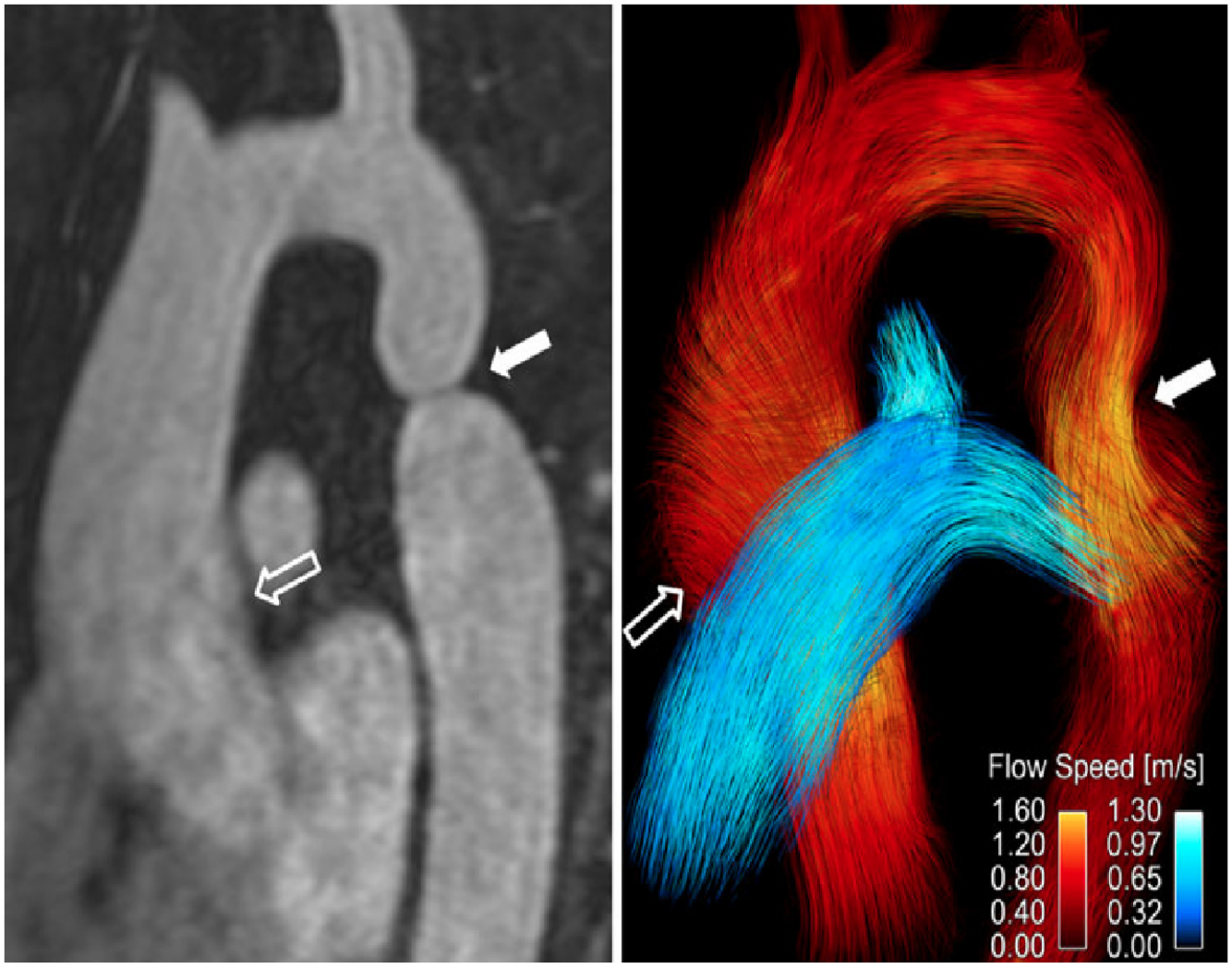
Advances in Quantitative MRI: Acquisition, Estimation, and Application

Gopal Nataraj

Dissertation Defense
March 23, 2018

Dept. of Electrical Engineering and Computer Science
University of Michigan

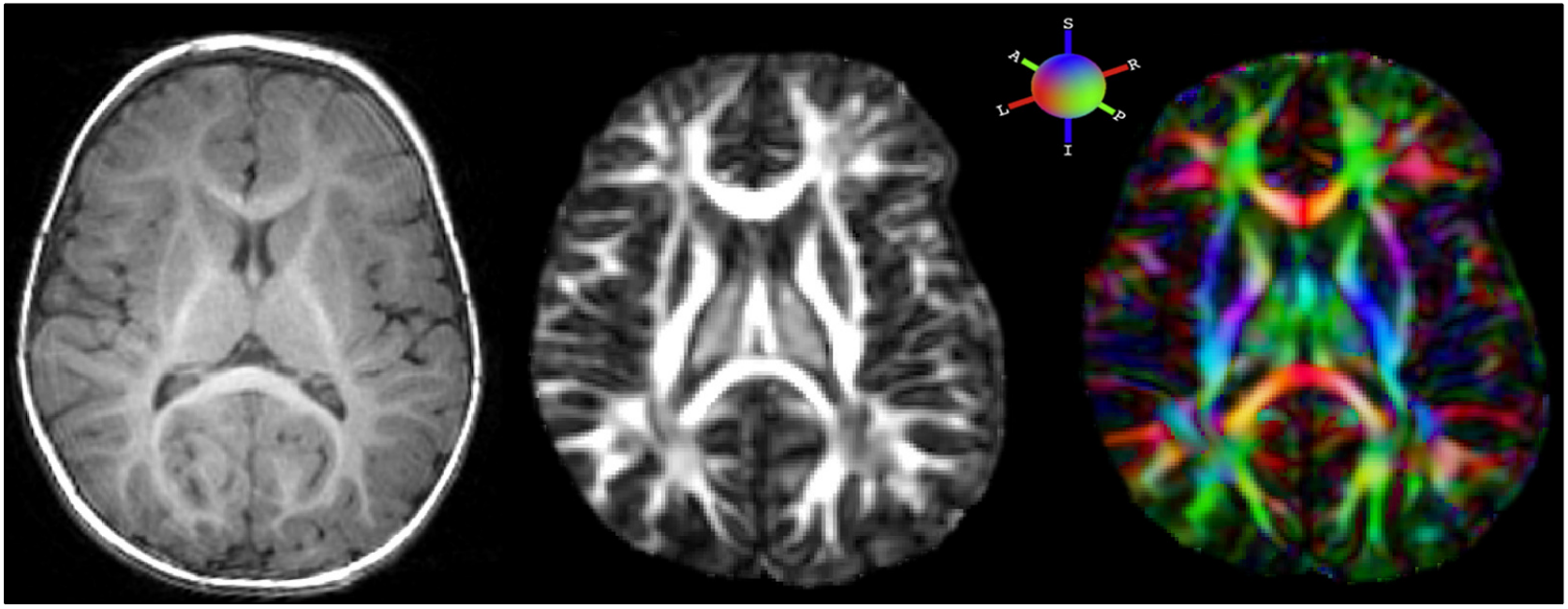
Example: flow imaging



qualitative quantitative[1]

[1] figure borrowed from [Hope et al., 2013]

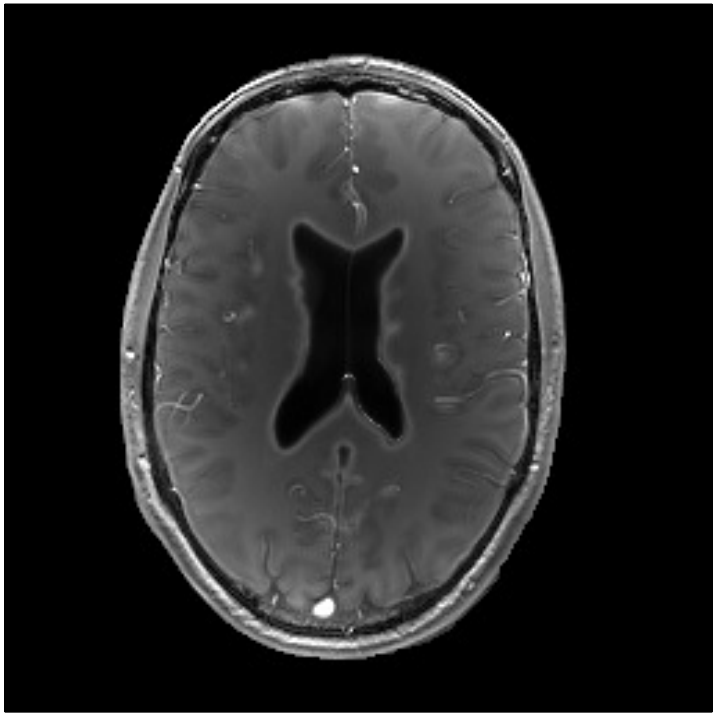
Example: diffusion imaging



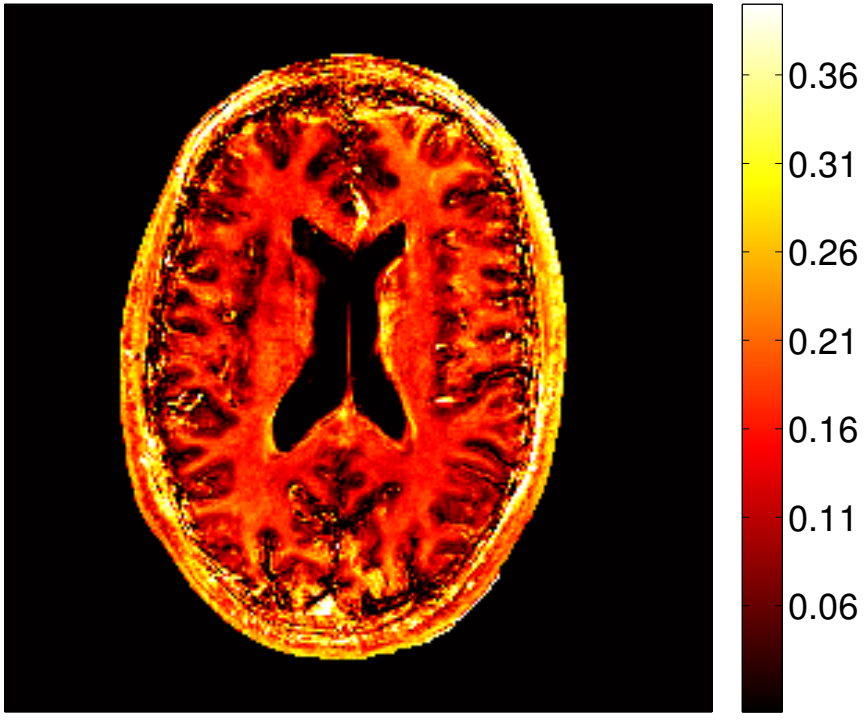
qualitative fractional anisotropy (FA) directional FA[2]

[2] figure borrowed from www.diffusion-imaging.com

Example: myelin water imaging



qualitative



fast-relaxing fraction[3]

[3] figure adapted from [Nataraj et al., 2017a]

Quantitative MRI (QMRI)

Goal: rapidly and accurately localize biomarkers from MR data

Quantitative MRI (QMRI)

Goal: rapidly and accurately localize biomarkers from MR data

- biomarker measurable tissue property (*e.g.*, flow rate)
  that indicates a biological process (*e.g.*, blockage)
  characteristic to specific disorders (*e.g.*, stroke)

Quantitative MRI (QMRI)

Goal: rapidly and accurately localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, flow rate) that indicates a biological process (*e.g.*, blockage) characteristic to specific disorders (*e.g.*, stroke)
- localize — produce quantitative MR images

Quantitative MRI (QMRI)

Goal: rapidly and accurately localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, flow rate) that indicates a biological process (*e.g.*, blockage) characteristic to specific disorders (*e.g.*, stroke)
- localize — produce quantitative MR images
- accurately — physically realistic signal models
- rapidly — fast acquisition, fast estimation

Quantitative MRI (QMRI)

Goal: rapidly and accurately localize biomarkers from MR data

- biomarker — measurable tissue property (*e.g.*, flow rate) that indicates a biological process (*e.g.*, blockage) characteristic to specific disorders (*e.g.*, stroke)
- localize — produce quantitative MR images
- accurately — physically realistic signal models
- rapidly — fast acquisition, fast estimation

Challenge: rapidly vs. accurately often competing goals

- more accurate models typically depend on more markers
- precisely estimating more markers usually requires longer scans and more computation

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 accurate models and informative data,
  how can we rapidly 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 accurate models and informative data, how can we rapidly 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 accurate models and informative data, how can we rapidly 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

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

- $\mathbf{x} \in \mathbb{R}^L$ unknown 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* is a set of D scans that produces at each voxel a measurement vector $\mathbf{y} := [y_1, \ldots, y_D]^\mathsf{T}$ modeled as

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

- $\mathbf{x} \in \mathbb{R}^L$ unknown 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* is a set of D scans that produces at each voxel a measurement vector $\mathbf{y} := [y_1, \ldots, y_D]^\mathsf{T}$ modeled as

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

- $\mathbf{x} \in \mathbb{R}^L$ unknown 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

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

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

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

Maximum-likelihood (ML) estimates achieve CRB asymptotically
or (equivalently for Gaussian data) at sufficiently high SNR.
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., 2017b]

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

  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., 2017b]

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

- Bayesian scan design

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

Scan Design

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

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

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., 2017b]

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

- Bayesian scan design

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

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)
optimal max 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)
optimal max 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.

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 3D SPGR scans in 1m40s total
- Used for T_1, T_2 est from both candidate and reference profiles

Phantom Accuracy Results

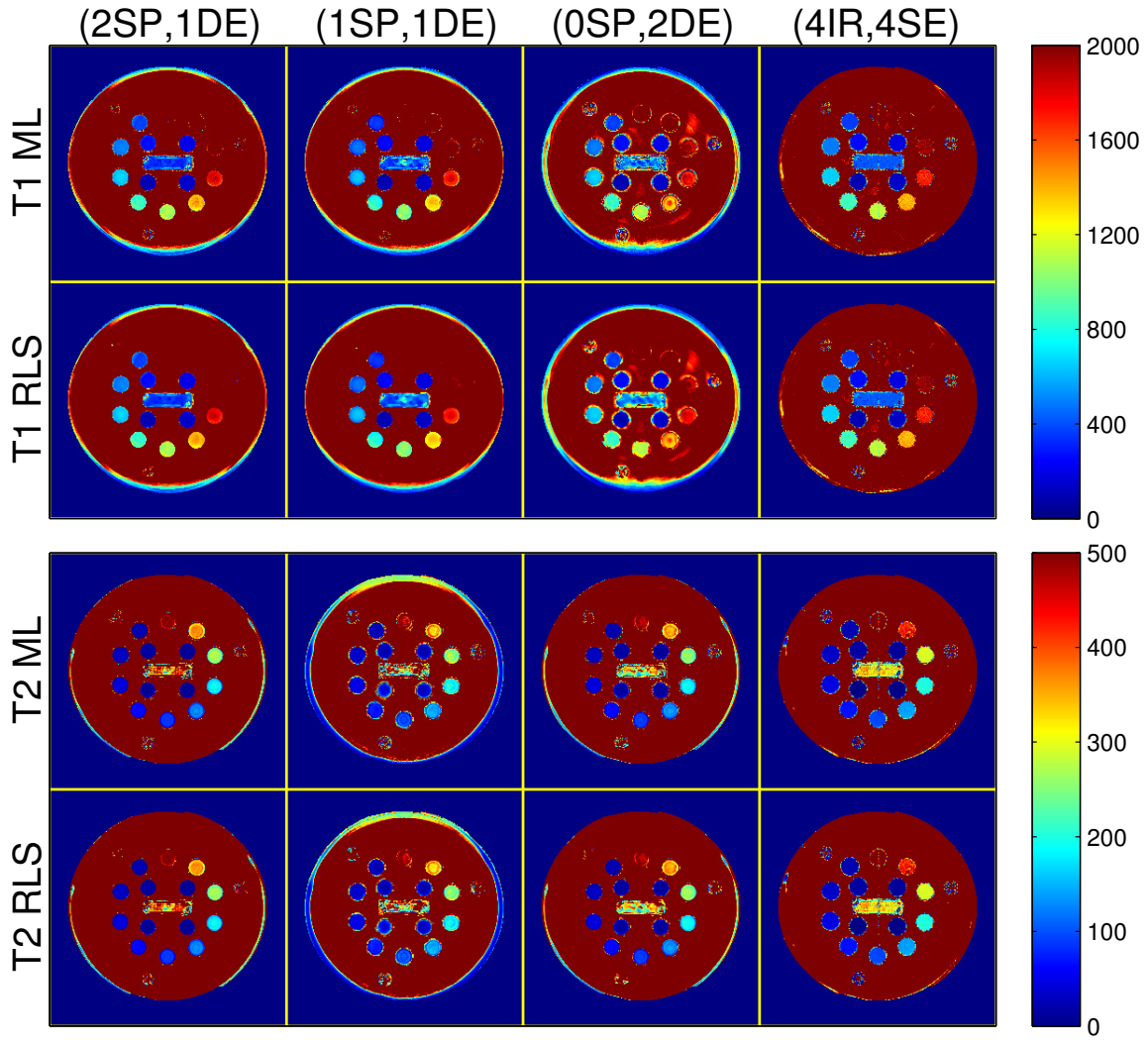
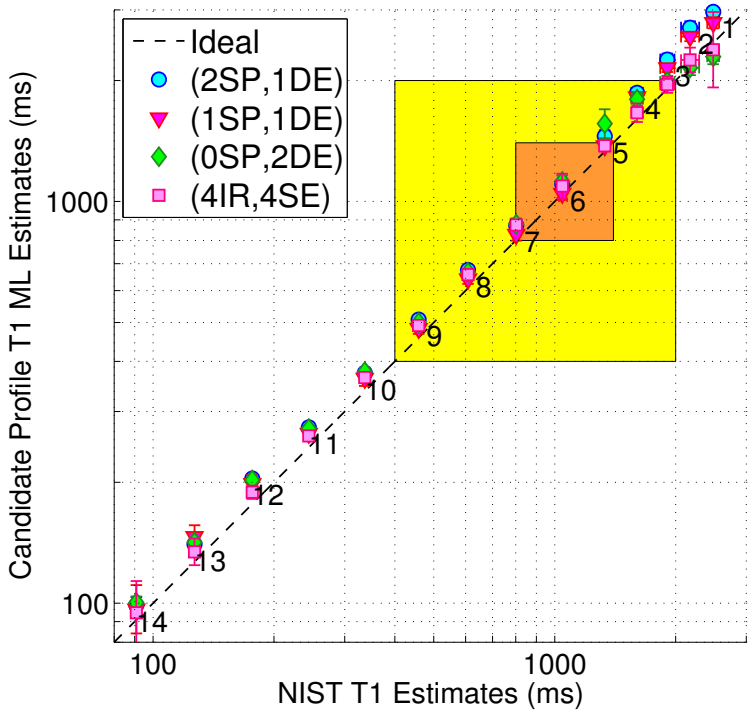
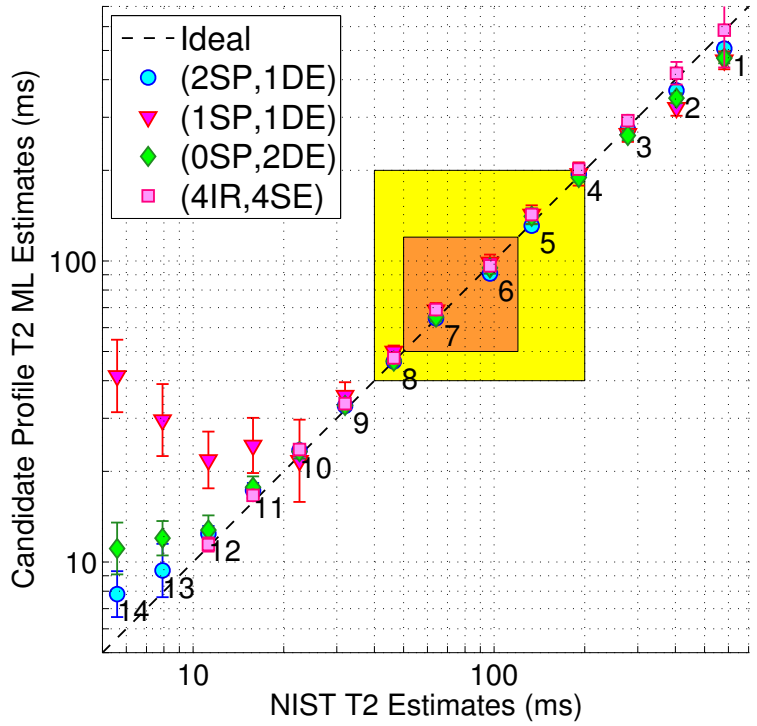


Figure 1: Colorbar ranges in ms.

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	60 ± 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
$\sqrt{\text{opt max cost}}$ estimate	8.9 ± 1.8	11 ± 2.6	$\mathbf{8.3 \pm 2.1}$

Table 1: 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	60 ± 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
$\sqrt{\text{opt max cost}}$ estimate	8.9 ± 1.8	11 ± 2.6	$\mathbf{8.3 \pm 2.1}$

Table 1: 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!

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
    - Phantom (and omitted simulation) results validate method as a predictor of unbiased estimation precision.

Summary

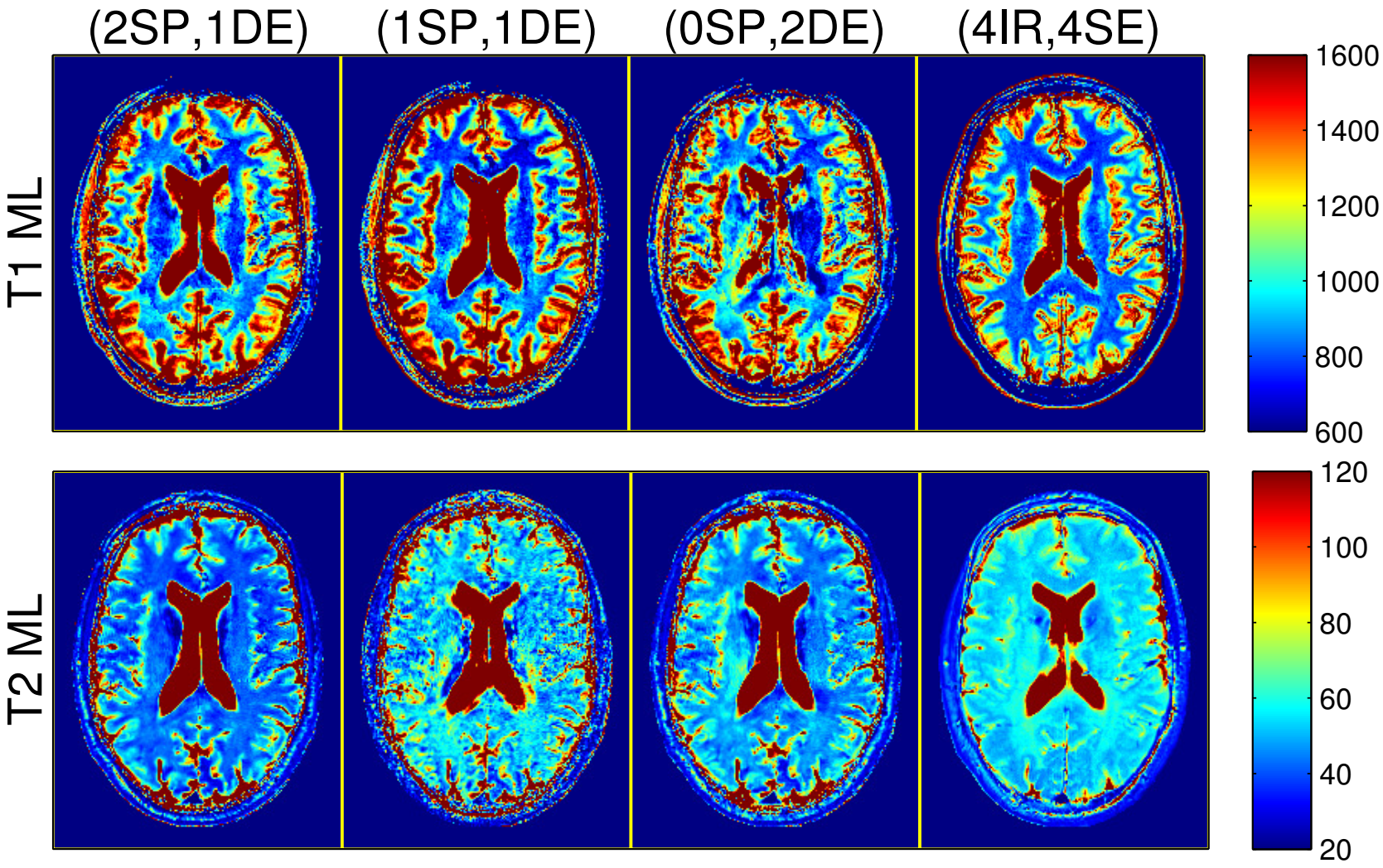


Figure 2: Colorbar ranges in ms.

Summary

Contributions

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

How to address in vivo model mismatch?

- More accurate *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 accurate models and informative data, how can we rapidly quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Signal Model

Given: at every voxel, measurement vector $\mathbf{y} \in \mathbb{C}^D$ modeled as

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

- $\mathbf{x} \in \mathbb{R}^L$ unknown 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 every voxel, measurement vector $\mathbf{y} \in \mathbb{C}^D$ modeled as

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

- $\mathbf{x} \in \mathbb{R}^L$ unknown 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: design fast voxel-by-voxel estimator $\widehat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\nu})$

Prior Work

Task: design fast voxel-by-voxel 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: design fast voxel-by-voxel 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: design fast voxel-by-voxel 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: design fast voxel-by-voxel 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
- dictionary-based grid search

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 $(\mathbf{x}_1, \boldsymbol{\nu}_1, \epsilon_1), \ldots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \epsilon_N)$ 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 $(\mathbf{x}_1, \boldsymbol{\nu}_1, \epsilon_1), \ldots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \epsilon_N)$ 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 $(\mathbf{x}_1, \boldsymbol{\nu}_1, \epsilon_1), \ldots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \epsilon_N)$ 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: Param Estimation via Regression with Kernels (PERK)

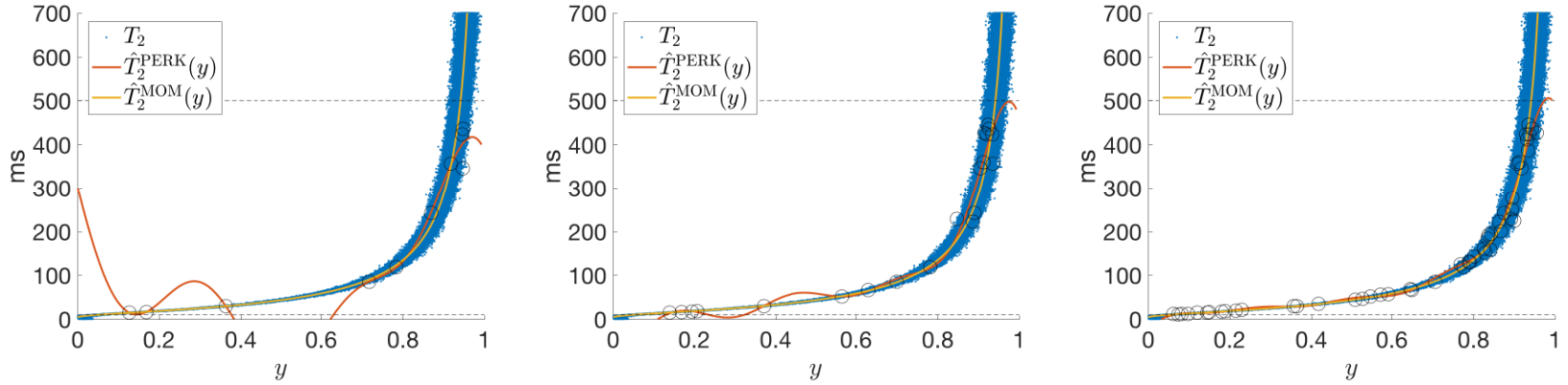
[Nataraj et al., 2018]

- restrict optimization to a certain rich function space $\mathbb{H}$
- optimal $\widehat{h}_l \in \mathbb{H}$ takes form $\widehat{h}_l = \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n)$

[Schölkopf et al., 2001]

PERK in a 1-D Toy Problem

Task: estimate T_2, given samples $(T_{21}, \epsilon_1, y_1), \ldots, (T_{2N}, \epsilon_N, y_N)$ simulated via model $y = \exp(-T_E/T_2) + \epsilon$ with fixed T_E



(a) $N \leftarrow 10$ **(b)** $N \leftarrow 20$ **(c)** $N \leftarrow 50$

Compare: $\widehat{T}_2^{\mathrm{PERK}}$ with method-of-moments (MOM) estimator

$$\widehat{T}_2^{\mathrm{MOM}}(\cdot) := -T_E/\log|\cdot|$$

(PERK more useful when good MOM estimator unavailable)

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (9)$$

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

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (9)$$

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

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (9)$$

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

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K}\mathbf{1}_N \right) \right) \quad (9)$$

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

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K} \mathbf{1}_N \right) \right) \quad (9)$$

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

- Perhaps, since (9) separable in $l \in \{1, \ldots, L\}$ by construction

PERK Solution

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

$$\widehat{x}_l(\cdot) = \mathbf{x}_l^{\mathsf{T}} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho_l \mathbf{I}_N)^{-1} \left(\mathbf{k}(\cdot) - \frac{1}{N} \mathbf{K} \mathbf{1}_N \right) \right) \quad (9)$$

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

- Perhaps, since (9) separable in $l \in \{1, \ldots, L\}$ by construction
- However, explicitly computing $\mathbf{K}$ may be undesirable...

PERK 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 $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{10}$$

PERK 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 $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{10}$$

Plugging (10) into PERK solution (9) and rearranging gives

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

PERK 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 $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{10}$$

Plugging (10) into PERK solution (9) and rearranging gives

$$\widehat{x}_l(\cdot) \approx \widehat{m}_{x_l} + \widehat{\mathbf{c}}_{x_l\tilde{\mathbf{z}}}^\mathsf{T} \left(\widehat{\mathbf{C}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho_l \mathbf{I}_Z \right)^{-1} (\tilde{\mathbf{z}}(\cdot) - \widehat{\mathbf{m}}_{\tilde{\mathbf{z}}}) \tag{11}$$

which is regularized Z-dimensional affine regression!

PERK 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 $\dim(\mathcal{Q}) \ll Z \ll N$

$$\mathbf{K} \approx \tilde{\mathbf{Z}}^\mathsf{T} \tilde{\mathbf{Z}}. \tag{10}$$

Plugging (10) into PERK solution (9) and rearranging gives

$$\widehat{x}_l(\cdot) \approx \widehat{m}_{x_l} + \widehat{\mathbf{c}}_{x_l\tilde{\mathbf{z}}}^\mathsf{T} \left(\widehat{\mathbf{C}}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho_l \mathbf{I}_Z \right)^{-1} (\tilde{\mathbf{z}}(\cdot) - \widehat{\mathbf{m}}_{\tilde{\mathbf{z}}}) \tag{11}$$

which is regularized Z-dimensional affine regression!

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

- Yes, *e.g.* for Gaussian $k(\mathbf{q}, \mathbf{q}') \leftarrow \exp\left(-\frac{1}{2} \left\| \mathbf{\Lambda}^{-1} (\mathbf{q} - \mathbf{q}') \right\|_2^2 \right)$
  [Rahimi and Recht, 2007]
- In such cases, can reduce from $\sim N^2$ to $\sim NZ$ computations

Experimental Setup

Demonstrated PERK for T_1, T_2 est from optim (2SP,1DE) scan

Experimental Setup

Demonstrated PERK for T_1, T_2 est from optim (2SP,1DE) scan

- PERK trained using $N \leftarrow 10^5$ samples from prior dist $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$
- To enable precise estimation, support of $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$ carefully chosen to coincide with min-max acquisition design support

Experimental Setup

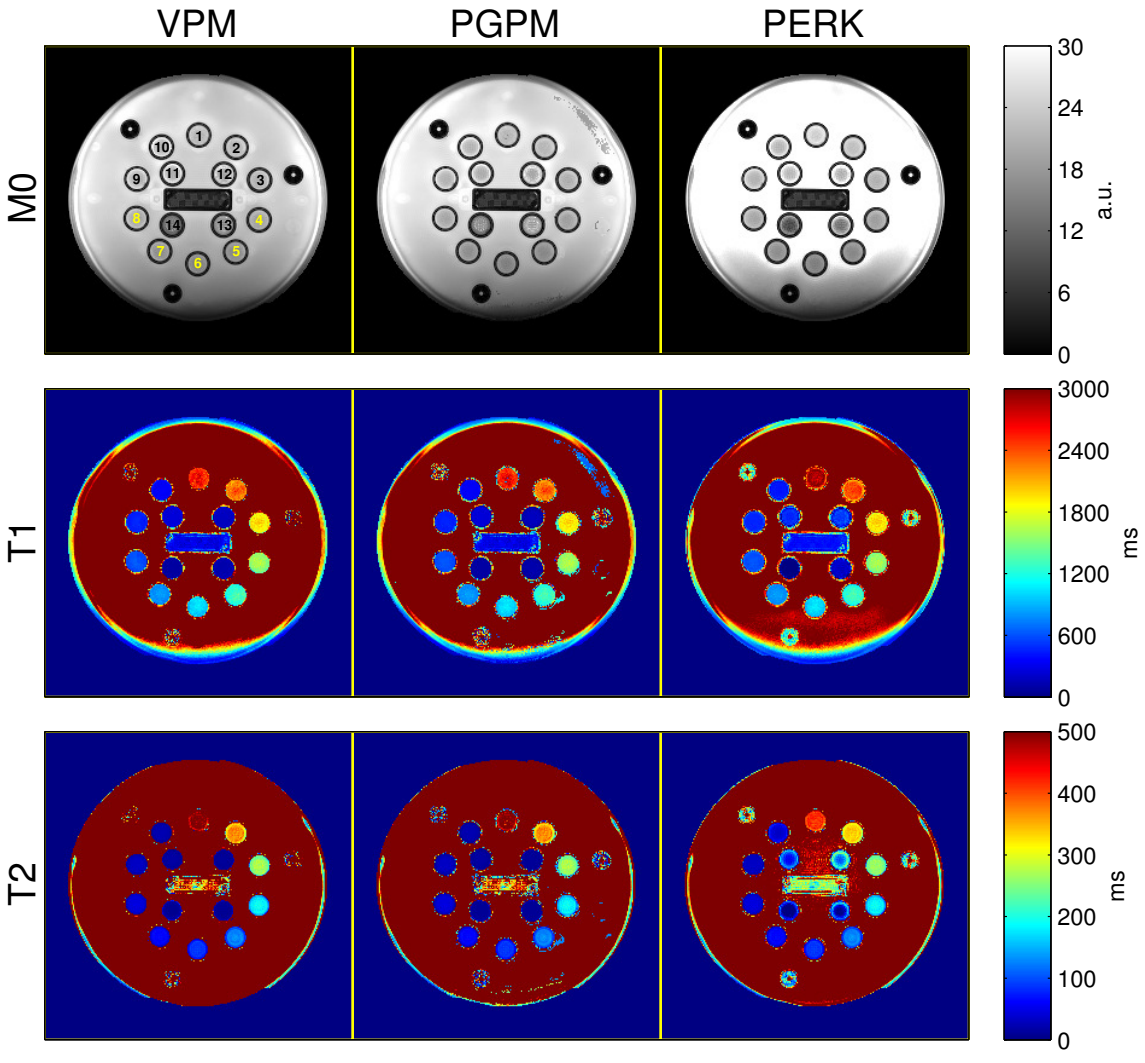
Demonstrated PERK for T_1, T_2 est from optim (2SP,1DE) scan

- PERK trained using $N \leftarrow 10^5$ samples from prior dist $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$
- To enable precise estimation, support of $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$ carefully chosen to coincide with min-max acquisition design support

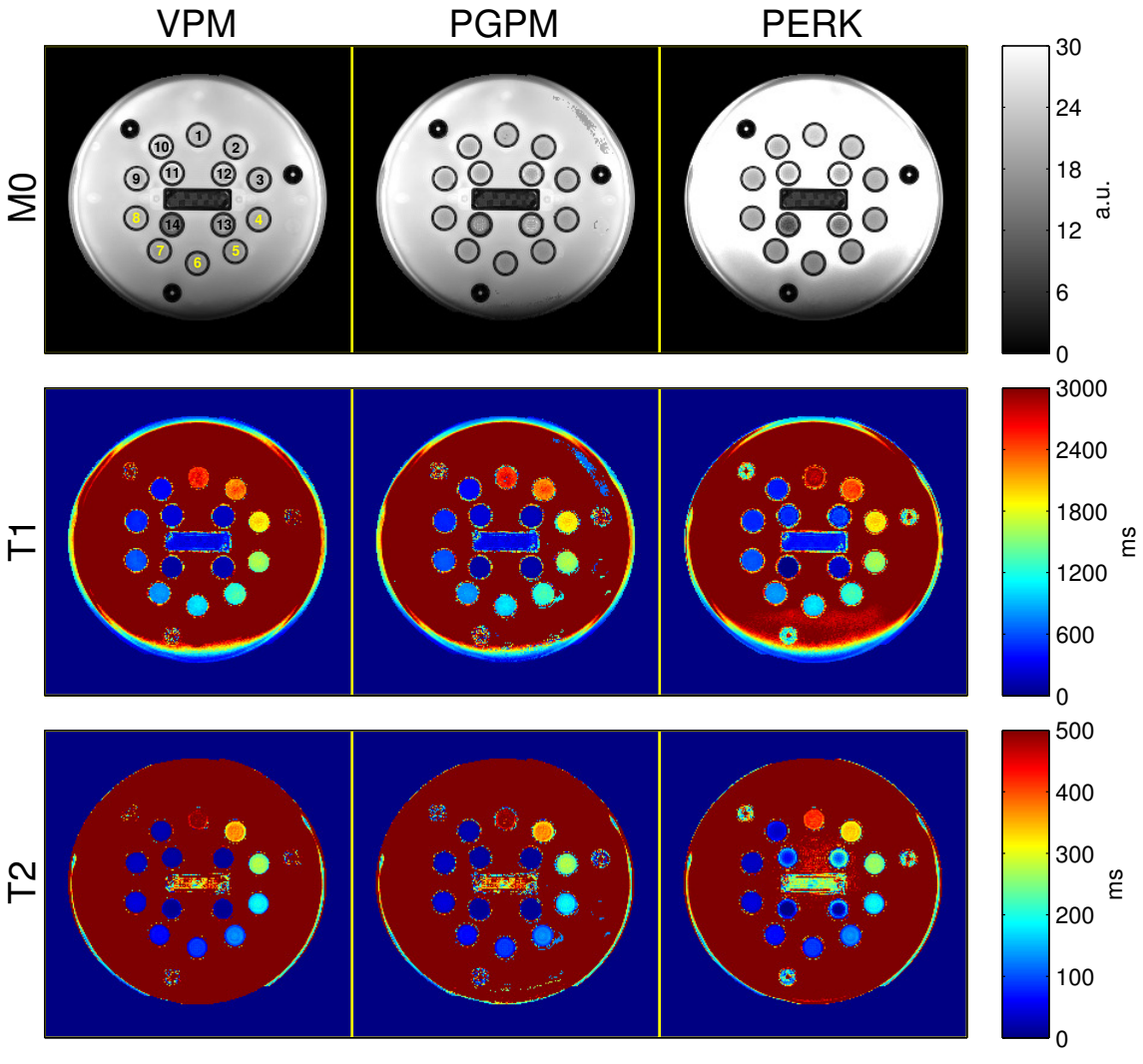
Compared PERK to two well-suited ML estimators:

- dictionary-based grid search estimator via variable projection method (VPM) [Golub and Pereyra, 2003]
- local optim estimator via preconditioned variant (PGPM) of classical gradient projection method [Rosen, 1960]

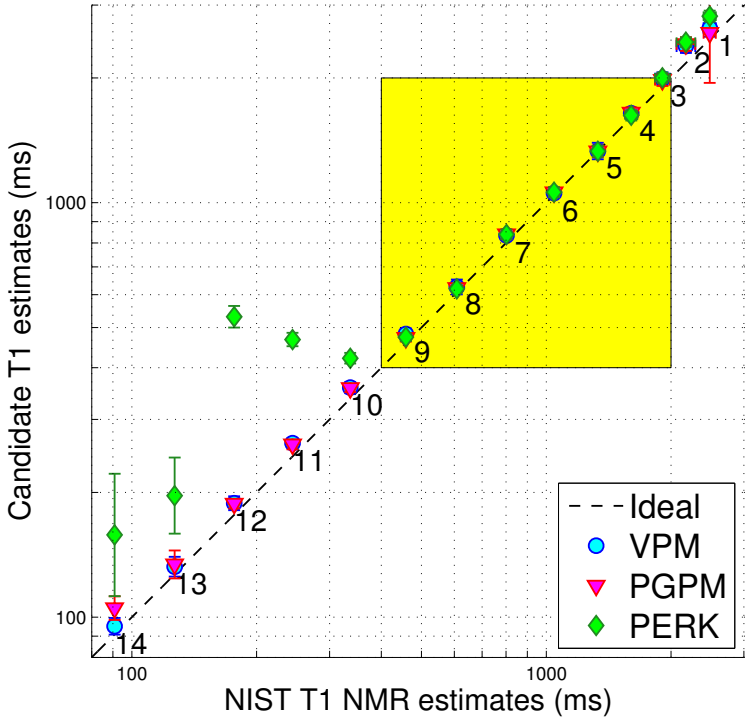
Phantom Results

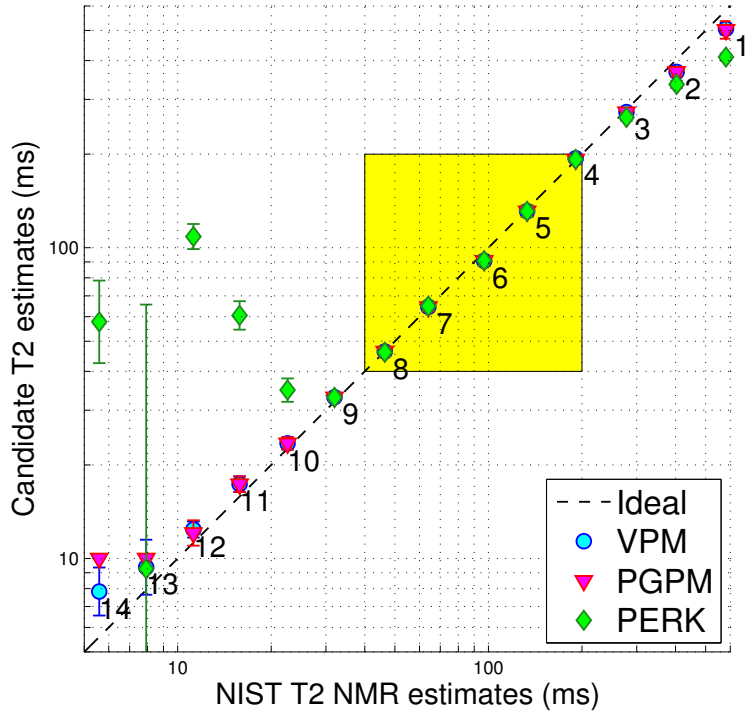


Phantom Results

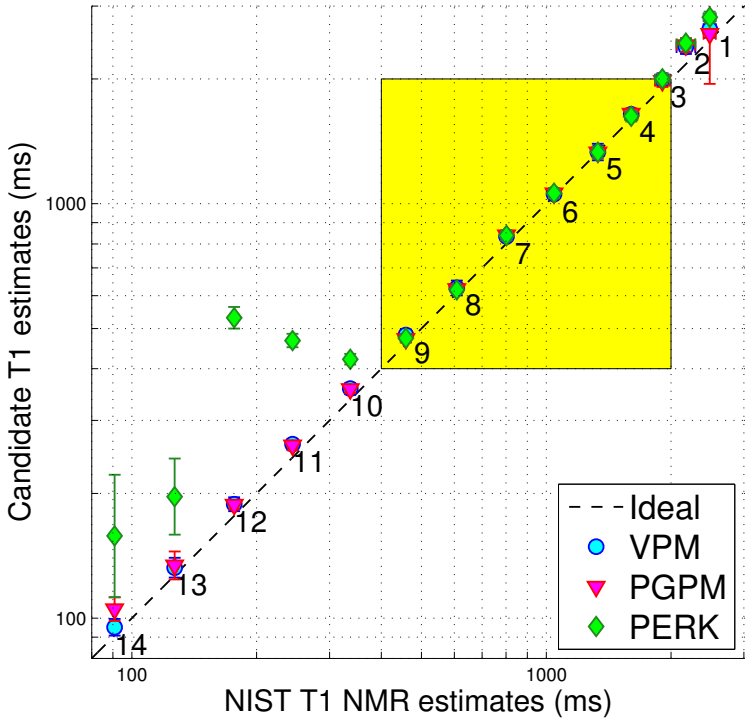


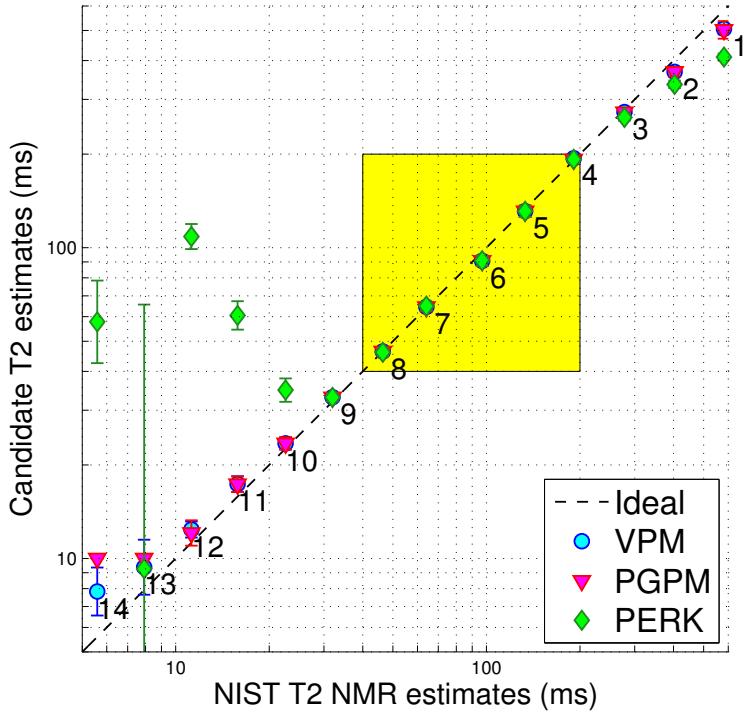
Phantom Results





Phantom Results

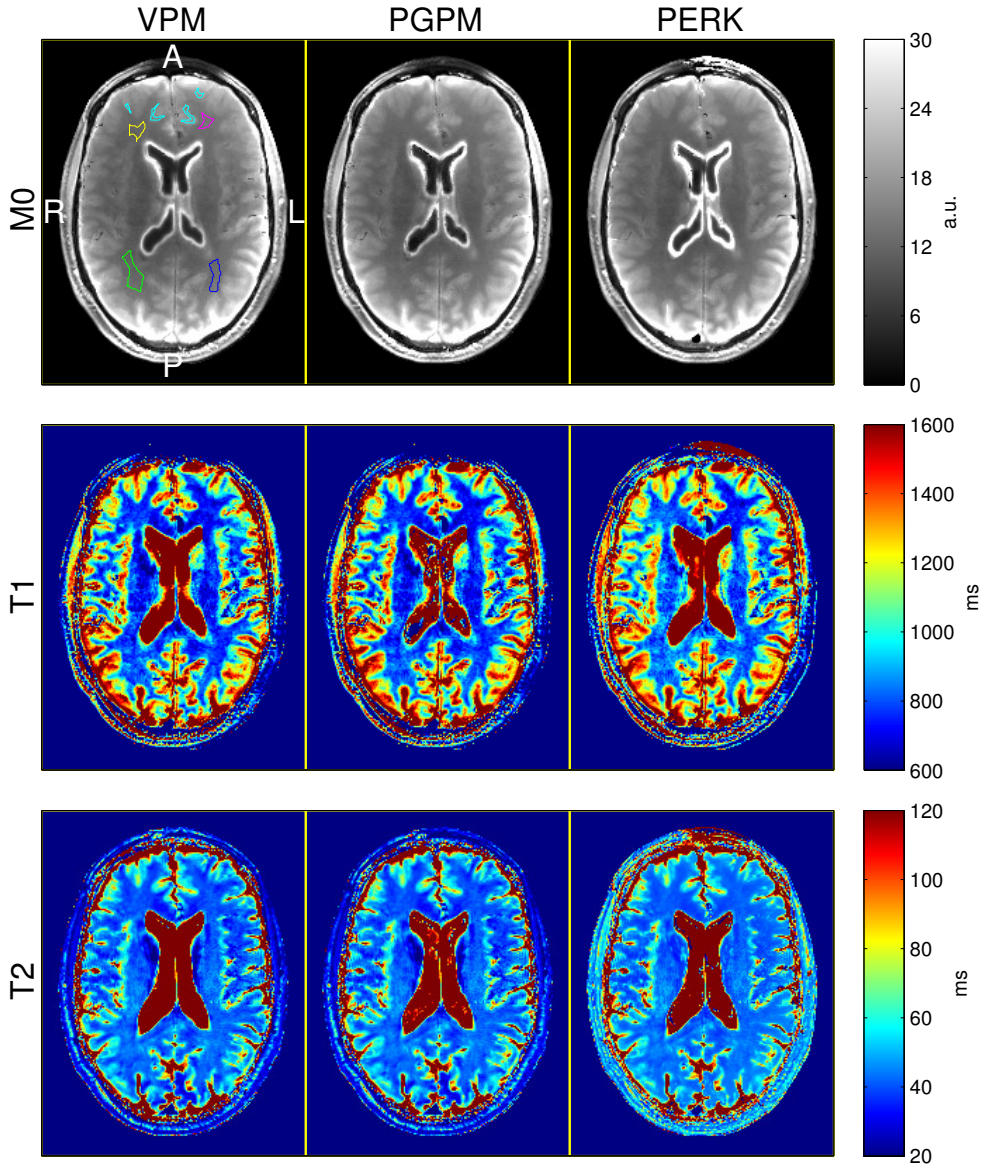




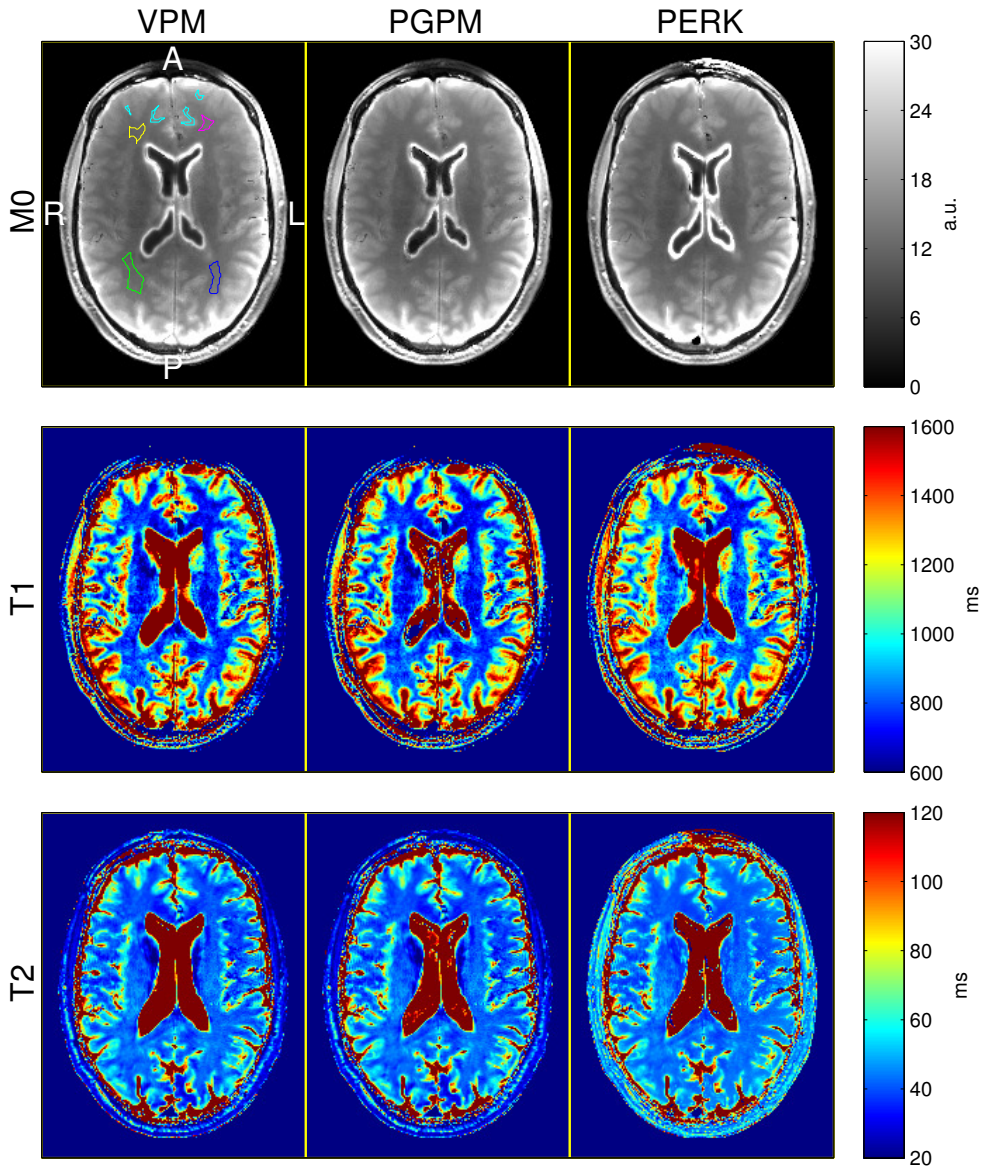
Within support of $p_{\mathbf{x},\boldsymbol{\nu}}$,

PERK and ML estimates agree excellently.

In vivo Results



In vivo Results



838s 2178s 4.2s train
1.6s test

Summary

Contributions

- **PERK**: fast, dictionary-free estimator for QMRI

Summary

Contributions

- **PERK**: fast, dictionary-free estimator for QMRI
- demonstrated PERK for T_1, T_2 estimation
  - Phantom (and omitted simulation) results show that PERK and ML estimators yield comparable accuracy/precision
  - *In vivo* PERK and ML estimates are comparable in WM/GM
  - **PERK is consistently at least** 140**x faster**

Summary

Contributions

- **PERK**: fast, dictionary-free estimator for QMRI
- demonstrated PERK for T_1, T_2 estimation
  - Phantom (and omitted simulation) results show that PERK and ML estimators yield comparable accuracy/precision
  - *In vivo* PERK and ML estimates are comparable in WM/GM
  - **PERK is consistently at least 140x faster**

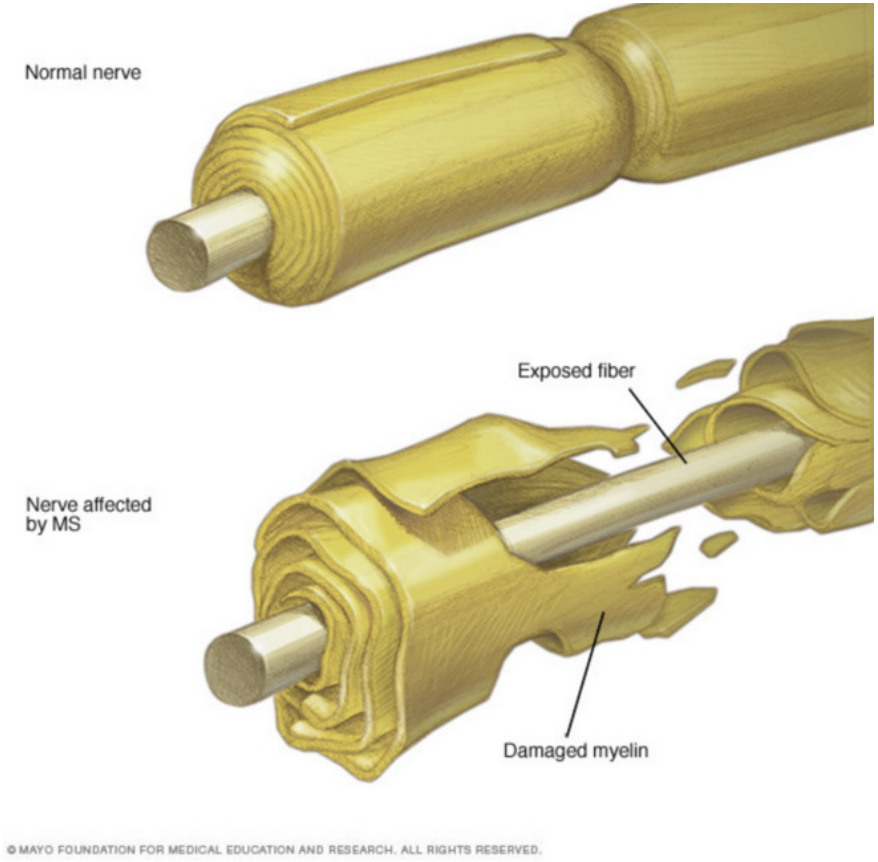
Can we exploit PERK's speed in a more compelling problem?

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 accurate models and informative data, how can we rapidly quantify these biomarkers?
- **Application** [Ch. 6]
  Using these tools, can we design a state-of-the-art biomarker?

Background[4]



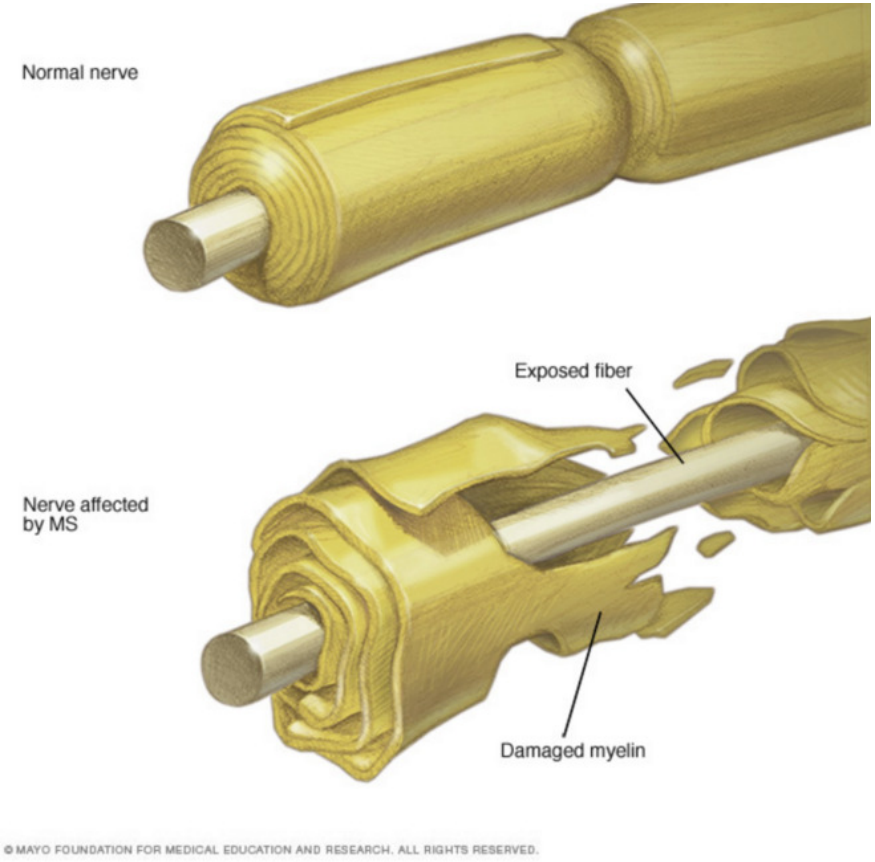
[4] figure borrowed from www.mayoclinic.org

Background[4]

Myelin water fraction (MWF):

- MR signal fraction arising from water trapped within myelin bilayers relative to total signal

[Mackay et al., 1994]

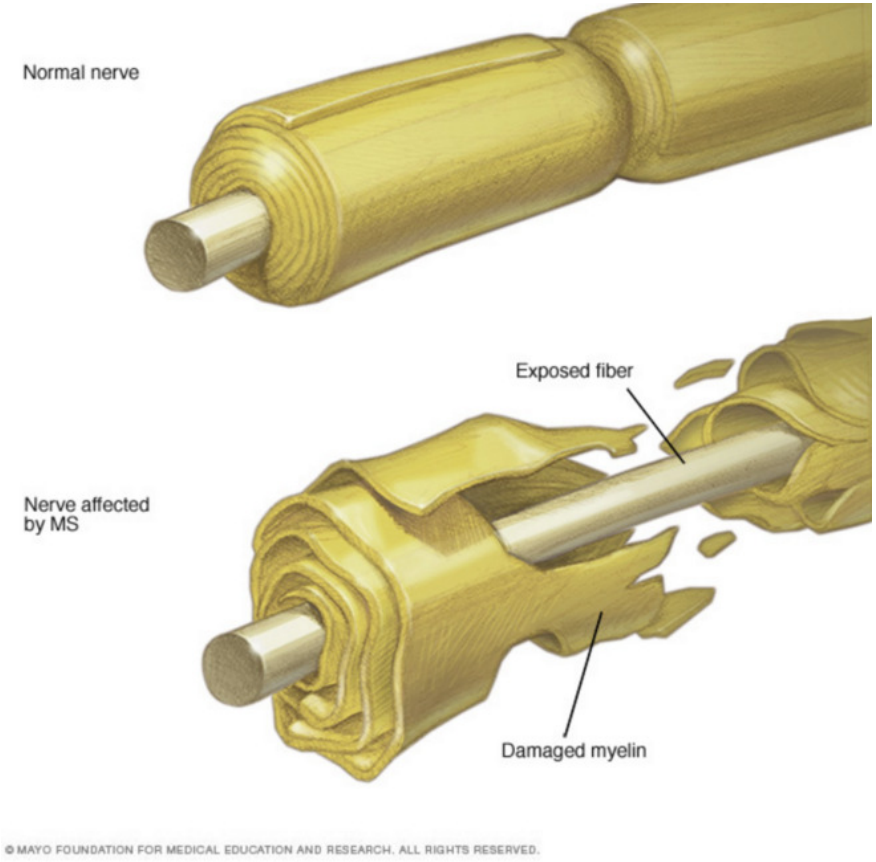


[4] figure borrowed from www.mayoclinic.org

Background[4]

Myelin water fraction (MWF):

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



[4] figure borrowed from www.mayoclinic.org

Previous MW imaging acquisitions

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

- Considered a gold-standard
- Speed-limited by long repetition times ($\sim$1-2s)

Previous MW imaging acquisitions

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

- Considered a gold-standard
- Speed-limited by long repetition times (∼1-2s)

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

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

Previous MW imaging acquisitions

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

- Considered a gold-standard
- Speed-limited by long repetition times ($\sim$1-2s)

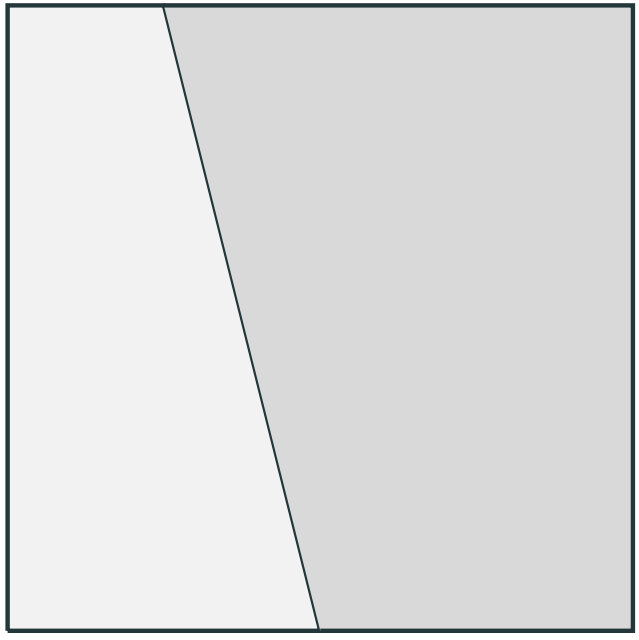
Combinations of fast steady-state scans using variable flip angles ("**mcDESPOT**") [Deoni et al., 2008]

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

Goal: fast, accurate MW content quantification in WM

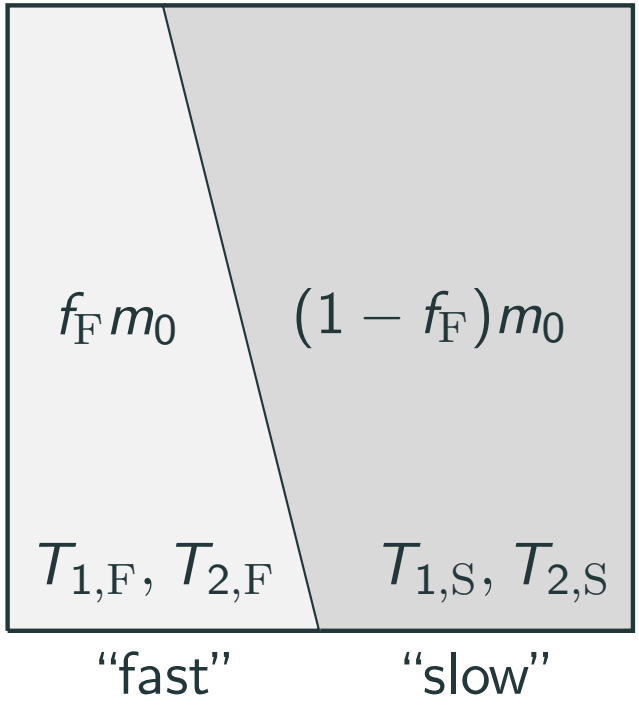
A voxel-scale MW content model

simple two-compartment model



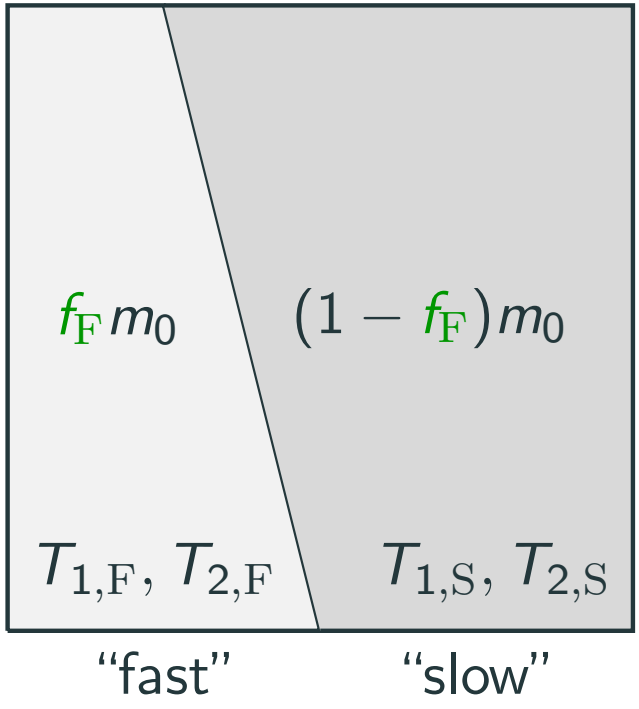
A voxel-scale MW content model

simple two-compartment model



A voxel-scale MW content model

simple two-compartment model



Take fast-relaxing fraction f_F as a simple measure of MW content

Multi-compartmental MR signal models

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

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

Multi-compartmental MR signal models

Two-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

Two-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 [Nataraj et al., 2017a]

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

Multi-compartmental MR signal models

Two-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 [Nataraj et al., 2017a]

- 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

Two-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 [Nataraj et al., 2017a]

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

For simplicity, we use short echo times and neglect exchange.

Bayesian Acquisition Design

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

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{12}$$

Bayesian Acquisition Design

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

$$\bar{\Psi}(\mathbf{P}) := \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{12}$$

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

Bayesian Acquisition Design

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

$$\bar{\Psi}(\mathbf{P}) := \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{12}$$

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

Bayesian Acquisition Design

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

$$\bar{\Psi}(\mathbf{P}) := \mathsf{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{12}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ transmit field spatial variation s^{t}
- $\mathbf{P}$ 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 Acquisition Design

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

$$\bar{\Psi}(\mathbf{P}) := \mathrm{E}_{\mathbf{x},\boldsymbol{\nu}} \left(\mathrm{tr} \left(\mathbf{W} \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \mathbf{W}^{\mathsf{T}} \right) \right) \tag{12}$$

- $\mathbf{x}$ $[f_{\mathrm{F}}, T_{1,\mathrm{F}}, T_{2,\mathrm{F}}, T_{1,\mathrm{S}}, T_{2,\mathrm{S}}, m_0]^{\mathsf{T}}$
- $\boldsymbol{\nu}$ transmit field spatial variation s^{t}
- $\mathbf{P}$ SPGR/DESS nominal flip angles, repetition times
- $\mathbf{W}$ $\mathrm{diag}\left(\left[(\mathrm{E}_{\mathbf{x},\boldsymbol{\nu}}(f_{\mathrm{F}}))^{-1}, \mathbf{0}_5^{\mathsf{T}} \right]^{\mathsf{T}} \right)$
- $\mathrm{E}_{\mathbf{x},\boldsymbol{\nu}}(\cdot)$ 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	–	–
DESS	$33.0, 18.3, 15.1$	$17.5, 30.2, 60.3$

Table 2: Optimized[5] Scan Parameters, $\check{\mathbf{P}}$

- Predicted f_{F} relative standard deviation in WM
  - Optimized SPGR/DESS: $\sqrt{\bar{\Psi}\left(\check{\mathbf{P}}\right)} = 0.425$
  - mcDESPOT: at least 1 [Lankford and Does, 2013]

[5] acquisition design reported by Mingjie Gao, UM

Simulation Studies

Applied PERK for f_{F} estimation from optimized DESS acquisition

Simulation Studies

Applied PERK for f_{F} estimation from optimized DESS acquisition

- PERK trained using $N \leftarrow 10^6$ samples from prior dist $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$
- To enable precise estimation, $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$ chosen similar to Bayesian scan design sampling dist, but with finite support

Simulation Studies

Applied PERK for f_{F} estimation from optimized DESS acquisition

- PERK trained using $N \leftarrow 10^6$ samples from prior dist $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$
- To enable precise estimation, $\mathrm{p}_{\mathbf{x},\boldsymbol{\nu}}$ chosen similar to Bayesian scan design sampling dist, but with finite support

Compared DESS PERK f_{F} estimates to:

- DESS ML f_{F} estimates
  (via VPM and unrealistically narrow grid search around truth)

Simulation Studies

Applied PERK for f_{F} estimation from optimized DESS acquisition

- PERK trained using $N \leftarrow 10^6$ samples from prior dist $\mathrm{p}_{\mathbf{x},\nu}$
- To enable precise estimation, $\mathrm{p}_{\mathbf{x},\nu}$ chosen similar to Bayesian scan design sampling dist, but with finite support

Compared DESS PERK f_{F} estimates to:

- DESS ML f_{F} estimates
  (via VPM and unrealistically narrow grid search around truth)
- MESE nonnegative least-squares (NNLS) MWF f_{M} estimates
  [Mackay et al., 1994]
- MESE regularized NNLS (RNNLS) f_{M} estimates
  [Whittall and MacKay, 1989]

Two-Compartment Simulation without Model Mismatch

Simulated data to arise from **two** water compartments
each with different nominal T_2 values but **same nominal T_1 value**

- DESS f_{F} estimates use known s^{t}
- MESE f_{M} estimates use known s^{t}, bulk T_1

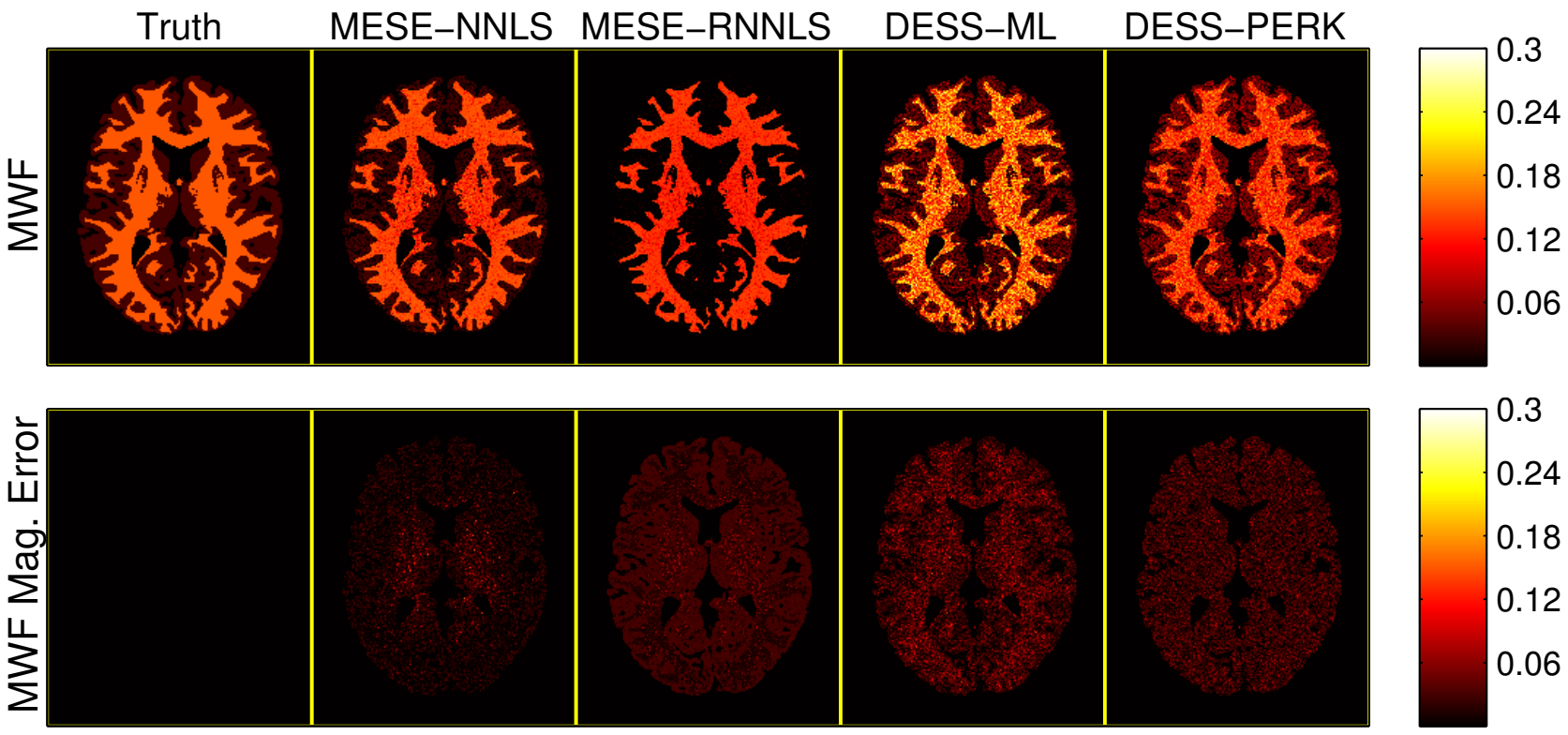
Two-Compartment Simulation without Model Mismatch

Simulated data to arise from **two** water compartments
each with different nominal T_2 values but **same nominal** T_1 **value**

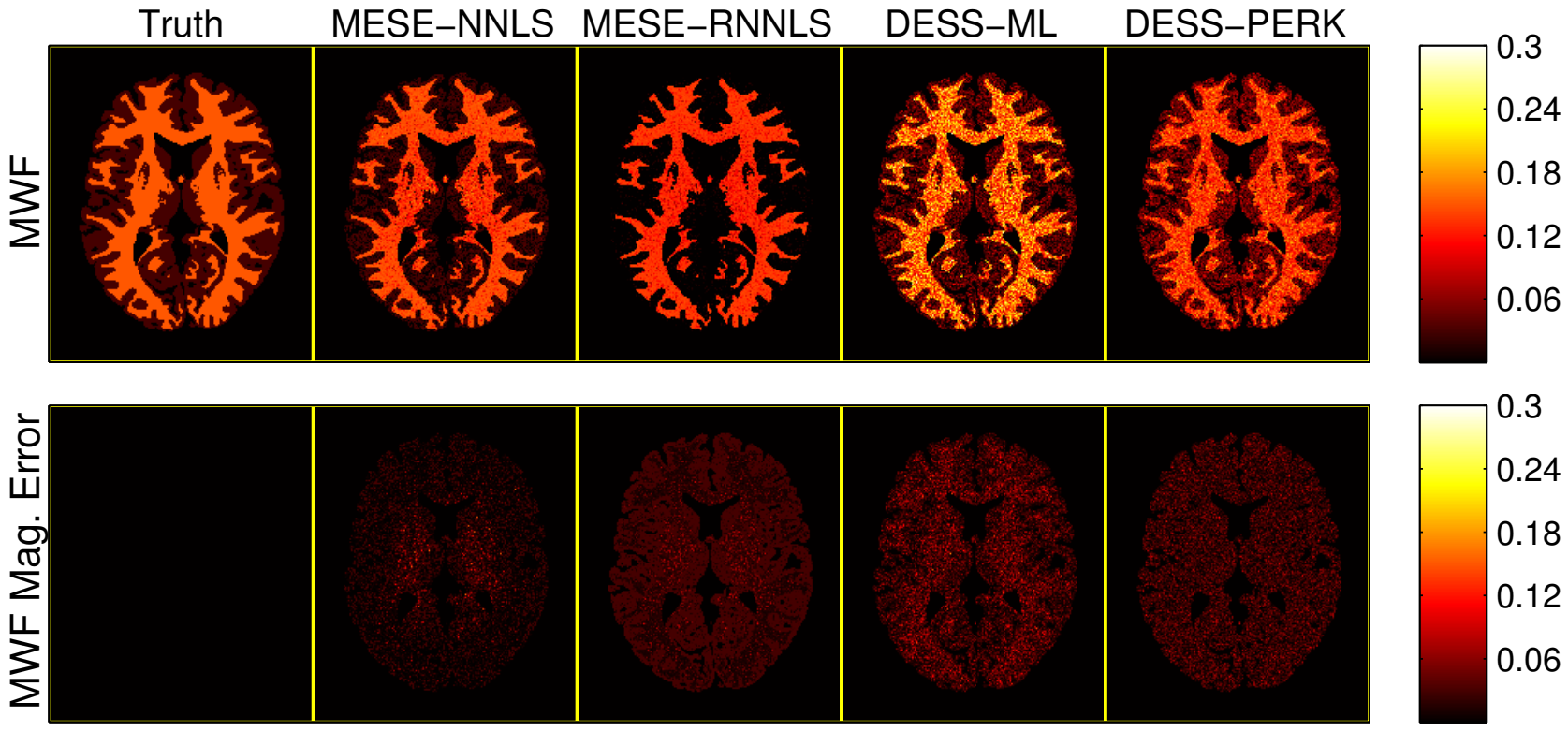
- DESS f_{F} estimates use known s^{t}
- MESE f_{M} estimates use known s^{t}, bulk T_1

Since no model mismatch, f_{F} and f_{M} estimates are comparable

Two-Compartment Simulation Result

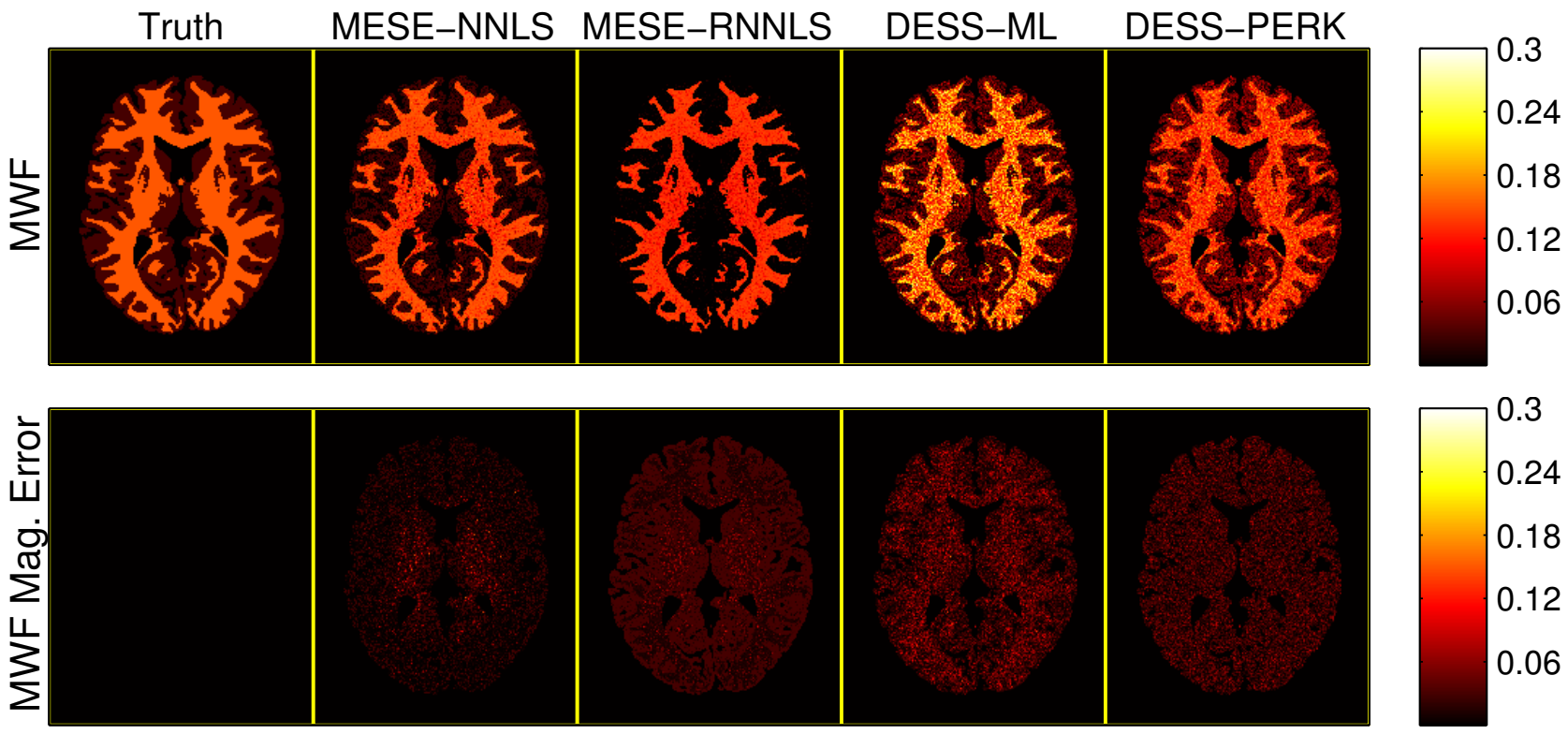


Two-Compartment Simulation Result



estimation time:		∼5h	33.8s train 1.0s test

Two-Compartment Simulation Result



estimation time:			∼5h	33.8s train 1.0s test
WM RMSE:	0.0225	0.0260	0.0433	0.0305

Three-Compartment Simulation with Model Mismatch

Next simulated data to arise from **three** water compartments each with different nominal T_2 **and** T_1 **values**

- DESS f_{F} estimators now incur bias
  due to two-compartment model assumption
- MESE f_{M} estimators now incur bias
  due to bulk-T_1 model assumption (significant for $T_{\mathrm{R}} \sim 1\mathrm{s}$)

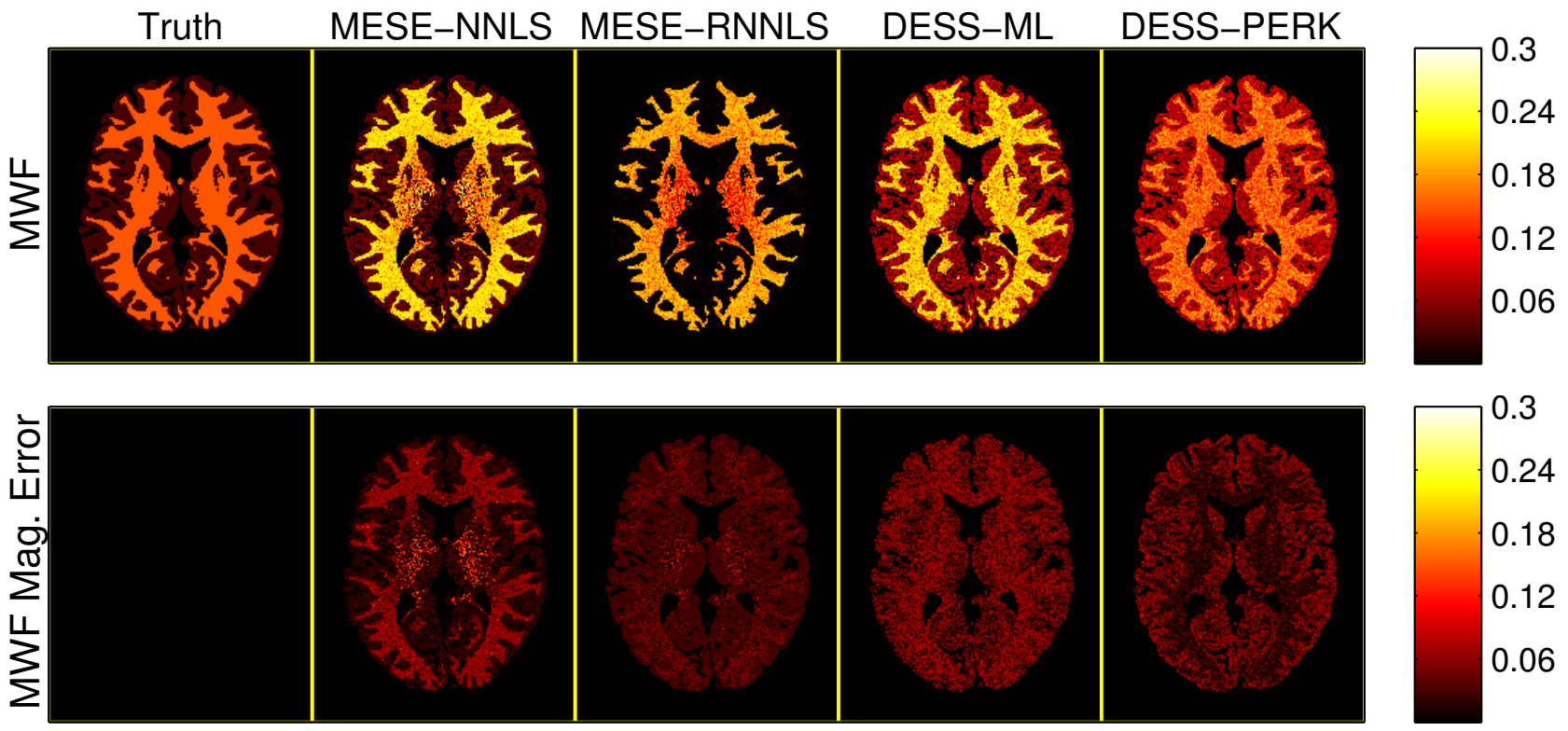
Three-Compartment Simulation with Model Mismatch

Next simulated data to arise from **three** water compartments each with different nominal T_2 **and** T_1 **values**

- DESS f_{F} estimators now incur bias due to two-compartment model assumption
- MESE f_{M} estimators now incur bias due to bulk-T_1 model assumption (significant for $T_{\mathrm{R}} \sim 1\mathrm{s}$)

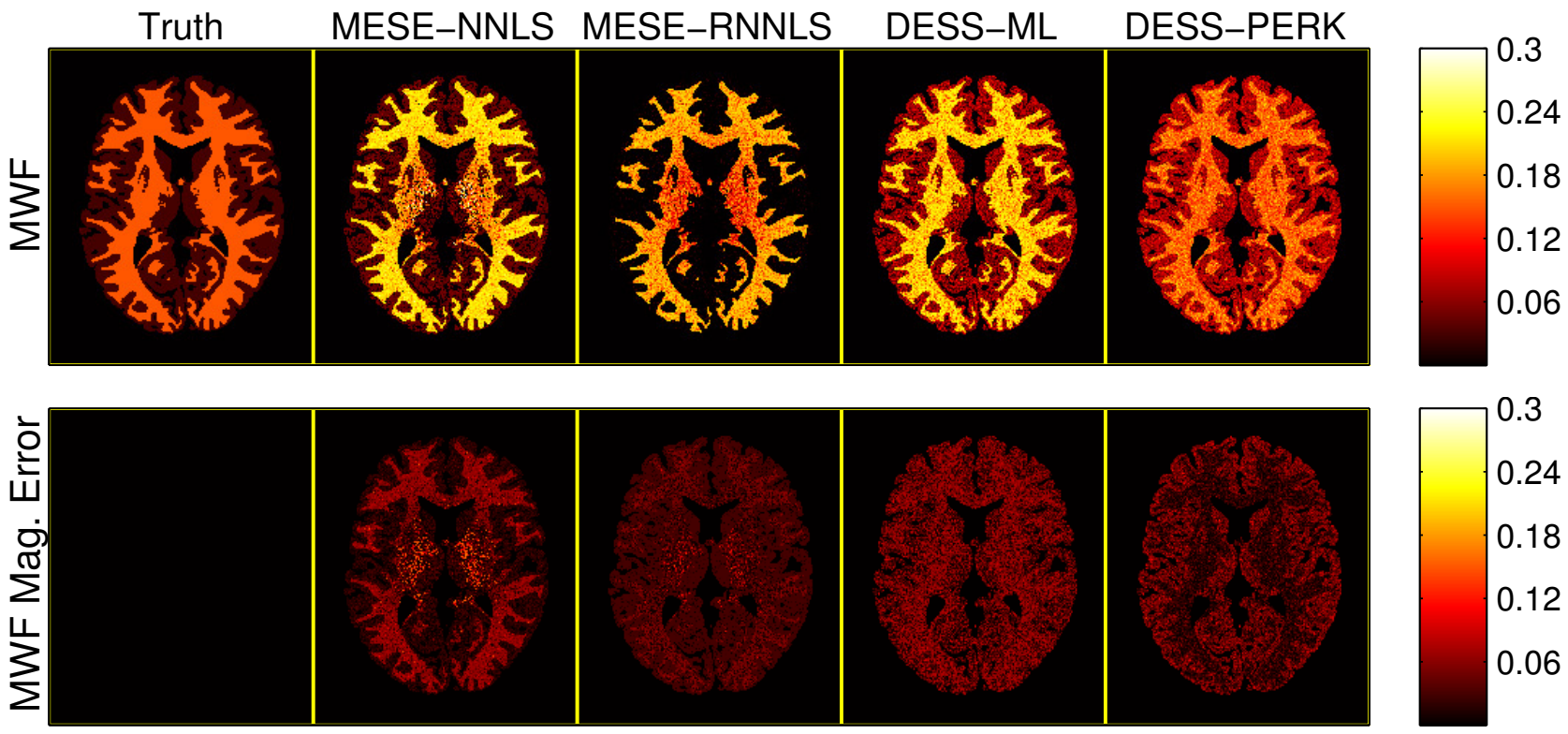
Due to model mismatch,
f_{F} and f_{M} estimates **need not be comparable**.

Three-Compartment Simulation Result



estimation time:		∼5h	34.2s train 1.1s test

Three-Compartment Simulation Result



estimation time:			∼5h	34.2s train 1.1s test
WM RMSE:	0.0618	0.0406	0.0559	0.0254

In vivo Experiment

In a single long study of a healthy volunteer:

- Precision-optimized DESS acquisition
- 32-echo MESE acquisition

- BS acquisition for separate s^{t} estimation
- Variable-flip SPGR acquisition for separate bulk T_1 estimation

In vivo Experiment

In a single long study of a healthy volunteer:

- Precision-optimized DESS acquisition
- 32-echo MESE acquisition
  - Used shaped refocusing pulses
    to suppress out-of-slab signal due to imperfect refocusing
  - Used shorter $T_{\mathrm{R}} \leftarrow 600\mathrm{ms}$ to limit scan time
    (compensated for incomplete recovery w/ separate bulk-T_1 est)
  - Repeated acquisition twice to increase SNR through averaging
- BS acquisition for separate s^{t} estimation
- Variable-flip SPGR acquisition for separate bulk T_1 estimation

In vivo Experiment

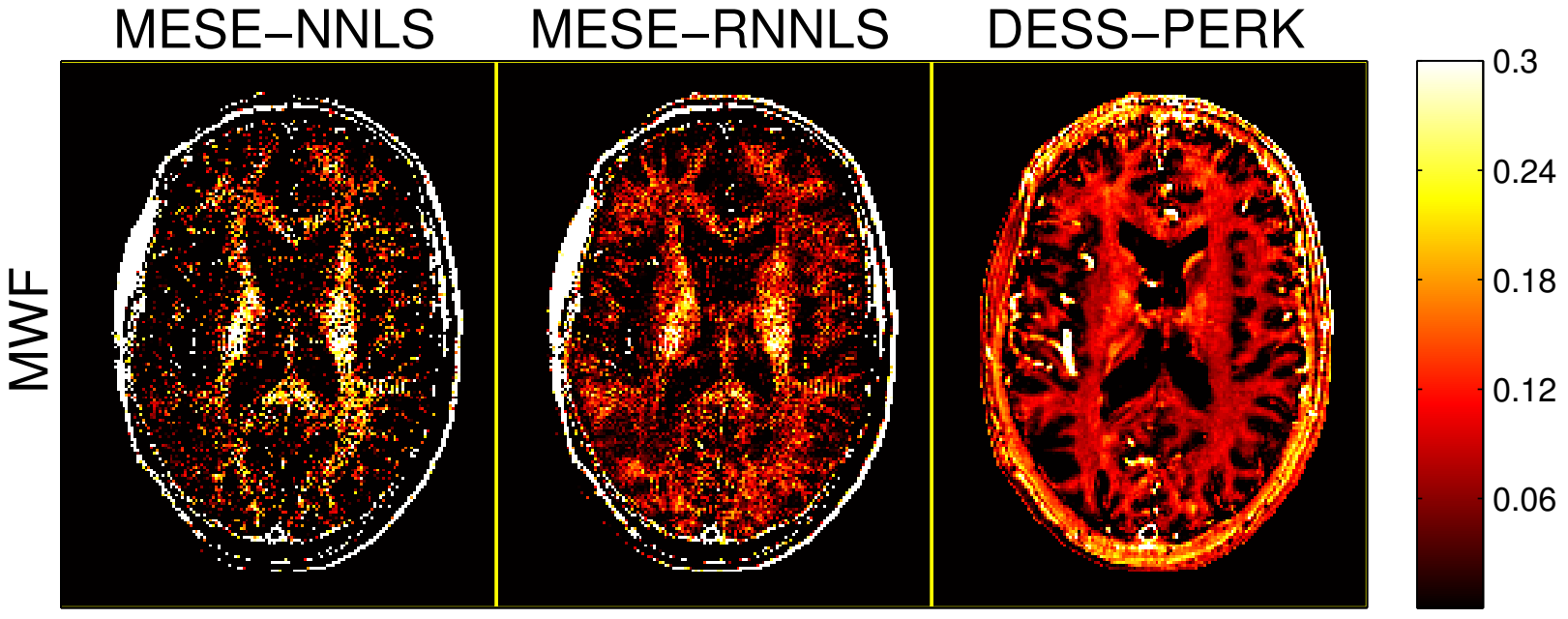
In a single long study of a healthy volunteer:

- Precision-optimized DESS acquisition
- 32-echo MESE acquisition
  - Used shaped refocusing pulses
    to suppress out-of-slab signal due to imperfect refocusing
  - Used shorter $T_{\mathrm{R}} \leftarrow 600\mathrm{ms}$ to limit scan time
    (compensated for incomplete recovery w/ separate bulk-T_1 est)
  - Repeated acquisition twice to increase SNR through averaging
- BS acquisition for separate s^{t} estimation
- Variable-flip SPGR acquisition for separate bulk T_1 estimation

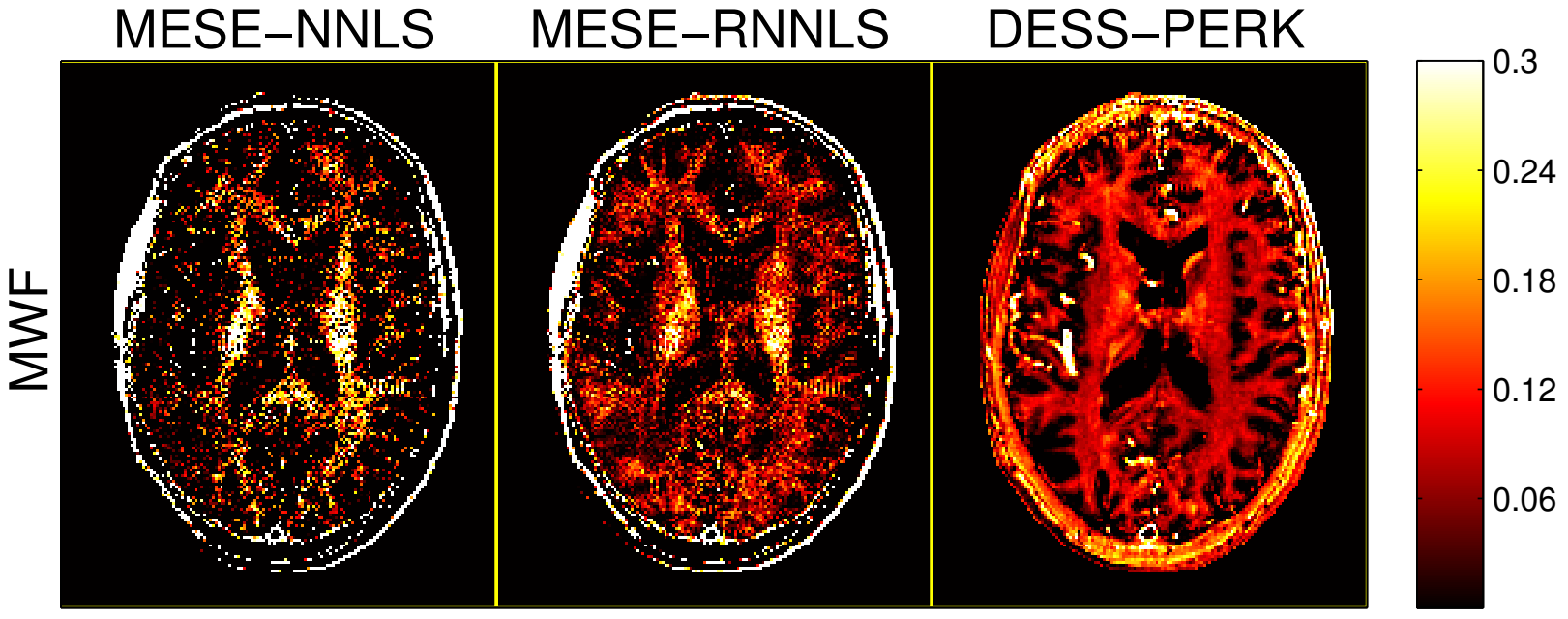
Compared DESS PERK f_{F} estimates to:

- MESE NNLS f_{M} estimates
- MESE RNNLS f_{M} estimates

In vivo Results



In vivo Results



scan time:	40m6s	40m6s	7m45s

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.
  - More realistic simulations demonstrate that both MESE f_{M} and DESS f_{F} estimates are sensitive to model mismatch.

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.
  - More realistic simulations demonstrate that both MESE f_{M} and DESS f_{F} estimates are sensitive to model mismatch.
  - *In vivo* experiments are the first to demonstrate lateral WM MW content estimates from a SS acquisition that are similar to conventional MESE MWF estimates.

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.
  - More realistic simulations demonstrate that both MESE f_{M} and DESS f_{F} estimates are sensitive to model mismatch.
  - *In vivo* experiments are the first to demonstrate lateral WM MW content estimates from a SS acquisition that are similar to conventional MESE MWF estimates.

Future Work

- Investigate DESS f_{F} accuracy in *ex vivo* studies
- Correlate with other myelin biomarkers

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.
  - More realistic simulations demonstrate that both MESE f_{M} and DESS f_{F} estimates are sensitive to model mismatch.
  - *In vivo* experiments are the first to demonstrate lateral WM MW content estimates from a SS acquisition that are similar to conventional MESE MWF estimates.

Future Work

- Investigate DESS f_{F} accuracy in *ex vivo* studies
- Correlate with other myelin biomarkers
- Exploit off-resonance for MW imaging

Summary

Contribution

- Fast SS MRI acquisition for precise MW imaging
  - Idealized simulations demonstrate that PERK and ML f_{F} estimates are comparable but PERK is more than 500$\times$ faster.
  - More realistic simulations demonstrate that both MESE f_{M} and DESS f_{F} estimates are sensitive to model mismatch.
  - *In vivo* experiments are the first to demonstrate lateral WM MW content estimates from a SS acquisition that are similar to conventional MESE MWF estimates.

Future Work

- Investigate DESS f_{F} accuracy in *ex vivo* studies
- Correlate with other myelin biomarkers
- Exploit off-resonance for MW imaging
- Combine PERK with image reconstruction

Overview

Advances in Quantitative MRI:

- **Acquisition** Design via Min-Max Optimization [Ch. 4]
- **Estimation** via Regression with Kernels [Ch. 5]
- **Application** to Myelin Water Imaging [Ch. 6]

Overview

Advances in Quantitative MRI:

- Estimation using Likelihood Models [Ch. 3]
- **Acquisition** Design via Min-Max Optimization [Ch. 4]
- **Estimation** via Regression with Kernels [Ch. 5]
- **Application** to Myelin Water Imaging [Ch. 6]
- Multiple-Dataset Complex Coil Combination [Appx. A]
- SS-Informed RF Pulse Design [Appx. B]

Acknowledgments

Acknowledgments

- advisors Jeff and Jon

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven
- funding — NIH P01 CA87634
  UM "M-Cubed" seed grant
  UM pre-doctoral fellowship

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven
- funding — NIH P01 CA87634
  UM "M-Cubed" seed grant
  UM pre-doctoral fellowship
- friends — "lab of Jeff" (LOJ)
  fMRI lab
  so many others, old and new...

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven
- funding — NIH P01 CA87634
  UM "M-Cubed" seed grant
  UM pre-doctoral fellowship
- friends — "lab of Jeff" (LOJ)
  fMRI lab
  so many others, old and new...
- roommates — Adam and Trey

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven
- funding — NIH P01 CA87634
  UM "M-Cubed" seed grant
  UM pre-doctoral fellowship
- friends — "lab of Jeff" (LOJ)
  fMRI lab
  so many others, old and new...
- roommates — Adam and Trey
- family — three generations here today!

Acknowledgments

- advisors — Jeff and Jon
- committee — Clay, Doug, and Scott
- collaborators — Mingjie, Steven
- funding — NIH P01 CA87634
  UM "M-Cubed" seed grant
  UM pre-doctoral fellowship
- friends — "lab of Jeff" (LOJ)
  fMRI lab
  so many others, old and new...
- roommates — Adam and Trey
- family — three generations here today!
- wife-to-be — Manisha

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., 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.


Golub, G. and Pereyra, V. (2003).
Separable nonlinear least squares: the variable projection method and its applications.
Inverse Prob., 19(2):R1–26.


Hope, M. D., Wrenn, S. J., and Dyverfeldt, P. (2013).
Clinical applications of aortic 4d flow imaging.
Curr. Cardiovasc. Imag. Rep., 6(2):128–39.

References ii

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.

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. (2017a).
Myelin water fraction estimation from optimized steady-state sequences using kernel ridge regression.
In *Proc. Intl. Soc. Mag. Res. Med.*, page 5076.

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

Nataraj, G., Nielsen, J.-F., Scott, C., and Fessler, J. A. (2018).
Dictionary-free MRI PERK: Parameter estimation via regression with kernels.
IEEE Trans. Med. Imag.
To appear.

References iii


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


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


Rosen, J. B. (1960).
The gradient projection method for nonlinear programming, Part I: Linear constraints.
SIAM J. Appl. Math., 8(1):181–217.


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


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.

References iv

Whittall, K. P. and MacKay, A. L. (1989).
Quantitative interpretation of NMR relaxation data.
J. Mag. Res., 84(1):134–52.

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.

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.

Numerical Simulation: Acquisition Design

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

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 $k : \mathcal{Q}^2 \mapsto \mathbb{R}$ that induces choice of RKHS $\mathbb{H}$

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

$$\left(\widehat{h}_l, \widehat{b}_l\right) \in \left\{ \arg \min_{\substack{h_l \in \mathbb{H} \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^{N} (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2 \right\} \tag{13}$$

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

$$\widehat{h}_l(\cdot) \equiv \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n) \tag{14}$$

- Plug (14) into (13); solve now instead for $\left(\widehat{a}_l, \widehat{b}_l\right)$; construct:

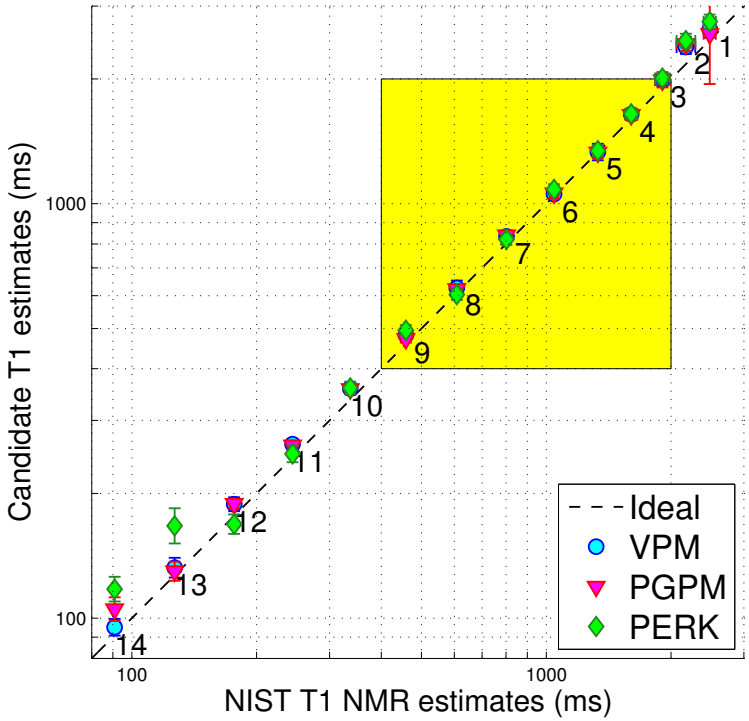
$$\widehat{x}_l(\cdot) = \sum_{n=1}^{N} \widehat{a}_{l,n} k(\cdot, \mathbf{q}_n) + \widehat{b}_l \tag{15}$$

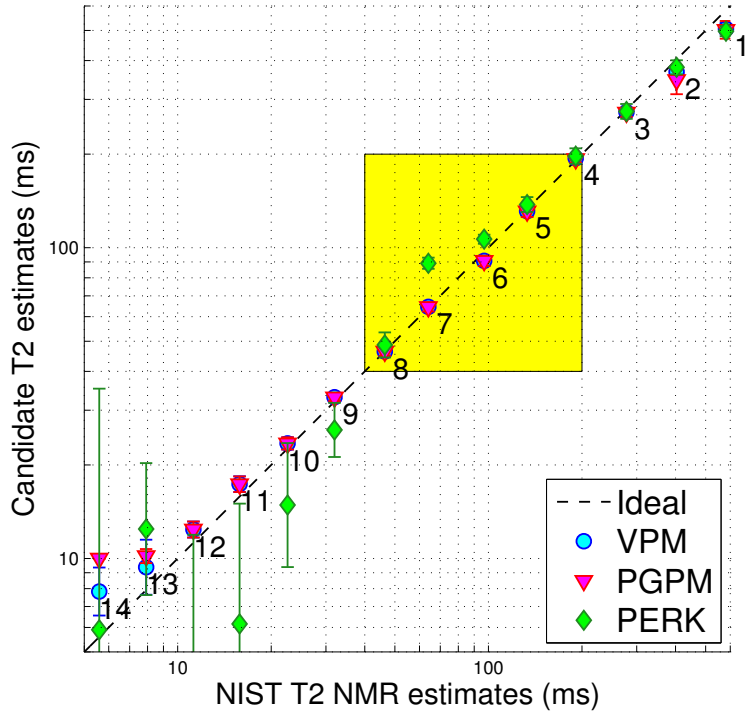
Numerical Simulation: PERK Estimation

	Truth	VPM	PGPM	PERK
WM T_1	832	832.1 $\pm$ 17.2 (17.2)	832.1 $\pm$ 16.2 (16.2)	833.0 $\pm$ 16.5 (16.5)
GM T_1	1331	1331.5 $\pm$ 31.1 (31.1)	1331.2 $\pm$ 29.7 (29.7)	1332.1 $\pm$ 30.4 (30.4)
WM T_2	79.6	79.61 $\pm$ 0.988 (0.988)	79.60 $\pm$ 0.952 (0.952)	79.46 $\pm$ 0.978 (0.989)
GM T_2	110.	110.02 $\pm$ 1.40 (1.40)	110.02 $\pm$ 1.35 (1.35)	109.91 $\pm$ 1.35 (1.35)

Table 4: Sample means $\pm$ sample standard deviations (RMSEs) of VPM, PGPM, and PERK m_0, T_1, T_2 estimates, computed in simulation over 7810 WM-like and 9162 GM-like voxels.

Mismatch in Scan Design vs. Sampling Dist Support





Widening supp($p_{\mathbf{x},\nu}$) degrades performance w/in scan design range.

Thus, scan design and param est should be considered in tandem.