Advances in Quantitative MRI: Acquisition, Estimation, and Applications

Gopal Nataraj

Dissertation Proposal
May 15, 2017

Dept. of Electrical Engineering and Computer Science
University of Michigan

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

- biomarker measurable tissue property (*e.g.*, elasticity) that characterizes a biological process (*e.g.*, sclerosis)

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, elasticity) that characterizes a biological process (*e.g.*, sclerosis)
- localize — produce quantitative MR images

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, elasticity) that characterizes a biological process (*e.g.*, sclerosis)
- localize — produce quantitative MR images
- rapidly — fast acquisition, fast estimation

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, elasticity) that characterizes a biological process (*e.g.*, sclerosis)
- localize — produce quantitative MR images
- rapidly — fast acquisition, fast estimation
- reliably — accurate signal models, precise estimation

Quantitative MRI (QMRI)

Goal: rapidly and reliably localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, elasticity) that characterizes a biological process (*e.g.*, sclerosis)
- localize — produce quantitative MR images
- rapidly — fast acquisition, fast estimation
- reliably — accurate signal models, precise estimation

Challenges (beyond conventional MRI):

- complicated, nonlinear signal models
- more data required, so longer scan times

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?
- **Estimation** [Ch. 5]
  Given data from an informative acquisition,
  how can we rapidly and accurately quantify these biomarkers?

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?
- **Estimation** [Ch. 5]
  Given data from an informative acquisition, how can we rapidly and accurately quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?
- **Estimation** [Ch. 5]
  Given data from an informative acquisition,
  how can we rapidly and accurately quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Signal Model

After reconstruction, single voxel y_d in dth image modeled as

$$y_d = s_d(\mathbf{x}; \boldsymbol{\nu}, \mathbf{p}_d) + \epsilon_d \tag{1}$$

- $\mathbf{x} \in \mathbb{R}^L$ latent free parameters
- $\boldsymbol{\nu} \in \mathbb{R}^K$ known parameters
- $\mathbf{p}_d \in \mathbb{R}^A$ acquisition parameters
- $s_d : \mathbb{R}^{L+K+A} \mapsto \mathbb{C}$ dth signal model
- $\epsilon_d \in \mathbb{C}$ noise $\sim \mathbb{C}\mathcal{N}\left(0, \sigma_d^2\right)$

Signal Model

A *scan profile* contains D voxels $\mathbf{y} := [y_1, \ldots, y_D]^\mathsf{T}$, modeled as

$$\mathbf{y} = \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) + \boldsymbol{\epsilon} \tag{1}$$

- $\mathbf{x} \in \mathbb{R}^L$ latent free parameters
- $\boldsymbol{\nu} \in \mathbb{R}^K$ known parameters
- $\mathbf{P} := [\mathbf{p}_1, \ldots, \mathbf{p}_D]$ acquisition parameter matrix
- $\mathbf{s} : \mathbb{R}^{L+K+AD} \mapsto \mathbb{C}^D$ vector signal model
- $\boldsymbol{\epsilon} \sim \mathbb{C}\mathcal{N}(\mathbf{0}_D, \boldsymbol{\Sigma})$ noise, with $\boldsymbol{\Sigma} := \mathrm{diag}(\sigma_1^2, \ldots, \sigma_D^2)$

Signal Model

A *scan profile* contains D voxels $\mathbf{y} := [y_1, \ldots, y_D]^\mathsf{T}$, modeled as

$$\mathbf{y} = \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) + \boldsymbol{\epsilon} \tag{1}$$

- $\mathbf{x} \in \mathbb{R}^L$ latent free parameters
- $\boldsymbol{\nu} \in \mathbb{R}^K$ known parameters
- $\mathbf{P} := [\mathbf{p}_1, \ldots, \mathbf{p}_D]$ acquisition parameter matrix
- $\mathbf{s} : \mathbb{R}^{L+K+AD} \mapsto \mathbb{C}^D$ vector signal model
- $\boldsymbol{\epsilon} \sim \mathbb{C}\mathcal{N}(\mathbf{0}_D, \boldsymbol{\Sigma})$ noise, with $\boldsymbol{\Sigma} := \mathrm{diag}(\sigma_1^2, \ldots, \sigma_D^2)$

Task: design $\mathbf{P}$ to enable precise unbiased estimation of $\mathbf{x}$

Towards an Objective Function

When $\mathbf{s}$ is analytic in $\mathbf{x}$ (as is typical),

Fisher information characterizes unbiased estimator precision:

$$\mathbf{F}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) := (\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}))^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{2}$$

Towards an Objective Function

When $\mathbf{s}$ is analytic in $\mathbf{x}$ (as is typical),
Fisher information characterizes unbiased estimator precision:

$$\mathbf{F}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) := (\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}))^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{2}$$

When $\mathbf{F}$ is invertible, Cramér-Rao Bound (CRB) [Cramér, 1946] ensures covariance of unbiased estimates $\widehat{\mathbf{x}}$ of $\mathbf{x}$ satisfy

$$\operatorname{cov}(\widehat{\mathbf{x}};\boldsymbol{\nu},\mathbf{P}) \succeq \mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{3}$$

Towards an Objective Function

When $\mathbf{s}$ is analytic in $\mathbf{x}$ (as is typical),
Fisher information characterizes unbiased estimator precision:

$$\mathbf{F}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) := (\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}))^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{2}$$

When $\mathbf{F}$ is invertible, Cramér-Rao Bound (CRB) [Cramér, 1946] ensures covariance of unbiased estimates $\widehat{\mathbf{x}}$ of $\mathbf{x}$ satisfy

$$\operatorname{cov}(\widehat{\mathbf{x}};\boldsymbol{\nu},\mathbf{P}) \succeq \mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{3}$$

Maximum-likelihood (ML) estimates achieve CRB asymptotically or equivalently (for Gaussian data) at sufficiently high SNR.

Towards an Objective Function

When $\mathbf{s}$ is analytic in $\mathbf{x}$ (as is typical),
Fisher information characterizes unbiased estimator precision:

$$\mathbf{F}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) := (\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}))^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{2}$$

When $\mathbf{F}$ is invertible, Cramér-Rao Bound (CRB) [Cramér, 1946]
ensures covariance of unbiased estimates $\widehat{\mathbf{x}}$ of $\mathbf{x}$ satisfy

$$\mathsf{cov}(\widehat{\mathbf{x}};\boldsymbol{\nu},\mathbf{P}) \succeq \mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P}). \tag{3}$$

Maximum-likelihood (ML) estimates achieve CRB asymptotically
or equivalently (for Gaussian data) at sufficiently high SNR.
Idea: choose $\mathbf{P}$ such that imprecision matrix $\mathbf{F}^{-1}$ "small"

Scan Design

Idea: choose $\mathbf{P}$ to minimize the objective

$$\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) = \mathrm{tr}\Big(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P})\mathbf{W}^{\mathsf{T}}\Big), \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a pre-selected diagonal matrix of weights.

Scan Design

Idea: choose $\mathbf{P}$ to minimize the objective

$$\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) = \mathrm{tr}\Big(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P})\mathbf{W}^{\mathsf{T}}\Big), \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a pre-selected diagonal matrix of weights.

Challenge: $\mathbf{x}, \boldsymbol{\nu}$ vary spatially

Scan Design

Idea: choose $\mathbf{P}$ to minimize the objective

$$\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) = \mathrm{tr}\Big(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P})\mathbf{W}^{\mathsf{T}}\Big), \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a pre-selected diagonal matrix of weights.

Challenge: $\mathbf{x}, \boldsymbol{\nu}$ vary spatially

Two problems considered:

- min-max scan design [Nataraj et al., 2017]

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P}\in\mathbb{P}} \max_{\substack{\mathbf{x}\in\mathbb{X}^{\mathrm{t}} \\ \boldsymbol{\nu}\in\mathbb{N}^{\mathrm{t}}}} \Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}), \right\} \tag{5}$$

  where $\mathbb{X}^{\mathrm{t}} \subseteq \mathbb{R}^{L}$ and $\mathbb{N}^{\mathrm{t}} \subseteq \mathbb{R}^{K}$ are "tight" ranges of interest and $\mathbb{P}$ is defined by acquisition/timing constraints

Scan Design

Idea: choose $\mathbf{P}$ to minimize the objective

$$\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) = \mathrm{tr}\Big(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P})\mathbf{W}^{\mathsf{T}}\Big), \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a pre-selected diagonal matrix of weights.

Challenge: $\mathbf{x}, \boldsymbol{\nu}$ vary spatially

Two problems considered:

- min-max scan design [Nataraj et al., 2017]

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P}\in\mathbb{P}} \max_{\substack{\mathbf{x}\in\mathbb{X}^{\mathrm{t}} \\ \boldsymbol{\nu}\in\mathbb{N}^{\mathrm{t}}}} \Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}), \right\} \tag{5}$$

- Bayesian scan design [§6.3]

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P}\in\mathbb{P}} \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}}(\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P})) \right\} \tag{6}$$

Scan Design

Idea: choose $\mathbf{P}$ to minimize the objective

$$\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}) = \mathrm{tr}\Big(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x};\boldsymbol{\nu},\mathbf{P})\mathbf{W}^{\mathsf{T}}\Big), \tag{4}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a pre-selected diagonal matrix of weights.

Challenge: $\mathbf{x}, \boldsymbol{\nu}$ vary spatially

Two problems considered:

- min-max scan design [Nataraj et al., 2017]

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P}\in\mathbb{P}} \max_{\substack{\mathbf{x}\in\mathbb{X}^{\mathrm{t}} \\ \boldsymbol{\nu}\in\mathbb{N}^{\mathrm{t}}}} \Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P}), \right\} \tag{5}$$

- Bayesian scan design [§6.3]

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P}\in\mathbb{P}} \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}}(\Psi(\mathbf{x};\boldsymbol{\nu},\mathbf{P})) \right\} \tag{6}$$

Detailed Example Study

Task: design fast acquisition for precise estimation of relaxation parameters T_1, T_2 in white/gray matter (WM/GM) of brain

Detailed Example Study

Task: design fast acquisition for precise estimation of relaxation parameters T_1, T_2 in white/gray matter (WM/GM) of brain

- Consider scan profiles consisting of two fast pulse sequences
  - Spoiled Gradient-Recalled Echo (SPGR) [Zur et al., 1991]
  - Dual-Echo Steady-State (DESS) [Redpath and Jones, 1988]

Detailed Example Study

Task: design fast acquisition for precise estimation of relaxation parameters T_1, T_2 in white/gray matter (WM/GM) of brain

- Consider scan profiles consisting of two fast pulse sequences
  - Spoiled Gradient-Recalled Echo (SPGR) [Zur et al., 1991]
  - Dual-Echo Steady-State (DESS) [Redpath and Jones, 1988]
- For each scan profile feasible under total time constraint:
  1. Let $\mathbf{s}$ model corresponding single-component signal
     - $\mathbf{x} \leftarrow [m_0, T_1, T_2]^{\mathsf{T}}$, where m_0 is a scale factor
     - $\boldsymbol{\nu} \leftarrow$ flip angle variation
     - $\mathbf{P} \leftarrow$ nominal flip angles, repetition times
  2. Optimize $\mathbf{P}$ subject to flip angle, sequence timing constraints
     - $\mathbf{W} \leftarrow \mathrm{diag}(0, 0.1, 1)$ emphasizes T_1, T_2 est roughly equally
     - $\mathbb{X}^{\mathrm{t}}$ chosen to focus on WM/GM at 3T field strength
     - $\mathbb{N}^{\mathrm{t}}$ chosen to allow 10% flip angle variation

Scan Profile Comparison

(#SPGR, #DESS) Profiles	(2, 1)	(1, 1)	(0, 2)
SPGR nom. flip (deg)	(15, 5)	15	–
DESS nom. flip (deg)	30	10	(35, 10)
SPGR rep. times (ms)	(12.2, 12.2)	13.9	–
DESS rep. times (ms)	17.5	28.0	(24.4, 17.5)
Optimized Cost	4.0	4.9	**3.5**

Scan Profile Comparison

(#SPGR, #DESS) Profiles	(2, 1)	(1, 1)	(0, 2)
SPGR nom. flip (deg)	(15, 5)	15	–
DESS nom. flip (deg)	30	10	(35, 10)
SPGR rep. times (ms)	(12.2, 12.2)	13.9	–
DESS rep. times (ms)	17.5	28.0	(24.4, 17.5)
Optimized Cost	4.0	4.9	**3.5**

Main finding: 2 DESS sequences can yield T_1, T_2 WM/GM estimates that are at least as precise as T_1, T_2 estimates from SPGR/DESS scan profiles, under this competitive time constraint.

Numerical Simulation

- Simulated many WM-like, GM-like voxel realizations
- Studied sample statistics of T_1, T_2 ML estimates $\widehat{T}_1^{\mathrm{ML}}$, $\widehat{T}_2^{\mathrm{ML}}$

Numerical Simulation

- Simulated many WM-like, GM-like voxel realizations
- Studied sample statistics of T_1, T_2 ML estimates $\widehat{T}_1^{\mathrm{ML}}$, $\widehat{T}_2^{\mathrm{ML}}$

Profile	$(2,1)$	$(1,1)$	$(0,2)$	Truth
WM $\widehat{T}_1^{\mathrm{ML}}$	830 ± 17	830 ± 15	830 ± 14	832
GM $\widehat{T}_1^{\mathrm{ML}}$	$1330 \pm 30.$	1330 ± 24	1330 ± 24	1331
WM $\widehat{T}_2^{\mathrm{ML}}$	$80. \pm 1.0$	$80. \pm 2.1$	79.6 ± 0.94	79.6
GM $\widehat{T}_2^{\mathrm{ML}}$	$110. \pm 1.4$	$110. \pm 3.0$	$110. \pm 1.6$	110

Table 1: $\widehat{T}_1^{\mathrm{ML}}$, $\widehat{T}_2^{\mathrm{ML}}$ sample means $\pm$ sample standard deviations

Experimental Setup

Candidate $(2,1)$, $(1,1)$, $(0,2)$ SPGR/DESS scan profiles

- Prescribed optimized nominal flip angles, repetition times
- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **1m37s** scan time for each profile

Experimental Setup

Candidate $(2,1)$, $(1,1)$, $(0,2)$ SPGR/DESS scan profiles

- Prescribed optimized nominal flip angles, repetition times
- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **1m37s** scan time for each profile

Reference scan profile

- Four inversion recovery (IR) scans for T_1 estimation
- Four spin-echo (SE) scans for T_2 estimation
- 256×256 matrix over $24 \times 24 \times 0.5$cm FOV
- Required **40m58s** scan time total

Experimental Setup

Candidate $(2,1)$, $(1,1)$, $(0,2)$ SPGR/DESS scan profiles

- Prescribed optimized nominal flip angles, repetition times
- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **1m37s** scan time for each profile

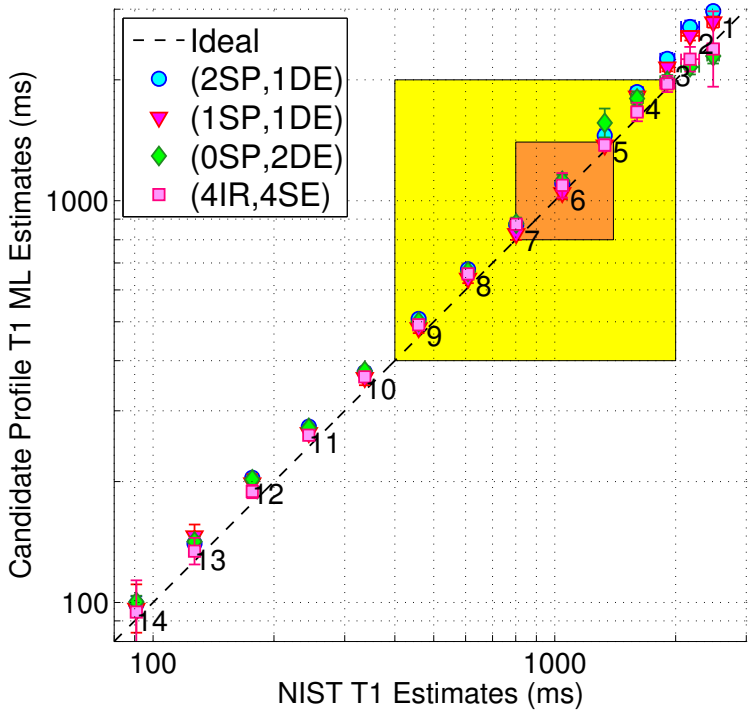
Reference scan profile

- Four inversion recovery (IR) scans for T_1 estimation
- Four spin-echo (SE) scans for T_2 estimation
- 256×256 matrix over $24 \times 24 \times 0.5$cm FOV
- Required **40m58s** scan time total

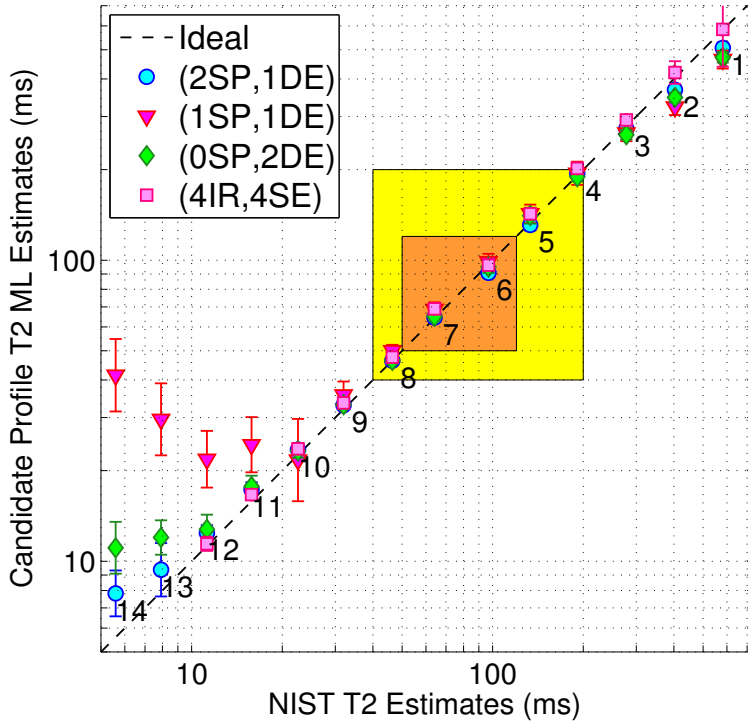
Bloch-Siegert (BS) acquisition for separate flip angle calibration

- Acquired 2 BS-shifted SPGR scans in 1m40s total
- Used for T_1, T_2 est from both candidate and reference profiles

Phantom Accuracy Results



(a) $\widehat{T}_1^{\mathrm{ML}}$ Estimates



(b) $\widehat{T}_2^{\mathrm{ML}}$ Estimates

Compared against NIST NMR measurements [Keenan et al., 2016]

Phantom Precision Results

- Repeated each profile 10 times
- Estimated T_1, T_2 std dev of typical voxel across repetitions

Phantom Precision Results

	(2, 1)	(1, 1)	(0, 2)
V5 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	50 ± 12	$40 \pm 10.$	39 ± 9.4
V6 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	70 ± 18	60 ± 15	70 ± 16
V7 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	60 ± 13	50 ± 13	50 ± 13
V5 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	2.6 ± 0.63	6 ± 1.4	3.5 ± 0.84
V6 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	1.9 ± 0.46	5 ± 1.1	2.3 ± 0.54
V7 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	1.4 ± 0.34	3.4 ± 0.80	1.5 ± 0.35

Table 2: Pooled sample standard deviations $\pm$ pooled standard errors of sample standard deviations (ms), from optimized SPGR/DESS profiles.

Phantom Precision Results

	(2, 1)	(1, 1)	(0, 2)
V5 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	50 ± 12	$40 \pm 10.$	39 ± 9.4
V6 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	70 ± 18	60 ± 15	70 ± 16
V7 $\widehat{\sigma}_{\widehat{T}_1^{\mathrm{ML}}}$	60 ± 13	50 ± 13	50 ± 13
V5 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	2.6 ± 0.63	6 ± 1.4	3.5 ± 0.84
V6 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	1.9 ± 0.46	5 ± 1.1	2.3 ± 0.54
V7 $\widehat{\sigma}_{\widehat{T}_2^{\mathrm{ML}}}$	1.4 ± 0.34	3.4 ± 0.80	1.5 ± 0.35

Table 2: Pooled sample standard deviations $\pm$ pooled standard errors of sample standard deviations (ms), from optimized SPGR/DESS profiles.

Similar trends across profiles of empirical vs. theoretical std dev!

In vivo Results

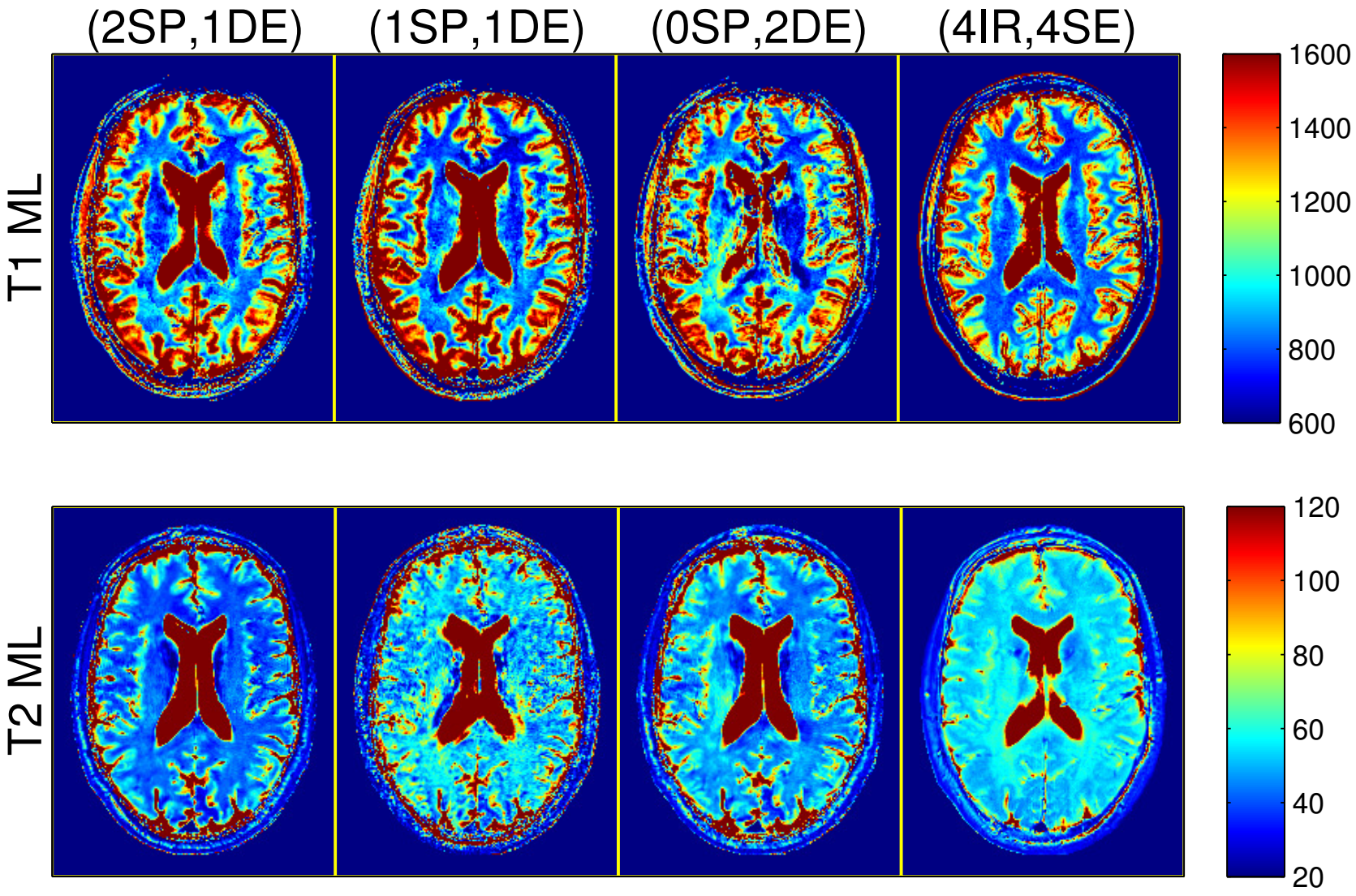


Figure 1: Colorbar ranges in ms.

Summary

Contributions

- MR scan design method for precise parameter estimation
- Fast SPGR/DESS scan profile for T_1, T_2 estimation in brain

Summary

Contributions

- MR scan design method for precise parameter estimation
- Fast SPGR/DESS scan profile for T_1, T_2 estimation in brain
    - Simulation and phantom results validate method as a predictor of unbiased estimation precision.

Summary

Contributions

- MR scan design method for precise parameter estimation
- Fast SPGR/DESS scan profile for T_1, T_2 estimation in brain
  - Simulation and phantom results validate method as a predictor of unbiased estimation precision.
  - *In vivo* results reveal discrepancies (especially in T_2 estimates), suggesting sensitivity to model mismatch.

Summary

Contributions

- MR scan design method for precise parameter estimation
- Fast SPGR/DESS scan profile for T_1, T_2 estimation in brain
  - Simulation and phantom results validate method as a predictor of unbiased estimation precision.
  - *In vivo* results reveal discrepancies (especially in T_2 estimates), suggesting sensitivity to model mismatch.

How to address model mismatch?

- More complete *in vivo* signal models
- More scalable parameter estimation

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?
- **Estimation** [Ch. 5]
  Given data from an informative acquisition,
  how can we rapidly and accurately quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Signal Model

Given: at each voxel, image sequence $\mathbf{y} \in \mathbb{C}^D$ modeled as

$$\mathbf{y} = \mathbf{s}(\mathbf{x}, \boldsymbol{\nu}) + \boldsymbol{\epsilon} \tag{7}$$

- $\mathbf{x} \in \mathbb{R}^L$ latent free parameters
- $\boldsymbol{\nu} \in \mathbb{R}^K$ known parameters
- $\mathbf{s} : \mathbb{R}^{L+K} \mapsto \mathbb{C}^D$ vector signal model
- $\boldsymbol{\epsilon} \in \mathbb{C}^D$ noise $\sim \mathbb{C}\mathcal{N}(\mathbf{0}_D, \boldsymbol{\Sigma})$

Signal Model

Given: at each voxel, image sequence $\mathbf{y} \in \mathbb{C}^D$ modeled as

$$\mathbf{y} = \mathbf{s}(\mathbf{x}, \boldsymbol{\nu}) + \boldsymbol{\epsilon} \tag{7}$$

- $\mathbf{x} \in \mathbb{R}^L$ latent free parameters
- $\boldsymbol{\nu} \in \mathbb{R}^K$ known parameters
- $\mathbf{s} : \mathbb{R}^{L+K} \mapsto \mathbb{C}^D$ vector signal model
- $\boldsymbol{\epsilon} \in \mathbb{C}^D$ noise $\sim \mathbb{C}\mathcal{N}(\mathbf{0}_D, \boldsymbol{\Sigma})$

Task: construct fast estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Prior Work

Task: construct fast estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Challenges:

- signal $\mathbf{s}$ often nonlinear in $\mathbf{x}$: non-convex inverse problems
- signal $\mathbf{s}$ might be difficult to write in closed form

Prior Work

Task: construct fast estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Challenges:

- signal $\mathbf{s}$ often nonlinear in $\mathbf{x}$: non-convex inverse problems
- signal $\mathbf{s}$ might be difficult to write in closed form

Conventional Approaches:

- gradient-based local optimization
  - initialization-dependent solution
  - requires signal gradients

Prior Work

Task: construct fast estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Challenges:

- signal $\mathbf{s}$ often nonlinear in $\mathbf{x}$: non-convex inverse problems
- signal $\mathbf{s}$ might be difficult to write in closed form

Conventional Approaches:

- gradient-based local optimization
  - initialization-dependent solution
  - requires signal gradients
- stochastic methods (*e.g.*, simulated annealing)
  - unclear convergence analysis [Bertsimas and Tsitsiklis, 1993]
  - several unintuitive tuning parameters

Prior Work

Task: construct fast estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Challenges:

- signal $\mathbf{s}$ often nonlinear in $\mathbf{x}$: non-convex inverse problems
- signal $\mathbf{s}$ might be difficult to write in closed form

Conventional Approaches:

- gradient-based local optimization
  - initialization-dependent solution
  - requires signal gradients
- stochastic methods (*e.g.*, simulated annealing)
  - unclear convergence analysis [Bertsimas and Tsitsiklis, 1993]
  - several unintuitive tuning parameters
- grid search *e.g.*, for 1-compartment relaxivity [Ch. 4]

Motivation

Grid search computational costs

	L	$\sim$number dictionary atoms
1-compartment relaxivity	3	$\sim 100^2$

Motivation

Grid search computational costs

	L	$\sim$number dictionary atoms
1-compartment relaxivity	3	$\sim 100^2$
flow velocity	4	$\sim 100^3$
diffusivity tensor	7	$\sim 100^6$
2-3 compartment relaxivity	6-10	$\sim 100^5 - 100^9$

Motivation

Grid search computational costs

	L	$\sim$number dictionary atoms
1-compartment relaxivity	3	$\sim 100^2$
flow velocity	4	$\sim 100^3$
diffusivity tensor	7	$\sim 100^6$
2-3 compartment relaxivity	6-10	$\sim 100^5 - 100^9$

Can we scale computation with L more gracefully?

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

- sample $(\mathbf{x}_1, \boldsymbol{\nu}_1, \boldsymbol{\epsilon}_1), \ldots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \boldsymbol{\epsilon}_N)$ from prior distributions
- simulate image data vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ via signal model $\mathbf{s}$

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

- sample $(\mathbf{x}_1, \boldsymbol{\nu}_1, \boldsymbol{\epsilon}_1), \ldots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \boldsymbol{\epsilon}_N)$ from prior distributions
- simulate image data vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ via signal model $\mathbf{s}$
- design *nonlinear* functions $\widehat{x}_l(\cdot) := \widehat{h}_l(\cdot) + \widehat{b}_l$ for $l \in \{1, \ldots, L\}$ that map each $\mathbf{q}_n := [\mathrm{Re}(\mathbf{y}_n)^\mathsf{T}, \mathrm{Im}(\mathbf{y}_n)^\mathsf{T}, \boldsymbol{\nu}_n^\mathsf{T}]^\mathsf{T} \in \mathcal{Q}$ to $x_{l,n} \in \mathbb{R}$

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

- sample $(\mathbf{x}_1, \nu_1, \epsilon_1), \ldots, (\mathbf{x}_N, \nu_N, \epsilon_N)$ from prior distributions
- simulate image data vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ via signal model $\mathbf{s}$
- design *nonlinear* functions $\widehat{x}_l(\cdot) := \widehat{h}_l(\cdot) + \widehat{b}_l$ for $l \in \{1, \ldots, L\}$ that map each $\mathbf{q}_n := [\mathrm{Re}(\mathbf{y}_n)^\mathsf{T}, \mathrm{Im}(\mathbf{y}_n)^\mathsf{T}, \boldsymbol{\nu}_n^\mathsf{T}]^\mathsf{T} \in \mathcal{Q}$ to $x_{l,n} \in \mathbb{R}$

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} \left(h_l(\mathbf{q}_n) + b_l - x_{l,n}\right)^2 \right\}$$

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

- sample $(\mathbf{x}_1, \nu_1, \epsilon_1), \ldots, (\mathbf{x}_N, \nu_N, \epsilon_N)$ from prior distributions
- simulate image data vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ via signal model $\mathbf{s}$
- design *nonlinear* functions $\widehat{x}_l(\cdot) := \widehat{h}_l(\cdot) + \widehat{b}_l$ for $l \in \{1, \ldots, L\}$ that map each $\mathbf{q}_n := [\mathrm{Re}(\mathbf{y}_n)^\mathsf{T}, \mathrm{Im}(\mathbf{y}_n)^\mathsf{T}, \boldsymbol{\nu}_n^\mathsf{T}]^\mathsf{T} \in \mathcal{Q}$ to $x_{l,n} \in \mathbb{R}$

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} \left(h_l(\mathbf{q}_n) + b_l - x_{l,n}\right)^2 \right\} \quad \text{ill-posed!}$$

Machine Learning for QMRI Parameter Estimation

Idea: learn a *nonlinear* estimator from simulated training data

- sample $(\mathbf{x}_1, \nu_1, \epsilon_1), \ldots, (\mathbf{x}_N, \nu_N, \epsilon_N)$ from prior distributions
- simulate image data vectors $\mathbf{y}_1, \ldots, \mathbf{y}_N$ via signal model $\mathbf{s}$
- design *nonlinear* functions $\widehat{x}_l(\cdot) := \widehat{h}_l(\cdot) + \widehat{b}_l$ for $l \in \{1, \ldots, L\}$ that map each $\mathbf{q}_n := [\mathrm{Re}(\mathbf{y}_n)^\mathsf{T}, \mathrm{Im}(\mathbf{y}_n)^\mathsf{T}, \boldsymbol{\nu}_n^\mathsf{T}]^\mathsf{T} \in \mathcal{Q}$ to $x_{l,n} \in \mathbb{R}$

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \in \mathbb{H} \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2 \right\} \quad (8)$$

Solution: solve a *kernel ridge regression* (KRR) problem

- restrict function space over which we optimize
- include function regularization

A Function Space over which Optimization is Tractable

Hilbert space: complete inner product function space

A Function Space over which Optimization is Tractable

Hilbert space: complete inner product function space

Reproducing kernel Hilbert space (RKHS)

Hilbert space $\mathbb{H}$ over input space $\mathcal{Q}$ with *reproducing property*

$$\langle h, k(\cdot, \mathbf{q}) \rangle_{\mathbb{H}} = h(\mathbf{q}), \qquad \forall h \in \mathbb{H}, \mathbf{q} \in \mathcal{Q}$$

for some $k : \mathcal{Q}^2 \mapsto \mathbb{R}$ called a reproducing kernel (RK)

A Function Space over which Optimization is Tractable

Hilbert space: complete inner product function space

Reproducing kernel Hilbert space (RKHS)

Hilbert space $\mathbb{H}$ over input space $\mathcal{Q}$ with *reproducing property*

$$\langle h, k(\cdot, \mathbf{q}) \rangle_{\mathbb{H}} = h(\mathbf{q}), \qquad \forall h \in \mathbb{H}, \mathbf{q} \in \mathcal{Q}$$

for some $k : \mathcal{Q}^2 \mapsto \mathbb{R}$ called a reproducing kernel (RK)

Relevant facts

- Bijection between RKHS $\mathbb{H}$ and RK k [Aronszajn, 1950]
- Function $k(\cdot, \mathbf{q}) \in \mathbb{H}$ called a *feature mapping*

Function Optimization over a RKHS

Choose: RK $k : \mathcal{Q}^2 \mapsto \mathbb{R}$, which induces choice of RKHS $\mathbb{H}$

Function Optimization over a RKHS

Choose: RK $k : \mathcal{Q}^2 \mapsto \mathbb{R}$, which induces choice of RKHS $\mathbb{H}$

- *Nonlinear* kernel corresponds to *nonlinear* estimation
- We use $k(\mathbf{q}, \mathbf{q}') \leftarrow \exp\left(-\frac{1}{2}\left\|\mathbf{\Lambda}^{-1}(\mathbf{q} - \mathbf{q}')\right\|_2^2\right)$

Function Optimization over a RKHS

Choose: RK $k : \mathcal{Q}^2 \mapsto \mathbb{R}$, which induces choice of RKHS $\mathbb{H}$

Solve: for each desired latent parameter $l \in \{1, \ldots, L\}$,

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \in \mathbb{H} \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2 \right\} \quad (9)$$

Function Optimization over a RKHS

Choose: RK $k : \mathcal{Q}^2 \mapsto \mathbb{R}$, which induces choice of RKHS $\mathbb{H}$

Solve: for each desired latent parameter $l \in \{1, \ldots, L\}$,

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \in \mathbb{H} \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2 \right\} \quad (9)$$

- Optimal $\widehat{h}_l$ over $\mathbb{H}$ takes form [Schölkopf et al., 2001]

$$\widehat{h}_l(\cdot) \equiv \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n) \quad (10)$$

Function Optimization over a RKHS

Choose: RK $k : \mathcal{Q}^2 \mapsto \mathbb{R}$, which induces choice of RKHS $\mathbb{H}$

Solve: for each desired latent parameter $l \in \{1, \ldots, L\}$,

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \in \mathbb{H} \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2 \right\} \quad (9)$$

- Optimal $\widehat{h}_l$ over $\mathbb{H}$ takes form [Schölkopf et al., 2001]

$$\widehat{h}_l(\cdot) \equiv \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n) \quad (10)$$

- Plug (10) into (9); solve now instead for $\left(\widehat{a}_l, \widehat{b}_l\right)$; construct:

$$\widehat{x}_l(\cdot) = \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n) + \widehat{b}_l \quad (11)$$

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ training pt regressands

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ — training pt regressands
- $\mathbf{K} := \begin{bmatrix} k(\mathbf{q}_1, \mathbf{q}_1) & \cdots & k(\mathbf{q}_1, \mathbf{q}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{q}_N, \mathbf{q}_1) & \cdots & k(\mathbf{q}_N, \mathbf{q}_N) \end{bmatrix}$ — Gram matrix

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ — training pt regressands
- $\mathbf{K} := \begin{bmatrix} k(\mathbf{q}_1, \mathbf{q}_1) & \cdots & k(\mathbf{q}_1, \mathbf{q}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{q}_N, \mathbf{q}_1) & \cdots & k(\mathbf{q}_N, \mathbf{q}_N) \end{bmatrix}$ — Gram matrix
- $\mathbf{M} := \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^{\mathsf{T}}$ — de-meaning operator

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M} (\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ — training pt regressands
- $\mathbf{K} := \begin{bmatrix} k(\mathbf{q}_1, \mathbf{q}_1) & \cdots & k(\mathbf{q}_1, \mathbf{q}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{q}_N, \mathbf{q}_1) & \cdots & k(\mathbf{q}_N, \mathbf{q}_N) \end{bmatrix}$ — Gram matrix
- $\mathbf{M} := \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^{\mathsf{T}}$ — de-meaning operator
- $\mathbf{k}(\cdot) := [k(\cdot, \mathbf{q}_1), \ldots, k(\cdot, \mathbf{q}_N)]^{\mathsf{T}}$ — nonlin kernel embedding

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M} (\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ — training pt regressands
- $\mathbf{K} := \begin{bmatrix} k(\mathbf{q}_1, \mathbf{q}_1) & \cdots & k(\mathbf{q}_1, \mathbf{q}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{q}_N, \mathbf{q}_1) & \cdots & k(\mathbf{q}_N, \mathbf{q}_N) \end{bmatrix}$ — Gram matrix
- $\mathbf{M} := \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^{\mathsf{T}}$ — de-meaning operator
- $\mathbf{k}(\cdot) := [k(\cdot, \mathbf{q}_1), \ldots, k(\cdot, \mathbf{q}_N)]^{\mathsf{T}}$ — nonlin kernel embedding

Can we scale computation with L more gracefully?

- Yes, in fact (12) separable in $l \in \{1, \ldots, L\}$ by construction

MRI Parameter Estimation via KRR

Non-iterative closed-form solution, for $l \in \{1, \ldots, L\}$:

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (12)$$

- $\mathbf{x}_l := [x_{l,1}, \ldots, x_{l,N}]^{\mathsf{T}}$ — training pt regressands
- $\mathbf{K} := \begin{bmatrix} k(\mathbf{q}_1, \mathbf{q}_1) & \cdots & k(\mathbf{q}_1, \mathbf{q}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{q}_N, \mathbf{q}_1) & \cdots & k(\mathbf{q}_N, \mathbf{q}_N) \end{bmatrix}$ — Gram matrix
- $\mathbf{M} := \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^{\mathsf{T}}$ — de-meaning operator
- $\mathbf{k}(\cdot) := [k(\cdot, \mathbf{q}_1), \ldots, k(\cdot, \mathbf{q}_N)]^{\mathsf{T}}$ — nonlin kernel embedding

Can we scale computation with L more gracefully?

- Yes, in fact (12) separable in $l \in \{1, \ldots, L\}$ by construction
- However, explicitly computing $\mathbf{K}$ may be undesirable...

KRR as High-Dimensional Affine Regression

Suppose there exists "approximate feature mapping" $\tilde{\mathbf{z}} : \mathcal{Q} \mapsto \mathbb{R}^Z$ such that $\tilde{\mathbf{Z}} := [\tilde{\mathbf{z}}(\mathbf{q}_1), \ldots, \tilde{\mathbf{z}}(\mathbf{q}_N)]$ has for $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{13}$$

KRR as High-Dimensional Affine Regression

Suppose there exists "approximate feature mapping" $\tilde{\mathbf{z}} : \mathcal{Q} \mapsto \mathbb{R}^Z$ such that $\tilde{\mathbf{Z}} := [\tilde{\mathbf{z}}(\mathbf{q}_1), \ldots, \tilde{\mathbf{z}}(\mathbf{q}_N)]$ has for $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{13}$$

Plugging (13) into KRR solution (12) and rearranging gives

$$\widehat{x}_l(\cdot) \approx \frac{1}{N} \mathbf{x}_l^\mathsf{T} \mathbf{1}_N + \frac{1}{N} \mathbf{x}_l^\mathsf{T} \mathbf{M} \tilde{\mathbf{Z}}^\mathsf{T} \left(\frac{1}{N} \tilde{\mathbf{Z}} \mathbf{M} \tilde{\mathbf{Z}}^\mathsf{T} + \rho_l \mathbf{I}_Z \right)^{-1} \left(\tilde{\mathbf{z}}(\cdot) - \frac{1}{N} \tilde{\mathbf{Z}} \mathbf{1}_N \right)$$

KRR as High-Dimensional Affine Regression

Suppose there exists "approximate feature mapping" $\tilde{\mathbf{z}} : \mathcal{Q} \mapsto \mathbb{R}^Z$ such that $\tilde{\mathbf{Z}} := [\tilde{\mathbf{z}}(\mathbf{q}_1), \ldots, \tilde{\mathbf{z}}(\mathbf{q}_N)]$ has for $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{13}$$

Plugging (13) into KRR solution (12) and rearranging gives

$$\widehat{x}_l(\cdot) \approx \widehat{m}_{x_l} + \widehat{\mathbf{c}}_{x_l\tilde{\mathbf{z}}}^\mathsf{T} \left(\widehat{\mathbf{C}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho_l \mathbf{I}_Z \right)^{-1} (\tilde{\mathbf{z}}(\cdot) - \widehat{\mathbf{m}}_{\tilde{\mathbf{z}}}) \tag{14}$$

which is regularized ("ridge") Z-dimensional affine regression!

KRR as High-Dimensional Affine Regression

Suppose there exists "approximate feature mapping" $\tilde{\mathbf{z}} : \mathcal{Q} \mapsto \mathbb{R}^Z$ such that $\tilde{\mathbf{Z}} := [\tilde{\mathbf{z}}(\mathbf{q}_1), \ldots, \tilde{\mathbf{z}}(\mathbf{q}_N)]$ has for $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{13}$$

Plugging (13) into KRR solution (12) and rearranging gives

$$\widehat{x}_l(\cdot) \approx \widehat{m}_{x_l} + \widehat{\mathbf{c}}_{x_l\tilde{\mathbf{z}}}^\mathsf{T} \left(\widehat{\mathbf{C}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho_l \mathbf{I}_Z \right)^{-1} (\tilde{\mathbf{z}}(\cdot) - \widehat{\mathbf{m}}_{\tilde{\mathbf{z}}}) \tag{14}$$

which is regularized ("ridge") Z-dimensional affine regression!

Does such a $\tilde{\mathbf{z}}$ exist and work well in practice?

- Yes, *e.g.* for "shift invariant" kernels (like our Gaussian) of form $k(\mathbf{q}, \mathbf{q}') \equiv k(\mathbf{q} - \mathbf{q}')$ [Rahimi and Recht, 2007]
- In such cases, can reduce from $\sim N^2$ to $\sim NZ$ computations

Model Selection

Online model selection: train *after* observing (unlabeled) test data

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known $\boldsymbol{\nu}$ — density estimation
- Noise covariance $\boldsymbol{\Sigma}$ — low-signal data regions

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known ν
- Noise covariance Σ
- Regularization parameters

density estimation

low-signal data regions

Bayesian perspective...

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known ν — density estimation
- Noise covariance Σ — low-signal data regions
- Regularization parameters — Bayesian perspective...

  - Assume prior $h_l(\cdot) + b_l \sim \mathcal{GP}(0(\cdot), k(\cdot,\cdot)), l \in \{1, \ldots, L\}$

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known $\boldsymbol{\nu}$ — density estimation
- Noise covariance $\boldsymbol{\Sigma}$ — low-signal data regions
- Regularization parameters — Bayesian perspective...

  - Assume prior $h_l(\cdot) + b_l \sim \mathcal{GP}(0(\cdot), k(\cdot,\cdot)), l \in \{1, \ldots, L\}$
  - Include latent parameter variability ϵ_{x_l}
    in observed regressand model $x_l(\mathbf{q}) = h_l(\mathbf{q}) + b_l + \epsilon_{x_l}$

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known $\boldsymbol{\nu}$ — density estimation
- Noise covariance $\boldsymbol{\Sigma}$ — low-signal data regions
- Regularization parameters — Bayesian perspective...

  - Assume prior $h_l(\cdot) + b_l \sim \mathcal{GP}(0(\cdot), k(\cdot,\cdot)), l \in \{1, \ldots, L\}$
  - Include latent parameter variability ϵ_{x_l} in observed regressand model $x_l(\mathbf{q}) = h_l(\mathbf{q}) + b_l + \epsilon_{x_l}$
  - If we further assume $\epsilon_{x_l} \sim \mathcal{N}(0, N\rho_l)$ then posterior mean function is KRR solution (12) [Rasmussen and Williams, 2005]
  - Reasonable to set $N\rho_l \leftarrow \mathrm{cov}(x_l) \approx \frac{1}{N}\mathbf{x}_l^{\mathsf{T}}\mathbf{M}\mathbf{x}_l$

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known $\boldsymbol{\nu}$ — density estimation
- Noise covariance $\boldsymbol{\Sigma}$ — low-signal data regions
- Regularization parameters — $\rho_l \leftarrow \frac{1}{N^2}\mathbf{x}_l^\mathsf{T}\mathbf{M}\mathbf{x}_l$
  - Assume prior $h_l(\cdot) + b_l \sim \mathcal{GP}(0(\cdot), k(\cdot,\cdot)), l \in \{1,\ldots,L\}$
  - Include latent parameter variability ϵ_{x_l} in observed regressand model $x_l(\mathbf{q}) = h_l(\mathbf{q}) + b_l + \epsilon_{x_l}$
  - If we further assume $\epsilon_{x_l} \sim \mathcal{N}(0, N\rho_l)$ then posterior mean function is KRR solution (12) [Rasmussen and Williams, 2005]
  - Reasonable to set $N\rho_l \leftarrow \mathrm{cov}(x_l) \approx \frac{1}{N}\mathbf{x}_l^\mathsf{T}\mathbf{M}\mathbf{x}_l$
- Kernel smoothing length-scale — $\boldsymbol{\Lambda} \leftarrow \mathrm{diag}\left(\sum_{n=1}^{N}\mathbf{q}_n\right)$

Model Selection

Online model selection: train *after* observing (unlabeled) test data

- Prior on known $\boldsymbol{\nu}$ — density estimation
- Noise covariance $\boldsymbol{\Sigma}$ — low-signal data regions
- Regularization parameters — $\rho_l \leftarrow \frac{1}{N^2}\mathbf{x}_l^\mathsf{T}\mathbf{M}\mathbf{x}_l$
  - Assume prior $h_l(\cdot) + b_l \sim \mathcal{GP}(0(\cdot), k(\cdot,\cdot)), l \in \{1,\ldots,L\}$
  - Include latent parameter variability ϵ_{x_l} in observed regressand model $x_l(\mathbf{q}) = h_l(\mathbf{q}) + b_l + \epsilon_{x_l}$
  - If we further assume $\epsilon_{x_l} \sim \mathcal{N}(0, N\rho_l)$ then posterior mean function is KRR solution (12) [Rasmussen and Williams, 2005]
  - Reasonable to set $N\rho_l \leftarrow \mathrm{cov}(x_l) \approx \frac{1}{N}\mathbf{x}_l^\mathsf{T}\mathbf{M}\mathbf{x}_l$
- Kernel smoothing length-scale — $\boldsymbol{\Lambda} \leftarrow \mathrm{diag}\left(\sum_{n=1}^{N}\mathbf{q}_n\right)$

Some parameters still require manual selection...

- Prior on $\mathbf{x}$ — from tissue properties
- Kernel shape — $k(\mathbf{q},\mathbf{q}') \leftarrow \exp\left(-\frac{1}{2}\left\|\boldsymbol{\Lambda}^{-1}(\mathbf{q}-\mathbf{q}')\right\|_2^2\right)$

Summary

Contribution

- Fast KRR method for nonlinear MRI parameter estimation

Summary

Contribution

- Fast KRR method for nonlinear MRI parameter estimation
  - Key insight: even with complicated MR signal models, can simulate training points "for free"
  - Convert *nonlinear estimation* problem into *nonlinear regression* problem that we solve in closed-form with kernels

Summary

Contribution

- Fast KRR method for nonlinear MRI parameter estimation
  - Key insight: even with complicated MR signal models, can simulate training points "for free"
  - Convert *nonlinear estimation* problem into *nonlinear regression* problem that we solve in closed-form with kernels

Ongoing work

- Performance analysis: how should N scale with L, Q?

Summary

Contribution

- Fast KRR method for nonlinear MRI parameter estimation
  - Key insight: even with complicated MR signal models, can simulate training points "for free"
  - Convert *nonlinear estimation* problem into *nonlinear regression* problem that we solve in closed-form with kernels

Ongoing work

- Performance analysis: how should N scale with L, Q?
- Exploit partially linear structure to incorporate scale invariance

Summary

Contribution

- Fast KRR method for nonlinear MRI parameter estimation
  - Key insight: even with complicated MR signal models, can simulate training points "for free"
  - Convert *nonlinear estimation* problem into *nonlinear regression* problem that we solve in closed-form with kernels

Ongoing work

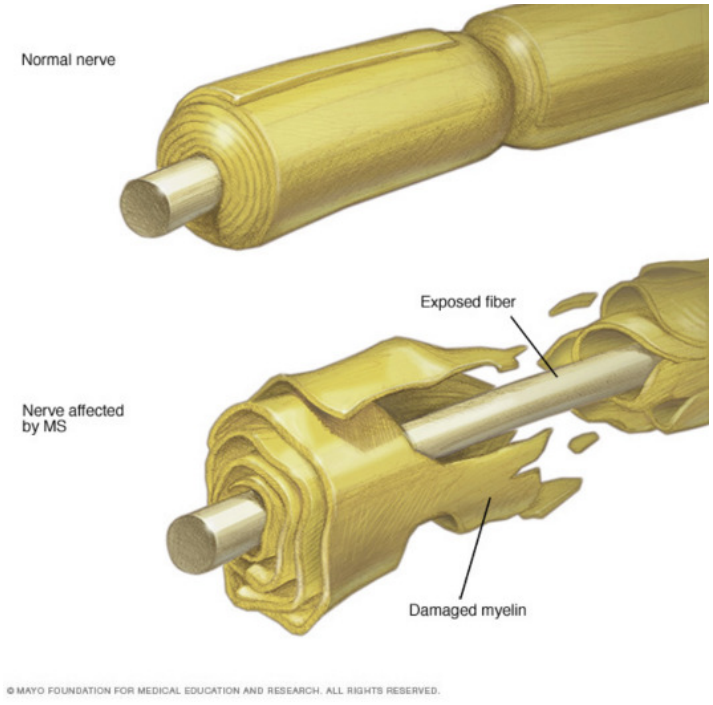
- Performance analysis: how should N scale with L, Q?
- Exploit partially linear structure to incorporate scale invariance
- Validation on a compelling problem...

Overview

Advances in Quantitative MRI:

- **Acquisition** [Ch. 4]
  How can we assemble fast, informative collections of scans to enable precise biomarker quantification?
- **Estimation** [Ch. 5]
  Given data from an informative acquisition, how can we rapidly and accurately quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Background

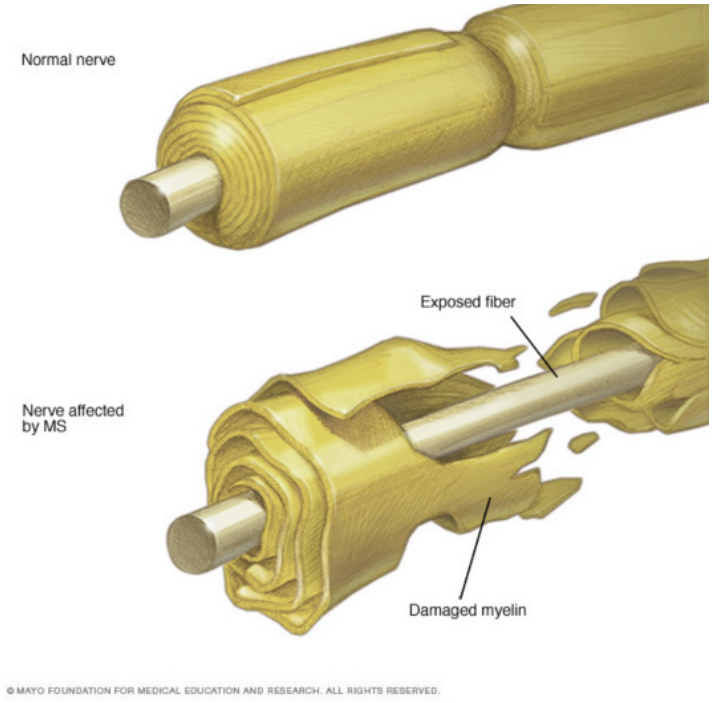


www.mayoclinic.org

Background

Myelin water fraction (MWF):

- Proportion of MR signal arising from water trapped within myelin bilayers, relative to total signal [Mackay et al., 1994]

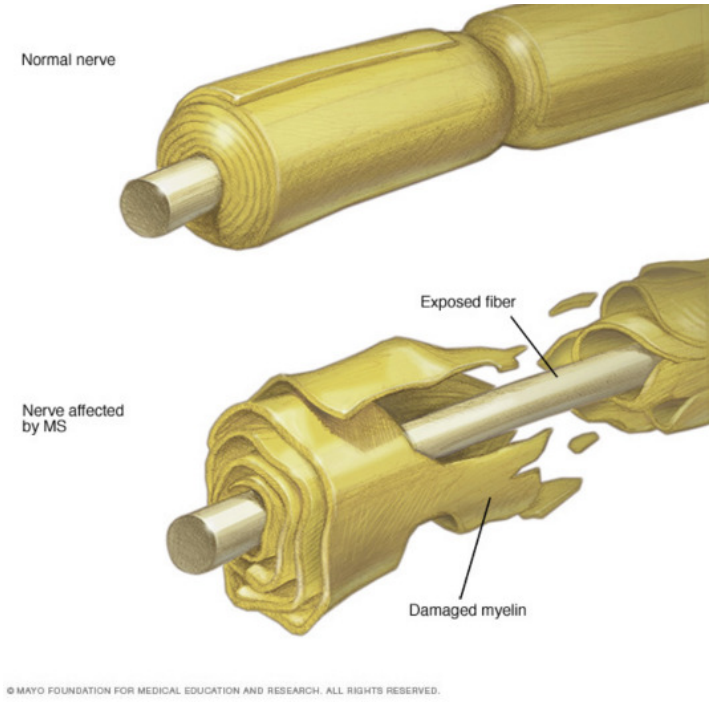


www.mayoclinic.org

Background

Myelin water fraction (MWF):

- Proportion of MR signal arising from water trapped within myelin bilayers, relative to total signal [Mackay et al., 1994]
- Correlates well with myelin content [Webb et al., 2003]



www.mayoclinic.org

Previous MWF imaging acquisitions

Multi-echo spin-echo (**MESE**) [Mackay et al., 1994]

- Gold-standard
- Speed-limited by long repetition times ($\sim$2s)

Previous MWF imaging acquisitions

Multi-echo spin-echo (**MESE**) [Mackay et al., 1994]

- Gold-standard
- Speed-limited by long repetition times (∼2s)

Combinations of fast steady-state scans using variable flip angles ("**mcDESPOT**") [Deoni et al., 2008]

- Whole-brain, high-resolution MWF imaging in ∼30m
- Disagree with MESE MWF estimates, [Zhang et al., 2015] likely due to insufficient precision [Lankford and Does, 2013]

Previous MWF imaging acquisitions

Multi-echo spin-echo (**MESE**) [Mackay et al., 1994]

- Gold-standard
- Speed-limited by long repetition times (∼2s)

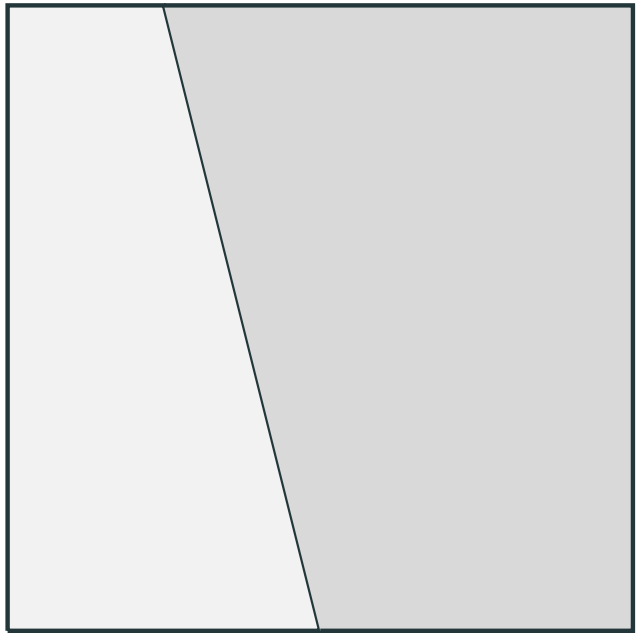
Combinations of fast steady-state scans using variable flip angles ("**mcDESPOT**") [Deoni et al., 2008]

- Whole-brain, high-resolution MWF imaging in ∼30m
- Disagree with MESE MWF estimates, [Zhang et al., 2015]
  likely due to insufficient precision [Lankford and Does, 2013]

Goal: fast, precise MWF quantification in WM

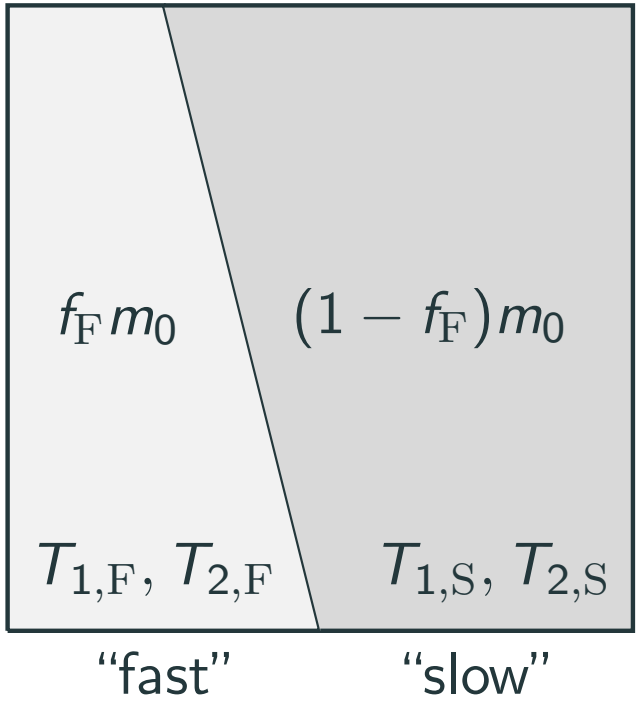
A voxel-scale MWF model

simple two-compartment model



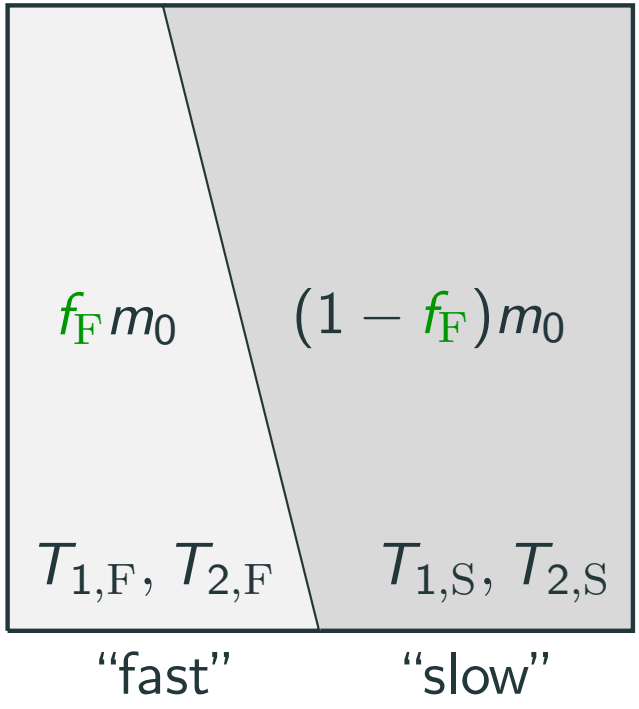
A voxel-scale MWF model

simple two-compartment model



A voxel-scale MWF model

simple two-compartment model



Take fast-relaxing fraction f_F as a simple proxy for MWF

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange
- neglected relaxation, precession, exchange during excitation

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange
- neglected relaxation, precession, exchange during excitation
- absorbing off-resonance effects into m_0 *implies* neglecting exchange between excitation and readout

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange
- neglected relaxation, precession, exchange during excitation
- absorbing off-resonance effects into m_0 *implies* neglecting exchange between excitation and readout

Two-compartment DESS model [§6.2.2]

- additional approximations required unless we assume time-independent diff in compartmental off-resonance freq

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange
- neglected relaxation, precession, exchange during excitation
- absorbing off-resonance effects into m_0 *implies* neglecting exchange between excitation and readout

Two-compartment DESS model [§6.2.2]

- additional approximations required unless we assume time-independent diff in compartmental off-resonance freq
- including exchange, closed-form solutions still elusive

Multi-compartmental MR signal models

2-compartment SPGR model [Spencer and Fishbein, 2000]

- included first-order physical exchange
- neglected relaxation, precession, exchange during excitation
- absorbing off-resonance effects into m_0 *implies* neglecting exchange between excitation and readout

Two-compartment DESS model [§6.2.2]

- additional approximations required unless we assume time-independent diff in compartmental off-resonance freq
- including exchange, closed-form solutions still elusive

For simplicity, we neglect exchange.

Bayesian MWF Scan Design

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where}$$

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^{\mathsf{T}} \right) \right) \tag{15}$$

Bayesian MWF Scan Design

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where}$$

$$\bar{\Psi}(\mathbf{P}) := \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^{\mathsf{T}} \right) \right) \tag{15}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ flip angle variation
- $\mathbf{P}$ SPGR/DESS nominal flip angles, repetition times

Bayesian MWF Scan Design

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where}$$

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{15}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ flip angle variation
- $\mathbf{P}$ SPGR/DESS nominal flip angles, repetition times
- $\mathbf{W}$ $\mathrm{diag}\left(\left[(\mathsf{E}_{\mathbf{x},\boldsymbol{\nu}}(f_{\mathrm{F}}))^{-1}, \mathbf{0}_5^{\mathsf{T}} \right]^{\mathsf{T}} \right)$

Bayesian MWF Scan Design

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where}$$

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{15}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ flip angle variation
- $\mathbf{P}$ SPGR/DESS nominal flip angles, repetition times
- $\mathbf{W}$ $\mathrm{diag}\left(\left[(\mathsf{E}_{\mathbf{x},\boldsymbol{\nu}}(f_{\mathrm{F}}))^{-1}, \mathbf{0}_5^{\mathsf{T}} \right]^{\mathsf{T}} \right)$
- $\mathsf{E}_{\mathbf{x},\boldsymbol{\nu}}(\cdot)$ approximated via empirical averages of samples drawn from separable prior

Bayesian MWF Scan Design

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where}$$

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{15}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ flip angle variation
- $\mathbf{P}$ SPGR/DESS nominal flip angles, repetition times
- $\mathbf{W}$ $\mathrm{diag}\left(\left[(\mathsf{E}_{\mathbf{x},\boldsymbol{\nu}}(f_{\mathrm{F}}))^{-1}, \mathbf{0}_5^{\mathsf{T}} \right]^{\mathsf{T}} \right)$
- $\mathsf{E}_{\mathbf{x},\boldsymbol{\nu}}(\cdot)$ approximated via empirical averages of samples drawn from separable prior
- $\mathbb{P}$ nom flip angle, total scan time constraints

Optimized SPGR/DESS Acquisition

	Optimized flip angles (deg)	Optimized rep. times (ms)
SPGR	$38.1, 12.9, 9.2, 33.5$	$50.2, 32.4, 16.4, 11.8$
DESS	$32.0, 40.3, 52.9$	$17.5, 98.0, 37.6$

Table 3: Optimized Scan Parameters, $\check{\mathbf{P}}$

- Predicted MWF relative standard deviation in WM
  - Optimized SPGR/DESS: $\sqrt{\bar{\Psi}\left(\check{\mathbf{P}}\right)} = 0.285$
  - mcDESPOT: at least 1 [Lankford and Does, 2013]

MWF Estimation via KRR

Simulation Setup

- At each voxel, generate ground-truth $\mathbf{x}, \boldsymbol{\nu}$
- Using 2-compartment SPGR/DESS model $\mathbf{s}$ with optimized (and now fixed) parameters $\check{\mathbf{P}}$, generate voxel data $\mathbf{y} \in \mathbb{C}^{10}$

MWF Estimation via KRR

Simulation Setup

- At each voxel, generate ground-truth $\mathbf{x}, \boldsymbol{\nu}$
- Using 2-compartment SPGR/DESS model $\mathbf{s}$ with optimized (and now fixed) parameters $\check{\mathbf{P}}$, generate voxel data $\mathbf{y} \in \mathbb{C}^{10}$

Use KRR to estimate just f_{F}

- Separable prior on $\mathbf{x}$: f_{F}, m_0 uniform; others log-uniform
- $N \leftarrow 10^6$ training points
- $Z \leftarrow 10^3$ kernel approximation order

MWF Estimation via KRR

Simulation Setup

- At each voxel, generate ground-truth $\mathbf{x}, \boldsymbol{\nu}$
- Using 2-compartment SPGR/DESS model $\mathbf{s}$ with optimized (and now fixed) parameters $\check{\mathbf{P}}$, generate voxel data $\mathbf{y} \in \mathbb{C}^{10}$

Use KRR to estimate just f_{F}

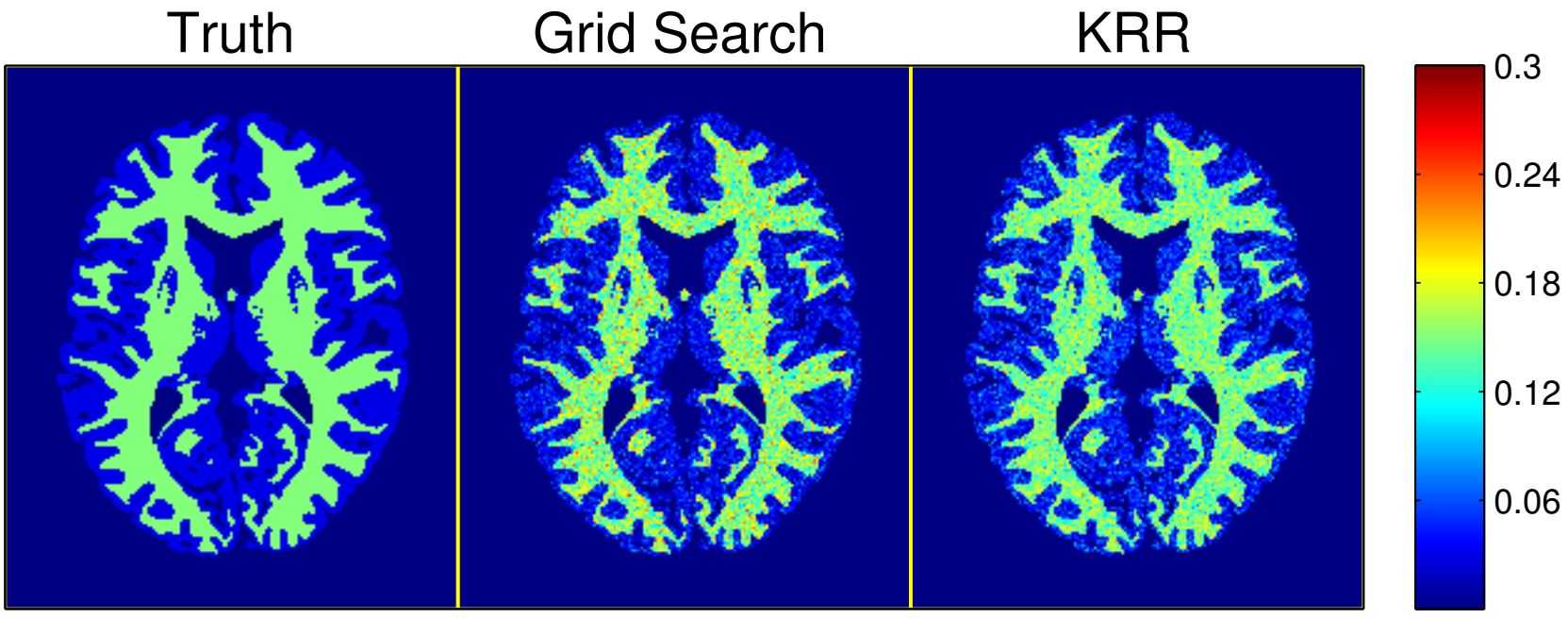
- Separable prior on $\mathbf{x}$: f_{F}, m_0 uniform; others log-uniform
- $N \leftarrow 10^6$ training points
- $Z \leftarrow 10^3$ kernel approximation order

Compare against grid search

- unconstrained search would require $\sim 100^5$ dictionary atoms
- we artificially constrain search here to limit computation

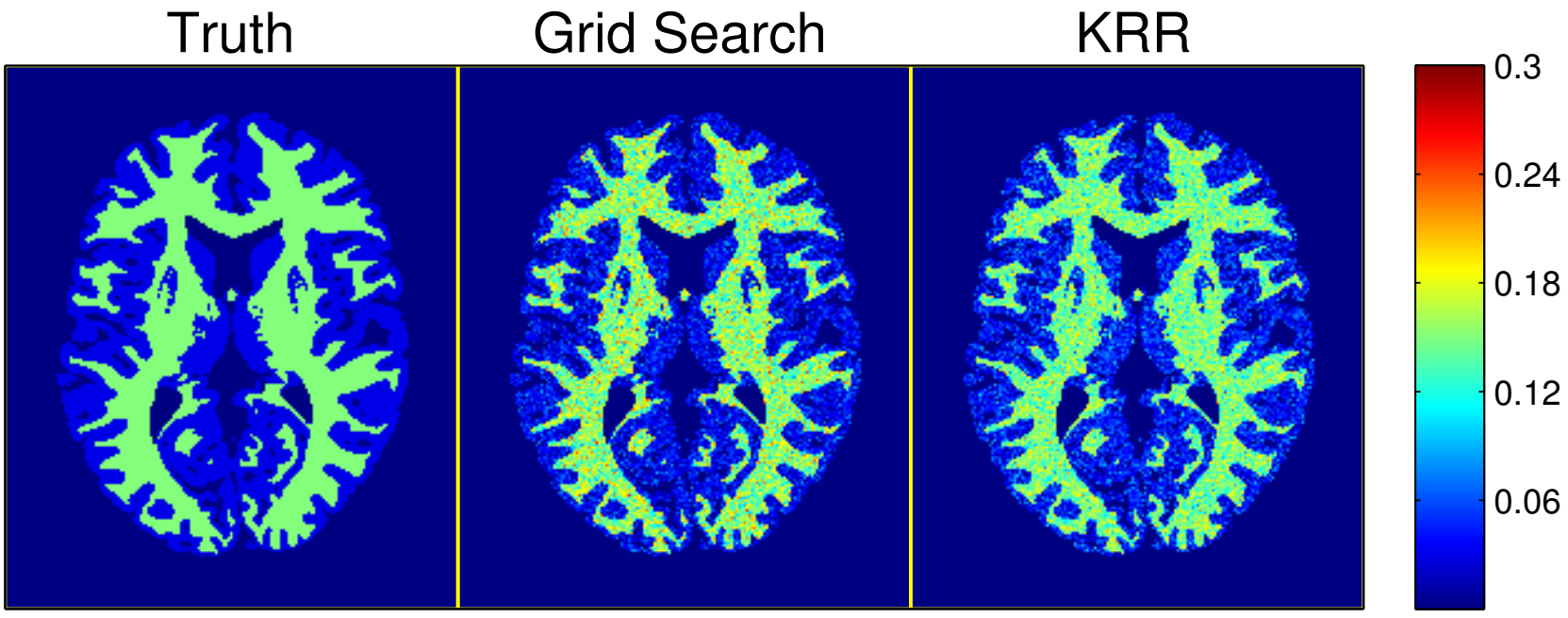
MWF Simulation Result

Fast-fraction f_F estimates, in simulation:



MWF Simulation Result

Fast-fraction f_{F} estimates, in simulation:



$\sim$4h 40s training, 2s testing

MWF Proof-of-concept Experimental Study

Acquired *in vivo* data using optimized MWF protocol

- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **11m48s** total (including BS scan)

MWF Proof-of-concept Experimental Study

Acquired *in vivo* data using optimized MWF protocol

- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **11m48s** total (including BS scan)

Rapidly estimated f_{F} as proxy for MWF

- Full-scale grid search intractable on typical desktop

MWF Proof-of-concept Experimental Study

Acquired *in vivo* data using optimized MWF protocol

- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **11m48s** total (including BS scan)

Rapidly estimated f_{F} as proxy for MWF

- Full-scale grid search intractable on typical desktop
- KRR training and testing took **35s** and **5s/slice**

MWF Proof-of-concept Experimental Study

Acquired *in vivo* data using optimized MWF protocol

- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **11m48s** total (including BS scan)

Rapidly estimated f_{F} as proxy for MWF

- Full-scale grid search intractable on typical desktop
- KRR training and testing took **35s** and **5s/slice**
- Iterative ML refinement took **29s/slice**

MWF Proof-of-concept Experimental Study

Acquired *in vivo* data using optimized MWF protocol

- Used $256 \times 256 \times 8$ 3D matrix over $24 \times 24 \times 4$cm FOV
- Required **11m48s** total (including BS scan)

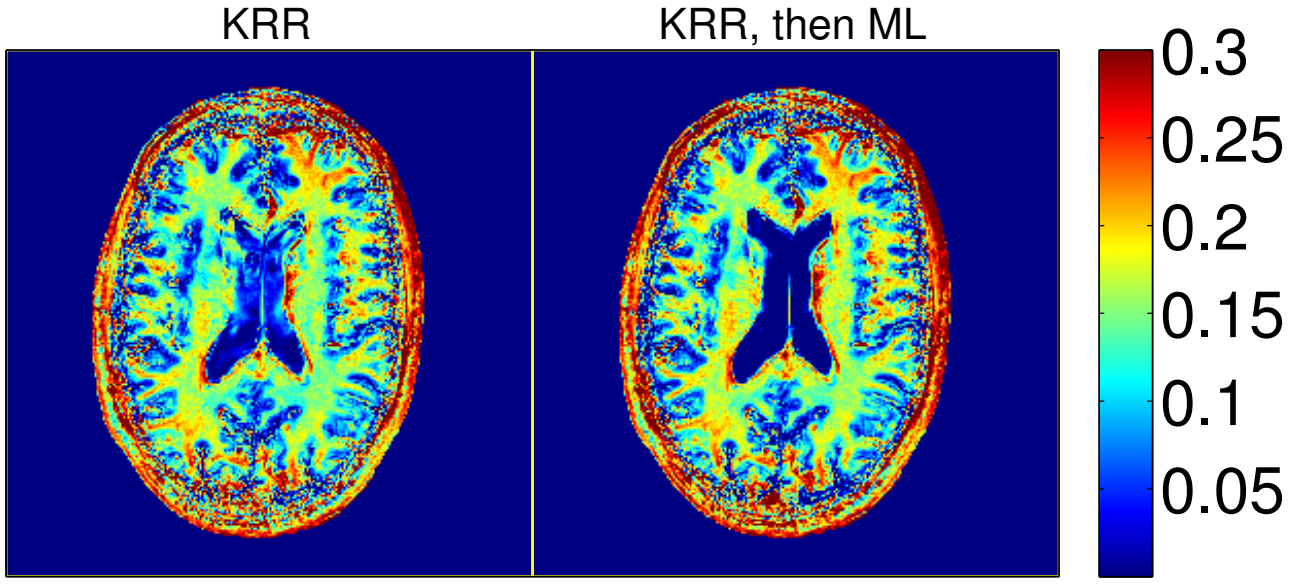
Rapidly estimated f_{F} as proxy for MWF

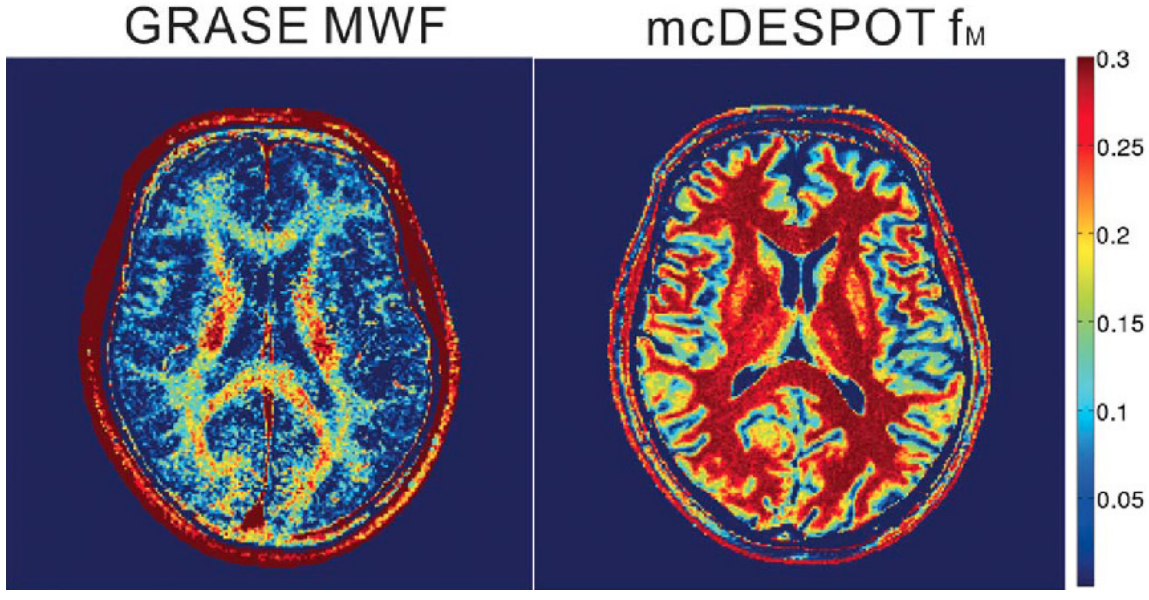
- Full-scale grid search intractable on typical desktop
- KRR training and testing took **35s** and **5s/slice**
- Iterative ML refinement took **29s/slice**

Compared (qualitatively) with results in [Zhang et al., 2015]

- GRASE: accelerated MESE acq [Prasloski et al., 2012]
- mcDESPOT: 9 SPGR, 18 bSSFP scans [Deoni, 2011]

MWF Proof-of-concept In Vivo Result





Summary

Contributions

- Two-compartment DESS signal model
- Fast acquisition for precise WM MWF estimation
- Proof-of-concept *in vivo* MWF images via KRR

Summary

Contributions

- Two-compartment DESS signal model
- Fast acquisition for precise WM MWF estimation
- Proof-of-concept *in vivo* MWF images via KRR

Ongoing work

- Systematic validation

Summary

Contributions

- Two-compartment DESS signal model
- Fast acquisition for precise WM MWF estimation
- Proof-of-concept *in vivo* MWF images via KRR

Ongoing work

- Systematic validation
- Further-optimized MWF acquisition

Future Work

Table 4: Timeline to Defense

start date	task
2017-05	validate KRR estimation
2017-08	prepare KRR journal paper
2017-10	validate f_{F} estimates from SPGR/DESS acquisition
2018-01	prepare fast MWF imaging journal paper
2018-03	defend dissertation

Future Work

Table 4: Timeline to Defense

start date	task
2017-05	validate KRR estimation
2017-08	prepare KRR journal paper
2017-10	validate f_{F} estimates from SPGR/DESS acquisition
2018-01	prepare fast MWF imaging journal paper
2018-03	defend dissertation

Longer-term research directions

- combine image reconstruction and KRR estimation
- correlate our MWF estimates with other myelin biomarkers
- apply KRR to other problems

References i


Cramér, H. (1946).
Mathematical methods of statistics.
Princeton Univ. Press, Princeton.


Aronszajn, N. (1950).
Theory of reproducing kernels.
Trans. Amer. Math. Soc., 68(3):337–404.


Bertsimas, D. and Tsitsiklis, J. (1993).
Simulated annealing.
Statistical Science, 8(1):10–15.


Deoni, S. C. L. (2011).
Correction of main and transmit magnetic field (B0 and B1) inhomogeneity effects in multicomponent-driven equilibrium single-pulse observation of T1 and T2.
Mag. Res. Med., 65(4):1021–35.

References ii

Deoni, S. C. L., Rutt, B. K., Arun, T., Pierpaoli, C., and Jones, D. K. (2008).
Gleaning multicomponent T1 and T2 information from steady-state imaging data.
Mag. Res. Med., 60(6):1372–87.

Keenan, K. E., Stupic, K. F., Boss, M. A., Russek, S. E., Chenevert, T. L., Prasad, P. V., Reddick, W. E., Cecil, K. M., Zheng, J., Hu, P., and Jackson, E. F. (2016).
Multi-site, multi-vendor comparison of T1 measurement using ISMRM/NIST system phantom.
In *Proc. Intl. Soc. Mag. Res. Med.*, page 3290.

Lankford, C. L. and Does, M. D. (2013).
On the inherent precision of mcDESPOT.
Mag. Res. Med., 69(1):127–36.

References iii


Mackay, A., Whittall, K., Adler, J., Li, D., Paty, D., and Graeb, D. (1994).
In vivo visualization of myelin water in brain by magnetic resonance.
Mag. Res. Med., 31(6):673–7.


Nataraj, G., Nielsen, J.-F., and Fessler, J. A. (2017).
Optimizing MR scan design for model-based T1, T2 estimation from steady-state sequences.
IEEE Trans. Med. Imag., 36(2):467–77.


Prasloski, T., Rauscher, A., MacKay, A. L., Hodgson, M., Vavasour, I. M., Laule, C., and Mädler, B. (2012).
Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence.
NeuroImage, 63(1):533–9.

References iv


Rahimi, A. and Recht, B. (2007).
Random features for large-scale kernel machines.
In *NIPS*.


Rasmussen, C. E. and Williams, C. K. I. (2005).
Gaussian processes for machine learning (adaptive computation and machine learning).
MIT Press.


Redpath, T. W. and Jones, R. A. (1988).
FADE-A new fast imaging sequence.
Mag. Res. Med., 6(2):224–34.


Schölkopf, B., Herbrich, R., and Smola, A. J. (2001).
A generalized representer theorem.
In *Proc. Computational Learning Theory (COLT)*, pages 416–426.
LNCS 2111.

References v

Spencer, R. G. and Fishbein, K. W. (2000).
Measurement of spin-lattice relaxation times and concentrations in systems with chemical exchange using the one-pulse sequence: breakdown of the Ernst model for partial saturation in nuclear magnetic resonance spectroscopy.
J. Mag. Res., 142(1):120–35.

Webb, S., Munro, C. A., Midha, R., and Stanisz, G. J. (2003).
Is multicomponent T2 a good measure of myelin content in peripheral nerve?
Mag. Res. Med., 49(4):628–45.

Zhang, J., Kolind, S. H., Laule, C., and MacKay, A. L. (2015).
Comparison of myelin water fraction from multiecho T2 decay curve and steady-state methods.
Mag. Res. Med., 73(1):223–32.

References vi


Zur, Y., Wood, M. L., and Neuringer, L. J. (1991).
Spoiling of transverse magnetization in steady-state sequences.
Mag. Res. Med., 21(2):251–63.