A variable splitting based algorithm for Fast Multi-Coil Blind Compressed Sensing MRI reconstruction

Sampada Bhave[1], Sajan Goud Lingala[2] and Mathews Jacob[1]

***Abstract*— Recent work on blind compressed sensing (BCS) has shown that exploiting sparsity in dictionaries that are learnt directly from the data at hand can outperform compressed sensing (CS) that uses fixed dictionaries. A challenge with BCS however is the large computational complexity during its optimization, which limits its practical use in several MRI applications. In this paper, we propose a novel optimization algorithm that utilize variable splitting strategies to significantly improve the convergence speed of the BCS optimization. The splitting allows us to efficiently decouple the sparse coefficient, and dictionary update steps from the data fidelity term, resulting in subproblems that take closed form analytical solutions, which otherwise require slower iterative conjugate gradient algorithms. Through experiments on multi coil parametric MRI data, we demonstrate the superior performance of BCS over conventional CS schemes, while achieving convergence speed up factors of over 10 fold over the previously proposed implementation of the BCS algorithm.**

I. INTRODUCTION

Over the recent years, compressed sensing (CS) schemes have shown considerable potential to accelerate MRI acquisition. CS exploits sparse representation of data in a known dictionary bases. For instance, wavelet bases in [1] and temporal Fourier bases in [2] have been used in static and dynamic MRI applications. A challenge in using such predetermined dictionaries often lies with the misfit between the representation and the data; many coefficients are often required for an accurate representation. For instance, in free breathing perfusion MRI, many temporal Fourier bases are required to represent the temporal dynamics of the data, thereby restricting the maximum achievable acceleration factor.

Recently, several researchers have proposed to jointly estimate the sparse representations and the dictionaries from the under sampled data at hand. Dictionaries containing atoms of one-dimensional non-orthogonal temporal bases [3] or two-dimensional spatial patches [4] have been proposed for dynamic and static applications. These schemes termed as blind compressed sensing (BCS) have shown considerable promise over conventional CS schemes in several MRI applications such as dynamic contrast enhanced MRI [3], functional lung [5], parametric, and high resolution static MRI [4].

The BCS scheme is formulated as a constrained optimization problem consisting of linear combination of data fidelity l_2 norm, and a sparsity promoting norm on the coefficients subject to a Frobenius norm constraint on the dictionary. The single coil BCS algorithmic implementation in [3] relies on iterations between quadratic update steps of the coefficient and dictionary update steps. These steps were solved using slow iterative conjugate gradient algorithms due to the complexity in constructing the inverses of the matrices resulting from the data fidelity. The large computational complexity encountered during the optimization pose a challenge in dealing with large practical datasets such as multi-coil and three-dimensional applications.

In this work, we propose to employ variable splitting strategies [6] to efficiently decouple the coefficient update and dictionary update steps from the data fidelity term. Through the decoupling, we develop a computational efficient algorithm that efficiently cycles between different steps that have analytical closed form solutions. We demonstrate through experiments on multi coil parametric MRI data, the proposed algorithm obtains a significant speed up factor of ten fold over the previous implementation.

II. BACKGROUND

A. BCS Model representation

In parametric imaging, the k-space corresponds to several images acquired at different values of encoding parameters (such as echo time, spin lock time, flip angle) denoted by p. We model the multi-coil undersampled measurement as:

$$\mathbf{b}(\mathbf{k},p) = \underbrace{\mathbf{SFC}}_{\mathcal{A}}[\gamma(\mathbf{x},p)] + \mathbf{n}(\mathbf{k},p), \tag{1}$$

where $\mathbf{b}(\mathbf{k},p)$ represents the concatenated vector of the $\mathbf{k}-p$ measurements from all the coils. $\gamma(\mathbf{x},p);(\mathbf{x} = (x_1,y_1))$ denote the underlying images pertaining to different contrasts; $\mathbf{n}$ is additive noise. $\mathcal{A}$ is the operator that models coil sensitivity $\mathbf{C}$ and Fourier encoding $\mathbf{F}$ on a specified $\mathbf{k}-p$ sampling trajectory $\mathbf{S}$. The dataset is represented as $M \times N$ Casoriti matrix [7] $\mathbf{\Gamma}_{M\times N}$

$$\mathbf{\Gamma}_{M\times N} = \begin{pmatrix} \gamma(\mathbf{r}_1,c_1) & . & \gamma(\mathbf{r}_1,c_N) \\ . & . & . \\ . & . & . \\ \gamma(\mathbf{r}_M,c_1) & . & \gamma(\mathbf{r}_M,c_N) \end{pmatrix} \tag{2}$$

where M is the number of voxels in the image and N is the number of encoding parameters.

B. Image reconstruction

We model $\mathbf{\Gamma}$ as a product of spatial coefficients $\mathbf{U}_{M\times R}$ and dictionary $\mathbf{V}_{R\times N}$. The joint recovery of $\mathbf{U}$, $\mathbf{V}$ is

[1]Sampada Bhave and Mathews Jacob are with the Department of Electrical and Computer Engineering, The University of Iowa, Iowa, 52242 USA sampada-bhave at uiowa.edu and mathews-jacob at uiowa.edu

[2]Sajan Goud Lingala is with the Department of Electrical Engineering, University of Southern California, California, USA

formulated as a constrained optimization problem:

$$\underset{\mathbf{U},\mathbf{V}}{\arg\min} \underbrace{\|\mathcal{A}(\mathbf{UV}) - \mathbf{b}\|_F^2 + \lambda\|\mathbf{U}\|_{\ell_p}}_{\mathcal{C}(\mathbf{U},\mathbf{V})} \text{ s. t. } \|\mathbf{V}\|_F^2 < 1. \quad (3)$$

The first term ensures data consistency. The second term promotes sparsity on the spatial coefficients by using a non-convex $\ell_p; p < 1$ semi-norm prior on $\mathbf{U}$. A unit Frobenius norm is imposed on the dictionary $\mathbf{V}$ making the recovery problem well posed.

C. Algorithm 1: Without using variable splitting

We majorize an approximation of the ℓ_p penalty on $\mathbf{U}$ in Eq. (3) as $\|\mathbf{U}\|_{\ell_p} \approx \min_{\mathbf{L}} \frac{\beta}{2}\|\mathbf{U}-\mathbf{L}\|^2 + \|\mathbf{L}\|_{\ell_p}$, where $\mathbf{L}$ is an auxiliary variable. This approximation becomes exact as $\beta \to \infty$. We use augmented Lagrangian optimization scheme to enforce the constraint $\mathbf{V} = \mathbf{Q}$, where $\mathbf{Q}$ is the auxiliary variable for $\mathbf{V}$. Thus, the optimization problem is given by

$$\{\mathbf{U}^*, \mathbf{V}^*\} = \arg\min_{\mathbf{U},\mathbf{V}} \min_{\mathbf{Q},\mathbf{L}} \|\mathcal{A}(\mathbf{UV}) - \mathbf{b}\|_F^2 + \frac{\beta\lambda}{2}\|\mathbf{U}-\mathbf{L}\|_F^2 + \lambda\|\mathbf{L}\|_{\ell_p} \text{ s. t. } \|\mathbf{Q}\|_F^2 < 1, \mathbf{V} = \mathbf{Q} \quad (4)$$

We use an alternating strategy to solve for the variables $\mathbf{U}, \mathbf{V}, \mathbf{Q}$ and $\mathbf{L}$. The $\mathbf{L}$ and $\mathbf{Q}$ subproblems are solved analytically. Eq. 4 is quadratic in $\mathbf{U}$ and $\mathbf{V}$. Iterative methods like conjugate gradient are required to solve these subproblems due to the enormous size of A. In addition, for the majorization to well approximate the ℓ_p penalty, β needs to be a high value. At higher values of β, the condition number of these subproblems is significantly high resulting in slow convergence as many iterations of CG are required.

III. Proposed Algorithm : Using variable splitting

To improve convergence speed, Ramani and Fessler proposed the use of the technique of variable splitting to decouple the effect of coil sensitivities $\mathbf{C}$ and the regularization [6]. We introduce a novel optimization algortihm using variable splitting technique to accelerate the convergence of Eq. 4. First, we decouple the data fidelity term from sparse coefficients $\mathbf{U}$ and dictionary $\mathbf{V}$ by introducing a constraint $\mathbf{X} = \mathbf{UV}$ where $\mathbf{X}$ is the auxiliary variable for $\mathbf{UV}$. The optimization problem is of the form

$$\underset{\mathbf{U},\mathbf{V},\mathbf{X}}{\arg\min} \|\mathbf{SFCX} - \mathbf{b}\|_F^2 + \lambda\|\mathbf{U}\|_{\ell_p} \text{ s. t. } \mathbf{X} = \mathbf{UV}, \|\mathbf{V}\|_F^2 < 1 \quad (5)$$

We further decouple the coil sensitivities from $\mathbf{X}$ by introducing another constraint $\mathbf{Z}$=$\mathbf{CX}$ where $\mathbf{Z}$ is the auxiliary variable. The constrained optimization problem can be written as

$$\underset{\mathbf{U},\mathbf{V},\mathbf{X},\mathbf{Z}}{\arg\min} \|SF\mathbf{Z} - \mathbf{b}\|_F^2 + \lambda\|\mathbf{U}\|_{l_p} \text{ s.t. } \mathbf{X} = \mathbf{UV}, \|\mathbf{V}\|_F^2 < 1, \mathbf{Z} = \mathbf{CX} \quad (6)$$

We majorize an approximation of the ℓ_p penalty on $\mathbf{U}$ in (3) as $\|\mathbf{U}\|_{\ell_p} \approx \min_{\mathbf{L}} \frac{\beta}{2}\|\mathbf{U}-\mathbf{L}\|^2 + \|\mathbf{L}\|_{\ell_p}$, where $\mathbf{L}$ is an auxiliary variable. We enforce the constraints in Eq. 6 by using an augmented Lagrangian (AL) framework [8]. The associated AL function is written as

$$\begin{aligned}\mathcal{L}(\mathbf{U},\mathbf{V},\mathbf{L},\mathbf{Q},\mathbf{X},\mathbf{Z}) = \|\mathbf{SFZ} - \mathbf{b}\|_F^2 + \frac{\beta_X}{2}\|\mathbf{X}-\mathbf{UV}\|_F^2 \\ + \mathbf{\Lambda}_X'(\mathbf{X}-\mathbf{UV}) + \lambda\|\mathbf{L}\|_{\ell_p} + \frac{\lambda\beta_U}{2}\|\mathbf{U}-\mathbf{L}\|_F^2 \\ + \frac{\beta_V}{2}\|\mathbf{V}-\mathbf{Q}\|_F^2 + \mathbf{\Lambda}_V'(\mathbf{V}-\mathbf{Q}) \\ + \frac{\beta_Z}{2}\|\mathbf{Z}-\mathbf{CX}\|_F^2 + \mathbf{\Lambda}_Z'(\mathbf{Z}-\mathbf{CX}) \text{ s. t. } \|\mathbf{Q}\|_F^2 < 1\end{aligned} \quad (7)$$

Here $\mathbf{Q}$ is the auxiliary variable for $\mathbf{V}$, $\mathbf{\Lambda}_X, \mathbf{\Lambda}_V$ and $\mathbf{\Lambda}_Z$ are the Lagrange multipliers. $\beta_X, \beta_V, \beta_U$ and β_Z are the penalty parameters. We use an alternating strategy to solve for the variables $\mathbf{U}, \mathbf{V}, \mathbf{Q}, \mathbf{L}, \mathbf{X}$ and $\mathbf{Z}$. All of these subproblems are solved analytically as described below, by minimizing the Eq. 7 with respect to these variables one at a time assuming the other variables to be fixed.

L subproblem: Ignoring all the terms independent of $\mathbf{L}$, Eq. 7 can be written as

$$\arg\min_{\mathbf{L}} \frac{\lambda\beta_U}{2}\|\mathbf{U}-\mathbf{L}\|_F^2 + \lambda\|\mathbf{L}\|_{\ell_p} \quad (8)$$

The $\mathbf{L}$ subproblem is solved using shrinkage rule as

$$\mathbf{L}_{n+1} = \frac{\mathbf{U}}{|\mathbf{U}_n|}\left(|\mathbf{U}_n| - \frac{1}{\beta}|\mathbf{U}_n|^{p-1}\right)_+ \quad (9)$$

where '+' represents the soft thresholding operator defined as $(\tau)_+ = \max\{0, \tau\}$ and β_U is the penalty parameter.

U subproblem: The minimization of Eq. 7 with respect to $\mathbf{U}$ results in a quadratic subproblem which has an closed form solution given by

$$\arg\min_{\mathbf{U}} \frac{\beta_X}{2}\|\mathbf{X}-\mathbf{UV}\|_F^2 + \mathbf{\Lambda}_X'(\mathbf{X}-\mathbf{UV}) + \frac{\lambda\beta_U}{2}\|\mathbf{U}-\mathbf{L}\|_F^2 \quad (10)$$

The quadratic subproblem can be solved analytically as shown below.

$$\mathbf{U}_{n+1} = (\beta_{X_n}\mathbf{X}_n\mathbf{V}_n' + \mathbf{\Lambda}_X\mathbf{V}_n' + \lambda\beta_U\mathbf{L}_{n+1})\mathbf{H}_U^{-1} \quad (11)$$

$$\mathbf{H}_U = \beta_X\mathbf{V}_n\mathbf{V}_n' + \lambda\beta_U I \quad (12)$$

Q subproblem: The $\mathbf{Q}$ subproblem is obtained by minimizing Eq. 7 with respect to Q

$$\arg\min_{\mathbf{Q}} \frac{\beta_V}{2}\|\mathbf{V}-\mathbf{Q}\|_F^2 + \mathbf{\Lambda}_V'(\mathbf{V}-\mathbf{Q}) \text{ s. t. } \|\mathbf{Q}\|_F^2 < 1 \quad (13)$$

The above problem is solved using a projection scheme as specified in Eq. 14. If Frobenius norm of $\mathbf{Q}$ is less than unity, we set $\mathbf{Q} = \mathbf{V}$, else we scale $\mathbf{V}$ to have a unit Frobenius norm

$$\mathbf{Q}_{n+1} = \begin{cases} V_n & \|V_n\|_F^2 \leqslant 1 \\ \frac{1}{\|V_n\|_F}V_n & \text{else} \end{cases} \quad (14)$$

Note that $\mathbf{Q}_n$ is obtained by scaling $\mathbf{V}_n$ so that the Frobenius norm is unity.

V subproblem: The $\mathbf{V}$ subproblem is a quadratic subproblem as shown below.

$$\begin{aligned}\arg\min_{\mathbf{V}} \frac{\beta_X}{2}\|\mathbf{X}-\mathbf{UV}\|_F^2 + \mathbf{\Lambda}_X'(\mathbf{X}-\mathbf{UV}) \\ + \frac{\beta_V}{2}\|\mathbf{V}-\mathbf{Q}\|_F^2 + \mathbf{\Lambda}_V'(\mathbf{V}-\mathbf{Q})\end{aligned} \quad (15)$$

Minimization of the above equation with respect to $\mathbf{V}$ yields the following closed form solution

$$\mathbf{V}_{n+1} = \mathbf{H}_V^{-1}(\beta_X \mathbf{U}'_{n+1}\mathbf{X}_n + \mathbf{U}'_{n+1}\mathbf{\Lambda}_X + \beta_V \mathbf{Q}_{n+1} - \mathbf{\Lambda}_V) \tag{16}$$

$$\mathbf{H}_V = (\beta_X \mathbf{U}'_{n+1}\mathbf{U}_{n+1} + \beta_V I) \tag{17}$$

$\underline{\mathbf{X}}$ subproblem: Dropping all the terms independent of $\mathbf{X}$ we get

$$\begin{aligned}\arg\min_{\mathbf{X}} \ & \frac{\beta_X}{2}\|\mathbf{X} - \mathbf{U}\mathbf{V}\|_F^2 + \mathbf{\Lambda}'_X(\mathbf{X} - \mathbf{U}\mathbf{V}) \\ & + \frac{\beta_Z}{2}\|\mathbf{Z} - \mathbf{C}\mathbf{X}\|_F^2 + \mathbf{\Lambda}'_Z(\mathbf{Z} - \mathbf{C}\mathbf{X})\end{aligned} \tag{18}$$

The closed form solution to the above minimization problem is given by

$$\mathbf{X}_{n+1} = \mathbf{H}_X^{-1}(\beta_X \mathbf{U}_{n+1}\mathbf{V}_{n+1} - \mathbf{\Lambda}_X + \beta_Z \mathbf{C}'\mathbf{Z}_n + \mathbf{C}'\mathbf{\Lambda}_Z) \tag{19}$$

$$\mathbf{H}_X = \beta_X I + \beta_Z \mathbf{C}'\mathbf{C} \tag{20}$$

$\underline{\mathbf{Z}}$ subproblem: Writing the Eq. 7 with respect to $\mathbf{Z}$ (ignoring constants independent of $\mathbf{Z}$) we get,

$$\arg\min_{\mathbf{Z}} \ \|SF\mathbf{Z} - \mathbf{b}\|_F^2 + \frac{\beta_Z}{2}\|\mathbf{Z} - \mathbf{C}\mathbf{X}\|_F^2 + \mathbf{\Lambda}'_Z(\mathbf{Z} - \mathbf{C}\mathbf{X}) \tag{21}$$

This problem is a Fourier domain replacement problem which can be solved analytically as shown below.

$$\mathbf{Z}_{n+1} = \mathbf{F}' \left[\left(S + \frac{\beta_Z}{2} I \right)^{-1} \mathbf{F} \left(\frac{\beta_Z}{2} C\mathbf{X}_{n+1} - \frac{\mathbf{\Lambda}_Z}{2} + \mathbf{F}'\mathbf{S}\mathbf{b} \right) \right] \tag{22}$$

We update all the Lagrange multipliers using a steepest ascent method at each iteration as shown below.

$$\mathbf{\Lambda}_{V(n+1)} = \mathbf{\Lambda}_{Vn} + \beta_V \ (\mathbf{V}_{n+1} - \mathbf{Q}_{n+1}) \tag{23}$$

$$\mathbf{\Lambda}_{X(n+1)} = \mathbf{\Lambda}_{Xn} + \beta_X \ (\mathbf{X}_{n+1} - \mathbf{U}_{n+1}\mathbf{V}_{n+1}) \tag{24}$$

$$\mathbf{\Lambda}_{Z(n+1)} = \mathbf{\Lambda}_{Zn} + \beta_Z \ (\mathbf{Z}_{n+1} - \mathbf{C}_{n+1}\mathbf{X}) \tag{25}$$

The optimization algorithm in Eq. 6 is solved by cycling between the above subproblems. The matrix $\mathbf{H}_U$ and $\mathbf{H}_V$ are $R \times R$ and can be easily inverted. Since $\mathbf{C}'\mathbf{C}$ is a diagonal, $\mathbf{H}_X$ matrix is also diagonal and is therefore easily inverted. Splitting the k-space and coil sensitivities ($\mathbf{F}$ and $\mathbf{C}$ components) from $\mathbf{UV}$ in the data fidelity term has led to separate matrix inverses involving $\mathbf{F}'\mathbf{F}$ and $\mathbf{C}'\mathbf{C}$ which are easier to compute. Since, all the steps can be solved analytically, the convergence is much faster than conventional iterative conjugate gradient steps. Although β_X and β_Z parameters do not affect the final solution, they can affect the convergence rate. These parameters were chosen empirically. Since we use the augmented Lagrangian framework for enforcing the constraint on the dictionary, it is not necessary for β_V to tend to ∞ for the constraint to hold, allowing faster convergence. The quality of reconstruction is affected by β_U parameters as the non-convex penalty is enforced using majorization. As discussed earlier, the majorization is only exact when $\beta_U \to \infty$. We initialize β_U to a small value and gradually increment it when the cost in Eq. 3 stagnates to a threshold level of 10^{-2}.

The pseudo-code of the algorithm is shown below.

Algorithm III.1: FASTBCS($\mathbf{S}, \mathbf{F}, \mathbf{C}, \mathbf{b}, \lambda$)

Input : $\mathbf{b}$, $\beta_X > 0$, $\beta_Z > 0$, $\beta_U > 0$
while $|\mathcal{C}_n - \mathcal{C}_{n-1}| > 10^{-5}\,\mathcal{C}_n$
do
- Initialize $\beta_V > 0$
- **while** $|\mathbf{V} - \mathbf{Q}|^2 > 10^{-5}$
  do
  - Update $\mathbf{L}$:← $Eq.[9]$
  - Update $\mathbf{U}$:← $Eq.[11]$
  - Update $\mathbf{Q}$:← $Eq.[14]$
  - Update $\mathbf{V}$:← $Eq.[16]$
  - Update $\mathbf{Z}$:← $Eq.[22]$
  - Update $\mathbf{X}$:← $Eq.[19]$
  - Update $\mathbf{\Lambda}_V$:← $Eq.[23]$
  - Update $\mathbf{\Lambda}_X$:← $Eq.[24]$
  - Update $\mathbf{\Lambda}_Z$:← $Eq.[25]$
  - $\beta_V = 5 * \beta_V$
- **if** $|\mathcal{C}_n - \mathcal{C}_{n-1}| < 10^{-2}\,\mathcal{C}_n$ **then** $\{\beta_U = 50 * \beta_U$

return $(\mathbf{U}, \mathbf{V})$

IV. Experimental Evaluation

To study the convergence rate of both the algorithms, we acquired a single slice fully sampled 2D dataset on a Siemens 3T Trio scanner using a turbo spin echo (TSE) sequence with turbo factor (TF) of 8, matrix size = 128x128, FOV= 22x22cm^2, TR=2500ms, slice thickness =5mm, B_1 spin lock frequency=400Hz, bandwidth= 130Hz/pixel, and echo spacing of 12.2ms. A 12-channel phased array coil was used. $T_{1\rho}$ preparation pulse [9] and T_2 preparation pulse [10] were used prior to readout. The data was collected for 12 equispaced spin lock times (TSLs) and 12 equispaced echo times (TEs) values, both ranging from 10 ms to 120 ms. This provided a total of 24 parametric measurements. The dataset was undersampled using a Cartesian hybrid sampling pattern (acceleration of 1.5 using pseudo-random variable density sampling and an acceleration factor of 4 from uniform sampling pattern) giving a net acceleration of 6. The coil sensitivity maps were obtained using Walsh method for coil map estimation [11]. Both the algorithms were implemented in MATLAB on a quad core linux machine.

We compare the performance and convergence speed of both the algorithms using a Mean square error (MSE) metric given by

$$\text{MSE} = \left(\frac{\|\Gamma_{\mathbf{recon}} - \Gamma_{\mathbf{orig}}\|_F^2}{\|\Gamma_{\mathbf{orig}}\|_F^2} \right). \tag{26}$$

The regularization parameter λ of both the algorithms was chosen such that the error between reconstructions and the fully sampled data given by MSE was minimized. Comparisons were done for the optimal λ value of 0.05.

The reconstruction error vs CPU time is shown in Fig. 1(a). It is observed that proposed algorithm converges to almost the same solution in just 3 min. This is also demonstrated by the reconstructed images and $T_{1\rho}$ parameter maps shown in Fig. 2. Every time, when the β in Eq. 4 is incremented, the condition number of the $\mathbf{U}$ subproblem increases

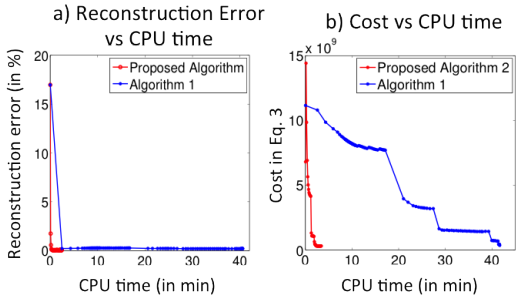

Fig. 1. Convergence plots: Reconstruction error vs CPU time and the cost vs CPU time plots are shown in a) and b) respectively. It is observed that proposed algorithm converges to same solution in about 3 min where as the Algorithm 1 takes roughly 40 min to converge. Note the threshold in both algorithms was set to 10^{-6}

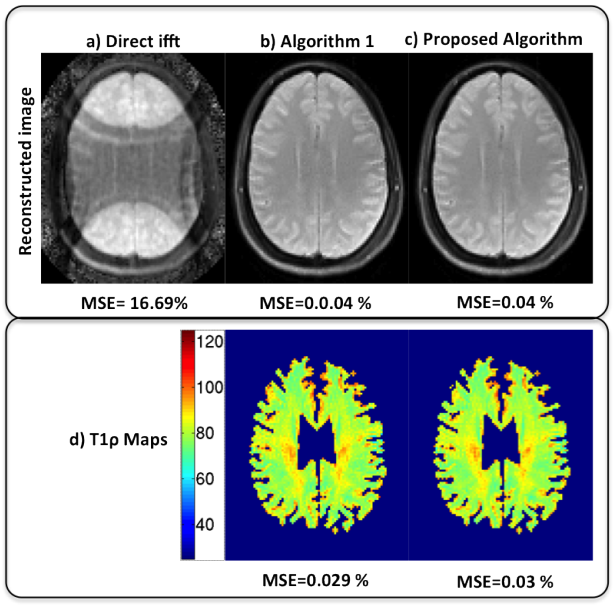

Fig. 2. Qualitative results: The first row shows reconstructed images using Direct ifft, Algorithm 1 and the Proposed algorithm in a-c. The second row shows $T_{1\rho}$ parameter maps for Algorithm 1 and proposed Algorithm. It is seen that proposed algorithm converges 10 times faster without degradation in quality as compared to the Algorithm 1.

and the CG algorithm needs many iterations to converge thus increasing the reconstruction time considerably. This behavior can be seen in Fig. 1(b). In contrast the proposed algorithm takes much lesser time as it solves the subproblems analytically. The proposed algorithm converges in about 146 secs ($\approx$ 3 min) while the Algorithm 1 takes about 2500 secs ($\approx$ 40 min) resulting in 10 fold acceleration without much degradation in image quality.

V. Conclusion

In this paper , we have demonstrated the usage of simple splitting strategies to offer significant speed up in synthesis based dictionary learning optimization problems. We have demonstrated it to show considerable improvement in cases of parallel MRI, at least a speed up factor of 10 fold.

References

[1] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic resonance in medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[2] H. Jung, J. Park, J. Yoo, and J. C. Ye, "Radial k-t FOCUSS for high-resolution cardiac cine MRI," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 68–78, 2010.

[3] S. Lingala and M. Jacob, "Blind compressive sensing dynamic MRI." *IEEE transactions on medical imaging*, vol. 32, no. 6, p. 1132, 2013.

[4] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 5, pp. 1028–1041, 2011.

[5] S. G. Lingala, Y. Mohsin, J. D. Newell, J. C. Sieren, D. Thedens, P. Kollasch, and M. Jacob, "Accelerating dynamic imaging of the lung using blind compressed sensing," *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. Suppl 1, p. W27, 2014.

[6] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented lagrangian methods," *Medical imaging, iEEE Transactions on*, vol. 30, no. 3, pp. 694–706, 2011.

[7] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in *Joint Meeting of the 6th International Symposium on Non-invasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, 2007. NFSI-ICFBI 2007.* IEEE, 2007, pp. 181–182.

[8] D. P. Bertsekas, "Multiplier methods: a survey," *Automatica*, vol. 12, no. 2, pp. 133–145, 1976.

[9] S. R. Charagundla, A. Borthakur, J. S. Leigh, and R. Reddy, "Artifacts in T1ρ-weighted imaging: correction with a self-compensating spin-locking pulse," *Journal of Magnetic Resonance*, vol. 162, no. 1, pp. 113–121, 2003.

[10] J. H. Brittain, B. S. Hu, G. A. Wright, C. H. Meyer, A. Macovski, and D. G. Nishimura, "Coronary angiography with magnetization-prepared T2 contrast," *Magnetic resonance in medicine*, vol. 33, no. 5, pp. 689–696, 1995.

[11] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin, "Adaptive reconstruction of phased array MR imagery," *Magnetic Resonance in Medicine*, vol. 43, no. 5, pp. 682–690, 2000.