REGULARIZED ESTIMATION OF BLOCH-SIEGERT $|B_1^+|$ MAPS IN MRI

Hao Sun[⋆] *William A. Grissom*[†] *Jeffrey A. Fessler*[⋆]

[⋆] University of Michigan-Ann Arbor, Department of Electrical Engineering and Computer Science
[†]Vanderbilt University, Department of Biomedical Engineering

ABSTRACT

The Bloch-Siegert (BS) method is a fast method for finding RF transmit fields ($|B_1^+|$ maps) in MRI based on the phase shift induced by an off-resonance pulse. The phase calculation can be very noisy in regions with low image magnitude, leading to inaccurate estimations using the conventional method. This paper proposes a penalized likelihood (PL) estimator of the $|B_1^+|$ map in the image domain, and a fast optimization transfer algorithm for solving the problem. We validated our proposed method with both simulation and experimental data sets.

1. INTRODUCTION

Mapping the magnitude of the RF transmit magnetic field ($|B_1^+|$) is important for a variety of applications in MRI such as parallel transmit pulse design [1] and electrical property mapping [2]. Recently Sacolick et al. [3] proposed using the Bloch-Siegert (BS) shift for $|B_1^+|$ mapping. The method has the advantages of speed, relatively large dynamic range, and robustness to relaxation and off-resonance. However, the method of moments (MOM) estimator proposed in [3] can be inaccurate in regions with low image magnitude, because the BS method relies on the phase difference between two acquisitions, and the phase difference calculation can be quite noisy in those regions. Also, because the image magnitude is proportional to the tissue properties and the flip angle of the excitation pulse in the BS mapping sequence, and is not influenced by the BS encoding pulse itself, it is possible that those low magnitude regions will coincide with high $|B_1^+|$ values, resulting in noisy $|B_1^+|$ estimates where it is important for subsequent pulse design. This paper proposes a penalized likelihood estimator that is less sensitive to this type of problem, and develops and compares several optimization algorithms. We compare our methods to the conventional $|B_1^+|$ estimator using both simulation and experimental data sets.

Research supported in part by NIH R21EB012674; NIH R01NS58576

2. MODEL

2.1. Bloch Siegert B1 Mapping

A Bloch-Siegert $|B_1^+|$ mapping sequence acquires two images; each one is a gradient echo image with an excitation pulse followed by an off-resonance pulse (Fermi pulse is often used [3]. The off-resonance frequency of the Fermi pulses in those two acquisitions are often set to be opposite in practice to eliminate the first-order Bloch-Siegert phase shift dependence on the B0 effect [3]. Define $\mathbf{b} = [|B_{1,1}^+|, ..., |B_{1,N}^+|]^T$ to be the vector of unknown $|B_1^+|$ map values, and $\mathbf{f} = [f_1, ..., f_N]^T$ to be the unknown complex-valued image in the absence of a Fermi pulse, where N is the number of pixels. The mathematical model for the complex signal at spatial location j in these two images is:

$$\begin{aligned} y_j^1 &= f_j e^{i\mathrm{k}b_j^2} + \epsilon_j^1 \\ y_j^2 &= f_j e^{-i\mathrm{k}b_j^2} + \epsilon_j^2 \end{aligned} \tag{1}$$

where k is a known constant determined by the pulse shape and off-resonance frequency of the pulse, and ϵ_j is the complex Gaussian noise. The method of moments estimator in current use is given by:

$$\hat{|B_{1,j}^+|} = \sqrt{\frac{\angle(y_j^1 y_j^{2*})}{2\mathrm{k}}} \tag{2}$$

The phase calculation can be dominated by noise when f_j is small, leading to inaccurate estimation.

2.2. Penalized Likelihood Estimation

To improve the $|B_1^+|$ mapping, we propose a penalized-likelihood estimator. The joint maximum likelihood estimate of $\mathbf{b}$ and $\mathbf{f}$ is the minimizer of the following cost function subject to $\mathbf{b} \geq 0$:

$$\begin{aligned} \Psi_{\mathrm{ML}}(\mathbf{b}, \mathbf{f}) =& \frac{1}{2} \sum_{i=1}^{N} |y_j^1 - f_j e^{\imath \mathrm{k} b_j^2}|^2 \\ &+ \frac{1}{2} \sum_{i=1}^{N} |y_j^2 - f_j e^{-\imath \mathrm{k} b_j^2}|^2 \end{aligned} \tag{3}$$

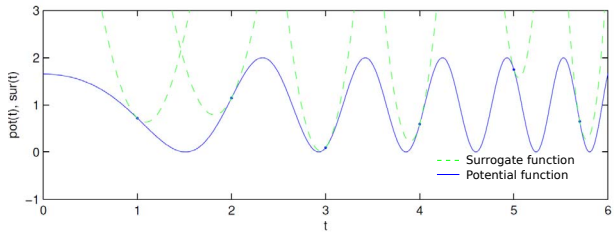

Fig. 1. Illustration of the potential function and its surrogate.

This function is quadratic and separable in f_j and the ML estimate for f_j (given $\mathbf{b}$) is:

$$\hat{f}_j = \frac{1}{2}(y_j^1 e^{-\imath \mathrm{k} b_j^2} + y_j^2 e^{\imath \mathrm{k} b_j^2}) \tag{4}$$

Substituting this into the cost function (3) yields

$$\Psi_{\mathrm{ML}}(\mathbf{b}) = \sum_{j=1}^{N} |y_j^1 y_j^2| [1 - \cos(2\mathrm{k} b_j^2 - \angle y_j^1 + \angle y_j^2)]. \tag{5}$$

The ML estimator of $\mathbf{b}$ ignores the prior knowledge that the $|B_1^+|$ map tends to be spatially smooth due to the physical nature of the transmit field. A natural approach to incorporating this characteristic is to add a roughness penalty to form the following penalized-likelihood cost function:

$$\Psi(\mathbf{b}) = \sum_{j=1}^{N} |y_j^1 y_j^2| [1 - \cos(2\mathrm{k} b_j^2 - \angle y_j^1 + \angle y_j^2)] + R(\mathbf{b}) \tag{6}$$

We estimate the $|B_1^+|$ map $\mathbf{b}$ by solving the following minimization problem:

$$\hat{\mathbf{b}} = \underset{\mathbf{b} \in \mathbb{R}^{\mathrm{N}}; \mathbf{b} \geq 0}{\arg\min}\, \Psi(\mathbf{b}) \tag{7}$$

The non-negativity constraint can be relaxed in practice since a pixel with negative estimated value typically has very low B_1^+ magnitude and can be set to 0 after solving the unconstrained problem.

3. MINIMIZATION ALGORITHMS

This section proposes to solve the optimization problem (7) using optimization transfer methods [4]. We consider several possible surrogate function designs.

3.1. Maximum Curvature for the ML Term

Let $t_j^2 = 2\mathrm{k} b_j^2$ and $\gamma_j = -\angle y_j^1 + \angle y_j^2$, then the maximum likelihood data fitting term can be expressed as

$$\Psi_{\mathrm{ML}}(\mathbf{t}) = \sum_{j=1}^{N} |y_j^1 y_j^2| \psi(t_j). \tag{8}$$

where $\psi(t) = 1 - \cos(t^2 + \gamma)$. The first order derivative of $\psi(t)$ is

$$\dot{\psi}(t) = 2t \sin(t^2 + \gamma), \tag{9}$$

which does not satisfy the Huber's conditions [5, P.184] because $\dot{\psi}(t)/t$ is not non-increasing for $t > 0$. Furthermore, the second derivative is

$$\begin{aligned} \ddot{\psi}(t) &= 4t^2 \cos(t^2 + \gamma) + 2\sin(t^2 + \gamma) \\ &= 2\sqrt{4t^4 + 1} \cos(t^2 + \gamma - \arctan(\frac{1}{2t^2})). \end{aligned}$$

which is unbounded as t goes to ∞, so a "maximum curvature" approach is also infeasible in theory. However, the $|B_1^+|$ field is typically bounded in practice; therefore, t is also bounded and we can design a "maximum curvature" by assuming a upper bound of t. Assuming $t \in [0, t_{\max}]$, then $\ddot{\psi}(t) \leq 2\sqrt{4t^4 + 1} \leq 2\sqrt{4t_{\max}^4 + 1}$, so we can use quadratic surrogate function with the following curvature:

$$\breve{c} = 2\sqrt{4t_{\max}^4 + 1} \tag{10}$$

In this approach, we must consider the box constraint [0, $t_{\max}$]; we can either project the solution to the feasible set in each iteration if using separate quadratic surrogate, or set a relative large $t_{\max}$ such that the estimated value in each iteration is always smaller than $t_{\max}$ in practice.

However, this maximum curvature approach may lead to slow convergence. Fig. 1 shows that the potential function ψ is an oscillating function with the same upper and lower bound in each cycle; a quadratic surrogate in one cycle is guaranteed to be a surrogate over the entire feasible domain, leading to the following quadratic surrogate design.

The function $\psi(t)$ has a local maximum at every $t_n^{\max} = \sqrt{2\pi n + \pi - \gamma}$ if $\gamma < \pi$, $t_{-1}^{\max} = 0$ if $\gamma > \pi$. For any $s \geq t_1^{\max}$, let $n(s) = \lfloor \frac{s^2 + \gamma - \pi}{2\pi} \rfloor$, then s is in the interval $[t_{n(s)}^{\max}, t_{n(s)+1}^{\max}]$, and therefore an upper bound on the curvature over this cycle containing s is

$$\breve{c}(s) = 2\sqrt{4(t_{n(s)+1}^{\max})^4 + 1}. \tag{11}$$

Thus we create a separable quadratic surrogate (SQS) [6] for the negative log-likelihood based on this upper bound.

3.2. Minimization of the Surrogate Function

Because $|B_1^+|$ maps are smooth, we use the following quadratic roughness penalty:

$$R(\mathbf{b}) = \frac{1}{2}\beta ||\mathbf{C}\mathbf{b}||^2 \tag{12}$$

where β is the regularization parameter, and $\mathbf{C}$ is the first-order 2D finite difference matrix. Combining with the surrogate function for the ML term, we get the following surrogate

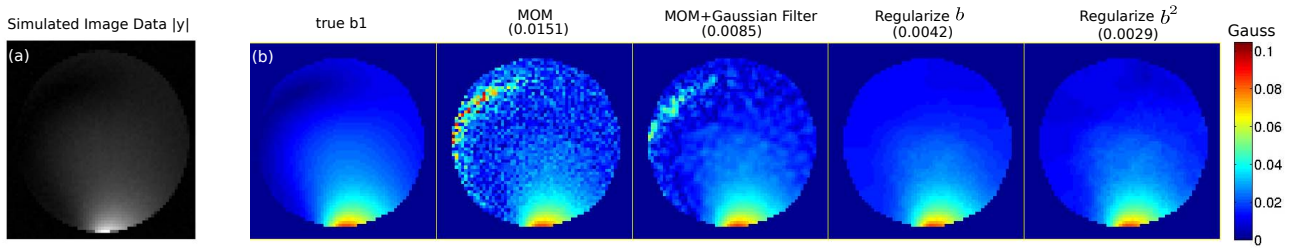

Fig. 2. Results for the simulated data set. (a) simulated image data (magnitude). (b) Estimated $|B_1^+|$ maps. The RMSE in Gauss for all methods are shown in the parentheses. Both PL estimators generate more accurate $|B_1^+|$ maps than the MOM and MOM+smoothing approach.

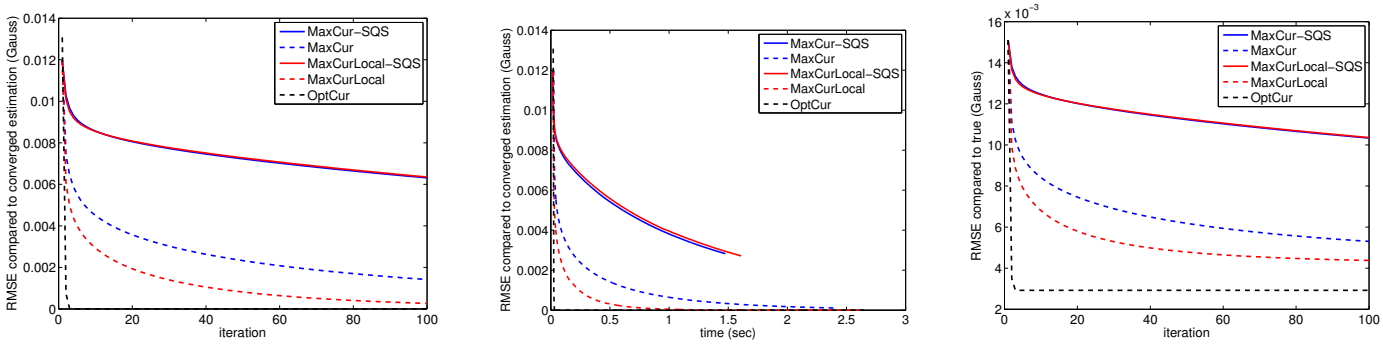

Fig. 3. Plots of RMSE in Gauss compared to the converged estimation with respect to iteration (a) and time (b). Plots of RMSE compared to the true $|B_1^+|$ with respect to iteration (c). Using true Hessian for the regularization term instead of SQS greatly improves the convergence rate. Using cost function (16) and its optimal curvature converges the fastest.

function for the cost function:

$$\Phi(\mathbf{b};\mathbf{b}^{(n)}) = \Psi(\mathbf{b}^{(n)}) + \nabla\Psi(\mathbf{b}^{(n)})(\mathbf{b}-\mathbf{b}^{(n)}) + \frac{1}{2}(\mathbf{b}-\mathbf{b}^{(n)})'\mathbf{D}(\mathbf{b}-\mathbf{b}^{(n)}) + \frac{1}{2}\beta\mathbf{b}'\mathbf{C}'\mathbf{C}\mathbf{b} \quad (13)$$

where $\mathbf{D}$ is the diagonal matrix with elements $d_j^{(n)} = 2\mathrm{k}|y_j^1 y_j^2|\check{c}(\sqrt{2\mathrm{k}b_j^{(n)}})$, where $\check{c}(\cdot)$ is defined in (10).

The Hessian matrix of the surrogate is $\mathbf{H}^{(n)} = \mathbf{D}^{(n)} + \beta\mathbf{C}'\mathbf{C}$. We investigate two ways to deal with this Hessian. First, we can design a diagonal majorizor for $\mathbf{C}'\mathbf{C}$, namely $\mathbf{C}'\mathbf{C} \leq 4\mathbf{I}$, leading to the following algorithm:

$$\mathbf{b}^{(n+1)} = \mathbf{b}^{(n)} - \mathrm{diag}\{\frac{1}{d_j^{(n)} + 4\beta}\}\nabla\Psi(\mathbf{b}^{(n)}) \quad (14)$$

Alternatively, since $\mathbf{C}'\mathbf{C}$ has a sparse banded structure, we can calculate $\mathbf{H}^{-1}\nabla\Psi(\mathbf{b}^{(n)})$ efficiently by sparse Cholesky factorization techniques [7], leading to the following Huber's algorithm:

$$\mathbf{b}^{(n+1)} = \mathbf{b}^{(n)} - {\mathbf{H}^{(n)}}^{-1}\nabla\Psi(\mathbf{b}^{(n)}) \quad (15)$$

3.3. Alternative Formulation

In the above approach, the curvature for the data fitting term is suboptimal, and it is not straight forward to find the optimal curvature. Alternatively, we can change the regularization term in (7) to a roughness penalty on $\mathbf{b}^2$ instead of $\mathbf{b}$. Letting $\mathbf{x} = \mathbf{b}^2$, the problem to solve becomes:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^N;\mathbf{x}\geq 0} \sum_{j=1}^{N} |y_j^1 y_j^2| [1 - \cos(2\mathrm{k}x_j - \angle y_j^1 + \angle y_j^2)] + R(\mathbf{x}) \quad (16)$$

where $R(\cdot)$ is the same quadratic roughness penalty defined in Eq. (12). The data fitting term in this formulation satisfies Huber's condition [8], and we can therefore design a SQS with optimal curvature as in [8]:

$$\check{c}(x_j) = 4\mathrm{k}^2|y_j^1 y_j^2|\frac{\sin(s_j^{(n)})}{s_j^{(n)}} \quad (17)$$

where $s_j^{(n)} = (2\mathrm{k}x_j - \angle y_j^1 + \angle y_j^2) \mod \pi \in [-\pi, \pi]$. Then we get the following Huber's algorithm:

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} - {\mathbf{H}^{(n)}}^{-1}\nabla\Psi(\mathbf{x}^{(n)}) \quad (18)$$

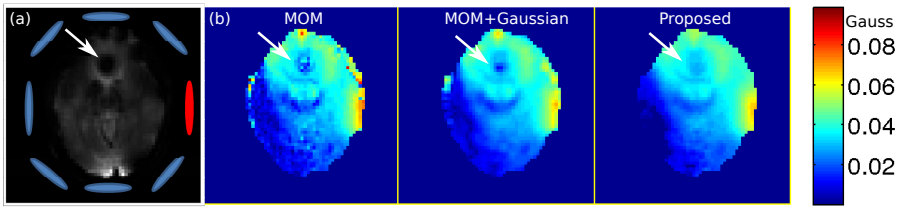

Fig. 4. Results for the *in vivo* data set: (a) acquired image with BS encoding from one of 8 channel transmit coil (red ellipse), (b) estimated $|B_1^+|$ map from the method of moments and the proposed penalized likelihood method (16). Artifact (arrow) due to T2* signal drop is greatly reduced in the proposed method. Also, our method removes the popcorn noise observed around the periphery of the head, and that noise can significantly affect subsequent RF pulse designs.

where $\mathbf{H}^{(n)} = \mathbf{D}^{(n)} + \beta \mathbf{C}'\mathbf{C}$, and where $\mathbf{D}^{(n)}$ is a diagonal matrix with element $\mathbf{D}_{jj}^{(n)} = \check{c}(x_j^{(n)})$.

This formulation differs from Eq. (7) in the sense that we are regularizing the spatial variation of $\mathbf{b}^2$ instead of $\mathbf{b}$, so now $R(\mathbf{b}^2) = \beta \sum_{j=1}^{N} (b_j^2 - b_{j-1}^2)^2 = \beta \sum_{j=1}^{N} (b_j + b_{j-1})^2 (b_j - b_{j-1})^2$. Compared to $R(\mathbf{b})$, we are putting more regularization in the region with high transmit field strength, which may be undesired. One could compensate for this effect by adding a weighting matrix in the regularization if needed.

4. SIMULATION AND EXPERIMENTAL RESULTS

We compared the penalized likelihood methods with the method of moments with both simulated and measured data. The simulated data were synthesized for one channel of a 8-channel transmit array in the image domain by applying the $\sin |B_1^+(\vec{r})|$ magnitude weighting and BS phase shift to an uniform ball phantom image and then adding Gaussian white noise with 60 dB SNR. Figure 2(b) shows the simulated true $|B_1^+|$ map and reconstruction maps using different methods. The MOM estimate suffers from large noise, especially in the region with low transmit magnitude. Applying Gaussian low pass filter to the MOM estimate is a conventional way to improve the result in practice, but still shows mismatch in the region with low transmit magnitude. Both PL estimates using formulation (7) and (16) generate more accurate maps. Figure 3 shows convergence plot using different minimization methods with respect to iteration (a) and time (b). In the maximum curvature approach (10), we assume maximum $|B_1^+|$ = 0.2 Gauss, which is larger than the maximum $|B_1^+|$ in this simulated data and a reasonable upper limit in practice. Using the local maximum curvature (11) has faster convergence rate than using the global maximum curvature (10). The Cholesky approach is implemented in MALTAB using "\", which converges faster than the separable quadratic surrogate approach, with respect to iteration and time. The fastest algorithm is using formulation (16) with the optimal curvature and the Cholesky factorization, which converges in just 3 iteration. Figure 2(c) shows the error plot compared to the true $|B_1^+|$ map for each method.

The proposed methods were also validated with real experiment data. Data were acquired with a 8 channel transmit/receive array in a GE 3.0 T scanner. A 12 ms, $\pm$ 4KHz Bloch-Siegert encoding pulse was transmitted in one of the 8 coil (see, red ellipse in Fig. 4(a)). We used 64x64 spin warp readout, 24 cm FOV, TE = 15 ms. The acquired image is shown in Fig. 4(a), which shows a dark hole due to large T2* signal drop. Figure 4(b) shows the $|B_1^+|$ estimates using the MOM, MOM with Gaussian low pass filter, and the proposed method (16). We observe a steep $|B_1^+|$ drop in the frontal sinus region from the MOM based estimation (see the arrow), which is unexpected because the $|B_1^+|$ should change relatively smoothly in brain in 3T. The proposed approach greatly reduced this artifact. Also, our method removes the popcorn noise observed around the periphery of the head, and that noise can significantly affect subsequent RF pulse designs. We tried adding weighting matrix to the regularization in (16) to compensate the difference between $R(\mathbf{b}^2)$ and $R(\mathbf{b})$, but it made little difference in both simulation and experimental data sets (not shown).

5. CONCLUSION

We proposed a penalized likelihood estimator for Bloch-Siegert $|B_1^+|$ mapping in the image domain, and compared several optimization algorithms to solve this problem. By penalizing $|B_1^+|^2$ instead of $|B_1^+|$, we can find an optimal curvature quadratic surrogate and solve the problem efficiently using Huber's algorithm with Cholesky factorization technique. The proposed method is validated in both simulated data set and *in vivo* data, showing reduced noise and artifact compared to the conventional MOM based methods. In the future, we plan to combine the proposed estimation method with lower energy Fermi pulses to reduce the specific absorption rate (SAR) without sacrificing the B1 mapping accuracy. Also, we will try to extend our method to the optimized Bloch-Siegert encoding pulses [9, 10].

6. REFERENCES

[1] W. Grissom, C. Yip, Z. Zhang, V. A. Stenger, J. A. Fessler, and D. C. Noll, "Spatial domain method for the design of RF pulses in multi-coil parallel excitation," *Magn. Reson. Med.*, vol. 56, no. 3, pp. 620–9, Sept. 2006.

[2] Ulrich Katscher, Tobias Voigt, Christian Findeklee, Peter Vernickel, Kay Nehrke, and O Dossel, "Determination of electric conductivity and local sar via b1 mapping," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1365–1374, 2009.

[3] L. I. Sacolick, F. Wiesinger, I. Hancu, and M. W. Vogel, "B1 mapping by Bloch-Siegert shift," *Magn. Reson. Med.*, vol. 63, no. 5, pp. 1315–22, May 2010.

[4] 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. 2411–22, Oct. 2007.

[5] P. J. Huber, *Robust statistics*, Wiley, New York, 1981.

[6] H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, vol. 44, no. 11, pp. 2835–51, Nov. 1999.

[7] M. J. Allison and J. A. Fessler, "Accelerated computation of regularized field map estimates," in *Proc. Intl. Soc. Magn. Reson. Med.*, 2012, p. 0413.

[8] 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.

[9] Mohammad Mehdi Khalighi, Brian K Rutt, and Adam B Kerr, "Adiabatic RF pulse design for Bloch-Siegert B_1^+ mapping," *Magnetic Resonance in Medicine*, vol. 70, no. 3, pp. 829–835, 2013.

[10] Marcin Jankiewicz, John C Gore, and William A Grissom, "Improved encoding pulses for Bloch-Siegert $|B_1^+|$ mapping," *Journal of Magnetic Resonance*, vol. 226, pp. 79–87, 2013.