Reconstruction Algorithms for MRI

Berkin Bilgic
17 December 2012
Outline

- Magnetic Resonance Imaging (MRI)
Magnetic Resonance Imaging (MRI)

- Non-invasive imaging, great versatility

structural imaging
## Magnetic Resonance Imaging (MRI)

- Non-invasive imaging, great versatility

Patient with Glioma

- T2-Weighted
- T1-Weighted
- T1-Weighted + Gadolinium

- diffusion imaging
- susceptibility mapping

[1] Burnet et al. Radiotherapy & Oncology '07
Magnetic Resonance Imaging (MRI)

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

- Non-invasive imaging, great versatility
- Inherently slow, protocol takes $\geq 30$ 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

Proton Density  T2-Weighted  T1-Weighted

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):

<table>
<thead>
<tr>
<th>Method</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>SparseMRI</td>
<td>9.4%</td>
</tr>
<tr>
<td>Proposed</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

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

- Proton Density
- T2-Weighted
- T1-Weighted

Proposed: Joint Reconstruction
Bilgic et al. MRM’11

- Difference to fully-sampled: 9.4%
- 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

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

Proposed

Lee et al.
ISMRM’10

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

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
Total Variation prior

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

\[
\min_{\text{img}} \| F_\Omega \cdot \text{img} - \text{data} \|_2^2 + \lambda \cdot \| G \cdot \text{img} \|_1
\]

Undersampled DFT \rightarrow \text{image} \rightarrow \text{k-space samples} \rightarrow \text{gradient operator}
Total Variation prior

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

\[
\min_{img} \| F_\Omega \cdot img - data \|_2^2 + \lambda \cdot \| G \cdot img \|_1
\]

undersampled DFT \hspace{2cm} image \hspace{2cm} k-space samples \hspace{2cm} Total Variation

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\textsuperscript{1} contains such multi-contrast data.

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.

\[1\] 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.

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

T1 weighted

![Brain MRI Image]
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.
To couple multi-contrast signals,

- take the $\ell_2$ norm across the contrast dimension,
- then apply $\ell_1$ regularization to the combination,

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

- Data consistency for $L$ images
- $\ell_2$ across contrasts in transform domain
- $\ell_1$ over combination

Prior: few non-zero rows

$$
\begin{pmatrix}
(\Psi x)_1 \\
\vdots \\
(\Psi x)_L
\end{pmatrix}
$$

- $\ell_2$ across contrasts
Joint reconstruction with $\ell_1$-$\ell_2$ regularization

- To couple multi-contrast signals,
  - take the $\ell_2$ norm across the contrast dimension,
  - then apply $\ell_1$ regularization to the combination,

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

- $\ell_2$ across contrasts in transform domain
- $\ell_1$ over combination
- Data consistency for $L$ images

$\begin{bmatrix}
\ell_2 \text{ combo} \\
\ell_1 \text{ over combo}
\end{bmatrix}$
**Joint reconstruction with \( \ell_1-\ell_2 \) regularization**

- To couple multi-contrast signals,
  - take the \( \ell_2 \) norm across the contrast dimension,
  - then apply \( \ell_1 \) regularization to the combination,

\[
\sum_{i=1}^{L} \| F_{Ω} x_i - y_i \|_2^2 + \lambda \cdot \sum_{j=1}^{N} \left( \sum_{i=1}^{L} (\Psi x)_{i,j}^2 \right)^{1/2}
\]

- \( \ell_2 \) across contrasts in transform domain
- \( \ell_1 \) over combination

- **M-FOCUSS\(^1\)** is an iteratively reweighted \( \ell_2 \) regularization algorithm that solves this optimization problem

---

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

\[ F_{\Omega} x = y \]

- \( F_{\Omega} \): partial Fourier transform
- \( x \): image to be estimated
- \( y \): undersampled k-space data
Observation model – sparse representation

\[ V F_{\Omega} x = V y \]

\[ V = (1 - e^{-2\pi j \omega / n}) \]

\( k \)-space representation of differencing: \( x_i - x_{i-1} \)
Observation model – sparse representation

\[ F_{\Omega} \delta = \tilde{y} \]

\( \delta \): image gradient to be estimated
\( \tilde{y} \): modified k-space data
\[ F_{\Omega} \delta = \tilde{y} \]

\( \delta \): image gradient to be estimated

\( \tilde{y} \): modified k-space data
Data likelihood

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

$$p(\tilde{y} \mid \delta, \sigma^2) \sim \mathcal{N}(F_\Omega \delta - \tilde{y}, \sigma^2)$$

Gaussian likelihood
Prior distribution on gradient coefficients

- Bayesian CS places hyperparameters $\gamma$ on each pixel,

$$p(\delta_i \mid \gamma_i) \sim \mathcal{N}(0, \gamma_i)$$

Gaussian prior

- So that $i^{th}$ pixel is a zero-mean Gaussian with variance $\gamma_i$
Bayesian CS places hyperparameters $\gamma$ on each pixel,

$$p(\delta_i \mid \gamma_i) \sim \mathcal{N}(0, \gamma_i)$$

Gaussian prior

- So that $i^{th}$ pixel is a zero-mean Gaussian with variance $\gamma_i$

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

$$p(\delta \mid \gamma) \sim \prod_i \mathcal{N}(0, \gamma_i)$$
Using the likelihood and the prior, we invoke Bayes’ Rule to arrive at the posterior,

\[ p(\delta | \tilde{y}, \gamma) \propto p(\delta | \gamma) \cdot p(\tilde{y} | \delta) \]
Using the likelihood and the prior, we invoke Bayes’ Rule to arrive at the posterior,

\[
p(\delta | \tilde{y}, \gamma) \propto p(\delta | \gamma) \cdot p(\tilde{y} | \delta)
\]

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

\[ p(\delta | \tilde{y}, \gamma) \sim \mathcal{N}(\mu, \Sigma) \]

\[ \mu = \Gamma F_\Omega^H A^{-1} \tilde{y} \]

\[ \Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma \]
Using the likelihood and the prior, we invoke Bayes’ Rule to arrive at the posterior,

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

$$\mu = \Gamma F_\Omega^H A^{-1} \tilde{y}$$

$$\Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma$$

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

$$A^{-1} = (\sigma^2 I + F_\Omega \Gamma F_\Omega^H)^{-1} \rightarrow 10^4 \times 10^4 \text{ matrix inversion}$$
Using the likelihood and the prior, we invoke Bayes’ Rule to arrive at the posterior,

\[ p(\delta | \tilde{y}, \gamma) \sim \mathcal{N}(\mu, \Sigma) \]

\[ \mu = \Gamma F_{\Omega}^H A^{-1} \tilde{y} \]

\[ \Sigma = \Gamma - \Gamma F_{\Omega}^H A^{-1} F_{\Omega} \Gamma \]

\[ \Gamma = diag(\gamma) \]

\[ A^{-1} = \left( \sigma^2 I + F_{\Omega} \Gamma F_{\Omega}^H \right)^{-1} \]

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:**

\[
\mu = \Gamma F_\Omega^H A^{-1} \tilde{y}
\]

\[
\Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma
\]

**Maximization step:**

\[
\gamma_i = \frac{|\mu_i|^2}{1 - \sum_{ii} / \gamma_i}
\]

**EM algorithm for optimization**

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

\[
\mu = \Gamma F_\Omega^H A^{-1} \tilde{y} \\
\Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma
\]

**Expectation step:**

**Maximization step:**

\[
\gamma_i = \frac{|\mu_i|^2}{1 - \sum_{ii} / \gamma_i}
\]

EM algorithm for optimization

- Expectation-maximization algorithm\(^1\) is used to estimate the hyperparameters and the posterior iteratively,

**Expectation step:**

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

\[
\Sigma = \Gamma - \Gamma \mathbf{F}_\Omega^H \mathbf{A}^{-1} \mathbf{F}_\Omega \Gamma
\]

**Maximization step:**

\[
\gamma_i = \frac{\left|\mathbf{\mu}_i\right|^2}{1 - \sum_{ii}/\gamma_i}
\]

EM algorithm for optimization

- Expectation-maximization algorithm\(^1\) is used to estimate the hyperparameters and the posterior iteratively,

**Expectation step:**

\[
\mu = \Gamma F_{\Omega}^H A^{-1} \tilde{y} \\
\Sigma = \Gamma - \Gamma F_{\Omega}^H A^{-1} F_{\Omega} \Gamma
\]

**Maximization step:**

\[
\gamma_i = \frac{|\mu_i|^2}{1 - \Sigma_{ii}/\gamma_i} \quad \rightarrow \text{for a single image}
\]

EM algorithm for optimization

- Expectation-maximization algorithm\(^1\) is used to estimate the hyperparameters and the posterior iteratively,

**Expectation step:**

\[
\mu = \Gamma F_\Omega^H A^{-1} \widetilde{y} \\
\Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma
\]

**Maximization step:**

\[
\gamma_i = \frac{\|\mu_1, \ldots, \mu_L\|^2}{L - L \cdot \Sigma_{ii}/\gamma_i} \quad \rightarrow \text{for } L \text{ images jointly}
\]

EM algorithm for optimization

- Expectation-maximization algorithm\(^1\) is used to estimate the hyperparameters and the posterior iteratively,

**Expectation step:**

\[
\mu = \Gamma F_\Omega^H A^{-1} \tilde{y}
\]

\[
\Sigma = \Gamma - \Gamma F_\Omega^H A^{-1} F_\Omega \Gamma
\]

**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

---

SRI24 Atlas

k-space, 100% of Nyquist rate

Inverse FFT  Error: 0% RMSE
k-space, 25 % of Nyquist rate

SparseMRI\(^1\) Error: 9.4 % RMSE

[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
$k$-space, 25\% of Nyquist rate

Joint BCS  Error: 2.3\% RMSE

SparseMRI  9.4\%
M-FOCUSS  3.2\%
Joint Bayes  2.3\%
TSE Scans: *in vivo* acquisition

$k$-space
100% of Nyquist rate

Inverse FFT

Error: 0% RMSE
$k$-space, 40% of Nyquist rate

Error: 9.4% RMSE

SparseMRI

[1] Lustig et al. MRM 2007
$k$-space, 40 % of Nyquist rate

M-FOCUSS

Error: 5.1 % RMSE
$k$-space, 40% of Nyquist rate

**Errors:**

- **SparseMRI:** 9.4%
- **M-FOCUSS:** 5.1%
- **Joint Bayes:** 3.6%

**Error:** 3.6% RMSE

**Methods:**

- **Joint BCS**
Demonstrated improved reconstruction quality for multi-contrast imaging by exploiting similarity across contrasts.

Proposed to use two methods for joint reconstruction:

- M-FOCUSS: $\ell_1-\ell_2$ regularization
- Bayesian CS: common variance
<table>
<thead>
<tr>
<th>Method</th>
<th>Speed</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>SparseMRI(^1)</td>
<td>~ minutes</td>
<td>good</td>
</tr>
<tr>
<td>M-FOCUSS</td>
<td>~ minutes</td>
<td>better</td>
</tr>
<tr>
<td>Bayesian CS</td>
<td>~ hours</td>
<td>best</td>
</tr>
</tbody>
</table>

\(^1\) 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

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 $\sim 10^5 \times 10^5$, 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 $\sim 10^5 \times 10^5$, ongoing research

- **Multi-modal Imaging**
  - Extend joint reconstruction to PET / MRI\(^1\) etc.

[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

RF pulses

Gradient

RF signal

Spins are in phase

Dephasing

Rephasing

Signal
Diffusion imaging – moving water molecules

Excite → diffusion encoding → receive

RF pulses:
- 90° pulse
- 180° pulse

Gradient:
- Dephasing
- Rephasing

RF signal:
- Spins are in phase
- Spins are not in phase
- Signal is decreased

Protons move
Weight the diffusion in the desired direction of space using magnetic gradients in 3-D
Diffusion imaging

q-space

90°  180°  RF signal

90°  180°  RF signal

90°  180°  RF signal
Diffusion imaging

- Image intensity attenuation is dependent on water diffusion in each direction
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

  \[ r \sim 10 \mu m \ll 1 \text{ mm (voxel size)} \]

- Tensor representation:

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

\[ D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} \]
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:
  \[ \text{prob}(r, \Delta) \propto \exp \left( -\frac{r^T D^{-1} r}{4\Delta} \right) \]
  - \( r \sim 10 \mu m \)

- Tensor representation:

- Isotropic tissue
- Fibrous tissue

- DTI estimation

- MR signal detected:
  \[ S(g) = S(0) \cdot \exp(-b \cdot \hat{g}^T D \hat{g}) \]
  \[ b \propto G^2 \delta^2 (\Delta - \delta/3) \]
  \( \hat{g} \): unit vector along \( g \)
**Diffusion Tensor Imaging (DTI)**

- Model the water diffusion as Gaussian:
  
  \[
  \hat{g}^T D \hat{g} = \frac{1}{b} \cdot \ln \left( \frac{S(0)}{S(g)} \right)
  \]

  \[
  S(g) = S(0) \cdot \exp(-b \cdot \hat{g}^T D \hat{g})
  \]

- Tensor estimation:
  
  \[
  D = \begin{bmatrix}
  D_{xx} & D_{xy} & D_{xz} \\
  D_{xy} & D_{yy} & D_{yz} \\
  D_{xz} & D_{yz} & D_{zz}
  \end{bmatrix}
  \]

  6 unknowns
Diffusion Tensor Imaging (DTI)

- Model the water diffusion as Gaussian:

\[
\begin{align*}
\mathbf{S}(\mathbf{g}) &= \mathbf{S}(0) \cdot \exp(-b \cdot \hat{\mathbf{g}}^T \mathbf{D} \hat{\mathbf{g}}) \\
\hat{\mathbf{g}}^T \mathbf{D} \hat{\mathbf{g}} &= \frac{1}{b} \cdot \ln \left( \frac{\mathbf{S}(0)}{\mathbf{S}(\mathbf{g})} \right)
\end{align*}
\]

- Tensor estimation:

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

6 unknowns

- At least 6 DWI + 1 non-DWI acquisitions are needed for DTI estimation.
Tensor visualization

- CSF: isotropic
- White matter: anisotropic
1. Define “seed points”
2. Launch the 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

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]

\[
\min_p \| F_\Omega p - q \|_2^2 + \alpha \cdot \| \Psi p \|_1 + \beta \cdot TV(p)
\]

Undersampled q-space of a single voxel

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

**Step 1:** Create dictionary from a training pdf dataset \([P]\)

\[
\min_{P,D} \sum_i \|x_i\|_0 \quad \text{subject to} \quad \|P - DX\|_F^2 \leq \epsilon
\]

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

**Step 2:** Use dictionary to impose sparsity constraint

\[
\min \|x\|_1 \quad \text{such that} \quad F_\Omega DX = q
\]

FOCUSS[2] was used to provide parameter free recon

Methods

- 3 healthy volunteers, 3T Siemens Skyra
Methods

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

\[
G_{\text{max}} = 300 \, \text{mT/m} \\
\text{Conventional} = 45 \, \text{mT/m}
\]

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

† MAGNETOM Skyra CONNECTOM system (Siemens Healthcare)

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 \( \delta \)

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 \( \delta \)

Shorter echo time, higher signal

Methods

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

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, $b_{\text{max}} = 8000$ s/mm$^2$
- 515 q-space points, 50 min scan time
- Number of voxels $= 96 \times 96 \times 57 \approx 500.000$

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, bmax = 8000 s/mm²
- 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
3 healthy volunteers, 3T Siemens Skyra
Connectom gradients, 64-chan head coil
2.3 mm isotropic, \( \text{bmax} = 8000 \text{ s/mm}^2 \)
515 q-space points, 50 min scan time
Number of voxels = \( 96 \times 96 \times 57 \approx 500.000 \)

One dictionary trained with data from each subject
Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, $b_{\text{max}} = 8000 \text{ s/mm}^2$
- 515 q-space points, 50 min scan time
- Number of voxels $= 96 \times 96 \times 57 \approx 500.000$

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

Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, \( b_{\text{max}} = 8000 \text{ s/mm}^2 \)
- 515 q-space points, 50 min scan time
- Number of voxels = \( 96 \times 96 \times 57 \approx 500,000 \)

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

Comparison of methods:

i. Wavelet + TV (Menzel et al [2])
ii. L1-FOCUSS (apply L1 penalty on pdfs)
iii. Dictionary-FOCUSS (proposed)
Methods

- 3 healthy volunteers, 3T Siemens Skyra
- Connectom gradients, 64-chan head coil [1]
- 2.3 mm isotropic, $b_{\text{max}} = 8000$ s/mm$^2$
- 515 q-space points, 50 min scan time
- Number of voxels $= 96 \times 96 \times 57 \approx 500,000$
- 10 average collected at 5 q-space points
Low-noise data, serve as ground truth

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, $b_{\text{max}} = 8000 \text{ s/mm}^2$
- 515 q-space points, 50 min scan time
- Number of voxels = $96 \times 96 \times 57 \approx 500.000$
- 10 average collected at 5 q-space points
  Low-noise data, serve as ground truth

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

2. Menzel MI et al MRM 2011
Subject A, pdf reconstruction error

Wavelet+TV  $\ell_1$-FOCUSS

Acceleration
$R = 3$

- Wavelet+TV @ $R=3$: 15.8% error
- $\ell_1$-FOCUSS @ $R=3$: 15.0% error
Subject A, pdf reconstruction error

Wavelet+TV  15.8% RMSE
\(\ell_1\)-FOCUSS  15.0% RMSE

Acceleration

R = 3

Trained on subject A  7.8% RMSE
Trained on subject B  7.8% RMSE
Trained on subject C  8.2% RMSE

Dictionary-FOCUSS

Trained on subject A  7.8% RMSE
Trained on subject B  7.8% RMSE
Trained on subject C  8.2% RMSE

Dictionary @ R=3  7.8% error
\(\ell_1\)-FOCUSS @ R=3  15.0% error
Wavelet+TV @ R=3  15.8% error

Dictionary @ R=3  7.8% error
Subject A, pdf reconstruction error

Wavelet+TV

$\ell_1$-FOCUSS

Acceleration
R = 3

$\ell_1$-FOCUSS @ R=3  15.0% error

Dictionary @ R=3  7.8% error

Dictionary @ R=5  8.9% error

Dictionary-FOCUSS

Trained on subject A
7.8% RMSE

Trained on subject B
7.8% RMSE

Trained on subject C
8.2% RMSE

Wav+TV @ R=3  15.8% error
Subject A, pdf reconstruction error

Wavelet+TV  \( \ell_1 \)-FOCUSS

Wavelet+TV @ R=3  \( \ell_1 \)-FOCUSS @ R=3

Dictionary @ R=3  Dictionary @ R=5

Dictionary @ R=9

<table>
<thead>
<tr>
<th>Acceleration</th>
<th>R = 3</th>
<th>R = 5</th>
<th>R = 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained on subject A</td>
<td>7.8% RMSE</td>
<td>8.9% RMSE</td>
<td>10.0% RMSE</td>
</tr>
<tr>
<td>Trained on subject B</td>
<td>7.8% RMSE</td>
<td>8.9% RMSE</td>
<td>10.0% RMSE</td>
</tr>
<tr>
<td>Trained on subject C</td>
<td>8.2% RMSE</td>
<td>9.3% RMSE</td>
<td>10.4% RMSE</td>
</tr>
</tbody>
</table>

15.8% RMSE  15.0% RMSE
Subject A, pdf reconstruction error

Wavelet+TV  \(\ell_1\)-FOCUSS

\begin{align*}
\text{Acceleration} & \quad R = 3 \\
15.8\% \text{ RMSE} & \quad 15.0\% \text{ RMSE}
\end{align*}

\begin{align*}
\text{Acceleration} & \quad R = 5 \\
& \quad \text{Trained on subject A} \\
& \quad 7.8\% \text{ RMSE} \\
& \quad \text{Trained on subject B} \\
& \quad 7.8\% \text{ RMSE} \\
& \quad \text{Trained on subject C} \\
& \quad 8.2\% \text{ RMSE}
\end{align*}

\begin{align*}
\text{Acceleration} & \quad R = 9 \\
& \quad \text{Dictionary-FOCUSS} \\
& \quad 8.9\% \text{ RMSE} \\
& \quad \text{Dictionary-FOCUSS} \\
& \quad 8.9\% \text{ RMSE} \\
& \quad \text{Dictionary-FOCUSS} \\
& \quad 9.3\% \text{ RMSE}
\end{align*}

Wav+TV @ R=3 \quad 15.8\% \text{ error}

\(\ell_1\)-FOCUSS @ R=3 \quad 15.0\% \text{ error}

Dictionary @ R=3 \quad 7.8\% \text{ error}

Dictionary @ R=5 \quad 8.9\% \text{ error}

Dictionary @ R=9 \quad 10.0\% \text{ error}

\text{108} \quad 10.0\% \text{ RMSE} \quad 10.0\% \text{ RMSE} \quad 10.4\% \text{ RMSE}
q-space reconstructions at q=[5,0,0]

- Wavelet+TV
- $l_1$-FOCUSS
- Dict-FOCUSS
- Fully-sampled

% RMSE in q-space vs. missing q-space directions

$q=[5,0,0]$
Fully-sampled 10 average

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

- Wavelet+TV
- $\ell_1$-FOCUSS
- Dict-FOCUSS

% RMSE in q-space

- Dictionary-FOCUSS
- Wavelet + TV
- L1-FOCUSS

Missing q-space directions increasing |q|

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

- Wavelet+TV
  - poor performance

- $\ell_1$-FOCUSS
  - same $\ell_2$ norm as 10 average
  - good performance

- Dict-FOCUSS
  - 10 average

- Fully-sampled

% RMSE in q-space

Increasing $|q|$
- 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

![Graph showing RMSE comparison](image)

Lower RMSE than acquired data

Denoising effect [1]

Tractography solutions for subject A

Fully-sampled data

Dictionary-FOCUSS recon with 3-fold acceleration
Tractography solutions for subject A

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

1. Yendiki A et al
   Front Neuroinform 2011
Mean FA error = 3%
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 $\approx$ 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 $\ell_1$, applies $\ell_2$ 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 $\ell_1$, applies $\ell_2$ 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: $\ell_2$ regularization

- Dictionary-FOCUSS iteratively solves

$$\min \| x \|_1 \text{ such that } \mathbf{F}_\Omega \mathbf{D} x = q$$
Dictionary-FOCUSS iteratively solves

$$min \| x \|_1 \text{ such that } F_\Omega D x = q$$

Instead, consider

$$min \| F_\Omega D x - q \|_2^2 + \lambda \cdot \| x \|_2^2$$
PINV: $\ell_2$ regularization

- Dictionary-FOCUSS iteratively solves
  \[ \min \|x\|_1 \text{ such that } F_\Omega Dx = q \]

- Instead, consider
  \[ \min \|F_\Omega Dx - q\|_2^2 + \lambda \cdot \|x\|_2^2 \]

- Solution: \( \tilde{x} = ((F_\Omega D)^H F_\Omega D + \lambda I)^{-1} (F_\Omega D)^H q \)
PINV: $\ell_2$ regularization

- Dictionary-FOCUSS iteratively solves
  \[
  \min \| x \|_1 \quad \text{such that} \quad F_{\Omega} D x = q
  \]

- Instead, consider
  \[
  \min \| F_{\Omega} D x - q \|_2^2 + \lambda \cdot \| x \|_2^2
  \]

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

Singular Value Decomposition: $F_{\Omega} D = U \Sigma V^H$
PINV: $\ell_2$ regularization

- Dictionary-FOCUSS iteratively solves
  \[
  \min \|x\|_1 \text{ such that } F_\Omega Dx = q
  \]
- Instead, consider
  \[
  \min \|F_\Omega Dx - q\|_2^2 + \lambda \cdot \|x\|_2^2
  \]
- Solution:  
  \[
  \tilde{x} = (\left((F_\Omega D)^H F_\Omega D + \lambda I\right)^{-1}(F_\Omega D)^H q
  \]
  
  \[
  F_\Omega D = U\Sigma V^H
  \]
  
  \[
  \Sigma^+ = (\Sigma^H \Sigma + \lambda I)^{-1} \Sigma^H
  \]
  
  compute once
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:

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

$$
\mathbf{ZZ}^H = \mathbf{Q}\Lambda\mathbf{Q}^H
$$
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:

\[
\mathbf{Z} = \mathbf{P} - \mathbf{p}_{\text{mean}}
\]

\[
\mathbf{Z\mathbf{Z}}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H
\]

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

\[
\mathbf{pca} = \mathbf{Q}_T^H(\mathbf{p} - \mathbf{p}_{\text{mean}})
\]

$T$ - 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:

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

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

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

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

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

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

$$
\min \| F_\Omega p_T - q \|_2^2
$$
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,
  \[ \min ||F_\Omega p_T - q||^2 \]

- In PCA coordinates,
  \[ \min_{pca} ||F_\Omega Q_T pca - (q - F_\Omega p_{mean})||^2 \]
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,

$$\min_{F\Omega p_T - q} 2$$

- In PCA coordinates,

$$\min_{pc\alpha} \| F\Omega Q_T p\alpha - (q - F\Omega p_{mean}) \|^2_2$$

- Closed-form solution:

$$\tilde{p\alpha} = \text{pinv}(F\Omega Q_T)(q - F\Omega p_{mean})$$

    \[\text{compute once}\]
Selection of regularization parameters

- PINV: selection of $\lambda$

\[
\min \| F_\Omega D x - q \|_2^2 + \lambda \cdot \| x \|_2^2
\]
Selection of regularization parameters

- PINV: selection of $\lambda$

$$
\min \left\| F_\Omega D x - q \right\|_2^2 + \lambda \cdot \| x \|_2^2
$$

- PCA: selection of PCA dimension $T$ in $Q_T$

$$
\min_{pca} \left\| F_\Omega Q_T pca - (q - F_\Omega p_{mean}) \right\|_2^2
$$
Selection of regularization parameters

- PINV: selection of $\lambda$
  
  $$
  \min \| F_{\Omega} D x - q \|_2^2 + [\lambda] \cdot \| x \|_2^2
  $$

- PCA: selection of PCA dimension $T$ in $Q_T$
  
  $$
  \min_{pca} \| F_{\Omega} Q_T pca - (q - F_{\Omega} p_{mean}) \|_2^2
  $$

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

- Find $\lambda$ and $T$ that yields the lowest reconstruction error on $P$
Subject A, pdf reconstruction error

Wavelet+TV  \(\ell_1\)-FOCUSS

\[
\begin{array}{cc}
\text{15.8\% RMSE} & \text{15.0\% RMSE} \\
\end{array}
\]

Dict-FOCUSS PINV PCA

\[
\begin{array}{ccc}
\text{7.8\% RMSE} & \text{8.1\% RMSE} & \text{8.7\% RMSE} \\
\end{array}
\]

\[
\begin{array}{cc}
\text{1190 min} & \text{26 min} \\
\end{array}
\]

Recon Time

\[
\begin{array}{ccc}
\text{530 min} & \text{0.6 min} & \text{0.4 min} \\
\end{array}
\]

Acceleration

\(R = 3\)
Subject A, pdf reconstruction error

Slice 40

Wavelet+TV $\ell_1$-FOCUSS

15.8% RMSE 15.0% RMSE

Acceleration
R = 3

1190 min 26 min Recon Time

Dict-FOCUSS PINV PCA

7.8% RMSE 8.1% RMSE 8.7% RMSE

8.9% RMSE 8.9% RMSE 9.6% RMSE

Recon Time: 530 min 0.6 min 0.4 min

Acceleration
R = 5

143
Subject A, pdf reconstruction error

Slice 40

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE</th>
<th>Recon Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet+TV</td>
<td>15.8%</td>
<td>1190 min</td>
</tr>
<tr>
<td>$l_1$-FOCUSS</td>
<td>15.0%</td>
<td>26 min</td>
</tr>
<tr>
<td>Dict-FOCUSS</td>
<td>7.8%</td>
<td>530 min</td>
</tr>
<tr>
<td>PINV</td>
<td>8.1%</td>
<td>0.6 min</td>
</tr>
<tr>
<td>PCA</td>
<td>8.7%</td>
<td>0.4 min</td>
</tr>
<tr>
<td>PINV</td>
<td>8.7%</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>9.6%</td>
<td></td>
</tr>
<tr>
<td>PINV</td>
<td>9.6%</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>10.2%</td>
<td></td>
</tr>
<tr>
<td>PINV</td>
<td>10.2%</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>11.2%</td>
<td></td>
</tr>
</tbody>
</table>

Acceleration $R = 3$

Acceleration $R = 5$

Acceleration $R = 9$
Subject A, recon error across slices

<table>
<thead>
<tr>
<th>Slice no</th>
<th>Avg % RMSE in PDFs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>44</td>
</tr>
</tbody>
</table>

- Dictionary-FOCUSS
- PINV
- PCA
Comparison to Low-Noise 10 avg Data

![Comparison to Low-Noise 10 avg Data](image)

- (Fully-sampled 1avg) vs. 10avg
- Dictionary-FOCUSS vs. 10avg
- Tikhonov vs. 10avg
- PCA vs. 10avg
- Wavelet+TV vs. 10avg
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 $\chi$: degree of magnetization of a material when placed in a magnetic field
Susceptibility of Tissue

- Susceptibility $\chi$: degree of magnetization of a material when placed in a magnetic field.

- Diamagnetic: $\chi < 0$
- Paramagnetic: $\chi > 0$

$\chi_{\text{water}} = -9 \text{ ppm}$

$\chi_{\text{iron}} \gg 0$

$\chi = 0$
Susceptibility of Tissue

- Susceptibility $\chi$: degree of magnetization of a material when placed in a magnetic field

$$\chi_{\text{water}} = -9 \text{ ppm}$$

$$\chi_{\text{iron}} \gg 0$$

$$\chi = 0$$

- Susceptibility of brain tissue is $\approx -9$ ppm

- Tissues with increased iron deposition are relatively paramagnetic $\rightarrow \chi$ is more positive

Susceptibility of Tissue

- Susceptibility $\chi$: degree of magnetization of a material when placed in a magnetic field

  $\chi_{\text{water}} = -9$ ppm
  
  $\chi_{\text{iron}} >> 0$

  $\chi = 0$

- Susceptibility of brain tissue is $\approx -9$ ppm

- Tissues with increased iron deposition are relatively paramagnetic $\rightarrow \chi$ 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 $\chi$: degree of magnetization of a material when placed in a magnetic field

\[ \chi_{\text{water}} = -9 \text{ ppm} \]

- Variations in tissue susceptibility affects the magnetic field

\[ \Delta \chi \rightarrow \text{magnetic field perturbation} \]
Susceptibility of Tissue

- Susceptibility $\chi$: degree of magnetization of a material when placed in a magnetic field

$\chi_{\text{water}} = -9$ ppm

- Variations in tissue susceptibility affects the magnetic field

- Field perturbation causes a change in MR signal phase

$\Delta \chi \rightarrow$ magnetic field perturbation  $\rightarrow \Delta \phi$

$\chi_{\text{iron}} \gg 0$

estimate  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 \textit{et al}., PNAS 2007
\(^2\) Fan AP \textit{et al}., ISMRM 2010
\(^3\) Liu T \textit{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 \(\chi\) from the unwrapped phase \(\phi\) involves solving an inverse problem,

\[
\delta = F^{-1}DF\chi
\]

\(F\): Discrete Fourier Transform  
\(D\): susceptibility kernel  
\(\delta = \frac{\phi}{\gamma \cdot TE \cdot B_0}\): 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 \(\chi\) from the unwrapped phase \(\phi\) involves solving an inverse problem,

\[
\delta = F^{-1}DF\chi
\]

measured \hspace{1cm} to be estimated

\(^1\) Duyn JH et al., PNAS 2007 \hspace{1cm} \(^2\) Fan AP et al., ISMRM 2010 \hspace{1cm} \(^3\) 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 \(\chi\) from the unwrapped phase \(\phi\) involves solving an inverse problem, \(\delta = F^{-1}DF\chi\)

- The inversion is made difficult by zeros on a conical surface in susceptibility kernel \(D\)

\[
D = \frac{1}{3} - \frac{k_z^2}{k^2}
\]

\(^1\) Duyn JH \textit{et al.}, PNAS 2007  \(^2\) Fan AP \textit{et al.}, ISMRM 2010  \(^3\) Liu T \textit{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 \(\chi\) from the unwrapped phase \(\phi\) involves solving an inverse problem, \(\delta = F^{-1}DF\chi\)

- Undersampling is due to physics

- Not in our control

\(^1\) Duyn JH et al., PNAS 2007  \(^2\) Fan AP et al., ISMRM 2010  \(^3\) Liu T et al., ISMRM 2010
Regularized Inversion for QSM
Regularized Inversion for QSM

\[ \delta = F^{-1} D F \chi \]

- Solving for \( \chi \) by convolving with the inverse of \( D \) is not possible, as it diverges along the magic angle
Regularized Inversion for QSM

\[ \delta = \text{inv} D F \chi \]

- Solving for \( \chi \) by convolving with the inverse of \( D \) is not possible, as it diverges along the magic angle.
- Use inverse problem formulation, apply regularization.

\[ \log |D^{-1}| \]

\[ |D| \]
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 \( \sim 10 \times \) 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. Remove background phase

Phase accrued due to air-tissue interfaces needs to be removed
This background component is $\sim 10 \times$ larger than tissue phase

\[ initial \ phase \quad background \ phase^1 \]

\[ -0.8 \ ppm \quad 0.8 \ ppm \quad -0.8 \ ppm \quad 0.8 \ ppm \]

\[ ^1 \text{Liu T, NMR in Biomedicine 2011} \]
Phase Processing

- Several processing steps are required to obtain the tissue phase
  
  i. Mask out the skull
  ii. Unwrap the phase
  
  iii. Remove background phase

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

  This background component is \( \sim 10 \times \) larger than tissue phase

![Initial phase](image1)

![Background phase](image2)

![Tissue phase](image3)
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 $\sim 10 \times$ larger than tissue phase

- Now we can solve for $\chi$ from tissue phase $\delta$

  $$\delta = F^{-1}DF\chi$$
L1 Regularized Susceptibility Inversion

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

\[ \text{Find } \chi \text{ such that } \delta = F^{-1}DF\chi \]

- Susceptibility values are tied to the magnetic properties of the underlying tissues; hence they vary smoothly within anatomical boundaries.
We seek the susceptibility map that matches the observed tissue phase,

Find $\chi$ such that $\delta = F^{-1}DF\chi$

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 $\chi$
L1 Regularized Susceptibility Inversion

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

\[
\chi_{tissue} = \arg \min_{\chi} \left\| \delta - F^{-1}DF\chi \right\|^2 + \lambda \cdot \| G\chi \|_1
\]

- Data consistency
- $\ell_1$ over gradients
We solve for the susceptibility distribution with a convex program,

\[ \chi_{tissue} = \arg\min_{\chi} \| \delta - F^{-1}DF\chi \|_{2}^{2} + \lambda \cdot \| G\chi \|_{1} \]

Here, \( \lambda \) serves as a regularization parameter that adjusts the smoothness of the solution.
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
**Striatal ROIs**

**Brain Stem ROIs**

### Elderly >> Young Iron Deposition

- **Putamen** $p=0.0004$
- **Globus Pallidus** $p=0.001$
- **Red Nucleus** $p=0.002$
- **Substantia Nigra** $p=0.003$

---

**Average QSM for the Young**

**Average QSM for the Elderly**
QSM vs. Postmortem

- QSM results correlate well with published postmortem results\(^1\), with \( \text{Rho} = 0.881, p = 0.0198 \)

\(^1\) 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.

![FDRI comparison](image-a)

![QSM comparison](image-b)
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
  
  
  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 \((k_x, k_y, k_z)\)
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 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.

\[
SNR = V_{\text{size}} \times \sqrt{T_{\text{acq}}}
\]
**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.

\[ SNR = V_{size} \times \sqrt{T_{acq}} \]
**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.

\[
SNR = V_{\text{size}} \times \sqrt{T_{\text{acq}}}
\]
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.

\[
SNR = V_{\text{size}} \times \sqrt{T_{\text{acq}}}
\]

Water suppression + lipid suppression
Lipid signals in Spectroscopy

- Voxel sizes in spectroscopy are typically large \(\sim 1\text{cm}^3\)
- 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 \( \sim 1 \text{cm}^3 \)
- 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

![Diagram showing lipid layer, low resolution, and lipid ringing](image-url)
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

---

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**
  - Lipid and metabolite spectra are approximately orthogonal
  - Inside the brain, inner product of metabolites and lipids should be small

---

1 Lee & Adalsteinsson ISMRM 2010
Orthogonality of metabolite and lipid spectra

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

![Graph showing metabolite and lipid spectra](image)
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)

  \[ \text{Compute the projection of metabolite signal onto the lipid spectra and the orthogonal component} \]

  \[ \text{meta}_\parallel = \frac{\text{metabolite}^H \text{lipid}}{\|\text{lipid}\|_2^2} \cdot \text{lipid} \]

  \[ \text{meta}_\perp = \text{metabolite} - \text{meta}_\parallel \]
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

\[
\frac{\| \text{meta} \|_2^2}{\| \text{meta}_\perp \|_2^2} = 7.5\% 
\]
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 $x_{lipid}$

  $$x_{lipid} = M_{lipid}F_{high}^{-1}y_{high}$$

  - $M_{lipid}$: lipid mask
  - $y_{high}$: high-res k-space data
  - $F_{high}$: high-res DFT operator
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 $x_{lipid}$
  \[ x_{lipid} = M_{lipid} F_{high}^{-1} y_{high} \]

- Compute the dual-density image (combine $x_{lipid}$ with low-res CSI)
  \[ x_{dual} = F_{high}^{-1} \left( (F_{high} - F_{low}) x_{lipid} + y_{low} \right) \]

$y_{low}$: low-res k-space data
$F_{low}$: low-res DFT operator
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 $x_{lipid}$
  $$x_{lipid} = M_{lipid} F_{high}^{-1} y_{high}$$

- Compute the dual-density image (combine $x_{lipid}$ with low-res CSI)
  $$x_{dual} = F_{high}^{-1} \{(F_{high} - F_{low})x_{lipid} + y_{low}\}$$

- Make a lipid-basis matrix whose columns are lipid spectra in $x_{dual}$ and enforce orthogonality between metabolites and lipids
  $$x_{basic} = \arg\min_x \left\| F_{high} x - y_{dual} \right\|_2^2 + \lambda \cdot \sum_{i \in M_{brain}} \left\| L_{dual}^H x_i \right\|_1$$

$L_{dual}$ : lipid-basis matrix
$M_{brain}$ : brain mask
Proposed method - II

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

![Diagram showing lipid layer and lipid spectra]
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

![Diagram showing k-space sampling and undersampling]

- 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 \( \sim \) 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 \( \ell_1 \) penalty in space and frequency:

\[
\begin{align*}
W_{j,j}^t &= \text{diag} \left( |x_j^t|^{1/2} \right) \\
q^t &= \text{argmin}_q \| q \|_2^2 \quad \text{such that} \quad M_{\Omega} F_{\text{high}} W^t q = y_{\text{high}} \\
x^{t+1} &= W^t q^t
\end{align*}
\]

For iteration number \( t = 1, \ldots, T \),

- \( M_{\Omega} \): k-space undersampling mask
- \( x^{T+1} \): CS recon for high-res lipid image

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 cm² FOV inside the brain
Lipid Maps at $TE = 50$ ms

(a) Gold standard
$20$ avg$_{high}$, $R_{high} = 1$

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

20 avg for metabolites at Nyquist rate
20 avg for lipid at Nyquist rate

k-space sampling

15 dB

-35 dB
Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg\textsubscript{high}, R\textsubscript{high}=1

(b) Proposed 1
2 avg\textsubscript{high}, R\textsubscript{high}=1

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\(_{\text{high}}\), \(R_{\text{high}}=1\)

(b) Proposed 1
2 avg\(_{\text{high}}\), \(R_{\text{high}}=1\)

(c) Proposed 2
2 avg\(_{\text{high}}\), \(R_{\text{high}}=10\)

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
Lipid Maps at TE = 50 ms

(a) Gold standard
20 $\text{avg}_{\text{high}}, R_{\text{high}}=1$

(b) Proposed 1
2 $\text{avg}_{\text{high}}, R_{\text{high}}=1$

(c) Proposed 2
2 $\text{avg}_{\text{high}}, R_{\text{high}}=10$

(d) Lipid-basis penalty

Lipid-basis applied to 0.5cc data

20 $\text{avg}$ for metabolites at Nyquist rate

k-space sampling
Lipid Maps at TE = 50 ms

(a) Gold standard
20 avg\textsubscript{high}, R\textsubscript{high}=1

(b) Proposed 1
2 avg\textsubscript{high}, R\textsubscript{high}=1

(c) Proposed 2
2 avg\textsubscript{high}, R\textsubscript{high}=10

(d) Lipid-basis penalty
(e) Dual-density
Lipid Maps at TE = 50 ms

(a) Gold standard
20 $\text{avg}_{\text{high}}, R_{\text{high}}=1$

(b) Proposed 1
2 $\text{avg}_{\text{high}}, R_{\text{high}}=1$

(c) Proposed 2
2 $\text{avg}_{\text{high}}, R_{\text{high}}=10$

(d) Lipid-basis penalty

(e) Dual-density

(f) No lipid suppression
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
  - Spectra from OVS: in black
  - Reconstructed spectra: in blue
  - NAA RMSE = 12.9%

- **Proposed 1**
  - 2 avg\(_{\text{high}}\), \(R_{\text{high}}=1\)
  - NAA RMSE = 11.5%

- **Proposed 2**
  - 2 avg\(_{\text{high}}\), \(R_{\text{high}}=10\)
  - NAA RMSE = 12.9%

- **Lipid-basis penalty**
  - NAA RMSE = 14.7%
Comparison with Outer Volume Suppression, TE = 50ms

- Spectra from OVS: in black
- Reconstructed spectra: in blue
Comparison with Outer Volume Suppression, TE = 50ms

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

Ground Truth NAA

Outer Volume Suppression

Proposed 1
2 avg\(_{\text{high}}\), R\(_{\text{high}}\)=1
NAA RMSE = 11.5%

Proposed 2
2 avg\(_{\text{high}}\), R\(_{\text{high}}\)=10
NAA RMSE = 12.9%
Lipid-basis penalty vs. gold-standard

Proposed 2 (2 avg\textsubscript{high}, R\textsubscript{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
  \[ p(\delta | \gamma) \sim \mathcal{N}(0, \gamma) \]
  this does not constitute a sparse prior

- To promote sparsity, Gamma priors are placed over the variances \( \gamma \)
  \[ p(\gamma | a, b) \sim \Gamma(\gamma^{-1} | a, b) \]
Marginal prior on signal coefficients promotes sparsity

- We can marginalize over $\gamma$ and obtain the *marginal* prior

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

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

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