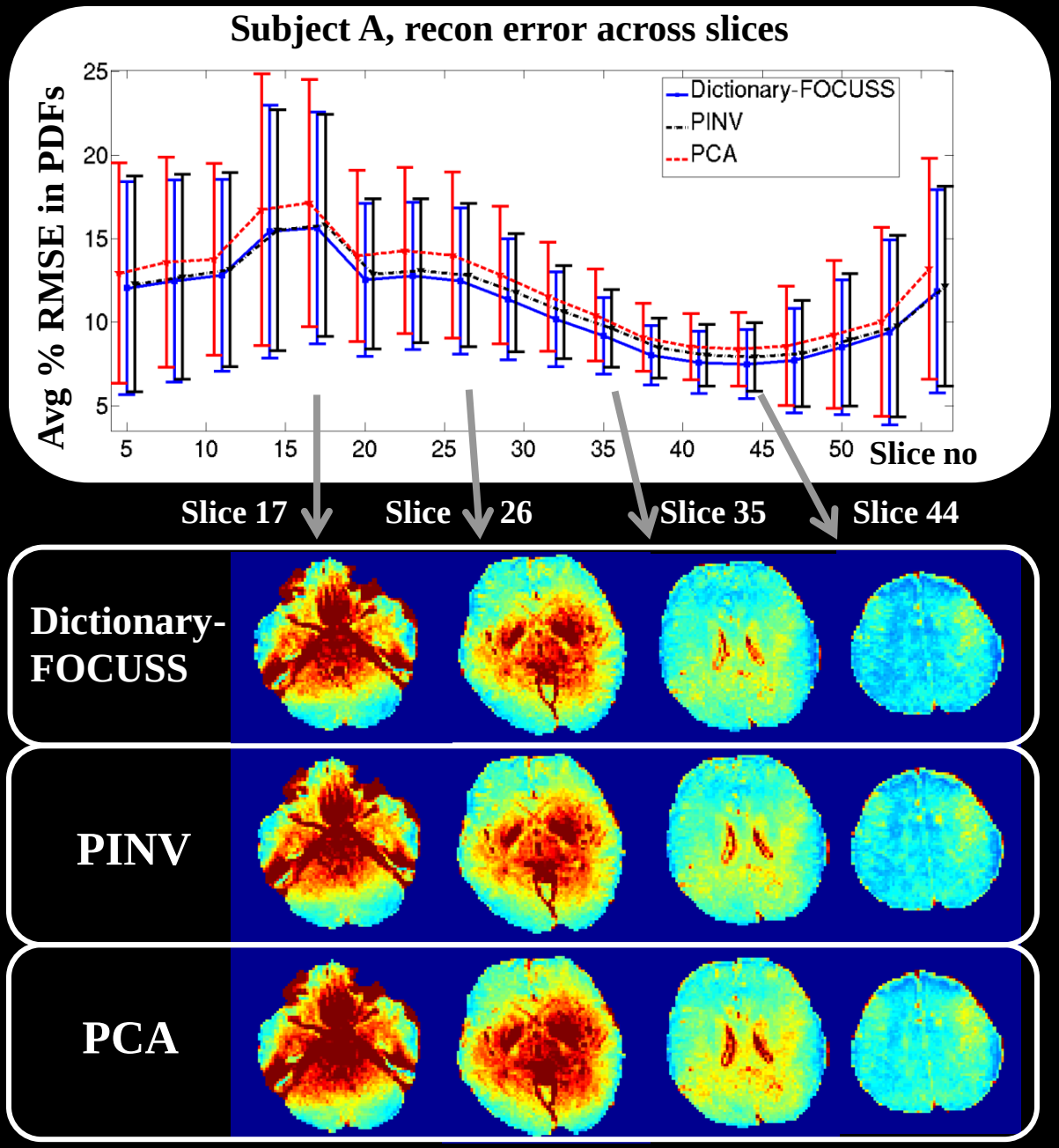
Slice 40



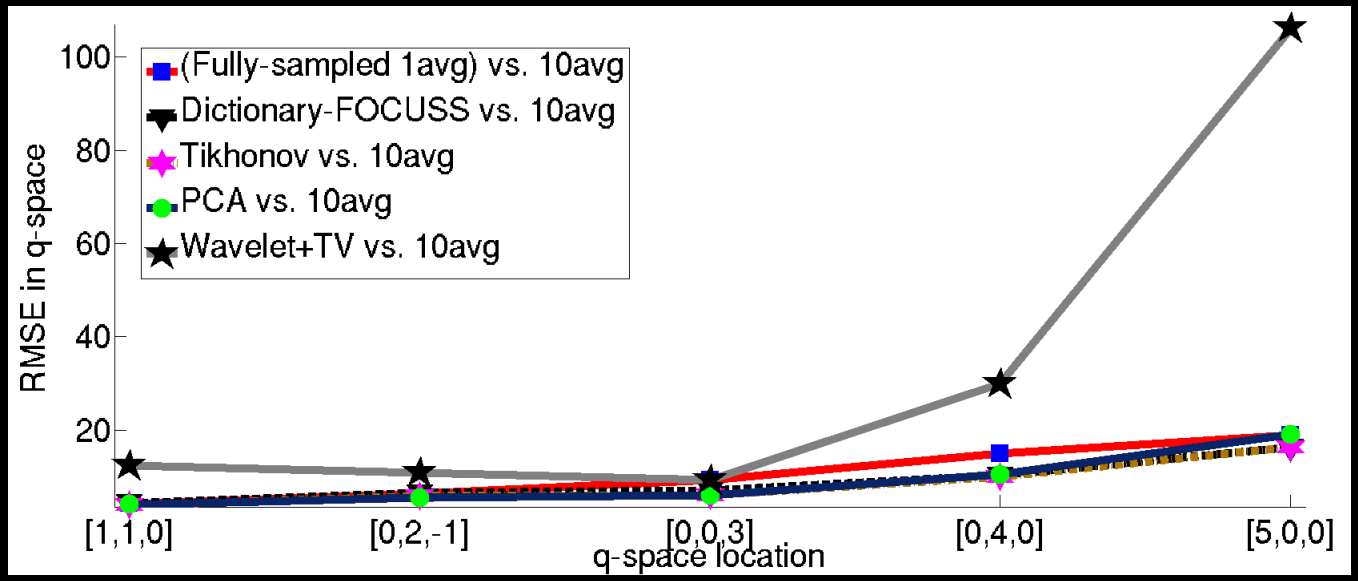
Subject A, pdf reconstruction error

Slice 40

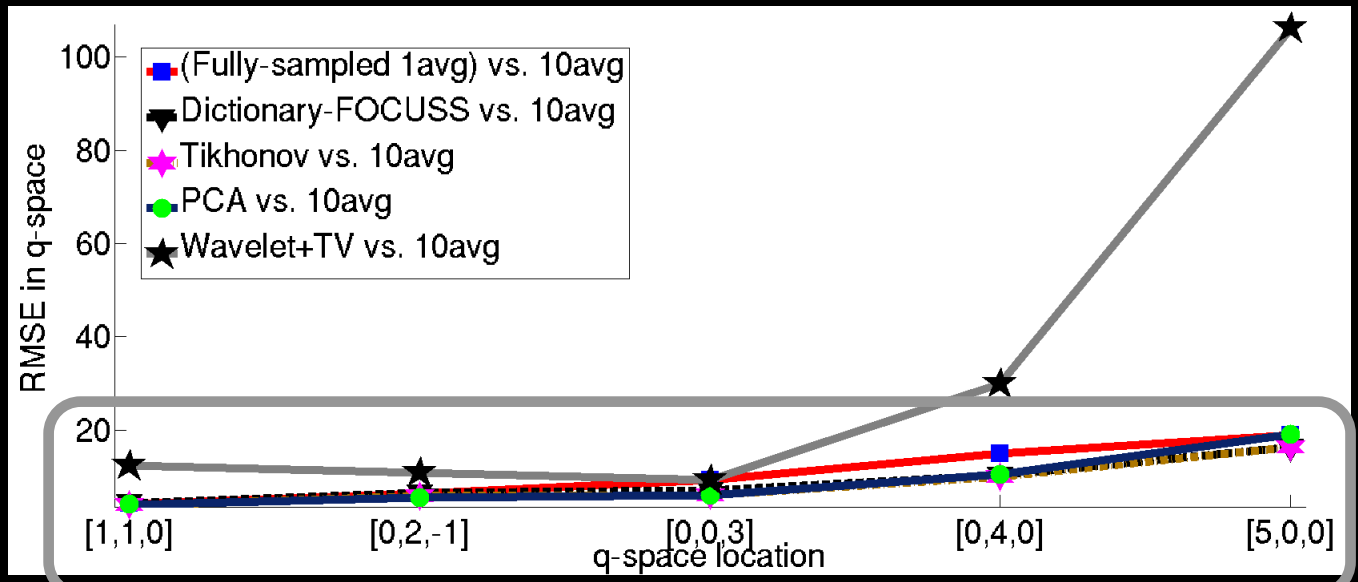


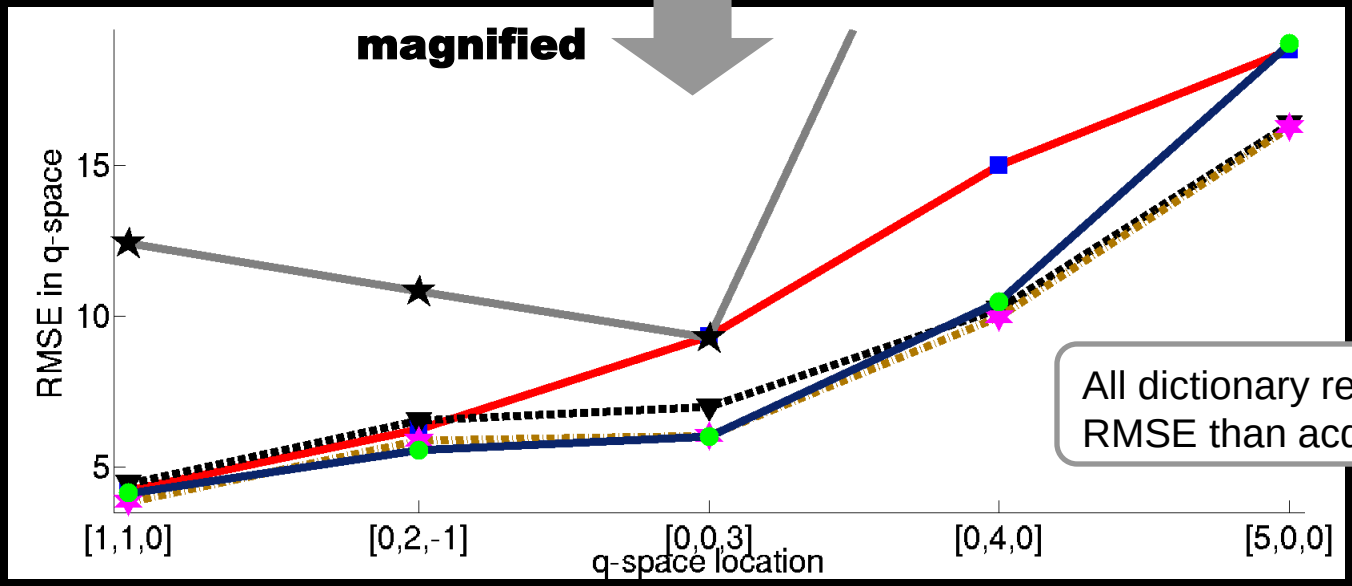

Subject A, recon error across slices
Avg % RMSE in PDFs
Slice no
Dictionary-FOCUSS
PINV
PCA
Slice 17
Slice 26
Slice 35
Slice 44
Dictionary-FOCUSS
PINV
PCA
20 %
0 %

Comparison to Low-Noise 10 avg Data



Comparison to Low-Noise 10 avg Data





All dictionary recons have lower RMSE than acquired data

Outline

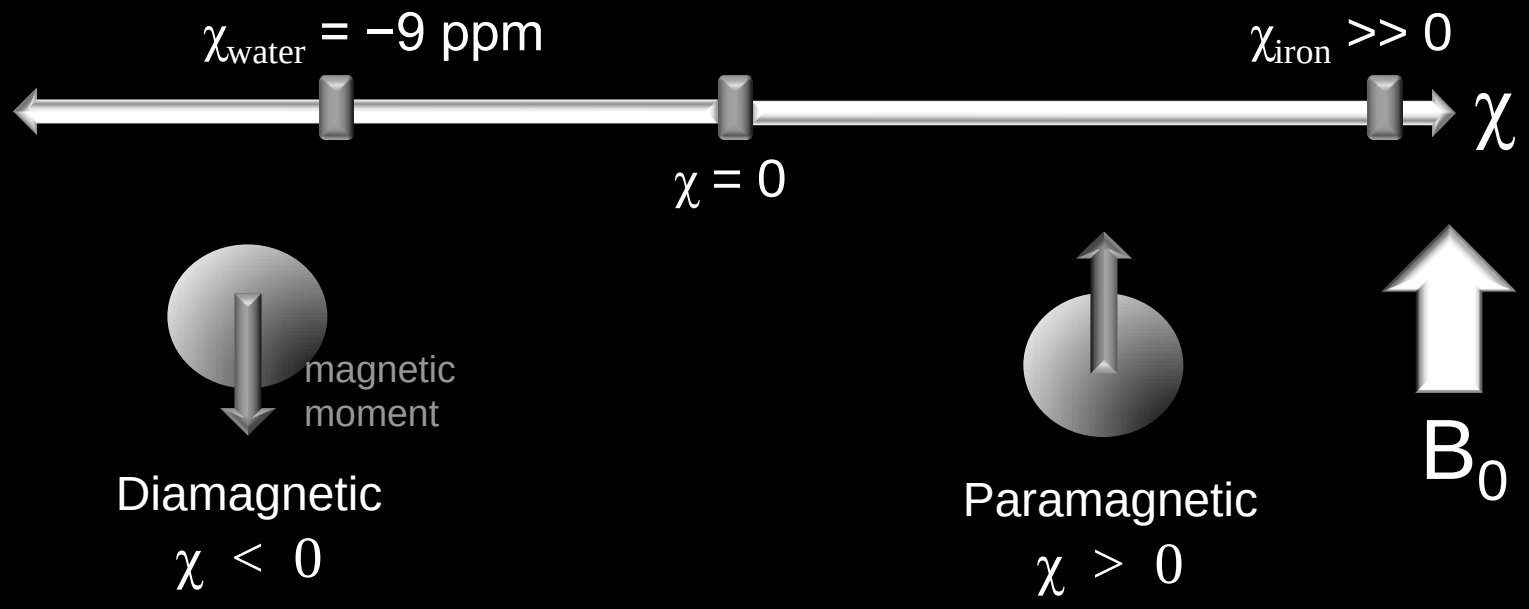
- Problems that were addressed, why they are worth solving
- Contribution to the field
- In particular,
  - Joint reconstruction of similar images
  - Accelerated Diffusion Spectrum Imaging
  - Quantifying tissue iron concentration
  - Lipid artifact suppression for Spectroscopic Imaging
    - Postpone to closed session

Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field

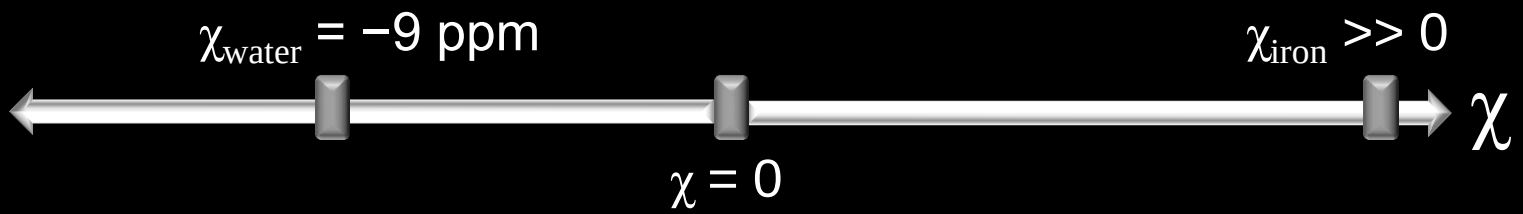
Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field



Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field

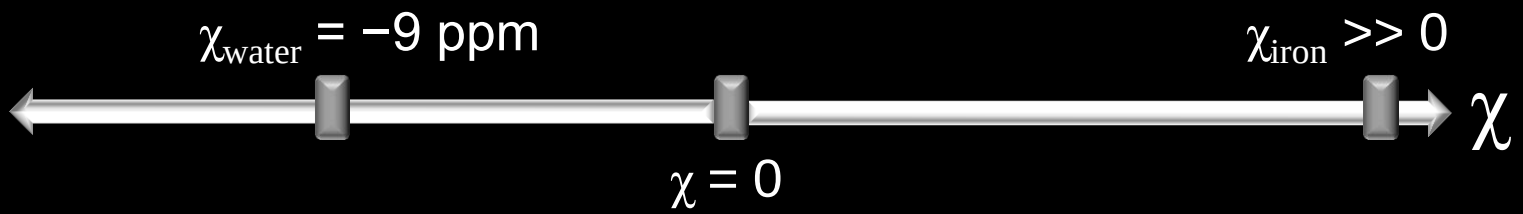


- Susceptibility of brain tissue is ≈ −9 ppm
- Tissues with increased iron deposition are relatively paramagnetic → χ **is more positive**

[1] Halgren & Sourander, J. Neurochem, 1960

Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field



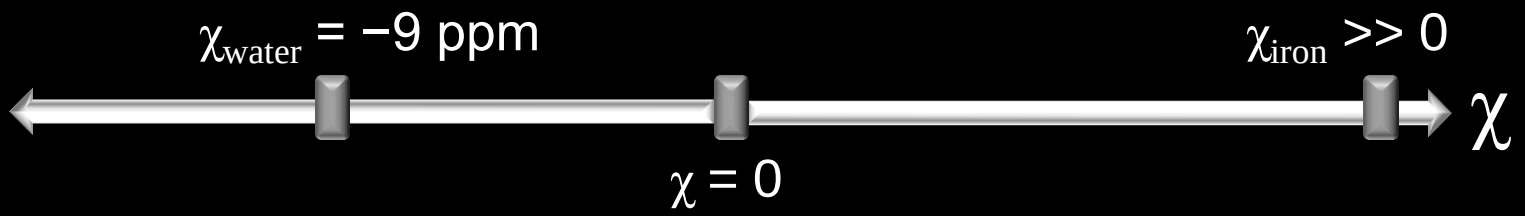
- Susceptibility of brain tissue is ≈ −9 ppm
- Tissues with increased iron deposition are relatively paramagnetic → **χ is more positive**
- Excessive iron concentration occurs in a variety of degenerative diseases[1],

  e.g. Alzheimer's, multiple sclerosis, Parkinson's

[1] Halgren & Sourander, J. Neurochem, 1960

Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field

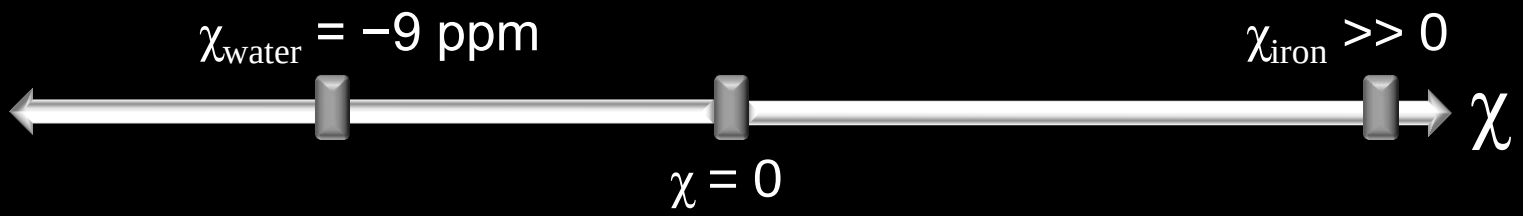


- Variations in tissue susceptibility affects the magnetic field

$\Delta\chi \rightarrow$ magnetic field perturbation

Susceptibility of Tissue

- Susceptibility χ : degree of magnetization of a material when placed in a magnetic field



- Variations in tissue susceptibility affects the magnetic field
- Field perturbation causes a change in MR signal phase

$$\underbrace{\Delta\chi}_{\text{estimate}} \rightarrow \text{magnetic field perturbation} \rightarrow \underbrace{\Delta\varphi}_{\text{measured}}$$

Quantitative Susceptibility Mapping (QSM)

- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility with applications such as,
  - Tissue contrast enhancement[1]
  - Estimation of venous blood oxygenation[2]
  - Quantification of tissue iron concentration[3]

[1] Duyn JH *et al*., PNAS 2007 [2] Fan AP *et al*., ISMRM 2010 [3] Liu T *et al*., ISMRM 2010

Quantitative Susceptibility Mapping (QSM)

- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility with applications such as,
  - Tissue contrast enhancement[1]
  - Estimation of venous blood oxygenation[2]
  - Quantification of tissue iron concentration[3]

- Estimation of the susceptibility map χ from the unwrapped phase φ involves solving an inverse problem,

$$\boldsymbol{\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi}$$

$\mathbf{F}$: Discrete Fourier Transform

$\mathbf{D}$: susceptibility kernel

$\boldsymbol{\delta} = \dfrac{\boldsymbol{\varphi}}{\gamma \cdot TE \cdot B_0}$: normalized field map

[1] Duyn JH *et al*., PNAS 2007 [2] Fan AP *et al*., ISMRM 2010 [3] Liu T *et al*., ISMRM 2010

Quantitative Susceptibility Mapping (QSM)

- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility with applications such as,
  - Tissue contrast enhancement[1]
  - Estimation of venous blood oxygenation[2]
  - Quantification of tissue iron concentration[3]

- Estimation of the susceptibility map χ from the unwrapped phase φ involves solving an inverse problem,

$$\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$$

δ: measured; χ: to be estimated

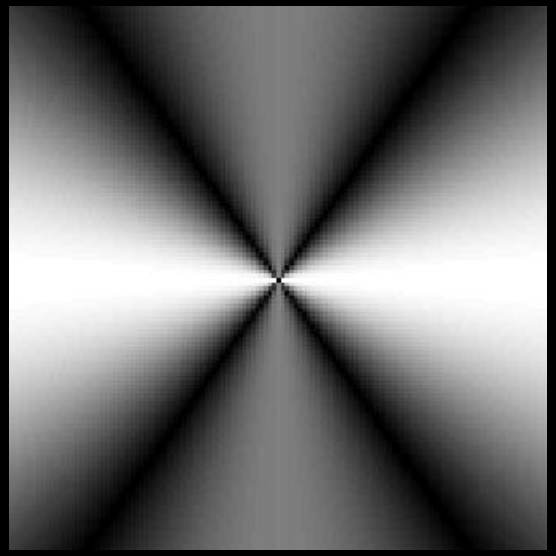
[1] Duyn JH *et al*., PNAS 2007 [2] Fan AP *et al*., ISMRM 2010 [3] Liu T *et al*., ISMRM 2010

Quantitative Susceptibility Mapping (QSM)

- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility with applications such as,
  - Tissue contrast enhancement[1]
  - Estimation of venous blood oxygenation[2]
  - Quantification of tissue iron concentration[3]
- Estimation of the susceptibility map χ from the unwrapped phase φ involves solving an inverse problem, $\boldsymbol{\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi}$
- The inversion is made difficult by zeros on a conical surface in susceptibility kernel $\mathbf{D}$

$$\mathbf{D} = \frac{1}{3} - \frac{{k_z}^2}{k^2}$$

|D|

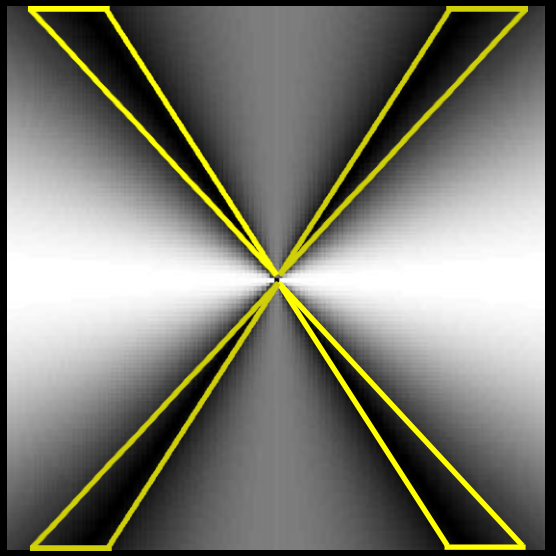


[1] Duyn JH *et al*., PNAS 2007 [2] Fan AP *et al*., ISMRM 2010 [3] Liu T *et al*., ISMRM 2010

Quantitative Susceptibility Mapping (QSM)

- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility with applications such as,
  - Tissue contrast enhancement[1]
  - Estimation of venous blood oxygenation[2]
  - Quantification of tissue iron concentration[3]
- Estimation of the susceptibility map χ from the unwrapped phase φ involves solving an inverse problem, $\boldsymbol{\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi}$
- Undersampling is due to physics
- Not in our control

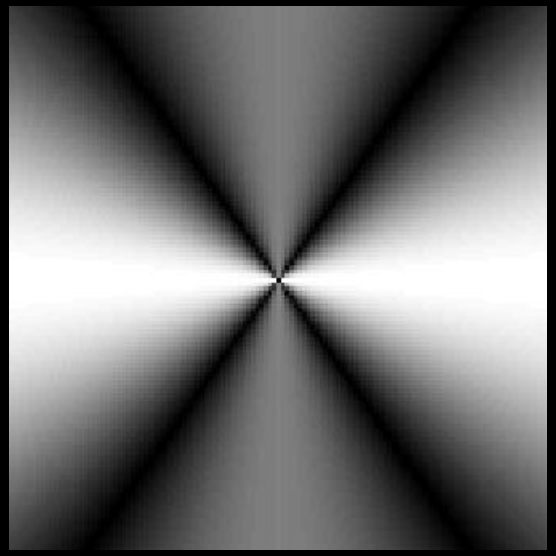
|D|



[1] Duyn JH *et al*., PNAS 2007 [2] Fan AP *et al*., ISMRM 2010 [3] Liu T *et al*., ISMRM 2010

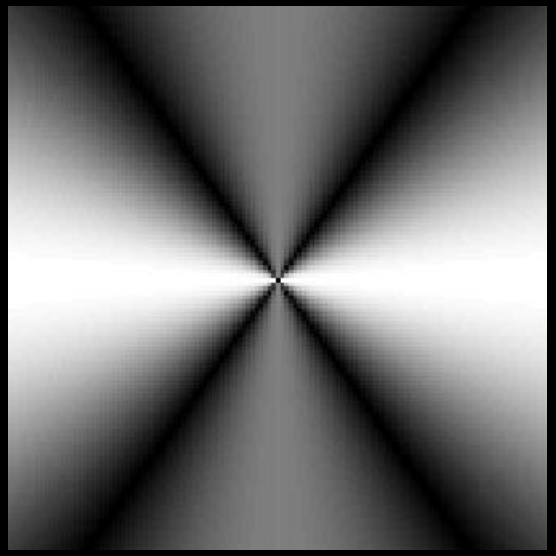
Regularized Inversion for QSM

|D|

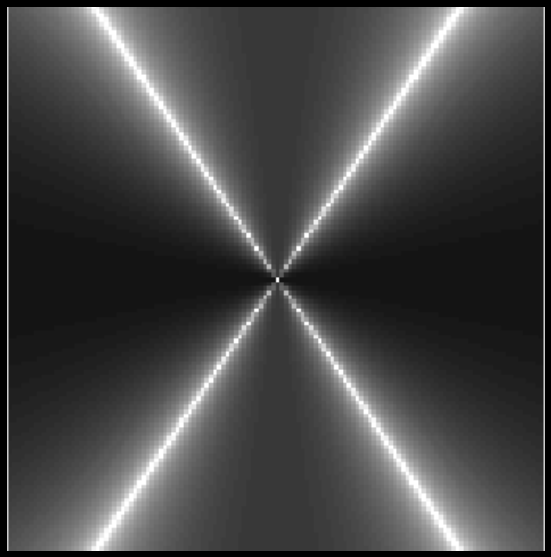


Regularized Inversion for QSM

|D|



$\log|\mathbf{D}^{-1}|$

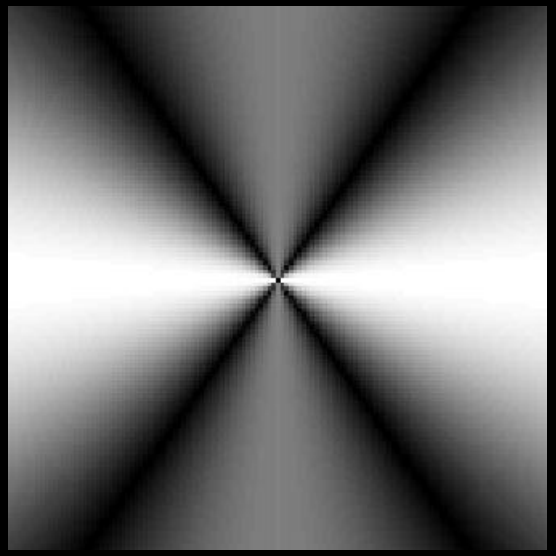


$$\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$$

- Solving for χ by convolving with the inverse of $\mathbf{D}$ is not possible, as it diverges along the magic angle

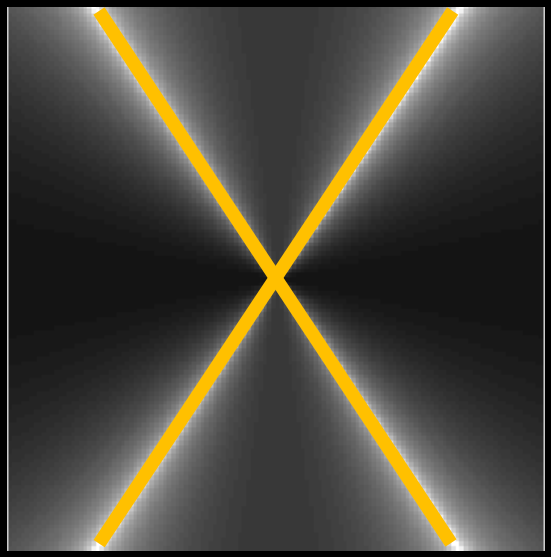
Regularized Inversion for QSM

|D|



$\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$

$\log|\mathbf{D}^{-1}|$



diverges to ∞

- Solving for χ by convolving with the inverse of $\mathbf{D}$ is not possible, as it diverges along the magic angle
- Use inverse problem formulation, apply regularization

Phase Processing

- Several processing steps are required to obtain the tissue phase

Phase Processing

- Several processing steps are required to obtain the tissue phase
  i. Mask out the skull

Using FSL Brain Extraction Tool[1]

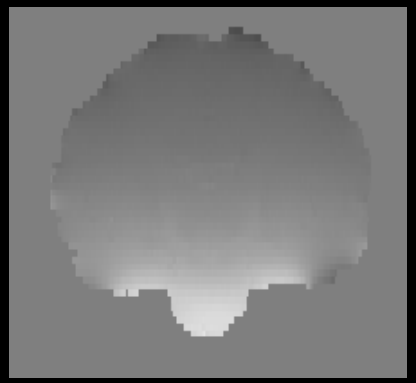
[1] Smith SM, Hum. Brain Mapp. 2002

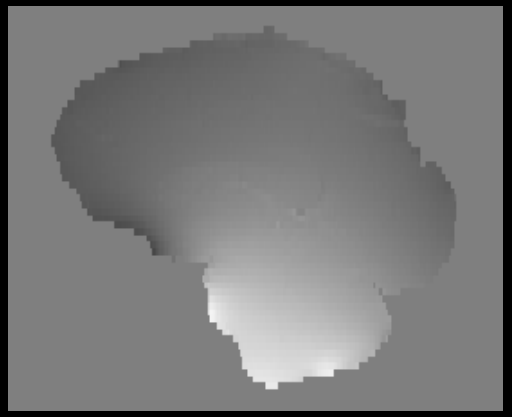
Phase Processing

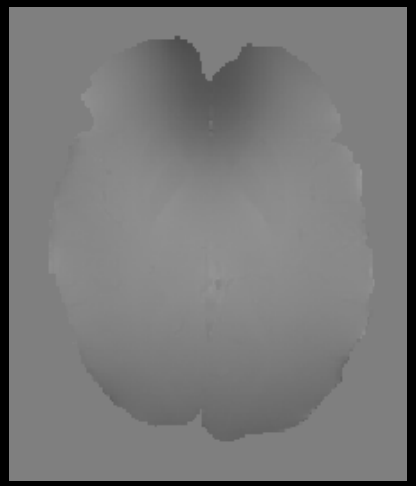
- Several processing steps are required to obtain the tissue phase

  i. Mask out the skull

  ii. Unwrap the phase







Using FSL PRELUDE[1]

[1] Jenkinson M, MRM 2003

Phase Processing

- Several processing steps are required to obtain the tissue phase

  i. Mask out the skull
  ii. Unwrap the phase

  iii. Remove background phase

  Phase accrued due to air-tissue interfaces needs to be removed

  This background component is ~10× larger than tissue phase

Phase Processing

- Several processing steps are required to obtain the tissue phase

  i. Mask out the skull
  ii. Unwrap the phase

  iii. <u>Remove background phase</u>

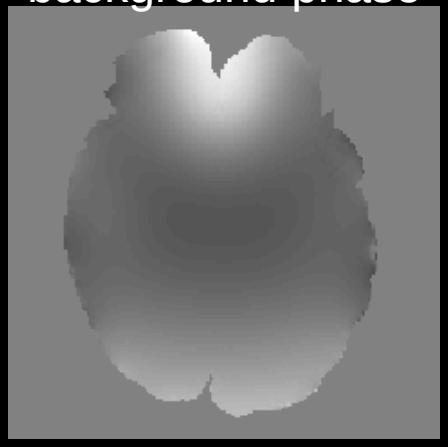
  Phase accrued due to air-tissue interfaces needs to be removed

  This background component is ~10× larger than tissue phase

initial phase

-0.8 ppm 0.8 ppm

background phase[1]



-0.8 ppm 0.8 ppm

[1] Liu T, NMR in Biomedicine 2011

Phase Processing

- Several processing steps are required to obtain the tissue phase

  i. Mask out the skull
  ii. Unwrap the phase

  iii. <u>Remove background phase</u>

  Phase accrued due to air-tissue interfaces needs to be removed

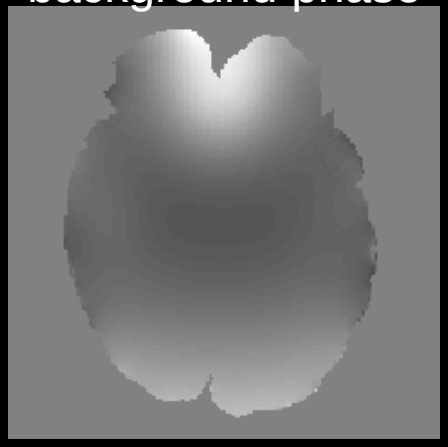
  This background component is ~10× larger than tissue phase

initial phase

-0.8 ppm 0.8 ppm

−

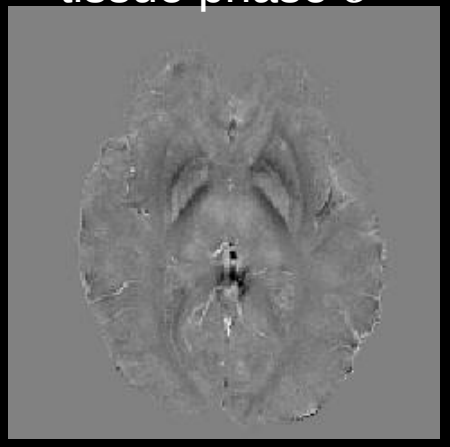
background phase



-0.8 ppm 0.8 ppm

=

tissue phase δ



-0.1 ppm 0.1 ppm

Phase Processing

- Several processing steps are required to obtain the tissue phase

  i. Mask out the skull
  ii. Unwrap the phase

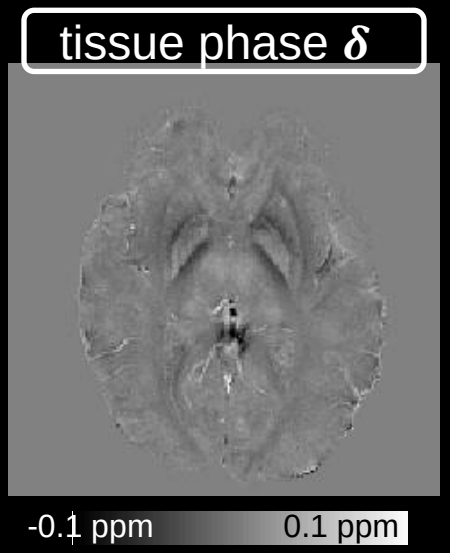
  iii. Remove background phase

  Phase accrued due to air-tissue interfaces needs to be removed

  This background component is ~10× larger than tissue phase

- Now we can solve for χ from tissue phase δ

$$\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$$



L1 Regularized Susceptibility Inversion

- We seek the susceptibility map that matches the observed tissue phase,

$$\text{Find } \boldsymbol{\chi} \text{ such that } \boldsymbol{\delta} = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi}$$

- Susceptibility values are tied to the magnetic properties of the underlying tissues; hence they vary smoothly within anatomical boundaries.

L1 Regularized Susceptibility Inversion

- We seek the susceptibility map that matches the observed tissue phase,

$$\text{Find } \chi \text{ such that } \delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$$

- Susceptibility values are tied to the magnetic properties of the underlying tissues; hence they vary smoothly within anatomical boundaries.

- Model the susceptibility map to be approximately piece-wise constant,

- Invoke sparsity inducing L1 norm on spatial gradients of χ

L1 Regularized Susceptibility Inversion

- We solve for the susceptibility distribution with a convex program,

$$\boldsymbol{\chi}_{tissue} = argmin_{\chi} \underbrace{\left\| \boldsymbol{\delta} - \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\boldsymbol{\chi} \right\|_2^2}_{\text{data consistency}} + \lambda \cdot \underbrace{\|\mathbf{G}\boldsymbol{\chi}\|_1}_{\ell_1 \text{ over gradients}}$$

L1 Regularized Susceptibility Inversion

- We solve for the susceptibility distribution with a convex program,

$$\boldsymbol{\chi}_{tissue} = argmin_{\chi} \underbrace{\left\| \boldsymbol{\delta} - \mathbf{F^{-1}DF}\boldsymbol{\chi} \right\|_2^2}_{\text{data consistency}} + \lambda \cdot \underbrace{\|\mathbf{G}\boldsymbol{\chi}\|_1}_{\ell_1 \text{ over gradients}}$$

- Here, λ serves as a regularization parameter that adjusts the smoothness of the solution

Tissue iron deposition in young and elderly subjects

- We used QSM to test the hypothesis that,

  iron deposition in striatal and brain stem nuclei is greater in older than younger adults

- Subjects:

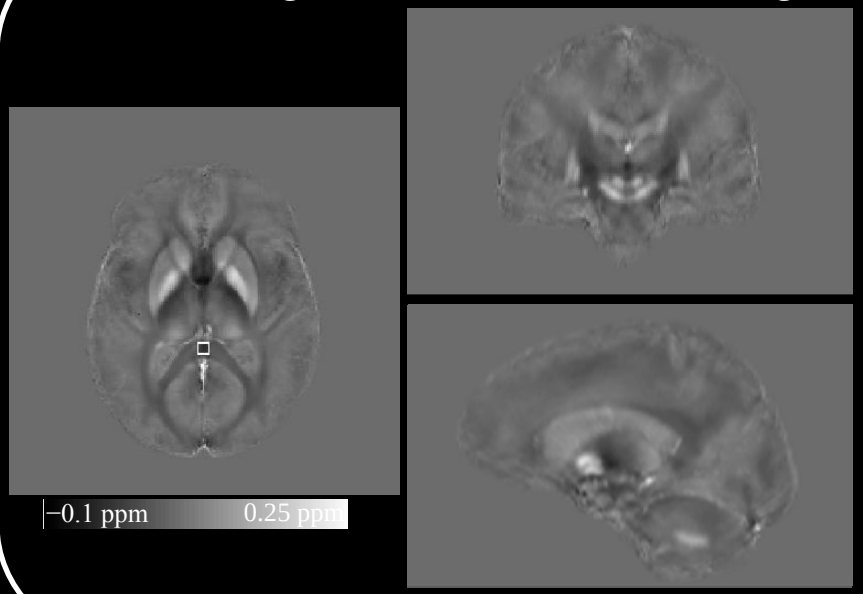
  11 younger adults (age = 24.0 ± 2.5) and

  12 elderly adults (age = 74.4 ± 7.6)

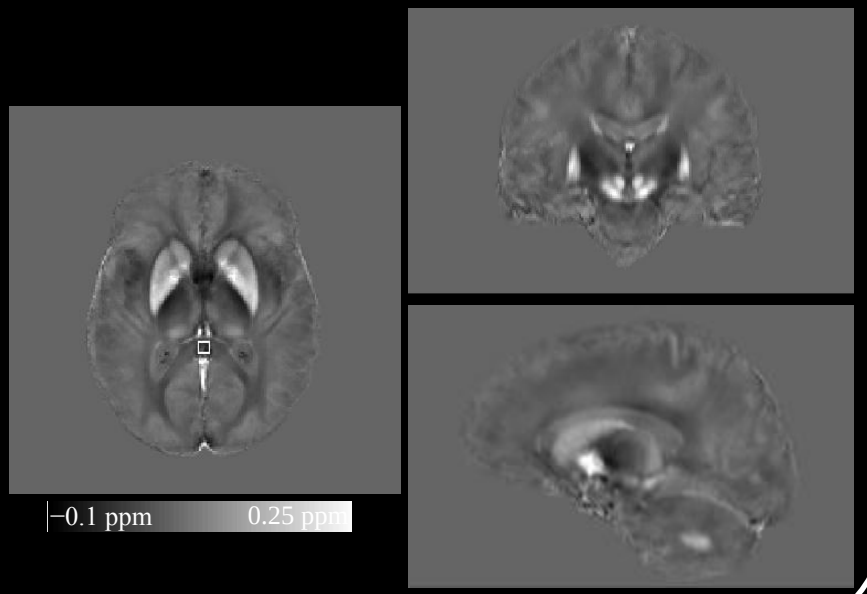
- Data:

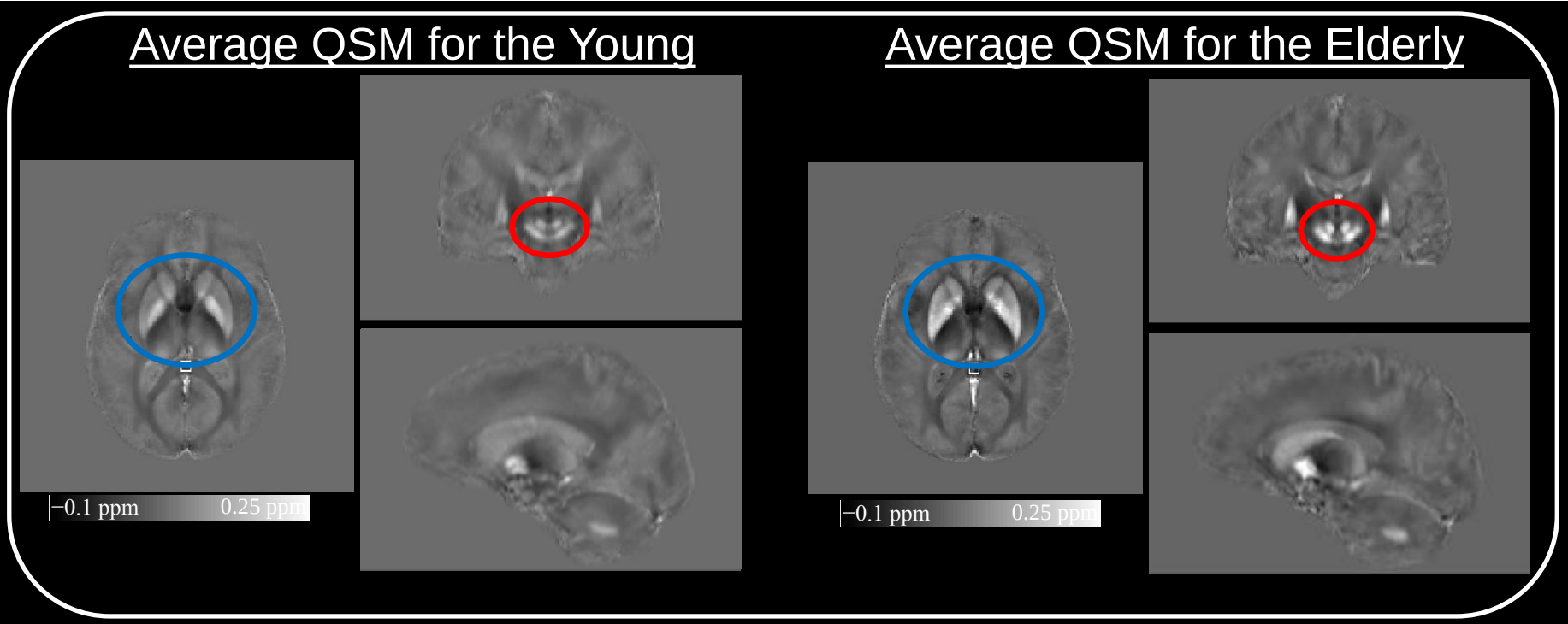
  Susceptibility Weighted 3D SPGR at 1.5 T

Average QSM for the Young

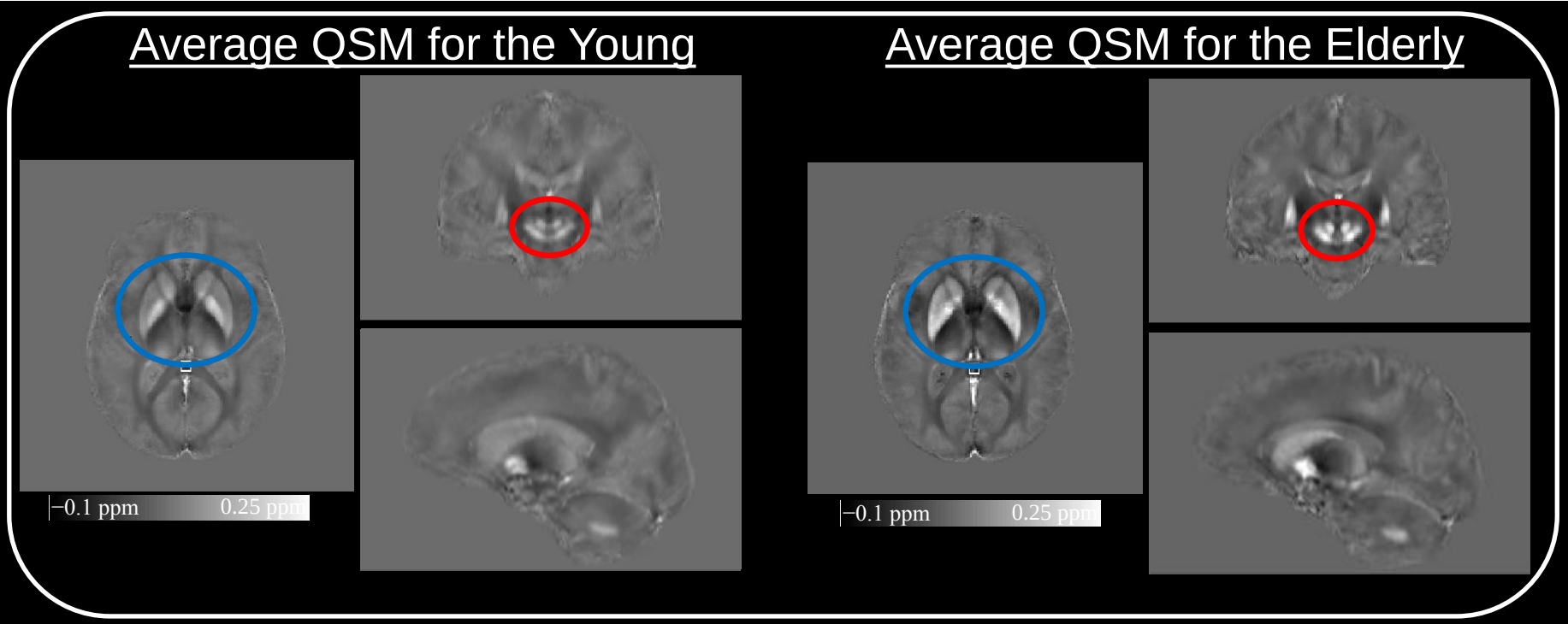


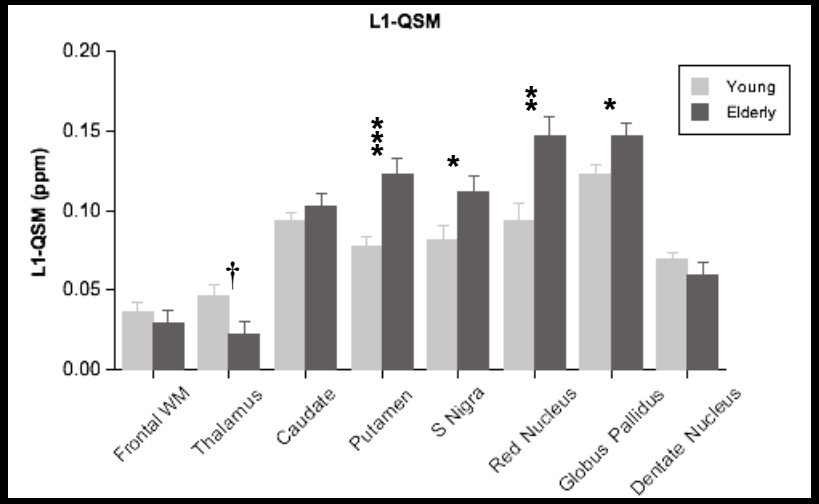
Average QSM for the Elderly





Striatal ROIs **Brain Stem ROIs**



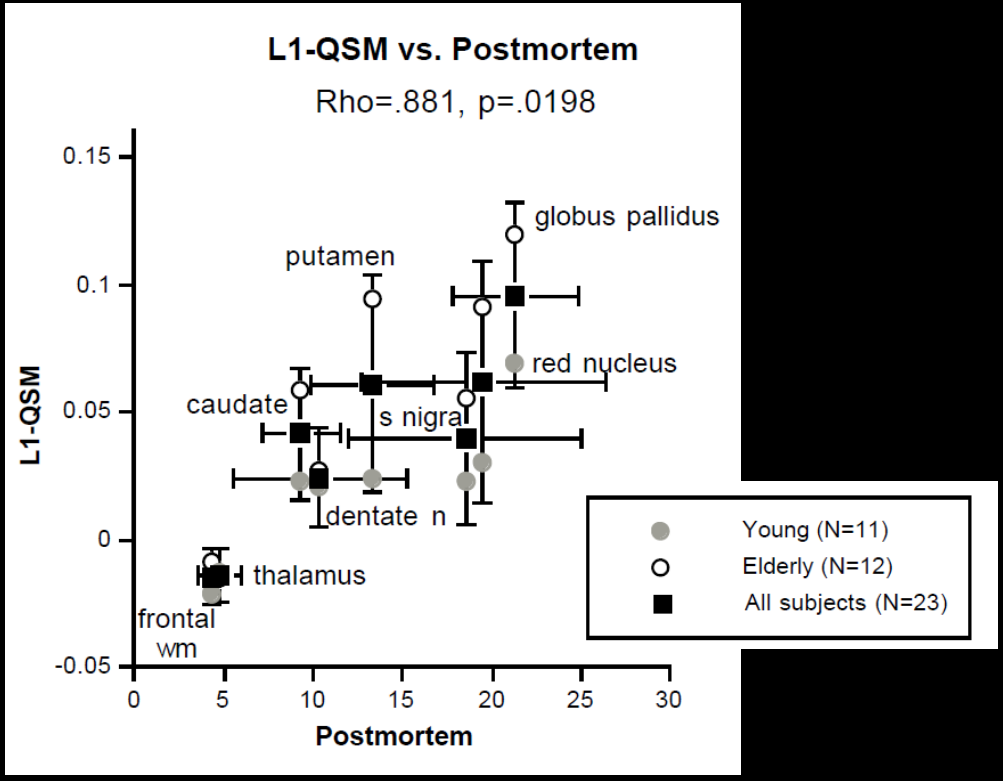


Elderly >> Young Iron Deposition

Putamen	p=0.0004	Striatal
Globus Pallidus	p=0.001	
Red Nucleus	p=0.002	Brain Stem
Substantia Nigra	p=0.003	

QSM vs. Postmortem

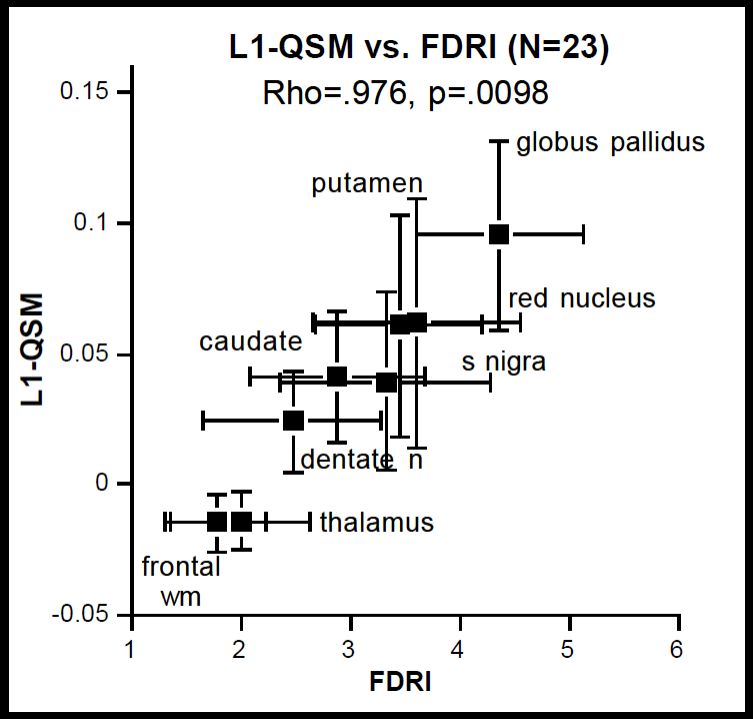
- QSM results correlate well with published postmortem results[1], with *Rho* = 0.881, *p* = 0.0198



[1] Hallgren, B. and Sourander, P. 1958, Journal of Neurochem

QSM vs. FDRI

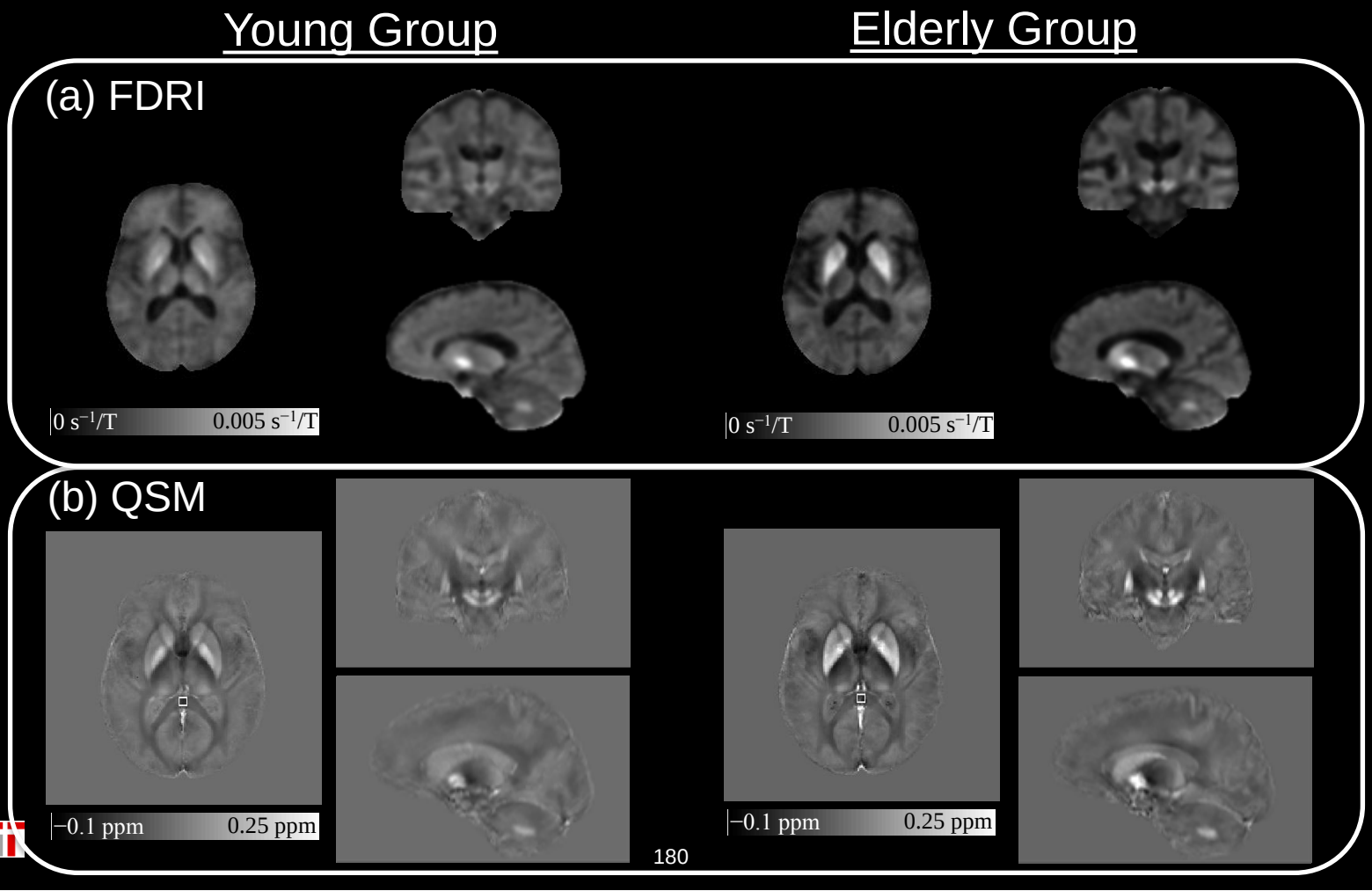
- Field-Dependent Relaxation Rate Increase (FDRI)[1] is another iron quantification method that requires data acquisition at two different field strengths.
- QSM is strongly correlated with FDRI results, with *Rho* = 0.976, *p* = 0.0098



[1] Bartzokis, G. *et al*., 1993, Magn Res Med

QSM vs. FDRI

- QSM requires data acquisition at a single field strength, and has much higher spatial resolution, enabling iron quantification in vessels.

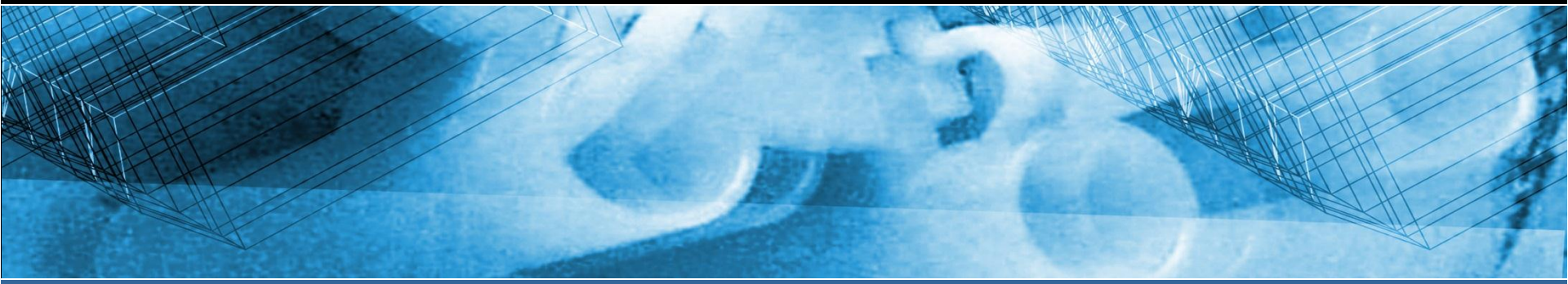


Conclusion

- Proposed algorithms that
  - Provide faster data acquisition in structural imaging and Diffusion Spectrum Imaging
  - Allow quantitative mapping of tissue susceptibility
  - Suppress lipid artifacts in MR spectroscopic imaging

Conclusion

- Proposed algorithms that
  - Provide faster data acquisition in structural imaging and Diffusion Spectrum Imaging
  - Allow quantitative mapping of tissue susceptibility
  - Suppress lipid artifacts in MR spectroscopic imaging
- **Thank you all for coming!**



Publications

Joint Reconstruction

- Journal:

  Multi-contrast Reconstruction with Bayesian Compressed Sensing

  B. Bilgic, V.K. Goyal, E. Adalsteinsson

  Magnetic Resonance in Medicine, 2011

- Conference Abstract:

  Joint Bayesian Compressed Sensing for Multi-contrast Reconstruction

  B. Bilgic, V.K. Goyal, E. Adalsteinsson

  ISMRM 2011, *oral presentation*

  Joint Bayesian Compressed Sensing with Prior Estimate

  B. Bilgic, E. Adalsteinsson

  ISMRM 2012, *oral presentation*

DSI

- Journal:

  Accelerated Diffusion Spectrum Imaging with Compressed Sensing using Adaptive Dictionaries

  B. Bilgic, K. Setsompop, J. Cohen-Adad, A. Yendiki, L.L. Wald, E. Adalsteinsson

  Magnetic Resonance in Medicine, 2012

  Accelerated Diffusion Spectrum Imaging with Compressed Sensing using Adaptive Dictionaries

  B. Bilgic, I. Chatnuntawech, K. Setsompop, S.F. Cauley, L.L. Wald, E. Adalsteinsson

  IEEE Trans on Medical Imaging, *submitted*

- Conference Paper:

  Accelerated Diffusion Spectrum Imaging with Compressed Sensing using Adaptive Dictionaries

  B. Bilgic, K. Setsompop, J. Cohen-Adad, V. Wedeen, L. Wald, E. Adalsteinsson

  MICCAI 2012, *oral presentation*

- Conference Abstract:

  Fast DSI Reconstruction with Trained Dictionaries

  B. Bilgic, I. Chatnuntawech, K. Setsompop, S.F. Cauley, L.L. Wald, E. Adalsteinsson

  ISMRM 2013, *submitted*

QSM

- Journal:

  MRI Estimates of Brain Iron Concentration in Normal Aging Using Quantitative Susceptibility Mapping

  B. Bilgic, A. Pfefferbaum, T. Rohlfing, E.V. Sullivan, E. Adalsteinsson

  NeuroImage, 2012

- Conference Abstract:

  Quantitative Susceptibility Map Reconstruction with Magnitude Prior

  B. Bilgic, A.P. Fan, E. Adalsteinsson

  ISMRM 2011, *oral presentation*

  Regularized QSM in Seconds

  B. Bilgic, I. Chatnuntawech, A.P. Fan, E. Adalsteinsson

  ISMRM 2013, *submitted*

Lipid Suppression

- Journal:

  Lipid Suppression in CSI with Spatial Priors and Highly Undersampled Peripheral k-space

  B. Bilgic, B. Gagoski, T. Kok, E. Adalsteinsson

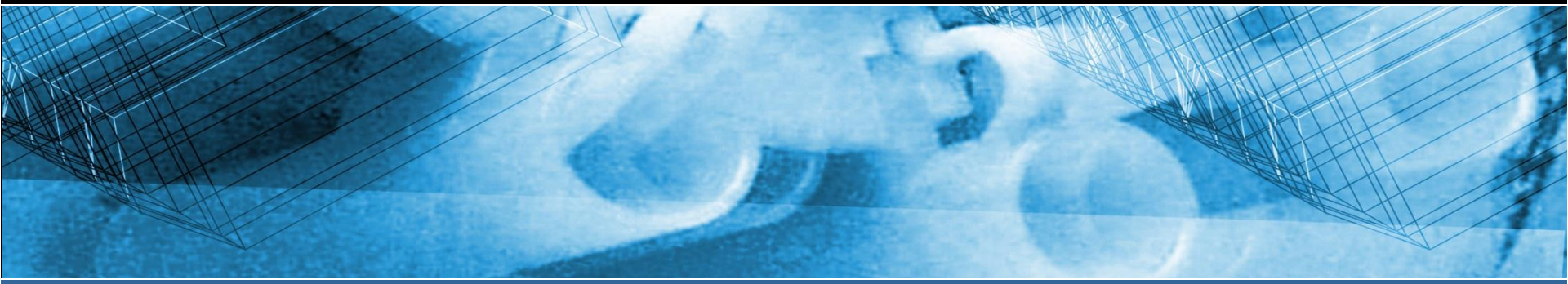
  Magnetic Resonance in Medicine, 2012

- Conference Abstract:

  Lipid Suppression in CSI with Highly-Undersampled Peripheral k-Space and Spatial Priors

  B. Bilgic, B. Gagoski, E. Adalsteinsson

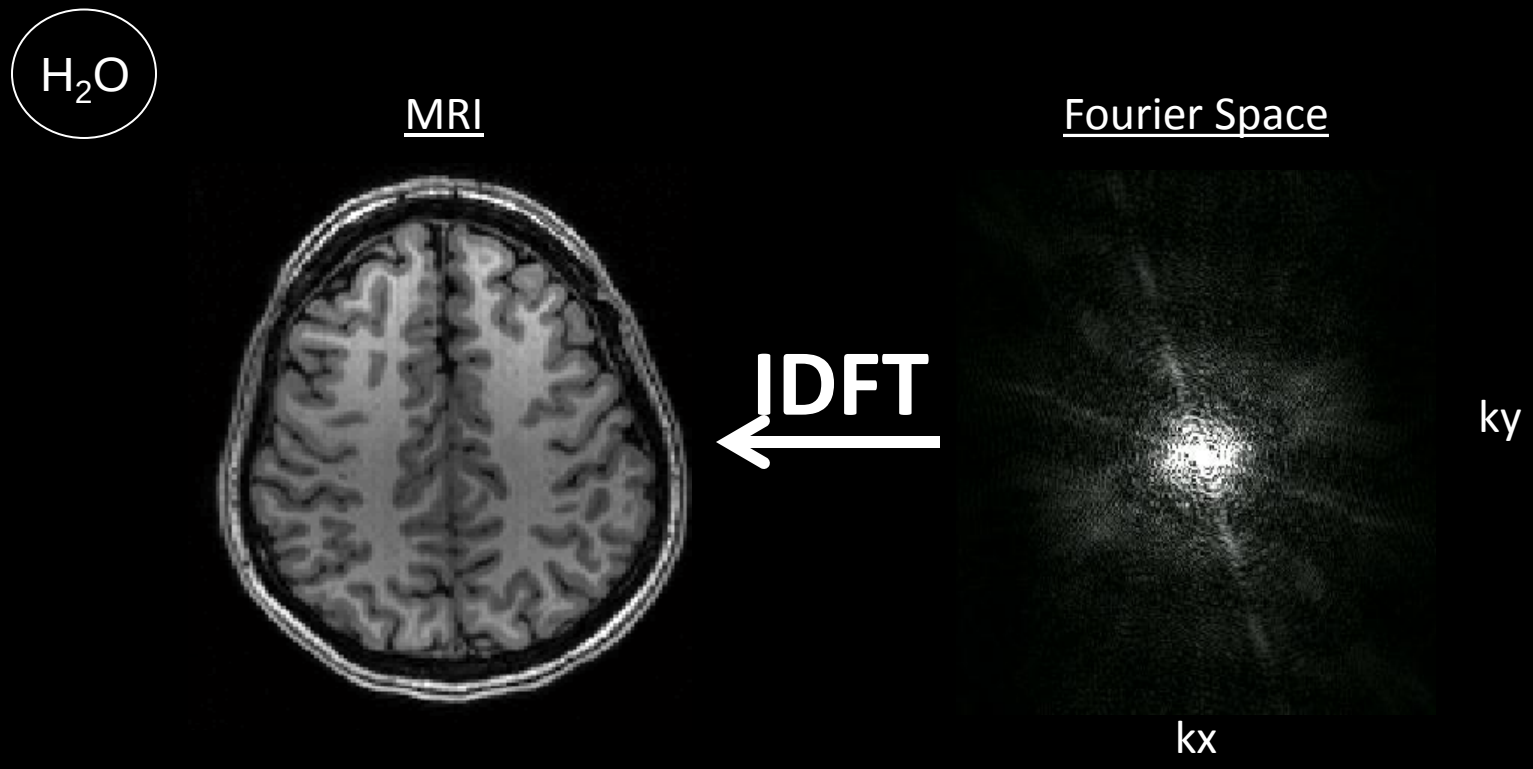
  ISMRM 2012, *poster presentation*



Lipid artifact suppression for Spectroscopic Imaging

MRI and MRSI

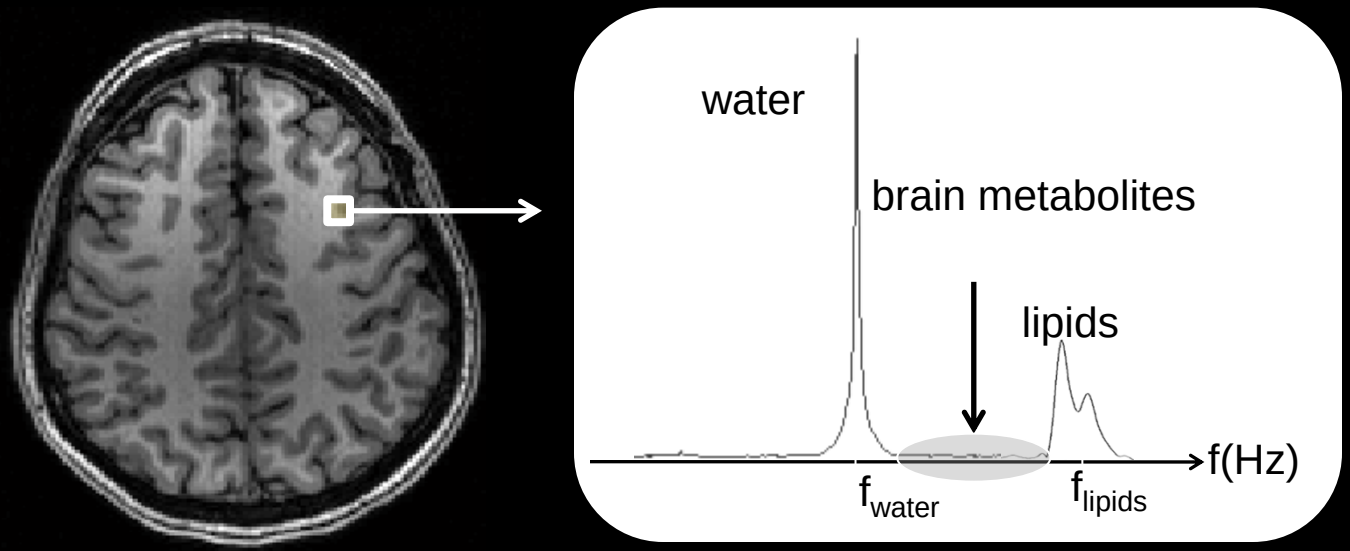
- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- Data are collected in (kx,ky,kz)



MRI and MRSI

- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- MR *Spectroscopic* Imaging (MRSI) or Chemical Shift Imaging (CSI) provides *spatial and* ***spectral*** encoding

MRI

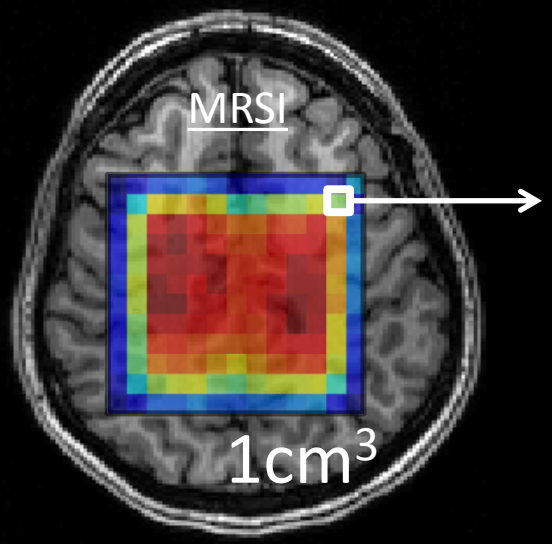


MRI and MRSI

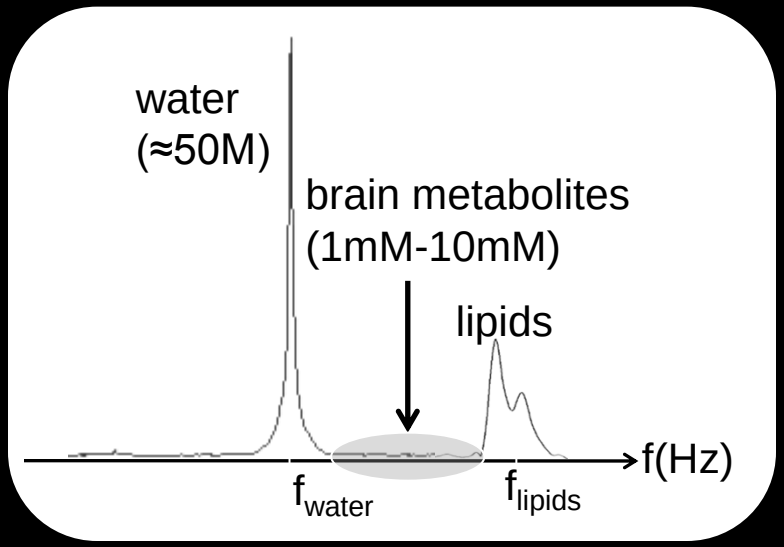
- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- MR *Spectroscopic* Imaging (MRSI) or Chemical Shift Imaging (CSI) provides *spatial and* ***spectral*** encoding

$$SNR = V_{size} \times \sqrt{T_{acq}}$$

MRI



1mm³

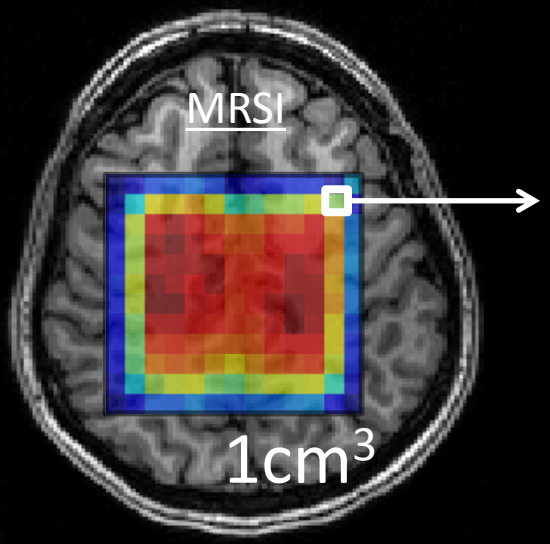


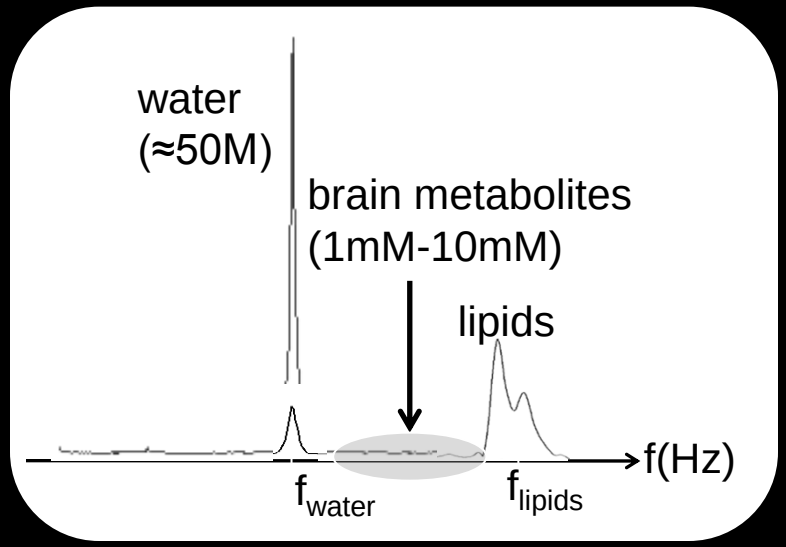
MRI and MRSI

- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- MR *Spectroscopic* Imaging (MRSI) or Chemical Shift Imaging (CSI) provides *spatial and* ***spectral*** encoding

$$SNR = V_{size} \times \sqrt{T_{acq}}$$

MRI





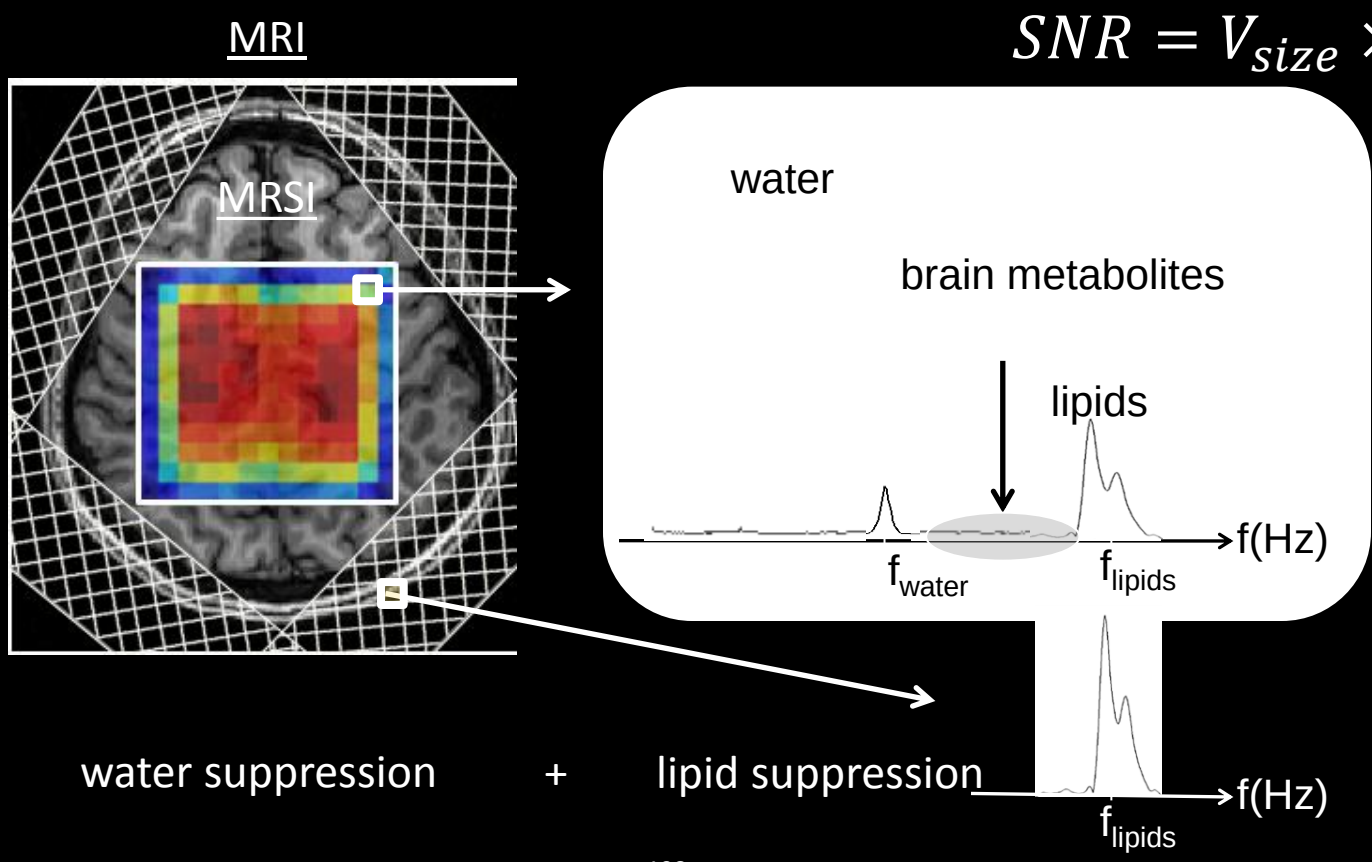
1mm3

water suppression

MRI and MRSI

- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- MR *Spectroscopic* Imaging (MRSI) or Chemical Shift Imaging (CSI) provides *spatial and* ***spectral*** encoding

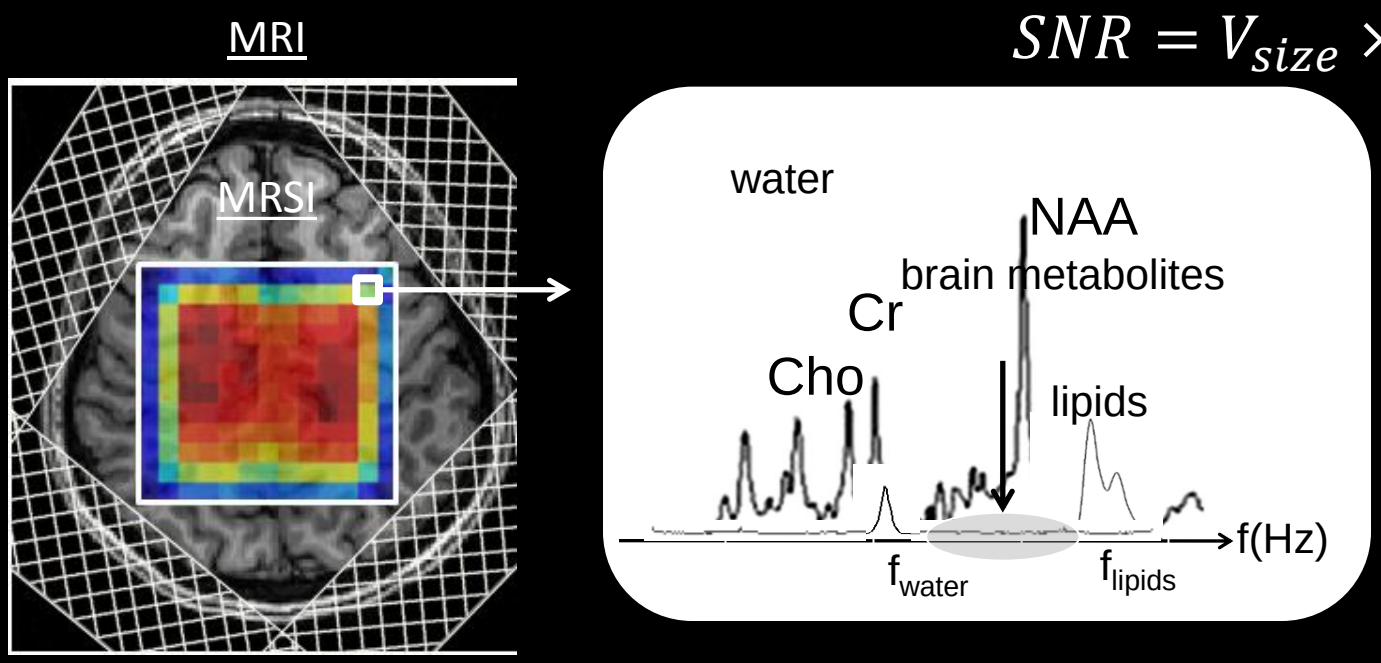
$$SNR = V_{size} \times \sqrt{T_{acq}}$$



MRI and MRSI

- Magnetic Resonance (MR) Imaging enables *spatial* encoding of the human tissue
- MR *Spectroscopic* Imaging (MRSI) or Chemical Shift Imaging (CSI) provides *spatial and* ***spectral*** encoding

$$SNR = V_{size} \times \sqrt{T_{acq}}$$



water suppression + lipid suppression

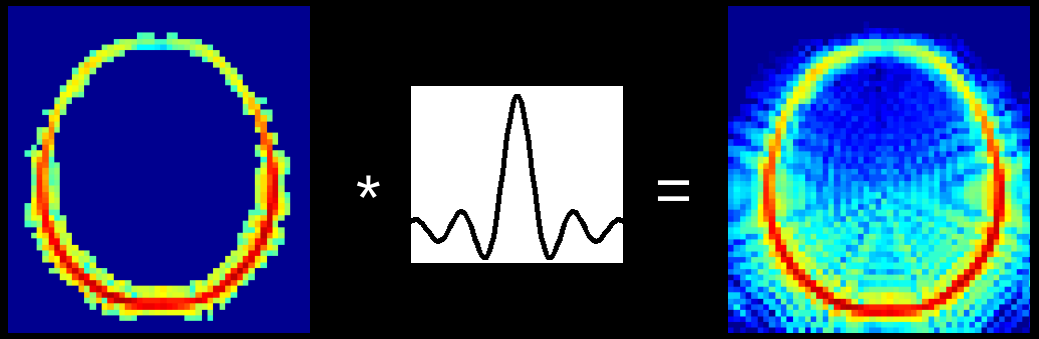
Lipid signals in Spectroscopy

- Voxel sizes in spectroscopy are typically large ~1cm^3
- This aims to increase the SNR of brain metabolites
- Encoding space and resonance frequency within reasonable scan time also limits the spatial resolution

Lipid signals in Spectroscopy

- Voxel sizes in spectroscopy are typically large ~1cm^3
- This aims to increase the SNR of brain metabolites
- Encoding space and resonance frequency within reasonable scan time also limits the spatial resolution

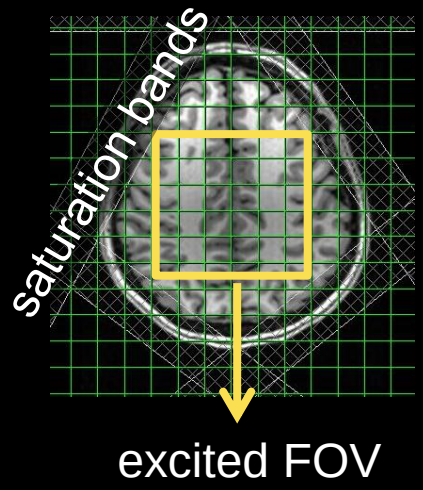
- Poor spatial resolution causes subcutaneous lipids to contaminate the metabolites inside the brain

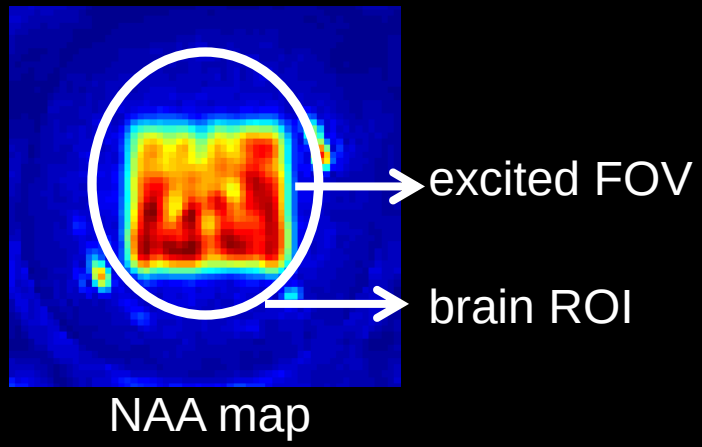


lipid layer low resolution lipid ringing

Previously proposed lipid suppression methods

- Outer Volume Suppression (OVS)[1,2,3]
  - Excites a rectangular field-of-view (FOV) inside the brain
  - Peripheral brain regions cannot be mapped

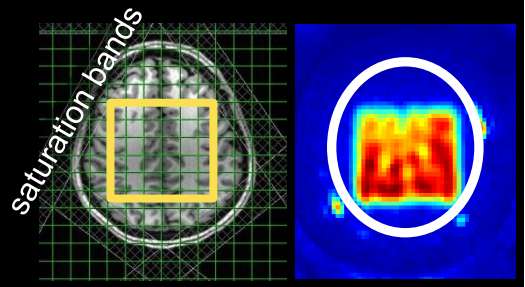




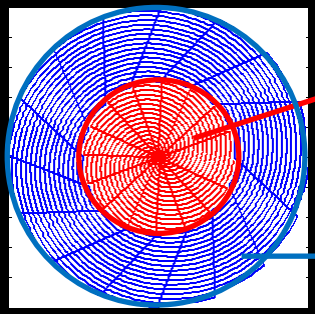
[1] Duyn et al. Radiology 1993 [2] Le Roux et al. JMRI 1998 [3] Luo et al. MRM 2001

Previously proposed lipid suppression methods

- Outer Volume Suppression (OVS)
  - Excites a rectangular FOV inside the brain
  - Peripheral brain regions cannot be mapped



- Dual-Density reconstruction[1,2,3]
  - Obtain center k-space with multiple avg for metabolites, high k-space with 1 avg for lipids which have strong signal
  - High frequency lipid information reduces ringing

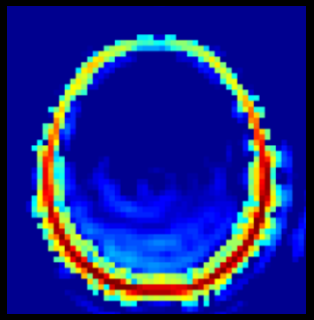


k-space sampling

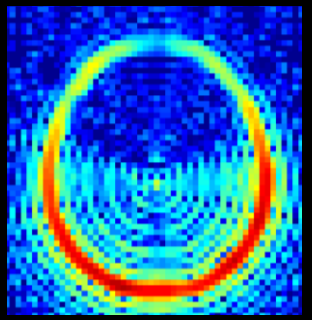
multi avg for metabolites at Nyquist rate

1 avg for lipid at Nyquist rate

Lipid Maps



dual-density

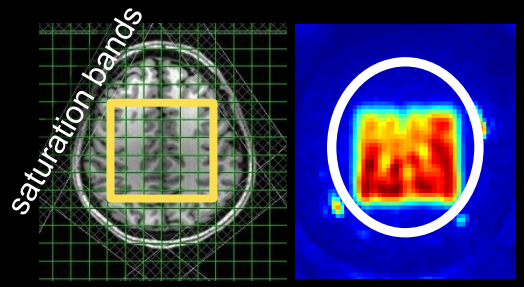


using only center k-space

[1] Hu et al. IEEE T.Med.Img. 1991 [2] Metzger et al. MRI 1999 [3] Sarkar et al. MRI 2002

Previously proposed lipid suppression methods

- Outer Volume Suppression (OVS)
  - Excites a rectangular FOV inside the brain
  - Peripheral brain regions cannot be mapped



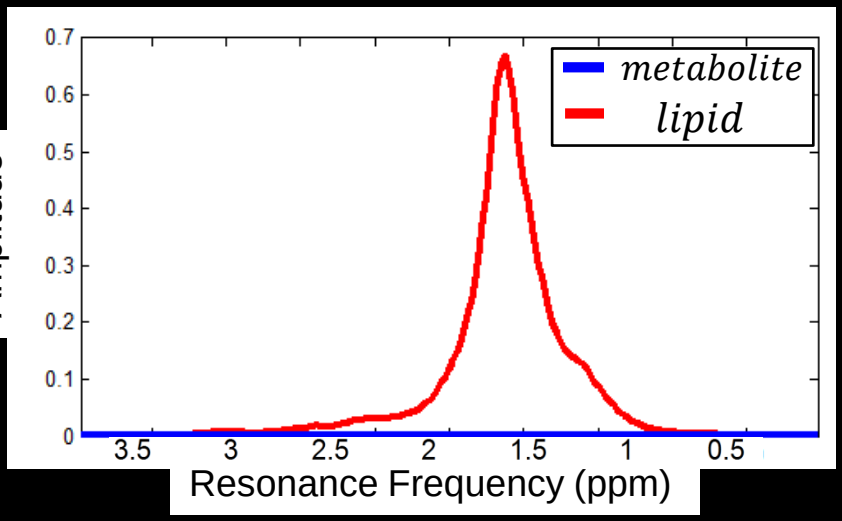
- Dual-Density reconstruction
  - Obtain center k-space with multiple avg for metabolites, high k-space with 1 avg for lipids which have strong signal
  - High frequency lipid information reduces ringing

- Lipid-basis penalty[1]
  - Lipid and metabolite spectra are approximately orthogonal
  - Inside the brain, inner product of metabolites and lipids should be small

[1] Lee & Adalsteinsson ISMRM 2010

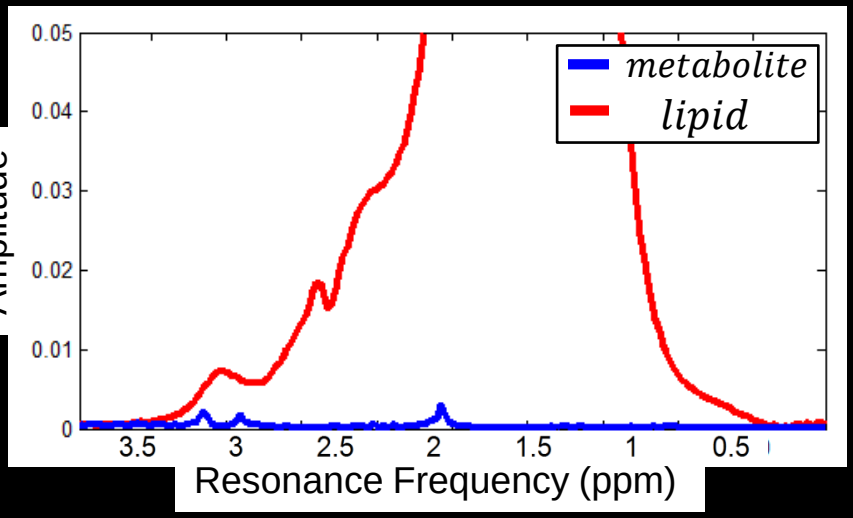
Lipid-basis penalty

- Orthogonality of metabolite and lipid spectra
  - Consider a metabolite spectra (taken from the OVS scan) and a lipid spectra (from non-lipid suppressed acquisition)



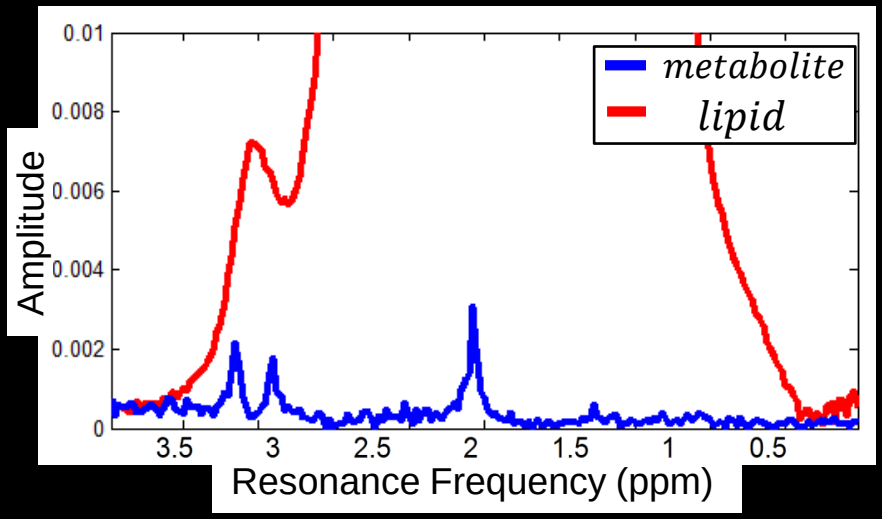
Lipid-basis penalty

- Orthogonality of metabolite and lipid spectra
  - Consider a metabolite spectra (taken from the OVS scan) and a lipid spectra (from non-lipid suppressed acquisition)



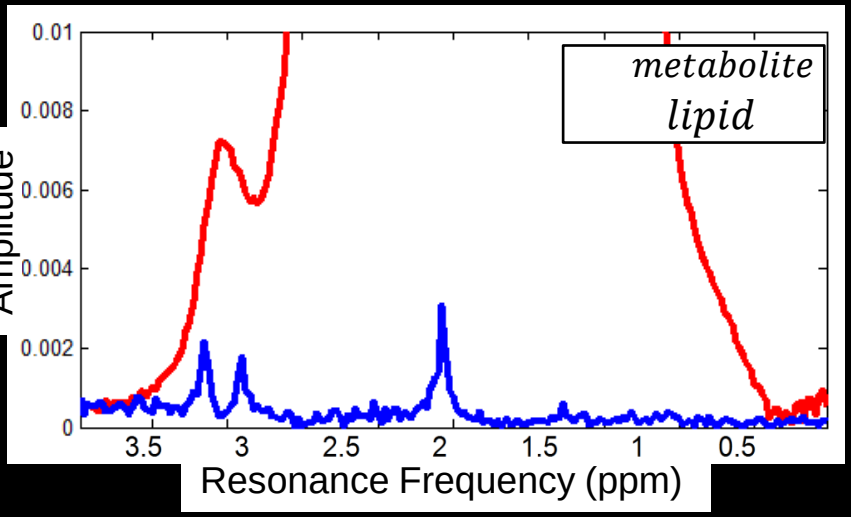
Lipid-basis penalty

- Orthogonality of metabolite and lipid spectra
  - Consider a metabolite spectra (taken from the OVS scan) and a lipid spectra (from non-lipid suppressed acquisition)

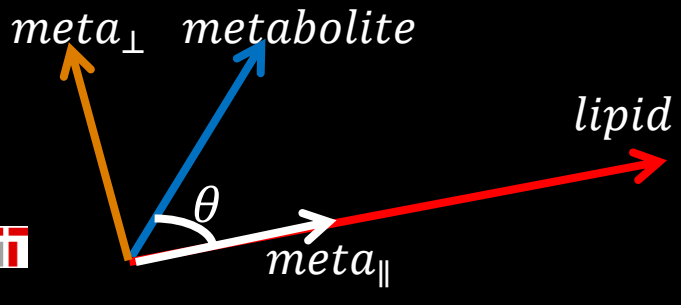


Lipid-basis penalty

- Orthogonality of metabolite and lipid spectra
  - Consider a metabolite spectra (taken from the OVS scan) and a lipid spectra (from non-lipid suppressed acquisition)



  - Compute the projection of metabolite signal onto the lipid spectra and the orthogonal component

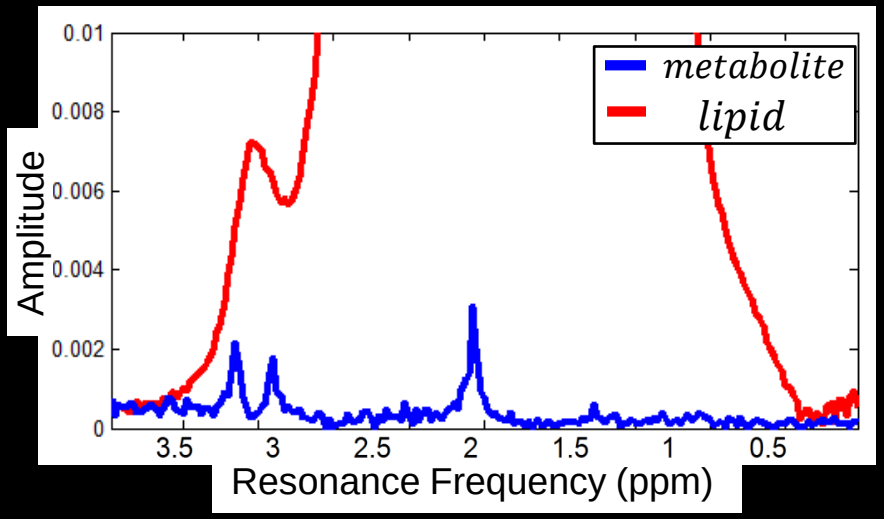


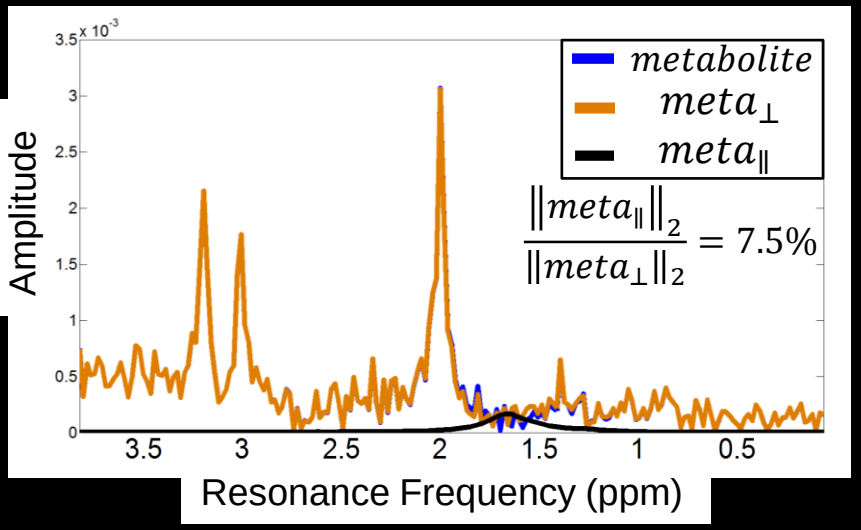
$$meta_\parallel = \frac{metabolite^H lipid}{\|lipid\|_2^2} \cdot lipid$$

$$meta_\perp = metabolite - meta_\parallel$$

Lipid-basis penalty

- Orthogonality of metabolite and lipid spectra
  - Consider a metabolite spectra (taken from the OVS scan) and a lipid spectra (from non-lipid suppressed acquisition)





  - Compute the projection of metabolite signal onto the lipid spectra and the orthogonal component
  - The projection is negligibly small, confirming the orthogonality approximation

Proposed method - I

- Combining dual-density and lipid-basis penalty
  - In addition to multiple avg low-resolution CSI acquisition, obtain 1−2 avg high-resolution lipid data
  - Apply iterative lipid-basis penalty

Proposed method - I

- Combining dual-density and lipid-basis penalty
  - In addition to multiple avg low-resolution CSI acquisition, obtain 1−2 avg high-resolution lipid data
  - Apply iterative lipid-basis penalty

- Form high-resolution, masked lipid image $\boldsymbol{x}_{lipid}$

$$\boldsymbol{x}_{lipid} = \mathbf{M}_{lipid}\mathbf{F}_{high}^{-1}\boldsymbol{y}_{high}$$

$\mathbf{M}_{lipid}$: lipid mask
$\boldsymbol{y}_{high}$: high-res k-space data
$\mathbf{F}_{high}$: high-res DFT operator

Proposed method - I

- Combining dual-density and lipid-basis penalty
  - In addition to multiple avg low-resolution CSI acquisition, obtain 1−2 avg high-resolution lipid data
  - Apply iterative lipid-basis penalty

- Form high-resolution, masked lipid image $\boldsymbol{x}_{lipid}$

$$\boldsymbol{x}_{lipid} = \mathbf{M}_{lipid}\mathbf{F}_{high}^{-1}\boldsymbol{y}_{high}$$

- Compute the dual-density image (combine $\boldsymbol{x}_{lipid}$ with low-res CSI)

$$\boldsymbol{x}_{dual} = \mathbf{F}_{high}^{-1}\left\{\left(\mathbf{F}_{high} - \mathbf{F}_{low}\right)\boldsymbol{x}_{lipid} + \boldsymbol{y}_{low}\right\}$$

$\boldsymbol{y}_{low}$: low-res k-space data
$\mathbf{F}_{low}$: low-res DFT operator

Proposed method - I

- Combining dual-density and lipid-basis penalty
  - In addition to multiple avg low-resolution CSI acquisition, obtain 1−2 avg high-resolution lipid data
  - Apply iterative lipid-basis penalty

- Form high-resolution, masked lipid image $\boldsymbol{x}_{lipid}$

$$\boldsymbol{x}_{lipid} = \mathbf{M}_{lipid}\mathbf{F}_{high}^{-1}\boldsymbol{y}_{high}$$

- Compute the dual-density image (combine $\boldsymbol{x}_{lipid}$ with low-res CSI)

$$\boldsymbol{x}_{dual} = \mathbf{F}_{high}^{-1}\left\{\left(\mathbf{F}_{high} - \mathbf{F}_{low}\right)\boldsymbol{x}_{lipid} + \boldsymbol{y}_{low}\right\}$$

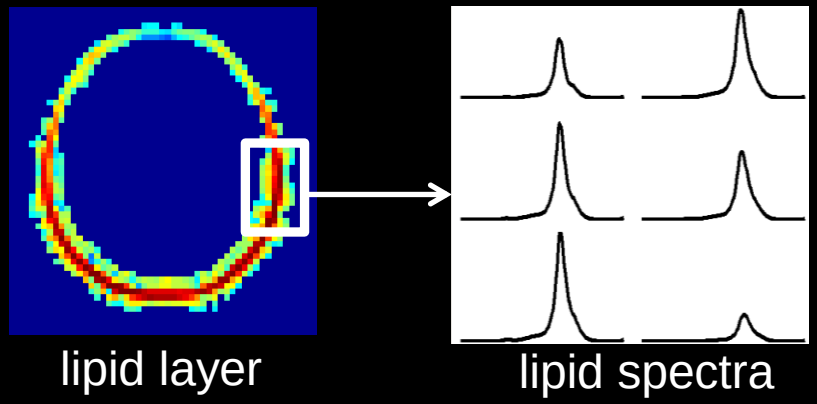
- Make a lipid-basis matrix whose columns are lipid spectra in $\boldsymbol{x}_{dual}$ and enforce orthogonality between metabolites and lipids

$$\boldsymbol{x}_{basic} = \operatorname{argmin}_{\boldsymbol{x}} \left\|\mathbf{F}_{high}\boldsymbol{x} - \boldsymbol{y}_{dual}\right\|_2^2 + \lambda \cdot \sum_{i \in \mathbf{M}_{brain}} \left\|\mathbf{L}_{dual}{}^{H}\, \boldsymbol{x}_i\right\|_1$$

$\mathbf{L}_{dual}$: lipid-basis matrix
$\mathbf{M}_{brain}$: brain mask

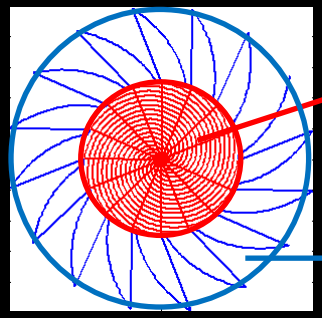
Proposed method - II

- Obtaining the high-res lipid image with compressed sensing
  - Lipid layer is ~sparse in space and in frequency



Proposed method - II

- Obtaining the high-res lipid image with compressed sensing
  - Lipid layer is ~sparse in space and in frequency
  - In addition to acquiring just 1−2 averages, substantially undersample the high-resolution scan to estimate lipid layer



k-space sampling

multi avg for
metabolites
at Nyquist rate

1 avg for lipid
10-fold undersampling

Proposed method - II

- Obtaining the high-res lipid image with compressed sensing
  - Lipid layer is ~sparse in space and in frequency
  - In addition to acquiring just 1−2 averages, substantially undersample the high-resolution scan to estimate lipid layer
  - Compute the lipid image with FOCUSS[1] algorithm that imposes ℓ_1 penalty in space and frequency:

For iteration number $t = 1, \ldots T$,

$$\mathrm{W}_{j,j}^{t} = diag\left(\left|x_j^t\right|^{1/2}\right)$$

$$\boldsymbol{q}^t = \mathrm{argmin}_{\boldsymbol{q}} \|\boldsymbol{q}\|_2^2 \quad \text{such that} \quad \mathbf{M}_\Omega \mathbf{F}_{high} \mathbf{W}^t \boldsymbol{q} = \boldsymbol{y}_{high}$$

$$\boldsymbol{x}^{t+1} = \mathbf{W}^t \boldsymbol{q}^t$$

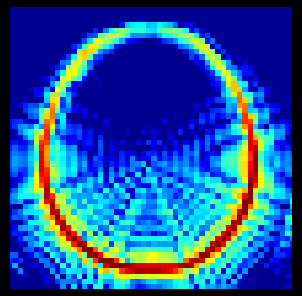
$\mathbf{M}_\Omega$: k-space undersampling mask

$\boldsymbol{x}^{T+1}$: CS recon for high-res lipid image

[1] Gorodnitsky & Rao IEEE T.Signal Proc. 1997

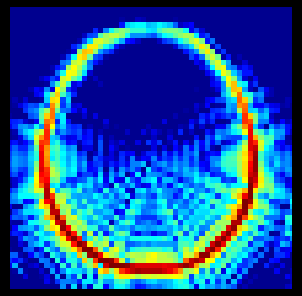
Data Acquisition

- In Vivo whole brain excitation
  - No lipid suppression, TE = 50 ms
  - Voxel size = 0.16 cc, 20 averages, in 33 min
  - CHESS for water suppression, PRESS-box excites whole FOV

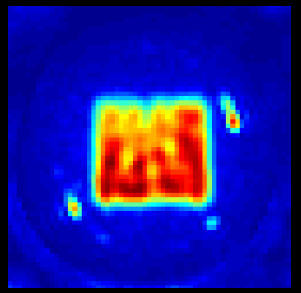


Data Acquisition

- In Vivo whole brain excitation
  - No lipid suppression, TE = 50 ms
  - Voxel size = 0.16 cc, 20 averages, in 33 min
  - CHESS for water suppression, PRESS-box excites whole FOV



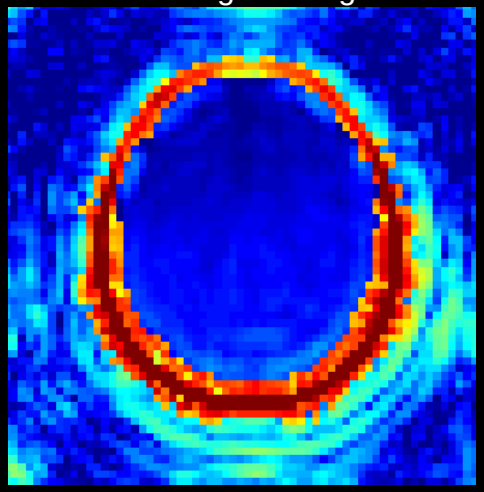
- Outer Volume Suppression acquisition
  - Voxel size = 0.5 cc, 20 averages, in 11 min
  - OVS bands null the lipid signals
  - PRESS-box excites 9×9 cm^2 FOV inside the brain



Lipid Maps at TE = 50 ms

(a) Gold standard

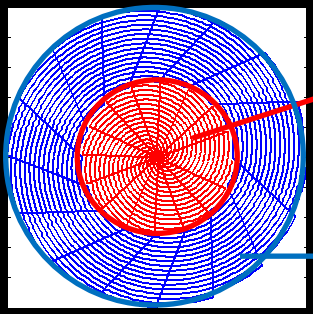
20 avg_{high}, R_{high}=1



15 dB

-35 dB

To serve as gold standard, lipid-basis penalty is applied to 20 average, 0.16 cc data



k-space sampling

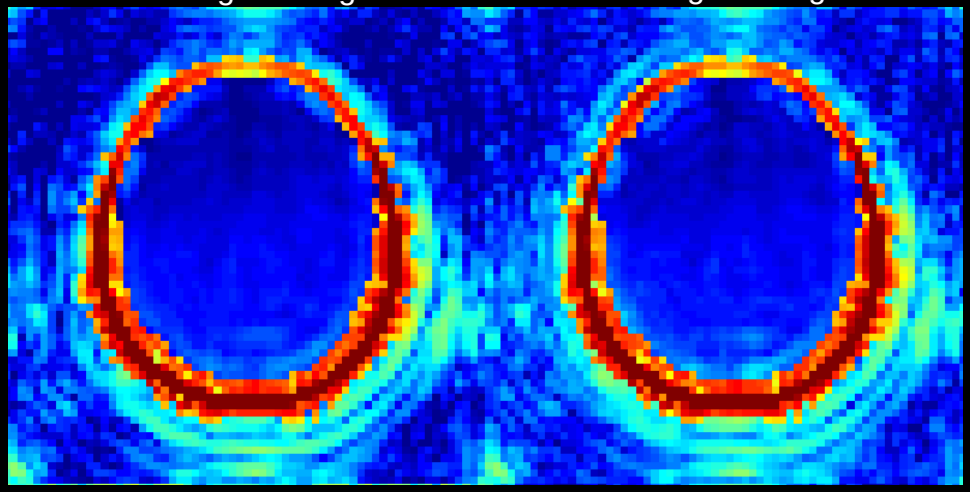
20 avg for metabolites at Nyquist rate

20 avg for lipid at Nyquist rate

Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg_{high}, R_{high}=1

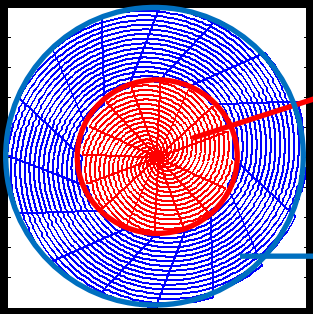
(b) Proposed 1
2 avg_{high}, R_{high}=1



15 dB

-35 dB

Proposed 1 : high-res k-space with 2 avg



20 avg for metabolites at Nyquist rate

2 avg for lipid at Nyquist rate

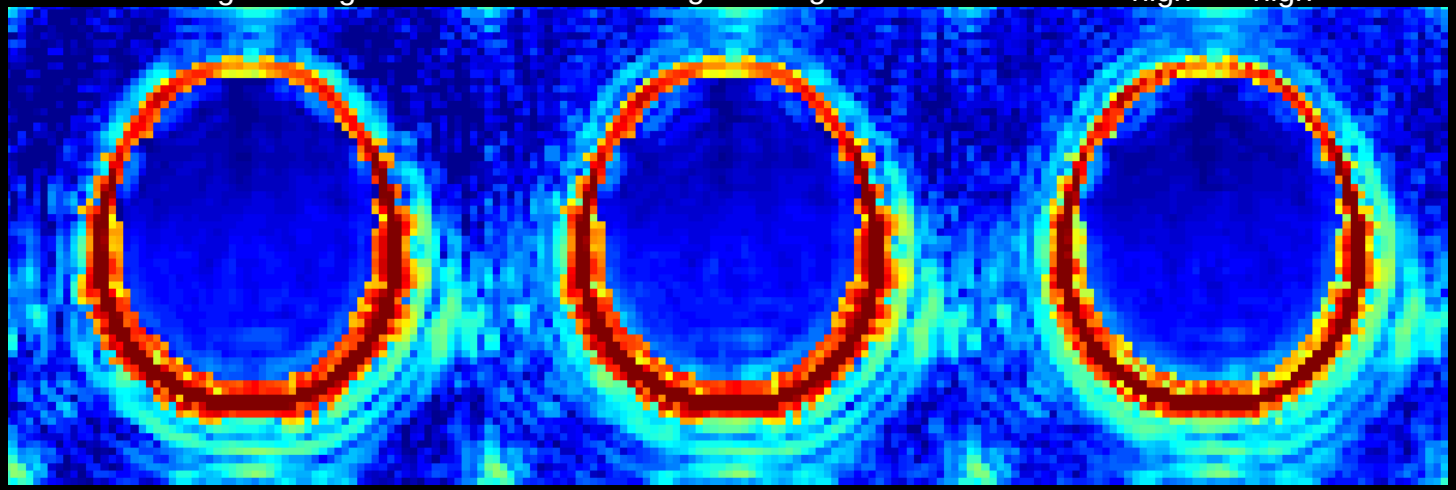
k-space sampling

Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg_{high}, R_{high}=1

(b) Proposed 1
2 avg_{high}, R_{high}=1

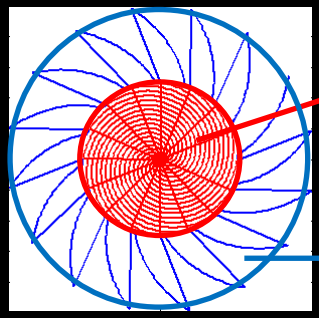
(c) Proposed 2
2 avg_{high}, R_{high}=10



15 dB

-35 dB

Proposed 2 : high-res k-space with 2 avg
10-fold undersampling



20 avg for
metabolites
at Nyquist rate

2 avg for lipid
10-fold undersampling

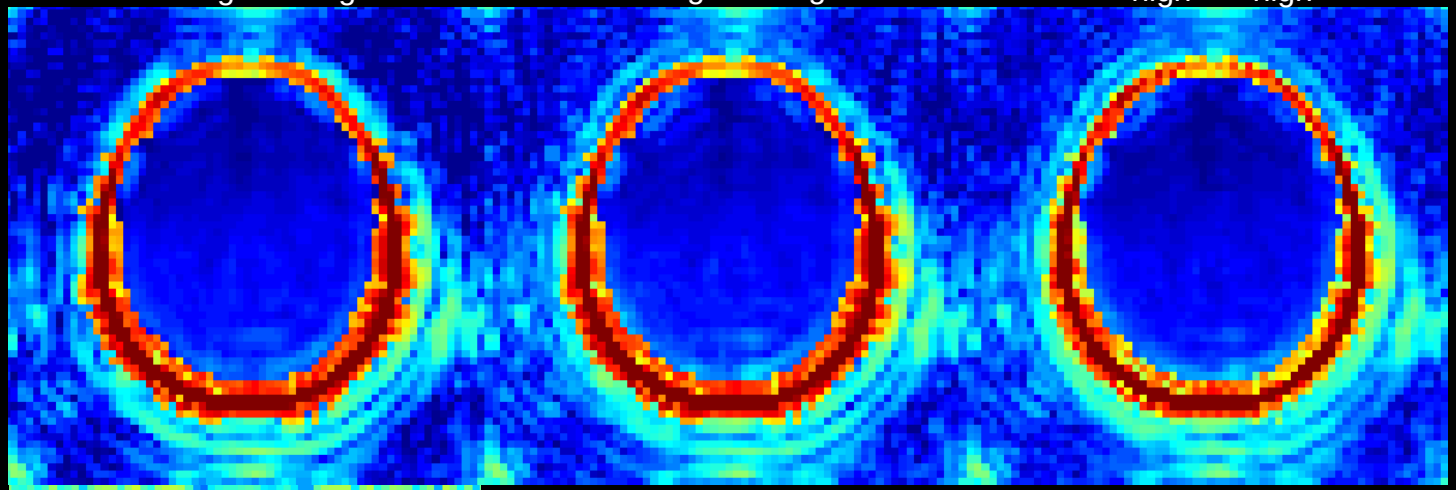
k-space sampling

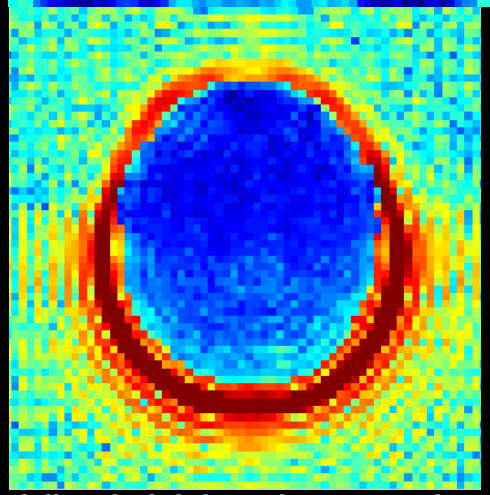
Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg_{high}, R_{high}=1

(b) Proposed 1
2 avg_{high}, R_{high}=1

(c) Proposed 2
2 avg_{high}, R_{high}=10



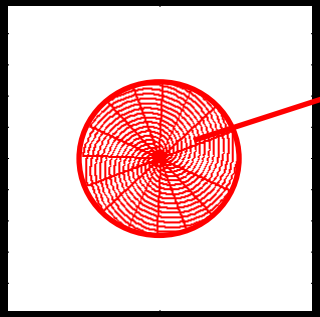


(d) Lipid-basis penalty

15 dB

-35 dB

Lipid-basis applied to 0.5cc data



20 avg for
metabolites
at Nyquist rate

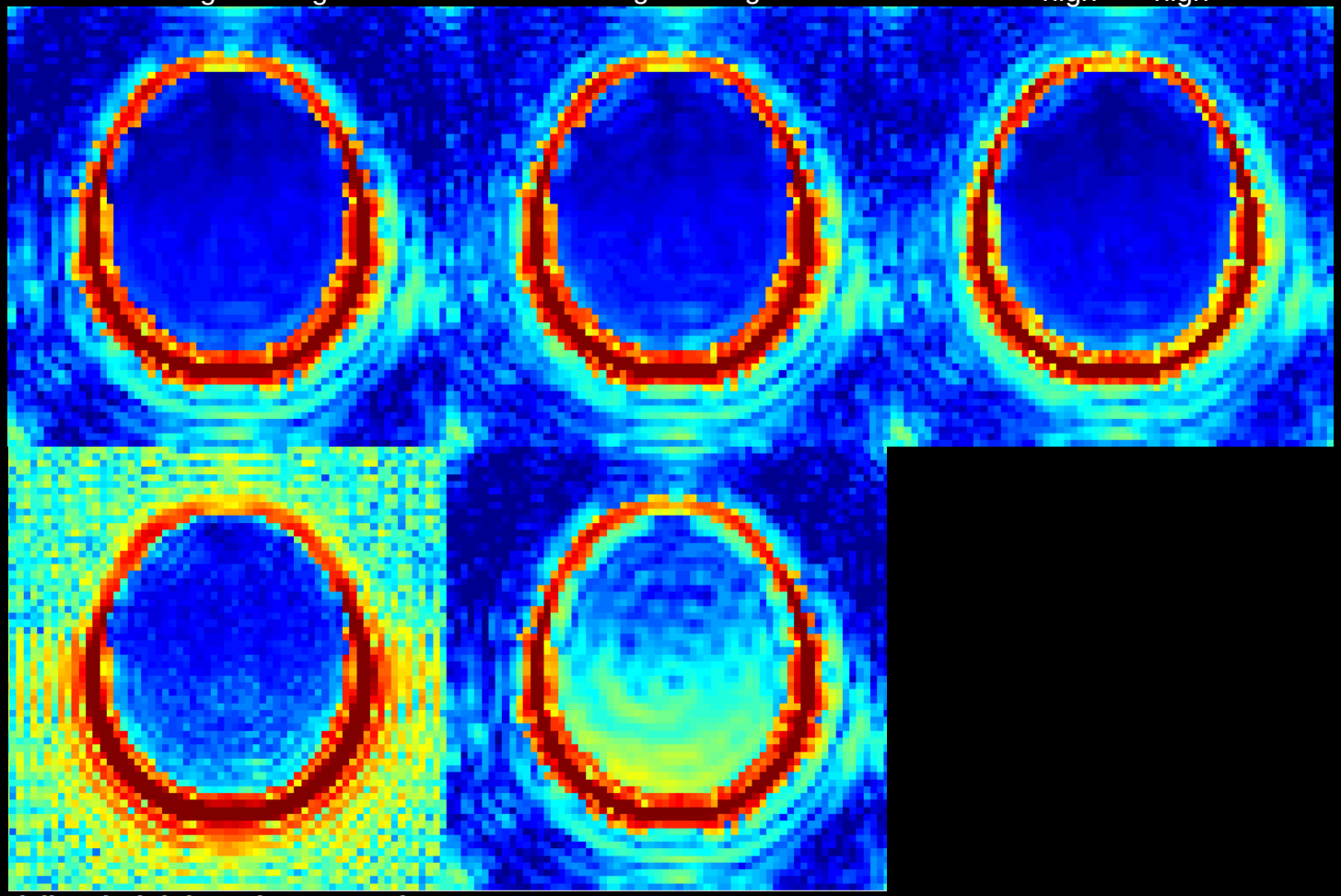
k-space sampling

Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg_{high}, R_{high}=1

(b) Proposed 1
2 avg_{high}, R_{high}=1

(c) Proposed 2
2 avg_{high}, R_{high}=10



15 dB

-35 dB

(d) Lipid-basis penalty

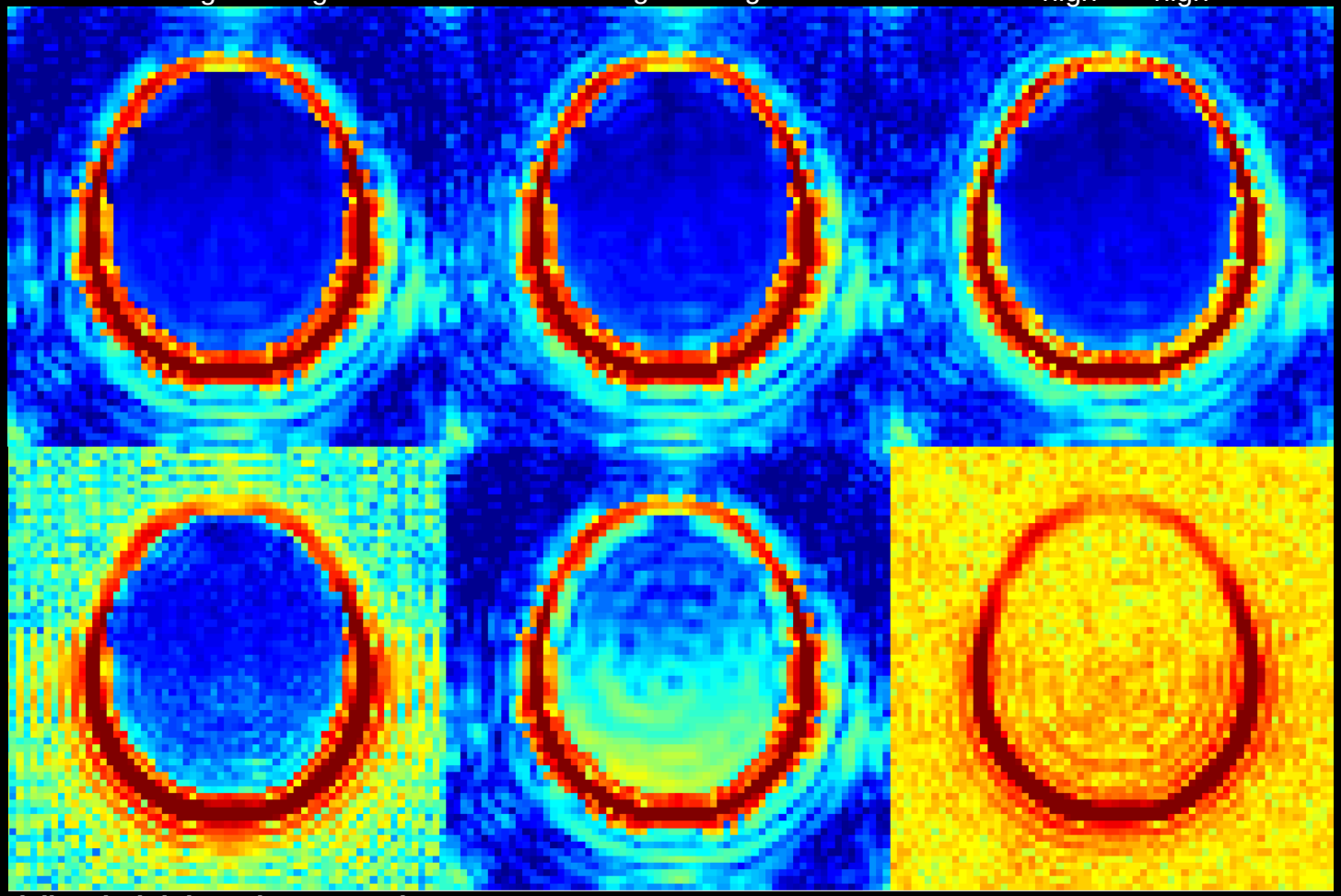
(e) Dual-density

Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg_{high}, R_{high}=1

(b) Proposed 1
2 avg_{high}, R_{high}=1

(c) Proposed 2
2 avg_{high}, R_{high}=10



15 dB

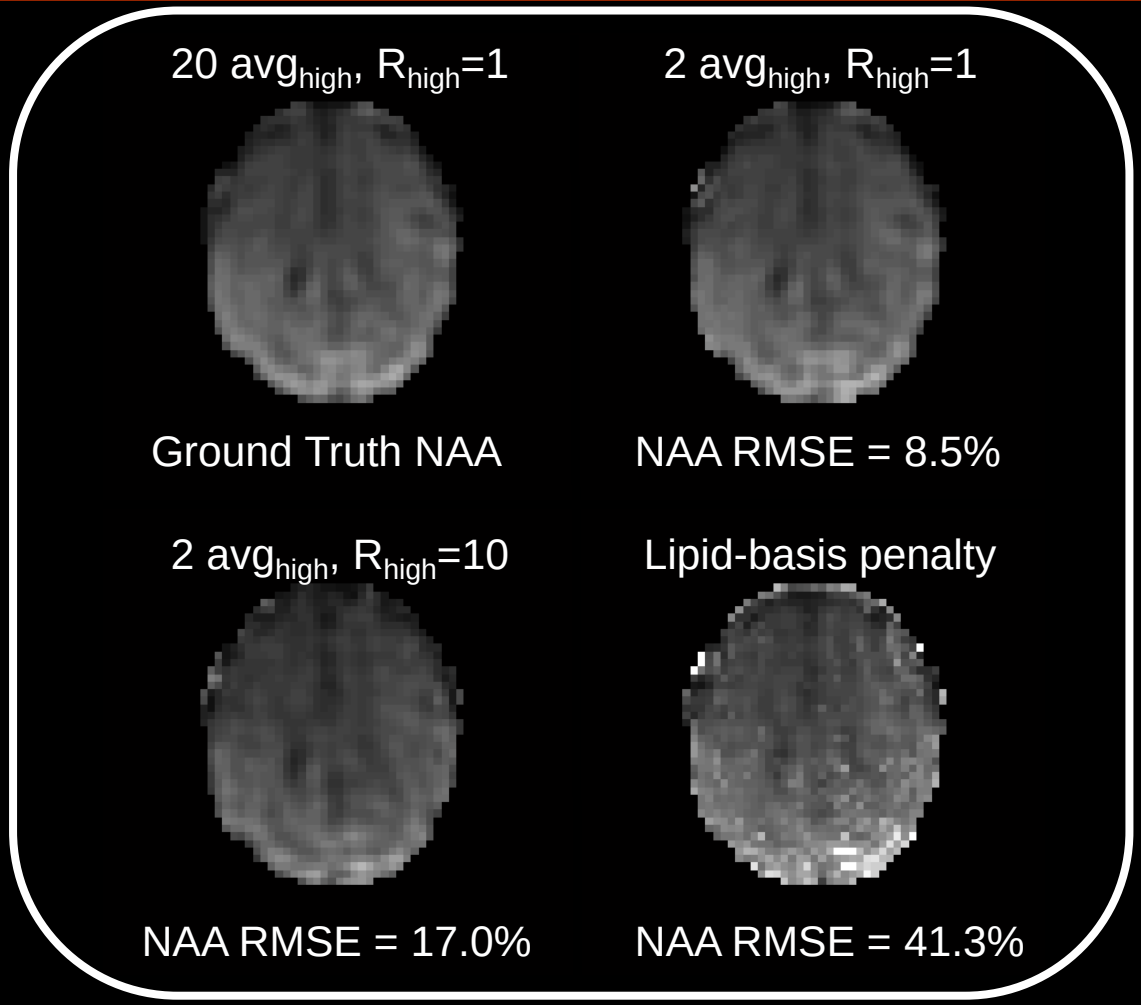
-35 dB

(d) Lipid-basis penalty

(e) Dual-density

(f) No lipid suppression

NAA Maps at TE = 50 ms

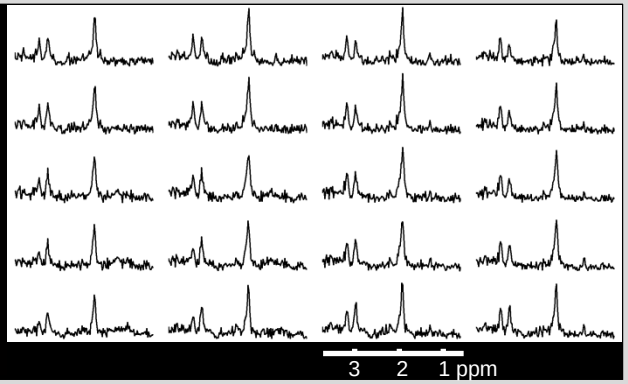


- Taking the NAA map from Gold Standard as reference, proposed methods have 4.9 and 2.4 times less error relative to lipid-basis method

Comparison with Outer Volume Suppression, TE = 50ms

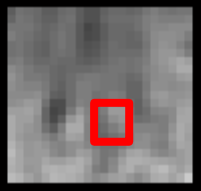
Outer Volume Suppression

Ground Truth NAA

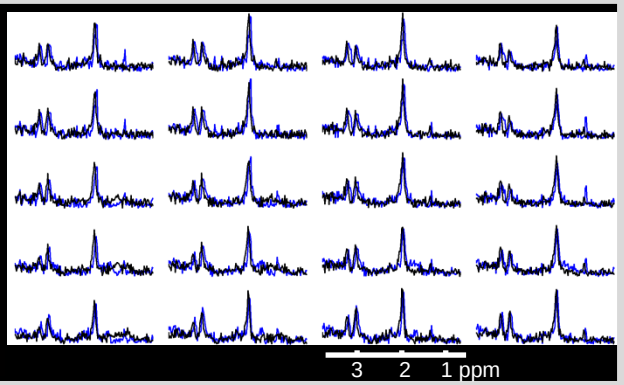


Proposed 1

2 avg_{high}, R_{high}=1

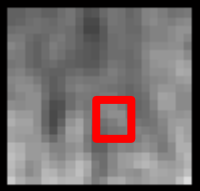


NAA RMSE = 11.5%

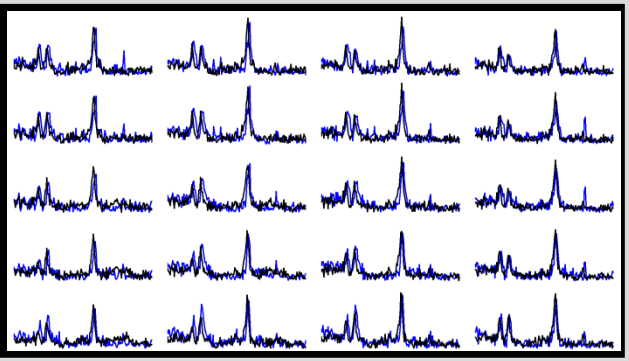


Proposed 2

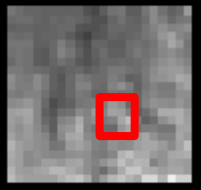
2 avg_{high}, R_{high}=10



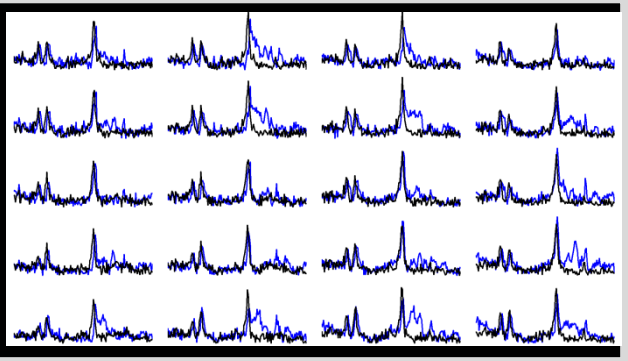
NAA RMSE = 12.9%



Lipid-basis penalty



NAA RMSE = 14.7%

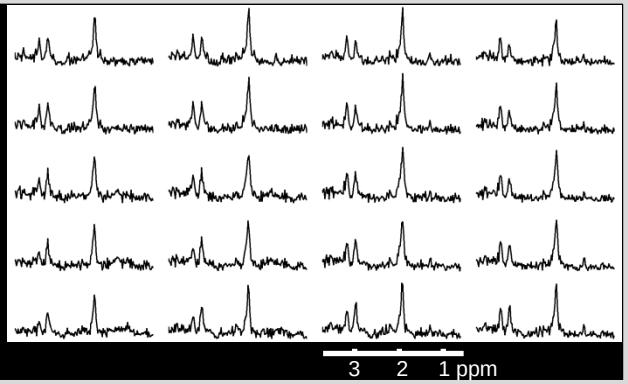


- Spectra from OVS: in black
- Reconstructed spectra: in blue

Comparison with Outer Volume Suppression, TE = 50ms

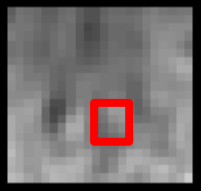
Outer Volume Suppression

Ground Truth NAA

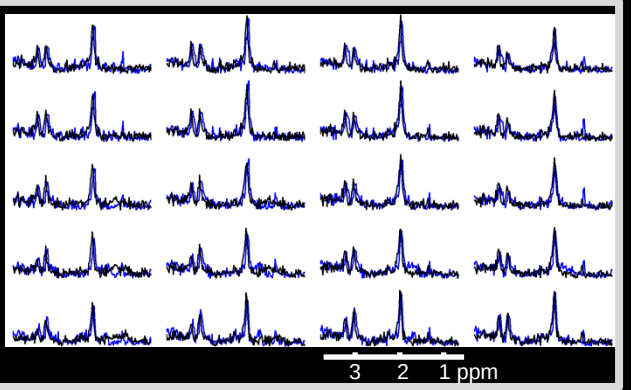


Proposed 1

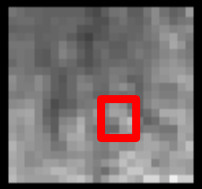
2 avg_{high}, R_{high}=1



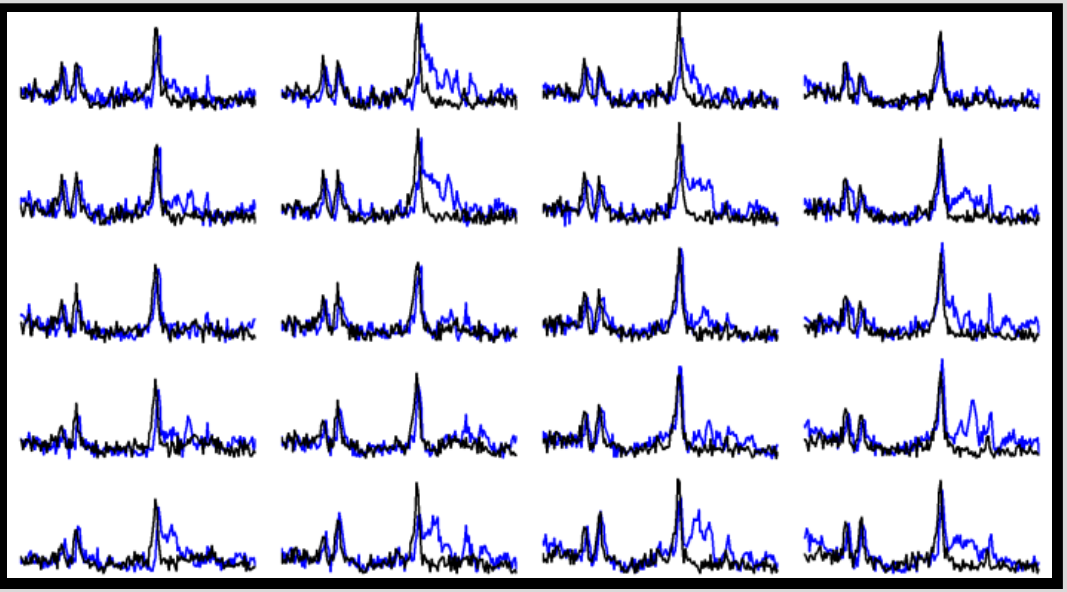
NAA RMSE = 11.5%



Lipid-basis penalty



NAA RMSE = 14.7%

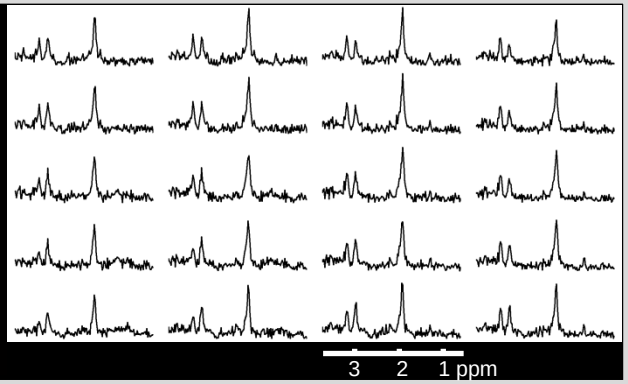


- Spectra from OVS: in black
- Reconstructed spectra: in blue

Comparison with Outer Volume Suppression, TE = 50ms

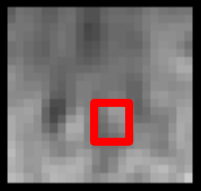
Outer Volume Suppression

Ground Truth NAA

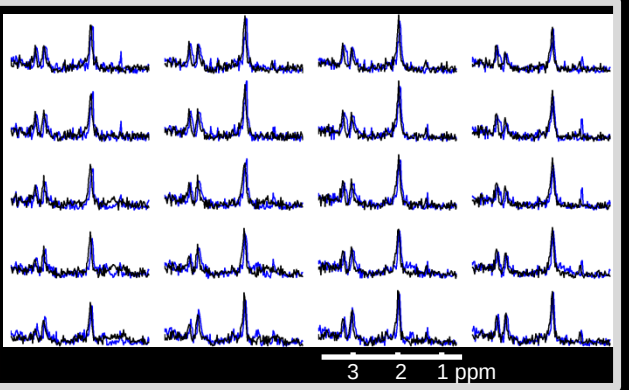


Proposed 1

2 avg_{high}, R_{high}=1

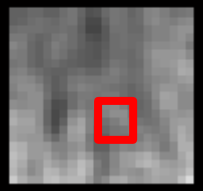


NAA RMSE = 11.5%

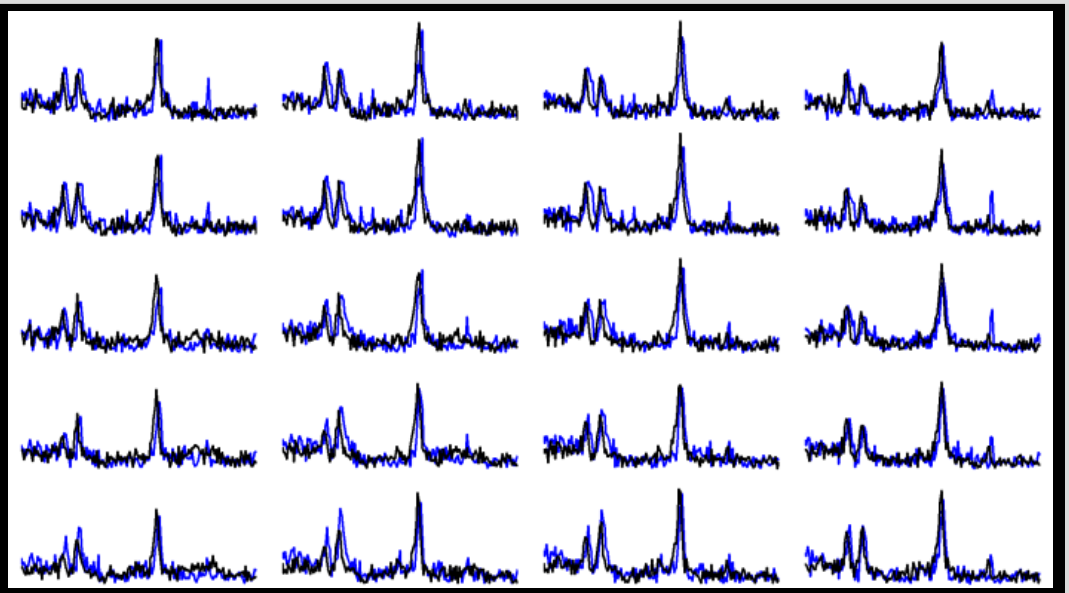


Proposed 2

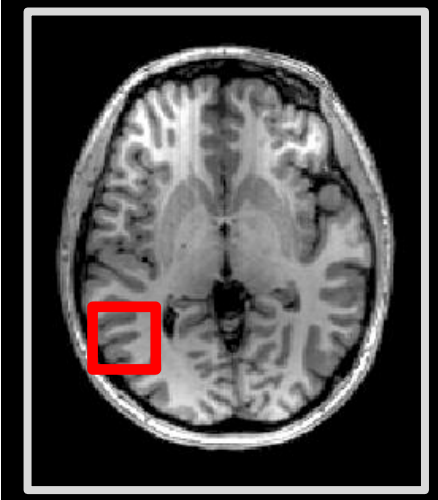
2 avg_{high}, R_{high}=10



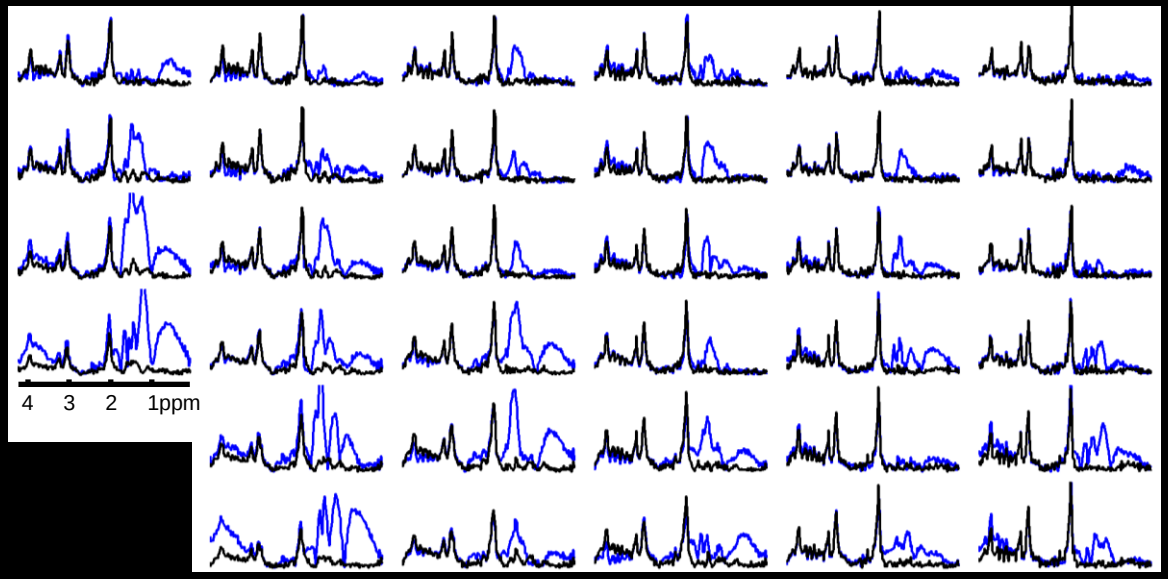
NAA RMSE = 12.9%



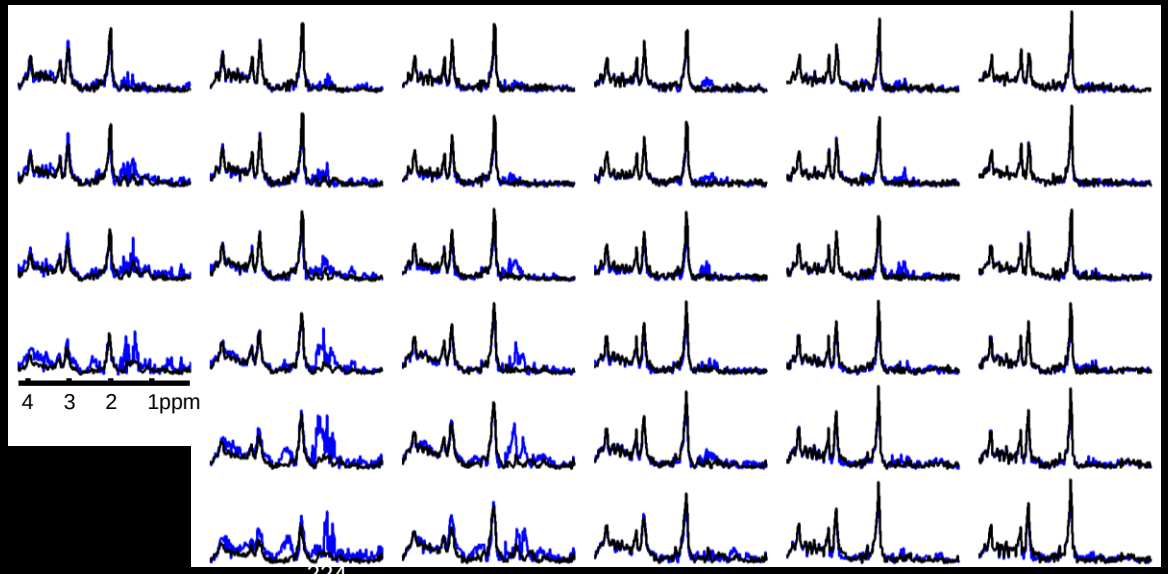
- Spectra from OVS: in black
- Reconstructed spectra: in blue

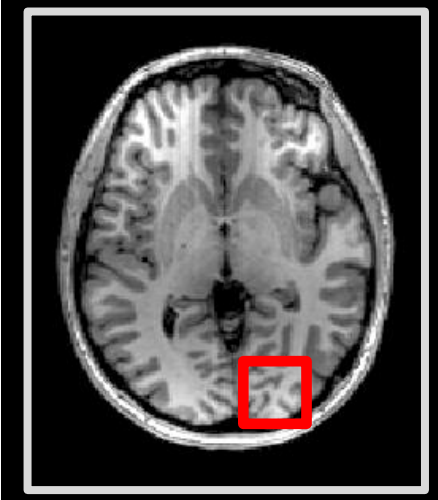


Lipid-basis penalty vs. gold-standard

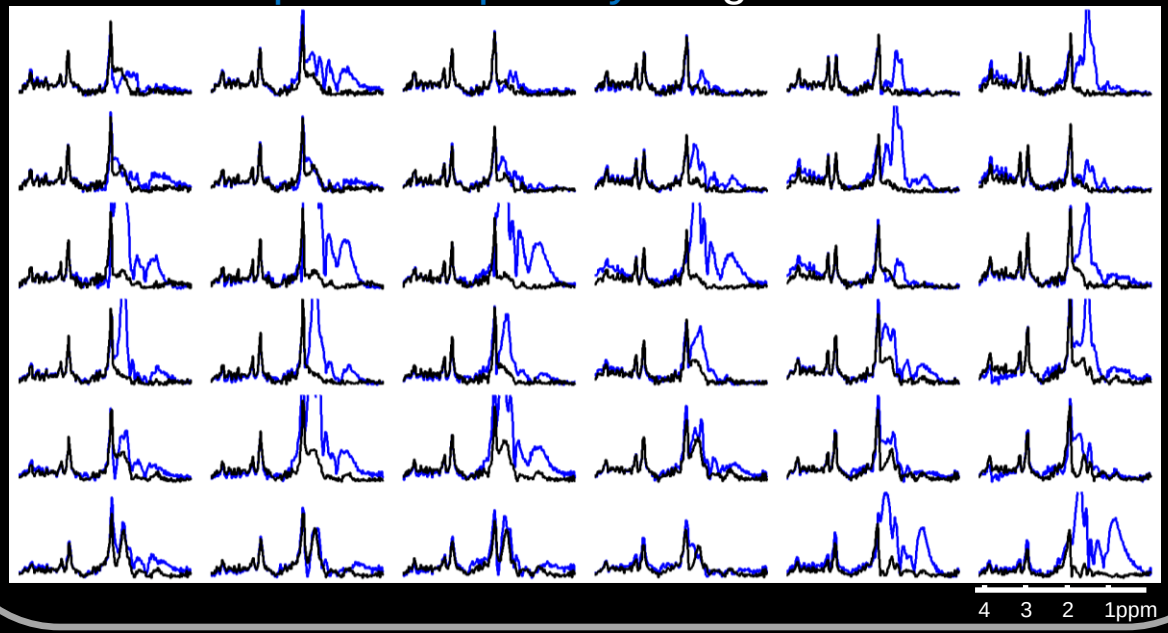


Proposed 2 (2 avg_{high}, R_{high}=10) vs. gold-standard

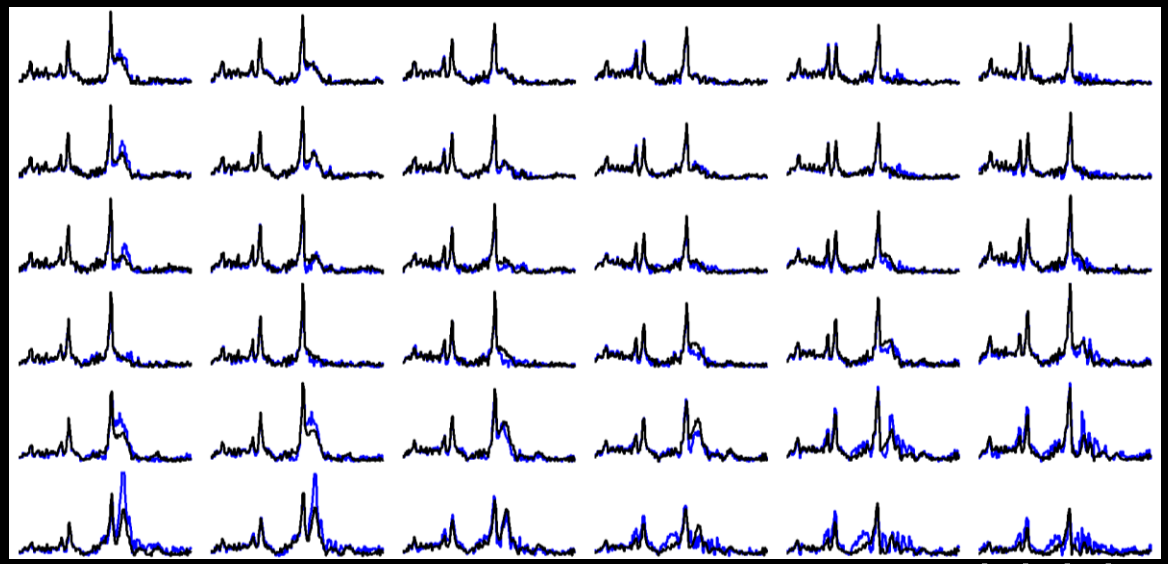


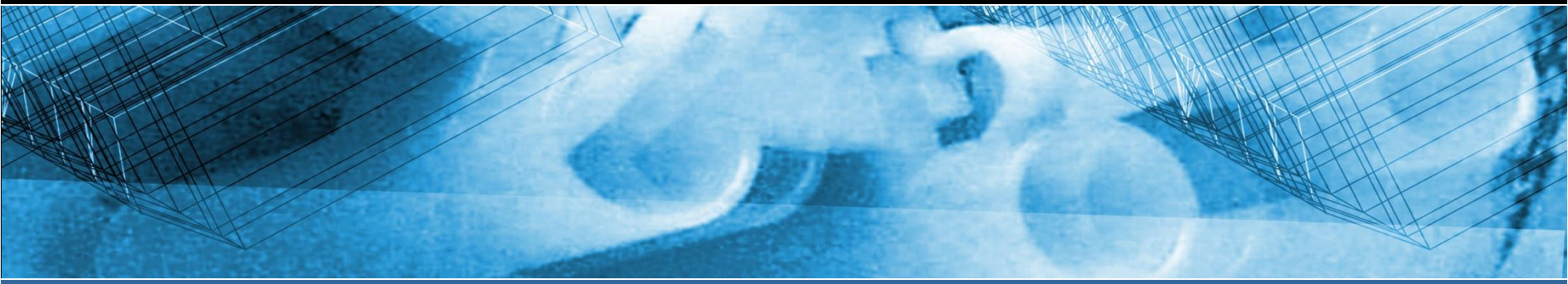


Lipid-basis penalty vs. gold-standard



Proposed 2 (2 avg_{high}, R_{high}=10) vs. gold-standard





Bayesian CS: Marginal prior

Prior on the signal coefficients

- Gradient coefficients are modeled as zero mean Gaussians

$$\mathrm{p}(\delta|\gamma) \sim \mathcal{N}(0, \gamma)$$

this does not constitute a sparse prior

- To promote sparsity, Gamma priors are placed over the variances γ

$$\mathrm{p}(\gamma|a, b) \sim \Gamma(\gamma^{-1}|a, b)$$

Marginal prior on signal coefficients promotes sparsity

- We can marginalize over γ and obtain the *marginal* prior

$$\mathrm{p}(\delta|a,b) = \int \mathrm{p}(\delta|\gamma) \cdot \mathrm{p}(\gamma|a,b) \cdot d\gamma$$

- This turns out to be a Student-*t* distribution. Using a non-informative prior for variances with $a = b = 0$,

$$\mathrm{p}(\delta) \propto \frac{1}{|\delta|}$$