FAST MR IMAGE RECONSTRUCTION WITH ORTHOGONAL WAVELET REGULARIZATION VIA SHIFT-VARIANT SHRINKAGE

Matthew J. Muckley and Jeffrey A. Fessler

Department of Biomedical Engineering, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI

ABSTRACT

Algorithms with Lipschitz bounds such as ISTA and FISTA are useful for solving optimization problems with sparsity-promoting regularizers. However, they can be slow in applications that involve shift-variant system matrices. One example of such an application is MRI with multiple sensitivity coils. We propose a reconstruction algorithm for wavelet regularized SENSE MR image reconstruction that exploits the spatial localization of the wavelet basis and the shift-variant behavior of the MR system matrix to accelerate algorithm convergence. Our results indicate that the proposed method is faster than state-of-the-art variable splitting algorithms in terms of convergence speed for a SENSE-type reconstruction problem even when the variable splitting methods are tuned carefully. Unlike variable splitting methods, the proposed method requires no convergence parameter tuning.

***Index Terms*—** MR Image Reconstruction, FISTA, Majorize-Minimize, Parallel MRI

1. INTRODUCTION

The amount of data collected in magnetic resonance imaging (MRI) directly corresponds to the scan time. As a result, undersampling strategies can minimize patient time in the scanner and reduce costs. Parallel imaging based on SENSitivity Encoding (SENSE) [1] facilitates undersampling by exploiting variations in coil sensitivities to remove aliasing patterns caused by undersampling. If the image is assumed to be sparse in some transform domain, then compressed sensing assumptions can be used to facilitate further accelerations [2]. Compressed sensing can be implemented by promoting sparsity in a transform domain through methods such as ℓ_1 regularization.

There are a wide array of algorithms for minimizing cost functions with ℓ_1 regularization. The current state-of-the-art methods are typically based on variable splitting techniques [3, 4, 5, 6]. These techniques rely on reformulating an unconstrained optimization problem as a constrained optimization problem. An augmented Lagrangian (AL) function can be formed for the constrained problem and the algorithm can proceed within the AL formalism [7]. A drawback of AL-based methods is that they require the tuning of penalty parameters that one must select empirically or based on heuristics [6, 8].

An alternative to variable splitting algorithms are majorize-minimize algorithms, which work by forming a surrogate for the cost function and minimizing the surrogate. When a majorize-minimize algorithm is used, momentum techniques can provide further acceleration [9]. The convergence rate of these methods is related to the magnitude of a *Lipschitz constant* that upper bounds the curvature of the Hessian of the data fit term in the original cost function. However, this can be a loose bound when forming a majorizing cost function, particularly in shift-variant problems. In this work, we develop a tighter bound for when the regularizing matrix is an orthogonal wavelet transform. For brevity we only prove the Haar wavelet case, although the theory generalizes to any orthogonal wavelet.

2. THEORY

2.1. Problem Formulation

We are interested in MR image reconstruction problems where the reconstructed image, $\hat{\mathbf{x}}$ is estimated by finding the minimizer of a convex cost function:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \frac{1}{2} \left\| \mathbf{y} - \mathbf{F}\mathbf{S}\mathbf{x} \right\|_2^2 + \beta \left\| \mathbf{W}\mathbf{x} \right\|_1, \qquad (1)$$

where $\mathbf{F}$ is a block diagonal matrix with each block having the same down-sampled DFT operator and $\mathbf{S}$ is a block column matrix with diagonal blocks. Defining $\mathbf{A} = \mathbf{F}\mathbf{S}$, we note that $\mathbf{S}$ gives $\mathbf{A}$ a highly shift-variant nature, a property that we will consider in our algorithm design. $\mathbf{W}$ is an orthogonal wavelet transform. We call $f(\mathbf{x}) = \frac{1}{2} \left\| \mathbf{y} - \mathbf{A}\mathbf{x} \right\|_2^2$ the data fit term and $R(\mathbf{x}) = \left\| \mathbf{W}\mathbf{x} \right\|_1$ the regularizer. For this cost function, the parameter β must be selected by the user to balance trade-offs between the data fit term and the regularizer. Monte Carlo techniques have been developed for estimating these parameters that perform well under mean-squared error metrics [10], but we assume it is fixed.

This work was supported by the University of Michigan MCubed program and NIH grants 1P01 CA87634 and R01 NS 058576.

2.2. VarFISTA: Shift-variant FISTA

FISTA algorithms require decoupling the effects of the system matrix, which can be accomplished via a *majorize-minimize* procedure. This procedure involves forming a surrogate (i.e. majorizer) for the original cost function and then minimizing the surrogate. To develop the majorizer, we will review standard techniques for synthesis majorizer design with a slight modification. Due to the assumption of orthogonal wavelets, we have $\mathbf{x} = \mathbf{W}^H\mathbf{z}$ and we can reformulate the problem in the wavelet basis as

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \ \frac{1}{2}\left\|\mathbf{y} - \mathbf{A}\mathbf{W}^H\mathbf{z}\right\|_2^2 + \beta\left\|\mathbf{z}\right\|_1, \tag{2}$$

where we will now refer to $f(\mathbf{z}) = \frac{1}{2}\left\|\mathbf{y} - \mathbf{A}\mathbf{W}^H\mathbf{z}\right\|_2^2$ as the data fit term. If a surrogate, $\phi_k(\mathbf{z})$ satisfies the following two conditions, then decreasing the surrogate will decrease the original cost function [11]:

$$f(\mathbf{z}^{(k)}) = \phi_k(\mathbf{z}^{(k)}) \tag{3}$$

$$f(\mathbf{z}) \leq \phi_k(\mathbf{z}). \tag{4}$$

We use the index, k, since the surrogate and current estimate may vary with iteration within an iterative algorithm. Defining $\mathbf{B} = \mathbf{A}\mathbf{W}^H$, one can rewrite $f(\mathbf{z})$ around a current estimate, $\mathbf{z}^{(k)}$, as follows:

$$\begin{aligned} f(\mathbf{z}) = f(\mathbf{z}^{(k)}) &+ \mathrm{Re}\{(\mathbf{B}^H(\mathbf{B}\mathbf{z}^{(k)} - \mathbf{y}))^H(\mathbf{z} - \mathbf{z}^{(k)})\} \\ &+ \frac{1}{2}(\mathbf{z} - \mathbf{z}^{(k)})^H\mathbf{B}^H\mathbf{B}(\mathbf{z} - \mathbf{z}^{(k)}), \end{aligned} \tag{5}$$

where $\mathrm{Re}\{\cdot\}$ takes the real part of its input argument. If we have $\mathbf{B}^H\mathbf{B} \preceq \mathbf{D}$ for some diagonal matrix, $\mathbf{D}$, we can write

$$\begin{aligned} f(\mathbf{z}) \leq \phi_k(\mathbf{z}) &= f(\mathbf{z}^{(k)}) \\ &+ \mathrm{Re}\{(\mathbf{B}^H(\mathbf{B}\mathbf{z}^{(k)} - \mathbf{y}))^H(\mathbf{z} - \mathbf{z}^{(k)})\} + \frac{1}{2}\left\|\mathbf{z} - \mathbf{z}^{(k)}\right\|_{\mathbf{D}}^2 \\ &\equiv \frac{1}{2}\left\|\mathbf{z} - (\mathbf{z}^{(k)} - \mathbf{D}^{-1}\mathbf{B}^H(\mathbf{B}\mathbf{z}^{(k)} - \mathbf{y}))\right\|_{\mathbf{D}}^2 \\ &:= \psi_k(\mathbf{z}), \end{aligned} \tag{6}$$

where the second-to-last step comes from completing the square and dropping irrelevant constants. Standard techniques use $\mathbf{D} = L\mathbf{I}$, where L is the Lipschitz constant. We use a more general $\mathbf{D}$. Decreasing $\psi_k(\mathbf{z})$ is the same as decreasing $\phi_k(\mathbf{z})$, which satisfies the monotonicity conditions. This suggests an iterative algorithm for minimizing the cost function in (1):

$$\begin{aligned} \mathbf{z}^{(k+1)} &= \underset{\mathbf{z}}{\operatorname{argmin}} \ \eta(\mathbf{z}, \mathbf{z}^{(k)}) := \{\psi_k(\mathbf{z}) + \beta\left\|\mathbf{z}\right\|_1\} \\ &= \mathit{shrink}(\mathbf{z}^{(k)} - \mathbf{D}^{-1}\mathbf{B}^H(\mathbf{B}\mathbf{z}^{(k)} - \mathbf{y}), \beta\mathbf{d}^{-1}), \end{aligned} \tag{7}$$

where the shrinkage function is defined as $\mathit{shrink}(\mathbf{y}, \beta) = \mathrm{diag}\{\frac{y_j}{|y_j|}\}(|\mathbf{y}| - \beta\mathbf{1})_+$ where $|\cdot|$ denotes the absolute value function and $(\cdot)_+$ sets negative values to 0. Iteratively applying (7) qualifies as a majorize-minimize algorithm, and as such it can be accelerated via momentum techniques [9]. This gives a scheme, which is our proposed method for solving orthogonal wavelet-regularized MR image reconstruction problems (Fig. 1). The inclusion of momentum

Fig. 1. VarFISTA: Shift-variant FISTA

initialize $k = 0, \mathbf{z}^{(0)} = \mathbf{W}\mathbf{x}^{(0)}, \mathbf{u}^{(0)} = \mathbf{z}^{(0)}, \mathbf{D}$
while $k < K$ **do**
  $\tau^{(k+1)} = (1 + \sqrt{1 + 4(\tau^{(k)})^2})/2$
  $\mathbf{z}^{(k+1)} = \underset{\mathbf{z}}{\operatorname{argmin}} \ \eta(\mathbf{z}, \mathbf{u}^{(k)})$
  $\mathbf{u}^{(k+1)} = \mathbf{z}^{(k+1)} + \frac{\tau^{(k)}-1}{\tau^{(k+1)}}(\mathbf{z}^{(k+1)} - \mathbf{z}^{(k)})$
  $k = k + 1$
end while
$\hat{\mathbf{x}} = \mathbf{W}^H\mathbf{z}^{(K)}$

is shown in the tracking of the previous update of $\mathbf{z}$, the use of $\eta(\mathbf{z}, \mathbf{u}^{(k)})$ to update $\mathbf{z}$, and the introduction of a $\tau^{(k)}$ factor. We also investigated the use of a momentum restart scheme, which helps convergence when the current estimate is close to the solution [12] (our restart scheme used $\mathrm{Re}\{\langle \mathbf{u}^{(k+1)} - \mathbf{x}^{(k+1)}, \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \rangle\}$ to estimate the angle). The convergence rate of the proposed method depends on the diagonal matrix, $\mathbf{D}$, the design of which is discussed in the next section.

2.3. Diagonal majorizer design

We focus on designing a small $\mathbf{D}$ such that $\mathbf{D} \succeq \mathbf{B}^H\mathbf{B}$, which we also expect will speed convergence. As stated earlier, we have $\mathbf{B} = \mathbf{F}\mathbf{S}\mathbf{W}^H$, where $\mathbf{F}$ is a Fourier operator, $\mathbf{S}$ is a block column matrix with diagonal blocks, and $\mathbf{W}$ is an orthogonal Haar wavelet transform. The shift-variant aspects of MRI are introduced by the $\mathbf{S}$ matrix. Thus we have $\mathbf{B}^H\mathbf{B} = \mathbf{W}\mathbf{S}^H\mathbf{F}^H\mathbf{F}\mathbf{S}\mathbf{W}^H$. In general we can use $\mathbf{F}^H\mathbf{F} \preceq M\mathbf{I}$, where M is the maximum eigenvalue of $\mathbf{F}^H\mathbf{F}$. In the Cartesian MRI case, $\mathbf{F}^H\mathbf{F}$ is a projection matrix that projects on to the span of the k-space points sampled in the experiment, so $M = 1$ provided unitary DFTs are used. In non-Cartesian MRI, M can be estimated offline via power iteration since $\mathbf{F}$ is not object-dependent, so we can now write

$$\mathbf{B}^H\mathbf{B} \preceq M\mathbf{W}\mathbf{S}^H\mathbf{S}\mathbf{W}^H \tag{8}$$

We note that $\mathbf{S}^H\mathbf{S}$ is diagonal, so to get our final diagonal majorizer for $\mathbf{B}^H\mathbf{B}$ we need to consider the matrix, $\mathbf{W}\mathbf{S}^H\mathbf{S}\mathbf{W}^H$. To do this we will exploit Property 1.

Property 1. *Let* $\mathbf{W}$ *be an orthogonal wavelet transform,* $\mathbf{T}$ *be a diagonal matrix with real, positive entries on the diagonal, and* $\mathbf{t}$ *be a vector with entries from the diagonal of* $\mathbf{T}$*. Let* $\mathcal{S}_q$ *be the set over which the qth wavelet coefficient of the output of* $\mathbf{W}$ *has support. Define* $\mathbf{D} = \mathit{diag}\{d_q\}$ *where*

$d_q = max_{l \in \mathcal{S}_q}(t_l)$, *where* t_l *is the* l*th entry of* $\mathbf{t}$. *Then we have* $\mathbf{D} \succeq \mathbf{W}\mathbf{T}\mathbf{W}^H$.

For brevity we focus on the case of Haar wavelets here; the derivation can be generalized to any unitary matrix by recognizing that an inner product between two compactly supported vectors can be computed over either vector's support. A Haar wavelet transform can be expressed as a cascade, i.e. $\mathbf{W} = \mathbf{W}_N ... \mathbf{W}_1$, where N denotes the number of levels of the decomposition. With a suitable permutation, each matrix for each step of the cascade is block diagonal, i.e.

$$\mathbf{W}_1 = \mathbf{I} \otimes \mathbf{W}_{1,s}, \tag{9}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{W}_{1,s}$ is a unitary matrix that applies a transform to a small patch of the image. For 2D Haar wavelets, we have

$$\mathbf{W}_{1,s} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{10}$$

In 3D $\mathbf{W}_{1,s}$ would be an 8×8 unitary matrix. A property of Haar wavelets is that over any individual step of the cascade, the blocks do not overlap. Let $\mathbf{T}_j$ denote the jth patch of the diagonal matrix, $\mathbf{T}$; then we have

$$\mathbf{W}_1 \mathbf{T} \mathbf{W}_1^H = \begin{bmatrix} \mathbf{W}_{1,s}\mathbf{T}_1\mathbf{W}_{1,s}^H & & \\ & \ddots & \\ & & \mathbf{W}_{1,s}\mathbf{T}_J\mathbf{W}_{1,s}^H \end{bmatrix}. \tag{11}$$

We note that $\mathbf{T}_j \preceq d_j \mathbf{I}$ where $d_j = \max\{\mathbf{T}_j\}$ and $\max\{\cdot\}$ takes the elementwise maximum from a matrix with real entries, and so

$$\begin{aligned} \mathbf{W}_1 \mathbf{T} \mathbf{W}_1^H &\preceq \begin{bmatrix} \mathbf{W}_{1,s} d_1 \mathbf{I} \mathbf{W}_{1,s}^H & & \\ & \ddots & \\ & & \mathbf{W}_{1,s} d_J \mathbf{I} \mathbf{W}_{1,s}^H \end{bmatrix} \\ &= \begin{bmatrix} d_1\mathbf{I} & & \\ & \ddots & \\ & & d_J\mathbf{I} \end{bmatrix} := \mathbf{D}_1, \end{aligned} \tag{12}$$

using the assumption of unitary wavelets. We now have $\mathbf{W}_N ... \mathbf{W}_1 \mathbf{T} \mathbf{W}_1^H ... \mathbf{W}_N^H \preceq \mathbf{W}_N ... \mathbf{W}_2 \mathbf{D}_1 \mathbf{W}_2^H ... \mathbf{W}_N^H$. We apply the same procedure to $\mathbf{W}_2 \mathbf{D}_1 \mathbf{W}_2^H$. Since $\mathbf{W}_2$ is effectively an identity operator over the detail coefficients produced by $\mathbf{W}_1$ and $\mathbf{D}_1$ is diagonal, only the approximation coefficients from $\mathbf{W}_1$ need to be dealt with further. Applying the same procedure for $\mathbf{W}_2$ as for $\mathbf{W}_1$ over the approximation coefficients from $\mathbf{W}_1$ leads to taking maximums over larger patch sizes. Recursively applying this procedure through $\mathbf{W}_N$ gives Property 1. Applying Property 1 with $\mathbf{T} = M\mathbf{S}^H\mathbf{S}$ is our choice for designing a separable majorizer in SENSE-type MRI with orthogonal wavelet regularization.

3. EXPERIMENTS

3.1. Experimental setup

We reproduced the *in vivo* experiments from a previous study that used variable splitting techniques for SENSE-type MRI [6]. We compared the proposed VarFISTA method to state-of-the-art variable splitting methods for MR image reconstruction problems involving orthogonal Haar and Daubechies D4 wavelets. We also show results using FISTA with $\mathbf{D} = L\mathbf{I}$ (denoted simply FISTA), to illustrate the benefits provided by a more general diagonal majorizer. The state-of-the-art variable splitting methods were each of the AL-P1 type in a previous work [6], which had similar speed to AL-P2 but is simpler to optimize for examining optimal performance since it has fewer penalty parameters.

To track convergence, we computed the following normalized residual as a function of iteration:

$$\xi(k) = 20 \log_{10} \left(\frac{\left\| \mathbf{x}^{(k)} - \mathbf{x}^{(\infty)} \right\|_2}{\left\| \mathbf{x}^{(\infty)} \right\|_2} \right), \tag{13}$$

where $\mathbf{x}^{(\infty)}$ is a "converged" solution obtained by running many thousands of iterations of VarFISTA (the convergence theory in [9] generalizes). We also stored the time at which the kth estimate was formed and in our figures we plot $\xi(\cdot)$ as a function of CPU time instead of iteration. We choose to do this since iterations of the proposed majorize-minimize methods and the variable splitting methods have drastically different compute times due to the conjugate gradient (CG) subroutine. AL-P1 iterations were about five times as slow due to the five CG subiterations. The methods both used identical subroutines for matrix multiplication. Regularization parameters were selected to give visually appealing solutions, although in practice the regularization parameter could be estimated via Monte Carlo SURE methods, which have previously been proposed for MRI [10].

3.2. In vivo human brain experiments

For the *in vivo* experiment, a 3D data set was acquired on a GE 3T scanner with an 8-channel head coil with acquisition parameters $T_R = 25$ ms, $T_E = 5.172$ ms, and voxel size 1 mm $\times$ 1.35 mm $\times$ 1 mm. The data matrix size was $256 \times 144 \times 128$ uniformly spaced samples. Sensitivity maps were estimated using quadratic regularized least squares routine similar to [13]. The data was sampled in the Fourier domain using a Poisson disk sampling scheme [14] with a fully sampled center (32-by-32 block), which has been demonstrated to be useful in compressed sensing MRI applications [15]. Only 20% of the full DFT sampling was used for reconstruction.

Figure 2c and Figure 2d show the results for Haar and Daubechies D4 wavelets, respectively. VarFISTA with restart is the fastest method in all cases, showing the benefits of combining a general diagonal majorizer and momentum restarting. The VarFISTA method without restart is fast in early

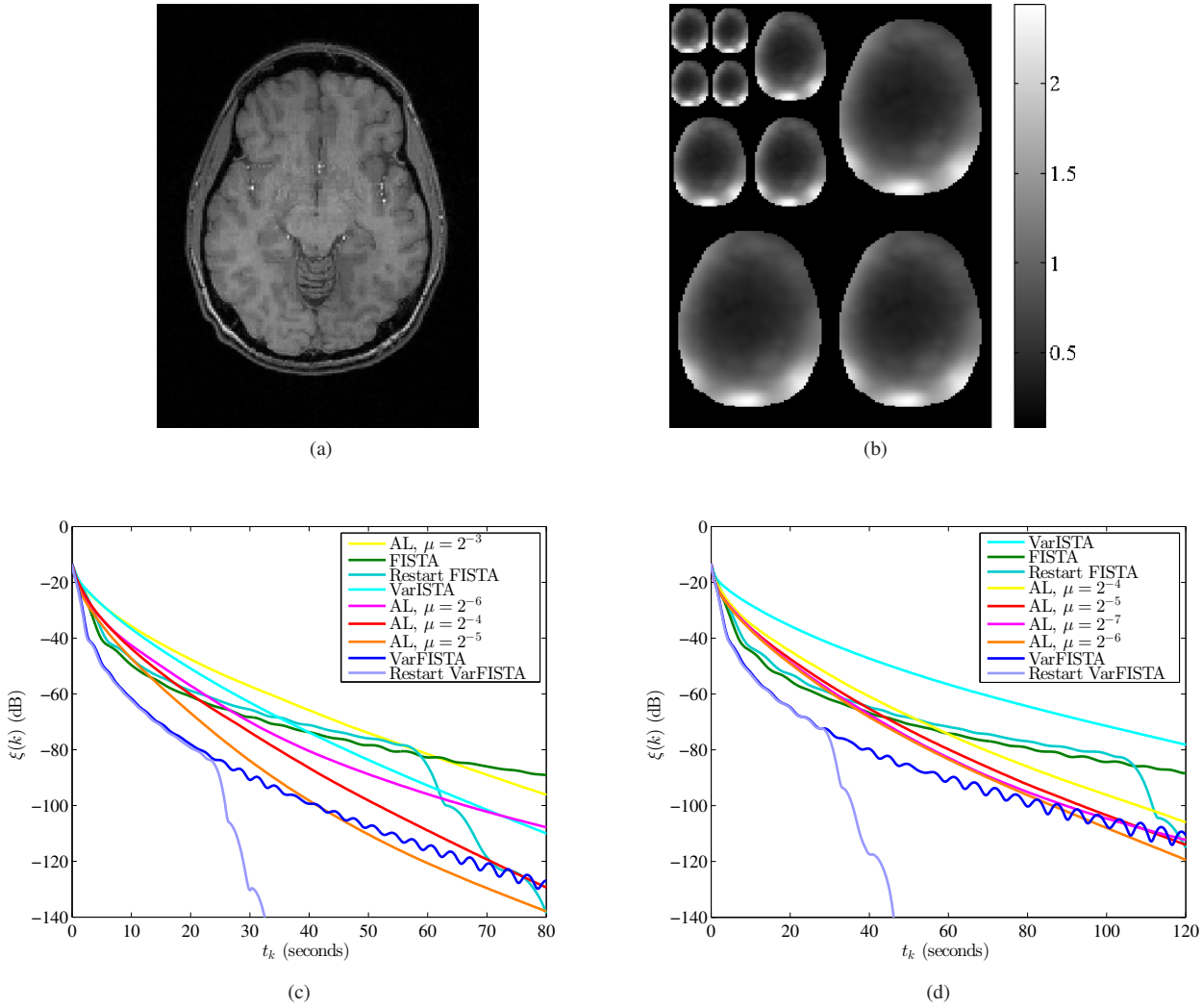

Fig. 2. Images corresponding to the wavelet regularized *in vivo* experiment. (a) $\mathbf{x}^{(\infty)}$ for the Haar wavelet regularized experiment. (b) The diagonal elements of $\mathbf{D}$ for the Haar wavelet regularized reconstruction problem arranged into a wavelet decomposition with areas outside the brain have been masked for presentation. (c) Convergence plot comparing the proposed method with variable splitting methods for Haar wavelets. The proposed method with momentum restarting is faster than the other methods. (d) Another convergence plot with Daubechies D4 wavelets.

iterations, but slows down when nearing the solution. FISTA with restart is slow in early iterations; the AL methods are nearing -120 dB before FISTA undergoes its first restart. We show a variety of AL tuning parameters to demonstrate that AL methods seem to universally be slower than the proposed method, although we cannot make any theoretical conclusions since we do not know of any method to optimally select the variable splitting parameter.

4. CONCLUSION

The convergence speed of AL splitting methods depends heavily on the choice of penalty parameter. We currently know of no sufficient conditions on selecting these parameters to ensure optimal convergence speed. We have shown that when tight diagonal majorizers can be designed in the basis of the regularizer for synthesis-type problems, majorize-minimize algorithms with momentum restarting can surpass current state-of-the-art variable splitting methods even with penalty parameter optimization, whereas the proposed method requires no parameter tuning. Analysis-type regularizers may give better reconstructions than the synthesis regularizers used here [16, 17]; we have generalized the techniques shown here for the analysis regularizers of total variation and undecimated wavelets in manuscript submitted for journal publication [18].

5. REFERENCES

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

[2] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Sig. Proc. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.

[3] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–43, 2009.

[4] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE J. Sel. Top. Sig. Proc.*, vol. 4, no. 2, pp. 288–97, Apr. 2010.

[5] M. V. Afonso, José. M. Bioucas-Dias, and Mário. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Im. Proc.*, vol. 19, no. 9, pp. 2345–56, Sep. 2010.

[6] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, Mar. 2011.

[7] D. P. Bertsekas, "On the Goldstein-Levitin-Polyak gradient projection method," *IEEE Trans. Auto. Control*, vol. 21, no. 2, pp. 174–84, Apr. 1976.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. & Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[9] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[10] S. Ramani, Z. Liu, J. Rosen, J.-F. Nielsen, and J. A. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods," *IEEE Trans. Im. Proc.*, vol. 21, no. 8, pp. 3659–72, Aug. 2012.

[11] 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.

[12] B. O'Donoghue and E. Candès, "Adaptive restart for accelerated gradient schemes," *Found. Computational Math.*, 2014.

[13] S. L. Keeling and R. Bammer, "A variational approach to magnetic resonance coil sensitivity estimation," *Applied Mathematics and Computation*, vol. 158, no. 2, pp. 53–82, Nov. 2004.

[14] D. Dunbar and G. Humphreys, "A spatial data structure for fast Poisson-disk sample generation," *ACM Trans. on Graphics*, vol. 25, no. 3, pp. 503–8, Jul. 2006, sIGGRAPH.

[15] M. Lustig, M. Alley, S. Vasanawala, D. L. Donoho, and J. M. Pauly, "L1 SPIR-iT: Autocalibrating parallel imaging compressed sensing," in *Proc. Intl. Soc. Mag. Res. Med.*, 2009, p. 334.

[16] I. W. Selesnick and M. A. T. Figueiredo, "Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors," in *Proc. SPIE 7446 Wavelets XIII*, 2009, p. 74460D, wavelets XIII.

[17] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Prob.*, vol. 23, no. 3, pp. 947–68, Jun. 2007.

[18] M. J. Muckley, D. C. Noll, and J. A. Fessler, "Fast parallel MR image reconstruction via B1-based, adaptive restart, iterative soft thresholding algorithms (BARISTA)," *IEEE Trans. Med. Imag.*, 2014, submitted.