

Published in final edited form as:

*Proc IEEE Int Symp Biomed Imaging*. 2013 December 31; 2013: 1–4. doi:10.1109/ISBI.2013.6556397.

## MODEL-BASED MR PARAMETER MAPPING WITH SPARSITY CONSTRAINT

Bo Zhao<sup>\*,†</sup>, Fan Lam<sup>\*,†</sup>, Wenmiao Lu<sup>†</sup>, and Zhi-Pei Liang<sup>\*,†</sup>

<sup>\*</sup>Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

<sup>†</sup>Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign

### Abstract

MR parameter mapping (e.g.,  $T_1$  mapping,  $T_2$  mapping, or  $T_2^*$  mapping) is a valuable tool for tissue characterization. However, its practical utility has been limited due to long data acquisition time. This paper addresses this problem with a new model-based parameter mapping method, which utilizes an explicit signal model and imposes a sparsity constraint on the parameter values. The proposed method enables direct estimation of the parameters of interest from highly undersampled, noisy  $k$ -space data. An algorithm is presented to solve the underlying parameter estimation problem. Its performance is analyzed using estimation-theoretic bounds. Some representative results from  $T_2$  brain mapping are also presented to illustrate the performance of the proposed method for accelerating parameter mapping.

### Index Terms

parameter mapping; model-based reconstruction; parameter estimation; sparsity

## 1. INTRODUCTION

MR parameter mapping (e.g.,  $T_1$  mapping,  $T_2$  mapping, and  $T_2^*$  mapping) often involves collecting a sequence of images  $\{I_m(\mathbf{x})\}$  with variable parameter-weightings. The parameter-weighted images  $I_m(\mathbf{x})$  are related to the measured  $\mathbf{k}$ -space data by

$$d_m(\mathbf{k}) = \int I_m(\mathbf{x}) \exp(-i2\pi \mathbf{k} \cdot \mathbf{x}) d\mathbf{x} + n_m(\mathbf{k}) \quad (1)$$

where  $n_m(\mathbf{k})$  denotes complex white Gaussian noise. Conventional parameter mapping methods sample  $\mathbf{k}$ -space at the Nyquist rate in the acquisition of each  $d_m(\mathbf{k})$  from which the  $I_m(\mathbf{x})$  are reconstructed, which are followed by parameter estimation from  $I_m(\mathbf{x})$ . One major practical limitation of these parameter mapping methods is long data acquisition time. To alleviate this problem, a number of methods have been proposed to enable parameter mapping from undersampled data. One approach to parameter mapping with sparse sampling is to reconstruct  $I_m(\mathbf{x})$  from undersampled data using various constraints (e.g., sparsity constraint [1] or partial separability constraint (PS) [2]), which is followed by parameter estimation from  $I_m(\mathbf{x})$ . Several successful examples of this approach are described in [3–7]. Another approach is to directly estimate the parameter map from the undersampled  $\mathbf{k}$ -space data, bypassing the image reconstruction step completely [8–10]. This approach

formulates the parameter mapping problem as a statistical parameter estimation problem, which allows for easier performance characterization. Our proposed method uses this approach but also allows sparsity constraint to be effectively used for improved parameter estimation. The advantages of using the sparsity constraint in this context are analyzed theoretically and demonstrated empirically in this paper.

The rest of paper is organized as follows: Section 2 describes the proposed method in detail, including problem formulation, optimization algorithm, and estimation-theoretic characterization; Section 3.1 illustrates the potential benefits of applying sparsity constraint directly to the parameter domain with an example of  $T_2$  brain mapping; Section 4 contains the conclusion of the paper.

## 2. PROPOSED METHOD

We assume that the parameter-weighted image sequence consists of  $M$  image frames, for each of which, a finite number of measurements, denoted as  $\mathbf{d}_m \in \mathbb{C}^{P_m}$ , are collected. For convenience, we use a discrete image model, in which  $\mathbf{I}_m$  is an  $N \times 1$  vector. In this setting, the imaging equation (1) can be written as

$$\mathbf{d}_m = \mathbf{F}_m \mathbf{I}_m + \mathbf{n}_m, \quad (2)$$

$m = 1, \dots, M$ , where  $\mathbf{F}_m \in \mathbb{C}^{P_m \times N}$  denotes the Fourier encoding matrix, and  $\mathbf{n}_m \in \mathbb{C}^{P_m}$  denotes the complex white Gaussian noise with variance  $\sigma^2$ .

### 2.1. Problem Formulation

**2.1.1. Signal model**—In parameter mapping, the parameter-weighted images  $I_m(\mathbf{x})$  can be written as [8–10]

$$I_m(\mathbf{x}) = \rho(\mathbf{x}) \phi(\theta(\mathbf{x}), \boldsymbol{\gamma}_m), \quad (3)$$

where  $\rho(\mathbf{x})$  represents the spin density distribution,  $\theta(\mathbf{x})$  is the desired parameter map (e.g.,  $T_1$ -map,  $T_2$ -map, or  $T_2^*$ -map), and  $\boldsymbol{\gamma}_m$  contains the user-specified parameters for a given data acquisition sequence (e.g., echo time  $T_E$ , repetition time  $T_R$ , flip angle  $\alpha$ , etc.). The exact mathematical form of (3) is known for a chosen parameter mapping experiment. For example, for a variable flip angle  $T_1$ -mapping experiment, (3) can be written as

$$I_m(\mathbf{x}) = \rho(\mathbf{x}) \frac{\sin \alpha_m (1 - e^{-T_R/T_1(\mathbf{x})})}{1 - \cos \alpha_m e^{-T_R/T_1(\mathbf{x})}}, \quad (4)$$

where  $T_1(\mathbf{x})$  is the parameter map of interest,  $\alpha_m$  and  $T_R$  are pre-selected data acquisition parameters. We can, therefore, assume that  $\phi$  is a known function in (3). After discretization, it can be written as

$$\mathbf{I}_m = \boldsymbol{\Phi}_m(\boldsymbol{\theta}) \boldsymbol{\rho}, \quad (5)$$

where  $\boldsymbol{\Phi}_m \in \mathbb{C}^{N \times N}$  is a diagonal matrix with  $[\boldsymbol{\Phi}_m]_{n,n} = \phi(\boldsymbol{\theta}_n, \boldsymbol{\gamma}_m)$ ,  $\boldsymbol{\theta}_n$  denotes the parameter value at the  $n$ th voxel, and  $\boldsymbol{\rho} \in \mathbb{C}^{N \times 1}$  contains the spin density values. Note that in (5),  $\mathbf{I}_m$  linearly depends on  $\boldsymbol{\rho}$ , but nonlinearly depends on  $\boldsymbol{\theta}$ .

Substituting (5) into (2) yields the following observation model

$$\mathbf{d}_m = \mathbf{F}_m \boldsymbol{\Phi}_m(\boldsymbol{\theta}) \boldsymbol{\rho} + \mathbf{n}_m. \quad (6)$$

We can determine  $\theta$  and  $\rho$  from the measured data,  $\{\mathbf{d}_m\}$ , directly based on (6) without reconstructing  $\{I_m(\mathbf{x})\}$ . Under the assumption that  $\mathbf{n}_m$  are white Gaussian noise, the maximum likelihood (ML) estimate of  $\rho$  and  $\theta$  is given by [8–10]

$$\{\hat{\rho}_{\text{ML}}, \hat{\theta}_{\text{ML}}\} = \underset{\rho, \theta}{\operatorname{argmin}} \sum_{m=1}^M \|\mathbf{d}_m - \mathbf{F}_m \Phi_m(\theta) \rho\|_2^2. \quad (7)$$

This is a standard nonlinear least squares problem, which can be solved using a number of numerical algorithms [11], although it may have multiple local minima.

**2.1.2. Sparsity constraint**—It is well known that in parameter mapping, the values of  $\theta$  are tissue-dependent. Since the number of tissue types is relatively small compared to the number of voxels, we can apply a sparsity constraint to  $\theta$  with an appropriate sparsifying transform. Incorporating the sparsity constraint into the ML estimation framework in (7) yields:

$$\{\hat{\rho}_{\text{sML}}, \hat{\theta}_{\text{sML}}\} = \underset{\rho, \theta}{\operatorname{argmin}} \sum_{m=1}^M \|\mathbf{d}_m - \mathbf{F}_m \Phi_m(\theta) \rho\|_2^2, \text{ s.t. } \|\mathbf{W}\theta\|_0 \leq K, \quad (8)$$

where  $\mathbf{W}$  is a chosen sparsifying transform (e.g., wavelet transform),  $\|\cdot\|_0$  represents the  $\ell_0$  pseudo norm, and  $K$  is a given sparsity level. For simplicity, we assume that  $\mathbf{W}$  is an orthonormal transform in this paper. Under this assumption, we can solve the following equivalent formulation:

$$\{\hat{\rho}_{\text{sML}}, \hat{\mathbf{c}}_{\text{sML}}\} = \underset{\rho, \mathbf{c}}{\operatorname{argmin}} \sum_{m=1}^M \|\mathbf{d}_m - \mathbf{F}_m \Phi_m(\mathbf{W}^H \mathbf{c}) \rho\|_2^2, \text{ s.t. } \|\mathbf{c}\|_0 \leq K, \quad (9)$$

where  $\mathbf{c} = \mathbf{W} \theta$  contains the transform domain coefficients.

## 2.2. Optimization Algorithm

We present a practical and efficient algorithm to solve (9). Note that (9) is a nonlinear optimization problem with a smooth, non-convex cost function and an explicit sparsity constraint. A number of greedy pursuit algorithms have recently been developed to address this type of problems (e.g., [12–14]). These algorithms are mostly generalizations of greedy algorithms for compressive sensing with linear measurements. Here we adapt one efficient algorithm, the Gradient Support Pursuit (GraSP) algorithm [14], to solve (9).

GraSP is an iterative algorithm that utilizes the gradient of the cost function to identify the candidate support of the sparse coefficients, and then solves the nonlinear optimization problem constrained to the identified support. Specifically, the procedure for solving (9) can be summarized in Algorithm 1. We denote the cost function in (9) as  $\Psi(\mathbf{c}, \rho)$ , and the solution at the  $n$ th iteration as

**Algorithm 1**

The Gra SP Algorithm for Solving (9).

---

Input:  $\Psi(\mathbf{c}, \boldsymbol{\rho})$  and sparsity level  $K$   
 Initialization:  $\{\mathbf{c}^{(0)}, \boldsymbol{\rho}^{(0)}\}$   
 while “not converged” do

1. evaluate the partial gradient:  $\mathbf{z} = \frac{\partial \Psi}{\partial \mathbf{c}}|_{\{\hat{\mathbf{c}}^{(n)}, \hat{\boldsymbol{\rho}}^{(n)}\}}$
2. identify directions:  $\mathcal{Z} = \text{supp}(\mathbf{z}_{2K})$
3. merge supports:  $\mathcal{T} = \mathcal{Z} \cup \text{supp}(\hat{\mathbf{c}}^{(n)})$
4. minimize over support:  
 $\{\hat{\mathbf{c}}, \hat{\boldsymbol{\rho}}^{(n+1)}\} = \text{argmin} \Psi(\mathbf{c}, \boldsymbol{\rho})$  s.t.  $\mathbf{c}|_{\mathcal{T}^c} = 0$
5. prune estimate:  $\hat{\mathbf{c}}^{(n+1)} = \hat{\mathbf{c}}_K$

---

$\{\hat{\mathbf{c}}^{(n)}, \hat{\boldsymbol{\rho}}^{(n)}\}$ . At the  $(n+1)$ th iteration, we first compute the partial derivative of  $\Psi$  with respect to  $\mathbf{c}$  at  $\{\hat{\mathbf{c}}^{(n)}, \hat{\boldsymbol{\rho}}^{(n)}\}$ :

$$\mathbf{z} = \frac{\partial \Psi}{\partial \mathbf{c}} = \mathbf{W} \sum_{m=1}^M \frac{\partial \Phi_m}{\partial \boldsymbol{\theta}} \text{Re}\{\mathbf{F}_m^H [\mathbf{F}_m \Phi_m(\boldsymbol{\theta}) \boldsymbol{\rho} - \mathbf{d}_m]\}, \quad (10)$$

where  $\frac{\partial \Phi_m}{\partial \boldsymbol{\theta}}$  is a diagonal matrix with  $\left[ \frac{\partial \Phi_m}{\partial \boldsymbol{\theta}} \right]_{n,n} = \frac{\partial \phi(\boldsymbol{\theta}_n, \boldsymbol{\gamma}_m)}{\partial \boldsymbol{\theta}_n}$ . Secondly, we identify a support set  $\mathcal{Z}$  associated with the  $2K$  largest entries of  $\mathbf{z}$ , i.e.,  $\mathcal{Z} = \text{supp}(\mathbf{z}_{2K})$  (Intuitively, minimization over  $\mathcal{Z}$  would lead to the most effective reduction in the cost function value). Then we merge  $\mathcal{Z}$  with  $\text{supp}(\hat{\mathbf{c}}^{(n)})$  to form a combined support set  $\mathcal{T}$ , over which we minimize  $\Psi$ . Finally, after obtaining the solution  $\tilde{\mathbf{c}}$ , we only keep its largest  $K$  entries and set other entries to zero, i.e.,  $\hat{\mathbf{c}}^{(n+1)} = \tilde{\mathbf{c}}_K$ . The above procedure is repeated until the relative change between two consecutive iterations is smaller than some pre-specified threshold.

The algorithm is computationally efficient: at each iteration it only involves gradient evaluation and the solution of a support constrained problem. For the support-constrained optimization problem, its computational complexity is much smaller than (7) due to the reduced number of unknowns. It has been shown in [14] that under certain theoretical conditions (the generalization of restricted isometry property in linear measurement model), GraSP has guaranteed convergence. However, note that verifying these theoretical conditions for a specific problem is typically NP-hard. In solving (9), we estimate the initial solution  $\{\mathbf{c}^{(0)}, \boldsymbol{\rho}^{(0)}\}$  from a low-resolution image sequence, which have consistently yielded good empirical results.

**2.3. Estimation-Theoretic Characterization**

The formulation (8) or (9) performs parameter estimation with sparsity constraint. In this subsection, we derive the sparsity-constrained Cramer-Rao lower bound (CRLB) [15,16] on  $\boldsymbol{\theta}$ , which provides useful insights into the benefits of incorporating the sparsity constraint into the parameter estimation problem. Furthermore, it can also be used as a benchmark to evaluate our solution algorithm.

We first derive the sparsity-constrained CRLB on  $\mathbf{c}$ , from which we can obtain a bound on  $\boldsymbol{\theta}$ . Assume that  $\mathbf{c}$  is  $K$ -sparse, i.e.,  $\mathbf{c} \in \Omega = \{\mathbf{c} \in \mathbb{R}^N \mid \text{supp}(\mathbf{c}) = K\}$  and  $K$  is given. Based on (6)

and  $\mathbf{c} = \mathbf{W}\boldsymbol{\theta}$ , the sparsity constrained CRLB for any locally unbiased estimator  $\hat{\mathbf{c}}$  can be expressed as

$$\text{Cov}(\hat{\mathbf{c}}) \geq \mathbf{Z}_s, \quad (11)$$

where

$$\mathbf{Z}_s = [\mathbf{E}_N \mathbf{0}_N] \mathbf{A} (\mathbf{A}^T \mathbf{J} \mathbf{A})^\dagger \mathbf{A}^T \begin{bmatrix} \mathbf{E}_N \\ \mathbf{0}_N \end{bmatrix} \in \mathbb{R}^{N \times N}, \quad (12)$$

$\mathbf{J} \in \mathbb{R}^{2N \times 2N}$  is the Fisher information matrix (FIM),  $\mathbf{A} = \begin{bmatrix} \tilde{\mathbf{E}}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_N \end{bmatrix}$ ,  $\mathbf{E}_N$  is the  $N \times N$  identity matrix,  $\tilde{\mathbf{E}}_K$  is an  $N \times K$  sub-matrix of  $\mathbf{E}_N$  whose columns are selected based on the support of  $\mathbf{c}$ . We can simplify the expression of  $\mathbf{Z}_s$  in (12). Let the partitioned FIM be

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{c,c} & \mathbf{J}_{c,\rho} \\ \mathbf{J}_{\rho,c} & \mathbf{J}_{\rho,\rho} \end{bmatrix} \text{ and } \mathbf{A}^T \mathbf{J} \mathbf{A} = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix}, \text{ where}$$

$\mathbf{G}_{11} = \tilde{\mathbf{E}}_K^T \mathbf{J}_{cc} \tilde{\mathbf{E}}_K$ ,  $\mathbf{G}_{12} = \tilde{\mathbf{E}}_K^T \mathbf{J}_{c\rho}$ ,  $\mathbf{G}_{21} = \mathbf{G}_{12}^T$ , and  $\mathbf{G}_{22} = \mathbf{J}_{\rho\rho}$ . Using the pseudo-inverse of the partitioned Hermitian matrix [17], it can be shown that<sup>1</sup>

$$\mathbf{Z}_s = \tilde{\mathbf{E}}_K (\mathbf{G}_{11}^\dagger + \mathbf{G}_{11}^\dagger \mathbf{G}_{12} \mathbf{Q}_s^\dagger \mathbf{G}_{21} \mathbf{G}_{11}^\dagger) \tilde{\mathbf{E}}_K^T, \quad (13)$$

where  $\mathbf{Q}_s = \mathbf{G}_{22} - \mathbf{G}_{21} \mathbf{G}_{11}^\dagger \mathbf{G}_{12}$ .

With  $\hat{\boldsymbol{\theta}} = \mathbf{W}^H \hat{\mathbf{c}}$ , the bound on  $\hat{\boldsymbol{\theta}}$  can be written as

$$\text{Cov}(\hat{\boldsymbol{\theta}}) \geq \mathbf{W}^H \mathbf{Z}_s \mathbf{W}. \quad (14)$$

Taking the diagonal entries of the covariance matrix, we can obtain the bound on the variance of  $\hat{\boldsymbol{\theta}}$  at each voxel as

$$\text{Var}(\hat{\theta}_n) \geq [\mathbf{W}^H \mathbf{Z}_s \mathbf{W}]_{n,n}. \quad (15)$$

Under the assumption that  $\hat{\boldsymbol{\theta}}_{\text{sML}}$  is locally unbiased in (8), the righthand side of (14) also provides a lower bound on  $\text{Cov}(\hat{\boldsymbol{\theta}}_{\text{sML}})$ . Note that this assumption can approximately hold with a finite number of measurements and a reasonable signal-to-noise ratio.

Excluding the sparsity constraint, we can calculate the unconstrained CRLB as follows:

$$\text{Cov}(\hat{\boldsymbol{\theta}}) \geq \mathbf{W}^H \mathbf{Z} \mathbf{W}, \quad (16)$$

where  $\mathbf{Z} = (\mathbf{J}_{c,c}^\dagger + \mathbf{J}_{c,c}^\dagger \mathbf{J}_{c,\rho} \mathbf{Q}^\dagger \mathbf{J}_{\rho,c} \mathbf{J}_{c,c}^\dagger)$  and  $\mathbf{Q} = \mathbf{J}_{\rho,\rho} - \mathbf{J}_{\rho,c} \mathbf{J}_{c,c}^\dagger \mathbf{J}_{c,\rho}$ . It has been shown in [15, 16] that  $\mathbf{W}^H \mathbf{Z} \mathbf{W} \geq \mathbf{W}^H \mathbf{Z}_s \mathbf{W}$ . Therefore, incorporating sparsity constraint can be theoretically demonstrated to be beneficial. In the next section, we will calculate the above bounds in a numerical brain phantom to illustrate the potential benefits of using the sparsity constraint in (8).

<sup>1</sup>The formula here has already taken into account the case that the FIM is singular. This happens when the null signal intensity appear in the background.

### 3. RESULTS

In this section, we use a  $T_2$  mapping example with a numerical brain phantom [18] to illustrate the estimation-theoretic bounds and empirical performance of the proposed method. In this example,  $\varphi(\theta, \gamma_m) = \exp(-\gamma_m R_2)$ , where  $R_2$  is the relaxation rate (i.e., the reciprocal of  $T_2$ ), and  $\gamma_m = T_{E,m}$  is the echo time. The  $R_2$  map of the brain phantom is illustrated in Fig. 1. We used a multi-echo spin echo acquisition with a total number of 16 echoes and 10 ms echo spacing. We define the acceleration factor (AF) as

$NM/(\sum_{m=1}^M P_m)$ , and the signal to noise ratio (SNR) as the ratio of the signal intensity (in a region of the white matter) to the noise standard deviation.

#### 3.1. Performance Bounds

We compute the estimation-theoretic bounds in (14) and (16) to analyze the potential benefits of incorporating the sparsity constraint. Specifically, we take the diagonal entries of the covariance matrix to extract the bounds on the variances for each voxel. We simulated a data set that has 1) a mono-exponential signal model, and 2) a  $R_2$  map that is sparse in the wavelet domain<sup>2</sup>. Data acquisition was performed at  $AF = 4$  and  $SNR = 28\text{dB}$ . To reduce the memory burden, we resized the original brain phantom to a smaller scale.

In addition to evaluating the performance bounds, we also performed a Monte Carlo study (with 300 trials) to compute the empirical MSE of the ML estimator in (7) and the proposed estimator in (9), respectively. The true sparsity level was used to set the parameter  $K$  for (9). Furthermore, we considered an oracle ML estimator that assumes complete knowledge of the exact sparse support of the  $R_2$  map. The MSE of the proposed estimator in (8) should be bounded below by that of the oracle estimator.

The performance bound and empirical MSE<sup>3</sup> are shown in Fig.2. As can be seen, the CRLB varies with respect to different brain tissues. With the wavelet domain sparsity constraint, the constrained CRLB is significantly lower than the CRLB and it becomes much less tissue-dependent. The theoretical prediction matches well with the simulation results. The proposed method yields much lower MSE than the ML estimator, demonstrating again the benefits of introducing the sparsity constraint for improving parameter estimation. Note that the performance of the proposed method is very close to that of the oracle estimator, which matches well with the constrained CRLB.

#### 3.2. Empirical performance

To demonstrate the empirical performance of the proposed method, we simulated the phantom in a way that both signal model mismatch (multi-exponential model caused by partial volume effect) and sparsity model mismatch (wavelet coefficients are approximately sparse) exist. We acquired two sets of  $\mathbf{k}$ -space measurements at  $AF = 2.67$  and 4, both of which had  $SNR = 24\text{dB}$ . Here, we manually chose the sparsity level  $K = 0.2N$  for the proposed method. Note that selecting  $K$  in a more principled way is worth of further study.

We compared the proposed method with a dictionary learning-based compressed sensing reconstruction [3] (referred to as CS), which only takes into account the temporal relaxation process. The reconstructed  $R_2$  maps are shown in Fig.3, along with the normalized root-mean-square-error (NRMSE) listed below the reconstructions. As can be seen, when  $AF = 4$ , the CS reconstruction shows several artifacts (marked by arrows), although these artifacts

<sup>2</sup>We kept 20% of the largest wavelet coefficients (of the Haar wavelet transform) of the original  $R_2$  map.

<sup>3</sup>For the three estimators, we observed empirically that the bias is much smaller than the variance so that the MSE is dominated by the variance.

were significantly reduced at the lower AF. In contrast, the proposed method produced higher-quality parameter maps at both high and low acceleration levels. The observations are consistent with the values of NRMSE.

## 4. CONCLUSION

This paper presented a new method to directly reconstruct parameter maps from highly undersampled, noisy  $\mathbf{k}$ -space data, utilizing an explicit signal model while imposing a sparsity constraint on the parameter values. A greedy pursuit algorithm was described to solve the underlying optimization problem. The benefit of incorporating sparsity constraint is analyzed theoretically using estimation-theoretic bounds and also illustrated empirically in a  $T_2$  mapping example. The proposed method should prove useful for fast MR parameter mapping with sparse sampling.

## Acknowledgments

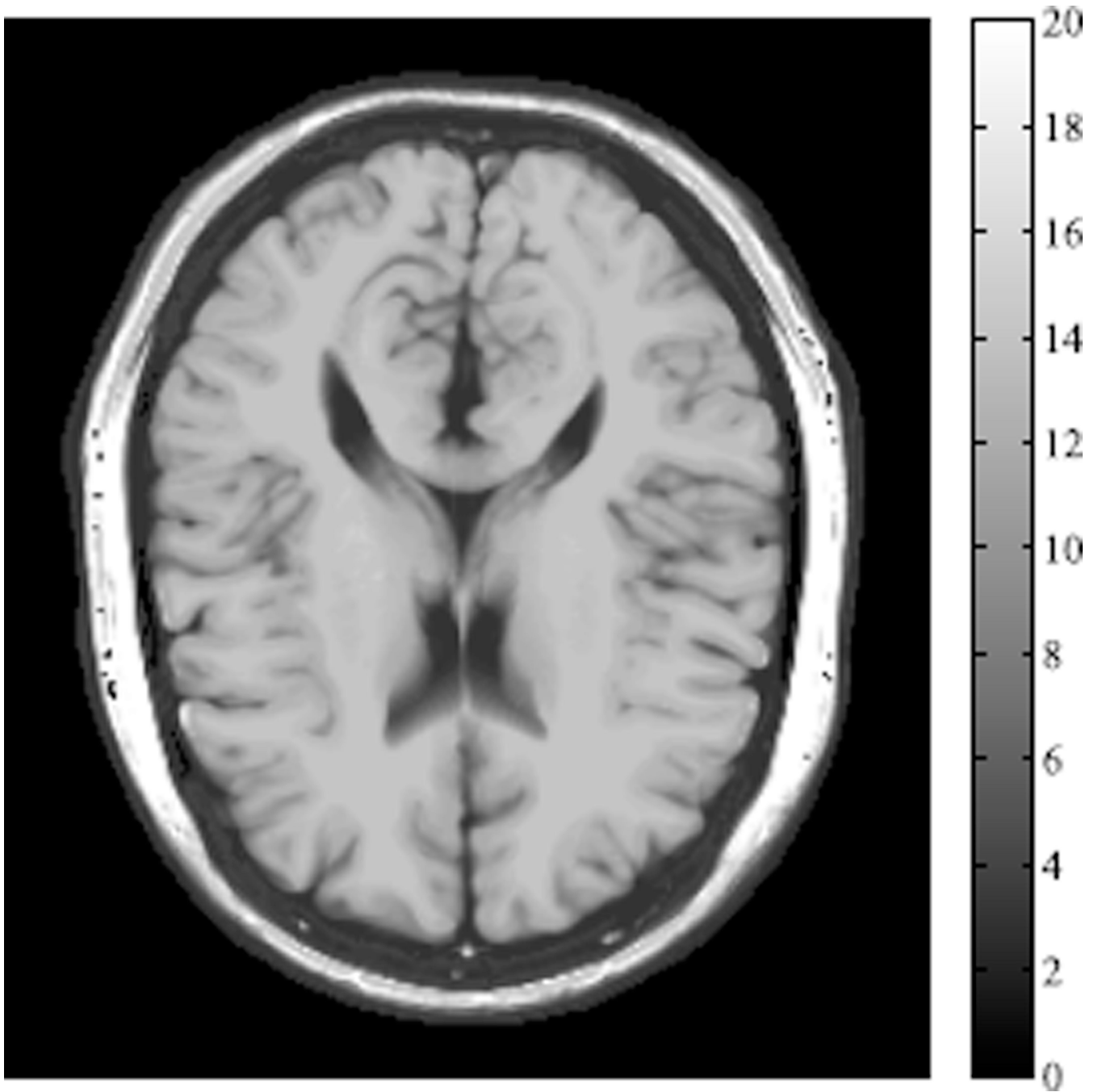
The work presented in this paper was supported in part by NIH-P41-EB015904, NIH-P41-EB001977 and NIH-1RO1-EB013695.

## REFERENCES

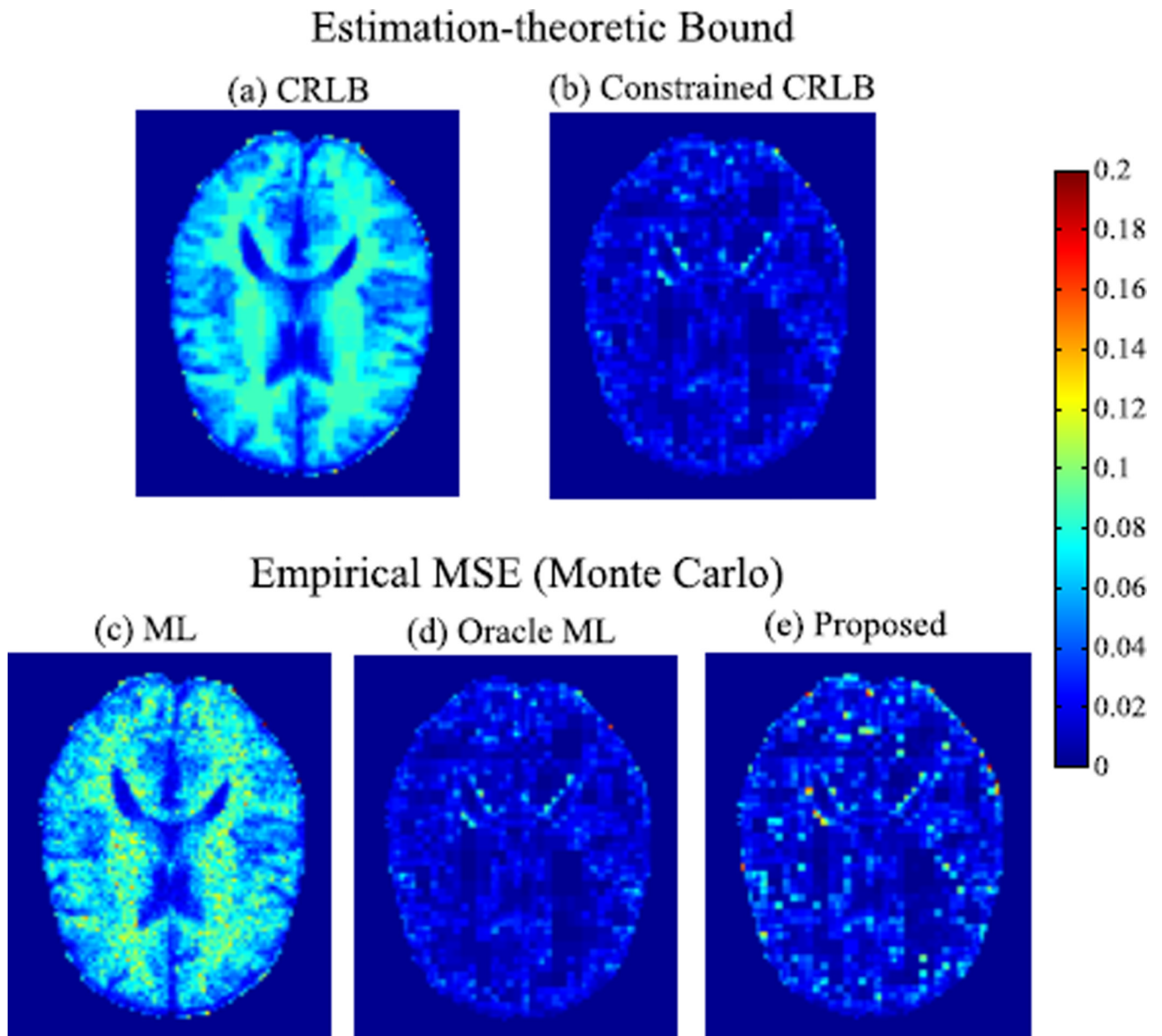
1. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* 2007; 58:1182–1195. [PubMed: 17969013]
2. Liang Z-P. Spatiotemporal imaging with partially separable functions. *Proc. IEEE Int. Symp. Biomed. Imaging.* 2007:988–991.
3. Doneva M, Bornert P, Eggers H, Stehning C, Senegas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magn. Reson. Med.* 2010; 64:1114–1120. [PubMed: 20564599]
4. Feng L, Otazo R, Jung H, Jensen JH, Ye JC, Sodickson DK, Kim D. Accelerated cardiac  $T_2$  mapping using breath-hold multiecho fast spin-echo pulse sequence with k-t FOCUSS. *Magn. Reson. Med.* 2011; 65(6):1661–1669. [PubMed: 21360737]
5. Petzschner FH, Ponce IP, Blaimer M, Jakob PM, Breuer FA. Fast MR parameter mapping using k-t principal component analysis. *Magn. Reson. Med.* 2011; 66:706–716. [PubMed: 21394772]
6. Huang C, Graff CG, Clarkson EW, Bilgin A, Altbach MI.  $T_2$  mapping from highly undersampled data by reconstruction of principal component coefficient maps using compressed sensing. *Magn. Reson. Med.* 2012; 67(5):1355–1366. [PubMed: 22190358]
7. Zhao B, Lu W, Liang Z-P. Highly accelerated parameter mapping with joint partial separability and sparsity constraints. *Proc. Int. Symp. Magn. Reson. Med.* 2012:2233.
8. Haldar JP, Hernando D, Liang Z-P. Super-resolution reconstruction of MR image sequences with contrast modeling. *Proc. IEEE Int. Symp. Biomed. Imaging.* 2009:266–269.
9. Block K, Uecker M, Frahm J. Model-based iterative reconstruction for radial fast spin-echo MRI. *IEEE Trans. Med Imaging.* 2009; 28:1759–1769. [PubMed: 19502124]
10. Sumpf TJ, Uecker M, Boretius S, Frahm J. Model-based nonlinear inverse reconstruction for  $T_2$  mapping using highly undersampled spin-echo MRI. *J Magn Reson Imaging.* 2011; 34:420–428. [PubMed: 21780234]
11. Nocedal, J.; Wright, SJ. *Numerical Optimization.* (2nd Edition). New York: NY: Springer; 2006.
12. Beck A, Eldar YC. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. 2012 preprint.
13. Blumensath T. Compressed sensing with nonlinear observations and related nonlinear optimization problems. 2012 preprint.
14. Bahmani S, Raj B, Boufounos P. Greedy sparsity-constrained optimization. 2012 preprint.
15. Gorman JD, Hero AO. Lower bounds for parametric estimation with constraints. *IEEE Trans. Inf. Theory.* 1990; (6):1285–1301.

16. Ben-Haim Z, Eldar YC. The Cramer-Rao bound for estimating a sparse parameter vector. *IEEE Trans. Signal Process.* 2010; (6):3384–3389.
17. Rohde CA. Generalized inverses of partitioned matrices. *J. Soc. Indust. Appl. Math.* 1965; 13:1033–1035.
18. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med Imaging.* 1998; 17:463–468. [PubMed: 9735909]

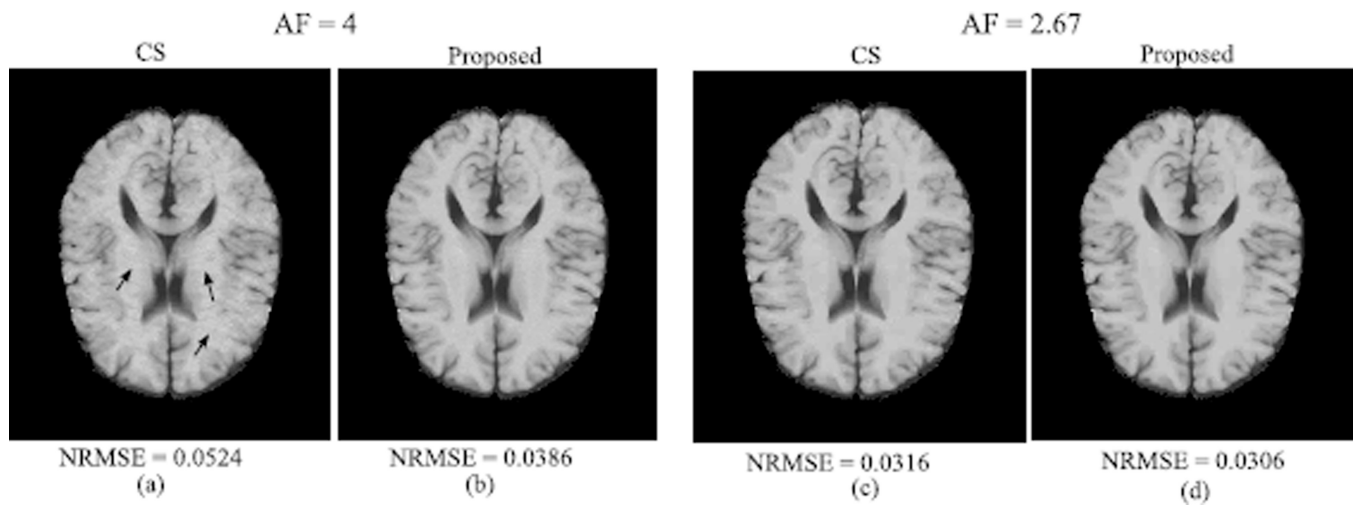




**Fig. 1.**  
The  $R_2$  map of the numerical brain phantom.



**Fig. 2.** Estimation-theoretic bounds and empirical MSE of the ML, oracle ML, and proposed estimators. Note that the background, skull, and the scalp are not region of interest for our study, and thus they were set to zero.



**Fig. 3.** (a)–(b) Reconstructed  $R_2$  maps at AF = 4; (c)–(d) Reconstructed  $R_2$  maps at AF = 2.67.