

Fast Regularized Reconstruction Tools for QSM and DSI

Berkin Bilgic^{1,2}, Itthi Chatnuntawech¹, Kawin Setsompop^{2,3}, Audrey P. Fan¹, Stephen F. Cauley², Lawrence L. Wald^{2,4}, E. Adalsteinsson^{1,4}

¹MIT, Cambridge, MA USA

²Martinos Center for Biomedical Imaging, Charlestown, MA, USA

³Harvard Medical School, Boston, MA, USA

⁴Harvard-MIT Health Sciences and Technology, Cambridge, MA USA



L2-Regularized Reconstruction

- L2-regularized recon admits closed-form solutions that can be computed efficiently
- Matlab tools that achieve dramatic speed-up relative to iterative algorithms will be presented



L2-Regularized Reconstruction

- L2-regularized recon admits closed-form solutions that can be computed efficiently
- Matlab tools that achieve dramatic speed-up relative to iterative algorithms will be presented
- Two representative applications are considered,
 - Quantitative Susceptibility Mapping (QSM)
 - □ Diffusion Spectrum Imaging (DSI)



- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility χ
- Susceptibility correlates well with tissue iron concentration, especially in iron rich deep gray matter structures [1,2]





- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility χ
- Susceptibility correlates well with tissue iron concentration, especially in iron rich deep gray matter structures [1,2]
- Susceptibility mapping requires the solution of an inverse problem,





- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility χ
- Susceptibility correlates well with tissue iron concentration, especially in iron rich deep gray matter structures [1,2]
- Susceptibility mapping requires the solution of an inverse problem,

$$\mathbf{F}^H \mathbf{D} \mathbf{F} \mathbf{\chi} = \mathbf{\phi}$$
to be estimated measured





- Quantitative Susceptibility Mapping (QSM) aims to quantify tissue magnetic susceptibility χ
- Susceptibility correlates well with tissue iron concentration, especially in iron rich deep gray matter structures
- Susceptibility mapping requires the solution of an inverse problem,

$${f F}^H \, {f D} \, {f F} \, {m \chi} = {m \phi}$$
 ${f D} = {f D} \, {f D} \, {f D} \, {f D} \, {f C} \, {f D} \, {f D} \, {f C} \, {f D} \, {f C} \,$





 Solution of inverse problem is facilitated by regularization that imposes prior knowledge [1]

$$\chi = argmin_{\chi} \|\phi - \mathbf{F}^{H} \mathbf{D} \mathbf{F} \chi\|_{2}^{2} + \lambda \cdot \|\mathbf{G} \chi\|_{2}^{2}$$

$$\text{data consistency} \qquad \ell_{2} \text{ over gradients}$$



 Solution of inverse problem is facilitated by regularization that imposes prior knowledge [1]

$$\mathbf{G} = egin{bmatrix} \mathbf{G}_{m{x}} \ \mathbf{G}_{m{y}} \ \mathbf{G}_{m{z}} \end{bmatrix}$$

gradient in 3D





 Solution of inverse problem is facilitated by regularization that imposes prior knowledge [1]

Prior: underlying susceptibility map is smooth





 Solution of inverse problem is facilitated by regularization that imposes prior knowledge [1]

Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$

 The minimizer can be computed efficiently given that the matrix inversion is rapidly performed





Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$



Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$

 Gradient in image space can be represented in k-space by multiplication with a diagonal matrix E

$$\mathbf{G} = \mathbf{F}^H \mathbf{E} \mathbf{F}$$

where
$$\mathbf{E}(i, i) = 1 - e^{(-2\pi\sqrt{-1}k(i, i)/N)}$$



Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$

 Gradient in image space can be represented in k-space by multiplication with a diagonal matrix E

G = **F**^H**E F** where **E**(*i*, *i*) =
$$1 - e^{(-2\pi\sqrt{-1}k(i,i)/N)}$$

■ **E** is simply the k-space representation of the difference operator $\delta_i - \delta_{i-1}$



Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$

 Gradient in image space can be represented in k-space by multiplication with a diagonal matrix E

G = **F**^H**E F** where **E**(*i*, *i*) =
$$1 - e^{(-2\pi\sqrt{-1}k(i,i)/N)}$$

With this formulation, closed-form solution becomes

$$\chi = \mathbf{F}^H \mathbf{D} [\mathbf{D}^2 + \lambda \cdot (\mathbf{E}_x^2 + \mathbf{E}_y^2 + \mathbf{E}_z^2)]^{-1} \mathbf{F} \boldsymbol{\phi}$$
all matrices diagonal





Solution can be evaluated in closed-form

$$\chi = (\mathbf{F}^H \mathbf{D}^2 \mathbf{F} + \lambda \cdot \mathbf{G}^H \mathbf{G})^{-1} \mathbf{F}^H \mathbf{D} \mathbf{F} \boldsymbol{\phi}$$

 Gradient in image space can be represented in k-space by multiplication with a diagonal matrix E

G = **F**^H**E F** where **E**(*i*, *i*) =
$$1 - e^{(-2\pi\sqrt{-1}k(i,i)/N)}$$

With this formulation, closed-form solution becomes

$$\chi = \mathbf{F}^H \mathbf{D} [\mathbf{D}^2 + \lambda \cdot (\mathbf{E}_{\chi}^2 + \mathbf{E}_{y}^2 + \mathbf{E}_{z}^2)]^{-1} \mathbf{F} \boldsymbol{\phi}$$

Total cost: Two FFTs and multiplication of diagonal matrices





- Numerical Phantom
 - \square Three compartments (gray, white, CSF) with constant χ
 - □ Phase ϕ computed from true χ , and peak-SNR = 100 noise added
 - $lue{}$ Regularization parameter λ chosen to minimize RMSE in reconstructed χ



- Numerical Phantom
 - $lue{}$ Three compartments (gray, white, CSF) with constant χ
 - $lue{}$ Phase $oldsymbol{\phi}$ computed from true $oldsymbol{\chi}$, and peak-SNR = 100 noise added
 - lue Regularization parameter λ chosen to minimize RMSE in reconstructed χ
- In Vivo 3D SPGR
 - □ Healthy subject at 1.5T with resolution 0.94×0.94×2.5mm³
 - \square Regularization parameter λ chosen based on L-curve heuristic



- Numerical Phantom
 - \square Three compartments (gray, white, CSF) with constant χ
 - □ Phase ϕ computed from true χ , and peak-SNR = 100 noise added
 - \square Regularization parameter λ chosen to minimize RMSE in reconstructed χ
- In Vivo 3D SPGR
 - ☐ Healthy subject at 1.5T with resolution 0.94×0.94×2.5mm³
 - \blacksquare Regularization parameter λ chosen based on L-curve heuristic
- Comparison of methods
 - i. Iterative solution using Nonlinear Conjugate Gradient [1,2]
 - ii. Proposed closed-form solution





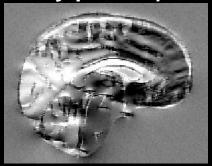
- Numerical Phantom
 - $lue{}$ Three compartments (gray, white, CSF) with constant χ
 - □ Phase ϕ computed from true χ , and peak-SNR = 100 noise added
 - \square Regularization parameter λ chosen to minimize RMSE in reconstructed χ
- In Vivo 3D SPGR
 - □ Healthy subject at 1.5T with resolution 0.94×0.94×2.5mm³
 - \blacksquare Regularization parameter λ chosen based on L-curve heuristic
- Comparison of methods
 - i. Iterative solution \rightarrow converges to closed-form solution
 - ii. Proposed closed-form solution





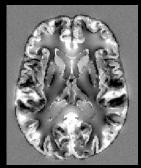
Numerical Phantom

Noisy phase ϕ



error due to noise: 5.9% RMSE





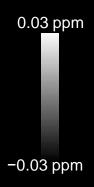
0.01 ppm -0.01 ppm

Closed-form QSM in 3.3 seconds

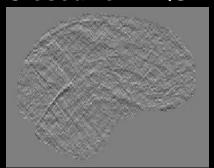


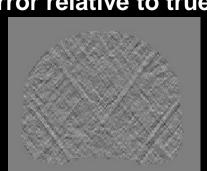


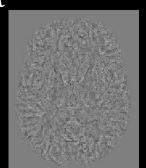


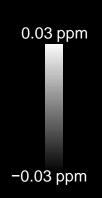


Closed-form QSM error relative to true χ







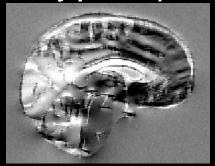






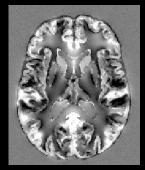
Numerical Phantom

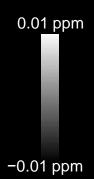
Noisy phase ϕ



error due to noise: 5.9% RMSE





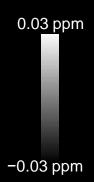


Closed-form QSM in 3.3 seconds





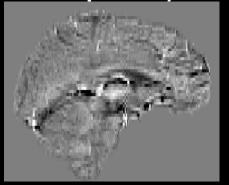


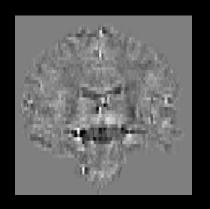


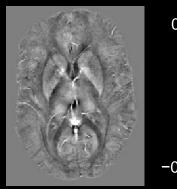
QSM Method	Recon Time	Error relative to true χ
Closed-form	3.3 seconds	17.4% RMSE
Iterative [1,2], 100 iters	65 minutes	18.0% RMSE

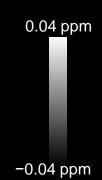




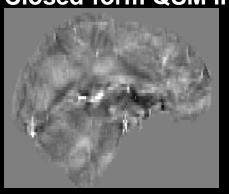


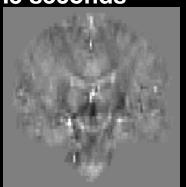


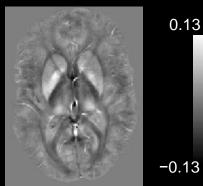


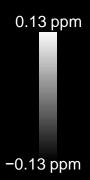


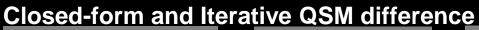
Closed-form QSM in 1.3 seconds











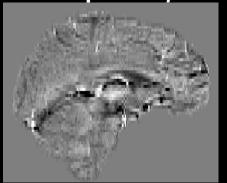


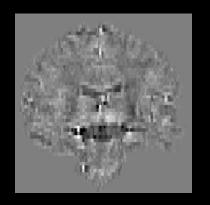


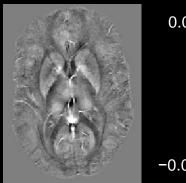


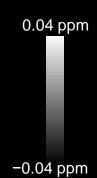
In Vivo QSM

Tissue phase ϕ

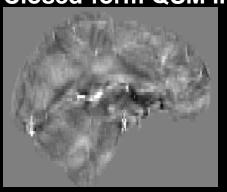


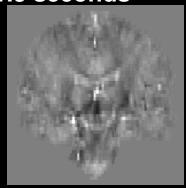


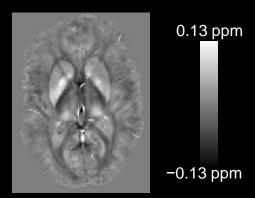




Closed-form QSM in 1.3 seconds







QSM Method	Recon Time
Closed-form	1.3 seconds
Iterative Conj Grad [1,2], 100 iters	29 minutes



Diffusion Spectrum Imaging (DSI)

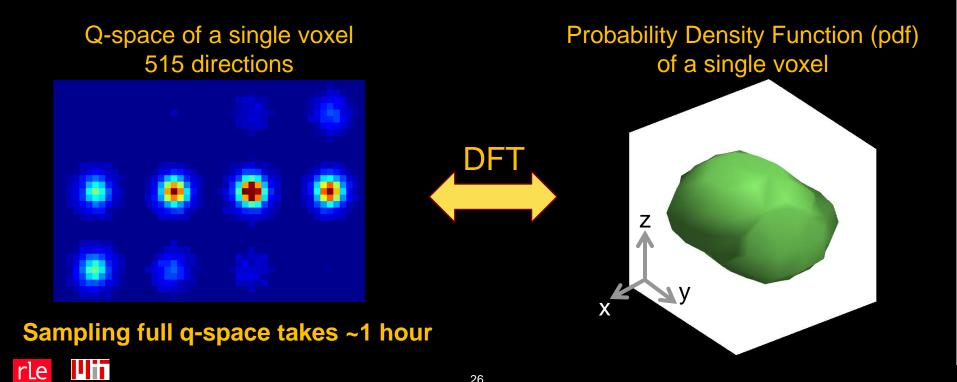
- Unlike tensor modeling, DSI offers a complete description of water diffusion
- And reveals complex distributions of fiber orientations
- DSI requires full sampling of q-space (DTI needs ≥7 points)





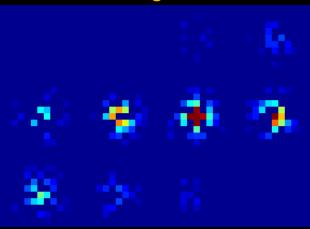
Diffusion Spectrum Imaging (DSI)

- 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)

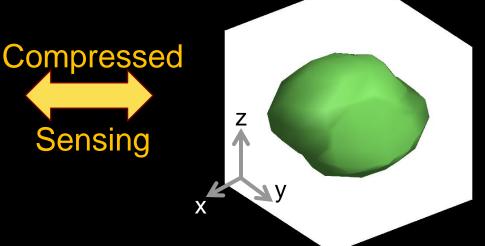


- To reduce scan time, undersample q-space
- Use sparsity prior to recon the pdfs via Compressed Sensing

Undersampled q-space of a single voxel



Probability Density Function (pdf) of a single voxel







Sensing

- To reduce scan time, undersample q-space
- Use sparsity prior to recon the pdfs via Compressed Sensing
 - i. Wavelet + Total Variation [1]

$$\min_{\boldsymbol{p}} \|\mathbf{F}_{\Omega}\boldsymbol{p} - \boldsymbol{q}\|_{2}^{2} + \alpha \cdot \|\mathbf{\Psi}\boldsymbol{p}\|_{1} + \beta \cdot \mathrm{TV}(\boldsymbol{p})$$
 undersampled pdf q-samples wavelet total variation





- To reduce scan time, undersample q-space
- Use sparsity prior to recon the pdfs via Compressed Sensing
 - i. Wavelet + Total Variation [1]

$$min_{\boldsymbol{p}} \|\mathbf{F}_{\Omega}\boldsymbol{p} - \boldsymbol{q}\|_{2}^{2} + \alpha \cdot \|\boldsymbol{\Psi}\boldsymbol{p}\|_{1} + \beta \cdot \mathrm{TV}(\boldsymbol{p})$$

- ii. Dictionary-FOCUSS [2]
 - □ Create a dictionary D from a training dataset of pdfs using K-SVD algorithm [3] → tailored for sparse representation
 - ☐ Impose sparsity constraint via FOCUSS algorithm [4] by solving

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





- To reduce scan time, undersample q-space
- Use sparsity prior to recon the pdfs via Compressed Sensing
 - i. Wavelet + Total Variation [1]

$$min_{\boldsymbol{p}} \|\mathbf{F}_{\Omega}\boldsymbol{p} - \boldsymbol{q}\|_{2}^{2} + \alpha \cdot \|\boldsymbol{\Psi}\boldsymbol{p}\|_{1} + \beta \cdot \mathrm{TV}(\boldsymbol{p})$$

- ii. Dictionary-FOCUSS [2]
 - □ Create a dictionary D from a training dataset of pdfs using K-SVD algorithm [3] → tailored for sparse representation
 - ☐ Impose sparsity constraint via FOCUSS algorithm [4] by solving

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



Dictionary transform coefficients

Iterative DSI Reconstruction

 Dictionary-FOCUSS [1] yields up to 2-times RMSE reduction using compared to Wavelet+TV





Iterative DSI Reconstruction

- Dictionary-FOCUSS [1] yields up to 2-times RMSE reduction using compared to Wavelet+TV
- Compressed Sensing recon is iterative, with processing times
 > 10 sec / voxel for both methods
- Full-brain recon for 10⁵ voxels: ~ 10 DAYS of computation



Iterative DSI Reconstruction

- Dictionary-FOCUSS [1] yields up to 2-times RMSE reduction using compared to Wavelet+TV
- Compressed Sensing recon is iterative, with processing times
 > 10 sec / voxel for both methods
- Full-brain recon for 10⁵ voxels: ~ 10 DAYS of computation
- Two L2-based methods that are 1000-fold faster with image quality similar to Dictionary-FOCUSS are proposed:
 - Tikhonov regularization
 - ii. Principal Component Analysis (PCA)





Tikhonov Regularization

Dictionary-FOCUSS iteratively solves

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

Instead, consider

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



Tikhonov Regularization

Dictionary-FOCUSS iteratively solves

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

Instead, consider

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

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





Tikhonov Regularization

Dictionary-FOCUSS iteratively solves

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

Instead, consider

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

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

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





Tikhonov Regularization

Dictionary-FOCUSS iteratively solves

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

Instead, consider

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

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

$$\mathbf{F}_{\Omega}\mathbf{D} = \mathbf{U}\mathbf{\Sigma}V^{H}$$
 $\mathbf{\Sigma}^{+}\mathbf{U}^{H}q$ $\mathbf{\Sigma}^{+} = (\mathbf{\Sigma}^{H}\mathbf{\Sigma} + \lambda\mathbf{I})^{-1}\mathbf{\Sigma}^{H}$ compute once





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





- 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 P
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - \boldsymbol{p}_{mean}$$
 $\mathbf{Z}\mathbf{Z}^H = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$





- PCA: approximates data points using a linear combo of them to retain the maximum variance in the dataset
- Start with a training set of pdfs P
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - p_{mean}$$
 $\mathbf{Z}\mathbf{Z}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$

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

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

T - dimensional pca coefficients





- 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 P
- Subtract the mean, diagonalize the covariance matrix:

$$\mathbf{Z} = \mathbf{P} - \boldsymbol{p}_{mean}$$
 $\mathbf{Z}\mathbf{Z}^H = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$

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

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

The location of pca in the pdf space,

$$p_T = \mathbf{Q}_T p c a + p_{mean}$$





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

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





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

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

In PCA coordinates,

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





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

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

In PCA coordinates,

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

Closed-form solution:

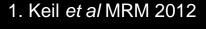
$$\widetilde{pca} = \operatorname{pinv}(\mathbf{F}_{\Omega}\mathbf{Q}_{T})(\mathbf{q} - \mathbf{F}_{\Omega}\mathbf{p}_{mean})$$
compute once





DSI Acquisition

- 2.3 mm isotropic with b_{max} = 8000 s/mm² at 3T
- Connectom gradients and 64-chan head coil [1]
- 515 q-space points collected in 50 min





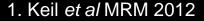


DSI Acquisition

- 2.3 mm isotropic with $b_{max} = 8000 \text{ s/mm}^2$ at 3T
- Connectom gradients and 64-chan head coil [1]
- 515 q-space points collected in 50 min
- Two subjects scanned → dictionary training is based on a subject different from the test subject
- Recon experiments at accelerations R = 3, 5 and 9







DSI Acquisition

- 2.3 mm isotropic with $b_{max} = 8000 \text{ s/mm}^2$ at 3T
- Connectom gradients and 64-chan head coil [1]
- 515 q-space points collected in 50 min
- Two subjects scanned → dictionary training is based on a subject different from the test subject
- Recon experiments at accelerations R = 3, 5 and 9
- Comparison of methods:
 - i. Wavelet + TV [2]
 - ii. Dictionary-FOCUSS [3]

- iii. Tikhonov regularization
- iv. PCA

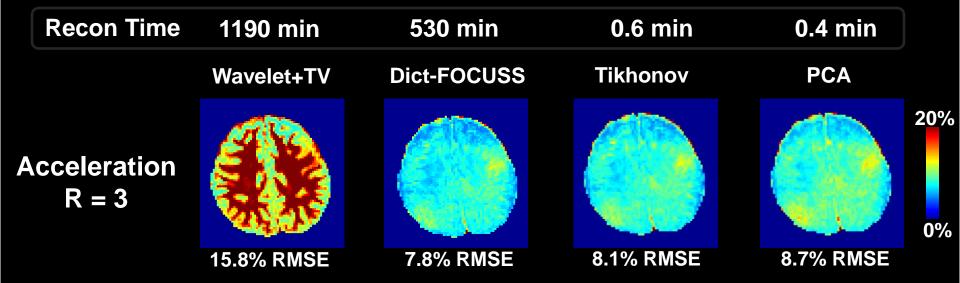
3. Bilgic et al MRM 2012



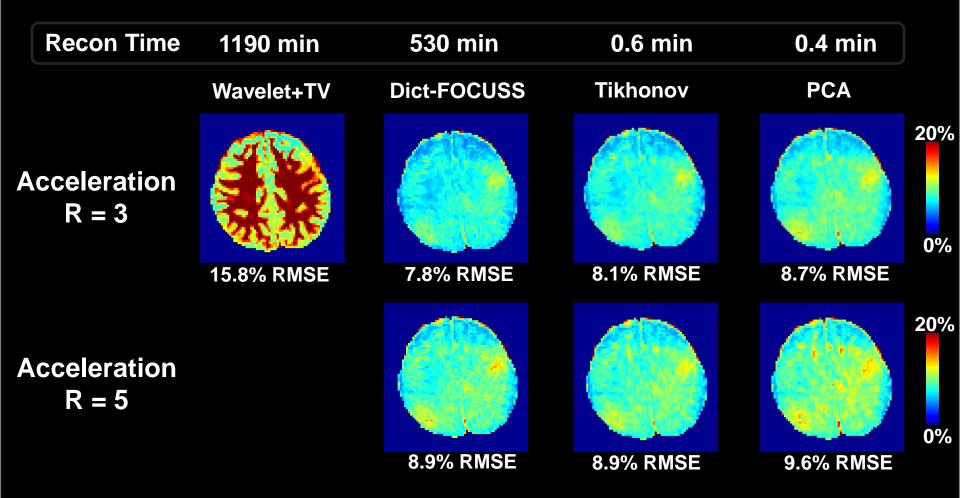
^{1.} Keil et al MRM 2012

^{2.} Menzel et al MRM 2011

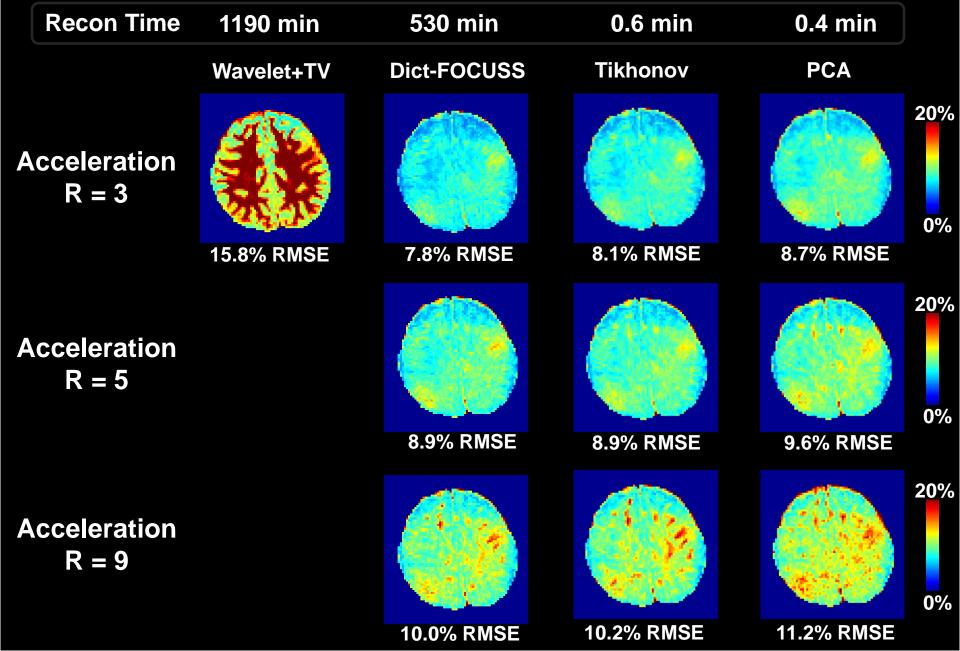
pdf reconstruction error maps



pdf reconstruction error maps



pdf reconstruction error maps



Conclusion

 For suitable applications, using L2-regularization can lead to fast and high-quality reconstructions



Conclusion

- For suitable applications, using L2-regularization can lead to fast and high-quality reconstructions
- Quantitative Susceptibility Mapping
 - □ Closed-form solution:1000-fold speed up obtained relative to state of the art



Conclusion

- For suitable applications, using L2-regularization can lead to fast and high-quality reconstructions
- Quantitative Susceptibility Mapping
 - □ Closed-form solution:1000-fold speed up obtained relative to state of the art
- ii. Diffusion Spectrum Imaging
 - Rather than enforcing sparsity, it seems that using a dictionary is the key to good reconstruction
 - 1000-fold speed up obtained relative to Compressed Sensing

Software Download:

http://web.mit.edu/berkin/www/software.html



