AN ENHANCED APPROACH FOR SIMULTANEOUS IMAGE RECONSTRUCTION AND SENSITIVITY MAP ESTIMATION IN PARTIALLY PARALLEL IMAGING

Meng Liu[⋆†∗] *Yunmei Chen*[⋆†] *Yuyuan Ouyang*[⋆] *Xiaojing Ye*[§] *Feng Huang*[¶]

[⋆] Department of Mathematics, University of Florida
[§] School of Mathematics, Georgia Institute of Technology
[¶] Phillips Research Asia Shanghai, China

ABSTRACT

We develop a variational model and a faster and robust numerical algorithm for simultaneous sensitivity map estimation and image reconstruction in partially parallel MR imaging with significantly under-sampled data. The proposed model uses a maximum likelihood approach to minimizing the residue of data fitting in the presence of independent Gaussian noise. The usage of maximum likelihood estimation dramatically reduces the sensitivity to the selection of model parameter, and increases the accuracy and robustness of the algorithm. Moreover, variable splitting based on the specific structure of the objective function, and alternating direction method of multipliers (ADMM) are used to accelerate the computation. The preliminary results indicate that the proposed method resulted in fast and robust reconstruction.

Index Terms— partially parallel imaging, primal-dual method, sensitivity estimation, maximum likelihood estimation, SENSE

1. INTRODUCTION

Partially Parallel Imaging (PPI) is an emerging technique in MR imaging. PPI reduces the acquisition time of the scanned object data by surrounding the scanned objects by multiple radiofrequency (RF) receivers and by acquiring a reduced number of k-space data for each RF receiver. SENSitivity Encoding (SENSE)[1, 2], one of the most commonly used image domain-based parallel imaging methods, utilizes the information of the coil sensitivities to separate aliased pixels resulted from under-sampled k-space data, satisfying $(M\mathcal{F}S_j u - f_j)$, $j = 1, 2, ..., K$, where K is the number of coil receivers, u is the underlying image, M is a mask(a binary matrix) presenting the trajectories to acquire partial k-space data, $\mathcal{F}$ is the Fourier transform, S_j is the sensitivity map for jth coil and f_j is the observed under-sampled k-space data.

In [3], Chen et, al. develop a variational model and a fast algorithm, referred as APD, to solve minimization problem

$$\min_u \lambda \|u\|_{\mathrm{TV}} + \sum_{j=1}^{K} \|M\mathcal{F}S_j u - f_j\|_2^2$$

by applying variable splitting with alternating direction method of multipliers (ADMM)[4] to decompose it into one subproblem involving Fourier transforms and another subproblem that can be treated by the primal-dual hybrid gradient method (PDHG)[5] scheme.

In PPI, however, the quality of SENSE reconstruction is highly depending on the accuracy of coil sensitivity maps and it is difficult to obtain artifact-free sensitivity maps. It is also important to have accurate sensitivity estimates in any low signal regions surrounding the object to avoid reconstruction artifacts[6]. Hence, It is argued that for successful parallel imaging, the choice of sensitivity estimation strategy is at least as important as the choice of reconstruction strategy[7]. Many common methods of coil sensitivity calibration has been to measure sensitivities directly using one or more low-resolution calibration images, obtained from low-frequency calibration data in k-space after division by a body coil image[8], or by the sum-of-squares (SoS) of low-resolution images from all the channels. It also introduces a possible source of error into the reconstruction, since it is difficult to ensure that the patient and coil array will be on the same positions during both the calibration scan and the accelerated data acquisition. Misregistrations or inconsistencies between the calibrated and the true sensitivities result in artifacts in the reconstructions. Coil sensitivities are also subject to noise processes. This causes sensitivity errors in regions of low signal and where the body coil data itself is noisy. The SoS technique uses fully sampled central k-space to obtain a relative sensitivity map. It involves pixel-wise division, which can be reasonably considered to yield nonidentically distributed but still fairly independent noise in the sensitivity maps.

However, the quality of SENSE reconstruction is highly depending on the accuracy of coil sensitivity maps. Based on the fact of difficulties on obtaining sufficiently accurate sensitivity maps, people need to find methods to update sensitivities to improve the quality of image reconstructions. Several iterative methods, such as JSENSE[9] by Ying et al., IRGN-TV/TGV [10, 11] by Uecker et al. and Knoll et al., and Sparse BLIP [12] by H. She et al., have been proposed to jointly reconstruct image and estimate sensitivity maps. JSENSE addresses the issue of sensitivity errors by iteratively correcting the sensitivities functions, represented by a simple polynomial. As well, IRGN-TV/TGV is a regularized nonlinear inversion framework based on a Newton-type method with variational penalties for auto-calibrated parallel imaging with arbitrary sampling patterns. Sparse BLIP is another joint method, which improves upon JSENSE by incorporating the concept of Compressed Sensing into the data consistency formulation and assumes the image and sensitivities are sparse in total variation (TV) of the spatial domain. These prior arts have demonstrated that the accuracy of coil sensitivity maps can be improved iteratively, and hence can improve the quality of SENSE reconstruction.

Keeling, et al. also propose a variational approach for PPI, named TVSENSE[13], using a high-order penalty for coil sensitivities and a TV like penalty,the Gauss-TV penalty for reconstructing

[∗] *Asterisk indicates corresponding author.*

[†] The research was partly supported by National Science Foundation, grant DMS1115568.

image. TVSENSE is formulated purely in image space as opposed to k-space, and it also jointly reconstructs image and estimates sensitivity maps. The numerical method in [13] is based on CGM, a Newton-type optimization framework introduced in [14]. However, the numerical result shows that TVSENSE is quite slower than many other algorithms of PPI such as IRGN and SENSE.

However, these algorithms still suffer one or both of the following two numerical problems. One is the sensitivity to the choice of regularization parameters, and the other is the high computational cost. As for the robustness, most algorithms of PPI reconstructions still suffer from two numerical problems. One is high computational cost, and the other is the sensitivity to the choice of regularization parameters. They require selecting parameters of regularizations manually, and the algorithms are sensitive to the selections of parameters. An inappropriate ratio, determined by the regularization parameters, between data fidelity and regularizations may result in either residual artifact or reduced spatial resolution. Intuitively, we wish that the regularization could be strong in the beginning and would gradually reduce when the fidelity term decreases. Therefore, it inspires us to borrow the idea of maximum likelihood estimation (MLE) on the residue of data fitting on k-space in the presence of independent Gaussian noise.

Motivated by all discussed above, we are interested in proposing a novel variational model, and a fast and robust algorithm, called TVL1SS, to simultaneously reconstruct the image and estimate the sensitivity maps. Except data fidelity, we utilize TV regularization to image and sparse representation, for instance by wavelet transform, to sensitivity maps as regularizations. Combined with MLE, we develop a self-adjusted technique to iteratively optimize the regularization ratio, $i.e.$ the balance between the regularization terms and the data fidelity term. The numerical approach is adopted from and based on [3] and [15]. By introducing auxiliary variables, the objective function is splitted into several subproblems, which can be solved by PDHG, 1-dimensional Shrinkage, and linear inverse problems with diagonal inversion matrices. All the subproblems are fast numerical schemes, which guarantee our algorithm is a fast approach. The usage of MLE dramatically reduces the difficulty of parameter decision and increases the practicability of regularized reconstruction techniques. The regularization parameter can be selected from a large range (the details will be given in section 3.2).

2. PROPOSED MODEL AND NUMERICAL METHOD

2.1. Proposed Model

We develop our model in this section. Assume that, for each j-th channel, the residues of $(M\mathcal{F}S_ju - f_j)$ on k-space indexed by r are identically independently distributed (i.i.d.), and obey a Gaussian distribution with zero-mean, standard deviation δ, which is to be optimized. The joint probability density function (PDF) of the residues in j-th channel is

$$P\left(\{(M\mathcal{F}S_ju - f_j)(r)|r \in \Omega\}\right) = \prod_{r\in\Omega} G_\delta\left((M\mathcal{F}S_ju - f_j)(r)\right), \quad (1)$$

where $\prod$ is the multiplication operator and

$$G_\delta(z) = \frac{1}{\sqrt{2\pi}\delta}\exp\{-\frac{z^2}{2\delta^2}\}. \quad (2)$$

The likelihood function, also assumed the independence of the residues of all the channels, $LE\left(\delta\right)$ is the joint PDF can be written as

$$LE(\delta) = \prod_{j=1}^{K}\frac{1}{\sqrt{2\pi}\delta}\exp\{-\|M\mathcal{F}S_ju - f_j\|_\Omega^2/2\delta^2\}.$$

Maximizing the likelihood enforces $(M\mathcal{F}S_ju - f_j)$ optimally following the Gaussian distribution. Instead of maximizing the likelihood, we minimize the negative log-likelihood function, given as follows:

$$-\log LE(\delta) = \frac{1}{2\delta^2}\sum_{j=1}^{K}\|M\mathcal{F}S_ju - f_j\|_2^2 + \frac{|\Omega|}{2}\ln\delta,$$

where $|\Omega|$ is the area of each k-space domain. Combined with regularizations on underlying image and sensitivity maps, we establish a variational model for simultaneous image reconstruction and sensitivity map estimation in PPI, as follows:

$$\min_{u,S,\delta}\lambda\|u\|_{\text{TV}} + \mu\sum_{j=1}^{K}\|\Psi^{\text{T}}S_j\|_1 + \frac{1}{2\delta^2}\sum_{j=1}^{K}\|M\mathcal{F}S_ju - f_j\|_2^2 + \frac{|\Omega|}{2}\ln\delta. \quad (3)$$

We minimize TV-norm of the reconstruction of image because it is edge-preserved and sparse in the finite difference domain. We propose $L1$-norm of sensitivity maps on wavelet domain because of the sparsity of sensitivity maps on wavelet domain.

Following the standard treatment we will vectorize an two-dimensional image u into one-dimensional column vector, $i.e.$ $u \in \mathbb{C}^N$, where N is the number of pixels of the image u. Each S_j is a $N \times N$ diagonal matrix in which the diagonal entries correspond the sensitivities of the j-th coil for all the pixels of the image. Then the (isotropic) TV norm is defined by

$$\|u\|_{\text{TV}} = \int_\Omega |Du| = \sum_{i=1}^{N}\|D_iu\|,$$

where for each $i = 1, 2, ..., N$, $D_i \in \mathbb{R}^{2\times N}$ has two nonzero entries in each row corresponding to finite difference approximations to partial derivatives of u at the i-th pixel along the coordinate axes. In addition, $\Psi = [\psi_1, ..., \psi_N] \in \mathbb{C}^{N\times N}$ is the wavelet transform matrix that sparsifies the underlying sensitivity maps.

2.2. Numerical Method

By the idea of variable splitting, introduce auxiliary variables and obtain a constrained problem,

$$\min_{u,S,\delta,t,v}\lambda\|u\|_{\text{TV}} + \mu\sum_{j=1}^{K}\sum_{i=1}^{N}|t_{ij}| + \frac{1}{2\delta^2}\sum_{j=1}^{K}\|M\mathcal{F}v_j - f_j\|_2^2 + \frac{|\Omega|}{2}\ln\delta, \quad (4)$$

subject to

$$\begin{aligned} t_{ij} &= \psi_i^{\text{T}}S_j, \text{ for } i = 1, 2, ..., N, \ j = 1, 2, ..., K, \\ v_j &= S_ju, \text{ for } j = 1, 2, ..., K. \end{aligned}$$

To optimize the minimization problem, we propose the following scheme, named TVL1SS, based on ADMM, as follows

$$\begin{cases} t_{ij}^{k+1} = \arg\min_{t_{ij}}|t_{ij}| + \alpha\left|t_{ij} - \psi_i^TS_j^k - d_{ij}^k\right|^2, \\ u^{k+1} = \arg\min_u \lambda\|u\|_{\text{TV}} + \frac{\sigma}{2}\sum_{j=1}^{K}\|S_j^ku - v_j^k + z_j^k\|^2, \\ v_j^{k+1} = \arg\min_{v_j}\|M\mathcal{F}v_j - f_j\|^2 + \sigma\delta^k\|S_j^ku^{k+1} - v_j + z_j^k\|^2, \\ S_j^{k+1} = \arg\min_{S_j}\alpha\mu\|\Psi^TS_j - t_j^{k+1} + d_j^k\|^2 \\ \qquad\qquad + \frac{\sigma}{2}\|S_ju^{k+1} - v_j^{k+1} + z_j^k\|^2, \\ \delta^{k+1} = \arg\min_\delta \frac{1}{2\delta^2}\sum_{j=1}^{K}\|M\mathcal{F}S_j^{k+1}u^{k+1} - f_j\|_2^2 + \frac{|\Omega|}{2}\ln\delta \\ d_{ij}^{k+1} = d_{ij}^k + \psi_i^TS_j^{k+1} - t_{ij}^{k+1}, \\ z_j^{k+1} = z_j^k + S_j^{k+1}u^{k+1} - v_j^{k+1}. \end{cases}$$

In this algorithm, each subproblem is easy to solve. For t-subproblem, 1-D Shrinkage can be applied. For u-subproblem, apply the framework of PDHG scheme. For v-subproblem and S-subproblem, we can solve them efficiently because the matrices to be taken the inversion is the product of Fourier transforms and a diagonal matrix or just a diagonal matrix. δ-subproblem is also a step of simple computation. Therefore the computation of TVL1SS is cheap and considerably fast.

3. NUMERICAL EXPERIMENTS

In this section, in order to examine if our work is improved from APD and performances comparably better than other existing algorithms for joint image reconstruction and sensitivity estimation, taking IRGN-TGV for example, we design two groups of experiments using the same acquired data with three different initial guesses of sensitivity maps to compare TVL1SS with APD and IRGN-TGV.

The first group of experiments tests the convergence speed and reconstruction quality of the proposed method for data sets with different initial guesses of the sensitivity maps. The second group tests the robustness and sensitivity of the proposed method to the choice of regularization parameters.

Briefly speaking, the constrained problem of APD is,

$$\min_u \lambda\|u\|_{\text{TV}} + \sum_{j=1}^{K} \|M\mathcal{F}v_j - f_j\|_2^2 \text{ subject to } v_j = S_j u,$$

and the ADMM of APD is

$$\begin{cases} v_j^{k+1} = \arg\min_{v_j} \|M\mathcal{F}v_j - f_j\|^2 + \alpha\|S_j u^{k+1} - v_j + b_j^k\|^2, \\ u^{k+1} = \arg\min_u \lambda\|u\|_{\text{TV}} + \alpha\sum_{j=1}^{K} \|S_j u - v_j^k + b_j^k\|^2, \\ b_j^{k+1} = b_j^k + S_j u^{k+1} - v_j^{k+1}, \ j = 1, ..., K. \end{cases}$$

In IRGN-TGV, the authors establish the relation between image, coil sensitivities and acquired k-space as the following function:

$$F(u, S) = (M\mathcal{F}S_1 u, ..., M\mathcal{F}S_K u)^T = f.$$

As known in [10], using the IRGN method, compute the solution $\delta x := (\delta S, \delta u)$ of a minimization problem in k-th step for given $x^k := (S^k, u^k)$:

$$\min_{\delta x} \frac{1}{2}\|F'(x^k)\delta x + F(x^k) - f\|^2 + \frac{\alpha_k}{2}\mathcal{W}(S^k + \delta S) + \beta_k\mathcal{R}(u^k + \delta u)$$

where $\alpha_k,\ \beta_k > 0$, and then setting $x^{k+1} = x^k + \delta x$, $\alpha_{k+1} := q_\alpha\alpha_k$ and $\beta_{k+1} := q_\beta\beta_k$ with $0 < q_\alpha, q_\beta < 1$. Here, $F'(x^k)$ is the Fréchet derivative of F evaluated at x^k. The term $\mathcal{W}(S) = \|Wc\|^2 = \|w \cdot \mathcal{F}S\|^2$ is a penalty of the high Fourier coefficients of the sensitivities and $\mathcal{R}$ is the TV regularization for the image. The authors apply the duality property of TV term and rewrite the minimization problem into a convex-concave saddle-point problem.

3.1. Data Acquisition

The original full-acquired k-space is a collection of sagittal Cartesian brain data acquired on a 3T system with 8-channel head coil (Invivo Corp., Gainesville, USA). The data acquisition parameters are field of view (FOV) 220 mm^2, size $512 \times 512 \times 8$, repetition time (TR) 2060ms, echo time (TE) 126ms, slice thickness 5mm, and flip angle 90°. The phase encoding direction is anterior-posterior. We reduced the image size from 512 to 256, and normalized the data sets such that the intensities of reference images have range $[0, 1]$. In our experiments, we artificially downsample the pseudofull k-space data using the Poisson mask with the reduction factor (RF) 4 (Figure 1a), $i.e.$ 25% sampling ratio.

3.2. Experimental Setting

3.2.1. Initial Guess of Sensitivity Maps and Image

For all tests, we set the initial guess u^0 to zero. We generally estimate the initial sensitivity maps S_j^0 by the low-resolution image $\tilde{u}_j$ generated by inverse Fourier transformed central k-space as part of under-sampled data,

$$(S_j^0)_i = \frac{(\tilde{u}_j)_i}{\sqrt{\sum_j |(\tilde{u}_j)_i|^2}}, \tag{5}$$

where $\tilde{u}_j$ is the Fourier transform of the central k-space data. The sensitivity maps are also normalized into the same range. The parameter setting and further discussion are based on the normalized data.

In our experiments, we estimate three different initial guesses of sensitivity maps, called SMAP1, SMAP2 and SMAP3, by the following ways.

For SMAP1, we simulate S_j^0 using the central 32×32 k-space data and applying inverse-Fourier-transform and (5). We add a complex-valued Gaussian noise (same level for both real and imaginary parts) with standard deviation 0.01 in magnitude to the pseudo full data.

For SMAP2 and SMAP3, S_j^0's are generated using central 64×64 and added complex-valued Gaussian noise (same level for both real and imaginary parts) with variance 0.01 and 0.025 in magnitude to the inverse-Fourier-transformed low-resolution images, $\tilde{u}_j$ in (5), from all the channels, respectively, before the sum-of-squares technique. In this case, the sensitivity maps are corrupted by the noise carried from the surface coil images and computation. Hence improving their accuracy during iterations is necessary for high quality reconstruction.

3.2.2. Termination Condition

The computation is set to be terminated when the relative change $\|u^k - u^{k-1}\|_2/\|u^k\|_2$ of the iteration reaches the prescribed stopping criterion 5×10^{-4}.

3.2.3. Computing Enviroment

All the experiments are implemented in MATLAB, R2010aSV, on a Macbook Pro with Mac OS X Lion operating system equipped a 2.66 Intel Core i7 processor and 4 GB 1067 MHz DDR3 memory. The sparsifying operator Ψ^T, discrete orthogonal wavelet transform, is set to use a level 4 Daubechies-4 (D4) wavelet transform in Rice Wavelet Toolbox (RWT) for MATLAB.

3.3. Experimental Results

3.3.1. Efficiency and Accuracy

The relative errors in the reconstructed image u, $\|u - \bar{u}\|_2/\|\bar{u}\|_2$, are indicated in the tables and figures. Here signal-to-noise ratio (SNR) is defined as $10\log_{10}(dv/mse)$, where mse is the mean of

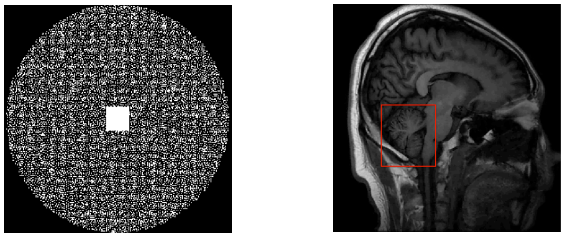
the squared difference between u and $\bar{u}$, $E[u-\bar{u}]^2$, and dv is the variance of $\bar{u}$, $Var(\bar{u}) = E[(\bar{u}-m)^2]$ with $m = E(\bar{u})$.

Compared with the performances of three algorithms on three different data sets. For SMAP2 and SMAP3, TVL1SS converges much faster than APD and IRGN-TGV do (see Table 1).

Data	Algorithms	SNR	Rel.err.	CPU Time(s)
SMAP1	APD	23.16	5.77%	16.09
	IRGN-TGV	13.31	16.44%	538.52
	TVL1SS	24.45	5.05%	71.53
SMAP2	APD	17.66	13.19%	213.78
	IRGN-TGV	14.61	14.09%	546.88
	TVL1SS	20.30	8.55%	29.81
SMAP3	APD	12.89	23.89%	459.68
	IRGN-TGV	14.80	13.79%	544.35
	TVL1SS	15.71	13.69%	82.35

Table 1: *Comparison of APD, IRGN-TGV and TVL1SS*

By Figure 2 and Table 1, first, the numerical results of TVL1SS are more accurate than the results of APD. In the experiment of SMAP1 with assumed well pre-estimated sensitivity maps, TVL1SS does not converge the fastest among the three algorithms, but the reconstruction quality of TVL1SS is better than the others, because APD applies fixed sensitivities while TVL1SS updates the sensitivity maps to achieve more accuracy , benefits from MLE and obtains more accurate reconstruction. Second, when sensitivity maps are not well pre-estimated, the reconstruction from both of APD and IRGN-TGV have more errors than results of TVL1SS. Compared with the L2-norm of weighted Fourier transform of sensitivities in IRGN-TGV, we use L1-norm of wavelet transform of sensitivities as the regularization. In the minimization problems, L2-norm requires all the entries to be uniformly small, but L1-norm allows sparse entries to be big and forces other entries to be zero.



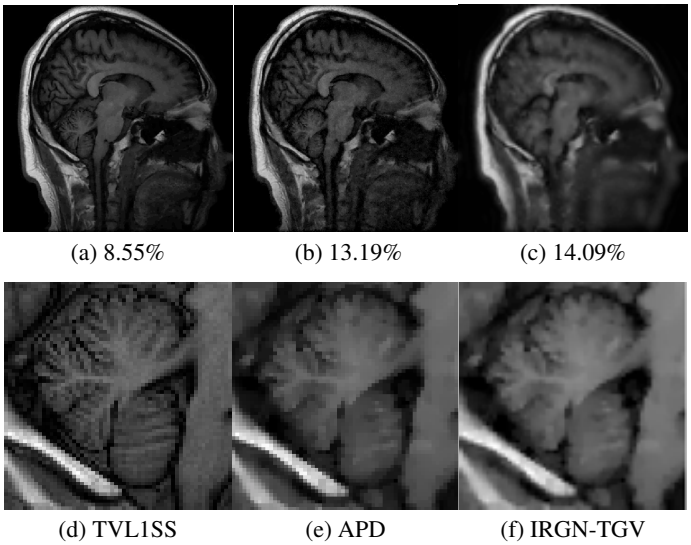
(a) Poisson Mask (b) Reference Image

Fig. 1. *(a) Poisson mask in k-space; (b) Reference Image with red rectangle indicating the zoomed-in box for comparison.*

3.3.2. Robustness

Next, we discuss the robustness of TVL1SS. While the algorithm iterating, the data fitting term is getting smaller, so is δ. Consequently, the relative weights of TV and wavelet regularization terms are reduced automatically along with the iterations, which helps to find the better data fitting. So the algorithm can be less sensitive to the parameters of regularizations.

In Table 2, we fix the parameter μ of $L1$-term as 1×10^{-2} in TVL1SS and fix the initial setting of α_0 as 0.1 in IRGN-TGV. We change λ of TV-term of TVL1SS and APD from 10^{-5} up to 10^{-2} and the initial setting of β_0 from 10^{-1} up to 100 in the same ratio to test the robustness of all the three algorithms from the different choices of parameters. We choose SMAP1 for example.



(a) 8.55% (b) 13.19% (c) 14.09%

(d) TVL1SS (e) APD (f) IRGN-TGV

Fig. 2. *The comparison on SMAP2 of TVL1SS, APD and IRGN-TGV. The first row shows the three numerical results obtained by three algorithm from SMAP2, corresponding to (a)-(c). The second row shows the zoomed-in boxes shown in Figure 1(b) of the reference.*

	λ	10^{-5}	10^{-4}	10^{-3}	10^{-2}
APD	SNR	23.16	22.90	20.93	15.40
	Rel. err.	5.77%	5.93%	7.36%	13.34%
w/o MLE	SNR	24.41	23.80	21.08	15.52
	Rel. err.	5.09%	5.31%	7.20%	13.09%
TVL1SS	SNR	24.43	24.43	24.45	24.41
	Rel. err.	5.07%	5.07%	5.05%	5.09%
	$\beta_0 = \lambda \times 10^4$	10^{-1}	1	10	100
IRGN-TGV	SNR	N.C	N.C.	13.31	12.60
	Rel. err.	N.C.	N.C.	16.44%	17.83%

Table 2:*Robustness of TVL1SS with SMAP1 (N.C.: not convergent)*

In the Table 2, the first and fifth rows indicate the choices of λ and β_0. The numerical results of TVL1SS (in the fourth row of Table 2) are quite similar, but the results of APD and IRGN-TGV vary a lot while parameter changing (the second and the last rows of Table 2). Even for IRGN-TGV, the algorithm does not converge when β_0 is not sufficiently large enough. On the other hand, we run TVL1SS without applying MLE, $i.e.$ the parameters are fixed and not self-adjusted (the third row of Table 2), which is also sensitive to the choice of the regularization parameters. In conclusion, TVL1SS is less sensitive to the choice of parameters, hence much more robust.

4. CONCLUSION

We enhance our previous work to joint estimation of coil sensitivity maps and image reconstruction, and result in faster convergence, higher accuracy and less sensitive to the selections of regularization parameters. Then it practically helps to make up the difficulties of reconstructions with low-quality coil sensitivity maps.

5. REFERENCES

[1] K. Pruessmann, W. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Mag. Reson. Med.*, vol. 46, pp. 638–51, 2001.

[2] K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Mag. Reson. Med.*, vol. 42, pp. 952–62, 1999.

[3] Y. Chen, W. Hager, F. Huang, D. Phan, X. Ye, and W. Yin, "Fast algorithm for image reconstruction with application to partially parallel MR imaging," *SIAM J. Imaging Sci.*, vol. 5, pp. 90–118, January 2012.

[4] J. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model Simul.*, vol. 8(2), pp. 337–69, 2009.

[5] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *CAM Report 08-34*, 2008.

[6] M. Allison, S. Ramani, and J. Fessler, "Regularized mr coil sensitivity estimation using augmented lagrangian methods," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*. IEEE, 2012.

[7] C. McKenzie, E. Yeh, M. Ohliger, M. Price, and D. Sodickson, "Self-calibration parallel imaging with automatic coil sensitivity extraction," *Mag. Reson. Med.*, vol. 47, pp. 529–38, 2002.

[8] A. Raj, Y. Wang, and R. Zabih, "A maximum likelihood approach to parallel imaging with coil sensitivity noise," *IEEE Trans. Med. Imaging*, vol. 26(8), pp. 1046–57, 2007.

[9] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in sense (JSENSE)," *Mag. Reson. Med.*, vol. 57, pp. 1196–1202, 2007.

[10] M. Uecker, T. Hohage, K. T. Block, and J. Frahm, "Image reconstruction by regularized nonlinear inversion–joint estimation of coil sensitivities and image content," *Mag. Reson. Med.*, vol. 60, pp. 674–82, 2008.

[11] F. Knoll, C. Clason, K. Bredies, M. Uecker, and R. Stollberger, "Parallel imaging with nonlinear reconstruction using variational penalties," *Mag. Reson. Med.*, vol. 67, pp. 34–41, 2012.

[12] H. She, R. Chen, D. Liang, E. DiBella, and L. Ying, "Sparse blip: Blind iterative parallel imaging reconstruction using compressed sensing," *Mag. Reson. Med.*, URL: http://dx.doi.org/10.1002/mrm.24716, ISSN: 1522-2594, 2013.

[13] S. Keeling, C. Clason, M. Hinterüller, F. Knoll, A. Laurain, and G. von Winckel, "An image space approach to Cartesian based parallel MR imaging with total variation regularization," *Med. Image Anal.*, vol. 16, pp. 189–200, 2012.

[14] T. Chan, G. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 20, pp. 1964–77, 1999.

[15] X. Ye, Y. Chen, and F. Huang, "Computational acceleration for MR image reconstruction in partially parallel imaging," *IEEE Trans. Med. Imaging*, vol. 30(5), pp. 1055–63, 2011.