MAGNETIC RESONANCE IMAGING GROUP
Prof. Elfar Adalsteinsson

http://www.rle.mit.edu/mri

- I am a grad student at the MRI group, working with Prof. Elfar Adalsteinsson

- We are affiliated with:

- I have been working on medical image reconstruction using probabilistic analysis & optimization methods.

Outline

1. **Joint Bayesian compressed sensing for multi-contrast reconstruction:**

   reconstruct images with different contrasts from undersampled data

2. **Quantitative Susceptibility Mapping with magnitude prior:**
   estimate tissue iron concentration from MRI signal phase

3. **Estimating brain iron concentration in normal aging using L1-QSM:**

   compare brain iron concentration in young & elderly subjects

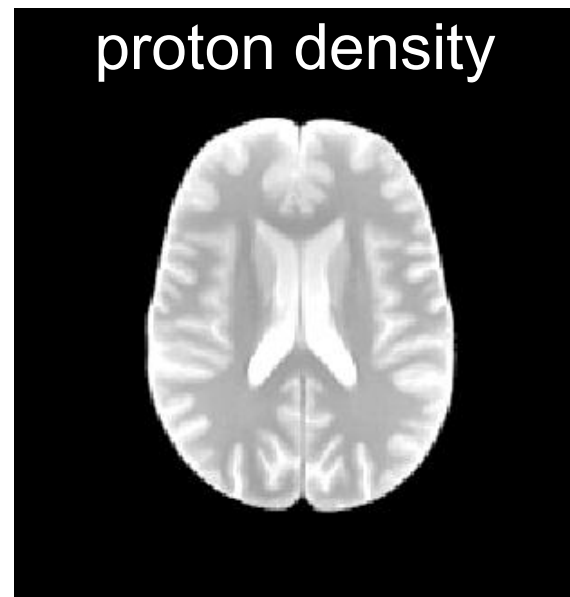# Joint Bayesian Compressed Sensing for Multi-contrast Reconstruction

Berkin Bilgic[1], Vivek K. Goyal[1], Elfar Adalsteinsson[1,2]

[1]EECS, MIT, Cambridge, MA, United States
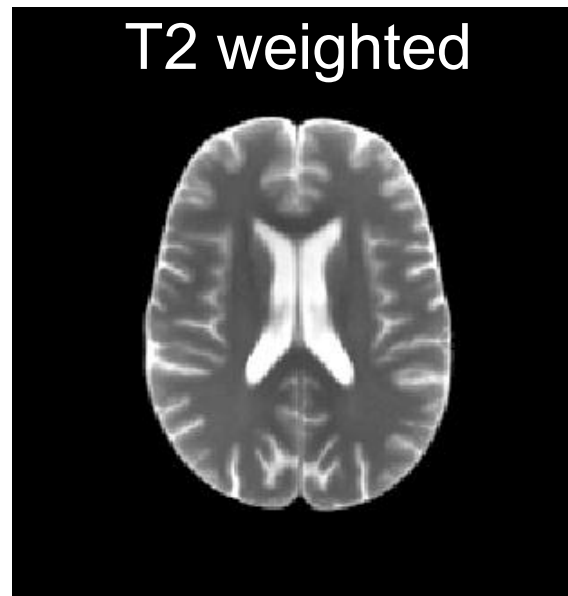[2]Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, United States

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,

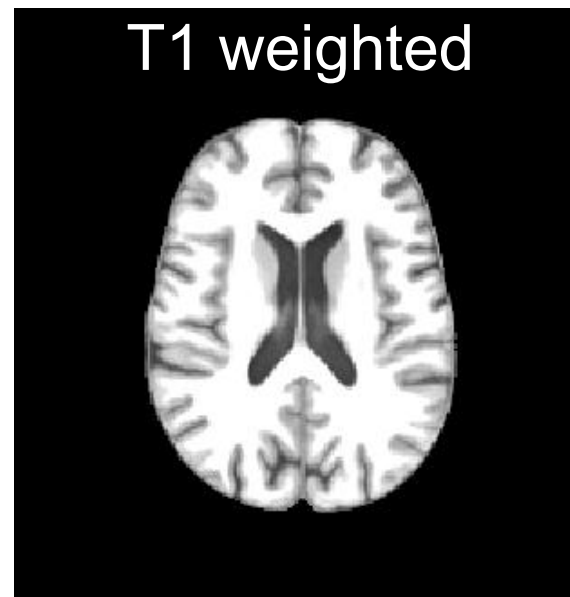
proton density

[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,
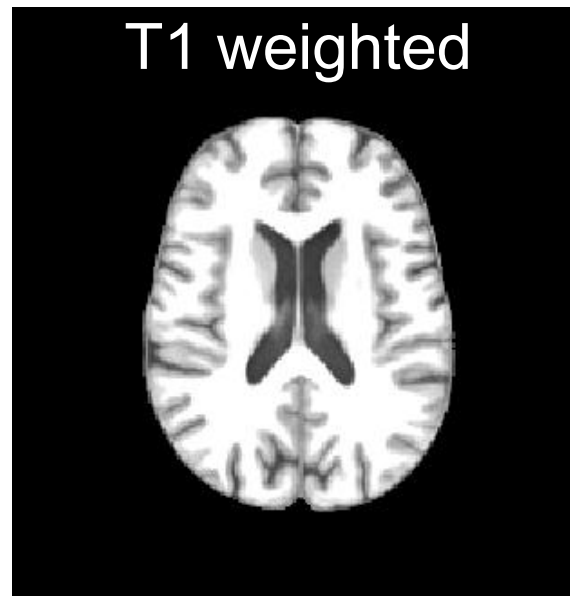

T2 weighted

[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,

T1 weighted

Undersampling the *k*-space

❖ To reduce data acquisition time, it is possible to collect a subset of *k*-space frequencies below the Nyquist rate due to

$$y = \mathbf{F}_\Omega x + n$$

$y \in \mathbb{C}^K$ is the undersampled k - space data,

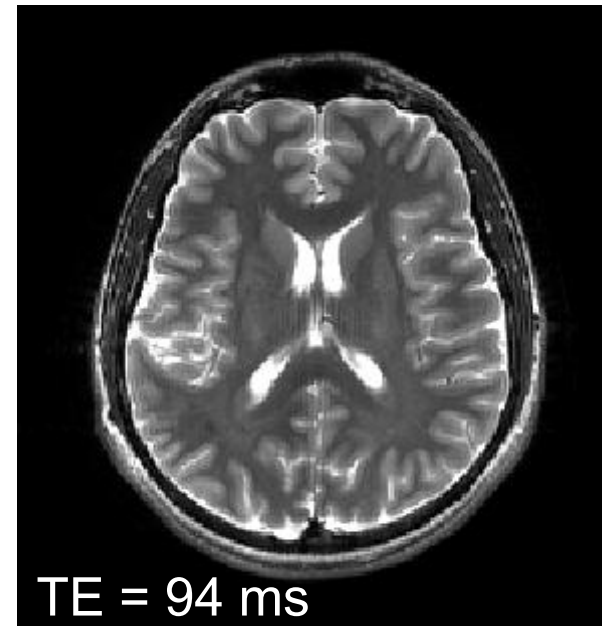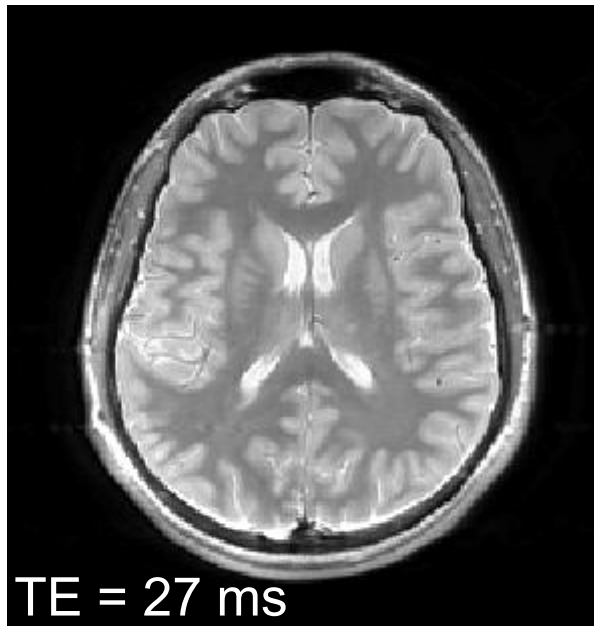$\mathbf{F}_\Omega \in \mathbb{C}^{K \times M}$ is the undersampled 2D - DFT matrix, with $K < M$

$x \in \mathbb{R}^M$ is the spatial image and,

$n \in \mathbb{C}^K$ is the noise in k - space

Undersampling the *k*-space

❖ To reduce data acquisition time, it is possible to collect a subset of *k*-space frequencies below the Nyquist rate due to

$$\boldsymbol{y} = \mathbf{F}_\Omega \boldsymbol{x} + \boldsymbol{n}$$

$\boldsymbol{y} \in \mathbb{C}^K$ is the undersampled k - space data,

$\mathbf{F}_\Omega \in \mathbb{C}^{K \times M}$ is the undersampled 2D - DFT matrix, with $K < M$

$\boldsymbol{x} \in \mathbb{R}^M$ is the spatial image and,

$\boldsymbol{n} \in \mathbb{C}^K$ is the noise in k - space

❖ This work aims to reconstruct multi-contrast data from undersampled acquisitions by making use of

- Bayesian Compressed Sensing theory and,
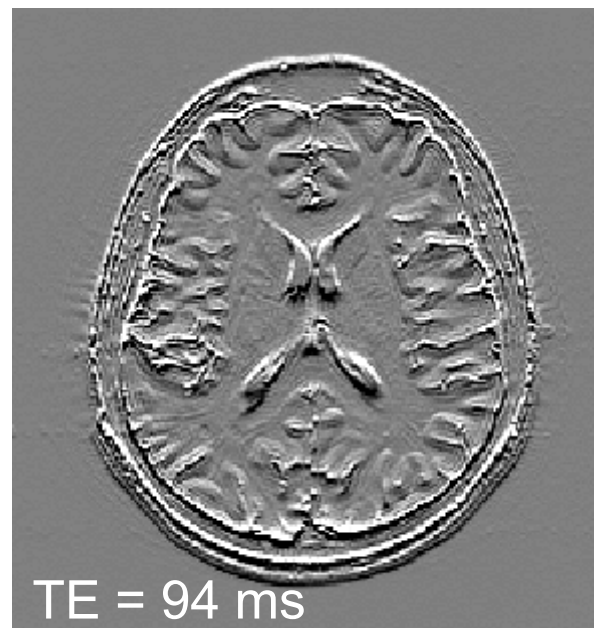- The similarity between the different contrast images.

Similarity of multi-contrast images

❖ Multi-contrast images possess unique properties, e.g. intensity levels at a given voxel

TE = 27 ms

TE = 94 ms

Similarity of multi-contrast images

❖ Multi-contrast images possess unique properties, e.g. intensity levels at a given voxel

❖ At the same time exhibit common features. We make use of the similarity in sparsity support under gradient transform

TE = 27 ms

TE = 94 ms

Similarity of multi-contrast images

❖ Multi-contrast images possess unique properties, e.g. intensity levels at a given voxel

❖ At the same time exhibit common features. We make use of the similarity in sparsity support under gradient transform

TE = 27 ms

TE = 94 ms
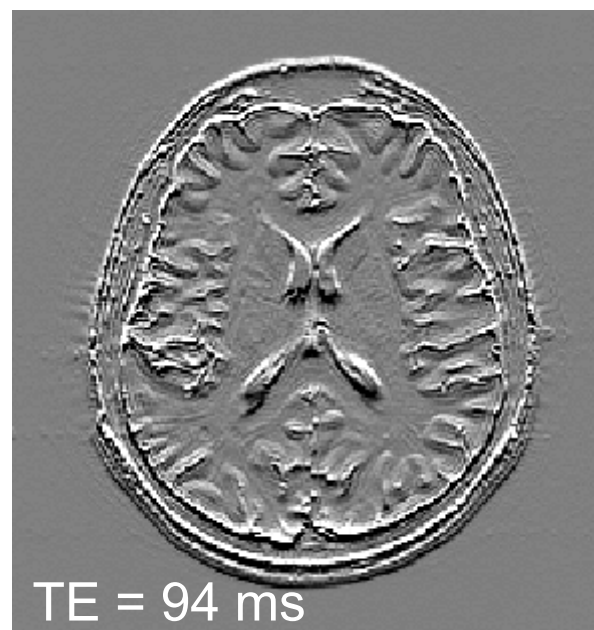
▪ Positions of non-zero coefficients are similar, even though there is no perfect overlap

Joint reconstruction algorithms

❖ We consider two joint reconstruction algorithms,

Multi-contrast Reconstruction

M-FOCUSS

Joint Bayesian CS
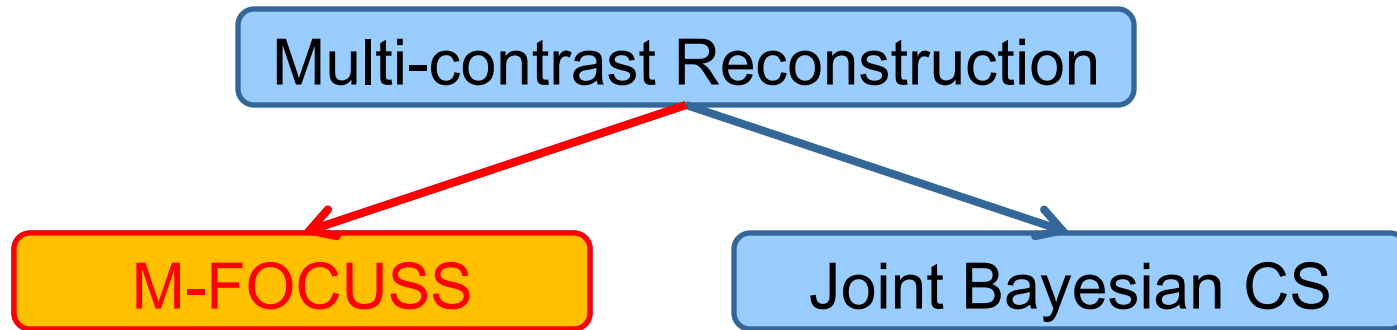
❖ And first introduce the M-FOCUSS method.

M-FOCUSS algorithm

❖ First approach is based on using an existing algorithm, **M-FOCUSS**[1] (**M**ultiple-**FOC**al **U**nderdetermined **S**ystem **S**olver) for joint reconstruction

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

M-FOCUSS algorithm

❖ First approach is based on using an existing algorithm, **M-FOCUSS**[1] (**M**ultiple-**FOC**al **U**nderdetermined **S**ystem **S**olver) for joint reconstruction

❖ M-FOCUSS places an ℓ_1 norm penalty on the gradient coefficients of each image, and an ℓ_2 norm penalty across the multi-contrast images

$$\min_{\boldsymbol{x_i}} \sum_{i=1}^{L} \left\| \mathbf{F}_\Omega \boldsymbol{x_i} - \boldsymbol{y_i} \right\|_2^2 + \lambda \cdot \sum_{j=1}^{M} \left(\sum_{i=1}^{L} \left| \partial \boldsymbol{x_{i,j}} \right|^2 \right)^{1/2}$$

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

M-FOCUSS algorithm

❖ First approach is based on using an existing algorithm, **M-FOCUSS**[1] (**M**ultiple-**FOC**al **U**nderdetermined **S**ystem **S**olver) for joint reconstruction

❖ M-FOCUSS places an ℓ_1 norm penalty on the gradient coefficients of each image, and an ℓ_2 norm penalty across the multi-contrast images

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

❖ As proposed, it is constrained to use the same undersampling pattern for each image

❖ And makes the strict assumption that the sparsity supports of the images are the same.

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

Joint reconstruction algorithms

❖ We consider two joint reconstruction algorithms,

Multi-contrast Reconstruction

M-FOCUSS
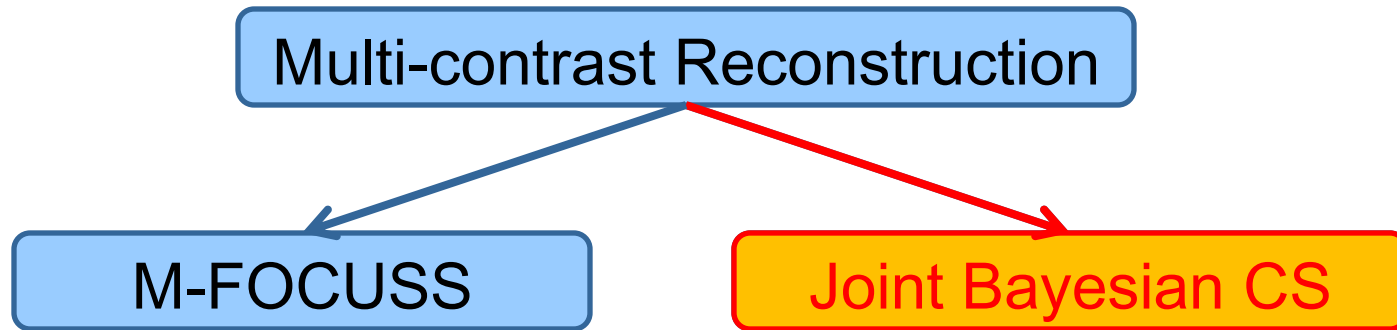
Joint Bayesian CS

❖ Next, we introduce our joint Bayesian reconstruction method.

❖ To obtain a sparse representation of the images $\{x_i\}_{i=1}^L$ with L different contrasts, we augment the undersampled k-space data $\{y_i\}_{i=1}^L$ as

$$\left(1 - e^{-2\pi j\omega/n}\right) \cdot y_i(\omega, \upsilon) = \mathbf{F}_{\Omega_i} \, \delta_i^x \equiv y_i^x$$

$\delta_i^x \in \mathbb{R}^M$ is i^{th} vertical image gradient

$y_i^x \in \mathbb{C}^{K_i}$ is the undersampled k - space data of δ_i^x

Sparse representation and data likelihood

❖ To obtain a sparse representation of the images $\{x_i\}_{i=1}^{L}$ with L different contrasts, we augment the undersampled k-space data $\{y_i\}_{i=1}^{L}$ as

$$\left(1 - e^{-2\pi j \omega/n}\right) \cdot y_i(\omega, \upsilon) = \mathbf{F}_{\Omega_i} \, \delta_i^x \equiv y_i^x$$

$\delta_i^x \in \mathbb{R}^M$ is i^{th} vertical image gradient

$y_i^x \in \mathbb{C}^{K_i}$ is the undersampled k-space data of δ_i^x

❖ Modeling the k-space noise to be Gaussian with zero mean and variance σ^2, the likelihood of observing the data becomes

$$Y_i^x = \left[\mathcal{R}e\left(y_i^x\right), \mathcal{I}m\left(y_i^x\right)\right]^T$$

$$\Phi_i = \left[\mathcal{R}e\left(\mathbf{F}_{\Omega_i}\right), \mathcal{I}m\left(\mathbf{F}_{\Omega_i}\right)\right]^T$$

$$p\left(Y_i^x / \delta_i^x, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-K_i} \exp\left(-\left\|Y_i^x - \Phi_i \delta_i^x\right\|_2^2 / 2\sigma^2\right)$$

Bayesian analysis for joint inference

❖ Next, we would like to impose a sparsity promoting prior distribution over the image gradients $\left\{\delta_i^x\right\}_{i=1}^L$ and $\left\{\delta_i^y\right\}_{i=1}^L$,

❖ And compute their posterior distribution with the Bayes' rule using this prior, the likelihood term and the observed *k*-space data $\left\{Y_i^x\right\}_{i=1}^L$ and $\left\{Y_i^y\right\}_{i=1}^L$

❖ At the same time, we would like to enable information sharing across the multi-contrast images.

Bayesian analysis for joint inference

❖ Next, we would like to impose a sparsity promoting prior distribution over the image gradients $\{\delta_i^x\}_{i=1}^{L}$ and $\{\delta_i^y\}_{i=1}^{L}$,

❖ And compute their posterior distribution with the Bayes' rule using this prior, the likelihood term and the observed *k*-space data $\{Y_i^x\}_{i=1}^{L}$ and $\{Y_i^y\}_{i=1}^{L}$

❖ At the same time, we would like to enable information sharing across the multi-contrast images.

❖ To this end, we carry out the inference within a hierarchical Bayesian model[1]

[1] Ji *et al*. IEEE T Signal Proces, 2009

Hierarchical Bayesian Model for joint inference

❖ At the bottom layer, we have the undersampled *k*-space observations, which are jointly parameterized by the hyperparameters on the layer above.

hyperparameters

coupled by hyperparameters α and $\alpha_0 = \sigma^{-2}$

k-space observations

Hierarchical Bayesian Model for joint inference

❖ At the bottom layer, we have the undersampled *k*-space observations, which are jointly parameterized by the hyperparameters on the layer above.

❖ We capture the similarity in the gradient domain by defining the hyperparameters *α* over the L gradient images

coupled by hyperparameters *α* and $\alpha_0 = \sigma^{-2}$

k-space observations

Hierarchical Bayesian Model for joint inference

❖ At the bottom layer, we have the undersampled *k*-space observations, which are jointly parameterized by the hyperparameters on the layer above.

❖ We capture the similarity in the gradient domain by defining the hyperparameters α over the L gradient images

❖ The hyperparameters are in turn controlled by the hyper-priors at the top level.

hyper-priors a, b, c, d

hyperparameters α $\alpha_0 = \sigma^{-2}$

δ_1^x N_1 δ_i^x δ_L^x N_L
Φ_1 + Φ_i + Φ_L +
Y_1 Y_i Y_L

Image 1   Image i   Image L

controlled by a, b, c, d

coupled by hyperparameters α and $\alpha_0 = \sigma^{-2}$

k-space observations

Prior on the signal coefficients

❖ The gradient coefficients are modeled to be drawn from a product of zero mean Gaussians

$$p\left(\delta_i^x \mid \boldsymbol{\alpha}\right) = \prod_{j=1}^{M} \mathcal{N}\left(\delta_{i,j}^x \mid 0, \alpha_j^{-1}\right)$$

and the precisions of the Gaussians are determined by $\boldsymbol{\alpha} \in \mathbb{R}^M$

❖ And Gamma priors are placed over the hyperparameters $\boldsymbol{\alpha}$

$$p(\boldsymbol{\alpha} \mid c, d) = \prod_{j=1}^{M} Ga(\alpha_j \mid c, d) \quad \text{where} \quad Ga\left(\alpha_j / c, d\right) = \frac{d^c}{\Gamma(c)} \alpha_j^{c-1} exp\left(-d\alpha_j\right)$$

Prior on the signal coefficients

❖ The gradient coefficients are modeled to be drawn from a product of zero mean Gaussians

$$p\left(\delta_i^x \mid \boldsymbol{\alpha}\right) = \prod_{j=1}^{M} \mathcal{N}\left(\delta_{i,j}^x \mid 0, \alpha_j^{-1}\right)$$

and the precisions of the Gaussians are determined by $\boldsymbol{\alpha} \in \mathbb{R}^M$

❖ And Gamma priors are placed over the hyperparameters $\boldsymbol{\alpha}$

$$p(\boldsymbol{\alpha} \mid c,d) = \prod_{j=1}^{M} Ga(\alpha_j \mid c,d) \qquad \text{where} \quad Ga\left(\alpha_j / c,d\right) = \frac{d^c}{\Gamma(c)} \alpha_j^{c-1} exp\left(-d\alpha_j\right)$$

❖ We can marginalize over the hyperparameters $\boldsymbol{\alpha}$ and obtain the *marginal* prior that enforces sparsity

$$p(\delta_{i,j}^x) = \int p(\delta_{i,j}^x / \alpha_j) p(\alpha_j \mid c,d) d\alpha_j$$
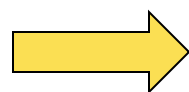
sharp peak at 0

$$p(\delta_{i,j}^x) \propto \frac{1}{|\delta_{i,j}^x|}$$

Student-*t*
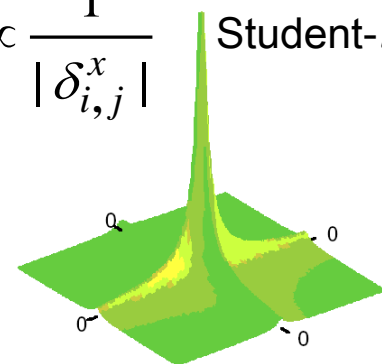
$c,d = 0$

26

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

$$p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0) = \frac{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)\, p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

$$\underbrace{p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0)}_{\text{posterior}} = \frac{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)\, p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

likelihood

posterior

$$\boxed{p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0)} = \frac{\boxed{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)} p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

$$\underbrace{p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0)}_{\text{posterior}} = \frac{\overbrace{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)}^{\text{likelihood}} \overbrace{p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}^{\text{prior}}}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

$$p(\delta_i^x \mid Y_i^x, \boldsymbol{\alpha}, \alpha_0) = \frac{\overset{\text{Gaussian}}{\boxed{p(Y_i^x \mid \delta_i^x, \alpha_0)}}\, \overset{\text{Gaussian}}{\boxed{p(\delta_i^x \mid \boldsymbol{\alpha})}}}{p(Y_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

also Gaussian

Gaussian   Gaussian

$$p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0) = \frac{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)\, p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

Computing the posterior for the signals

❖ Since the data likelihood and the signal prior are both Gaussian, the posterior for the gradient coefficients is also in the same family,

Gaussian   Gaussian

also Gaussian

$$p(\boldsymbol{\delta}_i^x \mid \boldsymbol{Y}_i^x, \boldsymbol{\alpha}, \alpha_0) = \frac{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\delta}_i^x, \alpha_0)\, p(\boldsymbol{\delta}_i^x \mid \boldsymbol{\alpha})}{p(\boldsymbol{Y}_i^x \mid \boldsymbol{\alpha}, \alpha_0)}$$

$$\boldsymbol{\delta}_i^x \approx \mathcal{N}\!\left(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\right)$$

$$\boldsymbol{\mu}_i = \alpha_0\,\boldsymbol{\Sigma}_i\,\boldsymbol{\Phi}_i^T\,\boldsymbol{Y}_i^x$$

$$\boldsymbol{\Sigma}_i = (\alpha_0\,\boldsymbol{\Phi}_i^T\,\boldsymbol{\Phi}_i + \mathbf{A})^{-1}$$

$$\mathbf{A} = diag(\alpha_1, \alpha_2, ..., \alpha_M)$$

We only need to estimate the α_i's

❖ We seek point estimates for the hyperparameters $\boldsymbol{\alpha}$ and α_0 in a maximum likelihood framework.

$$\max_{\boldsymbol{\alpha},\alpha_0}\boldsymbol{\mathcal{L}}(\boldsymbol{\alpha},\alpha_0)=\max_{\boldsymbol{\alpha},\alpha_0}\sum_{i=1}^{L}\log\ p\left(\boldsymbol{Y}_i^x\mid\boldsymbol{\alpha},\alpha_0\right)$$

❖ Summation over the L images enables **information sharing** while estimating the hyperparameters.

❖ Once the hyperparameters are estimated, the posterior for the gradient coefficients $\boldsymbol{\delta}_i^x$ is determined based only on its related *k*-space data $\boldsymbol{Y}_i^x$ due to,

$$\boldsymbol{\mu}_i=\alpha_0\,\boldsymbol{\Sigma}_i\boldsymbol{\Phi}_i^T\,\boldsymbol{Y}_i^x$$

❖ After estimating the vertical and horizontal gradients $\left\{\delta_i^x\right\}_{i=1}^{L}$ and $\left\{\delta_i^y\right\}_{i=1}^{L}$, we seek the images $\left\{x_i\right\}_{i=1}^{L}$ consistent with these and the *k*-space data $\left\{y_i\right\}_{i=1}^{L}$ in a Least Squares setting,

$$\hat{x}_i = \underset{x_i}{argmin} \ \left\|\partial_x x_i - \delta_i^x\right\|_2^2 + \left\|\partial_y x_i - \delta_i^y\right\|_2^2 + \lambda\left\|\mathbf{F}_{\Omega_i} x_i - y_i\right\|_2^2$$

$$for \ \ i = 1,...,L$$

where ∂_x and ∂_y are vertical and horizontal gradient operators

SRI24 Atlas
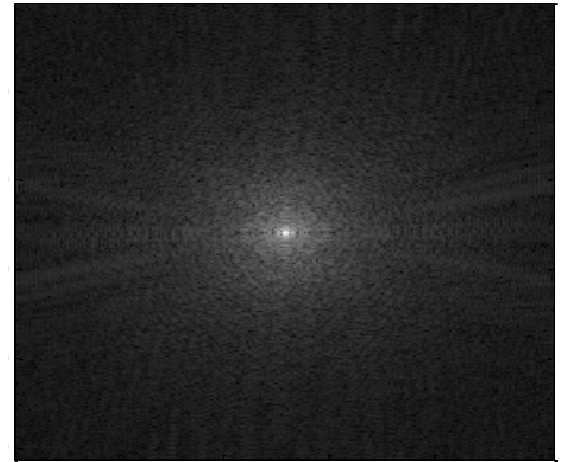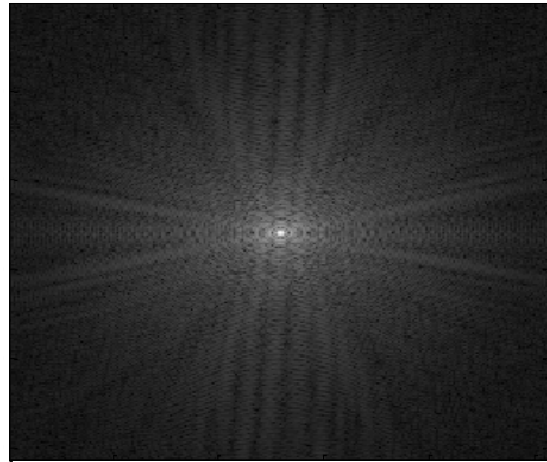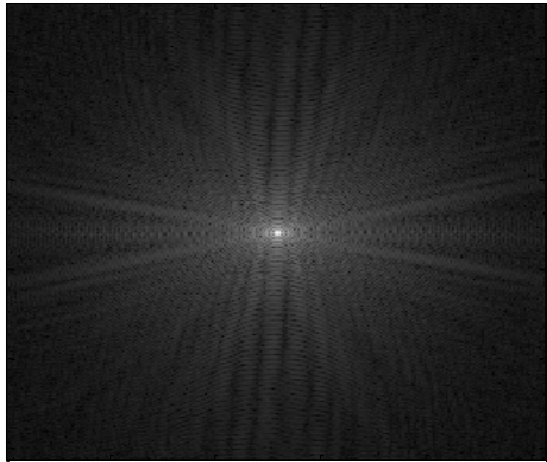
k-space, 100 % of Nyquist rate

Inverse FFT   Error: 0 % RMSE

Lustig *et al.* [yellow bar] 9.4 %
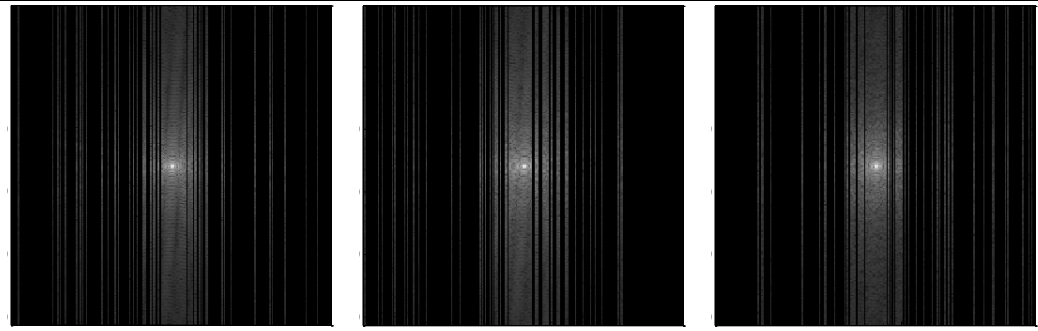
k-space, 25 % of Nyquist rate

Lustig *et al.*[1]   Error: 9.4 % RMSE
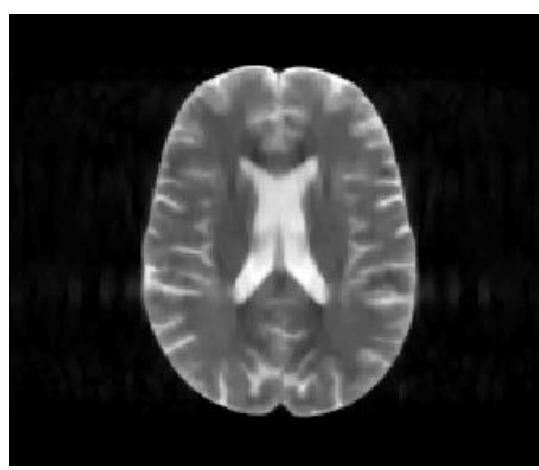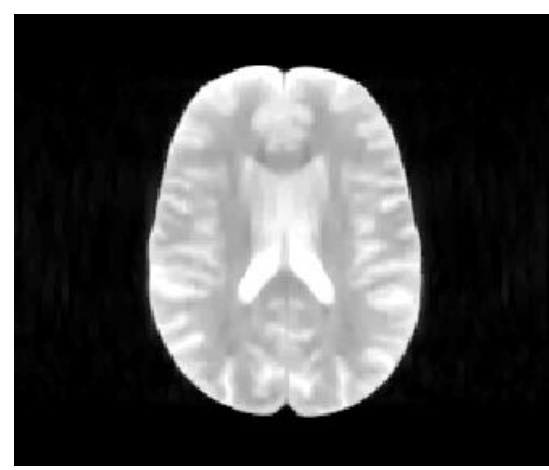
0.7

0
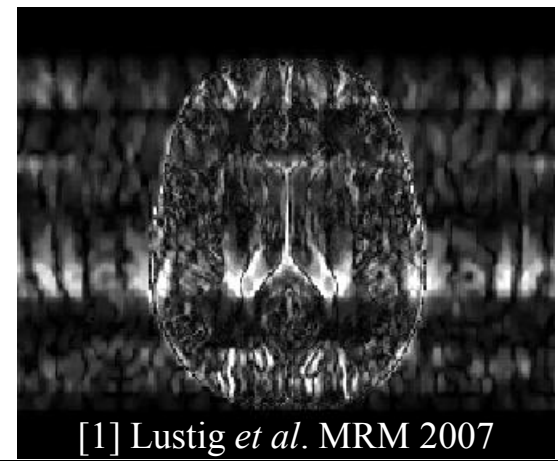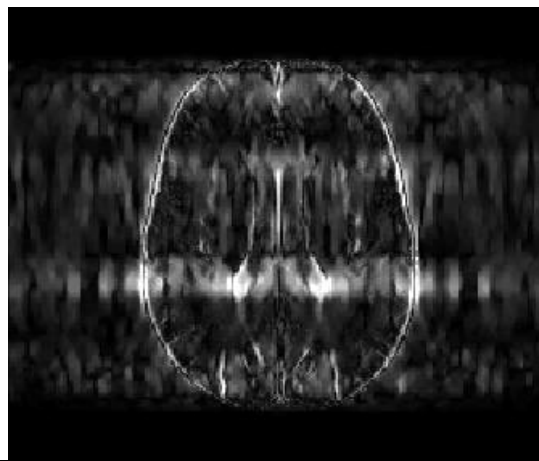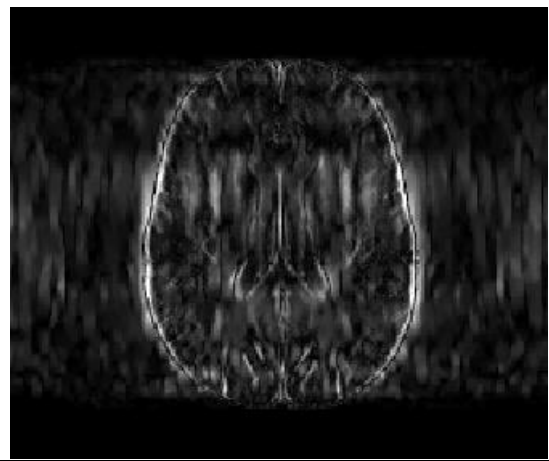
0.15

0

[1] Lustig *et al*. MRM 2007

k-space, 25 % of Nyquist rate

Same undersampling pattern

M-FOCUSS[1]   Error: 3.2 % RMSE

0.7

0

0.15

0

[1] Cotter *et al*. T Signal Proces, 2005
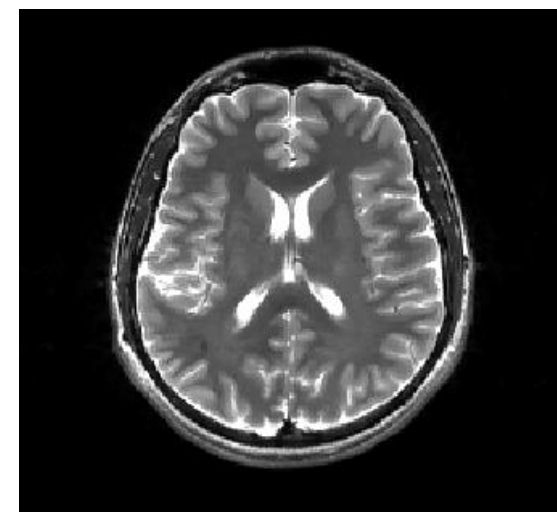
Different undersampling patterns
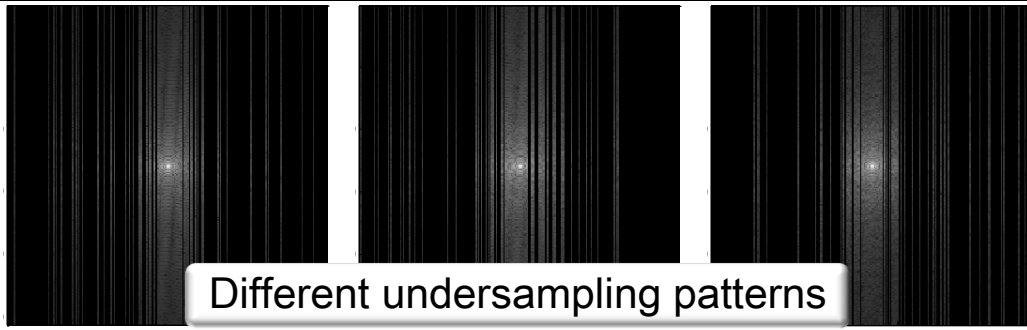
Lustig *et al.* 9.4 %
M-FOCUSS 3.2 %
Joint Bayes 2.3 %

k-space, 25 % of Nyquist rate

Our Bayesian CS   Error: 2.3 % RMSE
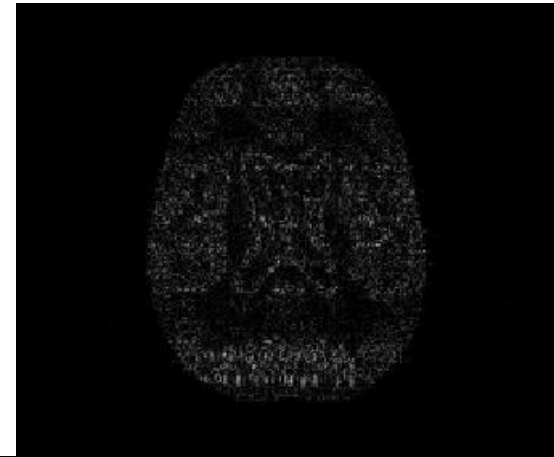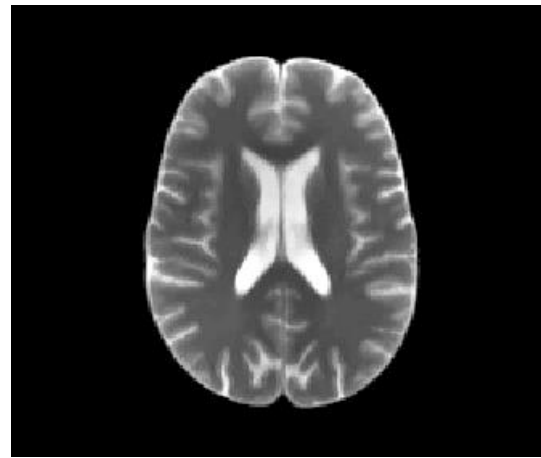
TSE Scans : *in vivo* acquisition

k-space
100 % of Nyquist rate

Inverse FFT

Error: 0 % RMSE

0.6

0

Lustig *et al.* ▬ 9.4 %

k-space, 40 % of Nyquist rate

Lustig *et al.*[1]

0.6

0

Error: 9.4 % RMSE

0.1

[1] Lustig *et al.* MRM 2007

0
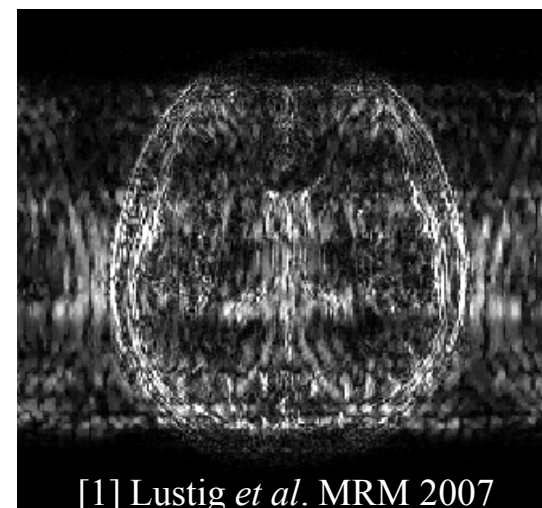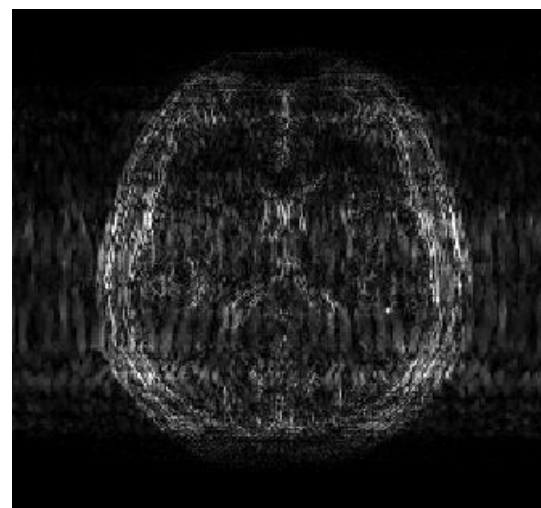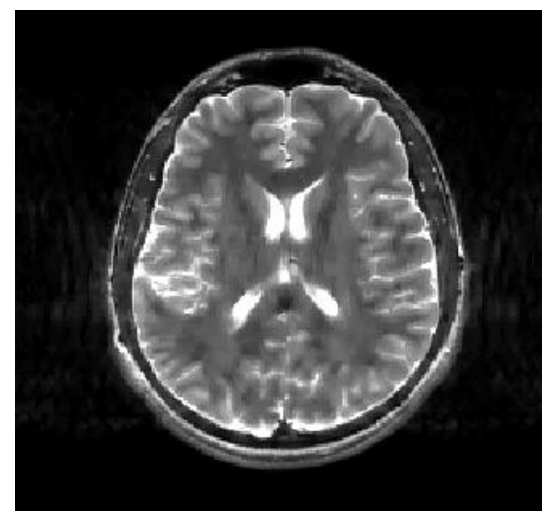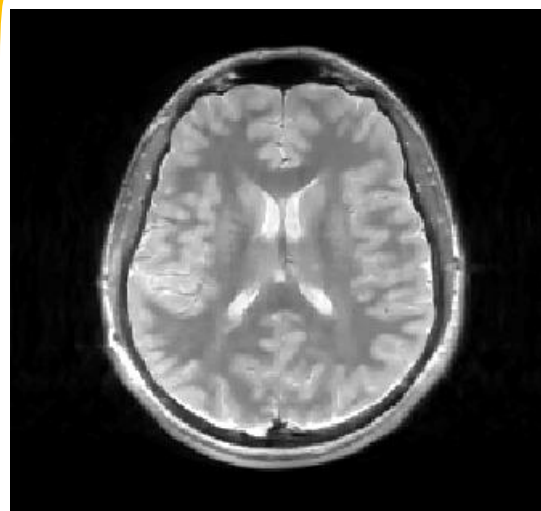
Lustig *et al.* 9.4 %

M-FOCUSS 5.1 %

k-space, 40 % of Nyquist rate

Same undersampling pattern

M-FOCUSS[1]

Error: 5.1 % RMSE

0.6

0

0.1

0

[1] Cotter *et al*. T Signal Proces, 2005
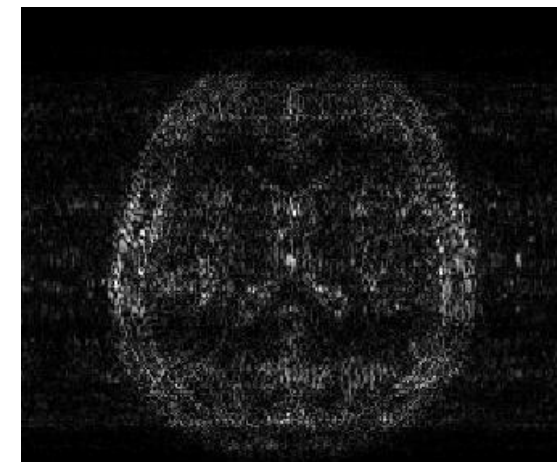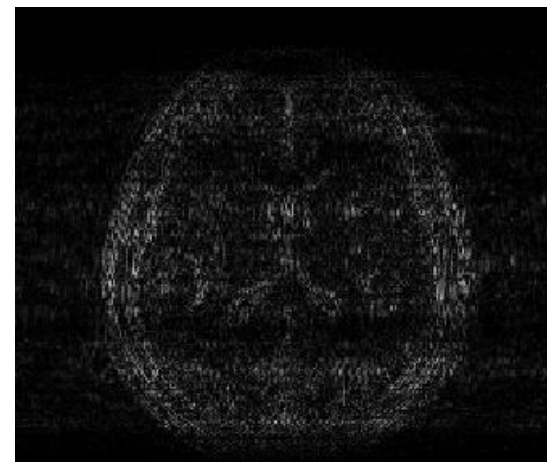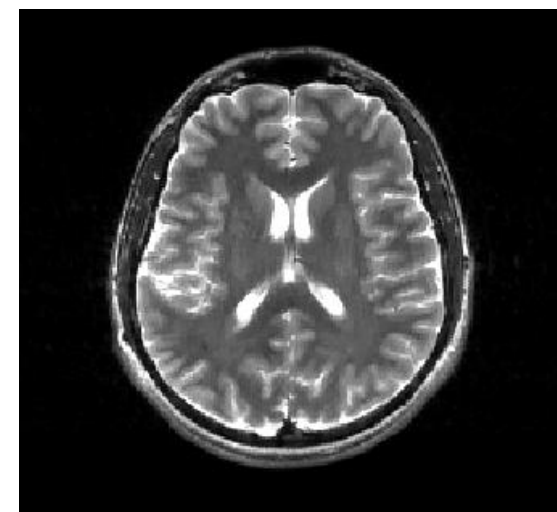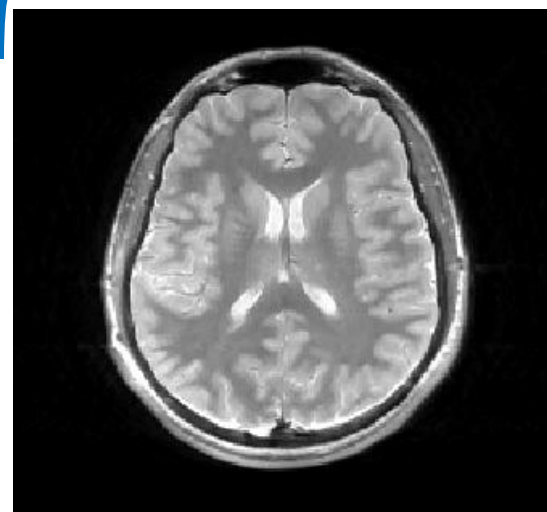
Lustig *et al.*   9.4 %

M-FOCUSS   5.1 %

Joint Bayes   3.6 %

k-space, 40 % of Nyquist rate

Different undersampling patterns

Our Bayesian CS

Error: 3.6 % RMSE
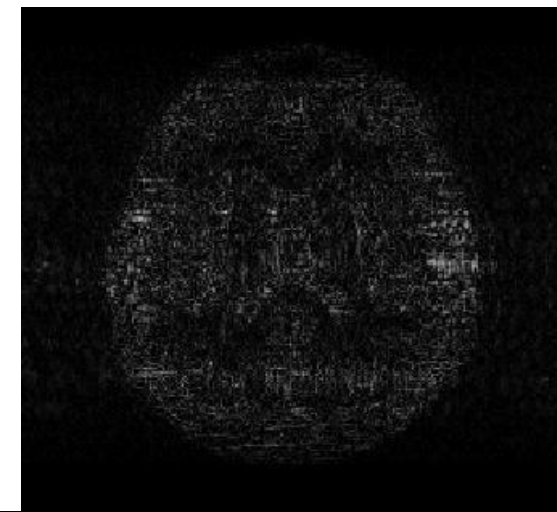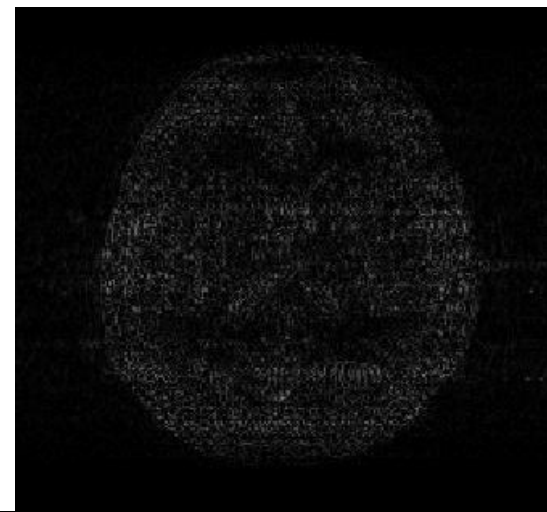
Extensions and Limitations

❖ We assumed the multi-contrast images to be real-valued. Extension to complex-valued images is possible by using a mirror-symmetric sampling pattern and separating real and imaginary parts of the images.

Extensions and Limitations

❖ We assumed the multi-contrast images to be real-valued. Extension to complex-valued images is possible by using a mirror-symmetric sampling pattern and separating real and imaginary parts of the images.

❖ Whereas the other two methods take under an hour, the Bayesian method takes about **20 hours** with this initial implementation.

❖ Current work is on increasing the reconstruction speed using

   ▪ Graphics Processing Cards (GPUs) on the hardware front, and

   ▪ Employing variational Bayesian analysis on the algorithm front

Other applications of joint reconstruction
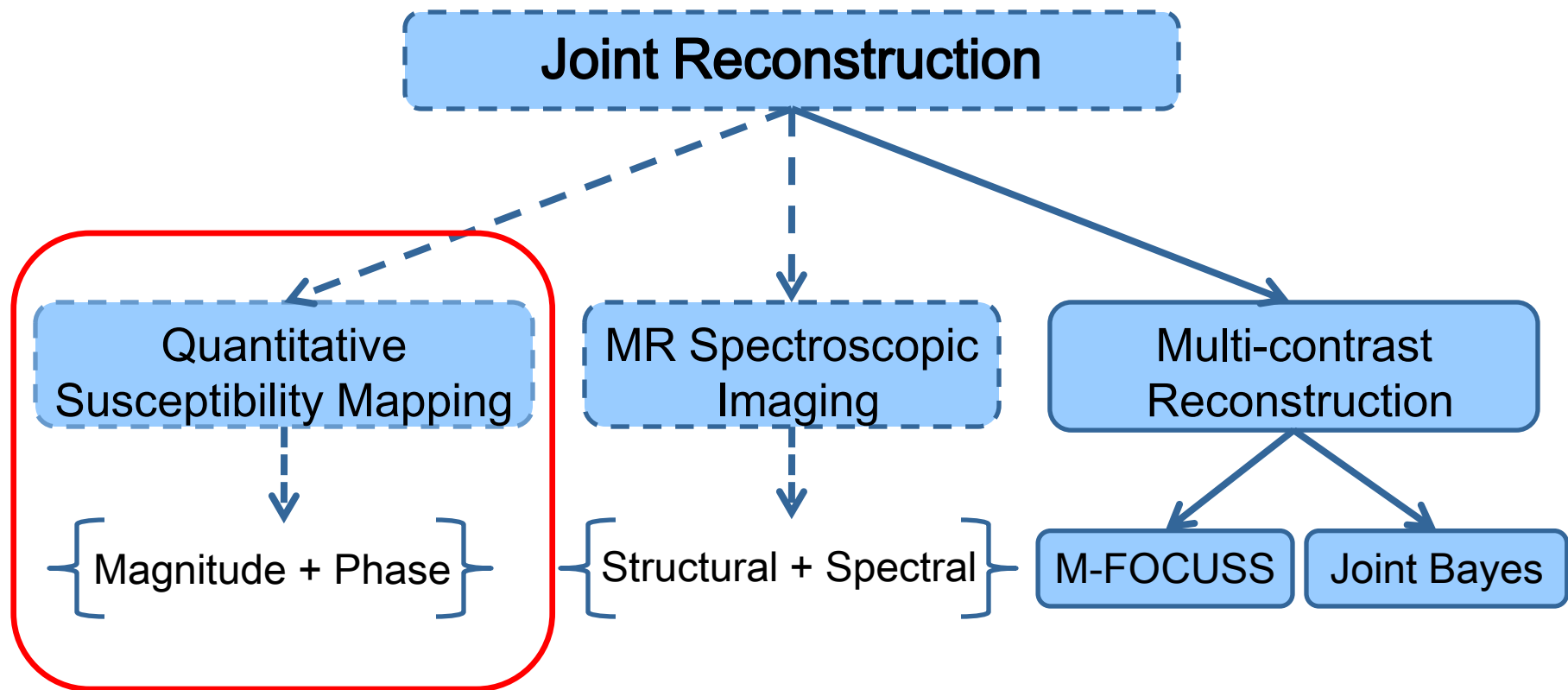
Conclusion

❖ We presented two joint reconstruction algorithms, M-FOCUSS and joint Bayesian CS, that significantly improved reconstruction quality of multi-contrast images from undersampled data.

❖ While joint Bayesian method reduced reconstruction errors by up to 4 times relative to a popular CS algorithm[1], current implementation suffers from long reconstruction times.

❖ M-FOCUSS is a notable candidate that trades off reconstruction quality and processing speed.

[1] Lustig *et al*. MRM 2007
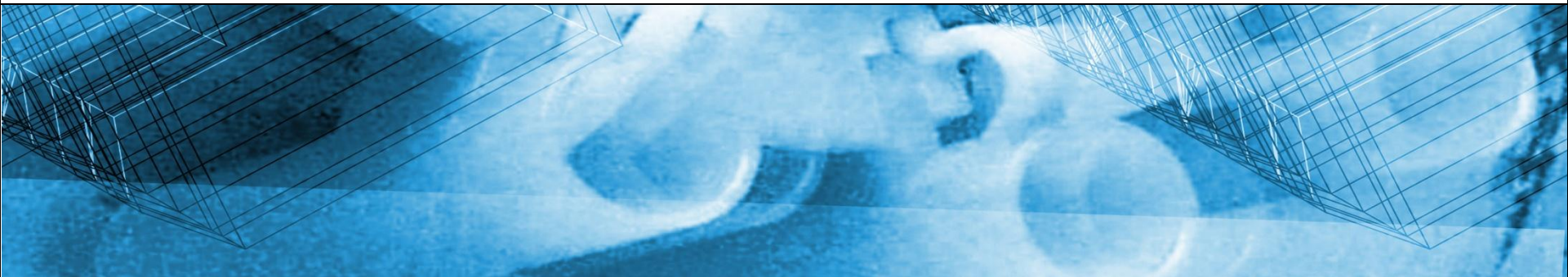
Quantitative Susceptibility Mapping with Magnitude Prior

Berkin Bilgic[1], Audrey P. Fan[1], Elfar Adalsteinsson[1,2]

[1]EECS, MIT, Cambridge, MA, United States
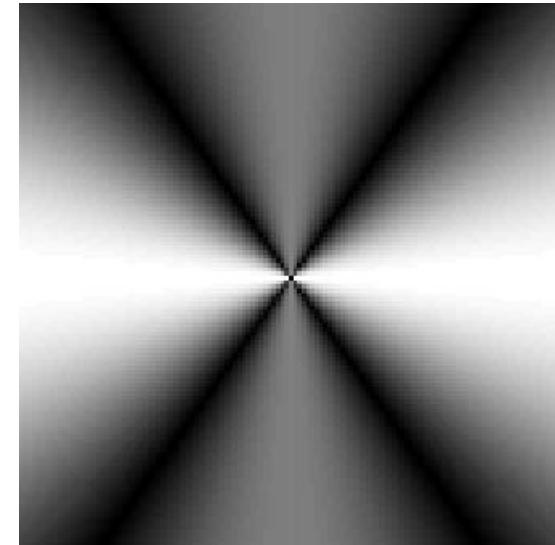[2]Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, United States

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,

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

$\mathbf{F}$: Discrete Fourier Transform matrix

$\mathbf{D}$: susceptibility kernel in k-space

$\delta = \dfrac{\varphi}{\gamma \cdot \boldsymbol{TE} \cdot \boldsymbol{B}_0}$: 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,

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

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

- The inversion is made difficult by zeros on a conical surface in susceptibility kernel $\mathbf{D}$

$$\mathbf{D} = \left(\frac{1}{3} - \frac{k_z^2}{k^2} \right)$$

$|\mathbf{D}|$

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

$|\mathbf{D}|$
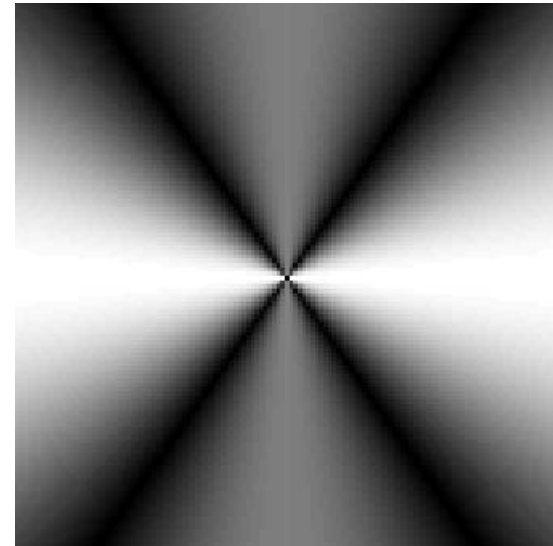
Regularized Inversion for QSM

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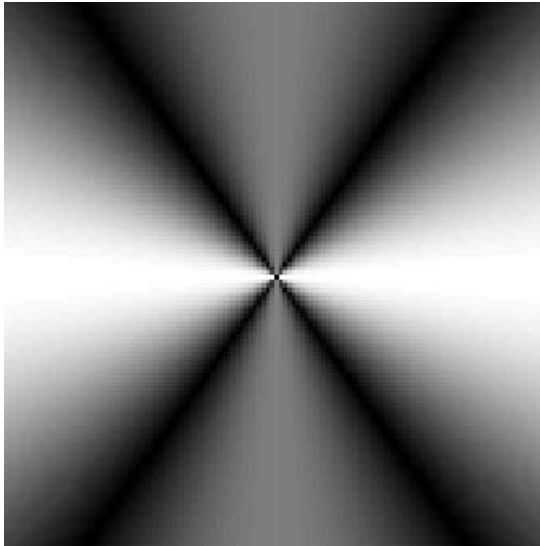
$\log|\mathbf{D}^{-1}|$

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

- 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

$|\mathbf{D}|$

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

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

diverges to ∞

- 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

True χ

Forward transform + noise

$$\mathbf{F}^{-1}\mathbf{D}\mathbf{F}$$

Noisy field map δ

sagittal view

sagittal view

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

- Spatial details that have frequency components at the magic angle lose conspicuity in the field map δ

Regularized Inversion for QSM

Forward
transform
+ noise

$$\mathbf{F}^{-1}\mathbf{DF}$$

Noisy
field map
δ

- Solving for χ by convolving with the inverse of **D** is not possible, as it diverges along the magic angle

- Spatial details that have frequency components at the magic angle lose conspicuity in the field map δ

- We propose to use regularization to facilitate the inversion
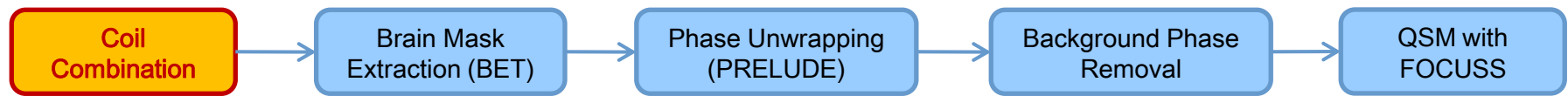
Phase-aware Coil Combination

- 3D GRE acquisition with phased array coils and body coil
- Normalize each channel image with the body coil
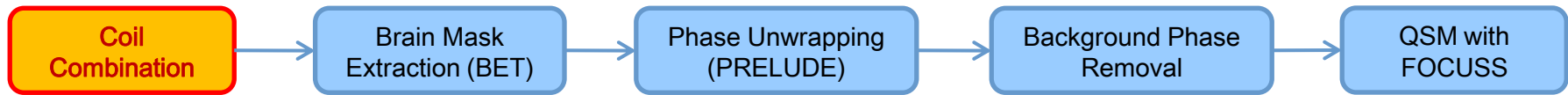
Phase-aware Coil Combination

magnitudes of the coil sensitivities
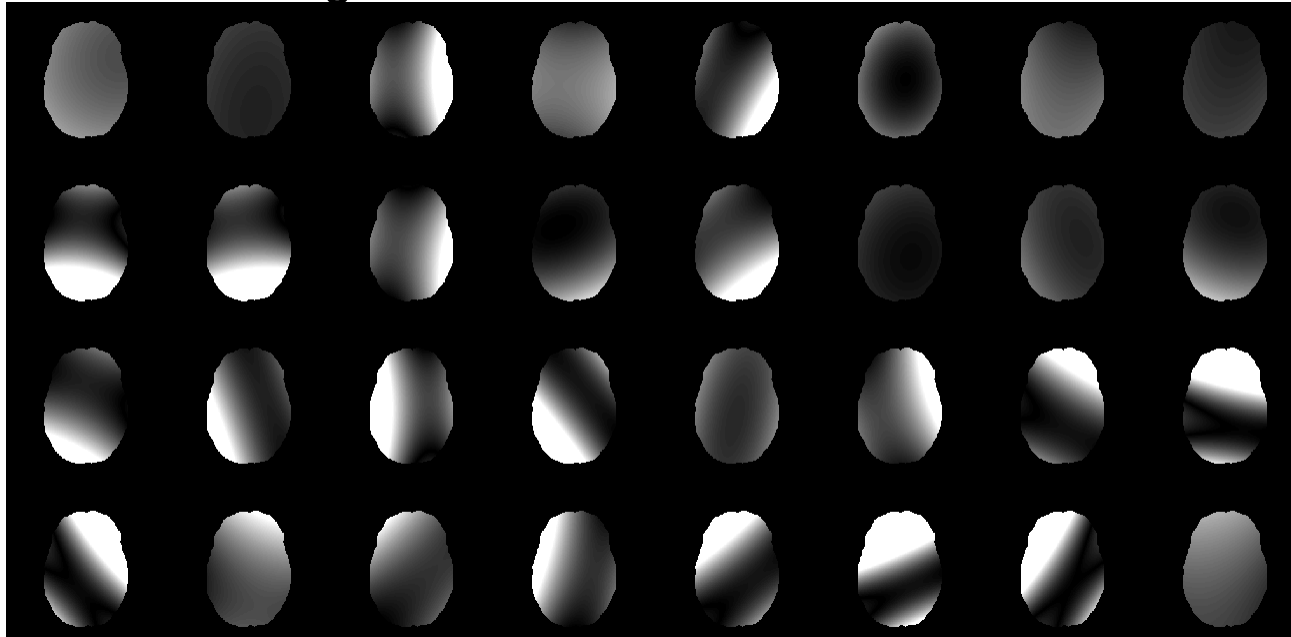
- 3D GRE acquisition with phased array coils and body coil

- Normalize each channel image with the body coil

- Fit 2nd order polynomials to the magnitude of the normalized images → magnitude of the coil sensitivities
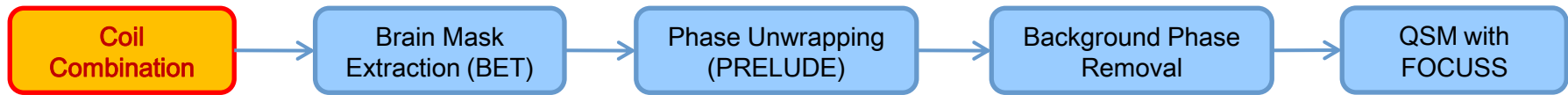
Phase-aware Coil Combination

phase of the coil sensitivities
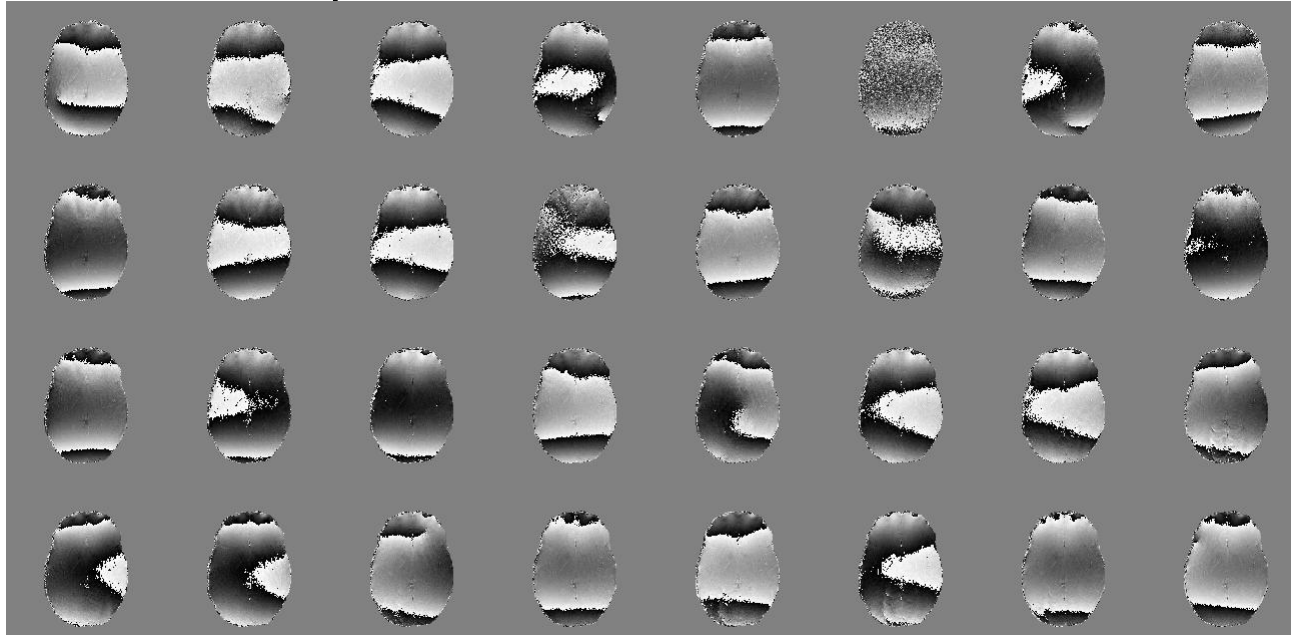
- 3D GRE acquisition with phased array coils and body coil

- Normalize each channel image with the body coil

- Fit 2nd order polynomials to the magnitude of the normalized images → magnitude of the coil sensitivities

- Phase of the normalized images → phase of the coil sensitivities
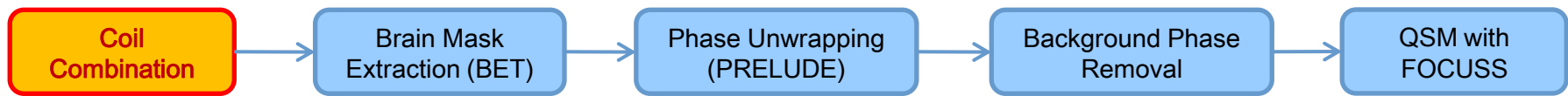
Phase-aware Coil Combination

magnitude of combined image
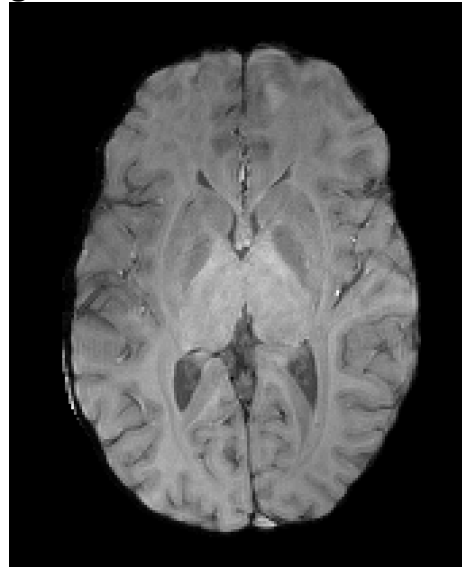
phase of combined image
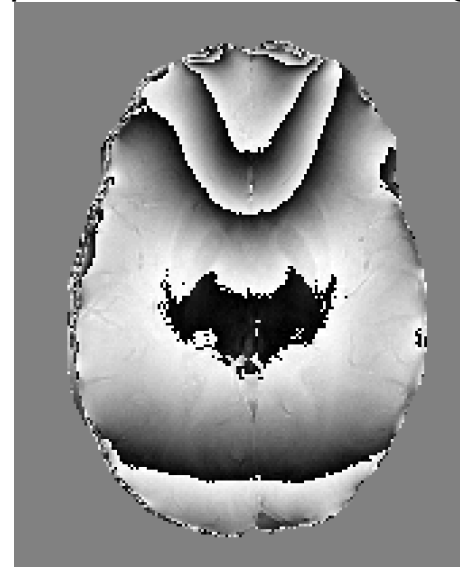
- 3D GRE acquisition with phased array coils and body coil

- Normalize each channel image with the body coil

- Fit 2nd order polynomials to the magnitude of the normalized images → magnitude of the coil sensitivities

- Phase of the normalized images → phase of the coil sensitivities

- Final image is obtained by least-squares coil combination

Brain Mask Extraction & Phase Unwrapping

- Brain mask was generated with the FSL Brain Extraction Tool[1]
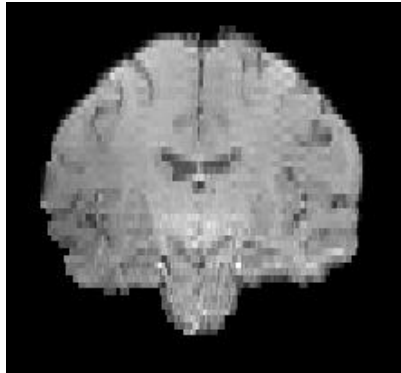
Brain Mask Extraction & Phase Unwrapping

- Brain mask was generated with the FSL Brain Extraction Tool[1]
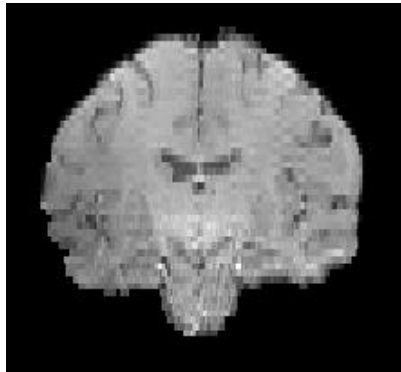
- Phase unwrapping was done with the FSL PRELUDE[2]
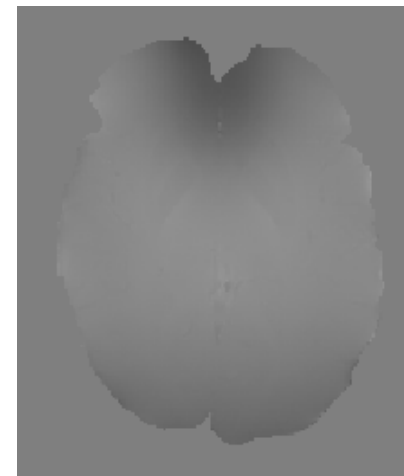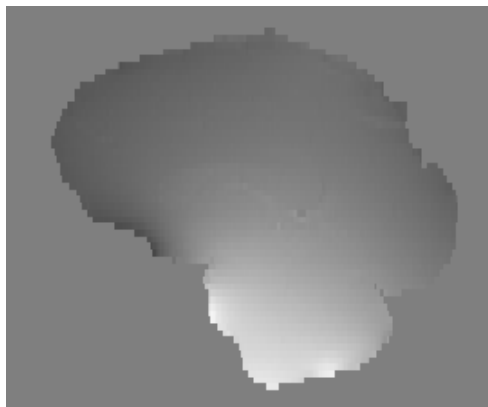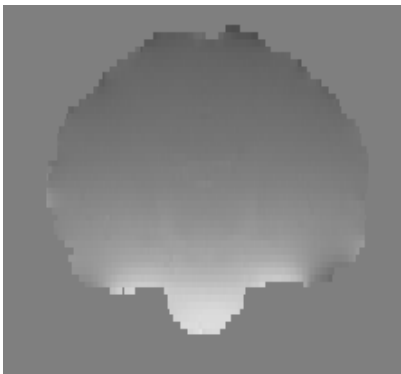
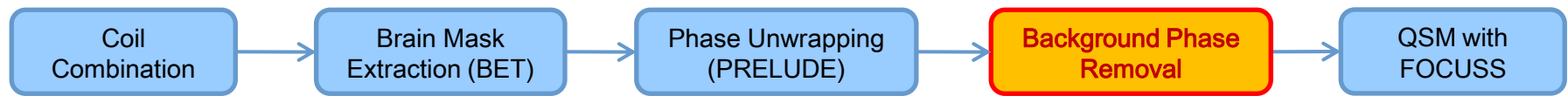[1] Smith SM, Hum. Brain Mapp. 2002   [2] Jenkinson M, MRM 2003
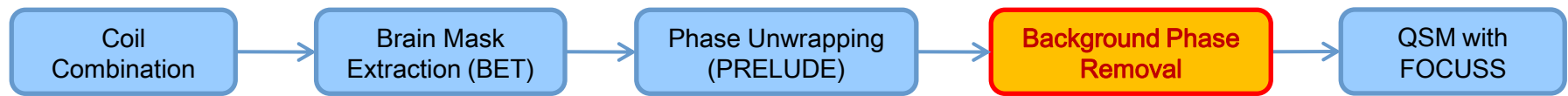
-30 rad   30 rad

Background Phase Removal

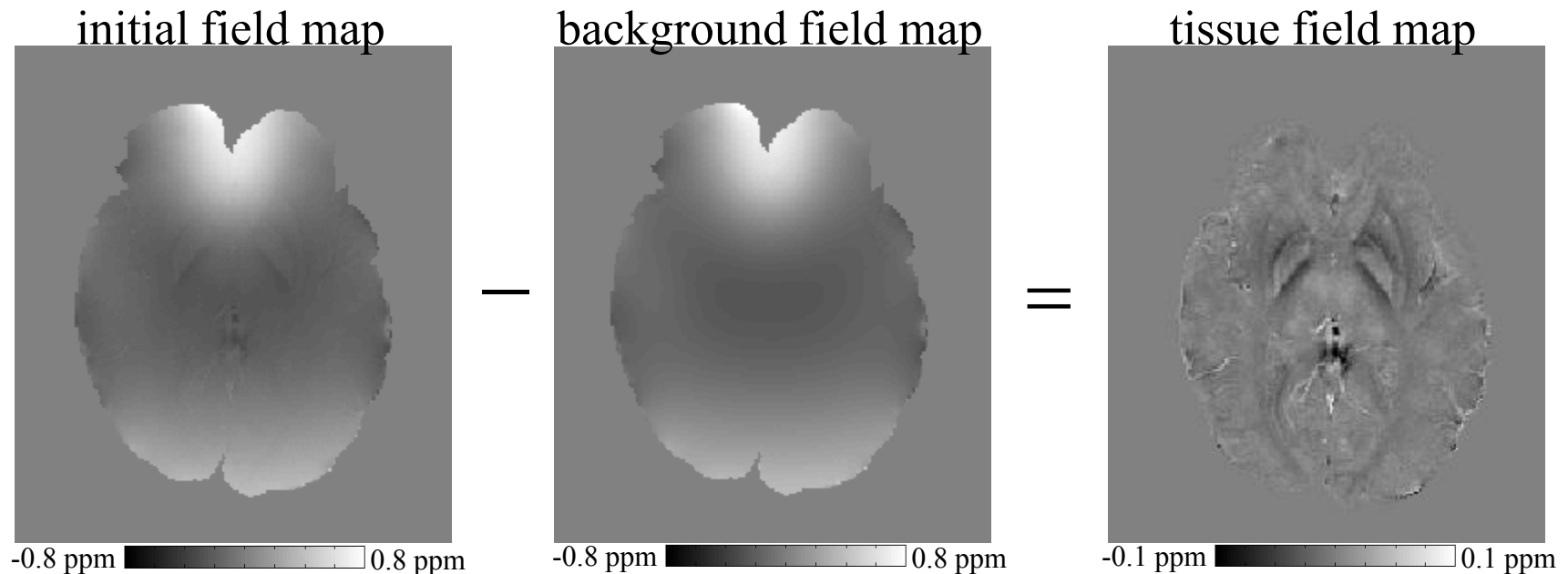- The background phase was estimated with the Effective Dipole Fitting method[1]

initial field map

background field map

-0.8 ppm ▮▮▮▮ 0.8 ppm

-0.8 ppm ▮▮▮▮ 0.8 ppm
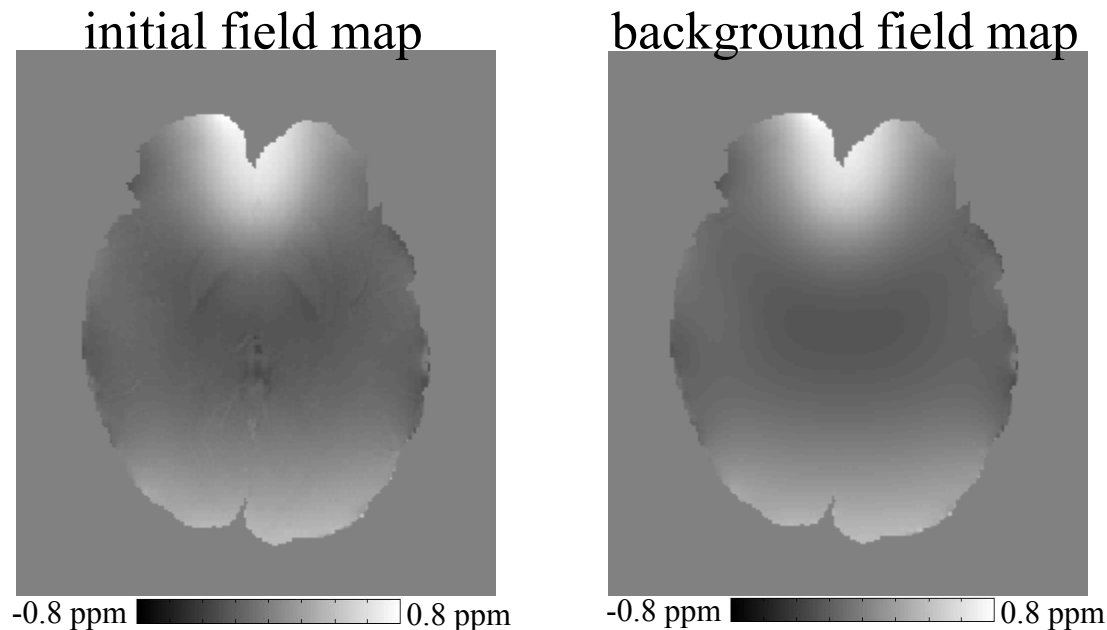
[1] Liu T *et al.*, ISMRM 2010

Background Phase Removal

- The background phase was estimated with the Effective Dipole Fitting method[1]

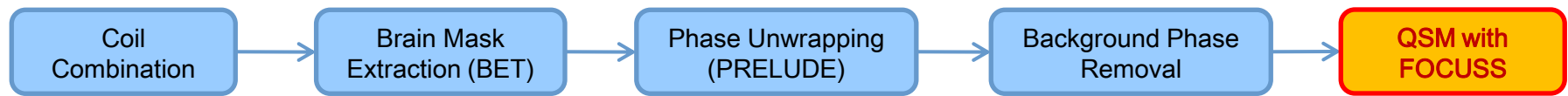- Subtracting the estimated background from the initial field map gives the tissue field map

initial field map — background field map = tissue field map

-0.8 ppm ▮▮ 0.8 ppm   -0.8 ppm ▮▮ 0.8 ppm   -0.1 ppm ▮▮ 0.1 ppm

[1] Liu T *et al.*, ISMRM 2010

66

www.rle.mit.edu

FOCUSS-QSM with magnitude prior

- The tissue field map δ is related to the susceptibility distribution χ via

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

FOCUSS-QSM with magnitude prior

- The tissue field map δ is related to the susceptibility distribution χ via
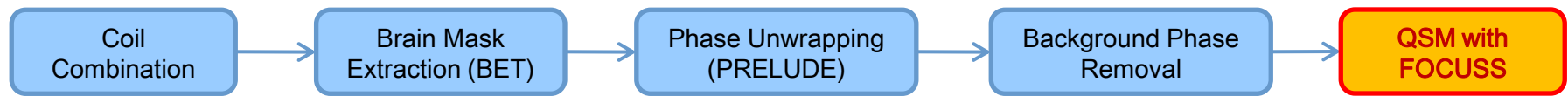
$$\delta = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}\chi$$

- Multiplying both sides with $\mathbf{V}_x\mathbf{F}$

$$\mathbf{V}_x\mathbf{F}\delta = \mathbf{V}_x\mathbf{D}\mathbf{F}\chi$$

where $\mathbf{V}_x$ is a diagonal matrix with $\mathbf{V}_x(\omega, \omega) = \left(1 - e^{-2\pi j\omega/n}\right)$
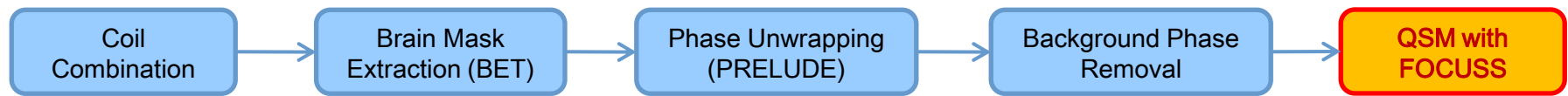
FOCUSS-QSM with magnitude prior

- The tissue field map $\boldsymbol{\delta}$ is related to the susceptibility distribution χ via

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

- Multiplying both sides with $\mathbf{V}_x\mathbf{F}$

$$\mathbf{V}_x\mathbf{F}\delta = \mathbf{V}_x\mathbf{D}\mathbf{F}\chi$$

where $\mathbf{V}_x$ is a diagonal matrix with $\mathbf{V}_x(\omega,\omega) = \left(1 - e^{-2\pi j\omega/n}\right)$

- This corresponds to taking the spatial gradient along the x axis

$$\mathbf{F}\left(\partial_x\delta\right) = \mathbf{D}\mathbf{F}\left(\partial_x\chi\right)$$
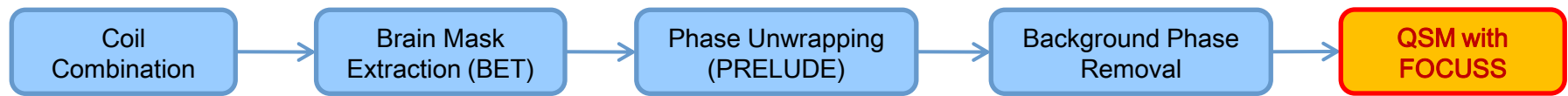
FOCUSS-QSM with magnitude prior

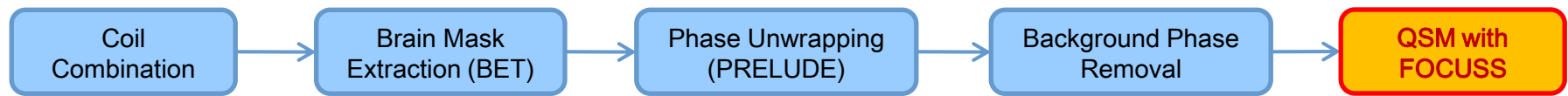- The gradient of the tissue field map δ is related to the gradient of the susceptibility distribution χ via

$$\mathbf{F}\left(\partial_x \delta\right) = \mathbf{DF}\left(\partial_x \chi\right)$$

- We solve for $\partial_x \chi$ with the FOCUSS algorithm[1]

at k^{th} iteration,

$$\mathbf{W}_k = \mathrm{diag}\left(\left|\partial_x \chi_{k-1}\right|^{1/2}\right)$$

[1] Gorodnitsky IF *et al.*, IEEE T. Signal Process. 1997

FOCUSS-QSM with magnitude prior

- The gradient of the tissue field map δ is related to the gradient of the susceptibility distribution χ via

$$\mathbf{F}\left(\partial_x \delta\right) = \mathbf{DF}\left(\partial_x \chi\right)$$

- We solve for $\partial_x \chi$ with the FOCUSS algorithm[1]

at k^{th} iteration,

$$\mathbf{W}_k = \operatorname{diag}\left(\left|\partial_x \chi_{k-1}\right|^{1/2}\right)$$

$$\boldsymbol{q}_k = \operatorname*{argmin}_{\boldsymbol{q}} \left\|\mathbf{F}\left(\partial_x \delta\right) - \mathbf{DFW}_k \boldsymbol{q}\right\|_2^2 + \lambda \left\|\boldsymbol{q}\right\|_2^2$$

[1] Gorodnitsky IF *et al.*, IEEE T. Signal Process. 1997

71

FOCUSS-QSM with magnitude prior

- The gradient of the tissue field map δ is related to the gradient of the susceptibility distribution χ via

$$\mathbf{F}\left(\partial_x \delta\right) = \mathbf{DF}\left(\partial_x \chi\right)$$

- We solve for $\partial_x \chi$ with the FOCUSS algorithm[1]

  at k^{th} iteration,

  $$\mathbf{W}_k = \operatorname{diag}\left(\left|\partial_x \chi_{k-1}\right|^{1/2}\right)$$

  $$\boldsymbol{q}_k = \underset{\boldsymbol{q}}{\operatorname{argmin}} \left\|\mathbf{F}\left(\partial_x \delta\right) - \mathbf{DFW}_k \boldsymbol{q}\right\|_2^2 + \lambda \left\|\boldsymbol{q}\right\|_2^2$$

  $$\partial_x \chi_k = \mathbf{W}_k \boldsymbol{q}_k$$

[1] Gorodnitsky IF *et al.*, IEEE T. Signal Process. 1997

FOCUSS-QSM with magnitude prior

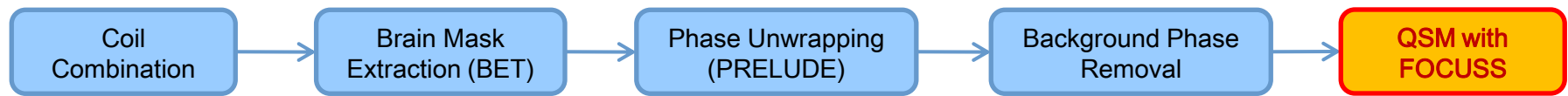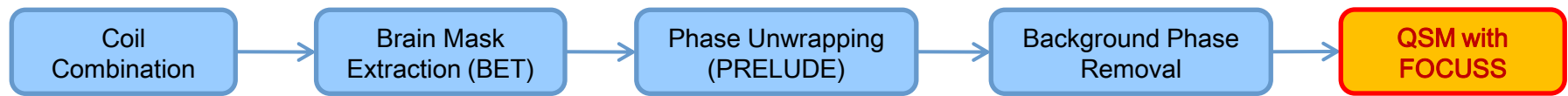- We expect the susceptibility distribution to share similar spatial gradients as the magnitude image.
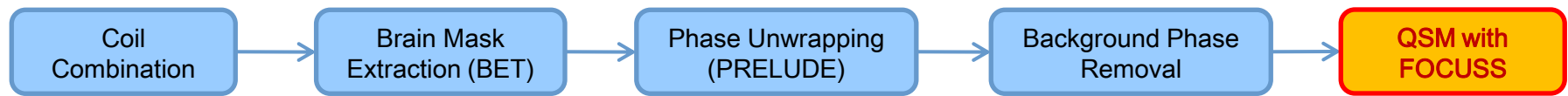
FOCUSS-QSM with magnitude prior

- We expect the susceptibility distribution to share similar spatial gradients as the magnitude image.

- To impose this prior, we modify the update equations as,

$$\mathbf{W}_{prior} = \text{diag}\left(\left|\partial_x \boldsymbol{m}\right|^{1/2}\right), \quad \boldsymbol{m}: \text{magnitude image}$$

at k^{th} iteration,

$$\mathbf{W}_k = \text{diag}\left(\left|\partial_x \chi_{k-1}\right|^{1/2}\right)$$

$$\boldsymbol{q}_k = \underset{\boldsymbol{q}}{\text{argmin}} \ \left\|\mathbf{F}\left(\partial_x \delta\right) - \mathbf{DF}\mathbf{W}_{prior}\mathbf{W}_k\boldsymbol{q}\right\|_2^2 + \lambda\left\|\boldsymbol{q}\right\|_2^2$$

$$\partial_x \chi_k = \mathbf{W}_{prior}\mathbf{W}_k\boldsymbol{q}_k$$
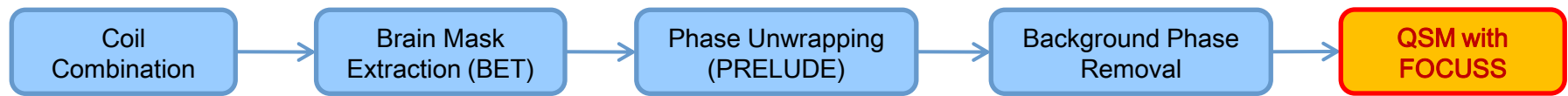
FOCUSS-QSM with magnitude prior

- We expect the susceptibility distribution to share similar spatial gradients as the magnitude image.

- Expressed in terms of $\partial_x \chi$,

$$\mathbf{W}_{prior} = \text{diag}\left(\left|\partial_x \boldsymbol{m}\right|^{1/2}\right), \quad \boldsymbol{m}: \text{magnitude image}$$

$$\partial_x \chi_k = \underset{\partial_x \chi}{\text{argmin}} \left\|\mathbf{F}\left(\partial_x \delta\right) - \mathbf{DF}\left(\partial_x \chi\right)\right\|_2^2 + \lambda \left\|\mathbf{W}_{prior}^{-1} \mathbf{W}_k^{-1}\left(\partial_x \chi\right)\right\|_2^2$$

FOCUSS-QSM with magnitude prior

- We expect the susceptibility distribution to share similar spatial gradients as the magnitude image.

- Expressed in terms of $\partial_x \chi$,

$$\mathbf{W}_{prior} = \mathrm{diag}\left(\left|\partial_x \boldsymbol{m}\right|^{1/2}\right), \quad \boldsymbol{m}: \text{magnitude image}$$

$$\partial_x \chi_k = \underset{\partial_x \chi}{\mathrm{argmin}} \left\|\mathbf{F}\left(\partial_x \delta\right) - \mathbf{DF}\left(\partial_x \chi\right)\right\|_2^2 + \lambda \left\|\mathbf{W}_{prior}^{-1}\mathbf{W}_k^{-1}\left(\partial_x \chi\right)\right\|_2^2$$

if $\partial_x \boldsymbol{m}_i$ is small, $\mathbf{W}_{prior}^{-1}(\boldsymbol{i},\boldsymbol{i})$ will be large and penalize $\partial_x \chi_i$ more

FOCUSS-QSM with magnitude prior

- After estimating the spatial gradients along x, y and z axes, the susceptibility distribution that matches these is found by solving a least squares problem,

$$\chi = \operatorname*{argmin}_{\theta} \sum_{r=x,y,z} \left\| \partial_r \theta - \partial_r \chi \right\|_2^2 + \beta \cdot \left\| \delta - \mathbf{F}^{-1} \mathbf{D} \mathbf{F} \theta \right\|_2^2$$

FOCUSS-QSM with magnitude prior

- After estimating the spatial gradients along x, y and z axes, the susceptibility distribution that matches these is found by solving a least squares problem,

$$\chi = \underset{\theta}{\arg\min} \sum_{r=x,y,z} \left\| \partial_r \theta - \partial_r \chi \right\|_2^2 + \beta \cdot \left\| \delta - \mathbf{F^{-1}} \mathbf{D} \mathbf{F} \theta \right\|_2^2$$

matching gradients · · · · · · · · · data consistency

QSM result: FOCUSS-QSM with magnitude prior

Noisy field map δ

QSM-FOCUSS inversion

χ with FOCUSS +prior

sagittal view

sagittal view
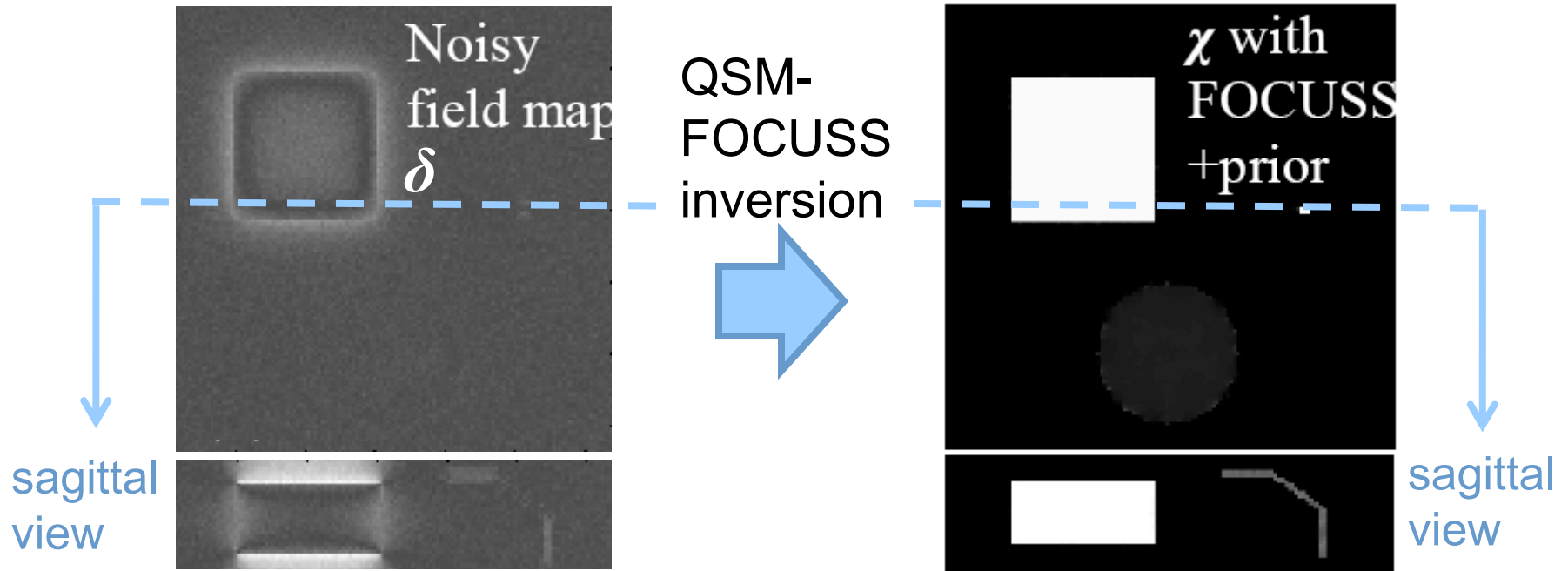
- Starting from the noisy field map δ, FOCUSS-QSM with magnitude prior yielded a susceptibility map with 1.3 % RMSE relative to true χ.

QSM result: FOCUSS-QSM with magnitude prior
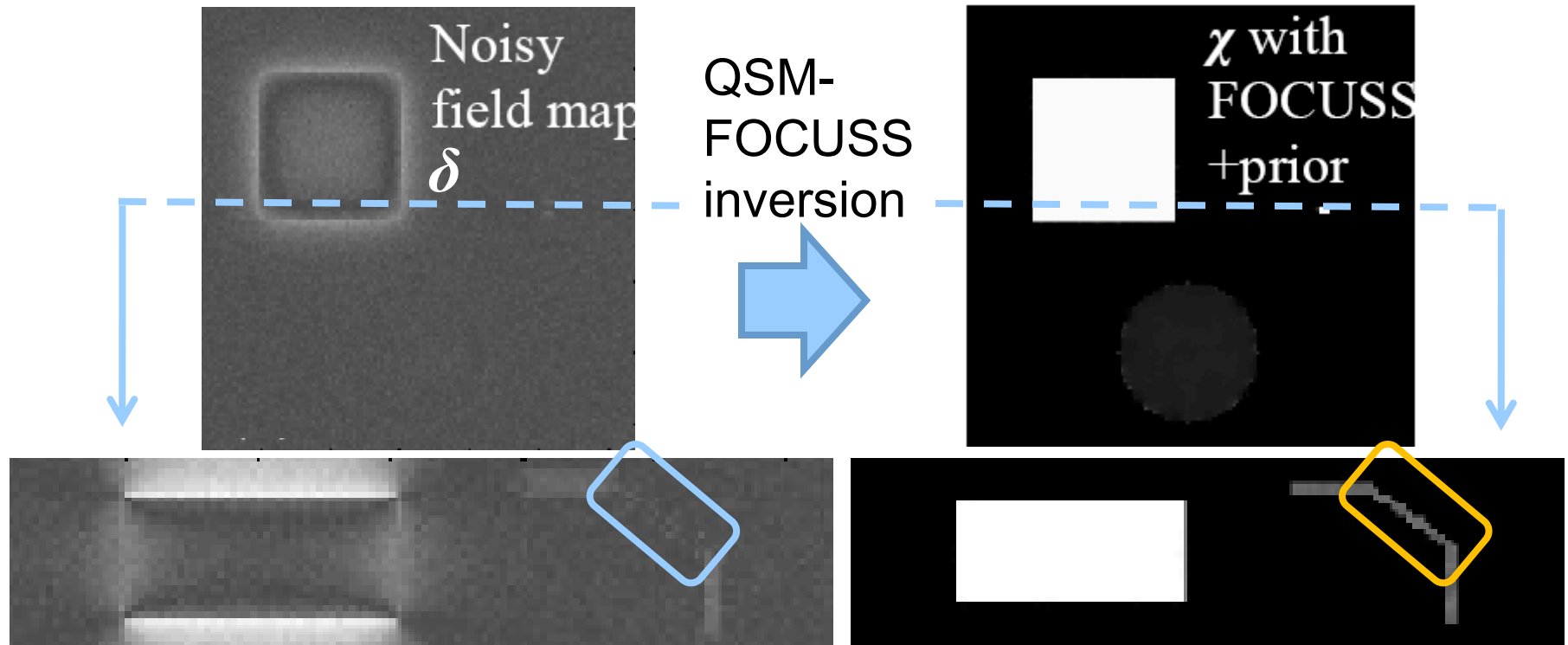
- The reconstructed susceptibility map managed to recover the vessel at the magic angle, which was virtually lost in the field map.

In vivo QSM result: FOCUSS-QSM with magnitude prior
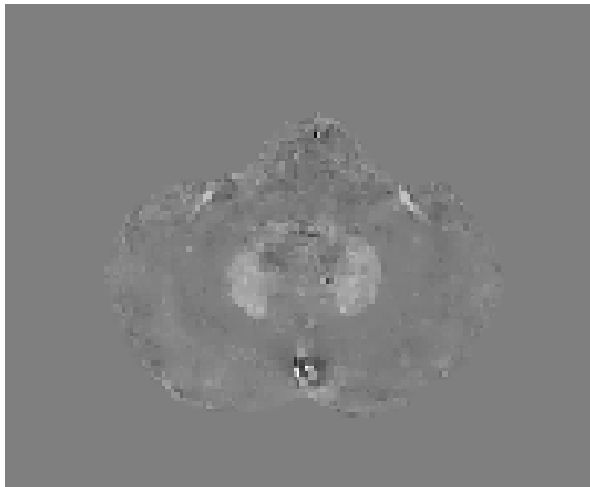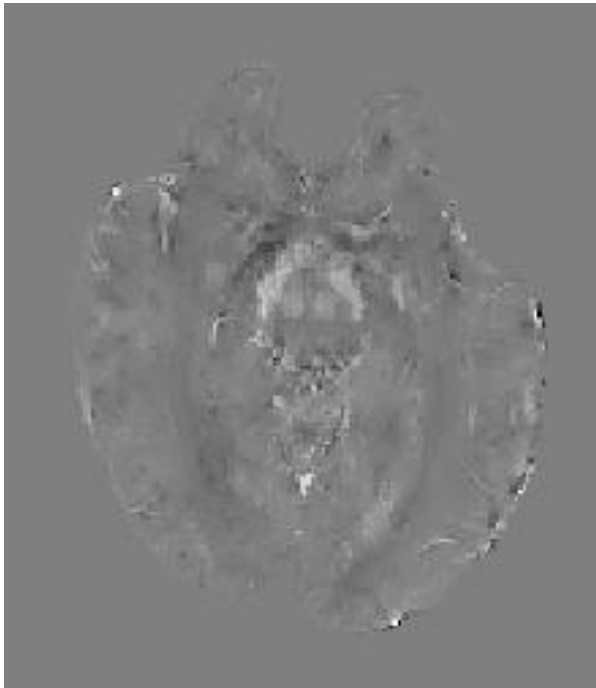
❖ 3D GRE acquisition at 3T

❖ 32 channel receive array

❖ 0.94x0.94x2.5 mm^3 resolution

❖ TE: 20 ms
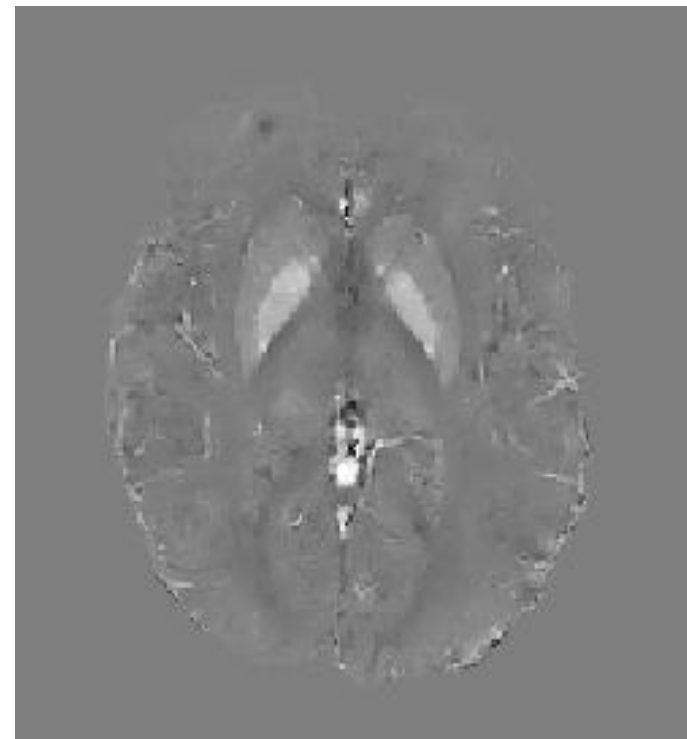
-0.3 ppm 0.3 ppm

In vivo QSM result: FOCUSS-QSM with magnitude prior
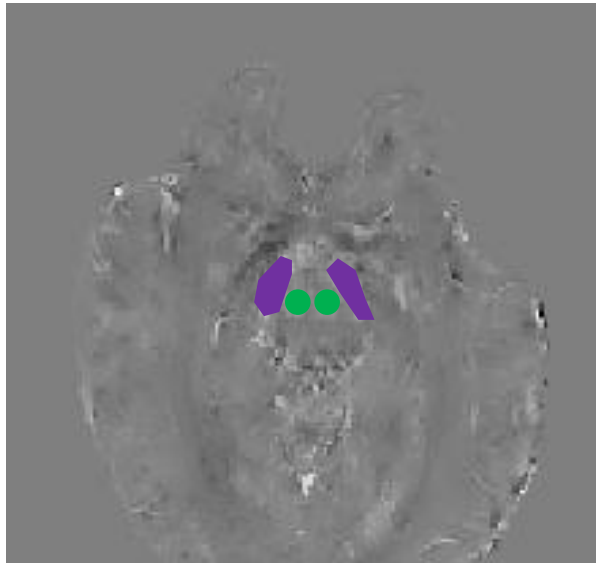
subs. nigra: 0.105 ppm

red nuclei: 0.045 ppm
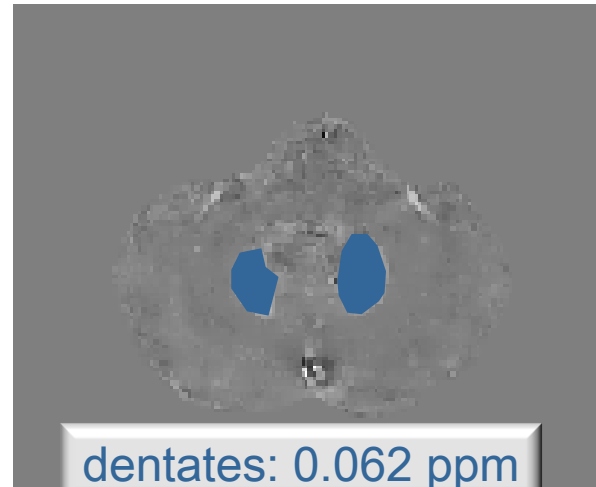
dentates: 0.062 ppm

Structure	$\Delta\chi$ [ppm]
Globus Pallidus	12.3
Substantia Nigra	10.5
Dentate	6.2
Red Nucleus	4.5
Putamen	3.2
Caudate	2.3

x 0.01 ppm, relative to χ_{CSF}
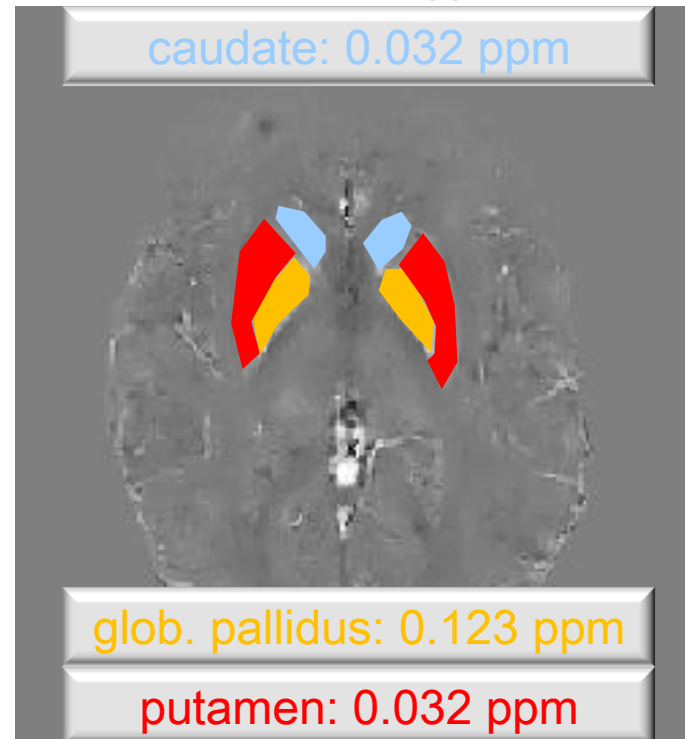
caudate: 0.032 ppm

glob. pallidus: 0.123 ppm

putamen: 0.032 ppm

-0.3 ppm ⬛ 0.3 ppm

In vivo QSM result: FOCUSS-QSM with magnitude prior

-0.3 ppm ▮▮▮▮▮▮ 0.3 ppm

www.rle.mit.edu
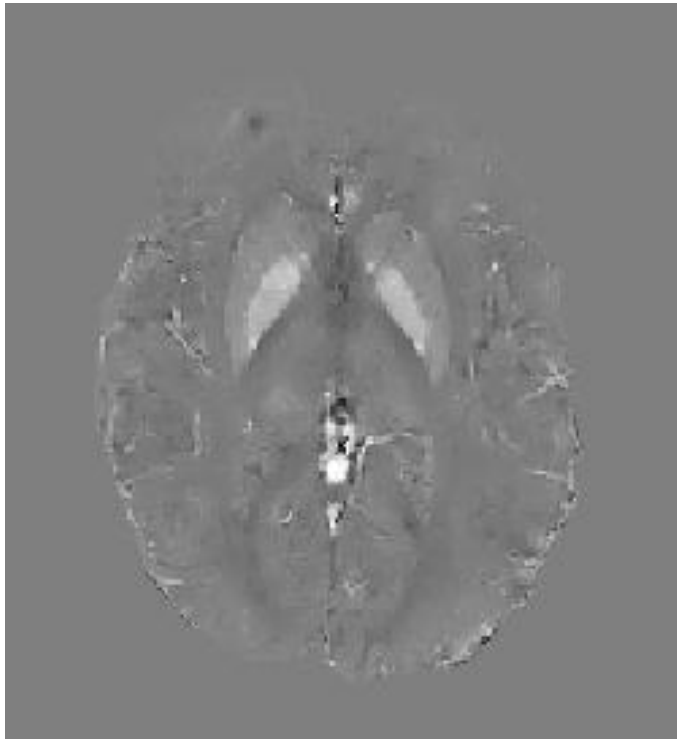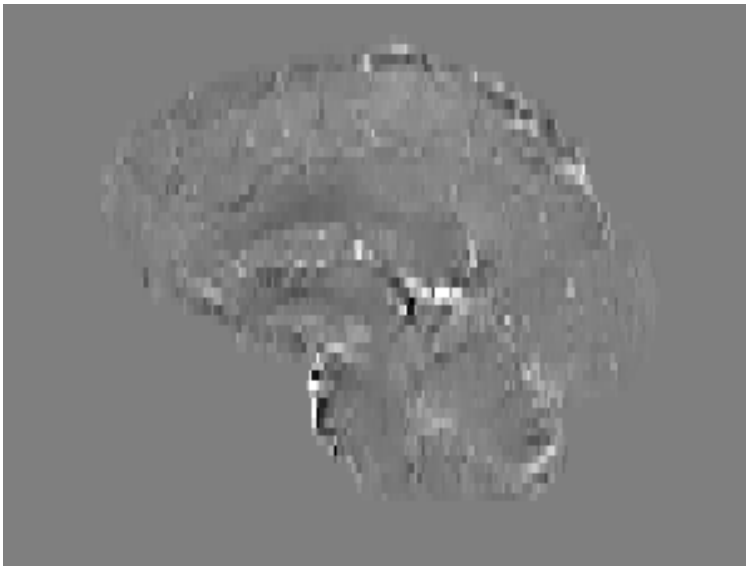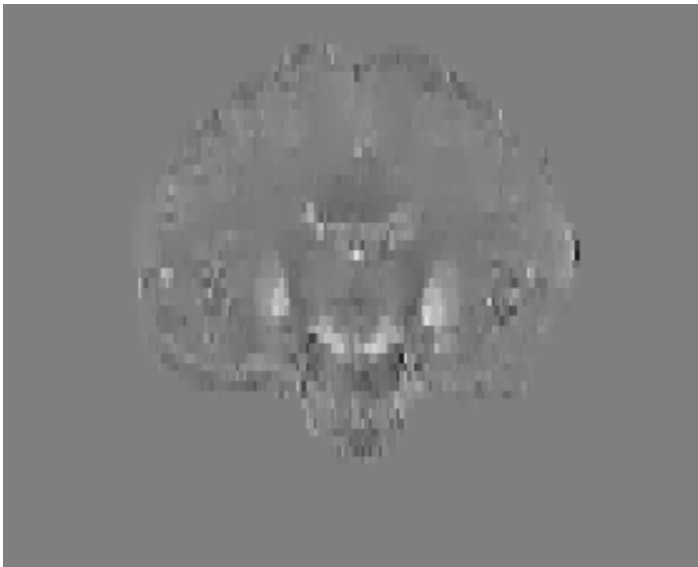
In vivo QSM result: FOCUSS-QSM with a prior
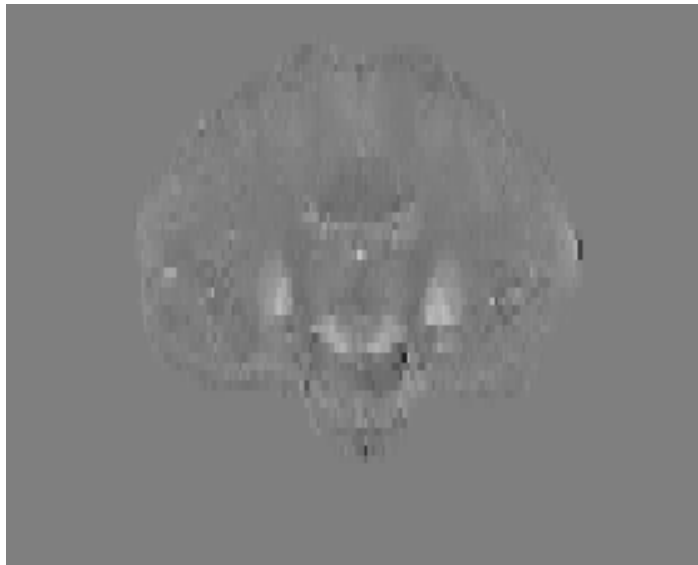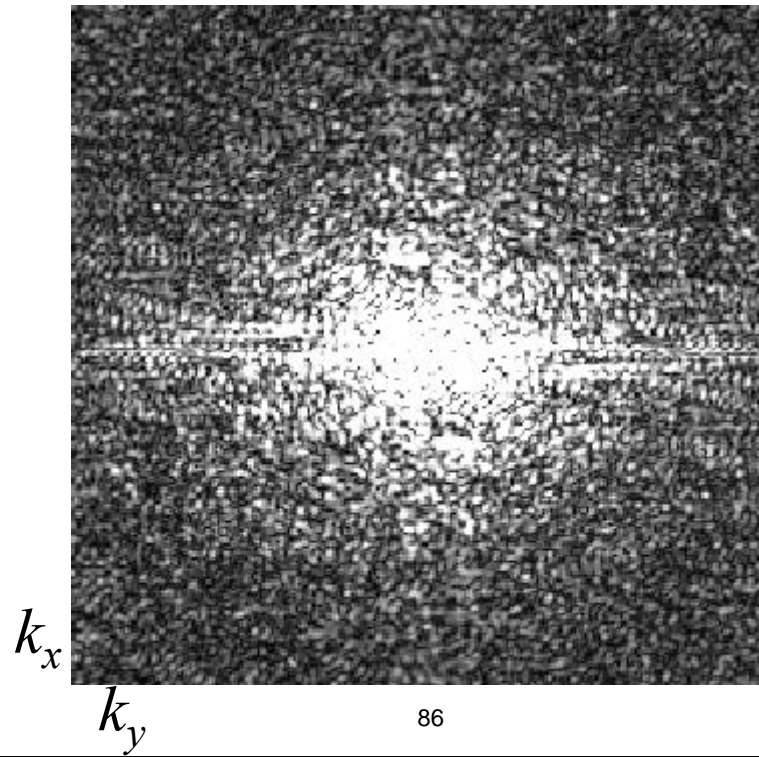

Vessels are less apparent without the magnitude prior

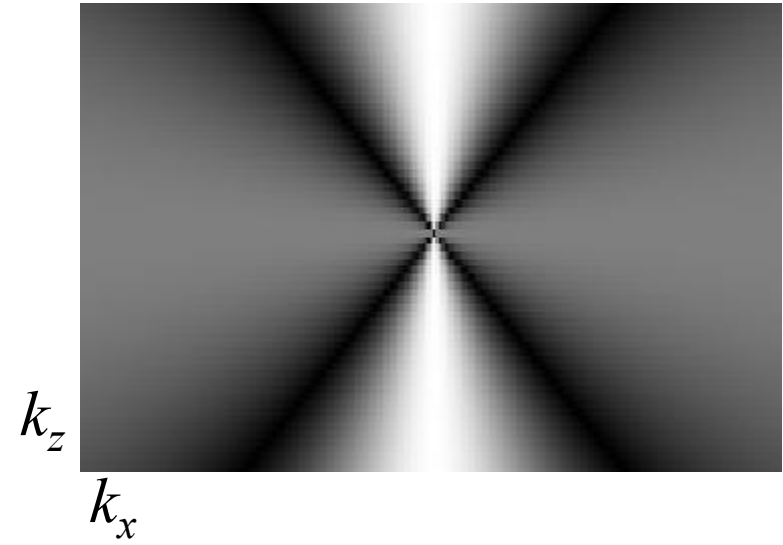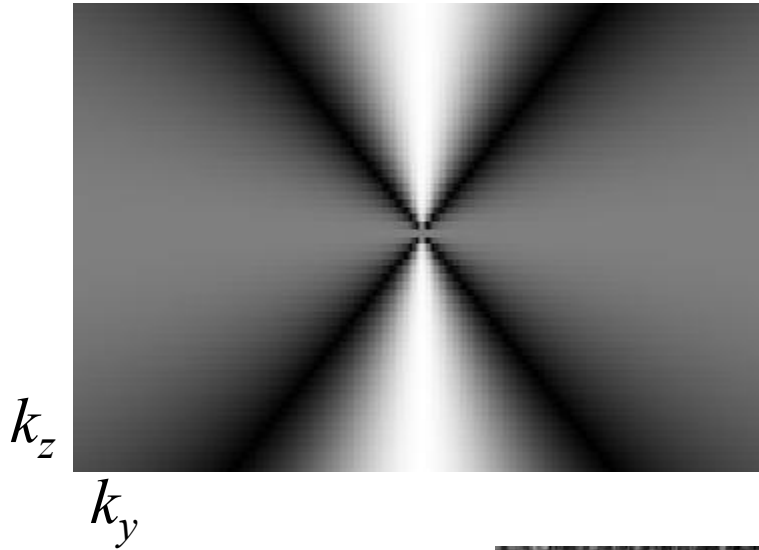-0.3 ppm ▬▬▬ 0.3 ppm

www.rle.mit.edu
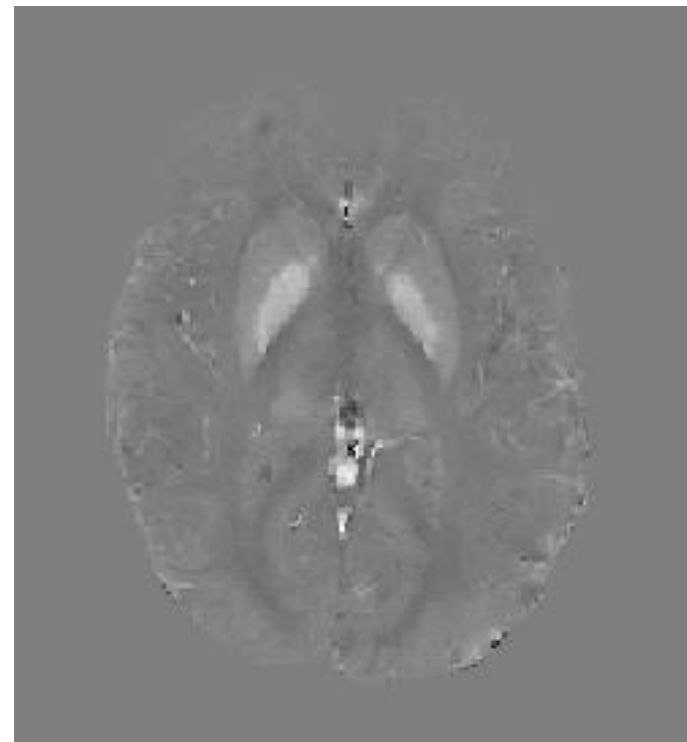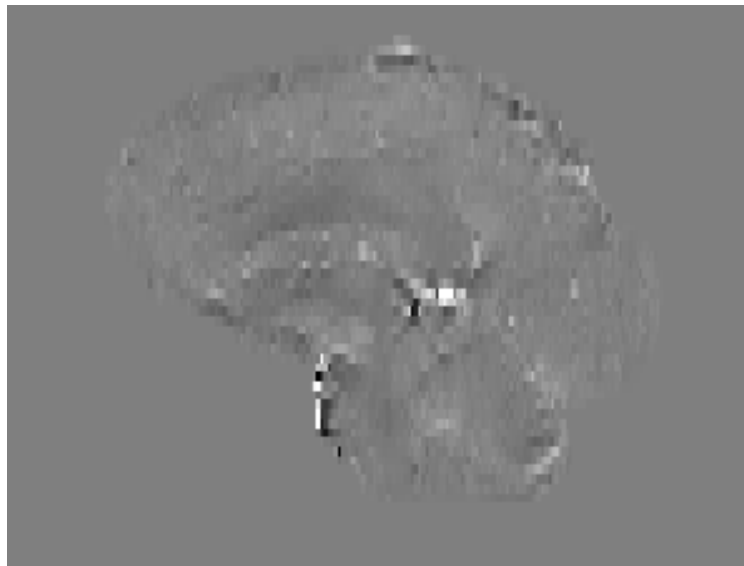
Corresponding Tissue Field Map:
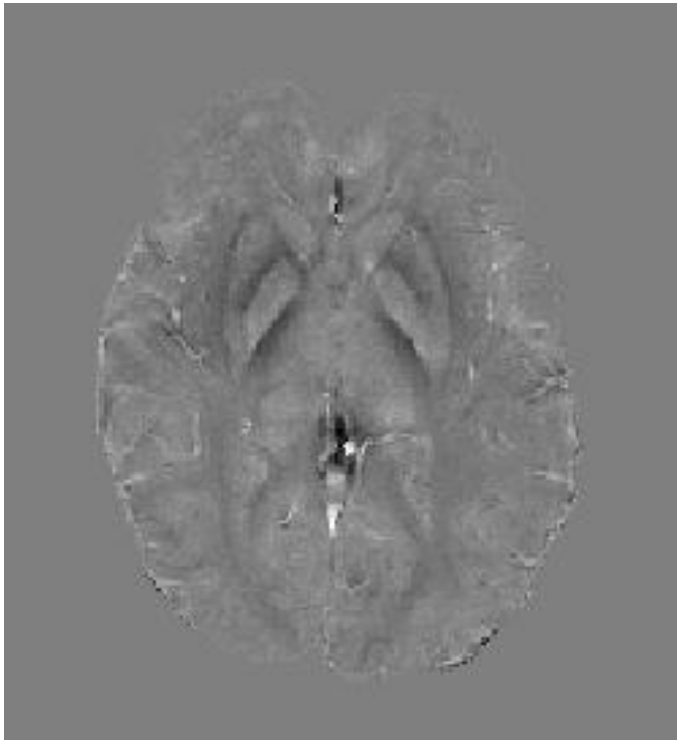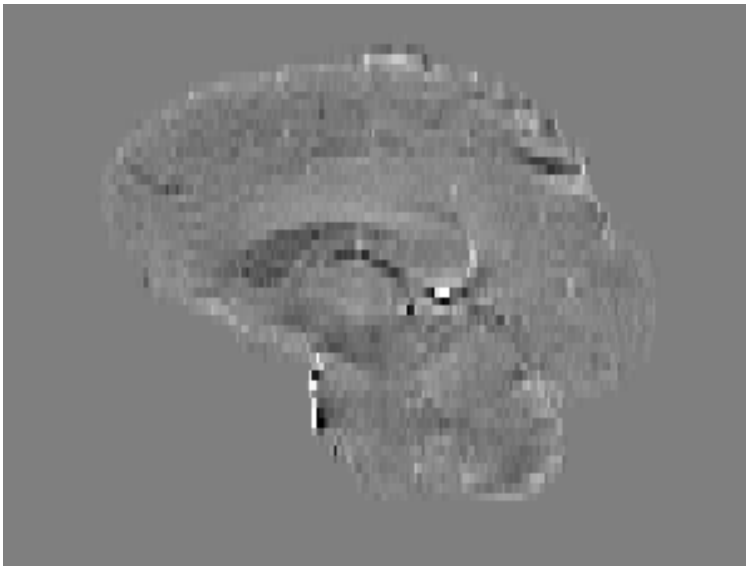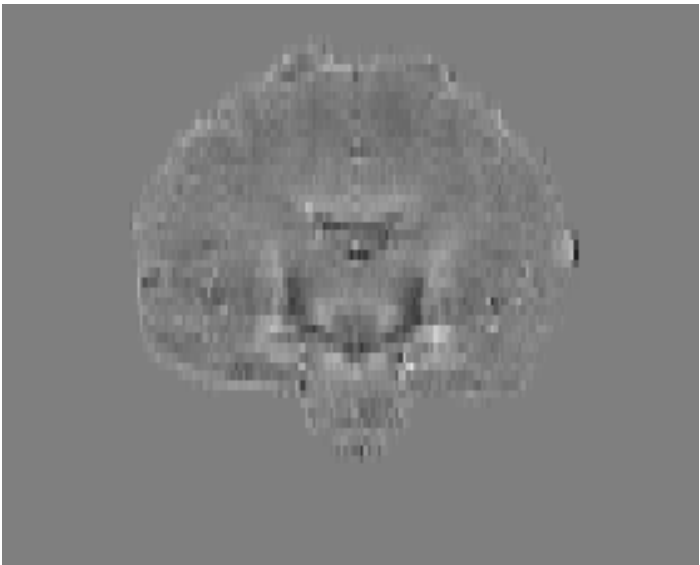
-0.1 ppm ▬▬▬▬▬ 0.1 ppm

www.rle.mit.edu

In vivo QSM result with magnitude prior in k-space:

k_z

k_y

k_z

k_x

k_x

k_y

In vivo QSM result with magnitude prior in k-space:

k_z

k_y

k_z

k_x

k_x

k_y
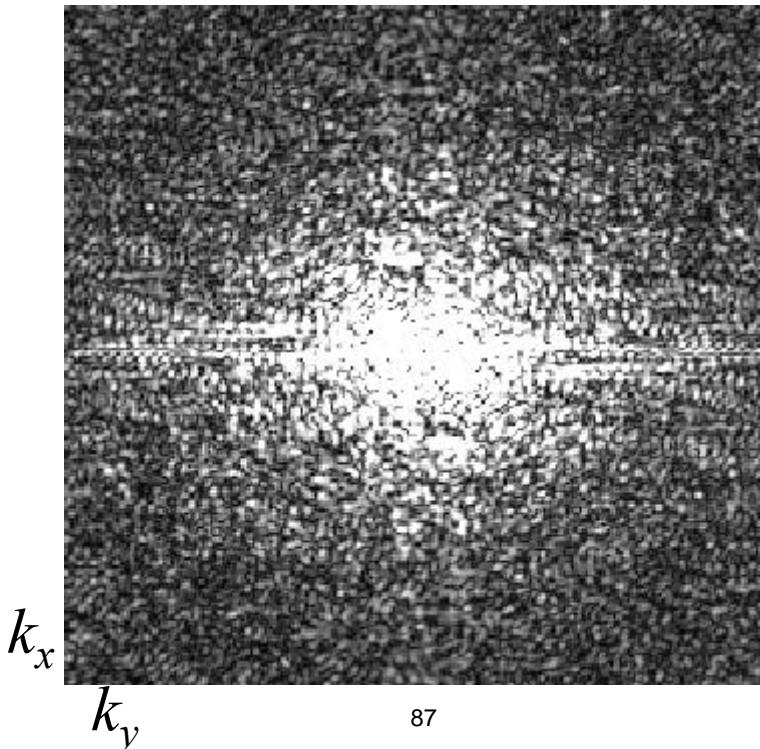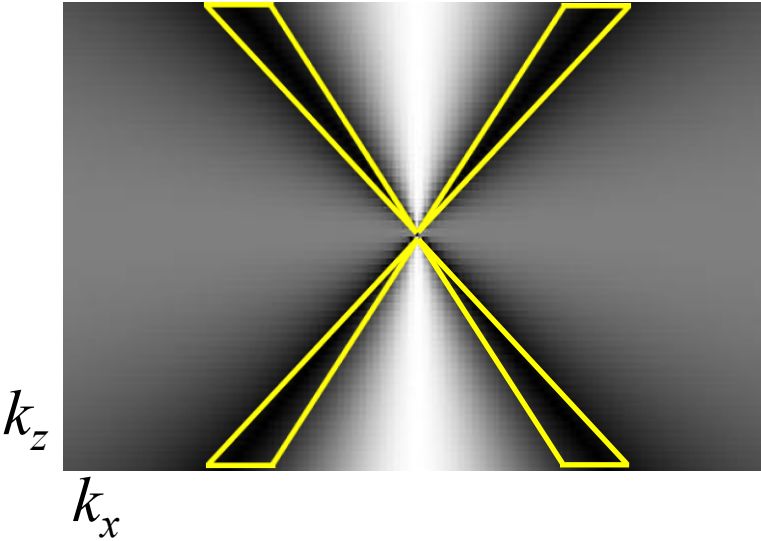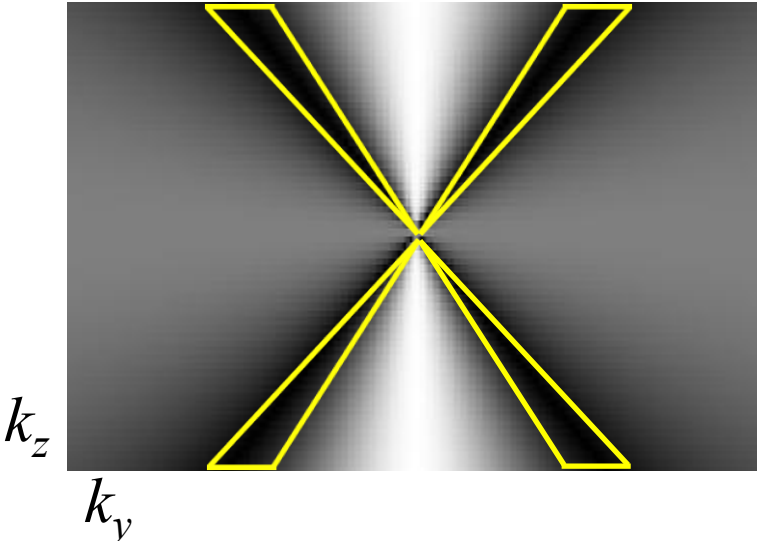
Potential drawbacks of FOCUSS-QSM

- ## Computation time:
    - ❖ Dipole fitting for background removal ≈ 2 hours
    - ❖ FOCUSS-QSM ≈ 1 hours
    - ❖ Total processing time ≈ **3 hours** for data of size [256x256x64]
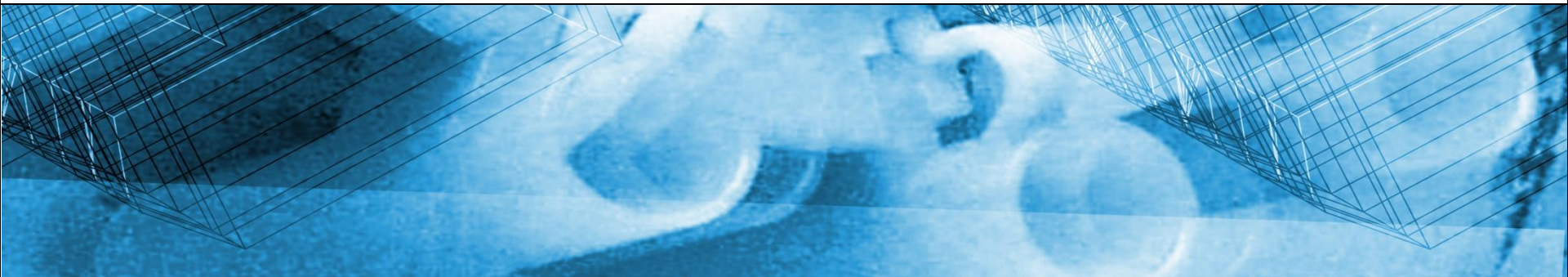
Potential drawbacks of FOCUSS-QSM

- Computation time:
  - ❖ Dipole fitting for background removal ≈ 2 hours
  - ❖ FOCUSS-QSM ≈ 1 hours
  - ❖ Total processing time ≈ **3 hours** for data of size [256x256x64]

- Solution:
  - ❖ Both algorithms solve Least Squares problems, Graphics Processing Card (GPU) implementation will greatly enhance the performance

Conclusion

- Starting with a multi-coil 3D GRE acquisition, we outlined coil combination and background phase elimination methods that yielded the tissue field map.

- We introduced a Quantitative Susceptibility Mapping algorithm that makes use of the magnitude image to facilitate the kernel inversion.

Estimating Brain Iron Concentration in Normal Aging using L1-QSM

Berkin Bilgic[1], Adolf Pfefferbaum[2,3], Torsten Rohlfing[2],
Edith V. Sullivan[3], and Elfar Adalsteinsson[1,4]

[1] EECS, MIT, Cambridge, MA, United States
[2] Neuroscience Program, SRI International, USA
[3] Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA
[4] Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, United States

L1 Regularized Susceptibility Inversion

- Again, we are seeking the susceptibility map that matches the observed tissue phase,

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

- The susceptibility values are tied to the paramagnetic properties of the underlying tissues; hence they vary smoothly across space within anatomical boundaries.

- Based on this, we model the susceptibility map to be approximately piece-wise constant,

- And formulate this belief by invoking sparsity inducing L1 norm on the spatial gradients of χ

L1 Regularized Susceptibility Inversion

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

$$\chi_{\text{tissue}} = \text{argmin}_{\chi} \|\boldsymbol{\delta} - \mathbf{F}^{-1}\mathbf{DF}\,\chi\|_2^2 + \lambda\,(\,\|\partial_x\,\chi\|_1 + \|\partial_y\,\chi\|_1 + \|\partial_z\,\chi\|_1\,)$$

- We call this method L1-QSM, for which λ serves as a regularization parameter that adjusts the smoothness of the solution

- This is essentially the same formulation as FOCUSS-QSM, but is less sophisticated as magnitude information is not used

Tissue iron deposition in young and elderly subjects

- Tissue susceptibility is a sensitive marker of iron concentration, however it is partially influenced by myelin, proteins etc.

- In this study, we used L1-QSM to test the hypothesis that, **iron deposition in striatal and brain stem nuclei would be 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 L1-QSM Result for the Elderly
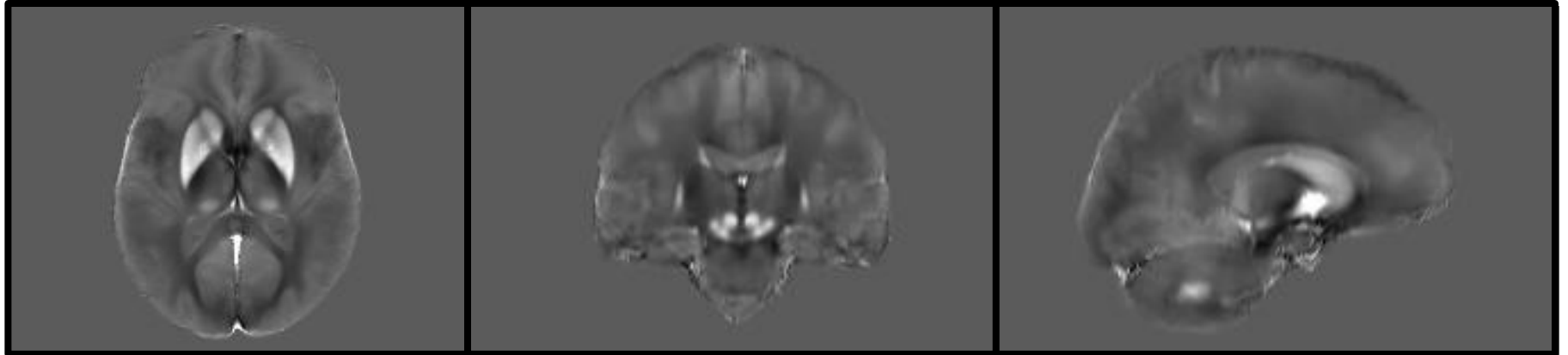
−0.1 ppm   0.16 ppm

Average L1-QSM Result for the Young
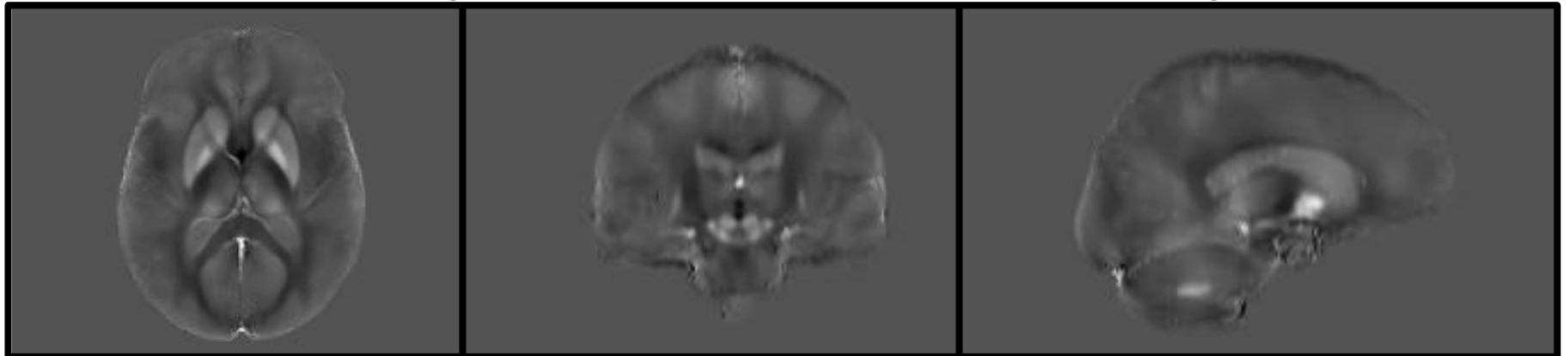
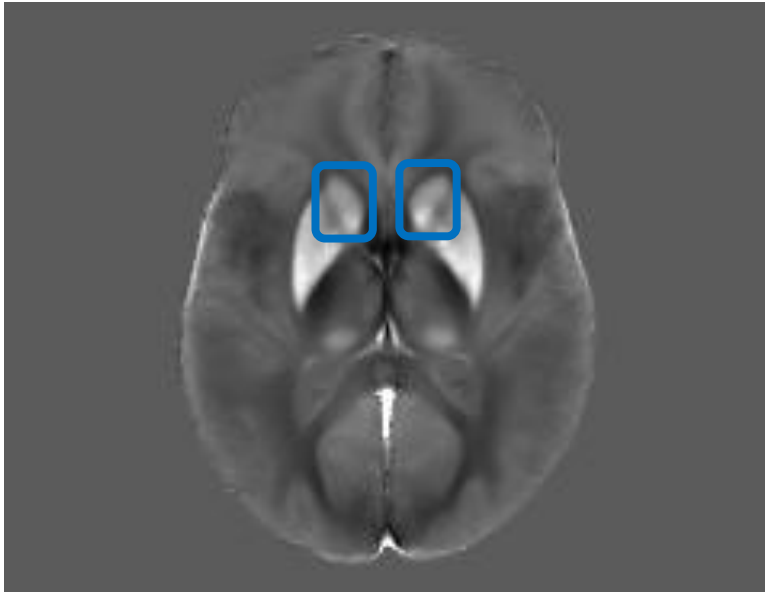# Average L1-QSM Result for the Elderly
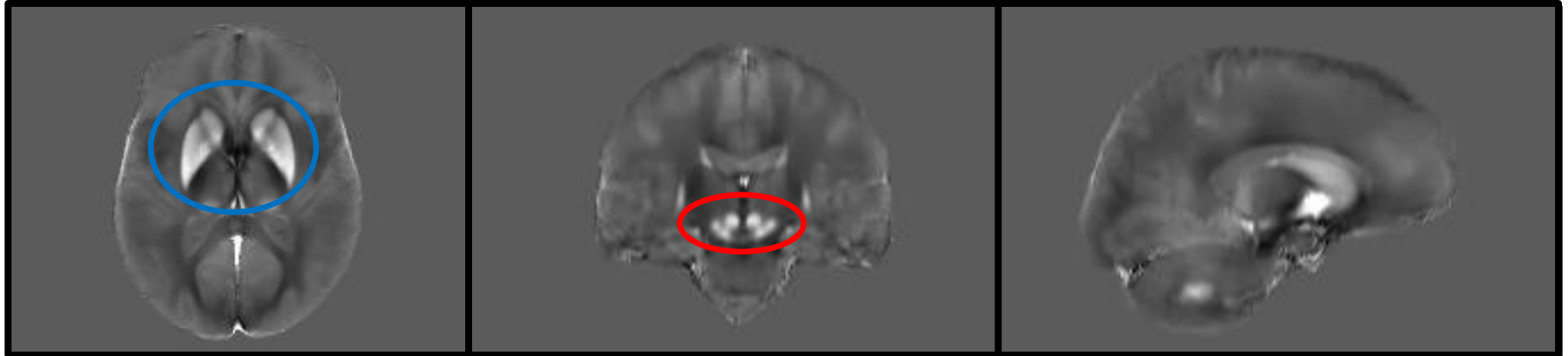
Striatal ROIs   **Brain Stem ROIs**

−0.1 ppm   0.16 ppm

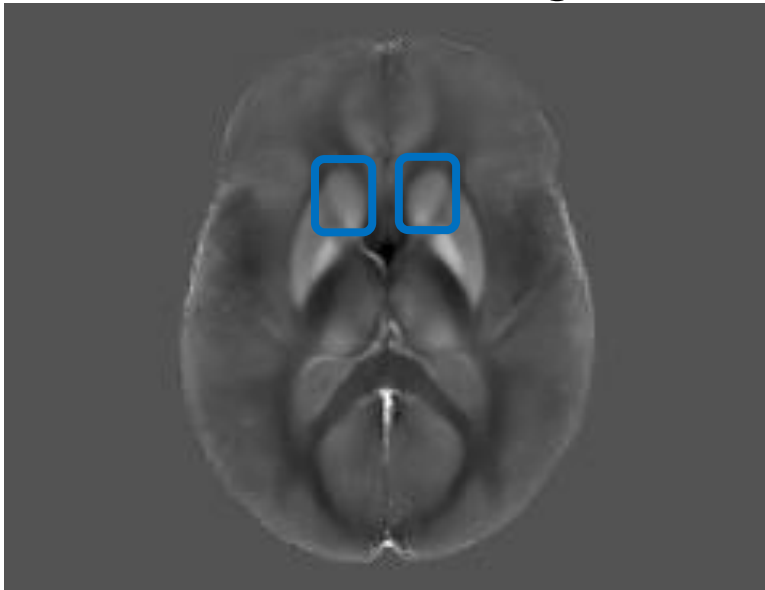# Average L1-QSM Result for the Young
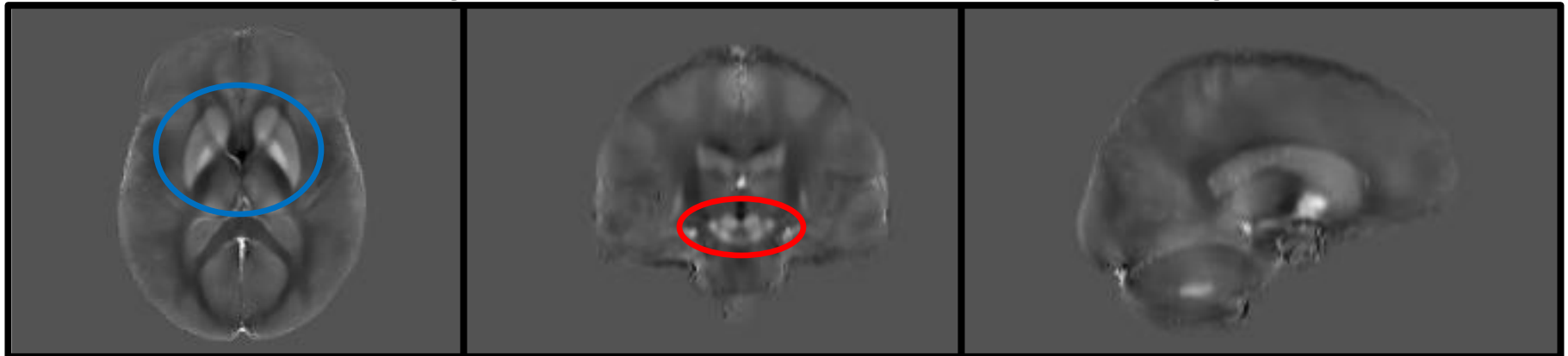
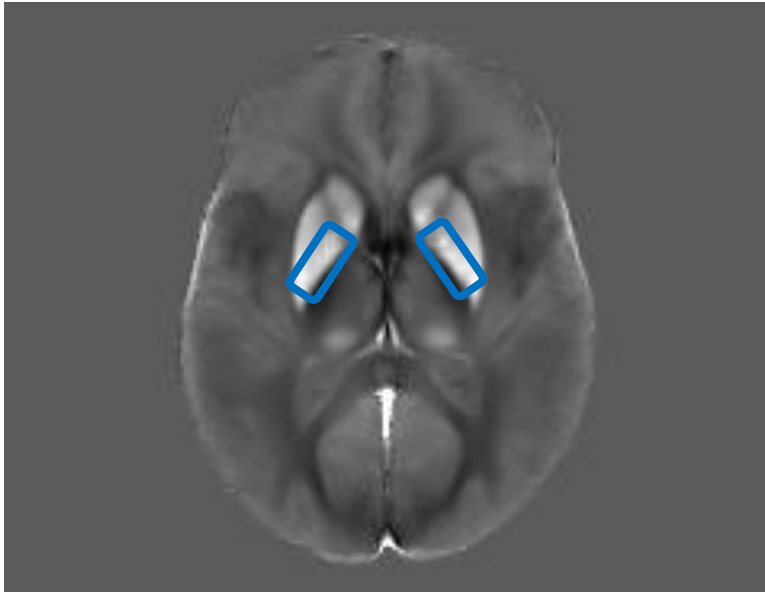# Average L1-QSM Result for the Elderly

Elderly caudate nucleus: 0.059 ppm

Average L1-QSM Result for the Young
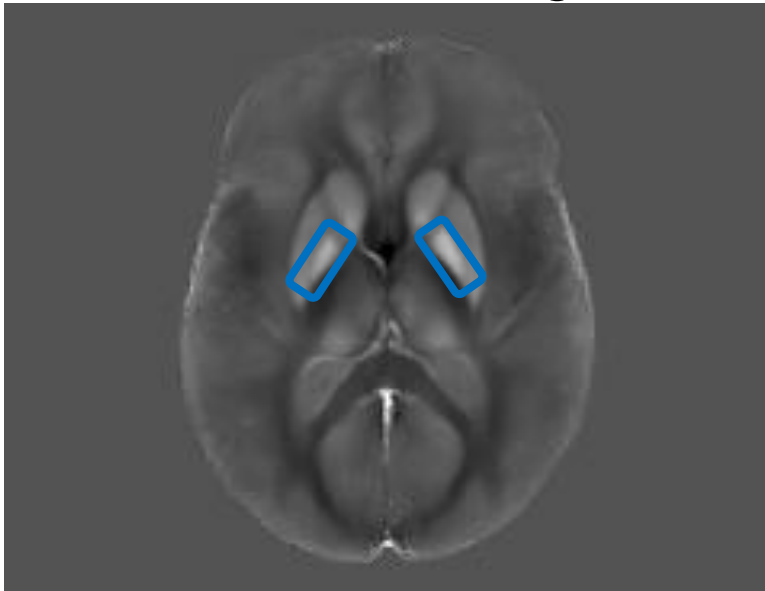

Young caudate nucleus: 0.023 ppm

t-test result: **p < 0.0001**
significant

Average L1-QSM Result for the Elderly
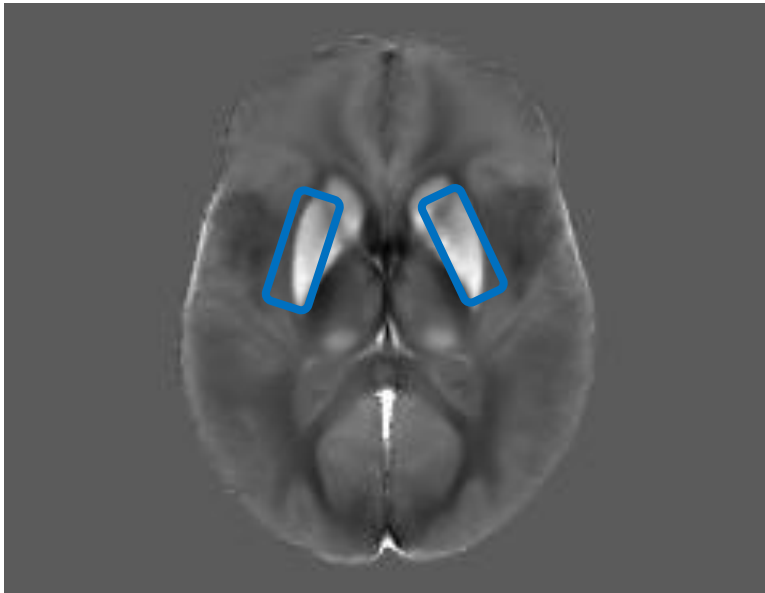
Elderly globus pallidus: 0.120 ppm

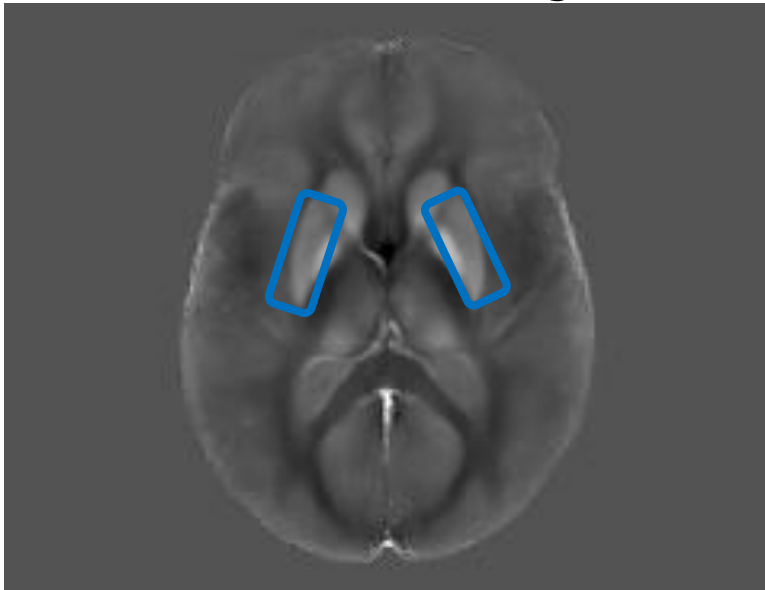Average L1-QSM Result for the Young

Young globus pallidus: 0.069 ppm

t-test result: **p < 0.0001**
significant

Average L1-QSM Result for the Elderly
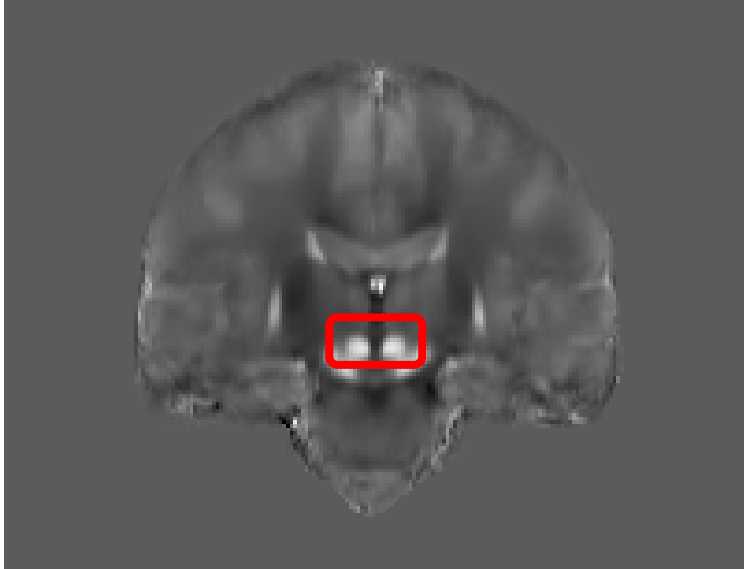

Elderly putamen: 0.095 ppm

Average L1-QSM Result for the Young

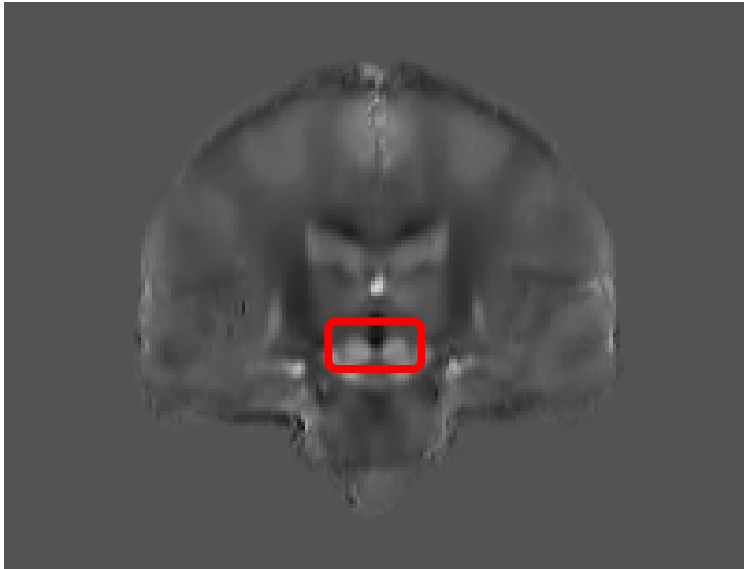
Young putamen: 0.024 ppm

t-test result: **p < 0.0001**
significant

Average L1-QSM Result for the Elderly

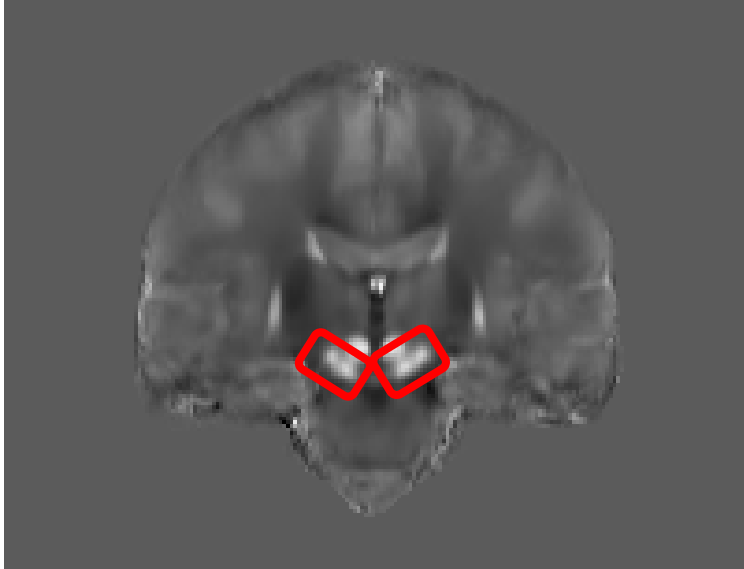Elderly red nucleus: 0.091 ppm

Average L1-QSM Result for the Young

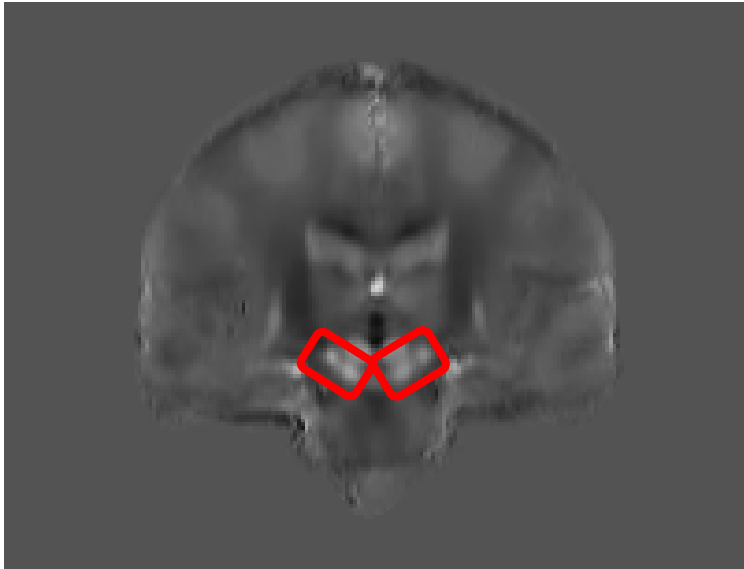Young red nucleus: 0.030 ppm

t-test result: **p = 0.0008
significant**

Average L1-QSM Result for the Elderly

Elderly substantia nigra: 0.055 ppm
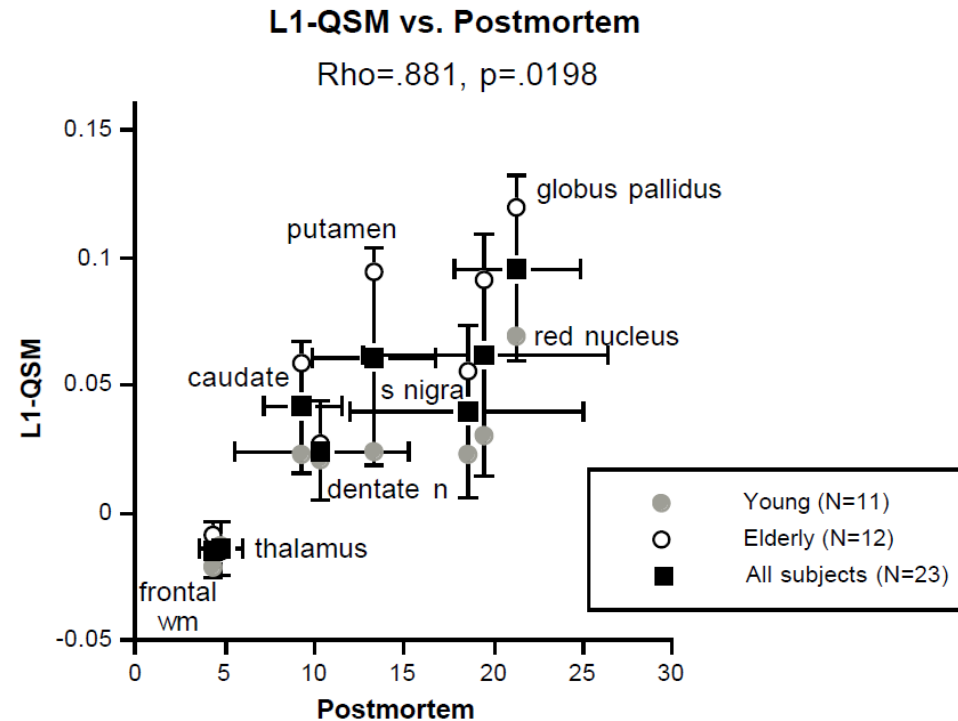
Average L1-QSM Result for the Young

Young substantia nigra: 0.023 ppm

*t*-test result: **p = 0.0178**
significant

- L1-QSM results correlate well with published postmortem results[1], with $Rho = 0.881$, $p = 0.0198$

L1-QSM vs. Postmortem
Rho=.881, p=.0198

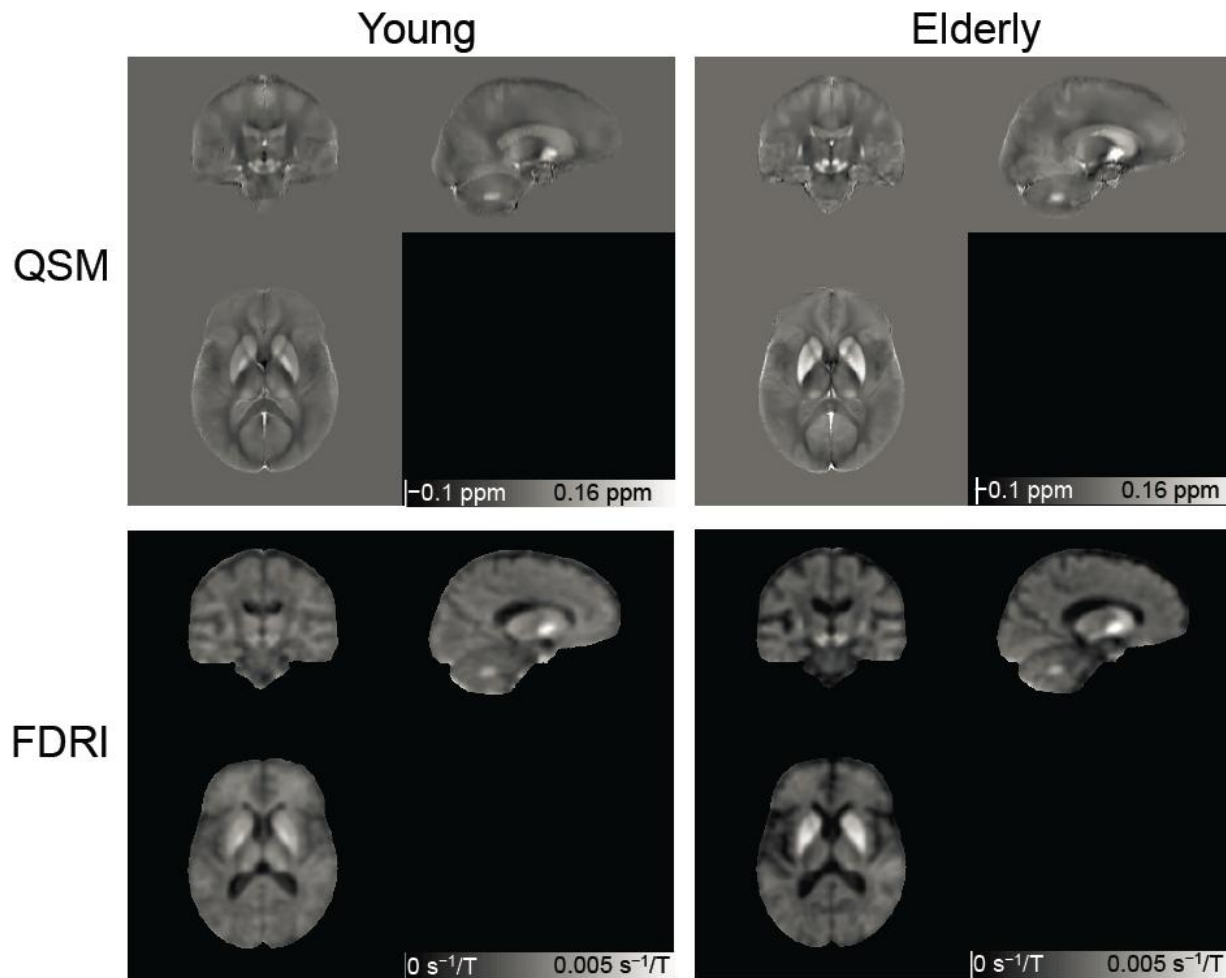[1] Hallgren, B. and Sourander, P. 1958, Journal of Neurochem

L1-QSM vs. FDRI

- Field-Dependent Relaxation Rate Increase (FDRI)[1] is another iron quantification that requires data acquisition at two different main field strengths.

- L1-QSM is strongly correlated with FDRI results, with $Rho = 0.976$, $p = 0.0098$

L1-QSM vs. FDRI (N=23)
Rho=.976, p=.0098

[1] Bartzokis, G. *et al.*, 1993, Magn Res Med
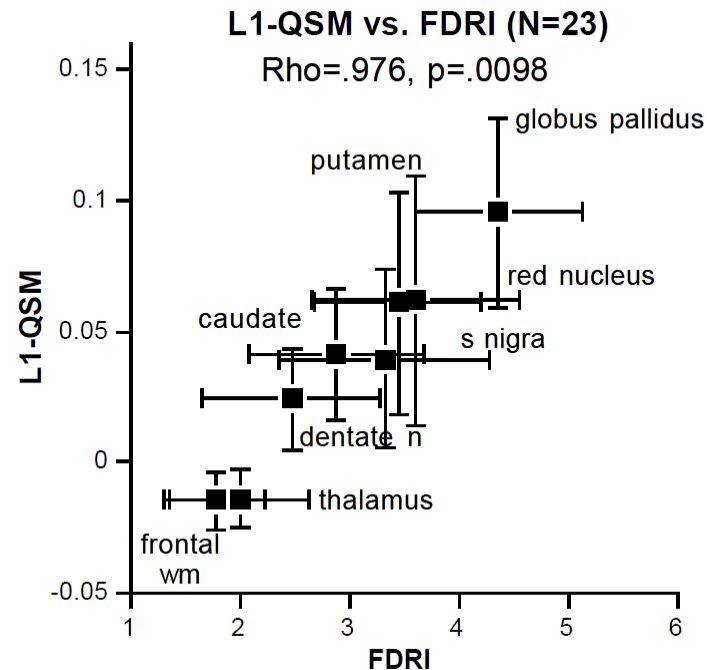
L1-QSM vs. FDRI

- L1-QSM requires data acquisition at a single main magnetic field strength, and has much higher spatial resolution, enabling iron quantification in vessels.

Effect of regularization parameter λ

(a) Over-regularized, $\lambda = 3 \cdot 10^{-3}$

(b) Under-regularized, $\lambda = 3 \cdot 10^{-4}$

-0.2 ppm 0.2 ppm

(c) Optimally regularized, $\lambda = 10^{-3}$