Regularized Field Map Estimation in MRI

Amanda K. Funai*, *Student Member, IEEE*, Jeffrey A. Fessler, *Fellow, IEEE*,
Desmond T. B. Yeo, *Student Member, IEEE*, Valur T. Olafsson, *Student Member, IEEE*, and
Douglas C. Noll, *Senior Member, IEEE*

Abstract—In fast magnetic resonance (MR) imaging with long readout times, such as echo-planar imaging (EPI) and spiral scans, it is important to correct for the effects of field inhomogeneity to reduce image distortion and blurring. Such corrections require an accurate *field map*, a map of the off-resonance frequency at each voxel. Standard field map estimation methods yield noisy field maps, particularly in image regions with low spin density. This paper describes regularized methods for field map estimation from two or more MR scans having different echo times. These methods exploit the fact that field maps are generally smooth functions. The methods use algorithms that decrease monotonically a regularized least-squares cost function, even though the problem is highly nonlinear. Results show that the proposed regularized methods significantly improve the quality of field map estimates relative to conventional unregularized methods.

Index Terms—Magnetic field inhomogeneity, susceptibility, penalized-likelihood estimation, phase unwrapping, regularization.

I. INTRODUCTION

MAGNETIC resonance (MR) imaging techniques with long readout times, such as echo-planar imaging (EPI) and spiral scans, suffer from the effects of field inhomogeneity that cause blur and image distortion. To reduce these effects via field-corrected MR image reconstruction, e.g., [1]–[5], one must have available an accurate estimate of the field map. A common approach to measuring field maps is to acquire two scans with different echo times, and then to reconstruct the images (without field correction) from those two scans. The conventional method is then to compute their phase difference and divide by the echo time difference $\triangle_1$. This model makes no account for noise and creates field maps that are very noisy in voxels with low spin density. Section II first introduces this model and then reviews standard approaches for this problem. A limitation of the standard two-scan approach to field mapping is that selecting the echo-time-difference $\triangle_1$ involves a trade off: if $\triangle_1$ is too large, then undesirable phase wrapping will occur, but if $\triangle_1$ is too small, the variance of the field map is large. One way to reduce the variance while also avoiding phase unwrapping procedures is to acquire more than two scans, e.g., one pair with a

small echo time difference and a third scan with a larger echo time difference. By using multiple echo readouts, the scan times may remain reasonable, at least for the modest spatial resolutions needed in functional magnetic resonance imaging (fMRI). Therefore, we present a general model that accommodates more than two scans and describe a regularized least-squares field map estimation method using those scans. Section III shows the improvements both in the estimated field maps and the reconstructed images using multiple scans. This is shown first with simulated results in Section III-A and then using real MR data in Section III-B.

II. MULTIPLE SCAN FIELDMAP ESTIMATION—THEORY

A. Reconstructed Image Model

The usual approach to measuring field maps in MRI is to acquire two scans of the object with slightly different echo times, and then to reconstruct images y^0 and y^1 (without field correction) from those two scans, e.g., [6]–[8]. We assume the following model for those undistorted reconstructed images is:

$$y_j^0 = f_j + \epsilon_j^0$$
$$y_j^1 = f_j e^{\iota \omega_j \triangle_1} + \epsilon_j^1 \tag{1}$$

where $\triangle_1$ denotes the echo-time difference, f_j denotes the underlying complex transverse magnetization in the jth voxel which is a function of the spin density, and ε_j denotes (complex) noise. The goal in field mapping is to estimate an (undistorted) field map, $\omega = (\omega_1, \ldots, \omega_N)$, from y^0 and y^1, whereas $f = (f_1, \ldots, f_N)$ is a nuisance parameter vector. This section reviews the standard approach for this problem, other approaches in the literature, and then describes a new and improved method.

B. Conventional Field Map Estimator

Based on (1), the usual field map estimator $\hat{\omega}_j$ uses the phase difference of the two images, computed as follows [1], [9]:

$$\hat{\omega}_j = \angle \left(y_j^{0*} y_j^1 \right) / \triangle_1. \tag{2}$$

This expression is a method-of-moments approach that would work perfectly in the absence of noise and phase wrapping, within voxels where $|f_j| > 0$. However, (2) can be very sensitive to noise in voxels where the image magnitude $|f_j|$ is small relative to the noise deviations. Furthermore, that estimate ignores our *a priori* knowledge that field maps tend to be smooth or piecewise smooth. Although one could try to smooth the above estimate using a low pass filter, usually many of the $\hat{\omega}_j$ values are severely corrupted so smoothing would further propagate such

errors (see Fig. 2 top right). Instead, we propose below to integrate the smoothing into the estimation of $\boldsymbol{\omega}$ in the first place, rather than trying to "fix" the noise in $\hat{\omega}$ by post processing.

C. Other Field Map Estimators

Although the conventional estimate (2) is most common, other methods for estimating field maps have appeared in the literature.

Different techniques have been proposed that incorporate field map acquisition with image acquisition ([7] for projection reconstruction and [10] for spiral scans). Chen *et al.* in [11] used multiple echos during EPI acquisition and used these distorted scans to create a final corrected undistorted image. Priest *et al.* in [12] used a two-shot EPI technique to obtain a field map for each image; this could prevent changes in the field map due to subject motion from being propagated through an entire fMRI time series.

Stand alone field map acquisition techniques have also been proposed. Windischberger *et al.* [13] used three echos and corrected for phase wrapping by classifying the degree of phase wrapping into seven categories. They then used linear regression to create a field map that is followed by median and Gaussian filtering. Reber *et al.* [14] used ten separate echo times and acquired distorted EPI images. They used a standard phase unwrapping technique of adding multiples of 2π and then spatially smoothed the image with a Gaussian filter. While these techniques both seek to use more echos to increase the accuracy of the field map, they have several disadvantages. Neither are based on a statistical model and, thus, do not consider any noise in developing their estimator. The filtering suggested by both techniques also adds additional blur. Aksit *et al.* [15] used three scans, the first two with a small echo time and no phase unwrapping and the third with a larger echo time. Two techniques were tried: 1) phase unwrapping by using the first two sets of data and 2) taking a Fourier transform to determine the EPI shift experienced. In phantom studies, using three scans yielded half to a third of the error of two scans. Because the estimator uses a linear fit, there is still error in voxels near phase discontinuities and along areas of large susceptibility differences.

An additional technique used to improve the conventional estimate is local (nonlinear) fitting, e.g., [16], [17]. While this can improve the conventional estimate, we desire a more statistical approach.

Our technique is unique in that it uses a statistical model using multiple scans and operates without the constraint of linearity. By using a penalized-likelihood cost function, we can easily adjust the regularization parameter to control the amount of smoothing without any additional filtering step. By using a field map derived from the first two echos as the initialization for the iterative method (assuming the two echos are close enough together), no phase unwrapping is required. Our model also takes into account R_2^* decay, which was ignored in previous multiple echo techniques.

D. Multiple Scan Model

We now generalize the conventional model (1) to the case of multiple scans, i.e., with more than one echo time difference. The reconstructed images are denoted here by $\boldsymbol{y}^0, \ldots, \boldsymbol{y}^L$,

where L is the number of echo time differences. Because we are using multiple echo time differences, R_2^* decay may no longer be negligible and should be included in our model. Our model for these images is

$$y_j^l = f_j e^{\iota \omega_j \triangle_l} e^{-R_j \triangle_l} + \varepsilon_j^l \qquad (3)$$

for $l = 0, \ldots, L$, where $\triangle_l$ denotes the echo time difference of the lth scan relative to the original scan, i.e., $(\triangle_0 = 0)$, where j denotes the voxel number and where R_j denotes the value of R_2^* for the jth voxel. As in most field map estimation methods, this model assumes implicitly there is no motion between the scans. As in (1), f_j denotes the complex transverse magnetization and ε_j^l denotes the (complex) noise. If we choose the $\triangle_l$ values carefully, this data model allows for a scan that is free or largely free of phase wraps but which gives a phase difference lower in signal-to–noise ratio (SNR), as well as scan(s) with wrapped phase but higher in SNR. Including the scan(s) with a larger echo time difference should help reduce noise in $\hat{\omega}_j$, whereas the wrap-free scan helps avoid the need for phase unwrapping tools.

E. Maximum-Likelihood Field Map Estimation

The conventional estimate (2) appears to disregard noise effects, so a natural alternative approach is to estimate $\boldsymbol{\omega}$ using a maximum likelihood (ML) method based on a statistical model for the measurements $\boldsymbol{y}$. In MR, the k-space measurements have zero-mean white Gaussian complex noise [18], and we furthermore assume here that the additive noise values in $\boldsymbol{y}$ in (3) have independent Gaussian distributions[1] with the same variance σ^2. Under these assumptions, the joint log likelihood for $\boldsymbol{f}$ and $\boldsymbol{\omega}$ given $\boldsymbol{y} = (y^0, \ldots, y^L)$ is

$$\log \mathsf{p}(\boldsymbol{y}; \boldsymbol{f}, \boldsymbol{\omega}) = \sum_{l=0}^{L} \log \mathsf{p}(\boldsymbol{y}^l; \boldsymbol{f}, \boldsymbol{\omega})$$
$$\equiv \frac{-1}{2\sigma^2} \sum_{j=1}^{N} \sum_{l=0}^{L} \left| y_j^l - f_j e^{\iota \omega_j \triangle_l} e^{-R_j \triangle_l} \right|^2$$

$$(4)$$

where "$\equiv$" denotes equality to within constants independent of $\boldsymbol{f}$ and $\boldsymbol{\omega}$. If the R_j values were known, the joint ML estimate of $\boldsymbol{f}$ and $\boldsymbol{\omega}$ could be solved by the following minimization problem:

$$\arg\min_{\boldsymbol{\omega} \in \mathbb{R}^N, \boldsymbol{f} \in \mathbb{C}^N} \sum_{j=1}^{N} \left\| \begin{bmatrix} y_j^0 \\ y_j^1 \\ \vdots \\ y_j^L \end{bmatrix} - \begin{bmatrix} 1 \\ e^{\iota \omega_j \triangle_1} e^{-R_j \triangle_1} \\ \vdots \\ e^{\iota \omega_j \triangle_L} e^{-R_j \triangle_L} \end{bmatrix} f_j \right\|^2. \qquad (5)$$

This problem is quadratic in f_j; minimizing over f_j yields the following ML estimate:

$$\hat{f}_j = \frac{\sum_{l=0}^{L} y_j^l e^{-\iota \omega_j \triangle_l} e^{-R_j \triangle_l}}{\sum_{l=0}^{L} e^{-2R_j \triangle_l}}. \qquad (6)$$

---

[1] Independence in image space is an approximation. The noise values in k-space data are statistically independent, but reconstruction may produce correlations, especially in scans with non-Cartesian k-space imaging.

Substituting this estimate back into the cost function (5) and simplifying considerably yields the following cost function used for ML estimation of $\boldsymbol{\omega}$:

$$\Psi_{\text{ML}}(\boldsymbol{\omega}) \equiv \sum_{j=1}^{N} \sum_{m=0}^{L} \sum_{n=0}^{L} \left| y_j^m y_j^n \right| \cdot w_j^{m,n}$$
$$\cdot \left[1 - \cos \left(\angle y_j^n - \angle y_j^m - \omega_j (\triangle_n - \triangle_m) \right) \right] \quad (7)$$

where $w_j^{m,n}$ is a weighting factor that depends on R_2^* as follows:

$$w_j^{m,n} = \frac{e^{-R_j(\triangle_m + \triangle_n)}}{\sum_{l=0}^{L} e^{-2R_j \triangle_l}}. \quad (8)$$

Similar weighting appeared in the weighted phase estimate proposed in [19] for angiography. The ML cost function Ψ_{ML} is periodic, similar to cost functions used in phase unwrapping problems, e.g., [20]. The cost function (7) appears to require either knowledge of or a good estimate of R_2^*. However, we note that

$$\left| \mathsf{E} \left[y_j^m \right] \right| = |f_j| |e^{-R_j \triangle_m}|$$

therefore, hereafter, we approximate $w_j^{m,n}$ as follows:

$$w_j^{m,n} \approx \frac{|y_j^m| |y_j^n|}{\sum_{l=0}^{L} |y_j^l|^2}. \quad (9)$$

This approximation does not require knowledge of R_2^* values.

There is no analytical solution for the minimizer, $\boldsymbol{\omega}$ in (7), except in the $L = 1$ case. Thus, iterative minimization methods are required, even for the ML estimator.

F. Special Case: $L = 1$ (Conventional Two Scans)

In the case where $L = 1$ usually $\triangle_1$ is chosen small enough that we can ignore R_2^* decay (i.e., let $R_2^* = 0$) and the ML cost function in (7) simplifies to

$$\Psi_{\text{ML}} \equiv \sum_{j=1}^{N} |y_j^0 y_j^1| \left[1 - \cos \left(\angle y_j^1 - \angle y_j^0 - \omega_j \triangle_1 \right) \right]. \quad (10)$$

The ML estimate is not unique here due to the possibility of phase wrapping. But ignoring that issue, the ML estimate of $\boldsymbol{\omega}$ is $\hat{\omega}_j = (\angle y_j^1 - \angle y_j^0)/\triangle_1$, because $1 - \cos(t)$ has a minimum at zero. This ML estimate is simply the usual estimate (2) once again to within multiples of 2π. Thus, the usual field mapping method (for $L = 1$) is in fact an ML estimator under the white Gaussian noise model. The more general cost function (7) for the field map ML estimator for $L > 1$ is new to our knowledge.

G. Penalized-Likelihood Field Map Estimation

The ML estimator ignores our *a priori* knowledge that field maps tend to be spatially smooth functions due to the physical nature of main field inhomogeneity and susceptibility effects.[2] (We note that this assumption does not address the presence of signal from fat). A natural approach to incorporating this characteristic is to add a regularizing roughness penalty to the cost function. Here, we regularize only the phase map $\boldsymbol{\omega}$ and not the

magnetization map $\boldsymbol{f}$; we expect $\boldsymbol{f}$ to be far less smooth because it contains anatomical details. Such regularization is equivalent to replacing ML estimation with the following penalized-likelihood (PL) estimator:

$$(\hat{\boldsymbol{\omega}}, \hat{f}) = \arg \max_{\boldsymbol{\omega}, \boldsymbol{f}} \sum_{l=0}^{L} \log \mathsf{p}(\boldsymbol{y}^l; \boldsymbol{f}) - \beta \mathsf{R}(\boldsymbol{\omega})$$

where $\mathsf{R}(\boldsymbol{\omega})$ is a spatial roughness penalty (or log prior in a Bayesian MAP philosophy). Based on (6) and (7), after solving for $\boldsymbol{f}$ and substituting it back in, the resulting regularized cost function has the form

$$\Psi_{\text{PL}}(\boldsymbol{\omega}) \triangleq \Psi_{\text{ML}}(\boldsymbol{\omega}) + \beta \mathsf{R}(\boldsymbol{\omega}) \quad (11)$$

where we use the approximation (9) for Ψ_{ML}. This cost function automatically gives low weight to any voxels where the magnitude $|y_j^m y_j^n|$ is small. For such voxels, the regularization term will have the effect of smoothing or extrapolating the neighboring values. Thus, this approach avoids the phase "outlier" problem that plagues the usual estimate (2) in voxels with low signal magnitude. If $\boldsymbol{\omega}$ corresponds to a $N_1 \times N_2$ field map $\omega_{n,m}$, then a typical regularizing roughness penalty uses the second-order finite differences between horizontal and vertical neighboring voxel values as follows:

$$\mathsf{R}(\boldsymbol{\omega}) = \sum_{n=1}^{N_1-1} \sum_{m=0}^{N_2-1} \psi(2\omega_{n,m} - \omega_{n-1,m} - \omega_{n+1,m})$$
$$+ \sum_{n=0}^{N_1-1} \sum_{m=1}^{N_2-1} \psi(2\omega_{n,m} - \omega_{n,m-1} - \omega_{n,m+1}) \quad (12)$$

where ψ is a convex "potential function." Here, we use the quadratic potential function, $\psi(t) = t^2/2$. In this paper, we used second-order differences for all results; we found that second-order finite differences are preferable to first-order differences because the resulting point spread function (PSF) tails decrease more rapidly even when the FWHM values are identical. A quadratic potential function has the advantage of being differentiable and easy to analyze, especially with Gaussian noise. Although quadratic regularization blurs edges, we assume the field map is smooth, so a more complicated potential function, such as using a Huber function [21], is not considered here.

Usually ψ is differentiable, so we can minimize the cost function $\Psi(\boldsymbol{\omega})$ either by conventional gradient descent methods or by optimization transfer methods [22]–[24]. In particular, in the usual case where $\dot{\psi}(t)/t$ is bounded by unity, then the following iteration is guaranteed to decrease $\Psi(\boldsymbol{\omega})$ monotonically:

$$\boldsymbol{\omega}^{(n+1)} = \boldsymbol{\omega}^{(n)} - \text{diag}\left\{ \frac{1}{d_j + \beta \cdot c} \right\} \nabla \Psi \left(\boldsymbol{\omega}^{(n)} \right) \quad (13)$$

where ∇ is the gradient of the cost function

$$c \triangleq \begin{cases} 4, & \text{regularization with first–order differences} \\ 16, & \text{regularization with second–order differences} \end{cases} \quad (14)$$

and

$$d_j \triangleq \sum_{m=0}^{L} \sum_{n=0}^{L} \left| y_j^m y_j^n \right| \cdot w_j^{m,n} \cdot (\triangle_n - \triangle_m)^2 \quad (15)$$

---

[2]There may be discontinuities at air/water boundaries. Even in this case, sharp boundaries can be problematic if there is motion between scans, further motivating the use of regularization.

using the approximation for w_j shown in (9). For examples in this paper, we used a similar minimization algorithm described in Appendix I because of its faster convergence properties.

To initialize $\omega^{(0)}$, we used the regularized ML estimate (11) based on the first two sets of data y^0 and y^1. We choose the echo times to avoid phase wrapping between these sets of data (this same idea is used in [15] in their three-point method). Therefore, there is no need to apply phase unwrapping algorithms—the algorithm will converge to a local minimizer in the "basin" of the initial estimate [22].

In [25], we considered approximating the $1 - \cos$ term in (11) with its second-order Taylor series to create a penalized weighted least squares (PWLS) cost function. A simplified PWLS approach where the weights were thresholded was also considered. Those models ignore any phase wrap that may occur when evaluating (2). They also have increased error with little computational benefit. Therefore, those simplified methods are not explored further in this paper.

H. Spatial Resolution Analysis of Field Map Estimation

To use the regularized method (11) the user must select the regularization parameter β, which could seem tedious if one used trial-and-error methods. Fortunately, it is particularly simple to analyze the spatial resolution properties for this problem, using the methods in [26] for example. We make the second-order Taylor series approximation for this analysis. The local frequency response of the estimator using second-order finite differences at the jth voxel can be shown to be

$$H(\Omega_X, \Omega_Y) \approx \frac{1}{1 + \frac{\beta}{d_j}\left(\Omega_X^2 + \Omega_Y^2\right)^p} \tag{16}$$

where Ω_X and Ω_Y are the discrete space Fourier transform (DSFT) frequency variables and where $p = 1$ for regularization based on first-order differences and $p = 2$ for second-order finite differences as in (12). (See [27] for related analyses.) From (16) we see that the spatial resolution at each voxel depends on the data through d_j. In areas with small signal magnitudes, there will be more smoothing, as desired. The spatial resolution (16) also depends on the $\triangle_l$ values being used. Data from scans with larger $\triangle_l$ values have lower $\hat{\omega}_j$ variance (see (17) below), and will be smoothed less. However, data from these scans will also be affected by R_2^* decay through $w_j^{m,n}$ if the data is not scaled to compensate for this factor. To simplify selecting β, we normalize the data by the median of the square root of (15) using the approximation (9) for w_j. Normalizing by this factor allows us to create a standard β to FWHM table or graph (e.g., Fig. 1). If this normalization were not applied, a similar figure would need to be calculated with each new data set (or at least with each new set of $\triangle_l$ values) or β would need to be chosen empirically. Normalizing based on the analytical result (16) enables us to use the same β for all scans.

We used the inverse 2-D DSFT of (16) to compute the PSF $h[n, m]$ and tabulate its FWHM as a function of β, assuming the previous corrections were made and that the pixel j has $d_j = 1$. Fig. 1 shows this FWHM as a function of $\log_2(\beta)$, for both $p = 1$ and $p = 2$. The FWHM increases monotonically with β, as expected, although the "knees" in the curve are curious. Nevertheless, one can use this graph to select the appropriate β given
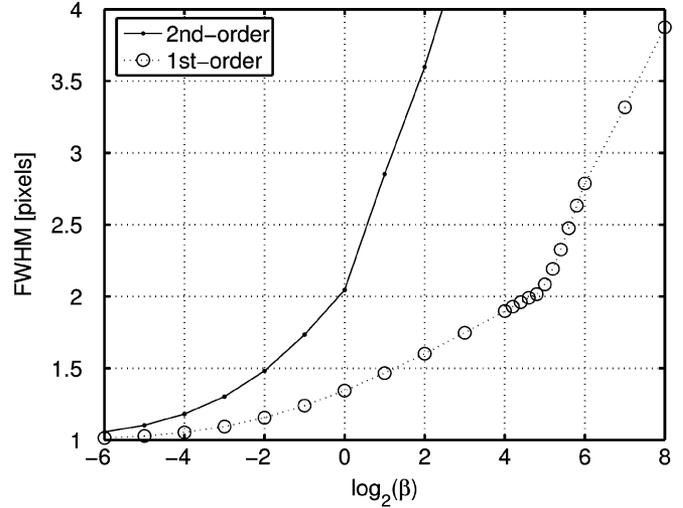

Fig. 1. Angularly averaged FWHM of PSF for field map estimation as a function of $\log_2 \beta$ for $d_j = 1$ in (16).

the desired spatial resolution in the estimated field map. The resulting spatial resolution will be inherently nonuniform, with more smoothing in the regions with low magnitudes and vice versa. One could explore modified regularization methods [26] to make the resolution uniform, but in this application nonuniform resolution seems appropriate since the goals include "interpolating" across signal voids.

I. Qualitative Example: $L = 1$

Fig. 2 shows an example of the data magnitude $|y_j^0|$ and the usual phase estimate based on $L = 1$ (2) which is very noisy. This is real data taken from a 3T MR scanner with $\triangle_1 = 2$ ms. The maximum value of $|\omega_j \cdot \triangle_1|$ is 1.61 radians in nonzero voxels, making the scan free of any phase wraps. Fig. 2 also shows the penalized-likelihood estimate based on (13) using two different values for β and using 150 iterations. Here, we can see the improvement from using a regularized estimator versus the conventional ML estimator. The effect of β on the smoothness of the estimate is also seen. The improvement seen is analyzed quantitatively in Section III. Fig. 2 also shows the effective FWHM (in pixels) of the regularized algorithm based on (16) for both values of β. Most of the image has a FWHM corresponding to the chosen β based on Fig. 1. Areas of low magnitude have a much higher FWHM (such as the sinuses) and areas of high magnitude have the lowest FWHM.

J. Theoretical Improvements Over Two Data Sets

Using more than two sets of data requires a longer data acquisition and also involves choosing the $\triangle_l$ values. Analyzing the theoretical improvements that may be attained by using multiple data sets can help determine when the increased acquisition time is warranted and can guide our choice of the $\triangle_l$ values. Therefore, we calculated the Cramér-Rao bound (CRB) for the model (3). This bound expresses the lowest achievable variance possible for an unbiased estimator based on a given model. Although a biased estimator (the penalized-likelihood estimator) is used in our implementation, the bound quantifies the maximal improvement possible based on the model and allows for a comparison on how close our implementation is to the ideal, unbiased case.
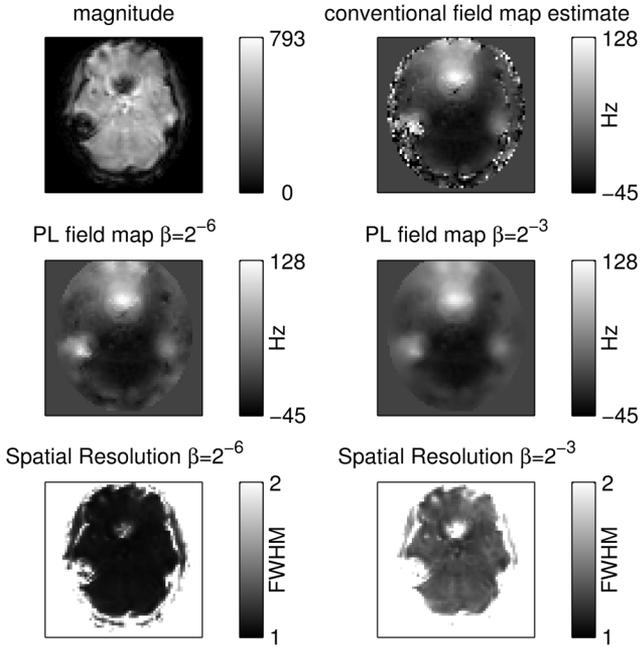
Fig. 2. Top row: magnitude image $|y_j|$, conventional field map estimate (2). Middle row: (field map estimates): penalized-likelihood estimate using (13) with $\beta = 2^{-6}$ (left) and $\beta = 2^{-3}$ (right). Bottom row: Maps of the spatial resolution at each pixel measured by the FWHM for $\beta = 2^{-6}$ (left) and $\beta = 2^{-3}$ (right).

Because there are multiple unknown parameters in these models $\boldsymbol{\theta} = (\omega_j, |f_j|, \angle f_j)$, the multiple parameter CRB must be used. In that case, the matrix CRB is

$$\mathsf{Cov}_{\boldsymbol{\theta}}\{\hat{\boldsymbol{\theta}}\} \geq \mathsf{F}^{-1}(\boldsymbol{\theta})$$

where $\mathsf{F}(\boldsymbol{\theta}) = -\mathsf{E}[\nabla_{\boldsymbol{\theta}}^2 \ln \mathsf{p}(Y; \boldsymbol{\theta})]$ is the Fisher information. Because f_j is a nuisance parameter, we focus on the CRB for the variance of ω_j, although the effect of f_j will be felt through the inversion of the Fisher matrix. For simplicity, we initially set R_2^* to 0 in the CRB derivations shown below.

Applying the CRB to the L echo-time difference model (3) yields, after considerable simplification, the expression

$$\mathsf{Var}_L\{\hat{\omega}_j\} \geq \frac{\sigma^2}{(L+1)\triangle_1^2 |f_j|^2 \lambda_L} \tag{17}$$

where, defining $\alpha_l = \triangle_l / \triangle_1$

$$\lambda_L \triangleq \left(\frac{1}{L+1}\sum_{l=0}^{L}\alpha_l^2\right) - \left(\frac{1}{L+1}\sum_{l=0}^{L}\alpha_l\right)^2.$$

The variance reduces, in general, as L is increased. The expression for λ_L is the "variance" of $\{\alpha_0, \alpha_1, \ldots \alpha_L\}$, measuring the variance between the echo time differences. Increasing the variance (spread) of the $\triangle_l$ values will decrease the overall variance of the field map estimate.

For the $L = 1$ (2 sets of data) model, $\lambda_1 = 1/4$ and (17) simplifies to

$$\mathsf{CRB}_1 \triangleq \frac{2\sigma^2}{\triangle_1^2 |f_j|^2}.$$

As expected, the field map variance decreases when the signal strength $|f_j|$, or echo time difference $\triangle_1$, increase. For an unbiased estimator based on the model (3) with $L = 2$ (three sets of data) one can show

$$\mathsf{CRB}_2 \triangleq \frac{\mathsf{CRB}_1}{4/3\,(\alpha_2^2 - \alpha_2 + 1)}. \tag{18}$$

Interestingly, simply using three scans, but using $\triangle_2 = \triangle_1$ (or $\alpha_2 = 1$), would reduce the variance by only $4/3$.

From (18), increasing α_2 should decrease the variance for an unbiased estimator. Making α_2 arbitrarily large, however, is not advisable for many reasons. A larger α_2 creates more phase-wrapping. Eventually, the wrapping will lead to intra-voxel aliasing and the desired improvement would be unattainable. Another problem with large values of α_l is the effect on the MR pulse sequence length. A large α_2 also causes much more R_2^* decay in the signal as shown in (7). Choosing optimal $\triangle_l$ values requires some knowledge of R_2^* decay. This can be seen more clearly in the CRB bounds for the model (3) with R_2^* decay included. For the $L = 1$ model, one can show

$$\mathsf{Var}_1\{\hat{\omega}_j\} \geq \mathsf{CRB}_1 \cdot \frac{1 + e^{2R_j\triangle_1}}{2}. \tag{19}$$

For the $L = 2$ (three sets of data) model

$$\mathsf{Var}_2\{\hat{\omega}_j\} \geq \frac{\sigma^2}{\triangle_1^2 |f_j|^2} \frac{1 + e^{-2R_j\triangle_1} + e^{-2R_j\triangle_1\alpha_2}}{b} \tag{20}$$

where

$$b \triangleq e^{-2R_j\triangle_1} + \alpha_2^2 e^{-2R_j\triangle_1\alpha_2} \\ + \left(1 + \alpha_2^2 - 2\alpha_2\right) e^{-2R_j\triangle_1(1+\alpha_2)}.$$

Using these expressions, we can optimize the $\triangle_l$ values, which will be inversely proportional to the value of R_2^*. In fact, for $L = 1$, one can show that the optimal choice is $\triangle_1^{\mathrm{opt}} = 1.11/R_j$. Therefore, small values of α_l based on the amount of R_2^* decay expected should be used.

III. EXPERIMENTS

A. Simulation: Comparison of $L = 1$ and $L = 2$ Methods

We compared the $L = 1$ and $L = 2$ methods with two examples. First, we used a simulated Gaussian true field map (Fig. 3) with a magnitude map equal to unity at all points. Second, we simulated a brain example. For the magnitude, we used a simulated normal T1-weighted brain image [28], [29]. We generated a simple field map consisting of a 4.8-cm-diameter sphere of air (centered around the nasal cavity) embedded in water using simple geometrical equations [30], [31], using a slice slightly above the sphere. Fig. 4 shows the field map and magnitude image $|f_j|$. We added complex Gaussian noise at many levels of SNR to the images. For this paper, we used the following definition of SNR:

$$\mathsf{SNR} = 20 \log \frac{\|\boldsymbol{f}\|}{\|\boldsymbol{y}^0 - \boldsymbol{f}\|}. \tag{21}$$

The SNR remains consistent even when varying R_2^*, L, or α.
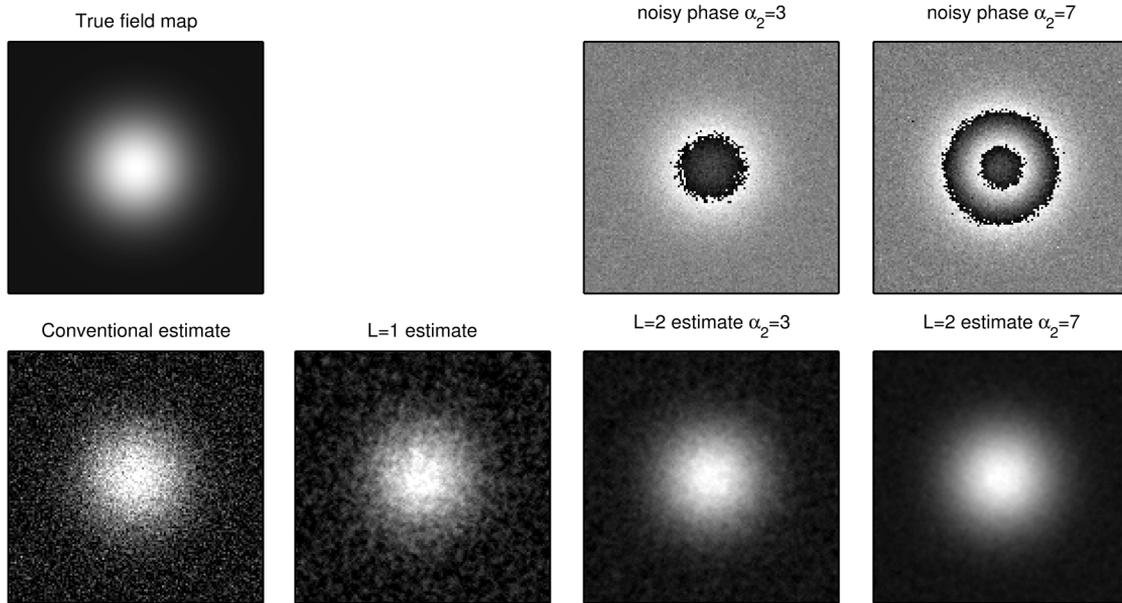
Fig. 3. Top: "True" field map for Gaussian example in Hz; Noisy ($\mathrm{SNR} = 10$ dB) wrapped phase $\angle(y_j^{0*} y_j^2)$ with $\alpha_2 = 3$, Noisy ($\mathrm{SNR} = 10$ dB) wrapped phase with $\alpha_2 = 7$. Bottom: Conventional estimate for $L = 1$, PL estimates for $L = 1$, $L = 2$ with $\alpha_2 = 3$, and $L = 2$ with $\alpha_2 = 7$. All field maps and estimates are shown on a colormap of $[-10 \ 128]$ Hz. The wrapped phase images are shown on a colormap of $[-\pi \ \pi]$.
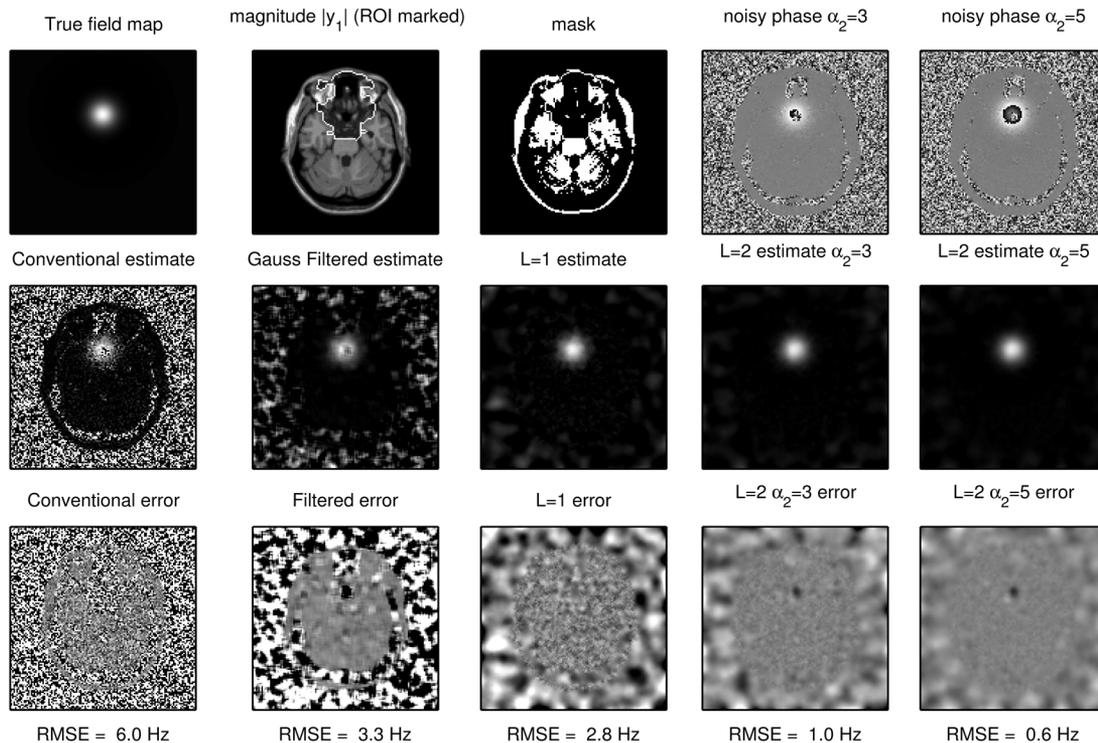

Fig. 4. Top: True field map and magnitude for brain example and mask, ($\mathrm{SNR} = 8.5$ dB) wrapped phase for $\alpha_2 = 3$ and $\alpha_2 = 5$ images. Center and Bottom: Conventional, Conventional convolved with a Gaussian filter, PL with two sets ($L = 1$), and PL with three sets ($L = 2$) for both $\alpha_2 = 3$ and $\alpha_2 = 5$ estimated field maps and their respective errors and RMSE. The wrapped phase images are shown on a colormap of $[-\pi \ \pi]$. All field maps and estimates are shown on a colormap of $[-2 \ 100]$ Hz. Field map errors are shown on a colormap of $[-15 \ 15]$ Hz.

We used $\triangle_1 = 2$ msec for both cases. For the $L = 2$ case we also varied α_2 to produce several $\triangle_2$ values. We used a uniform value of $R_2^* = 20 \ \mathrm{sec}^{-1}$ in generating our simulations.

The field map was reconstructed using the penalized-likelihood method (11) using normalization as described in Section II-H for both $L = 1$ and $L = 2$. The algorithm (13) was run at each SNR level for the $L = 1$ case and for the $L = 2$ case of data with varying values of α_2 using five realizations. We ran 300 iterations of the algorithm, using $\beta = 2^{-3}$.

We also applied the conventional estimator to our data. To reduce the noise, we convolved the conventional estimate with Gaussian filters of varying widths

($\sigma = 0.0625, 0.1250, \ldots, 3.125$). We chose the "optimal" σ based on the minimum masked root mean square error (RMSE). Choosing the optimal σ using the true field map gives the conventional estimate an advantage in this example unavailable in practice.

The RMSE (in Hertz) was computed between the "true" field map and the field map reconstructed using the PL method (11) and the conventional estimate. This RMSE was calculated in a masked region (pixels with magnitudes at least 20% of the maximum true magnitude).

Fig. 3 shows an example of the PL with $L = 1$ estimate compared to the PL with $L = 2$ estimate at $\alpha_2 = 3$ and $\alpha_2 = 7$ at an SNR of 10 dB. Qualitatively, we can see improvements with increases in both L and α_2. Fig. 4 shows similar results for the brain example.

The largest errors in these field maps occur where the magnitude is smallest. The RMSE is much higher using only the conventional method. We also calculated the RMSE in the sinus region of the brain (the region of interest (ROI) is shown in Fig. 4). We chose this ROI because the low magnitude makes the field map difficult to estimate here although the field inhomogeneity is also greatest here. The RMSE in this ROI was 61.1 Hz for the conventional estimate, 11.6 Hz for the Gaussian filtered estimate, 3.4 Hz for the $L = 1$ regularized estimate, and 1.9 Hz for the $L = 2$, $\alpha_2 = 3$ regularized estimate and 1.7 Hz for the $L = 2$, $\alpha_2 = 5$ regularized estimate. Overall, the filtered conventional estimate performed similar to the PL method with $L = 1$ over the masked region, but had higher error in the ROI. The PL method with $L = 2$ showed a decreased error in both the masked region and the ROI. We would expect even higher improvement over any practical Gaussian filtered estimate because a suboptimal σ would be used. The proposed regularized estimators are more accurate in pixels with low magnitude. Adding additional scans ($L > 1$) makes the PL estimate even more accurate.

Fig. 5 shows the improvement (defined as the RMSE for PL estimate with $L = 1$ divided by the error for PL estimate with $L = 2$) gained by using an additional set of data for the Gaussian example. For comparison, we also plotted the predicted improvement, given by the square root of the ratio of the expressions (19) and (20). The experimental gains are actually higher than the improvements anticipated as shown by the dotted lines (the predicted improvement) for some SNR values. Because this is a ratio of RMSEs and the amount of bias can vary between $L = 1$ and $L = 2$, the unbiased CRB provides a benchmark of expected ratios rather than an exact upper limit. Also, recall that (19) and (20) considered R_2^* to be a known value when, in fact, R_2^* is unknown and approximated through (9). The RMSE is low (in voxels with large magnitudes) at high SNRs using either $L = 1$ or $L = 2$. At lower SNRs, however, including in voxels with low magnitudes, using $L = 2$ and higher values of α_2 greatly reduces RMSE. We repeated these simulations with $R_2^* = 0$ (results not shown) and the empirical improvement almost exactly matched (18).

Fig. 6 shows the improvement gained by using an additional set of data for the brain image. For a low SNR (for example 10 dB), the improvements are close to expected. The brain image
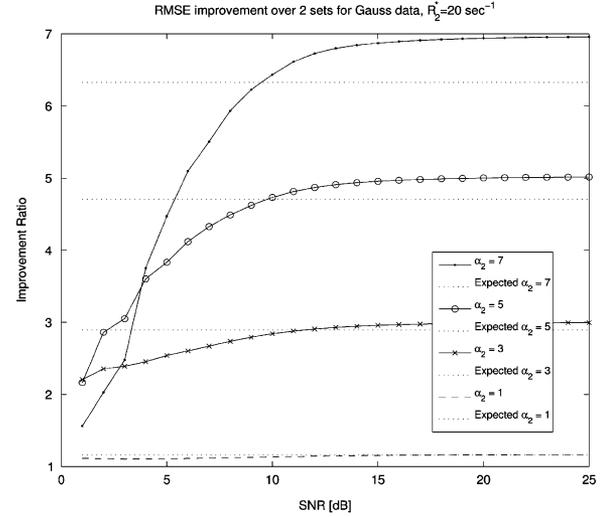

Fig. 5. Improvement in the RMSE for the Gaussian example by using three data sets rather than two sets. Expected improvements shown by dotted lines.
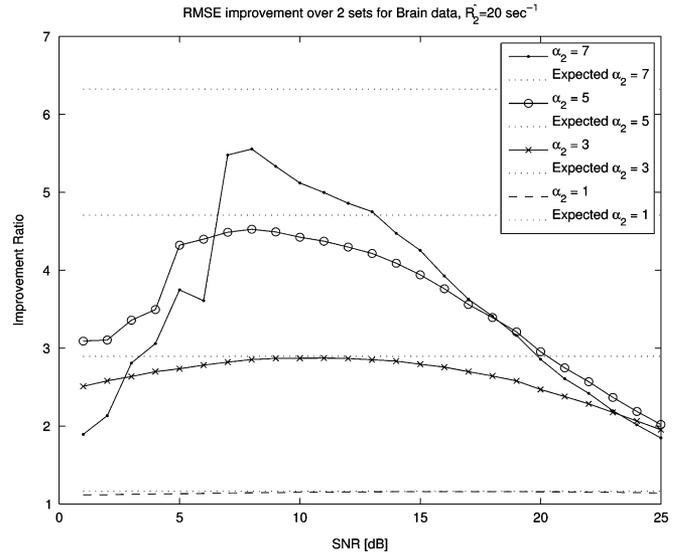

Fig. 6. Improvement in the RMSE for the brain example by using three data sets rather than two sets. Expected improvements shown by dotted lines.

has some areas where the magnitude is very low, making estimation using any method quite challenging. In addition, the field map phase itself is less smooth than in the Gaussian case, making the estimation more difficult. For a higher SNR (for example 20 dB), the three-set case still outperforms the two-set case substantially but by less than predicted by (18).

The RMSE has components of both bias error and variance in it, as shown below

$$\mathrm{RMSE(X)} = \sqrt{\mathrm{Var}\{X\} + \mathrm{bias}^2(X)}.$$

Therefore, we analyzed the bias and the standard deviation at a single representative $\mathrm{SNR} = 20$ dB and at $\alpha_2 = 1, 2, \ldots, 7$ using 500 iterations and 100 realizations for each factor. Fig. 7 compares the standard deviation for each α_2 relative to that at $\alpha_2 = 1$ and the empirical improvements were compared to
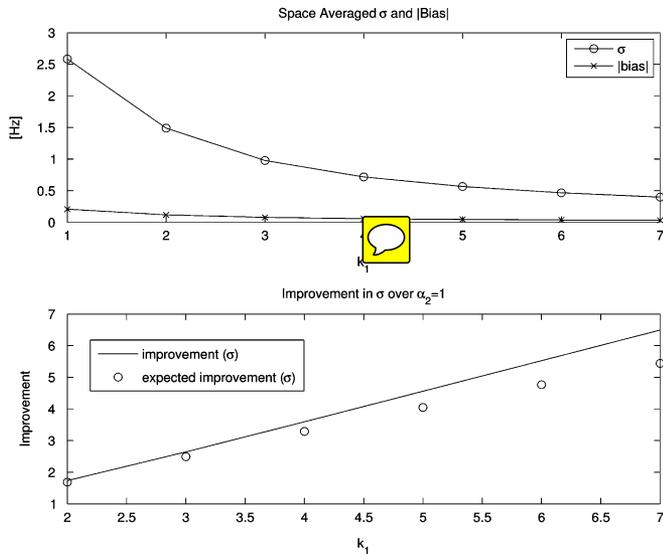
Fig. 7. Bias and RMSE improvement for Gaussian example. Top: Space-averaged σ and absolute bias for several α_2 values; Bottom: RMSE improvement, empirical and expected, over $\alpha_2 = 1$ for several α_2 values.
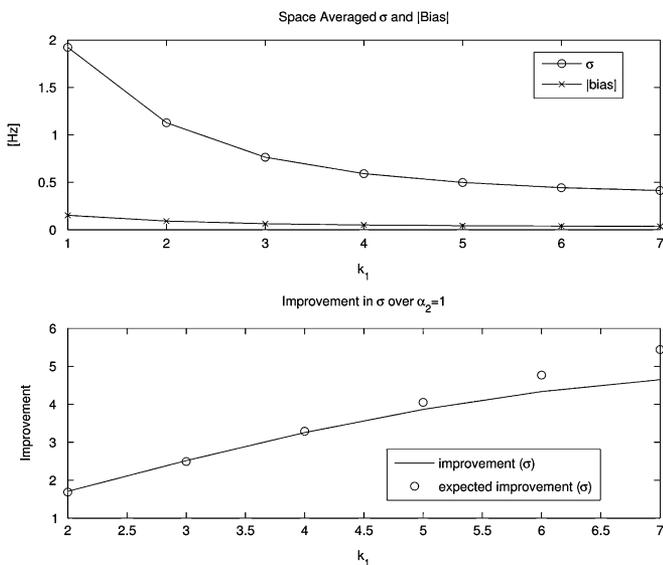

Fig. 8. Bias and RMSE improvement for brain example. Top: Space-averaged σ and absolute bias for several α_2 values; Bottom: RMSE improvement, empirical and expected, over $\alpha_2 = 1$ for several α_2 values.

those predicted by the CRB (20) for the Gaussian example. As expected, the improvements in variance are very close to predicted. Here, the bias is also very low at all levels of SNR—explaining the improvement seen in RMSE in Fig. 7.

Fig. 8 shows the bias and standard deviation for a single $\mathrm{SNR} = 20$ for the brain example. The empirical variances were close to those expected. The bias, however, introduced in part by the regularization, was nearly constant (independent of α). So for large values of α_2, the bias begins to dominate the variance in RMSE calculations, explaining Fig. 6.

Overall, the variance reductions in both examples due to using three echo times were close to the results predicted by the CRB. For low values of α_2 (i.e., five or less), the expected benefit using $L > 1$ holds even with a moderate value of R_2^*. The RMSE

reductions are largest at lower SNRs. For phase estimation, the local SNR depends on the spin density of each voxel as seen in (17). Voxels with lower spin density effectively have lower SNR. It is precisely in these voxels where using three or more scans has the greatest benefit.

B. MR Phantom Data: Application to Spiral Trajectories

To illustrate how improved field map estimation leads to improved reconstructed images, we used field maps produced by the conventional method (2) and produced by the PL method with three scans (11) to correct real spiral MR data for field inhomogeneities. We imaged a phantom with large field inhomogeneity. We used a spiral-out trajectory with a TE of 30 ms, TR of 2 s, and a flip angle of $90°$. We took six slices spaced 5 cm apart over the 15 cm field-of-view. First, we collected data to create the field maps (using eight interleaves to minimize the effect of the field inhomogeneity) at the original 30 ms, as well as at 32 ms ($\triangle_1 = 2$ ms) and at 34 ms ($\triangle_2 = 4$ ms) and at 40 ms ($\triangle_2 = 10$ ms). We took ten realizations for each echo difference. We reconstructed iteratively the resulting 64×64 pixel images in a masked region using [32]. Then, we used these images to create (for each slice) a conventional field map (2), a conventional field map blurred with a Gaussian filter, a PL field map with $L = 1$, a PL field map with $L = 2$, and $\alpha_2 = 2$, a PL field map with $L = 2$ and $\alpha_2 = 5$, and a PL field map with $L = 3$, (11). We used $\beta = 2^{-6}$ for the regularized iterative algorithm and $\sigma = 0.5$ for the Gaussian filter approach, approximately matching the FWHM of the two approaches. Finally, we collected one-shot spiral out data with $\mathrm{TE} = 30$ ms. This scan is thus much more affected by final inhomogeneity. We collected two realizations and then averaged them in k-space. We first reconstructed this data iteratively without a field map as in [32]. Uncorrelated field inhomogeneity causes a blurred image for spiral trajectories. Finally, we iteratively reconstructed this one-shot data with each of the field maps previously created as in [2].

Fig. 9 shows one representative slice. The regularized field maps are less noisy than the conventional one, especially in areas of low magnitude and along the edges. Fig. 9 illustrates the blur and distortion in the one-shot image reconstructed without a field map. The images reconstructed with a field map do not have this blur. Nevertheless, a noisy field map can cause error in the reconstructed image. For example, in Fig. 9, the image reconstructed with the conventional field map shows more artifacts than the eight-shot data or either of the images reconstructed with regularized field maps. Using the eight-shot data as "truth," we computed the normalized root mean square error (NRMSE) of each image and Table I shows the mean and variance over the ten realizations. We include data from two representative slices to show a range of values, although slice three is not shown. In addition, we calculated the NRMSE in the one-shot reconstructed images in pixels where the magnitude is less than 0.2 times the maximum pixel value of the eight-shot reconstructed image to see if the regularized field maps reduce errors in areas of the image with low magnitude. This is also reported in Table I. We use the norm of the eight-shot 30-ms image for normalization. The regularized iterative PL methods have a lower RMSE and much less variability than the other
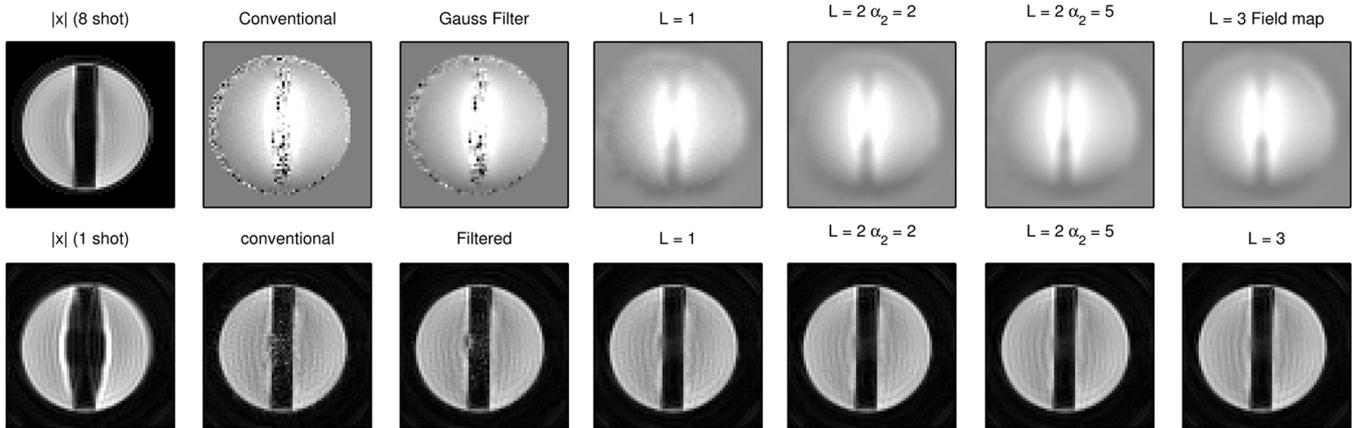
Fig. 9. First Slice—Top: Reconstructed eight-shot image, Conventional field map, Gaussian filtered field map, regularized field map $L = 1$, regularized field map $L = 2; \alpha_2 = 2$, regularized field map $L = 2 \alpha_2 = 5$, regularized field map $L = 3$. The field maps are displayed with a common color scale from -35 Hz to 50 Hz; Bottom: Reconstructed one-shot image with no field map and with each of the field maps above. The images are all on the same color scale. These are all from one representative realization.

TABLE I
PHANTOM NRMSE FOR TWO REPRESENTATIVE SLICES

Phantom NRMSE (%) for one realization							
	Slice One				Slice Three		
	Whole Image		Low Magnitude		Whole Image		Low Magnitude	
	μ	σ	μ	σ	μ	σ	μ	σ
No Field Map	31.1	-	4.8	-	20.4	-	2.9	-
Conventional	15.0	0.5	6.8	0.6	15.5	0.2	2.2	0.1
Gauss Filter	14.3	0.4	6.1	0.4	15.1	0.2	1.9	0.08
L=1	13.0	0.4	4.0	0.4	15.2	0.2	1.8	0.04
L=2 $\alpha_2 = 2$	13.1	0.4	4.1	0.4	14.8	0.1	1.8	0.03
L=2 $\alpha_2 = 5$	13.5	0.08	4.3	0.2	14.6	0.02	1.8	0.01
L=3	13.5	0.09	4.4	0.1	14.6	0.02	1.8	0.01

methods. Therefore, these regularized methods (especially with more than one echo time) give a very reliable estimate of the field map with little variability.

IV. DISCUSSION

We described a regularized method for field map estimation using two or more scans: the penalized-likelihood method (11). This method yields field maps that interpolate smoothly over regions with low spin density, thereby avoiding phase outliers that plague the conventional estimate (2). The method has been used with $L = 1$ (without full description) in [3], [33], [34].

Our analysis also shows that the conventional estimate (2) is in fact the ML estimate, a property that has previously gone unnoticed to our knowledge.

We also analyzed the spatial resolution properties of this method, leading to a practical procedure for choosing the regularization parameter to achieve a given desired spatial resolution.

We studied the CRB on the variance of the estimate for this method and found that our empirical simulation results for the PL method compared favorably, showing a reduction in the RMSE in comparison to using only two scans.

We collected real MR phantom data and created conventional and PL estimates of the field map which were used to reconstruct final images. The PL estimate reduces image artifacts caused by

the field inhomogeneity and has a reduced RMSE, especially in areas of very low magnitude where the conventional estimate has many errors. Omitting or using a poor field map estimate for image reconstruction can dramatically affect the final image quality.

As noted in Section II-D, our cost function assumes, as do most other field map estimation problems, that there is no motion between scans. While our analysis indicated that a larger L is better in terms of variance, motion could be a problem during the larger time required for L echo time differences. Practically, $L = 1$ or $L = 2$ are the most likely choices for L and here motion is less likely to be an issue. If a larger number of echo differences are desired, then the cost function could be further generalized to include a joint estimation of the field map and rigid motion parameters.

We have focused here on the case of a single receive coil. It is straightforward to generalize the method for phased array coils, cf. [35].

Although we did not estimate R_2^*, we used a simple weighting (9) in our algorithm to partially account for R_2^* decay; the improvements seen over estimation with two scans are still large, especially when using a small value of α_2.

While this method assumed the first two echo time differences were close enough to prevent phase wrapping, this method could, with proper intialization, extend to data with larger echo time differences and some phase wrapping. This is especially

interesting at higher field strengths where wrapping still exists at low echo time differences.

Overall, this method has potential to be a reliable estimator for MR field maps, able to utilize many scans to produce a good estimate. The general penalized-likelihood approach in this work is also applicable to estimating other parametric maps in MRI, such as relaxation maps [36] and sensitivity maps [37]. It may also be useful for phase unwrapping problems with noisy data. In some cases, it may be preferable to use edge-preserving regularization in (12), such as the Huber potential function [38].

Ultimately, this method is a tool that may help answer the main question of field mapping: how to best allocate scan time to achieve the most accurate field map. The preliminary CRB analysis guides choice of echo times given a set number of scans. In future work, we wish to further explore the relationship between number of echoes, signal to noise ratio, and spatial resolution.

APPENDIX I
MINIMIZATION ALGORITHMS

To minimize the cost function (11) developed in this paper, we need a method that will decrease it monotonically. The simple minimization algorithm shown in (13) is guaranteed to decrease $\Psi(\boldsymbol{\omega})$ monotonically; the proof that ensures monotonicity uses the fact that the second derivative of $1 - \cos t$ is bounded above by unity. This algorithm will converge to a local minimizer of $\Psi(\boldsymbol{\omega})$ within the "basin" that contains the initial estimate [39].

However, this simple minimization algorithm shown in (13) is only one possible option to minimize the cost function given in (11). In our implementation, we used an optimization transfer approach to refine the iterative algorithm [23], [39]. First express (11) as shown below

$$\Psi(\boldsymbol{\omega}) \triangleq \sum_{j=1}^{N} \sum_{m=0}^{L} \sum_{l=0}^{L} \varphi_{jml}(\omega_j) + \beta R(\boldsymbol{\omega}) \qquad (22)$$

where we define

$$\varphi_{jml}(\omega) \triangleq \left| y_j^m y_j^l \right| w^{m,l} \varphi \left(\omega(\triangle_l - \triangle_m) + \angle y_j^m - \angle y_j^l \right)$$

with

$$\varphi(t) \triangleq 1 - \cos(t).$$

To minimize this cost function, we adopt an optimization transfer approach, for which we need a surrogate function for $\varphi(s)$. In particular, we use the following parabola surrogate for φ:

$$\varphi(t) \leq q(t; s) \triangleq \varphi(s) + \dot{\varphi}(s)(t - s) + \frac{1}{2} \kappa_\varphi(\{s\}_{2\pi})(t - s)^2$$

where $\{s\}_{2\pi}$ denotes the principle value of s. Huber stated that parabola surrogate functions (which he called a comparison function) exist for φ that satisfy Huber's conditions [21, pp. 184–5]; the functions must be differentiable, symmetric, and have curvatures $(\kappa_\varphi(s))$ that are bounded and monotone non-increasing for $s > 0$. For $|s| \leq \pi, \varphi(s)$ shown above satisfies Huber's conditions. We note
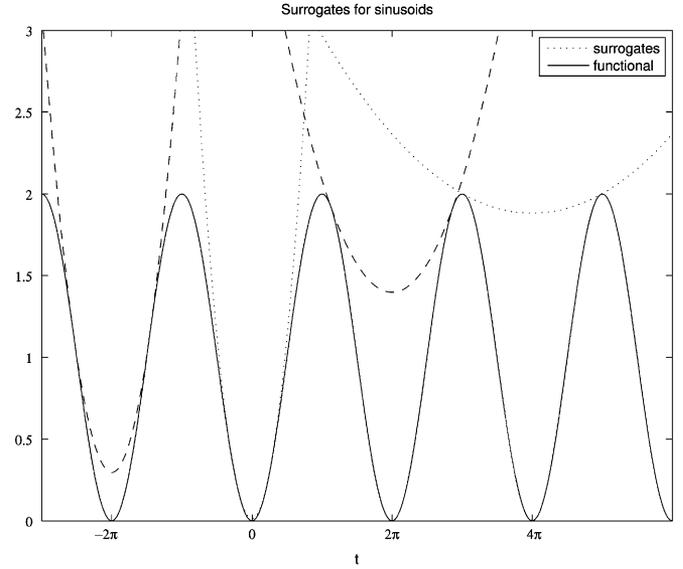
$$\dot{\varphi}(s) = \sin(s)$$

Fig. 10. Illustration of $\varphi(t)$ and quadratic surrogates for several values of s.

and

$$\kappa_\varphi(s) \triangleq \frac{\dot{\varphi}(s)}{s} = \frac{\sin(s)}{s}.$$

Substituting this curvature $\kappa_\varphi(s)$ into the expression for $\varphi_{jml}(\boldsymbol{\omega})$ gives us the following curvature for the parabola surrogate

$$\kappa_{\varphi,jml}(s) \triangleq \frac{\dot{\varphi}_{jml}(s)}{s} = \left| y_j^m y_j^l \right| w^{m,l} (\triangle_l - \triangle_m)^2 \frac{\sin(s)}{s}$$

which is bounded as $s \to 0$ and decreasing as $|s|$ increases. For values of $|s| > \pi$, we exploit the periodicity of φ and find an integer k such that $|s - k2\pi| \leq \pi$, i.e., the principal value of the phase s. Fig. 10 shows φ and parabola surrogates for several values of s. When s is an even multiple of π, the curvature κ_φ is the maximum curvature of φ. When s is an odd multiple of π, the curvature κ_φ is zero, and $\dot{\varphi}$ is also zero, so the surrogate function is a constant.

Aggregating such surrogates leads to the following surrogate function for the cost function: $\Psi(\boldsymbol{\omega})$

$$\phi^{(n)}(\boldsymbol{\omega}) \triangleq \sum_{j=1}^{N} \sum_{m=0}^{L} \sum_{l=0}^{L} q_{jml}^{(n)}(\omega_j) + \beta R(\boldsymbol{\omega})$$

where

$$q_{jml}^{(n)}(\omega) \triangleq \varphi_{jml}\left(\omega_j^{(n)}\right) + \dot{\varphi}_{jml}\left(\omega_j^{(n)}\right)\left(\omega - \omega_j^{(n)}\right) \\ + \frac{1}{2}\kappa_{\varphi,jml}\left(s_r^{(n)}\right)\left(\omega - \omega_j^{(n)}\right)^2$$

and where

$$s_r^{(n)} \triangleq \left(\omega_j^{(n)}|\triangle_l - \triangle_m| + \angle y_j^m - \angle y_j^l\right) \bmod \pi \in [-\pi, \pi].$$

If the roughness penalty $R(\boldsymbol{\omega})$ is a quadratic function, which is the natural choice for smooth phase maps, then the surrogate $\phi^{(n)}$ above is a quadratic function that can be minimized

easily by any classical method such as the conjugate gradient algorithm.

In our implementation, we used a separable quadratic surrogate algorithm to minimize this cost function [40]. Then, the following iteration, similar to that of (13), is guaranteed to decrease $\Psi(\boldsymbol{\omega})$ monotonically:

$$\boldsymbol{\omega}^{(n+1)} = \boldsymbol{\omega}^{(n)} - \text{diag}\left\{\frac{1}{\tilde{d}_j^{(n)} + \beta \cdot c}\right\} \nabla\Psi\left(\boldsymbol{\omega}^{(n)}\right) \quad (23)$$

where c was defined in (14) and where

$$\tilde{d}_j^{(n)} = \sum_{m=0}^{L}\sum_{l=0}^{L} \kappa_{\varphi,jml}\left(s_r^{(n)}\right).$$

The advantage of (23) over (13) is that $\tilde{d}_j^{(n)} \leq d_j$ in (15), so (23) will converge faster [41].

REFERENCES

[1] K. Sekihara, S. Matsui, and H. Kohno, "NMR imaging for magnets with large nonuniformities," *IEEE Trans. Med. Imag.*, vol. MI-4, no. 4, pp. 193–199, Dec. 1985.

[2] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 178–188, Feb. 2003.

[3] D. C. Noll, J. A. Fessler, and B. P. Sutton, "Conjugate phase MRI reconstruction with spatially variant sample density correction," *IEEE Trans. Med. Imag.*, vol. 24, no. 3, pp. 325–336, Mar. 2005.

[4] D. B. Twieg, "Parsing local signal evolution directly from a single-shot MRI signal: A new approach for fMRI," *Magn. Res. Med.*, vol. 50, no. 5, pp. 1043–1052, Nov. 2003.

[5] C. Barnet, J. Tsao, and K. P. Pruessmann, "Efficient iterative reconstruction for MRI in strongly inhomogeneous B_0," in *Proc. Intl. Soc. Magn. Res. Med.*, 2004, p. 347.

[6] P. Jezzard and R. S. Balaban, "Correction for geometric distortion in echo planar images from B_0 field variations," *Magn. Res. Med.*, vol. 34, no. 1, pp. 65–73, Jul. 1995.

[7] K. S. Nayak and D. G. Nishimura, "Automatic field map generation and off-resonance correction for projection reconstruction imaging," *Magn. Res. Med.*, vol. 43, no. 1, pp. 151–154, Jan. 2000.

[8] R. Cusack, M. Brett, and K. Osswald, "An evaluation of the use of magnetic field-maps to undistort echo-planar images," *NeuroImage*, vol. 18, no. 1, pp. 127–142, Jan. 2003.

[9] G. Glover, "Multipoint dixon technique for water and fat proton and susceptibility imaging," *J. Magn. Reson. Imag.*, vol. 1, no. 5, pp. 521–530, Sept. 1991.

[10] K. S. Nayak, C.-M. Tsai, C. H. Meyer, and D. G. Nishimura, "Efficient off-resonance correction for spiral imaging," *Magn. Res. Med.*, vol. 45, no. 3, pp. 521–524, Mar. 2001.

[11] N. Chen and A. M. Wyrwicz, "Correction for EPI distortions using multi-echo gradient-echo imaging," *Magn. Reson. Med.*, vol. 41, no. 6, pp. 1206–1213, Jun. 1999.

[12] A. N. Priest, E. D. Vita, D. L. Thomas, and R. J. Ordidge, "EPI distortion correction from a simultaneously acquired distortion map using TRAIL," *J. Mag. Res. Imag.*, vol. 23, no. 4, pp. 597–603, Apr. 2006.

[13] C. Windischberger, S. Robinson, A. Rauscher, M. Barth, and E. Moser, "Robust field map generation using a triple-echo acquisition," *J. Mag. Res. Imag.*, vol. 20, no. 4, pp. 730–734, Oct. 2004.

[14] P. J. Reber, E. C. Wong, R. B. Buxton, and L. Frank, "Correction of off-resonance-related distortion in echo planar imaging using EPI-based field maps," *Magn. Reson. Med.*, vol. 39, pp. 328–330, 1998.

[15] P. Aksit, J. A. Derbyshire, and J. L. Prince, "Three-point method for fast and robust field mapping for EPI geometric distortion correction," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Arlington, VA, 2007, pp. 141–144.

[16] E. Schneider and G. Glover, "Rapid in vivo proton shimming," *Magn. Res. Med.*, vol. 18, no. 2, pp. 335–347, Apr. 1991.

[17] P. Irarrazabal, C. H. Meyer, D. G. Nishimura, and A. Macovski, "Inhomogeneity correction using an estimated linear field map," *Magn. Res. Med.*, vol. 35, no. 2, pp. 278–282, Feb. 1996.

[18] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill, "Noise and filtration in magnetic resonance imaging," *Med. Phys.*, vol. 12, no. 5, pp. 586–591, Sept. 1985.

[19] M. A. Bernstein, M. Grgic, T. J. Brosnan, and N. J. Pelc, "Reconstructions of phase contrast, phased array multicoil data," *Mag. Reson. Med.*, vol. 32, no. 3, pp. 330–334, Sep. 1994.

[20] J. Leitao and M. Figueiredo, "Absolute phase image reconstruction: A stochastic non-linear filtering approach," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 868–882, Jun. 1997.

[21] P.-J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[22] M. W. Jacobson and J. A. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Im. Proc.*, vol. 16, no. 10, pp. 2411–2422, Oct. 2007.

[23] D. Böhning and B. G. Lindsay, "Monotonicity of quadratic approximation algorithms," *Ann. Inst. Stat. Math.*, vol. 40, no. 4, pp. 641–663, Dec. 1988.

[24] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions (with discussion)," *J. Computat. Graph. Statist.*, vol. 9, pp. 1–59, 2000.

[25] J. A. Fessler, D. Yeo, and D. C. Noll, "Regularized fieldmap estimation in MRI," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2006, pp. 706–709.

[26] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-Invariant tomographs," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1346–1358, Sep. 1996.

[27] M. Unser, A. Aldroubi, and M. Eden, "Recursive regularization filters: Design, properties, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 3, pp. 272–277, Mar. 1991.

[28] 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. 463–468, Jun. 1998.

[29] R. K.-S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085–1097, Nov. 1999.

[30] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. New York: Wiley, 1999.

[31] J. F. Schenck, "The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds," *Med. Phys.*, vol. 23, no. 6, pp. 815–850, Jun. 1996.

[32] J. A. Fessler and B. P. Sutton, "Nonuniform fast fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.

[33] C. Yip, J. A. Fessler, and D. C. Noll, "Iterative RF pulse design for multidimensional, small-tip-angle selective excitation," *Magn. Res. Med.*, vol. 54, no. 4, pp. 908–917, Oct. 2005.

[34] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Dynamic field map estimation using a spiral-in/spiral-out acquisition," *Magn. Res. Med.*, vol. 51, no. 6, pp. 1194–1204, Jun. 2004.

[35] K. Lu, T. T. Liu, and M. Bydder, "Optimal phase difference reconstruction: Comparison of two methods," *Mag. Res. Imag.*, vol. 26, no. 1, pp. 142–145, Jan. 2008.

[36] S. J. Garnier, G. L. Bilbro, J. W. Gault, and W. E. Snyder, "Magnetic resonance image restoration," *J. Math. Imag. Vis.*, vol. 5, no. 1, pp. 7–19, Feb. 1995.

[37] L. Ying, J. Sheng, and B. Liu, "Joint estimation of image and coil sensitivities in parallel MRI," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag.: Macro Nano*, Arlington, VA, 2006, pp. 17–20.

[38] D. F. Yu and J. A. Fessler, "Edge-preserving tomographic reconstruction with nonlocal regularization," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 159–173, Feb. 2002.

[39] M. W. Jacobson and J. A. Fessler, "Properties of MM algorithms on convex feasible sets: Extended version" Commun. Signal Process. Lab., Dept. Elect. Eng. Comput. Sci., Univ. Michigan, Ann Arbor, Tech. Rep. 353, Nov. 2004.

[40] H. Erdoğan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, vol. 44, no. 11, pp. 2835–2851, Nov. 1999.

[41] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "On complete data spaces for PET reconstruction algorithms," *IEEE Trans. Nucl. Sci.*, vol. 40, no. 4, pp. 1055–1061, Aug. 1993.