

Accelerated Regularized Estimation of MR Coil Sensitivities Using Augmented Lagrangian Methods

Michael J. Allison*, *Student Member, IEEE*, Sathish Ramani, *Member, IEEE*, and Jeffrey A. Fessler, *Fellow, IEEE*

Abstract—Several magnetic resonance (MR) parallel imaging techniques require explicit estimates of the receive coil sensitivity profiles. These estimates must be accurate over both the object and its surrounding regions to avoid generating artifacts in the reconstructed images. Regularized estimation methods that involve minimizing a cost function containing both a data-fit term and a regularization term provide robust sensitivity estimates. However, these methods can be computationally expensive when dealing with large problems. In this paper, we propose an iterative algorithm based on variable splitting and the augmented Lagrangian method that estimates the coil sensitivity profile by minimizing a quadratic cost function. Our method, ADMM-Circ, reformulates the finite differencing matrix in the regularization term to enable exact alternating minimization steps. We also present a faster variant of this algorithm using intermediate updating of the associated Lagrange multipliers. Numerical experiments with simulated and real data sets indicate that our proposed method converges approximately twice as fast as the preconditioned conjugate gradient method (PCG) over the entire field-of-view. These concepts may accelerate other quadratic optimization problems.

Index Terms—Augmented Lagrangian, coil sensitivity, finite differences, parallel imaging, quadratic minimization.

I. INTRODUCTION

ACCURATE radio-frequency coil sensitivity profiles are required in many parallel imaging applications (e.g., sensitivity encoding (SENSE) [1], simultaneous acquisition of spatial harmonics (SMASH) [2], and k - t SENSE [3]). Due to coil deformation during patient setup and dielectric coupling, these profiles must be determined at the time of acquisition [4]. One common approach is to perform a calibration scan prior to the parallel imaging acquisition in which images from a large body coil and multiple surface coils are acquired and reconstructed. Since the body coil has near uniform sensitivity, its image can be used in conjunction with a surface coil image to estimate the surface coil sensitivity profile.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by the National Institutes of Health (NIH) under Grant P01 CA87634. *Asterisk indicates corresponding author.*

M. J. Allison*, S. Ramani, and J. A. Fessler are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: mjalliso@umich.edu, sramani@umich.edu, fessler@umich.edu).

This paper has supplementary downloadable materials available at <http://ieeexplore.ieee.org>, provided by the author.

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The most straightforward method to estimate the coil sensitivity is to compute the ratio of the surface coil image voxel values (z_i) to the body coil image voxel values (y_i), z_i/y_i . However, ratio estimates can be corrupted by measurement noise, particularly in low signal regions. Furthermore, such estimates can have sharp discontinuities at object edges, contrary to the smooth nature of true coil sensitivity profiles [5]. It is also desirable to have reasonable sensitivity estimates in any low signal regions surrounding the object to avoid reconstruction artifacts that could arise due to patient motion [6]. The ratio estimator, however, does not extrapolate; thus, improved estimation methods can be beneficial.

One approach to generate smooth sensitivity estimates is to measure only the center of k -space [6]. Although simple, this approach does not accurately estimate sensitivities near object edges and can introduce Gibbs ringing artifacts. Filtering procedures have also been proposed including polynomial fitting [1], [7]–[9], wavelet denoising [10], and using thin-plate splines [11]. These methods do not completely eliminate the Gibbs ringing, while selecting a particular basis function is complicated by the varying size of low signal regions within the images [5], [12]. Furthermore, many of these methods disregard the non-stationary variance of the noise in the ratio estimates. In contrast, regularized estimation methods [5], [13], [14] provide smooth sensitivity estimates and are capable of extrapolation without explicit basis function selection or filtering. These methods, however, can be computationally expensive for large problems [5] and this cost is compounded by the large number of coils in some arrays [15]. Although sensitivity estimation can be performed off-line, the computational costs of regularized methods can increase the overall compute times of parallel imaging.

In this paper, we take a regularized approach and pose sensitivity estimation as the minimization of a quadratic cost function like in [5]. The large matrices in the cost function prevent one from computing a simple, non-iterative solution. Instead, iterative methods must be used for large data sets; however, traditional methods like conjugate gradient (CG) converge slowly for this problem [5], [16]. Augmented Lagrangian (AL) based minimization techniques [17], and the related Bregman iterations method [18], have been used to accelerate convergence in imaging problems such as denoising [19] and reconstruction [19]–[27]. Those papers primarily focus on problems that contain non-differentiable regularization terms such as those based on the ℓ_1 -norm. However, the underlying theory applies to a wide variety of optimization problems,

including the quadratic problem considered here. We therefore propose a new AL based method for estimating sensitivity profiles. To derive this method, we introduce a reformulation of the finite differencing matrix and a subsequent variable splitting that lead to an algorithm with exact alternating minimization steps. This algorithm is equivalent to an alternating direction method of multipliers (ADMM) [28] formulation, which provides a guarantee of convergence. We also explore a variation of this algorithm that updates the Lagrange multipliers between alternating minimization steps. Such variations have been found to improve the convergence rates of other AL based algorithms [29].

Section II presents the derivation of our ADMM algorithm and its intermediate updating variant. Section III compares the convergence speeds of these algorithms with those of CG based methods by performing experiments on both simulated and real data. Section IV discusses the results of these experiments and additional properties of the algorithms. Section V concludes by discussing other problems that have quadratic cost functions where our methods may provide an improvement over the traditional techniques.

II. MATERIALS AND METHODS

This section introduces our proposed methods for MR coil sensitivity estimation. We begin by posing the estimator as an optimization problem. We then outline the general approach used to solve this problem and present our specific algorithm, with variations, in detail.

A. Cost Function Formulation

Regularized methods for MR coil sensitivity estimation are both robust to noise and effective at extrapolating the estimate in regions of low signal [5], [14]. These methods avoid computing the quotient (z_i/y_i) by finding the minimizer of a cost function containing a data-fidelity term and a regularization term that promotes smoothness in the estimate. Similar to [5], [13], we estimate the sensitivity profile by minimizing a weighted sum of quadratic terms:

$$\hat{\mathbf{s}} \triangleq \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{z} - \mathbf{D}\mathbf{s}\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \|\mathbf{R}\mathbf{s}\|_2^2, \quad (1)$$

where $\mathbf{s} = [s_1, \dots, s_N]^T$ with $s_i \in \mathbb{C}$ denoting the desired coil sensitivity at the i th voxel and N denoting the number of voxels, $\mathbf{z} = [z_1, \dots, z_N]^T$ with $z_i \in \mathbb{C}$ denoting the surface coil image value at the i th voxel, $\mathbf{D} = \text{diag}\{y_i\}$ is a diagonal matrix containing the body coil image voxel values ($y_i \in \mathbb{C}$), $\mathbf{R} \in \mathbb{R}^{M \times N}$ is a finite differencing matrix for the case of non-periodic boundary conditions with M sets of finite differences, and $\lambda > 0$ is a regularization coefficient. Additionally, $\mathbf{W} = \text{diag}\{w_i\}$ is a diagonal weighting matrix (with $w_i \in [0, 1]$) that allows us to ensure that the estimate is based primarily on voxels that provide meaningful sensitivity information. Note that a finite differencing matrix with non-periodic boundary conditions is necessary as periodic boundary conditions introduce errors at the edges of the image that can propagate and corrupt the estimate near the object voxels.¹

¹See the supplementary material for additional details.

Equation (1) has a quadratic cost function and therefore has the closed-form solution $\hat{\mathbf{s}} = [\mathbf{D}^H \mathbf{W} \mathbf{D} + \lambda \mathbf{R}^H \mathbf{R}]^{-1} \mathbf{D}^H \mathbf{W} \mathbf{z}$ where \mathbf{X}^H denotes the Hermitian transpose of \mathbf{X} ; however, computing this solution is impractical due to the size and complexity of \mathbf{R} . Memory constraints further restrict the use of other direct methods, such as Cholesky factorization, for large problems like 3D data sets. Furthermore, standard iterative solution methods, such as CG, exhibit slow convergence for this problem even when using carefully selected preconditioners [30]. To address this, we propose an augmented Lagrangian method to minimize the cost function, the development of which consists of three stages [20]. First, we use variable splitting [25], [31] to convert the unconstrained optimization problem into an equivalent constrained problem, thereby decoupling the effects of the matrices in (1). Second, we introduce vector Lagrange multipliers and express the constrained problem in an AL framework. Third, we solve the resulting AL problem using an alternating minimization scheme.

B. ADMM–Circ: ADMM Sensitivity Estimation Algorithm with Circulant Substeps

Directly applying variable splitting to (1) results in an AL algorithm requiring an approximate solution for one of the alternating minimization steps due to the complexity of the finite differencing matrix \mathbf{R} [16]. The supplementary material for this paper presents one such algorithm, ADMM–CG. We can avoid this complication if we focus on traditional finite differencing matrices (those with spatially invariant stencils). For such regularizers, we can express the finite differencing matrix as $\mathbf{R} = \mathbf{B}\mathbf{C}$ where $\mathbf{C} \in \mathbb{R}^{M \times N}$ is a typical finite differencing matrix for the case of periodic boundary conditions, containing additional non-zero rows that penalize the differences between voxels on opposing boundaries of the image, and $\mathbf{B} \in \{0, 1\}^{M \times M}$ is a diagonal matrix that contains a mask to eliminate the effects of the added rows. The additional non-zero rows in \mathbf{C} ensure that $\mathbf{C}^H \mathbf{C}$ is block circulant with circulant blocks unlike \mathbf{R} . Fig. 1 illustrates these matrices for the case of 1D second-order finite differences. We then write the estimation problem in (1) as

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{z} - \mathbf{D}\mathbf{s}\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \|\mathbf{B}\mathbf{C}\mathbf{s}\|_2^2. \quad (2)$$

We introduce two splitting variables, $\mathbf{u}_0 \in \mathbb{C}^M$ and $\mathbf{u}_1 \in \mathbb{C}^N$, to this new formulation to decouple the matrices \mathbf{D} , \mathbf{B} , and \mathbf{C} . The resulting equivalent constrained optimization problem is

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}, \mathbf{u}_0, \mathbf{u}_1} \frac{1}{2} \|\mathbf{z} - \mathbf{D}\mathbf{u}_1\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \|\mathbf{B}\mathbf{u}_0\|_2^2 \quad (3)$$

s.t. $\mathbf{u}_1 = \mathbf{s}$ and $\mathbf{u}_0 = \mathbf{C}\mathbf{s}$.

Solving this constrained optimization problem is exactly equivalent to solving the unconstrained problem (1).

We express (3) in the more concise notation:

$$\arg \min_{\mathbf{s}, \mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{A}\mathbf{u}\|_2^2 \quad \text{s.t. } \mathbf{u} = \mathbf{T}\mathbf{s}, \quad (4)$$

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & & 0 \\ 0 & -1 & 2 & -1 & & 0 \\ \vdots & & & & \ddots & \\ 0 & 0 & 0 & 0 & & 0 \end{bmatrix}}_{\mathbf{R}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ \vdots & & & & \ddots & \\ 0 & 0 & 0 & 0 & & 0 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} 2 & -1 & 0 & 0 & \dots & -1 \\ -1 & 2 & -1 & 0 & & 0 \\ 0 & -1 & 2 & -1 & & 0 \\ \vdots & & & & \ddots & \\ -1 & 0 & 0 & 0 & & 2 \end{bmatrix}}_{\mathbf{C}}$$

Fig. 1. The matrices \mathbf{R} , \mathbf{B} , and \mathbf{C} for the case of 1D second-order finite differences. The top and bottom rows of \mathbf{C} compute the difference between the first and last pixels, hence the need for the mask \mathbf{B} .

where

$$\mathbf{u} \triangleq \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_0 \end{bmatrix}, \mathbf{T} \triangleq \begin{bmatrix} \mathbf{I} \\ \mathbf{C} \end{bmatrix}, \mathbf{h} \triangleq \begin{bmatrix} \mathbf{W}^{\frac{1}{2}} \mathbf{z} \\ \mathbf{0} \end{bmatrix}, \mathbf{A} \triangleq \begin{bmatrix} \mathbf{W}^{\frac{1}{2}} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \sqrt{\lambda} \mathbf{B} \end{bmatrix},$$

and $\mathbf{W}^{\frac{1}{2}} \triangleq \text{diag}\{\sqrt{w_i}\}$.

We then introduce two vectors of Lagrange multipliers, $\boldsymbol{\eta}_0 \in \mathbb{C}^M$ and $\boldsymbol{\eta}_1 \in \mathbb{C}^N$, and express (4) as an AL problem. We use the general AL formulation outlined in [20] that incorporates the Lagrange multiplier into the quadratic penalty term. This formulation is a natural extension of the traditional AL to the case of complex values and it simplifies the derivation of the subsequent alternating minimization steps. The resulting AL function-based minimization problem is

$$\arg \min_{\mathbf{s}, \mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{A}\mathbf{u}\|_2^2 + \frac{1}{2} \|\mathbf{u} - \mathbf{T}\mathbf{s} - \boldsymbol{\eta}\|_{\mathbf{V}}^2, \quad (5)$$

where

$$\boldsymbol{\eta} \triangleq \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_0 \end{bmatrix}, \mathbf{V} \triangleq \begin{bmatrix} \nu_1 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \nu_0 \mathbf{I} \end{bmatrix},$$

and $\nu_0, \nu_1 > 0$ are AL penalty parameters that influence the convergence rate of the algorithm but do not affect the final estimate [26].

Traditional AL methods would require jointly minimizing (5) over the vectors \mathbf{s} and \mathbf{u} ; however, such an approach is computationally expensive for typical image sizes. Instead, we use a block Gauss–Seidel type alternating minimization strategy that has been effective in solving other AL problems [19], [28] in which we alternate between minimizing (5) independently with respect to \mathbf{s} and \mathbf{u} as follows:

$$\mathbf{s}^{(j+1)} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{u}^{(j)} - \mathbf{T}\mathbf{s} - \boldsymbol{\eta}^{(j)}\|_{\mathbf{V}}^2, \quad (6)$$

$$\mathbf{u}^{(j+1)} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{A}\mathbf{u}\|_2^2 + \frac{1}{2} \|\mathbf{u} - \mathbf{T}\mathbf{s}^{(j+1)} - \boldsymbol{\eta}^{(j)}\|_{\mathbf{V}}^2. \quad (7)$$

Update (7) has a simple closed-form solution:

$$\mathbf{u}^{(j+1)} = (\mathbf{A}^H \mathbf{A} + \mathbf{V})^{-1} \left[\mathbf{A}^H \mathbf{h} + \mathbf{V}(\mathbf{T}\mathbf{s}^{(j+1)} + \boldsymbol{\eta}^{(j)}) \right]. \quad (8)$$

In fact, the block diagonal structures of \mathbf{A} and \mathbf{V} decouple the update of \mathbf{u} into two parallel updates in terms of \mathbf{u}_1 and \mathbf{u}_0 :

$$\mathbf{u}_1^{(j+1)} = \mathbf{D}_2^{-1} \left[\mathbf{D}^H \mathbf{W} \mathbf{z} + \nu_1 (\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_1^{(j)}) \right], \quad (9)$$

$$\mathbf{u}_0^{(j+1)} = \mathbf{B}_2^{-1} (\mathbf{C}\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_0^{(j)}), \quad (10)$$

where $\mathbf{B}_2 \triangleq \frac{\lambda}{\nu_0} \mathbf{B}^H \mathbf{B} + \mathbf{I}$ and $\mathbf{D}_2 \triangleq \mathbf{D}^H \mathbf{W} \mathbf{D} + \nu_1 \mathbf{I}$ are both diagonal matrices that are trivial to invert. The closed-form update for \mathbf{s} may at first appear more complicated to compute:

$$\begin{aligned} \mathbf{s}^{(j+1)} &= (\mathbf{T}^H \mathbf{V} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{V} \left(\mathbf{u}^{(j)} - \boldsymbol{\eta}^{(j)} \right) \\ &= (\nu_1 \mathbf{I} + \nu_0 \mathbf{C}^H \mathbf{C})^{-1} \\ &\quad \left[\nu_0 \mathbf{C}^H (\mathbf{u}_0^{(j)} - \boldsymbol{\eta}_0^{(j)}) + \nu_1 (\mathbf{u}_1^{(j)} - \boldsymbol{\eta}_1^{(j)}) \right]. \end{aligned} \quad (11)$$

However, since $\mathbf{C}^H \mathbf{C}$ is block circulant with circulant blocks, $\mathbf{C}^H \mathbf{C} = \mathbf{Q}^H \boldsymbol{\Phi} \mathbf{Q}$ where \mathbf{Q} is a (multidimensional) discrete Fourier transform (DFT) matrix and $\boldsymbol{\Phi}$ is a diagonal matrix containing the spectrum of the convolution kernel of $\mathbf{C}^H \mathbf{C}$. Substituting this decomposition into (11) yields:

$$\begin{aligned} \mathbf{s}^{(j+1)} &= \mathbf{Q}^H \boldsymbol{\Phi}_2^{-1} \mathbf{Q} \\ &\quad \left[\nu_0 \mathbf{C}^H (\mathbf{u}_0^{(j)} - \boldsymbol{\eta}_0^{(j)}) + \nu_1 (\mathbf{u}_1^{(j)} - \boldsymbol{\eta}_1^{(j)}) \right], \end{aligned} \quad (12)$$

where $\boldsymbol{\Phi}_2 \triangleq \nu_1 \mathbf{I} + \nu_0 \boldsymbol{\Phi}$. This formulation is simpler to compute since $\boldsymbol{\Phi}_2$ is a diagonal matrix and we implement \mathbf{Q} efficiently using fast Fourier transforms (FFTs).

Fig. 2 summarizes the resulting sensitivity profile estimation algorithm, ADMM–Circ. Each stage of the proposed algorithm consists of an exact, non-iterative update. Furthermore, it can be shown that the steps in this formulation are identical to those of an ADMM algorithm applied to the real valued case where we treat the complex valued terms as a stack of their real and imaginary components. As discussed in Section IV, this equivalence allows us to conclude that the ADMM–Circ algorithm converges to the solution of (1). In contrast, the parallel imaging reconstruction algorithm in [20] is an AL method that lacks a convergence proof due to the type of splitting used.

C. Alternating Minimization with Intermediate Updating

Updating the Lagrange multipliers $\boldsymbol{\eta}$ between each alternating minimization step has been shown to increase the convergence rates of several AL based algorithms [29]. We also explore this variation in our proposed algorithm by updating the relevant Lagrange multipliers after each alternating minimization step, Fig. 3. The resulting algorithm, ADMM–Circ–IU, requires no additional variables and the added updates (Step 2) are computationally inexpensive. Section IV describes the convergence properties of such adaptations.

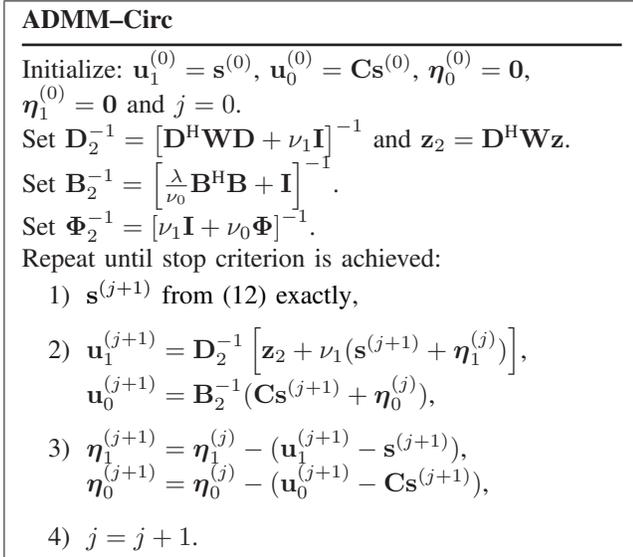


Fig. 2. Overview of the ADMM-Circ algorithm. Note that $\mathbf{C}\mathbf{s}^{(j+1)}$ only needs to be computed once per iteration.

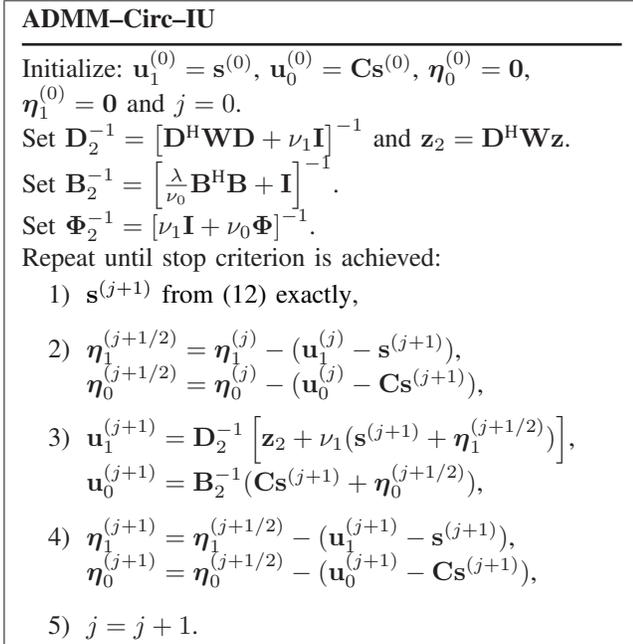


Fig. 3. The ADMM-Circ algorithm with intermediate Lagrange multiplier updating (ADMM-Circ-IU). Note that $\mathbf{C}\mathbf{s}^{(j+1)}$ only needs to be computed once per iteration.

D. Parameter Selection

Any regularized method requires the selection of the regularization parameter, λ in (1), which controls the smoothness of the sensitivity profile. We discuss how λ is selected for typical problems in Section III-C.

In addition, our proposed AL methods require that we specify values for the AL penalty parameters ν_0 and ν_1 . Following [20], we determine the parameter values using the condition numbers of the matrices requiring inversion in the alternating minimization steps. For both the ADMM-Circ method and its variation, we consider the matrices \mathbf{B}_2 , $\boldsymbol{\Phi}_2$, and \mathbf{D}_2 . We normalize the coil images before performing

the estimate; thus, the condition number of \mathbf{D}_2 , $\kappa(\mathbf{D}_2)$, is $(1 + \nu_1)/(d_{\min}^2 + \nu_1)$ where d_{\min}^2 is the smallest diagonal element of $\mathbf{D}^H\mathbf{W}\mathbf{D}$. Furthermore, effective weighting matrices should have near zero values to remove the effects of noise in the low signal regions of the body coil image. Thus, $d_{\min}^2 \approx 0$ and $\kappa(\mathbf{D}_2)$ does not typically depend on the data. We therefore set our parameters by considering the condition numbers of the other two matrices, $\kappa(\mathbf{B}_2)$ and $\kappa(\boldsymbol{\Phi}_2)$. Through extensive numerical simulation, we found that setting ν_0 such that $\kappa(\mathbf{B}_2) \in [225, 400]$ and then ν_1 such that $\kappa(\boldsymbol{\Phi}_2) \in [200, 1000]$ provided good convergence rates for a variety of data sets.

III. RESULTS

We evaluated our proposed sensitivity estimation methods using two very different data sets. The first experiment used simulated brain data whereas the second used real breast phantom data. Previous publications investigated the accuracy of similar regularized estimators [5]; however, there have been few comparisons with other methods concerning their effects on SENSE reconstruction quality. We therefore included an illustration of the improved SENSE reconstruction quality obtained from using regularized sensitivity estimates over standard techniques in the supplementary material. The focus of this paper is on accelerating these algorithms and thus, in this section, we compare the convergence speeds of our AL algorithms with those of conventional CG and PCG with the following circulant preconditioner (PCG-Circ):

$$\mathbf{P}_C = \mathbf{Q}^H (\mathbf{I} + \lambda\boldsymbol{\Omega}) \mathbf{Q}, \quad (13)$$

where $\boldsymbol{\Omega}$ is a diagonal matrix containing the spectrum of the convolution kernel of $\mathbf{R}^H\mathbf{R}$ [32].

We initialized each algorithm with a sensitivity profile comprising the standard ratio estimate over the object voxels and the mean magnitude and phase of these values over the non-object voxels. All of the algorithms were implemented in MATLAB (The MathWorks, Natick, MA, USA) and the experiments were run on a PC with a 2.66 GHz, quad-core Intel Xeon CPU.

We compared the convergence properties of the algorithms using the normalized ℓ_2 -distance between the current estimate, $\mathbf{s}^{(j)}$, and the minimizer of (1), $\hat{\mathbf{s}}$:

$$\mathcal{D}(\mathbf{s}^{(j)}) = \frac{\|\mathbf{s}^{(j)} - \hat{\mathbf{s}}\|_2}{\|\hat{\mathbf{s}}\|_2}. \quad (14)$$

We focused on 2D estimation problems so that we could use Cholesky factorization to determine a non-iterative ‘‘exact’’ solution to (1). Using this non-iterative solution for $\hat{\mathbf{s}}$ avoids favoring a specific iterative algorithm.

A. Cost Function Setup

In defining the estimation problem (1), we chose a second-order finite differencing matrix for \mathbf{R} as it resulted in more accurate sensitivity estimates than both first-order and fourth-order finite differences (results not shown). We used a binary mask, created by thresholding the body coil image, for the weighting matrix \mathbf{W} . This ensured that the majority of voxels

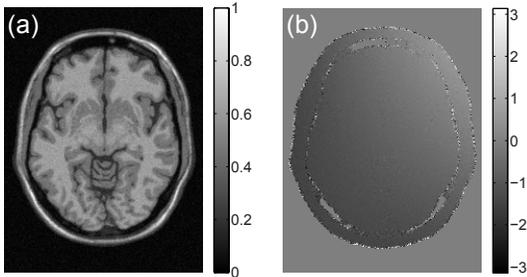


Fig. 4. The (a) magnitude and (b) phase (masked) of the body coil image for the simulated brain data.

in the object support were included in the estimate, while limiting the number of noisy, non-object voxels.

We selected the AL penalty parameters ν_0 and ν_1 for both experiments using the same set of condition numbers. In particular, we selected ν_0 and ν_1 such that $\kappa(\mathbf{B}_2) = 255$ and $\kappa(\Phi_2) = 650$ for both the ADMM-Circ and ADMM-Circ-IU algorithms.

B. Simulated Brain Data

Our first experiment used a 256×192 pixel, T1 weighted, transverse plane brain image from the BrainWeb database [33] (1 mm isotropic in-plane resolution, slice thickness = 1 mm, no noise). To create a more realistic MR image, we added a slowly varying phase component to the brain image. We then added complex random Gaussian noise to create a body coil image, \mathbf{y} , with a signal-to-noise ratio² (SNR) of 10. Fig. 4 presents the magnitude and phase of the resulting body coil image.

We simulated sensitivity profiles for four circular coils placed just outside the field-of-view (FOV) using an analytic method [34]. These sensitivities were then combined with our complex brain image and complex random Gaussian noise to create four surface coil images, \mathbf{z} , with SNRs of approximately 10. Fig. 5 presents the true sensitivities and their corresponding surface coil images.

We estimated the coil sensitivities using our proposed AL methods and the two CG methods. We set $\lambda = 2^5$ as this value produced accurate estimates (compared to the truth) over both the high intensity voxels and their surrounding regions. We ran 20 000 iterations of each method to ensure convergence. All of the algorithms converged to a normalized ℓ_2 -distance of less than -200 dB from, and appeared nearly identical to, the Cholesky based solution $\hat{\mathbf{s}}$. Fig. 6 presents the estimated coil sensitivities as well as their percentage difference to the truth. The convergence rates of the algorithms were similar for all four coils so we present the results for one representative coil. Fig. 7 plots $\mathcal{D}(\mathbf{s}^{(j)})$ with respect to both iteration and time for the bottom left coil in Fig. 5. ADMM-Circ-IU was the fastest algorithm, converging within $\mathcal{D}(\mathbf{s}^{(j)}) = 0.1\%$ in approximately 85 seconds. PCG-Circ was faster than ADMM-Circ with convergence times of nearly 130

²SNR = μ_o/σ_b where μ_o is the mean of the magnitudes of the non-zero object pixels in the true image and σ_b is the standard deviation of the background pixels in the noisy image.

and 165 seconds, respectively. Conventional CG took by far the longest time at 535 seconds.

C. Breast Phantom Data

Our second experiment used a breast phantom consisting of two containers plastered with vegetable shortening and filled with “Super Stuff” bolus material (Radiation Products Design Inc., Albertville, MN, USA). Calibration data was acquired using four surface coils and one body coil on a Philips 3T scanner ($T_R = 4.6$ ms, $T_E = 1.7$ ms, matrix = 384×96). We reconstructed four surface coil images and one body coil image, each 384×96 pixels, using an inverse FFT. Figs. 8 and 9(a) show the magnitudes of the body coil image and surface coil images, respectively. This data set presents several challenges for sensitivity estimation due to the placement of coils near the center of the FOV and because of large regions of low signal both within and outside the object.

To determine a suitable regularization parameter, λ , we first estimated the coil sensitivities using the CG method for several values of λ . We then performed two-fold accelerated SENSE reconstructions [1] using each set of estimated sensitivities and compared the resulting images to the body coil image (not shown). We selected $\lambda = 2^7$ as its corresponding reconstructed image had minimal artifacts and matched closely to the body coil image.

We estimated the coil sensitivities using our proposed AL methods and the two CG methods. We ran 20 000 iterations of each algorithm to ensure that convergence was achieved. Again, the resulting estimates all converged to a normalized ℓ_2 -distance of less than -200 dB from the Cholesky based solution $\hat{\mathbf{s}}$. Fig. 9(b) presents the estimated coil sensitivities. The convergence rates of the algorithms were similar for all four coils so we present the results for one representative coil. Fig. 10 plots $\mathcal{D}(\mathbf{s}^{(j)})$ with respect to both iteration and time for the bottom left coil in Fig. 9. ADMM-Circ-IU was again the fastest algorithm, converging within $\mathcal{D}(\mathbf{s}^{(j)}) = 0.1\%$ in approximately 50 seconds. Unlike in the brain data experiment, ADMM-Circ had a similar convergence rate to PCG-Circ with both algorithms requiring approximately 100 seconds. Conventional CG again took the longest time at 445 seconds.

IV. DISCUSSION

The sensitivity estimates generated by minimizing the cost function in (1) are smooth like true coil sensitivity profiles. As further discussed in the supplementary material, the sensitivity estimates of the brain data are highly accurate over the object and surrounding pixels. The largest errors are at the extreme corners of the image where there is no information about the true sensitivities. The flexibility of the regularized estimation method is highlighted in the breast phantom experiment by its ability to simultaneously estimate the sensitivity within both breasts and smoothly extrapolate over the regions in-between. This is particularly evident for the coils that have near uniform sensitivity over a single breast (the top right and bottom left coils in Fig. 9). As illustrated in the supplementary material, SENSE reconstructions performed with these sensitivity profiles were artifact free unlike those created using low-pass filter techniques.

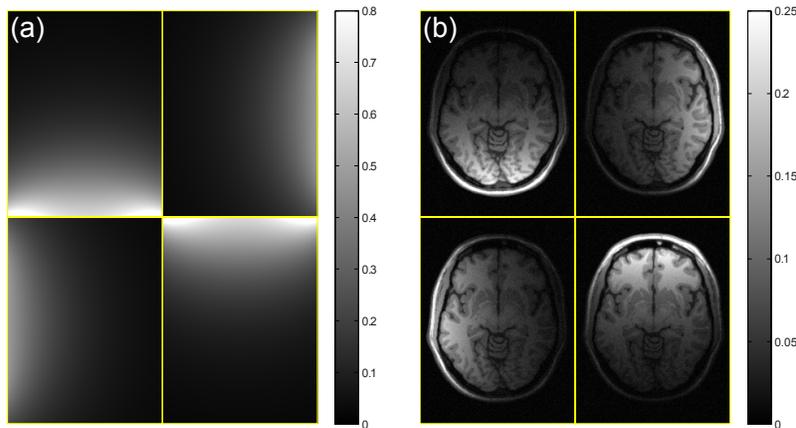


Fig. 5. The magnitudes of the (a) simulated sensitivity profiles and the (b) simulated surface coil images for the brain data.

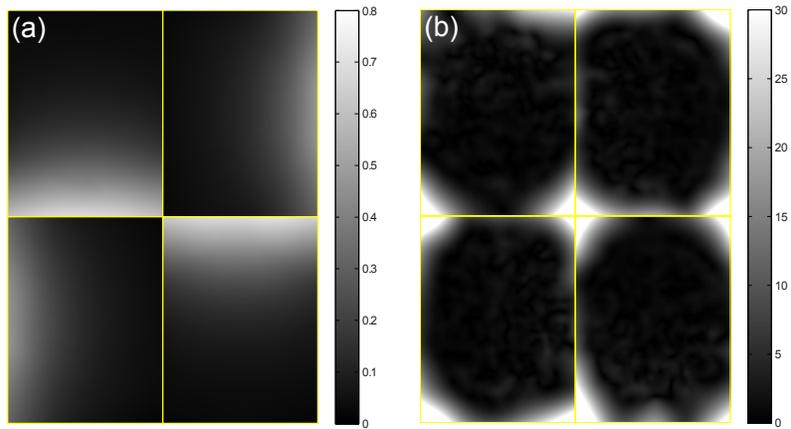


Fig. 6. The magnitudes of the (a) estimated sensitivity profiles and (b) their percentage difference to the true sensitivities for the simulated brain data.

ADMM–Circ–IU was the fastest method in all experiments requiring as little as half the time of PCG–Circ and a ninth the time of conventional CG. ADMM–Circ, although much faster than the CG based methods over the first few iterations, had similar convergence times to PCG–Circ in our breast experiment and was slower in our simulated brain experiment. Thus, using intermediate updating significantly accelerated our ADMM algorithm. The CG algorithm remained the slowest method in all experiments. Interestingly, the relative convergence rate of the PCG–Circ algorithm depended on the experimental data. This behavior is partly a result of the varying accuracy of the preconditioner used in the PCG algorithm. Specifically, the circulant preconditioner used an identity matrix in place of the weighted body coil image voxel intensities (i.e., $\mathbf{I} + \lambda\mathbf{R}^H\mathbf{R}$ for $\mathbf{D}^H\mathbf{W}\mathbf{D} + \lambda\mathbf{R}^H\mathbf{R}$). This approximation works best for images that have few low signal voxels as is apparent from the decreased performance of the PCG–Circ algorithm on the breast data compared to the simulated brain data which has a higher percentage of voxels with significant signal. In contrast, our proposed ADMM algorithms do not require such approximations and their convergence speeds are therefore more robust to differences in the data.

Table I presents the approximate number of complex multiplication and addition operations required by an iteration of each algorithm. For typical finite differencing matrices,

ADMM–Circ–IU, ADMM–Circ, and PCG–Circ require a similar number of operations, whereas traditional CG requires the fewest number of operations per iteration. The effect of these varying costs per iteration is highlighted by contrasting the convergence rates of each algorithm in terms of iteration and time as seen in Figs. 7 and 10. As with time, ADMM–Circ–IU needed approximately half as many iterations as PCG–Circ and ADMM–Circ. CG required significantly more iterations to converge than the other algorithms, offsetting any savings in cost per iteration.

The convergence curves for our ADMM methods exhibited non-monotonic behavior with respect to $\mathcal{D}(\mathbf{s}^{(j)})$. We found that the degree of non-monotonicity was influenced by the choice of AL penalty parameters, ν_0 and ν_1 . In fact, the parameter settings that provided the fastest convergence rates typically resulted in non-monotonicity in the $\mathcal{D}(\mathbf{s}^{(j)})$ plots.

All of our proposed algorithms converged to the solution of (1) in every experiment. As discussed after (12), our ADMM–Circ algorithm is equivalent to an ADMM algorithm with exact update steps. We can therefore conclude that this algorithm converges to the solution of (1) as per [28, Th. 8]. Our intermediate updating variant, ADMM–Circ–IU, does not have the exact formulation outlined in the hypotheses of [28, Th. 8]. However, a guarantee of convergence exists for similar ADMM variants with symmetric Lagrange multiplier

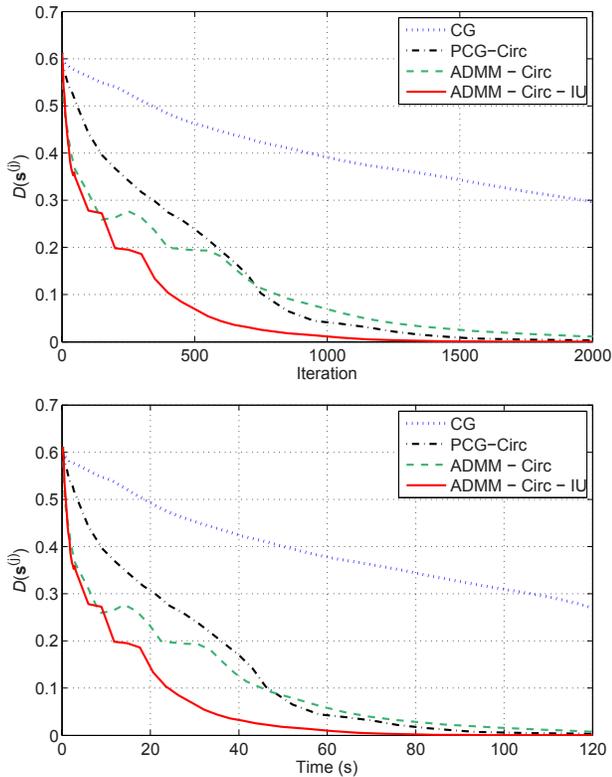


Fig. 7. Plots of the normalized ℓ_2 -distance between $\mathbf{s}^{(j)}$ and $\hat{\mathbf{s}}$, $\mathcal{D}(\mathbf{s}^{(j)})$, with respect to iteration (top) and time (bottom) for the bottom left brain data surface coil image in Fig. 5.



Fig. 8. The magnitude of the breast phantom body coil image.

updating [29]. We are currently investigating an extension of this proof to ADMM-Circ-IU.

The convergence rates of our proposed algorithms were robust to the particular choice of condition numbers used to determine the AL penalty parameters ν_0 and ν_1 . In fact, we used the same condition numbers for our two very different experiments. Furthermore, our fastest algorithm, ADMM-Circ-IU, remained faster than PCG-Circ for $\kappa(\mathbf{B}_2)$ values nearly two times larger and smaller than the optimal value and for $\kappa(\Phi_2)$ values three times larger or smaller than optimal. We also explored varying the λ value in our experiments and found

TABLE I
APPROXIMATE NUMBER OF COMPLEX ARITHMETIC OPERATIONS PER ITERATION FOR THE CASE OF SECOND-ORDER FINITE DIFFERENCES

Estimator	Number of Operations
ADMM-Circ-IU	$19N + 13M^a + 2 \cdot O_{\text{FFT}}^b$
ADMM-Circ	$17N + 11M + 2 \cdot O_{\text{FFT}}$
PCG-Circ	$23N + 11M + 2 \cdot O_{\text{FFT}}$
CG	$22N + 11M$

^a $M \approx 4N$ for 2D problems.

^b O_{FFT} denotes the cost of the FFT operations ($O(N \log(N))$).

that this set of condition numbers consistently worked well. The choice of the best condition numbers does not depend on the surface coil image. Therefore, if one wanted to fine-tune the convergence rate of the algorithms, a single coil of a multi-coil array would suffice.

It is a common practice in medical imaging to restrict estimates and reconstructions to within masked regions to improve both their computation time and quality over the region. If this is done for simple problems like our simulated brain data, which requires minimal interpolation within the object support, then the PCG-Circ algorithm estimating within a masked region will converge faster than our ADMM-Circ methods estimating over the entire FOV. However, this is not the case for more complicated problems like our breast phantom data. In particular, we found that our ADMM-Circ-IU algorithm, estimating over a full FOV, converged to $\mathcal{D}(\mathbf{s}^{(j)}) = 0.1\%$ at the same speed or faster than a PCG-Circ algorithm estimating within a masked region consisting of a convex hull³ surrounding the object support, Fig. 11. Furthermore, the quality of the unmasked ADMM-Circ-IU estimates was similar to that of the masked PCG estimates over the masked region. This is partially because the weighting matrix \mathbf{W} minimizes the impact of noisy voxels outside of the object support. A major disadvantage of masking is that the lack of an estimate outside the mask can lead to significant SENSE reconstruction artifacts if the object moves into this region during acquisition [6]. Thus, the mask would have to be carefully selected with this in mind. We therefore followed existing work [5] and focused on algorithms without support masks.

In addition to the algorithms presented in this work and the supplementary material, we also explored AL algorithms that incorporated simpler variable splittings. For instance, we introduced the single splitting variable $\mathbf{u}_0 = \mathbf{R}\mathbf{s}$ to (1) and similarly, $\mathbf{u}_0 = \mathbf{C}\mathbf{s}$ to (2). The AL formulations used to minimize the resulting cost functions had only two update equations. However, one of these equations required an approximate iterative solution and the resulting AL algorithms were highly sensitive to inaccuracies in the approximation. In fact, when using PCG for the approximate update step, the optimal number of inner PCG iterations was so large that the overall algorithms were slower than regular CG. Curiously, this is the type of splitting that is used in the popular split Bregman approaches [19], although there it is used in cases where $\mathbf{R}^H\mathbf{R}$ is circulant.

If the body coil data \mathbf{y} is not available, one could use the square-root of the sum-of-squares of the surface coil images in its place [8], [35], [36]. Our algorithms would remain the same and only the elements of \mathbf{D} would change. However, it may be more desirable in this situation to perform joint estimation of the final image and the sensitivity profiles (e.g., [8]). Such algorithms are more complicated to compute than (1) and might also benefit from an ADMM reformulation.

³See the supplementary material for an illustration of why a convex hull is required.

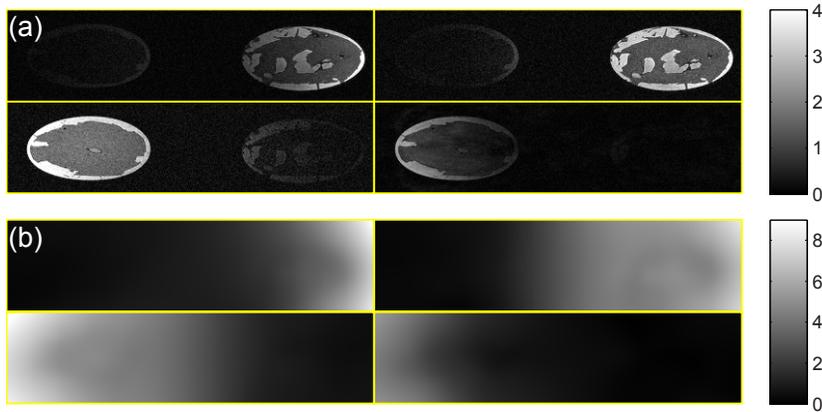


Fig. 9. The magnitudes of the (a) breast phantom surface coil images and the (b) corresponding estimated sensitivity profiles.

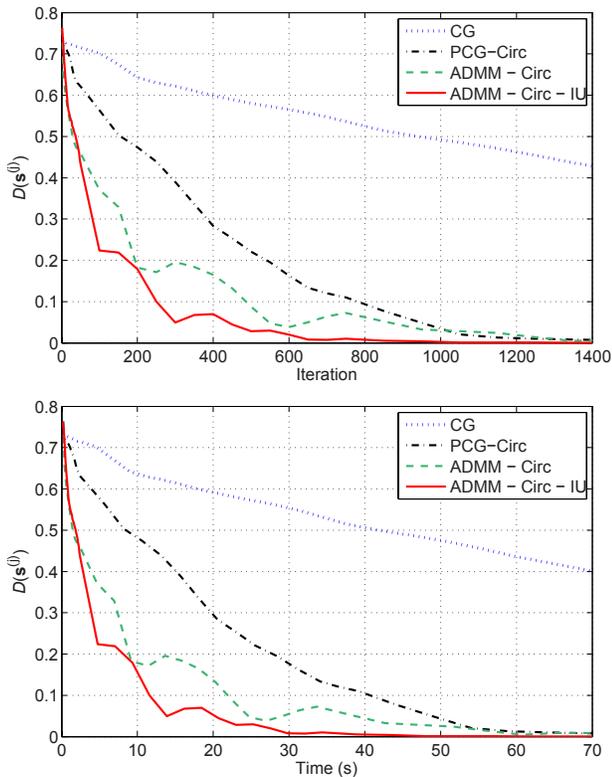


Fig. 10. Plots of the normalized ℓ_2 -distance between $\mathbf{s}^{(j)}$ and $\hat{\mathbf{s}}$, $\mathcal{D}(\mathbf{s}^{(j)})$, with respect to iteration (top) and time (bottom) for the bottom left breast data surface coil image in Fig. 9.

V. CONCLUSIONS

We developed a new iterative method, ADMM-Circ, using variable splitting and AL strategies that accelerates the regularized estimation of MR coil sensitivities. By separating the finite differencing matrix for the case of non-periodic boundary conditions into a finite differencing matrix for the case of periodic boundary conditions and a diagonal masking matrix, we were able to find a variable splitting strategy that resulted in an algorithm with exact update steps. Additionally, we demonstrated that intermediate updating of the Lagrange multipliers significantly accelerated our proposed AL algorithm. Our fastest method, ADMM-Circ-IU, had convergence

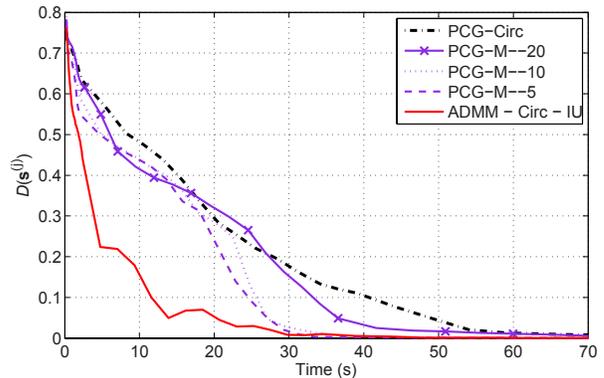


Fig. 11. Plots of $\mathcal{D}(\mathbf{s}^{(j)})$ with respect to time for ADMM-Circ-IU and PCG-Circ without masks, as well as PCG-Circ using masks with various degrees of dilation (5, 10, 20 pixels), applied to the bottom left breast data surface coil image in Fig. 9. For each case, the $\hat{\mathbf{s}}$ used in $\mathcal{D}(\mathbf{s}^{(j)})$ is the regularized solution for the appropriate mask.

speeds up to twice those of the PCG method with a circulant preconditioner.

More generally, this work illustrates how AL methods can be used to accelerate convergence for imaging problems with certain classes of quadratic cost functions. There are many areas in MR imaging where similar cost functions are used. For instance, B_0 and B_1 map estimation can be performed by minimizing cost functions over the image domain with quadratic regularization terms [37]–[39]. The application of similar acceleration techniques to these problems is currently being investigated [40].

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their insightful comments regarding practical aspects of the algorithms. We would also like to thank Dr. T. Chenevert and Dr. M. Ivancevic for providing the breast phantom data.

REFERENCES

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Mag. Res. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.

- [2] D. K. Sodickson and W. J. Manning, "Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays," *Mag. Res. Med.*, vol. 38, no. 4, pp. 591–603, Oct. 1997.
- [3] J. Tsao, P. Boesiger, and K. P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Mag. Res. Med.*, vol. 50, no. 5, pp. 1031–42, Nov. 2003.
- [4] S. O. Schönberg, O. Dietrich, M. F. Reiser, and A. L. Baert, *Parallel Imaging in Clinical MR Applications*, Medical Radiology Diagnostic Imaging Series. Springer, 2007.
- [5] S. L. Keeling and R. Bammer, "A variational approach to magnetic resonance coil sensitivity estimation," *Applied Mathematics and Computation*, vol. 158, no. 2, pp. 53–82, Nov. 2004.
- [6] D. J. Larkman and R. G. Nunes, "Parallel magnetic resonance imaging," *Phys. Med. Biol.*, vol. 52, no. 7, pp. R15–R55, Apr. 2007.
- [7] R. Bammer, R. Stollberger, M. J. Augustin, T. Seifert, S. Strasser-Fuchs, P. Wach, H. P. Hartung, and F. Fazekas, "Parallel imaging strategies for high-speed magnetic resonance diffusion imaging," in *Proc. SPIE 3978*, 2000, vol. 12.
- [8] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)," *Mag. Res. Med.*, vol. 57, no. 6, pp. 1196–1202, June 2007.
- [9] P. Vemuri, E. G. Kholmovski, D. L. Parker, and B. E. Chapman, "Coil sensitivity estimation for optimal SNR reconstruction and intensity inhomogeneity correction in phased array MR imaging," in *Information Processing in Medical Im.*, 2005, pp. 603–14.
- [10] F. H. Lin, Y. J. Chen, J. W. Belliveau, and L. L. Wald, "A wavelet-based approximation of surface coil sensitivity profiles for correction of image intensity inhomogeneity and parallel imaging reconstruction," *Hum. Brain Map.*, vol. 19, no. 2, pp. 96–111, Mar. 2003.
- [11] M. A. G. Ballester, Y. Machida, Y. Kassai, Y. Hamamura, and H. Sugimoto, "Robust estimation of coil sensitivities for RF subencoded acquisition techniques," in *Proc. Intl. Soc. Mag. Res. Med.*, 2001, p. 799.
- [12] J. Jin, F. Liu, E. Weber, Y. Li, and S. Crozier, "An electromagnetic reverse method of coil sensitivity mapping for parallel MRI - theoretical framework," *J. Magn. Reson.*, vol. 207, no. 1, pp. 59–68, 2010.
- [13] R. Bammer, M. Auer, S. L. Keeling, M. Augustin, L. A. Stables, R. W. Prokesch, R. Stollberger, M. E. Moseley, and F. Fazekas, "Diffusion tensor imaging using single-shot SENSE-EPI," *Mag. Res. Med.*, vol. 48, no. 1, pp. 128–36, July 2002.
- [14] F. Huang, Y. Chen, G. R. Duensing, J. Akao, A. Rubin, and C. Saylor, "Application of partial differential equation-based inpainting on sensitivity maps," *Mag. Res. Med.*, vol. 53, no. 2, pp. 388–97, Feb. 2005.
- [15] L. L. Wald, "Parallel Imaging Update: How Many Elements Do We Need?," in *Proc. Intl. Soc. Mag. Res. Med.*, 2006, p. 202.
- [16] M. J. Allison and J. A. Fessler, "An augmented Lagrangian method for MR coil sensitivity estimation," in *Proc. Intl. Soc. Mag. Res. Med.*, 2011, p. 2881.
- [17] D. P. Bertsekas, "Multiplier methods: A survey," *Automatica*, vol. 12, no. 2, pp. 133–45, Mar. 1976.
- [18] L. M. Bregman, "The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR Comp Math and Math Phys*, vol. 7, pp. 200–17, 1967.
- [19] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–43, 2009.
- [20] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, Mar. 2011.
- [21] B. Liu, K. King, M. Steckner, J. Xie, J. Sheng, and L. Ying, "Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations," *Mag. Res. Med.*, vol. 61, no. 1, pp. 145–52, Jan. 2009.
- [22] X. Ye, Y. Chen, and F. Huang, "Computational acceleration for MR image reconstruction in partially parallel imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1055–63, May 2011.
- [23] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition*, 2008, pp. 1–8.
- [24] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Im. Proc.*, vol. 20, no. 3, pp. 681–695, 2011.
- [25] M. V. Afonso, José M Bioucas-Dias, and Mário A T Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Im. Proc.*, vol. 19, no. 9, pp. 2345–56, Sept. 2010.
- [26] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 677–88, Mar. 2012.
- [27] M. A. T. Figueiredo and José M Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Trans. Im. Proc.*, vol. 19, no. 12, pp. 3133–45, Dec. 2010.
- [28] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, Apr. 1992.
- [29] R. Glowinski and P. L. Tallec, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, Soc. Indust. Appl. Math., 1989.
- [30] M. J. Allison, S. Ramani, and J. A. Fessler, "Regularized MR coil sensitivity estimation using augmented Lagrangian methods," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2012, pp. 394–7.
- [31] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imaging Sci.*, vol. 1, no. 3, pp. 248–72, 2008.
- [32] R. H. Chan, "Circulant preconditioners for Hermitian Toeplitz systems," *SIAM J. Matrix. Anal. Appl.*, vol. 10, no. 4, pp. 542–50, 1989.
- [33] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–8, June 1998.
- [34] M. I. Grivich and D. P. Jackson, "The magnetic field of current-carrying polygons: An application of vector field rotations," *Amer. J. Phys.*, vol. 68, no. 5, pp. 469–74, May 2000.
- [35] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller, "The NMR phased array," *Mag. Res. Med.*, vol. 16, no. 2, pp. 192–225, Nov. 1990.
- [36] C. A. McKenzie, E. N. Yeh, M. A. Ohliger, M. D. Price, and D. K. Sodickson, "Self-calibrating parallel imaging with automatic coil sensitivity extraction," *Mag. Res. Med.*, vol. 47, no. 3, pp. 529–38, Mar. 2002.
- [37] A. K. Funai, J. A. Fessler, D. T. B. Yeo, V. T. Olafsson, and D. C. Noll, "Regularized field map estimation in MRI," *IEEE Trans. Med. Imag.*, vol. 27, no. 10, pp. 1484–94, Oct. 2008.
- [38] A. Funai, J. A. Fessler, W. Grissom, and D. C. Noll, "Regularized B1+ map estimation in MRI," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2007, pp. 616–9.
- [39] V. T. Olafsson, D. C. Noll, and J. A. Fessler, "Fast joint reconstruction of dynamic R_2^* and field maps in functional MRI," *IEEE Trans. Med. Imag.*, vol. 27, no. 9, pp. 1177–88, Sept. 2008.
- [40] M. J. Allison and J. A. Fessler, "Accelerated computation of regularized field map estimates," in *Proc. Intl. Soc. Mag. Res. Med.*, 2012, p. 0413.

Accelerated Regularized Estimation of MR Coil Sensitivities Using Augmented Lagrangian Methods: Supplementary Material

Michael J. Allison*, *Student Member, IEEE*, Sathish Ramani, *Member, IEEE*, and Jeffrey A. Fessler, *Fellow, IEEE*

This document contains additional information related to the algorithms presented in [1]. In particular, Section S-I presents a second ADMM algorithm for sensitivity profile estimation. Section S-II presents an alternate formulation that leads to additional AL estimation algorithms with similar performance. Section S-III illustrates the improved SENSE reconstruction quality resulting from using regularized sensitivity estimates over traditional ratio based estimates. Section S-IV demonstrates why, when performing masked sensitivity estimation on breast data, a convex hull of the object voxels should be used for the estimation mask. Section S-V illustrates the importance of using a finite differencing matrix for the case of non-periodic boundary conditions in our cost function (1).

S-I. ADMM-CG: ADMM SENSITIVITY ESTIMATION ALGORITHM WITH CONJUGATE GRADIENT SUBSTEPS

In this section we present and evaluate a second AL algorithm, ADMM-CG, which does not use the reformulation of the finite differencing matrix as discussed in Section II-B of the manuscript. We begin with the derivation of the algorithm which uses the same techniques as in the manuscript. We then compare this new algorithm to the methods presented in the manuscript using the same data sets and briefly discuss its properties.

A. Method Derivation

We begin our derivation by introducing two new variables, $\mathbf{u}_0 \in \mathbb{C}^M$ and $\mathbf{u}_1 \in \mathbb{C}^N$, to the initial cost function in (1). The purpose of these variables is to isolate the finite differencing matrix \mathbf{R} from the diagonal matrix \mathbf{D} . The resulting constrained optimization problem is

$$\hat{\mathbf{s}} \triangleq \arg \min_{\mathbf{s}, \mathbf{u}_0, \mathbf{u}_1} \frac{1}{2} \|\mathbf{z} - \mathbf{D}\mathbf{u}_1\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \|\mathbf{u}_0\|_2^2 \quad (\text{S-1})$$

s.t. $\mathbf{u}_1 = \mathbf{s}$ and $\mathbf{u}_0 = \mathbf{R}\mathbf{s}$.

Solving this constrained optimization problem is exactly equivalent to solving the unconstrained problem (1).

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by the National Institutes of Health (NIH) under Grant P01 CA87634. *Asterisk indicates corresponding author.*

M. J. Allison*, S. Ramani, and J. A. Fessler are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: mjalliso@umich.edu, sramani@umich.edu, fessler@umich.edu).

As in (4), we can express (S-1) in the more concise notation:

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}, \mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{K}\mathbf{u}\|_2^2 \quad \text{s.t.} \quad \mathbf{u} = \mathbf{G}\mathbf{s}, \quad (\text{S-2})$$

where \mathbf{u} and \mathbf{h} were defined in (4),

$$\mathbf{G} \triangleq \begin{bmatrix} \mathbf{I} \\ \mathbf{R} \end{bmatrix}, \quad \text{and} \quad \mathbf{K} \triangleq \begin{bmatrix} \mathbf{W}^{\frac{1}{2}}\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \sqrt{\lambda}\mathbf{I} \end{bmatrix}.$$

We then tackle (S-2) using the previously described AL formalism and obtain the following AL function-based minimization problem:

$$\arg \min_{\mathbf{s}, \mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{K}\mathbf{u}\|_2^2 + \frac{1}{2} \|\mathbf{u} - \mathbf{G}\mathbf{s} - \boldsymbol{\eta}\|_{\mathbf{V}}^2, \quad (\text{S-3})$$

where $\boldsymbol{\eta}$ and \mathbf{V} were defined in (5).

Due to the complexity of jointly minimizing (S-3) over \mathbf{s} and \mathbf{u} , we again consider an alternating minimization scheme. In particular, we sequentially solve the following set of equations:

$$\mathbf{s}^{(j+1)} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{u}^{(j)} - \mathbf{G}\mathbf{s} - \boldsymbol{\eta}^{(j)}\|_{\mathbf{V}}^2, \quad (\text{S-4})$$

$$\mathbf{u}^{(j+1)} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{h} - \mathbf{K}\mathbf{u}\|_2^2 + \frac{1}{2} \|\mathbf{u} - \mathbf{G}\mathbf{s}^{(j+1)} - \boldsymbol{\eta}^{(j)}\|_{\mathbf{V}}^2. \quad (\text{S-5})$$

As with ADMM-Circ, the update equation for \mathbf{u} , (S-5), has a simple closed-form solution which can be decoupled into two parallel updates in terms of \mathbf{u}_1 and \mathbf{u}_0 due to the block diagonal structures of \mathbf{K} and \mathbf{V} :

$$\mathbf{u}_1^{(j+1)} = \mathbf{D}_2^{-1} \left[\mathbf{D}^H \mathbf{W} \mathbf{z} + \nu_1 (\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_1^{(j)}) \right], \quad (\text{S-6})$$

$$\mathbf{u}_0^{(j+1)} = \frac{\nu_0}{\nu_0 + \lambda} \left(\mathbf{R}\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_0^{(j)} \right). \quad (\text{S-7})$$

where $\mathbf{D}_2 \triangleq \mathbf{D}^H \mathbf{W} \mathbf{D} + \nu_1 \mathbf{I}$ is a diagonal matrix.

Equation (S-4) does not have an efficient closed-form solution due to the size and complexity of \mathbf{R} . Instead, we approximately solve (S-4) using several iterations of the preconditioned conjugate gradient (PCG) method with warm starting, the optimal number of which is determined empirically. We design the specific preconditioner, \mathbf{P} , by considering the closed-form solution of (S-4):

$$\mathbf{G}_2 \mathbf{s}^{(j+1)} = \mathbf{G}^H \mathbf{V} \left(\mathbf{u}^{(j)} - \boldsymbol{\eta}^{(j)} \right), \quad (\text{S-8})$$

where $\mathbf{G}_2 = \mathbf{G}^H \mathbf{V} \mathbf{G} = \nu_1 \mathbf{I} + \nu_0 \mathbf{R}^H \mathbf{R}$. Our goal is to create an easily invertible \mathbf{P} that preconditions \mathbf{G}_2 to obtain fast

convergence for this subproblem. For typical finite differencing matrices with non-periodic boundaries, $\mathbf{R}^H\mathbf{R}$ has a near block circulant with circulant blocks structure. We therefore approximate $\mathbf{R}^H\mathbf{R}$ in our preconditioner as $\mathbf{Q}^H\mathbf{\Omega}\mathbf{Q}$ where \mathbf{Q} is a (multidimensional) discrete Fourier transform (DFT) matrix and $\mathbf{\Omega}$ is a diagonal matrix containing the spectrum of the convolution kernel of $\mathbf{R}^H\mathbf{R}$ [2]. Our resulting preconditioner is

$$\mathbf{P} = \mathbf{Q}^H (\nu_1 \mathbf{I} + \nu_0 \mathbf{\Omega}) \mathbf{Q}. \quad (\text{S-9})$$

Fig. S-1 summarizes the resulting estimation algorithm composed of these update steps and the corresponding Lagrange multiplier updates, ADMM–CG. Note that the minimization in Step 1 is inexact, requiring an iterative solution; however, the optimal number of iterations is typically small. Furthermore, it can be shown that this algorithm is equivalent to an ADMM algorithm with an approximate update step for which the errors at each outer iteration can be made absolutely summable by using enough PCG iterations. We can therefore conclude that this algorithm converges to the solution of (1) as per [3, Th. 8].

ADMM–CG

Initialize: $\mathbf{u}_1^{(0)} = \mathbf{s}^{(0)}$, $\mathbf{u}_0^{(0)} = \mathbf{R}\mathbf{s}^{(0)}$, $\boldsymbol{\eta}_0^{(0)} = \mathbf{0}$, $\boldsymbol{\eta}_1^{(0)} = \mathbf{0}$ and $j = 0$.

Set $\mathbf{D}_2^{-1} = [\mathbf{D}^H\mathbf{W}\mathbf{D} + \nu_1\mathbf{I}]^{-1}$ and $\mathbf{z}_2 = \mathbf{D}^H\mathbf{W}\mathbf{z}$.

Repeat until stop criterion is achieved:

- 1) $\mathbf{s}^{(j+1)}$ from PCG solution of (S-4) using (S-9).
- 2) $\mathbf{u}_1^{(j+1)} = \mathbf{D}_2^{-1} \left[\mathbf{z}_2 + \nu_1(\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_1^{(j)}) \right]$,
 $\mathbf{u}_0^{(j+1)} = \frac{\nu_0}{\nu_0 + \lambda} \left(\mathbf{R}\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_0^{(j)} \right)$,
- 3) $\boldsymbol{\eta}_1^{(j+1)} = \boldsymbol{\eta}_1^{(j)} - (\mathbf{u}_1^{(j+1)} - \mathbf{s}^{(j+1)})$,
 $\boldsymbol{\eta}_0^{(j+1)} = \boldsymbol{\eta}_0^{(j)} - (\mathbf{u}_0^{(j+1)} - \mathbf{R}\mathbf{s}^{(j+1)})$,
- 4) $j = j + 1$.

Fig. S-1. Overview of the ADMM–CG algorithm. Note that $\mathbf{R}\mathbf{s}^{(j+1)}$ only needs to be computed once per iteration.

B. Alternating Minimization with Intermediate Updating

We also explored updating the Lagrange multipliers between each alternating minimization step. The resulting variation, ADMM–CG–IU, is presented in Fig. S-2. As with the ADMM–Circ–IU algorithm, this algorithm lacks a guarantee of convergence although such guarantees exist for similar intermediate updating algorithms [4].

C. Parameter Selection

The parameter selection strategy for our ADMM–CG based algorithms is similar to the strategy for ADMM–Circ because of the analogous structures of the alternating minimization steps. The major difference is that the update of \mathbf{u}_0 in (S-7) does not require the inversion of a matrix but rather a scalar term. In fact, this scalar term has the same form as $\kappa(\mathbf{B}_2)$ in the manuscript. Subsequently, we found that setting ν_0 such

ADMM–CG–IU

Initialize: $\mathbf{u}_1^{(0)} = \mathbf{s}^{(0)}$, $\mathbf{u}_0^{(0)} = \mathbf{R}\mathbf{s}^{(0)}$, $\boldsymbol{\eta}_0^{(0)} = \mathbf{0}$, $\boldsymbol{\eta}_1^{(0)} = \mathbf{0}$ and $j = 0$.

Set $\mathbf{D}_2^{-1} = [\mathbf{D}^H\mathbf{W}\mathbf{D} + \nu_1\mathbf{I}]^{-1}$ and $\mathbf{z}_2 = \mathbf{D}^H\mathbf{W}\mathbf{z}$.

Repeat until stop criterion is achieved:

- 1) $\mathbf{s}^{(j+1)}$ from PCG solution of (S-4) using (S-9).
- 2) $\boldsymbol{\eta}_1^{(j+1/2)} = \boldsymbol{\eta}_1^{(j)} - (\mathbf{u}_1^{(j)} - \mathbf{s}^{(j+1)})$,
 $\boldsymbol{\eta}_0^{(j+1/2)} = \boldsymbol{\eta}_0^{(j)} - (\mathbf{u}_0^{(j)} - \mathbf{R}\mathbf{s}^{(j+1)})$,
- 3) $\mathbf{u}_1^{(j+1)} = \mathbf{D}_2^{-1} \left[\mathbf{z}_2 + \nu_1(\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_1^{(j+1/2)}) \right]$,
 $\mathbf{u}_0^{(j+1)} = \frac{\nu_0}{\nu_0 + \lambda} \left(\mathbf{R}\mathbf{s}^{(j+1)} + \boldsymbol{\eta}_0^{(j+1/2)} \right)$,
- 4) $\boldsymbol{\eta}_1^{(j+1)} = \boldsymbol{\eta}_1^{(j+1/2)} - (\mathbf{u}_1^{(j+1)} - \mathbf{s}^{(j+1)})$,
 $\boldsymbol{\eta}_0^{(j+1)} = \boldsymbol{\eta}_0^{(j+1/2)} - (\mathbf{u}_0^{(j+1)} - \mathbf{R}\mathbf{s}^{(j+1)})$,
- 5) $j = j + 1$.

Fig. S-2. The ADMM–CG algorithm with intermediate Lagrange multiplier updating (ADMM–CG–IU). Note that $\mathbf{R}\mathbf{s}^{(j+1)}$ only needs to be computed once per iteration.

that the scalar $\frac{\nu_0 + \lambda}{\nu_0} \in [200, 400]$ and then setting ν_1 such that $\kappa(\mathbf{G}_2) \in [200, 1000]$ provided reasonable convergence rates.

D. Results

To evaluate our proposed ADMM–CG based algorithms, we performed the same experiments as in Section III of the manuscript. The cost function was setup as described in Section III-A and the same ratio based estimate was used to initialize the algorithms. We used a single PCG iteration for the approximate update to \mathbf{s} in the ADMM–CG based algorithms as this provided the fastest convergence rates with respect to time. We selected the AL penalty parameters ν_0 and ν_1 for ADMM–CG such that $\frac{\nu_0 + \lambda}{\nu_0} = 225$ and $\kappa(\mathbf{G}_2) = 600$. As further discussed in Section S-I-E, the optimal condition numbers for ADMM–CG–IU depended on the data and are therefore mentioned in the appropriate subsections.

1) *Simulated Brain Data:* We ran 20 000 iterations of the ADMM–CG based algorithms on the simulated brain data described in Section III-B of the manuscript. For the ADMM–CG–IU algorithm, we selected ν_0 and ν_1 such that $\frac{\nu_0 + \lambda}{\nu_0} = 375$ and $\kappa(\mathbf{G}_2) = 600$. Our proposed ADMM–CG and ADMM–CG–IU algorithms converged to a normalized ℓ_2 -distance of less than -200 dB from the Cholesky based solution to (1) and appeared nearly identical to Fig. 6. The convergence rates of the algorithms were similar for all four coils and thus we present the results for the same coil that was presented in Section III-B. Fig. S-3 plots $\mathcal{D}(\mathbf{s}^{(j)})$ with respect to both iteration and time for the ADMM–CG based algorithms as well as the algorithms evaluated in the manuscript. ADMM–CG–IU and ADMM–CG were both slower than PCG–Circ, but faster than conventional CG, reaching $\mathcal{D}(\mathbf{s}^{(j)}) = 0.1\%$ in approximately 145 and 185 seconds, respectively.

2) *Breast Phantom Data:* We also ran 20 000 iterations of the ADMM–CG based algorithms on the breast phantom data described in Section III-C of the manuscript. For the

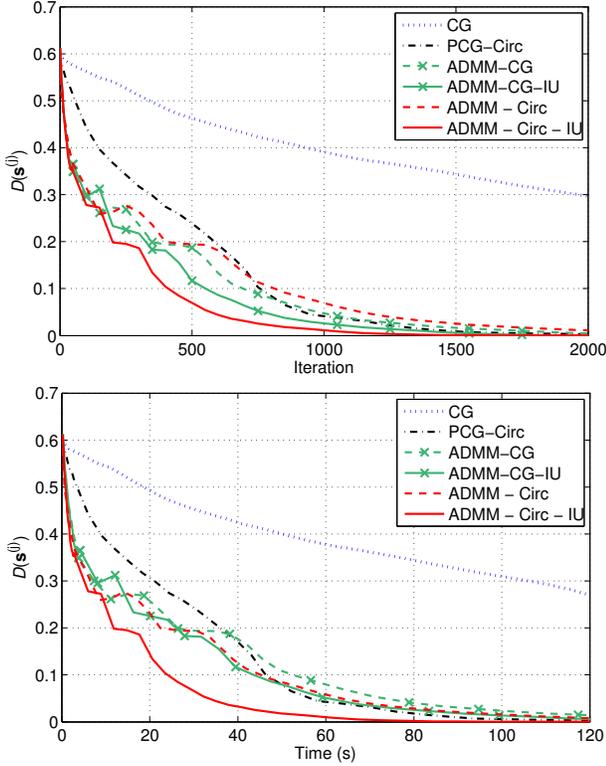


Fig. S-3. Plots of the normalized ℓ_2 -distance between $\mathbf{s}^{(j)}$ and $\hat{\mathbf{s}}$, $\mathcal{D}(\mathbf{s}^{(j)})$, with respect to iteration (top) and time (bottom) for the bottom left brain data surface coil image in Fig. 5.

ADMM-CG-IU algorithm, we selected ν_0 and ν_1 such that $\frac{\nu_0 + \lambda}{\nu_0} = 250$ and $\kappa(\mathbf{G}_2) = 600$. Again, both of our proposed algorithms converged to a normalized ℓ_2 -distance of less than -200 dB from the Cholesky based solution to (1) and appeared nearly identical to Fig. 9(b). The convergence rates of the algorithms were similar for all four coils and thus we present the results for the same coil that was presented in Section III-C. Fig. S-4 plots $\mathcal{D}(\mathbf{s}^{(j)})$ with respect to both iteration and time for the ADMM-CG based algorithms as well as the algorithms evaluated in the manuscript. ADMM-CG-IU was faster than both PCG-Circ and regular CG converging within $\mathcal{D}(\mathbf{s}^{(j)}) = 0.1\%$ in approximately 80 seconds. ADMM-CG was slower than its intermediate updating counterpart and PCG-Circ with a convergence time of nearly 120 seconds.

E. Discussion

The convergence rates with respect to iteration of the ADMM-CG based algorithms were close to their ADMM-Circ counterparts. However, the ADMM-CG based algorithms were much slower in time due to the added overhead of the PCG solution used to approximate Step 1. In fact, even when using only one iteration of PCG for this approximation, the per iteration costs of the ADMM-CG algorithms are much higher than those of the ADMM-Circ algorithms, Table S-1.

The convergence curves of the ADMM-CG based algorithms exhibit a higher rate of non-monotonic behavior than the ADMM-Circ algorithms. This is partly caused by the

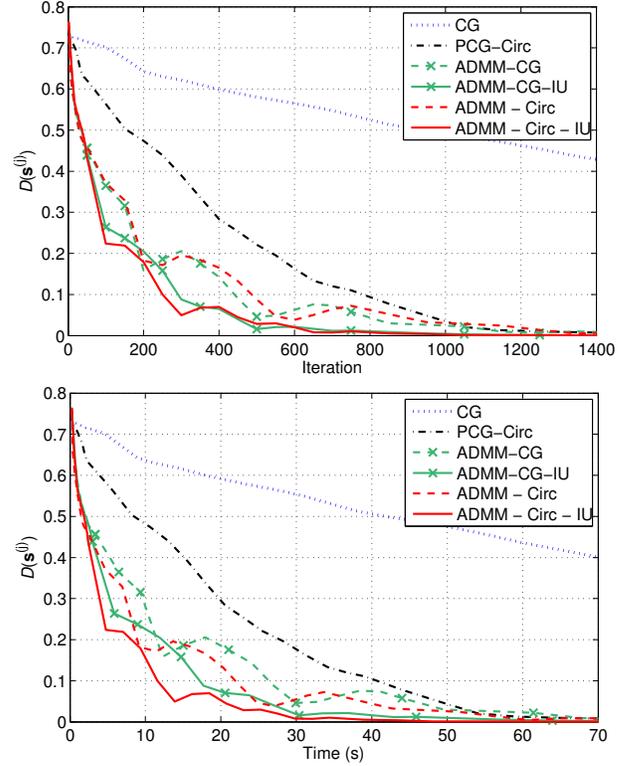


Fig. S-4. Plots of the normalized ℓ_2 -distance between $\mathbf{s}^{(j)}$ and $\hat{\mathbf{s}}$, $\mathcal{D}(\mathbf{s}^{(j)})$, with respect to iteration (top) and time (bottom) for the bottom left breast data surface coil image in Fig. 9.

approximate update in Step 1. If we run several more PCG sub-iterations in Step 1, the convergence curves with respect to iteration of the ADMM-CG algorithms appear similar to their ADMM-Circ counterparts (although much slower with respect to time). As with ADMM-Circ, the parameter settings that provided the fastest convergence rates typically resulted in non-monotonicity in the $\mathcal{D}(\mathbf{s}^{(j)})$ plots.

The proposed ADMM-CG-IU algorithm was faster than PCG-Circ in the breast phantom experiment, but slower in the simulated brain experiment. As discussed in Section IV of the manuscript, the relative speed of the PCG-Circ algorithm depends on the accuracy of the preconditioner in (13) and thus on the data. Contrarily, the preconditioning used for the approximation of Step 1 in the ADMM-CG algorithms does not depend on such differences. The ADMM-CG algorithm, although initially faster, converged slower than PCG-Circ in both experiments. Therefore, using intermediate updating also significantly accelerated the convergence rates of this ADMM algorithm. All of our proposed algorithms were significantly faster than traditional CG.

The convergence rates of our proposed ADMM-CG algorithms were robust to the choice of condition numbers used to determine the AL penalty parameters ν_0 and ν_1 . We found that the convergence rates remained similar for condition numbers that differed from the optimal values by up to fifty percent. The chosen condition numbers also worked well for a wide range of regularization parameter values λ . However, we found that

TABLE S-1
APPROXIMATE NUMBER OF COMPLEX ARITHMETIC OPERATIONS PER
ITERATION FOR THE CASE OF SECOND-ORDER FINITE DIFFERENCES

Estimator	Number of Operations
ADMM-CG-IU ^a	$28N + 18M + 2 \cdot O_{\text{FFT}}^{\text{b}}$
ADMM-CG ^a	$26N + 16M + 2 \cdot O_{\text{FFT}}$

^a Step 1 uses a single PCG iteration.

^b O_{FFT} denotes the cost of the FFT operations ($O(N \log(N))$).

the optimal condition numbers for ADMM-CG-IU depended on the data unlike for ADMM-CG and the ADMM-Circ based algorithms. Still, like the ADMM-Circ algorithms, the choice of the optimal condition numbers does not depend on the surface coil image. Therefore, if one wanted to fine-tune the convergence rate of the algorithms, a single coil of a multi-coil array would suffice.

S-II. SIMILAR SPLITTING

In formulating our proposed ADMM-Circ algorithm, we originally explored a different variable splitting strategy involving a double splitting within the regularization term of (1) [5], [6]:

$$\arg \min_{\mathbf{s}, \mathbf{u}_0} \frac{1}{2} \|\mathbf{z} - \mathbf{D}\mathbf{s}\|_{\mathbf{W}}^2 + \frac{\lambda}{2} \|\mathbf{B}\mathbf{u}_0\|_2^2 \quad \text{s.t. } \mathbf{u}_0 = \mathbf{C}\mathbf{u}_1 \text{ and } \mathbf{u}_1 = \mathbf{s}. \quad (\text{S-10})$$

This variable splitting led to update equations with nearly identical structures to those of ADMM-Circ. Furthermore, the resulting AL algorithm and its intermediate updating variation had similar convergence rates to their ADMM counterparts. However, analyzing the convergence properties of these algorithms was more complicated as they did not have ADMM structures. Thus, we focused on the ADMM formulations.

S-III. IMPROVED SENSE RECONSTRUCTION QUALITY

The advantages and accuracy of similar regularized sensitivity profile estimators have been discussed in previous papers [7], [8]; however, there has been limited investigation into their effects on SENSE reconstruction quality. We therefore compare the quality of the SENSE reconstructions created with the coil sensitivities estimated using the regularized method in (1) to those estimated using the commonplace ratio and ratio of low resolution images methods.

A. Simulated Brain Data

Our first experiment was performed using the simulated brain data outlined in Section III-B of the manuscript. We began by simulating a full resolution calibration scan using the same parameters as Figs. 4 and 5. Next, we estimated the coil sensitivities from the resulting body and surface coil images using our regularized method, the ratio of low resolution images method, and the ratio method.

We implemented the regularized method using our ADMM-Circ-IU algorithm with the same parameters as in Section III of the manuscript. The ratio of low resolution images method was implemented by taking a set number

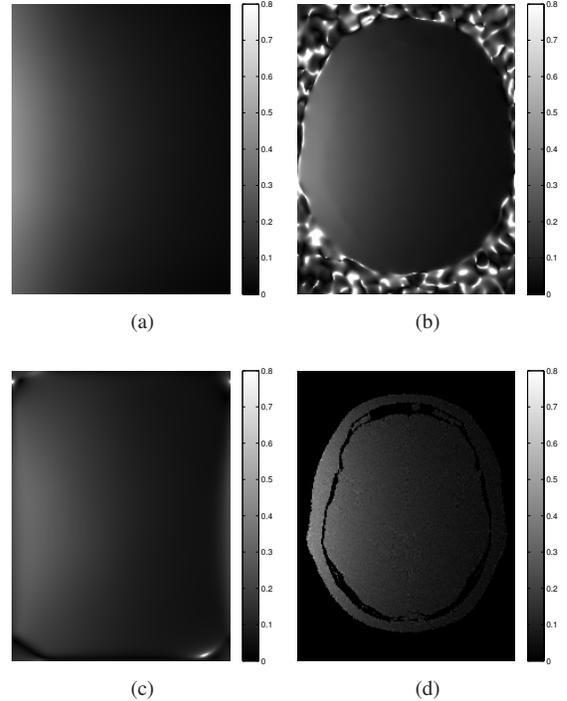


Fig. S-5. Example sensitivity profile estimates for the brain data using (a) the regularized method (b) the ratio of low resolution images method with the center 51×38 samples, (c) the ratio of low resolution images method with the center 13×9 samples, and (d) the conventional ratio method.

of samples from the center of k -space of each coil, zero padding to get 256×192 element matrices (corresponding to a $256 \text{ mm} \times 192 \text{ mm}$ FOV), and reconstructing low resolution body coil and surface coil images using inverse DFTs. Smooth sensitivity estimates were then obtained by taking the ratio of these low resolution images. We present the results for two different amounts of sampling. The first uses the center 13×9 k -space samples resulting in sensitivity estimates that extend smoothly to the image edges. The second uses the center 51×38 samples which was found to provide the best SENSE reconstruction quality for this method. In both cases we applied a Hamming window to the selected k -space data to reduce any Gibbs ringing artifacts. The conventional ratio estimate ($\hat{s}_i = z_i/y_i$) was masked to remove the highly corrupt estimates of the non-object pixels using a binary mask created by thresholding the body coil image. The resulting sensitivity profile estimates for a single, representative coil are presented in Fig. S-5.

As seen in Fig. 6 of the manuscript, the regularized estimate is very close to the true sensitivity, differing only at the corners of the image. The minor discrepancies at the corners of the estimates are in part due to selecting a regularization parameter that emphasized accuracy over the object pixels and their immediate surrounding area as well as from the fact that there is no information about the true sensitivity in this region of the image. The conventional ratio estimate is much noisier over the object pixels and has no extrapolation. Both low resolution ratio estimates are smooth over the object support with varying degrees of extrapolation into the surrounding

TABLE S-2
NRMSES BETWEEN THE TRUE BRAIN IMAGES (STATIONARY AND SHIFTED) AND THE SENSE RECONSTRUCTIONS

	Regularized	Low Res. Ratio		Ratio
		51×38	13×9	
Shift = 0	0.06	0.07	0.16	0.12
Shift = 2 pixels	0.06	0.07	0.16	0.16

regions. However, the implicit smoothing of these methods creates inaccuracies in the estimates near object edges and in areas with predominantly low signal. Furthermore, any voxels significantly beyond the extrapolated regions exhibit large estimation errors. The typical errors that result from Gibbs ringing artifacts [7] have been reduced by the additional windowing.

Data from the four surface coils were then simulated for every other vertical line in k -space. SENSE reconstructions [9] were performed using this undersampled data set and the various sensitivity profile estimates. We restricted the reconstruction to a masked region found by dilating the thresholded body coil image by two pixels. These reconstructions and their differences to the truth are presented in Fig. S-6.

The resulting normalized root-mean-square errors (NRMSE) between the SENSE reconstructions and the truth are presented in Table S-2. The regularized method led to the most accurate SENSE reconstruction in terms of NRMSE as well as the one with the fewest structural artifacts (beyond the amplified noise inherent to SENSE reconstruction). The low resolution ratio method with the center 13×9 samples led to the least accurate SENSE reconstruction. It is clear from the artifacts in the difference image that oversmoothing led to large inaccuracies in the sensitivity profile estimates at the object edges. The low resolution ratio method with the center 51×38 samples led to the second most accurate reconstruction. Although the effects are less severe than for the 13×9 case, there are again structural artifacts in the reconstructions due to inaccurate sensitivity estimates at the object boundaries and low signal regions within the brain. The conventional ratio method also led to significant artifacts in the SENSE reconstruction. Specifically, the lack of smoothing in the sensitivity estimates led to high noise in the SENSE reconstruction, while the lack of extrapolation resulted in aliased object edges within the final reconstruction.

B. Shifted Simulated Brain Data

One possible complication when performing SENSE imaging is if the patient moves between the calibration and acquisition scans. In such cases, poorly extrapolated sensitivity profile estimates will introduce significant artifacts into the reconstruction [10]. To evaluate the different sensitivity estimators under such circumstances, we simulated a set of two times undersampled surface coil images in which the brain was moved two pixels to the right with respect to the coil sensitivities and the field-of-view. We then reconstructed the image using the previously estimated coil sensitivities over an equally shifted masked region. These reconstructions and their differences to the shifted truth are presented in Fig. S-7.

The resulting NRMSEs between the SENSE reconstructions and the true shifted brain are presented in Table S-2. The regularized method again led to the most accurate SENSE reconstruction with similar NRMSE and a lack of structural artifacts. The low resolution ratio method with the center 51×38 samples led to the second most accurate SENSE reconstruction; however, the inaccuracies in the sensitivity estimates at the object edges resulted in larger artifacts due to the shift, particularly at the far right side of the brain. The low resolution ratio method with the center 13×9 samples again resulted in the worst SENSE reconstruction. The shift of two pixels to the right emphasized the inaccuracies in the estimates near the object edges by introducing even larger artifacts (not visible with the current contrast windowing). The SENSE reconstruction based on the conventional ratio method was significantly affected by the shift. In particular, the lack of any extrapolation in the estimated sensitivity profiles resulted in large artifacts within the object support.

C. High SNR Simulated Brain Data

To better illustrate the typical inaccuracies produced by the ratio of low resolution images estimation method, we repeated the previous SENSE reconstruction experiments using simulated brain data with a higher SNR of 20. The specific body coil and four surface coil images are presented in Fig. S-8. We performed sensitivity estimation using the regularized method and the ratio of low resolution images method with 51×38 samples. The resulting estimates for a representative coil are presented in Fig. S-9.

The sensitivity profile estimates are similar to those for the case of lower SNR brain data found in Fig. S-5. The regularized estimate is again very close to the true sensitivity differing only at the corners of the image. The ratio of low resolution images estimate is smooth over the object support and exhibits some extrapolation. However, there are noticeable inaccuracies in areas corresponding to regions of low signal within the brain.

We performed two-fold accelerated SENSE reconstructions with the higher SNR brain data and these sensitivity profile estimates. The results for both the case of no shift between calibration and scan, as well as a two pixel shift, are presented in Fig. S-10. As with the case of low SNR brain data, the reconstructions created using the regularized sensitivity estimate have very low error and no major structural artifacts. Furthermore, the two pixel shift had little effect on the reconstruction quality indicating accurate extrapolation within the estimate. In contrast, the reconstruction created using the low resolution ratio estimates had several large structural artifacts (indicated with a yellow arrow) that were a result of the inaccurate sensitivity profile estimates in regions of low signal. The two pixel shift increased these artifacts indicating inaccurate extrapolation within the sensitivity estimates.

D. Breast Phantom Data

We also compared the sensitivity estimation methods using our breast phantom data from Section III-C of the manuscript. In this case, we estimated the sensitivities of the breast

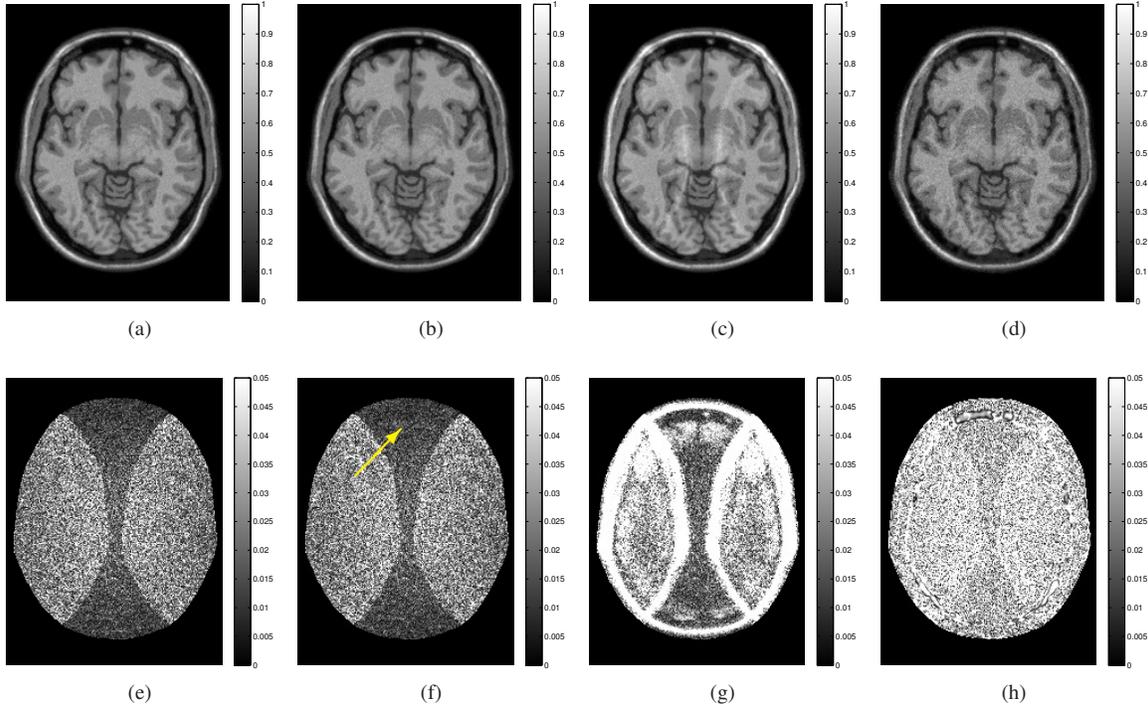


Fig. S-6. Resulting two-fold accelerated SENSE reconstructions for the brain data using (a) the regularized method (b) the ratio of low resolution images method with the center 51×38 samples, (c) the ratio of low resolution images method with the center 13×9 samples, and (d) the conventional ratio method sensitivity profile estimates. The corresponding differences to the truth are presented below in (e – h). The yellow arrow specifies an artifact in the SENSE reconstruction caused by inaccurate sensitivity estimates in a low signal region.

phantom images presented in Fig. 9 using the same four methods as before: the regularized method with $\lambda = 2^7$, the ratio method, and the low resolution ratio method with both the center 77×19 and center 19×5 samples zero padded to 384×96 elements. The resulting estimates for a representative coil are presented in Fig. S-11.

The regularized estimate is smooth over the entire field-of-view and closely matches the general trend in the ratio estimate. The low resolution ratio estimate with the center 77×19 samples is reasonably smooth over the object support with some extrapolation into the surrounding pixels. There are inaccuracies in the estimate near regions of low signal such as at the object edges and over the far right breast. The low resolution ratio estimate with the center 19×5 samples is smoother than the case of 77×19 samples and exhibits greater extrapolation. However, this estimate suffers from oversmoothing and is highly inaccurate at the object edges. Both of the low resolution ratio methods benefited from using a Hamming window to reduce the Gibbs ringing artifacts. The ratio estimate is very noisy over the object pixels and has no extrapolation.

To simulate the minor changes in the data that would occur between a calibration scan and an acquisition scan, we performed a SENSE reconstruction on a neighboring two-dimensional slice of our breast phantom data. The fully sampled body and surface coil images of this slice are presented in Fig. S-12. First, we undersampled the surface coil images by selecting every other vertical line in k -space. As was done for the simulated brain data, we then performed SENSE

reconstructions over a masked region using the previously estimated coil sensitivities. These reconstructions are presented in Fig. S-13.

The SENSE reconstruction resulting from the regularized estimate has very high quality and few visible artifacts when compared to the body coil image in Fig. S-12(a). The reconstruction resulting from the low resolution ratio estimate with the center 77×19 samples appears similar to that of the regularized estimate; however, the inner parts of the breasts are darker than in the body coil and surface coil images. This is largely a result of inaccurate sensitivity estimation in these low signal regions. In addition to the darkening artifact in the low signal regions of the image, the reconstruction resulting from the low resolution ratio estimate with the center 19×5 samples also has aliased edges of the breasts within the object support (indicated by a yellow arrow). These are a result of inaccurate sensitivity estimates at the object edges caused by oversmoothing. The reconstruction resulting from the conventional ratio estimate is very noisy and has several bright artifacts. This is due to inaccurate sensitivity estimation over the low signal pixels within the object support and a lack of extrapolation.

E. Discussion

From these experiments, we conclude that the regularized sensitivity estimation method outlined in (1), although more computationally expensive, provides improved sensitivity estimates for use in SENSE reconstructions compared to other commonly used non-parametric methods. Using a ratio of low

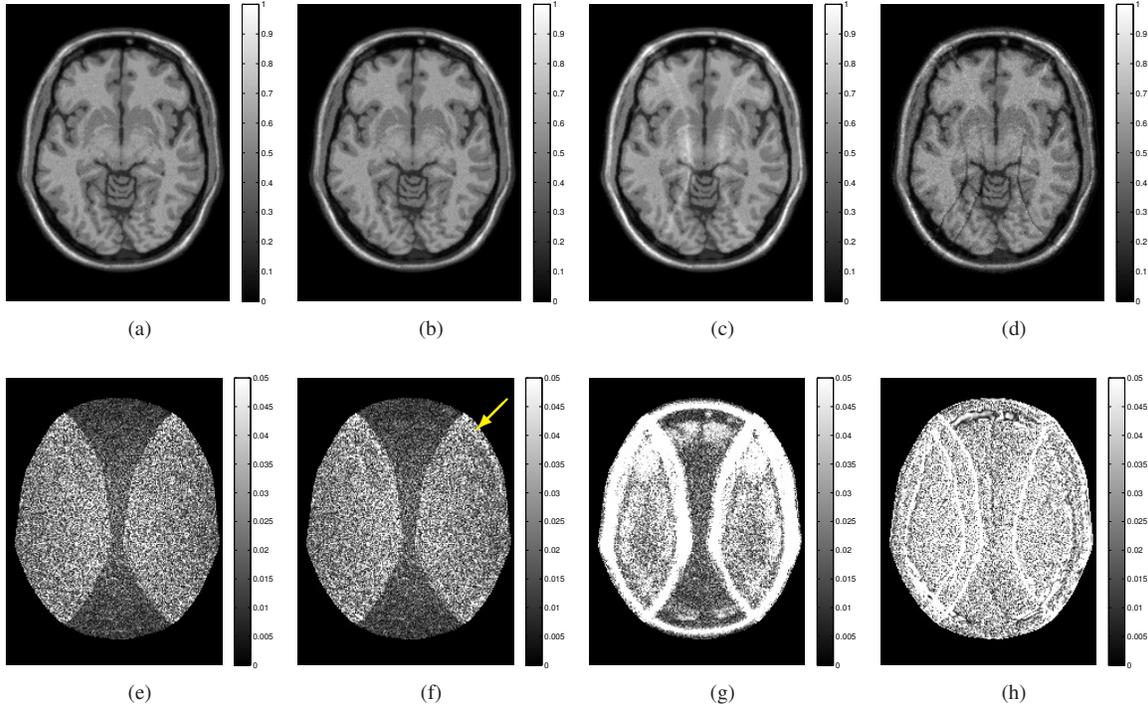


Fig. S-7. Resulting two-fold accelerated SENSE reconstructions of a brain shifted two pixels to the right with respect to Fig. S-6 using the previous (a) regularized method (b) the ratio of low resolution images method with the center 51×38 samples, (c) the ratio of low resolution images method with the center 13×9 samples, and (d) the conventional ratio method sensitivity profile estimates. The corresponding differences to the shifted truth are presented below in (e – h). The yellow arrow indicates an area with increased artifacts due to inaccuracies in the sensitivity estimates at the object edges.

resolution images provides reasonable estimates if the correct number of samples is selected. However, even after windowing to reduce the Gibbs ringing artifacts, these estimates are typically inaccurate at object edges and in areas of low signal. This results in artifacts in the SENSE reconstructions. The lack of smoothing and extrapolation in the conventional ratio method results in SENSE reconstructions that are very noisy and prone to large artifacts due to motion.

S-IV. ESTIMATION OVER A CONVEX HULL MASK

If estimating over a masked region in order to reduce computation time, the mask must be carefully selected to ensure accurate estimates. For data sets with spatially-contiguous support, such as the simulated brain in our manuscript, this is relatively trivial; however, this is not the case for more complicated data sets such as our breast phantom data whose field-of-view (FOV) contains several spatially distinct objects. Due to the underlying physics, the typical coil sensitivity profile should smoothly vary across the entire FOV and generally decrease with distance from the coil. However, using a tight mask isolates the estimate over each object and this can result in large errors for objects that have low signal or only a few pixels. This can be avoided by using a mask consisting of a convex hull containing the spatially distinct objects.

To illustrate this phenomenon, we considered another slice of our breast phantom data, Fig. S-14. This image has a small object to the left of the right breast (indicated by an arrow). We perform regularized estimation over a masked region consisting of spatially distinct objects as well as a masked

region consisting of a convex hull of these points. Fig. S-15 contains the two different masks and their corresponding sensitivity estimates. Fig. S-16 presents line profiles of the absolute value of the sensitivity estimates taken horizontally through the center of the FOV for both estimates.

Comparing the two estimates, it is clear that they are similar for regions with relatively high SNR; however, they differ greatly over the small object next to the right breast. When using a tight mask, the estimated sensitivity in this region is very high in comparison to the nearby breast which does not match the underlying physics. This inaccuracy is a result of the estimate in this region being based on only a few low signal pixels. In contrast, the convex hull estimate is smooth over the entire masked region and the estimate over the small object is more realistic. This is because such a mask enforces smoothness both within and between all of the objects in the FOV. Thus, a convex hull should be used for the estimation mask to avoid inaccuracies in the final estimates.

S-V. CIRCULANT VERSUS NON-CIRCULANT FINITE DIFFERENCING MATRICES

In this section we demonstrate the importance of using a finite differencing matrix for the case of non-periodic boundary conditions (\mathbf{R} or \mathbf{BC}) rather than a finite differencing matrix for the case of periodic boundary conditions (\mathbf{C}) in our cost function. Since $\mathbf{C}^H\mathbf{C}$ is block circulant with circulant blocks, we will refer to the matrix for the case of periodic boundary conditions as the “*circulant matrix*”. In contrast, we will refer

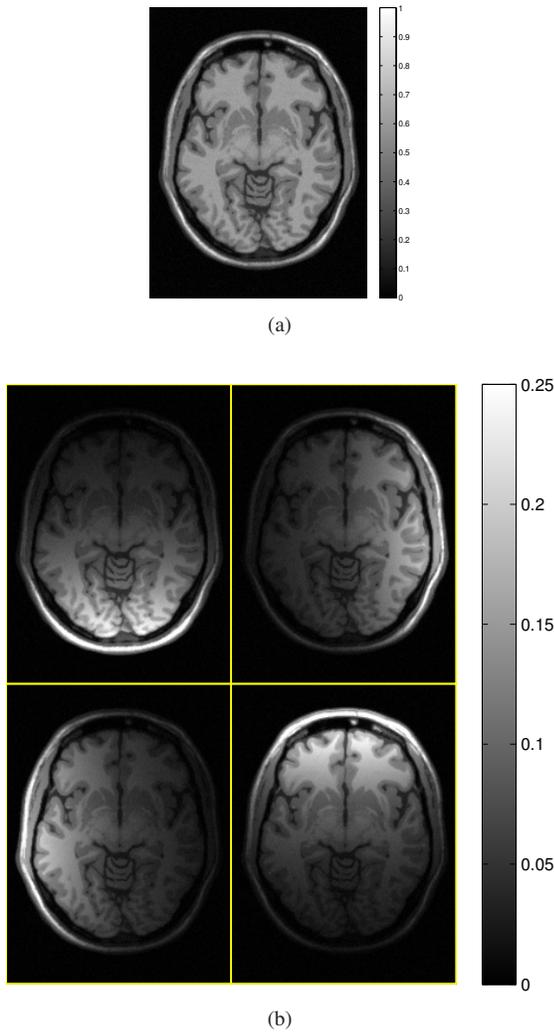


Fig. S-8. The magnitude of the fully sampled (a) body coil and (b) surface coil images for our high SNR simulated brain data.

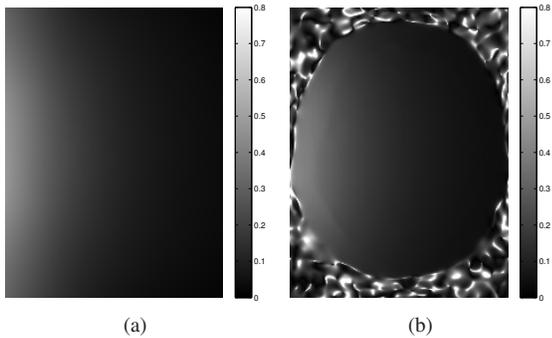


Fig. S-9. Example sensitivity profile estimates found for the high SNR brain data using (a) the regularized method and (b) the ratio of low resolution images method with the center 51×38 samples.

to the matrix for the case of non-periodic boundary conditions as the “*non-circulant matrix*”.

The receive coil is usually placed at or just beyond the boundary of the field-of-view. Since coil sensitivity is a physical phenomenon, its intensity will typically decrease with increased distance from the coil. However, if we use a circulant

finite differencing matrix, we will be penalizing differences in the estimated sensitivities at opposing boundaries of the volume. Since there is often little information about the sensitivity at the edges of the volume, this penalization will result in a sensitivity estimate that dips near the coil and rises at the opposite side of the field-of-view. This is a clear mismatch with the underlying physics of the problem. Furthermore, because of the lack of meaningful information outside of the object voxels, this error will propagate to the estimate at the edges of the object. These errors within (and just outside) the object support can generate significant artifacts in SENSE reconstructions (see Section S-III). Padding the image with zeros will not sufficiently remove this propagated estimation error. Thus, one must use a more realistic modeling assumption and select a non-circulant finite differencing matrix that avoids penalizing between opposite boundaries at the expense of increased complexity. To illustrate these claims, we recreated the estimates found in Section III of our manuscript using both the existing non-circulant finite differencing matrix (\mathbf{R}) and a circulant finite differencing matrix (\mathbf{C}).

A. Simulated Brain Data

We present the results for one coil of the simulated brain data. Fig. S-17 presents the body coil image, true sensitivity, and resulting surface coil image used in this experiment. Fig. S-18 presents the resulting estimates using both the non-circulant and circulant finite differencing matrices, as well as the percentage difference image for each estimate compared to the truth.

As stated before, the estimate using the circulant finite differencing matrix dips before the boundary near the coil and rises at the opposite edge of the field-of-view. This results in significant inaccuracies in the estimate at the image boundaries. In contrast, the estimate using the non-circulant matrix increases smoothly towards the image boundary closest to the coil. The overall estimation error is therefore much smaller and is confined to the outer corners of the image.

Fig. S-19 presents the same estimates as Fig. S-18, but masked in the spatial domain to highlight the error over the object support. In these images, we see that the error in the estimates from the non-circulant matrix has propagated to within the object support. This highly structured inaccuracy will cause large artifacts in SENSE reconstructions. In contrast, the error in the estimate from the circulant matrix is much lower over the entire object support and contains significantly less structure.

B. Padded Simulated Brain Data

We also padded the brain data in Fig. S-17 with zeros to get a 256×256 image (an addition of 32 pixels to both the left and right sides of the image). Fig. S-20 presents the resulting estimates, masked to highlight the error over the object support. Similar inaccuracies to before are present in the circulant matrix estimate, while the estimate using the non-circulant matrix continues to have low error. Thus, the zero padding did not sufficiently mitigate the corruption of the estimate caused by using a circulant matrix.

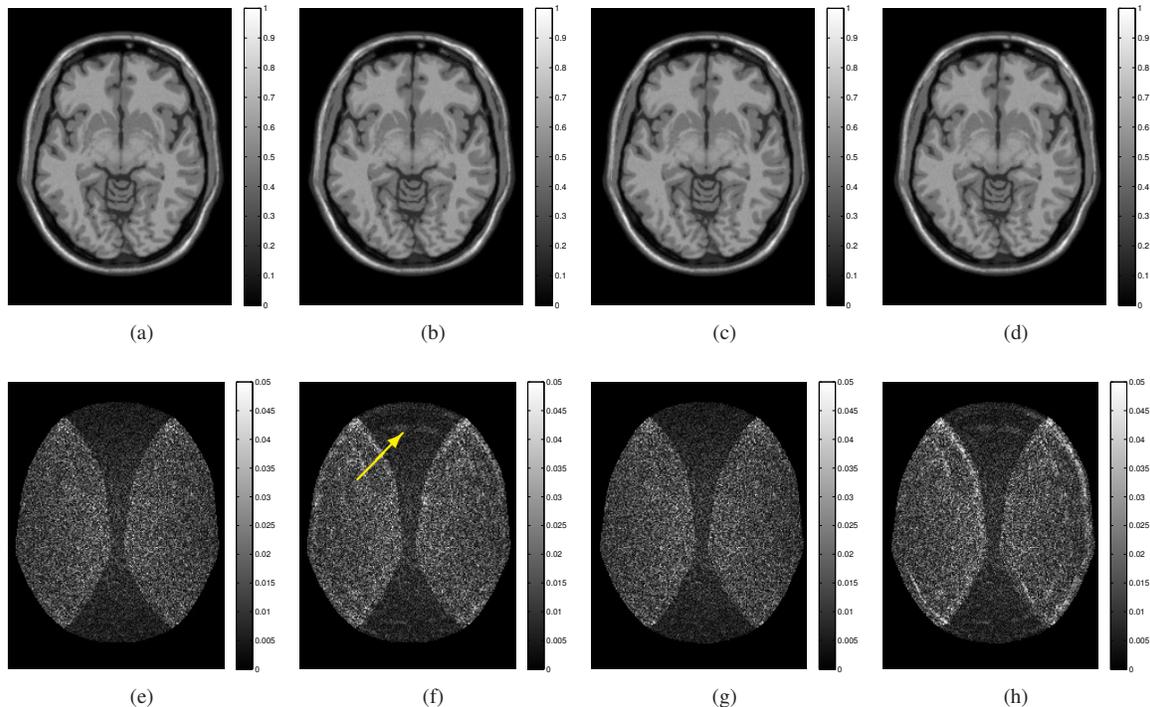


Fig. S-10. Resulting two-fold accelerated SENSE reconstructions of the high SNR simulated brain data. For the case of no shift, (a) and (b) are the reconstructions corresponding to the regularized method and ratio of low resolution images method, respectively. (c) and (d) are the corresponding reconstructions for the case of a two pixel shift. The difference to the truth or shifted truth for each reconstruction is presented below in (e – h).

C. Breast Phantom Data

We performed similar experiments on one coil of our breast phantom data found in Section III-C of the manuscript. Fig. S-21 presents the body coil and surface coil images used in this experiment.

Fig. S-22 presents the resulting estimates using both the non-circulant and circulant finite differencing matrices. As with the brain data, there is an unrealistic dip in the estimate near the coil and a rise at the opposing boundary when using the circulant finite differencing matrix, Fig. S-22(a). The estimate using the non-circulant matrix is more realistic, Fig. S-22(b).

Fig. S-23(a – b) presents the same estimates as Fig. S-22, but masked in the spatial domain to highlight the error over the object support. Fig. S-23(c) shows the difference between these two estimates. From these images, we see that the inaccuracies in the estimate at the boundaries of the image caused by the circulant finite differencing matrix propagated to within the object support. Thus, the need for a non-circulant finite differencing matrix is also evident for the case of real data.

D. Discussion

Using a finite differencing matrix for the case of periodic boundary conditions in our experiments caused substantial errors at the boundaries of the field-of-view and these propagated to within the object support. Furthermore, padding the images did not entirely mitigate the error. As seen in Section S-III, these types of errors can cause significant artifacts in SENSE

reconstructions. However, using a finite differencing matrix for the case of non-periodic boundary conditions avoided these errors. Thus, these experiments illustrate the need to use a non-circulant finite differencing matrix in the regularized estimator of (1).

REFERENCES

- [1] M. J. Allison, S. Ramani, and J. A. Fessler, “Accelerated regularized estimation of regularized MR coil sensitivities using augmented Lagrangian methods,” *IEEE Trans. Med. Imag.*, 2012, To appear as TMI-2012-0142.
- [2] R. H. Chan, “Circulant preconditioners for Hermitian Toeplitz systems,” *SIAM J. Matrix. Anal. Appl.*, vol. 10, no. 4, pp. 542–50, 1989.
- [3] J. Eckstein and D. P. Bertsekas, “On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, Apr. 1992.
- [4] R. Glowinski and P. L. Tallec, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, Soc. Indust. Appl. Math., 1989.
- [5] M. J. Allison and J. A. Fessler, “An augmented Lagrangian method for MR coil sensitivity estimation,” in *Proc. Intl. Soc. Mag. Res. Med.*, 2011, p. 2881.
- [6] M. J. Allison, S. Ramani, and J. A. Fessler, “Regularized MR coil sensitivity estimation using augmented Lagrangian methods,” in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2012, pp. 394–7.
- [7] S. L. Keeling and R. Bammer, “A variational approach to magnetic resonance coil sensitivity estimation,” *Applied Mathematics and Computation*, vol. 158, no. 2, pp. 53–82, Nov. 2004.
- [8] R. Bammer, M. Auer, S. L. Keeling, M. Augustin, L. A. Stables, R. W. Prokesch, R. Stollberger, M. E. Moseley, and F. Fazekas, “Diffusion tensor imaging using single-shot SENSE-EPI,” *Mag. Res. Med.*, vol. 48, no. 1, pp. 128–36, July 2002.
- [9] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: sensitivity encoding for fast MRI,” *Mag. Res. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.
- [10] D. J. Larkman and R. G. Nunes, “Parallel magnetic resonance imaging,” *Phys. Med. Biol.*, vol. 52, no. 7, pp. R15–R55, Apr. 2007.

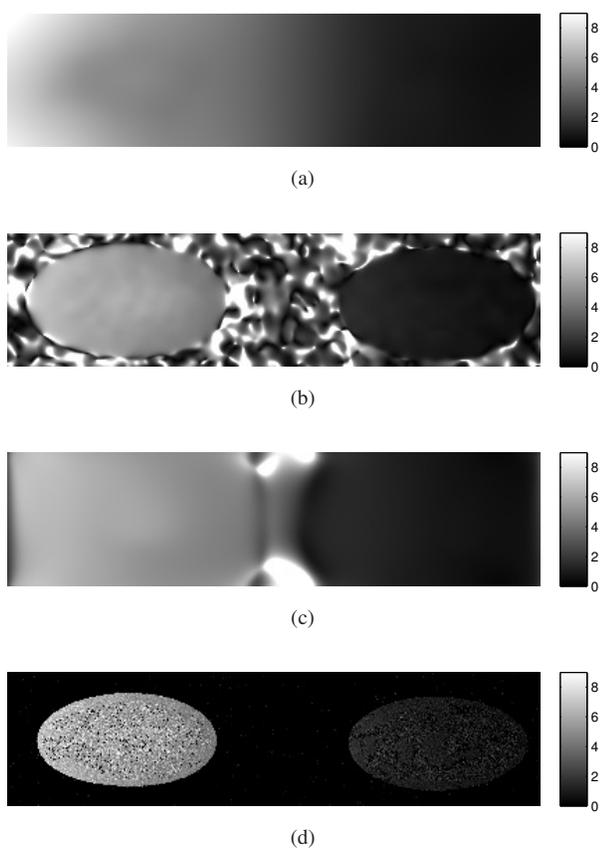


Fig. S-11. Example sensitivity profile estimates found for the breast phantom data using (a) the regularized method (b) the ratio of low resolution images method with the center 77×19 samples, and (c) the ratio of low resolution images method with the center 19×5 samples, and (d) the conventional ratio method.

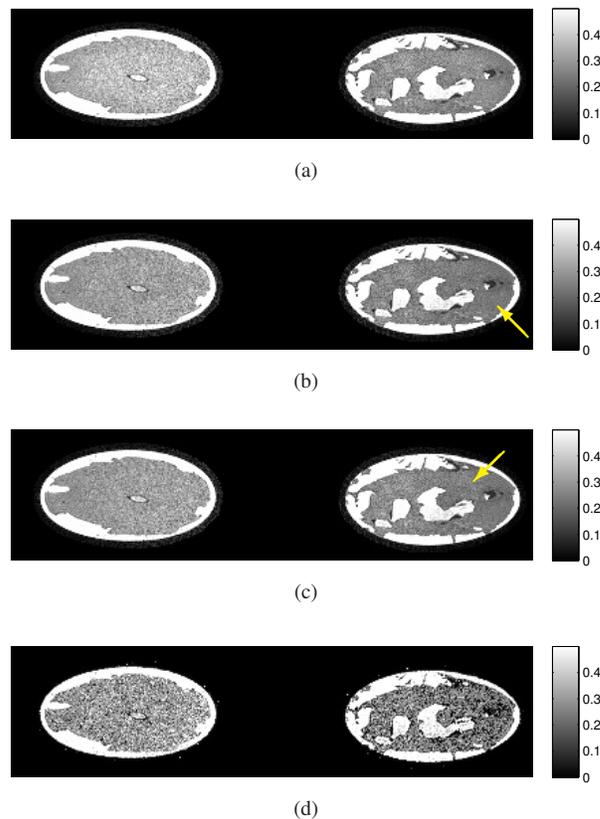


Fig. S-13. Resulting two-fold accelerated SENSE reconstructions of the neighboring slice of breast phantom data (Fig. S-12) using the previous (a) regularized method (b) ratio of low resolution images method with the center 77×19 samples, (c) ratio of low resolution images method with the center 19×5 samples, and (d) conventional ratio method sensitivity profile estimates. The arrow in (b) points to a dark region in the reconstruction, while the arrow in (c) points to a reconstruction artifact caused by inaccurate sensitivity estimation at object edges.

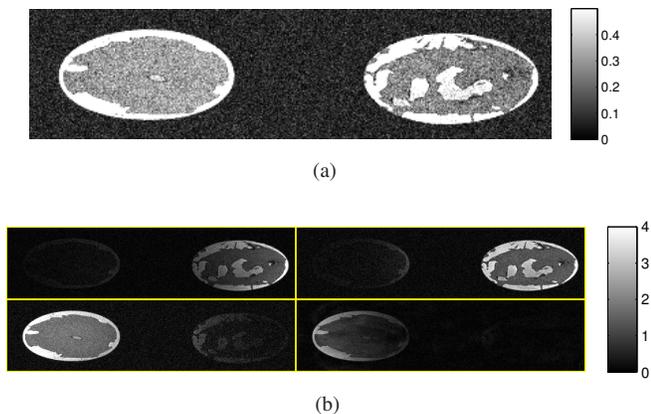


Fig. S-12. The magnitude of the fully sampled (a) body coil and (b) surface coil images for the neighboring two-dimensional slice of our breast phantom data.

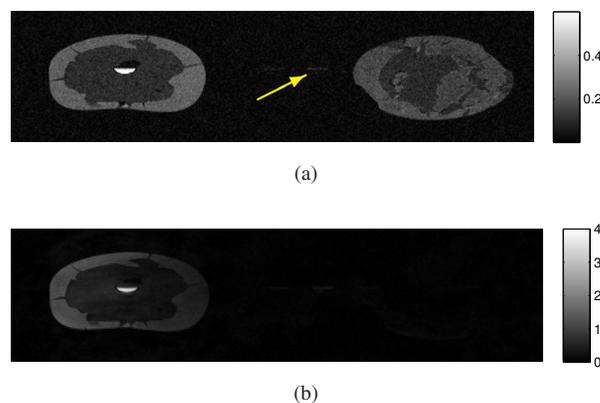


Fig. S-14. The magnitude of the (a) body coil and (b) surface coil image for an additional slice of our breast phantom data. The yellow arrow points to a small object within the FOV.

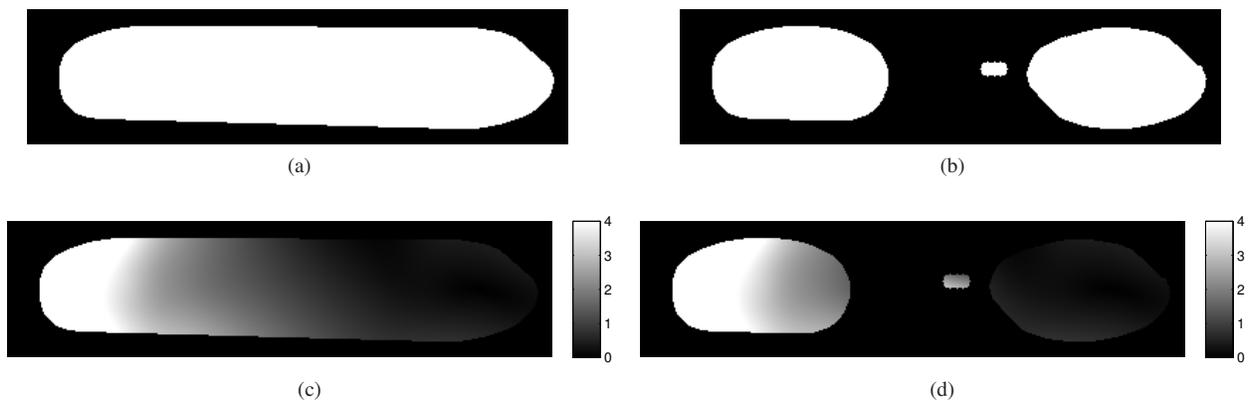


Fig. S-15. The masks for the cases of a (a) convex hull and (b) independent objects. The corresponding regularized estimates over the masked regions for the (c) convex hull and (d) independent objects.

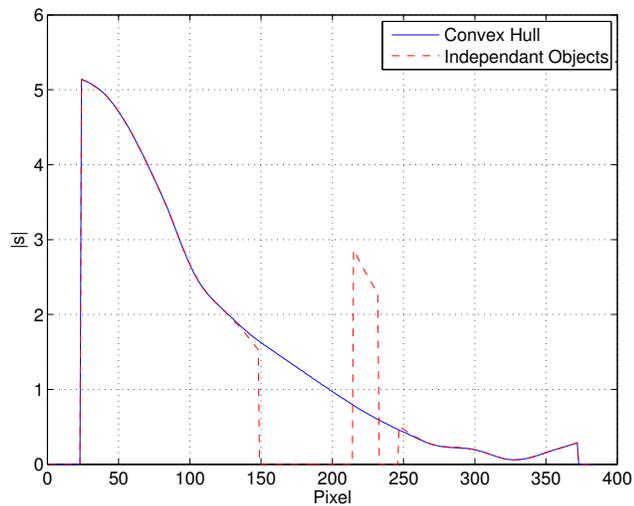


Fig. S-16. Horizontal line profiles taken through the center of the sensitivity estimates presented in Fig. S-15.

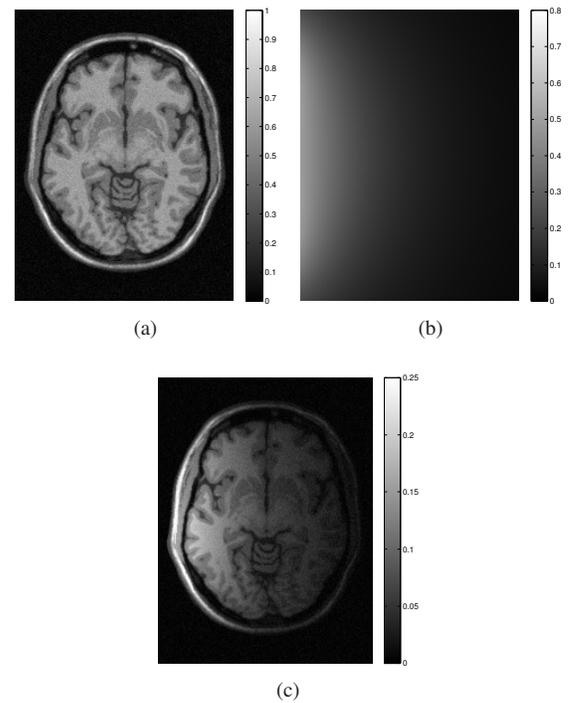


Fig. S-17. The (a) body coil, (b) true coil sensitivity, and (c) resulting surface coil magnitude images for the simulated brain data.

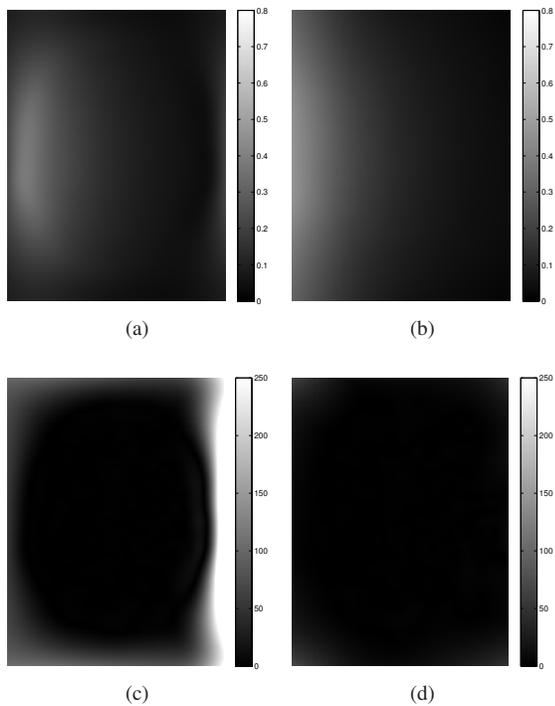


Fig. S-18. The resulting sensitivity estimates for the brain data using a (a) circulant matrix and a (b) non-circulant matrix. The percentage difference between the truth and the estimates from the (c) circulant matrix and the (d) non-circulant matrix are shown below.

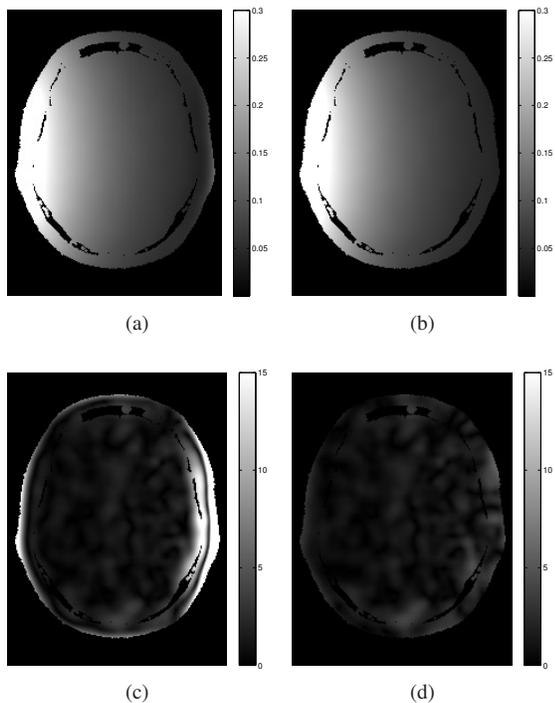


Fig. S-19. The same sensitivity estimates for the brain data as in Fig. S-18 but masked to highlight the error over the object support. (a) and (c) are the resulting estimate and percentage difference to the truth, respectively, resulting from a circulant matrix. (b) and (d) are the same but resulting from a non-circulant matrix.

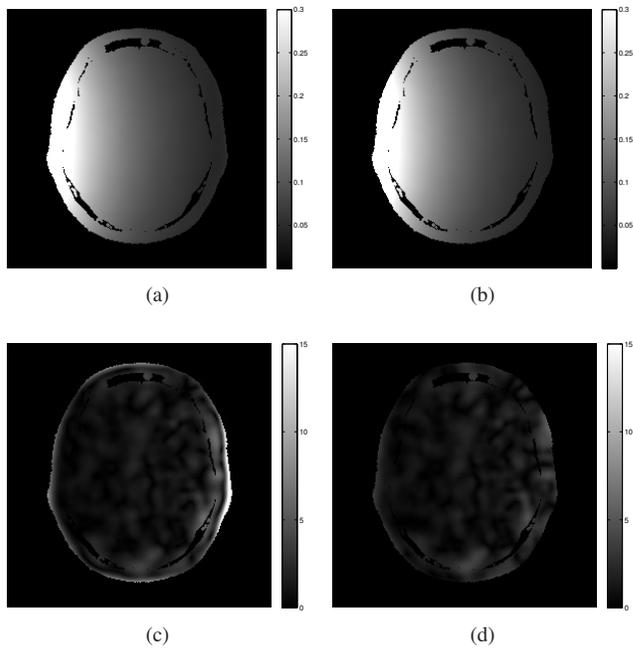


Fig. S-20. The masked sensitivity estimates for padded brain data generated using a (a) circulant matrix and a (b) non-circulant matrix. The masked percentage difference between the truth and the estimates from the (c) circulant matrix and the (d) non-circulant matrix are shown below.

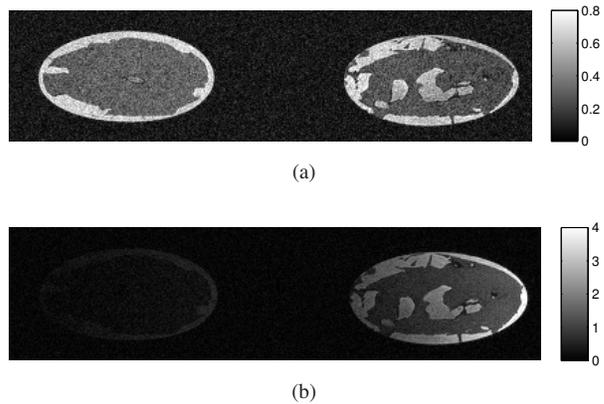


Fig. S-21. The (a) body coil and (c) surface coil magnitude images for the breast phantom data.

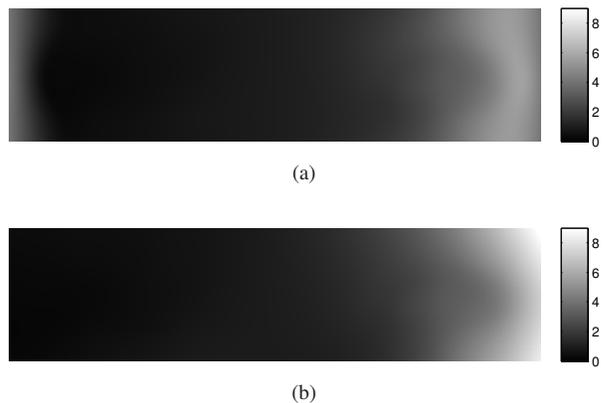


Fig. S-22. The sensitivity estimates for the breast phantom data generated using a (a) circulant matrix and a (b) non-circulant matrix.

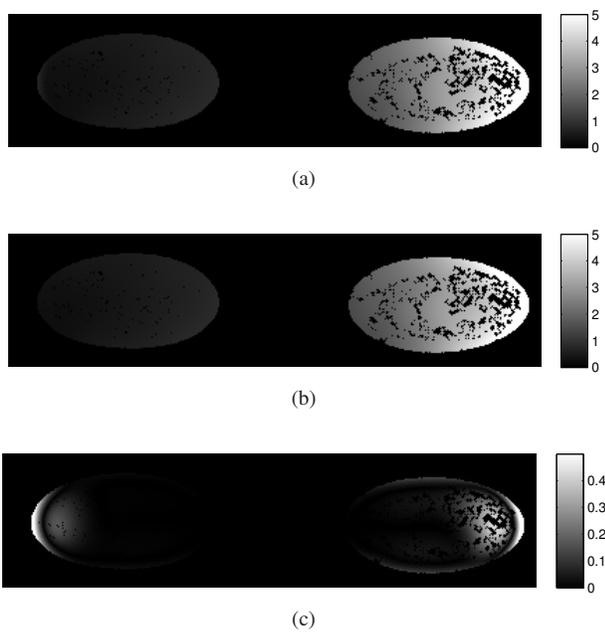


Fig. S-23. The same sensitivity estimates for the breast data as in Fig. S-22 but masked to highlight the error over the object support. The difference between the estimates in (a) and (b) is presented in (c).