A Convergence Proof of Projected Fast Iterative Soft-thresholding Algorithm for Parallel Magnetic Resonance Imaging

Xinlin Zhang, Hengfa Lu, Di Guo, Lijun Bao, Feng Huang, Xiaobo Qu*

Abstract—The boom of non-uniform sampling and compressed sensing techniques dramatically alleviates the prolonged data acquisition problem of magnetic resonance imaging. Sparse reconstruction, thanks to its fast computation and promising performance, has attracted researchers to put numerous efforts on it and has been adopted in commercial scanners. Algorithms for solving the sparse reconstruction models play an essential role in sparse reconstruction. Being a simple and efficient algorithm for sparse reconstruction, pFISTA has been successfully extended to parallel imaging, however, its convergence criterion is still an open question, confusing users on the setting of the parameter which assures the convergence of the algorithm. In this work, we prove the convergence of the parallel imaging version pFISTA. Specifically, the convergences of two well-known parallel imaging reconstruction models, SENSE and SPIRiT, solved by pFISTA are proved. Experiments on brain images demonstrate the validity of the convergence criterion. The convergence criterion proved in this work can help users quickly obtain the satisfy parameter that admits faithful results and fast convergence speeds.

Index Terms—Parallel imaging, image reconstruction, pFISTA, convergence analysis

I. INTRODUCTION

MAGNETIC resonance imaging (MRI) is a non-invasive, non-radioactive and versatile technique serving as a widely adopted and indispensable tool in medical diagnose, however, the slow imaging speed impedes its development. The advent of sparse sampling and compressed sensing (CS) theory [1]–[3] meets the eager demand of fast scan through sampling only a small amount of data points and recovering the missing data using well-developed reconstruction methods.

One principal assumption of CS lies in the transform domain sparsity of images. The sparse representation adopted to enable the image to be sparse plays a crucial role when designing reconstruction approach. Sparse representation approaches could be categorized into two main genres: orthogonal systems [3]–[5] and redundant systems [6]–[12]. The orthogonal systems favor theoretical analysis, fast algorithm design, and also time and memory-efficient computing. However, it exhibits insufficiency in sparsely representing diverse images. In contrast, designed to capture image-specific features, the redundant systems allow sparser image representation than the orthogonal systems do, thus eventually suggest better noise removal and artifacts suppression in applications.

Unlike the orthogonal system characterized by orthogonal basis, the redundant system is described by frame, mostly tight frame [13], [14], thus leading to two distinct kinds of reconstruction models, the synthetic model and the analysis model. Many researchers focus on designing efficient tight frames for MRI reconstruction and promising reconstructions are achieved [7], [9], [15], but at that time, few efforts had focused on MRI reconstruction models. Analysis model and synthesis model have different prior assumptions, analysis model is based on the assumption that the coefficients in transform domain of an MRI image are sparse, while synthesis model assumes that an MRI image can be formulated as a linear combination of sparse coefficients. Even with the same MRI data, sampling pattern, and sparse transform, the analysis model is observed to yield improved reconstruction results compared to the synthesis model [14], [16]. And it has been shown that the transition from synthetic model to analysis model comes the balanced model. In the context of MRI reconstruction, Liu et al. empirically explored the performance of the balanced model and observed that balanced model has a comparable reconstruction performance with the analysis model [16].

Analysis models, though enable better reconstructions with smaller errors, still has a compelling demand for fast algorithms that allows favorable convergence speed and fewer parameters. The alternating direction methods of multipliers (ADMM) [17], [18] can solve both synthesis and analysis models, however, it is vulnerable to parameter selections and is memory-demanding owing to the introduced dual variables. The iterative shrinkage thresholding algorithms (ISTA) [19] and its acceleration version - fast ISTA (FISTA) [20] are efficient and robust, nevertheless, they are limited to solve the synthesis model. Our group designed a projected iterative soft-thresholding algorithm (pISTA) and its acceleration version - pFISTA [14], by rewriting the analysis model into an equivalent synthesis-like one and calculating the proximal map of non-smooth terms in the objective function approximately. Being essen-
tially a variation of ISTA, pFISTA has only one adjustable parameter, and its convergence criterion has been provided in the paper [14]. Also, Liu et al. [14] theoretically proved that the pFISTA converges to a balanced model.

The pFISTA permits lower reconstruction error compared to FISTA, and faster convergence speed than the state-of-the-art methods, such as smoothing-based FISTA [14]. The pFISTA, however, is limited to tackle single-coil image reconstruction problem. Ting et al. independently proposed a computationally efficient balanced sparse reconstruction method in the context of parallel MRI under tight frame [21], named bFISTA, and applied bFISTA to two widely adopted parallel imaging models, sensitivity encoding (SENSE) and SPIRiT method [22] and iterative self-consistent parallel imaging reconstruction (SPIRiT) [23]. However, they did not provide proof of convergence of bFISTA, that is to say, in practice there is no guidance about how to choose the parameter, thus, and the algorithm users may encounter a problem of choosing a proper parameter to produce faithful results, and we will demonstrate this problem in Fig. 2 (Section III-B). Considering the importance of parallel imaging, it is necessary to give a clear mathematical proof of its convergence to assist setting a proper algorithm parameter.

In this work, we prove the sufficient conditions for the convergence of parallel imaging version pFISTA, and present convergence criteria and performances of applying pFISTA on solving two exemplars of parallel imaging reconstruction methods - SENSE and SPIRiT. In addition, we discuss the results of applying pFISTA on parallel reconstruction models under different tight frames. Last, for the unique parameter of pFISTA, we offer a recommended value to permit the fastest convergence speed as well as promising results.

The rest of the paper is organized as follows. In Section II, we introduce the notations. In Section III, we introduce some related works, firstly the pFISTA, and then SENSE and SPIRiT. In section IV, we prove that the parallel imaging version pFISTA converges under proper selection of algorithm parameter. And we offer the convergence criteria of pFISTA when applied to tackle SENSE and SPIRiT models. In Section V, we demonstrate the usefulness of the criteria we provided with multiple parallel imaging brain images. Finally, conclusions will be drawn in Section VI.

II. NOTATIONS

We first introduce notations used throughout this paper. We denote vectors by bold lowercase letters and matrices by bold uppercase letters. The transpose and conjugate transpose of a matrix are denoted by $X^T$ and $X^H$. For any vector $x$, $\|x\|_1$ and $\|x\|_2$ denote the $\ell_1$ and $\ell_2$ norm for vectors, respectively. For a matrix $X$, $\|X\|_2$ denotes the $\ell_2$ norm for matrix, which is the largest singular value of matrix $X$ and also the square root of the largest eigenvalue of the matrix $X^H X$.

Operators are denoted by calligraphic letters. Let $D_M$ denotes block diagonalization operator which places any $M$ matrices of the same size, $X_1, \cdots, X_M$, along the diagonal entris of a matrix with zeros:

$$D_M \begin{bmatrix} X_1 & 0 \\ \vdots & \ddots \\ 0 & X_M \end{bmatrix}.$$  \hspace{1cm} (1)

III. RELATED WORK

A. pFISTA for single-coil MRI reconstruction

An analysis model for single-coil sparse MRI reconstruction could be formulated as

$$\min_{x} \lambda \| \Psi x \|_1 + \frac{1}{2} \|y - UFx\|_2^2,$$  \hspace{1cm} (2)

where $x \in \mathbb{C}^N$ denotes the single-coil MR image data rearranged into a column vector, $y \in \mathbb{C}^M$ the undersampled k-space data, $U \in \mathbb{R}^{M \times N}$ the undersampling matrix, and $F \in \mathbb{C}^{N \times N}$ the discrete Fourier transform. $\Psi$ is a tight frame, and the constant $\lambda$ is the regularization parameter to balance the sparsity and data consistency.

To solve the problem (2), pFISTA rewrites the above-mentioned formula as a synthetic model as

$$\min_{\alpha \in \text{Range}(\Psi)} \lambda \| \alpha \|_1 + \frac{1}{2} \|y - UF \Psi \alpha\|_2^2,$$  \hspace{1cm} (3)

where $\Psi^*$ denotes the adjoint of $\Psi$, and specifically satisfies $\Psi \Psi^* = I$. $\alpha$ contains the coefficients of an image under the representation of a tight frame $\Psi^*$.

According to [14], the main iterations of pFISTA to solve the problem in Eq. (3) are

$$s^{(k+1)}_k = \Psi^* \gamma \lambda \left( \Psi^* s^{(k)}_k + \gamma F^H U^T \left( y - UF s^{(k)}_k \right) \right),$$  \hspace{1cm} (4)$$

$$t^{(k+1)} = \frac{1 + \sqrt{1 + 4(t^{(k)})^2}}{2},$$

$$s^{(k+1)}_k = s^{(k+1)}_k + \frac{t^{(k+1)} - 1}{t^{(k+1)}} (s^{(k+1)}_k - s^{(k)}_k),$$

where $T_{\gamma \lambda} (\cdot)$ is a point-wise soft-thresholding function defined as $T_{\gamma \lambda} (\alpha) = \max \{ |\alpha| - \gamma \lambda, 0 \} \cdot \alpha / |\alpha|$.

According to the Theorem 2 in the pFISTA paper [14], when the step size $0 < \gamma \leq 1$, the algorithm will converge. In addition, the larger $\gamma$ is, the faster pFISTA converges. Therefore, $\gamma = 1$ is recommended in pFISTA to produce promising reconstruction with the fastest convergence speed.

B. pFISTA for multi-coil MRI reconstruction

According to [21], we can formula analysis models for the parallel MRI reconstruction problem into a unified form as

$$\text{min}_{d} \lambda \| \Psi d \|_1 + \| y - Ad \|_2^2,$$  \hspace{1cm} (5)

where $d$ represents the desired image to be recovered, $y = [y_1; y_2; \cdots; y_J] \in \mathbb{C}^{MJ}$ the undersampled multi-coil k-space data rearranged into a column vector, and $y_J \in \mathbb{C}^M$ ($j = 1, 2, \cdots, J$) is the undersampled k-space data vector of $j$th coil, and $A$ the undersampling matrix in parallel MRI containing the Fourier transform with multi-coil modulation and undersampling.

For parallel MRI reconstruction methods based on different signal properties, the explicit expressions of Eq. (5) would
vary. Two reconstruction algorithms based on SENSE and SPIRiT are discussed in [21], however, the convergence of these two algorithms has not been proven. Thus, in this work, we first prove the convergence of pFISTA of solving the general parallel MRI reconstruction model, and then offer two concrete examples of multi-coils MRI analysis model, SENSE and SPIRiT, with convergence analysis. We first introduce how to tackle SENSE and SPIRiT using pFISTA.

Using pFISTA, we can get the solution of Eq. (7) by iteratively solve the following problems:

\[ x_c^{(k+1)} = \Psi^* T_{\gamma} \left( \Psi \left( x_c^{(k)} + \gamma R^T \tilde{F} H U_T \left( y - \tilde{U} F C R x_c^{(k)} \right) \right) \right), \]

\[ t^{(k+1)} = \frac{1 + \sqrt{1 + 4 (t(k))^2}}{2}, \]

\[ x_c^{(k+1)} = x_c^{(k+1)} + \frac{t_k - 1}{t_{k+1}} \left( x_c^{(k+1)} - x_c^{(k)} \right). \]

(8)

For simplicity, we call the pFISTA adopted to solve SENSE model pFISTA-SENSE.

2) pFISTA-SPIRiT: The SPIRiT [23] primarily bases on the assumption that each k-space data point of a given coil is a linear combination of the multi-coil data of its neighboring k-space points, and the weights of linear combination are estimate from auto-calibration signal (ACS) (Fig. 1 (b)). Let \( x = \left[ x_1; x_2; \ldots; x_J \right] \in \mathbb{C}^{N J} \) denote the multi-coil image data rearranged into a column vector, where \( x_j \in \mathbb{C}^{N J}, (j = 1, 2, \ldots, J) \) is the \( j \)-th coil image vector, then the calibration consistency in SPIRiT can be formulated as:

\[ \min \| \Psi x \|_1 + \frac{1}{2} \left\| \gamma \tilde{U} F x - y \right\|_2^2, \]

(10)

where the matrix \( G \in \mathbb{C}^{N J \times N J} \) is shown as below

\[ G = \begin{bmatrix}
G_{1,1} & G_{1,2} & \cdots & G_{1,J} \\
G_{2,1} & G_{2,2} & \cdots & G_{2,J} \\
\vdots & \vdots & \ddots & \vdots \\
G_{J,1} & G_{J,2} & \cdots & G_{J,J}
\end{bmatrix}. \]

Notice that strictly speaking, the \( \Psi \) in Eq. (10) should be written in the form of \( \Psi = D_J (\Psi, \ldots, \Psi) \) indicating that the \( \Psi \) is applied to each coil image. Here \( \Psi \) is still a tight frame which satisfies \( \Psi^* \Psi = I \), thus, we use only \( \Psi \) in the rest of the paper for simplicity.

In order to express Eq. (10) in the unified form shown in Eq. (5), we rewrite the Eq. (10) as:

\[ \min \lambda \| \Psi x \|_1 + \frac{1}{2} \left\| \frac{1}{\sqrt{N}} y - \sqrt{N} \left( G - I \right) \right\|_2^2 \]

(12)

Here, the undersampling matrix \( A \) in Eq. (5) has its explicit expression as \( A = \tilde{U} \tilde{F} \).

Using pFISTA, we can get the solution of Eq. (12) by iteratively solve the following problems:

\[ x^{(k+1)} = \Psi^* T_{\gamma, \lambda} \left( \Psi \left( x^{(k)} + \gamma RH^T \left( \tilde{U}^H \left( y - \tilde{U} F x^{(k)} \right) \right) \right) \right), \]

\[ \lambda (G - I)^{HT} (G - I) \tilde{F} \tilde{F} \]

\[ t^{(k+1)} = \frac{1 + \sqrt{1 + 4 (t(k))^2}}{2}, \]

\[ x^{(k+1)} = x^{(k+1)} + \frac{t_k - 1}{t_{k+1}} \left( x^{(k+1)} - x^{(k)} \right). \]

(13)
It has been shown that pFISTA-parallel embraces faster reconstruction speed than nonlinear conjugate gradient algorithm, rendering it having great potential in the clinic [21]. However, the convergence analysis of pFISTA-parallel is still an open problem. In other words, we don’t know explicitly in advance which γ can guarantee the algorithm to converge. Liu et al. [14] have proved that under the condition γ = 1, pFISTA for single-coil MRI reconstruction is guaranteed to converge. But if the same setting, γ = 1, is used in pFISTA-SENSE and pFISTA-SPIRiT, the algorithms may not converge (Fig. 2). This is because the sensitivity map or convolution kernel would affect the convergence property of pFISTA-parallel. We observed in experiments that a relatively large γ will lead to the divergence of pFISTA-parallel while a far smaller one results in the slow convergence of the algorithm (Fig. 2). And the range of γ allowing the algorithm to converge varies under different tested data. Therefore, we aim to offer a explicit rule about how to choose a proper γ of pFISTA-parallel to hold a fast convergence speed and promising results.

IV. CONVERGENCE ANALYSIS

In this section, we prove the convergence of pFISTA-parallel.

We present the analysis model of the parallel MRI reconstruction in a unified formula shown in Eq. (5) in which the undersampling matrix A has its explicit form A = \( \tilde{U}\tilde{F}\mathcal{CR} \) if the model is SENSE-based, and A = \( [\tilde{U} - \sqrt{\lambda_1}(G - I)]^T\tilde{F} \) if the model is SPIRiT-based. According to [14], [20], let \{\( \mathbf{d}^{(k)} \)\} be generated by pFISTA-parallel, and if the step size satisfies

\[
\gamma \leq \frac{1}{L(\gamma)},
\]

and \( \Psi \) is a tight frame, the sequence \{\( \mathbf{a}^{(k)} \)\} = \{\( \Psi\mathbf{d}^{(k)} \)\} converges to a solution of

\[
\min_{\mathbf{a}} \lambda ||\mathbf{a}|| + \frac{1}{2} ||\mathbf{y} - \mathbf{A}\Psi^*\mathbf{a}||^2 + \frac{1}{2\gamma} ||(\mathbf{I} - \Psi\Psi^*)\mathbf{a}||^2,
\]

with the speed

\[
F\left(\mathbf{a}^{(k)}\right) - F\left(\hat{\mathbf{a}}\right) \leq \frac{L(\gamma)}{\gamma(k+1)^2} ||\mathbf{a}^{(k)} - \hat{\mathbf{a}}||^2,
\]

where \( \hat{\mathbf{a}} \) is a solution of (15) and \( F(\cdot) \) is the objective function in (15) and \( L \) is the Lipschitz constant for the gradient term. Let \( \mathbf{B} = \mathbf{\Psi A}^H \mathbf{A}^{\Psi^*} - 1/\gamma \mathbf{\Psi} \mathbf{\Psi}^* \), we have

\[
L(\gamma) = \max_{i} \left( \left| e_i(\mathbf{B}) + \frac{1}{\gamma} \right| \right) = \max_{i} \left\{ \frac{1}{\gamma}, \left| e_i(\mathbf{A}^H \mathbf{A}) \right| \right\},
\]

where \( e_i(\cdot) \) denotes the \( i \)th eigenvalue of matrix. Now we just have to analyze the largest eigenvalue of matrix \( \mathbf{A}^H \mathbf{A} \) in different reconstruction problems. In the following, we will explicitly discuss the convergence of pFISTA-SENSE and pFISTA-SPIRiT.

A. Convergence of pFISTA-SENSE

In this section, we provide the sufficient conditions for the convergence of pFISTA-SENSE in the form of a theorem.

Theorem 1. Let \{\( \mathbf{x}^{(k)} \)\} be generated by pFISTA-SENSE, and if the step size satisfies

\[
\gamma \leq \frac{1}{c}, \quad c = \sum_{j=1}^{J} ||C_j||^2_2,
\]

and \( \Psi \) is a tight frame, the sequence \{\( \mathbf{\alpha}^{(k)} \)\} = \{\( \Psi\mathbf{d}^{(k)} \)\} converges to a solution of

\[
\min_{\mathbf{\alpha}} \lambda ||\mathbf{\alpha}|| + \frac{1}{2} ||\mathbf{y} - \tilde{\mathbf{U}}\tilde{\mathbf{FCR}}\mathbf{\Psi}^*\mathbf{\alpha}||^2 + \frac{1}{2\gamma} ||(\mathbf{I} - \mathbf{\Psi}^*\mathbf{\Psi})\mathbf{\alpha}||^2.
\]

Proof. In pFISTA-SENSE, we have \( \mathbf{A} = \tilde{\mathbf{U}}\tilde{\mathbf{FCR}} \), thus,

\[
\mathbf{A}^H \mathbf{A} = \mathbf{R}^T \mathbf{C}^H \tilde{\mathbf{F}}^H \tilde{\mathbf{U}}^T \tilde{\mathbf{FCR}}.
\]

Notice that \( \tilde{\mathbf{U}} \) and \( \tilde{\mathbf{F}} \) are block diagonal matrix, Eq. (20) can be rewritten as:

\[
\mathbf{A}^H \mathbf{A} = \mathbf{R}^T \mathbf{C}^H \tilde{\mathbf{F}}^H \tilde{\mathbf{U}}^T \tilde{\mathbf{FCR}} = \left[ \begin{array}{c} \mathbf{I} \\ \cdots \\ \mathbf{I} \end{array} \right] \left[ \begin{array}{c} \mathbf{C}_1^J \mathbf{F}^H \mathbf{U}^T \mathbf{FCR}_1 \\ \vdots \\ \mathbf{C}_J^J \mathbf{F}^H \mathbf{U}^T \mathbf{FCR}_J \end{array} \right] = \sum_{j=1}^J \mathbf{C}_j^J \mathbf{F}^H \mathbf{U}^T \mathbf{FCR}_j.
\]

Let \( \mathbf{Q} = \mathbf{F}^H \mathbf{U}^T \mathbf{UF} \), \( \sum_{j=1}^J \mathbf{C}_j^J \mathbf{Q} \mathbf{C}_j \) is a Hermitian matrix. For a Hermitian matrix, the largest eigenvalue is equal to the \( \ell_2 \) norm. In addition, notice that matrix \( \ell_2 \) norm satisfies triangle inequality and consistency property [24], we can find the upper bound of the largest eigenvalue of the matrix \( \sum_{j=1}^J \mathbf{C}_j^J \mathbf{Q} \mathbf{C}_j \):

\[
\max_{i} \left( \sum_{j=1}^J \mathbf{C}_j^J \mathbf{Q} \mathbf{C}_j \right) \leq \sum_{j=1}^J ||\mathbf{C}_j^J||_2 ||\mathbf{Q}||_2 ||\mathbf{C}_j||_2 \leq \sum_{j=1}^J ||\mathbf{C}_j^J||_2 ||\mathbf{Q}||_2 ||\mathbf{C}_j||_2.
\]

Here, the matrix \( \mathbf{F} \) is a unitary matrix, according to the unitary invariant of \( \ell_2 \) norm, we have

\[
||\mathbf{Q}||_2 = ||\mathbf{F}^H \mathbf{U}^T \mathbf{UF}||_2 = ||\mathbf{U}^T \mathbf{U}||_2,
\]

Fig. 2. Empirical convergence of pFISTA-SENSE (a) and pFISTA-SPIRiT (b) with different γ. The reconstruction experiments were carried out on a 32-coil brain image with 34% data acquired using a 1D Cartesian sampling pattern. The used data and the sampling pattern are presented in Fig. 3.
And $U^TU$ is a diagonal matrix with the diagonal elements 0 or 1, indicating that

$$\|Q\|_2 = 1.$$ (24)

With Eq. (24), we can further simplify Eq. (22)

$$\max_i e_i \left( \sum_{j=1}^J C_j^H Q C_j \right) \leq \sum_{j=1}^J \|C_j^H\|_2 \|Q\|_2 \|C_j\|_2
= \sum_{j=1}^J \|C_j^H\|_2 \|C_j\|_2
= \sum_{j=1}^J \|C_j\|_2^2.$$ (25)

Let $c = \sum_{j=1}^J \|C_j\|_2^2$, we have

$$L(\gamma) = \max_i \left\{ \frac{1}{\gamma} e_i \left( \sum_{j=1}^J C_j^H Q C_j \right) \right\} = \frac{1}{c}, \quad 0 < \gamma \leq \frac{1}{c},$$ (26)

$$L(\gamma) = \max_i \left\{ \frac{1}{\gamma} e_i \left( \sum_{j=1}^J C_j^H Q C_j \right) \right\} = c, \quad \gamma > \frac{1}{c}.$$ (27)

The Eq. (26) means that, when $0 < \gamma \leq 1/c$, one has $L(\gamma) = 1/\gamma_1$, which satisfies the convergence condition of pFISTA in Eq. (14); whereas when $\gamma > 1/c$, then $L(\gamma) = c > 1/\gamma$, which does not satisfy the convergence condition of pFISTA. In summary, when $0 < \gamma \leq 1/c$, the pFISTA-SENSE is guaranteed to converge. \hfill \Box

### B. Convergence of pFISTA-SPRiT

In this section, we provide the sufficient conditions for the convergence of pFISTA-SPRiT in the form of a theorem.

**Theorem 2.** Let $\{x^{(k)}\}$ be generated by pFISTA-SPRiT, and if the step size satisfies

$$c = 1 + \lambda_1 \left( \sum_{m=1}^J \sum_{n=1}^J \|G_{m,n}\|_2 + 1 \right)^2,$$ (27)

and $\Psi$ is a tight frame, the sequence $\{\alpha^{(k)}\} = \{\Psi d^{(k)}\}$ converges to a solution of

$$\min_\alpha \lambda_1 \|\alpha\|_1 + \frac{1}{2\gamma} \left\| (I - \Psi \Psi^*) \alpha \right\|_2^2
+ \frac{1}{2} \left\| Y - \left[ \begin{array}{c} \tilde{U} \\ -\sqrt{\lambda_1} (G - I) \end{array} \right] \tilde{F} \Psi^* \alpha \right\|_2^2.$$ (28)

**Proof.** In pFISTA-SPRiT, we have

$$A^H A = \tilde{F}^H \left[ \begin{array}{cc} U^T & -\sqrt{\lambda_1} (G - I)^H \end{array} \right] \left[ \begin{array}{c} \tilde{U} \\ -\sqrt{\lambda_1} (G - I) \end{array} \right] \tilde{F},$$ (29)

Since the $\tilde{F}$ is a unitary matrix, then the matrix $\left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)$ and the matrix $A^H A$ is unitary similar. And $\left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)$ is a Hermitian matrix, thus, the eigenvalue of $A^H A$ and $\left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)$ are the same. Now, we analyze the eigenvalue of matrix $\left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)$.

Indeed, once the kernel has been estimated using ACS, the matrix $G$ is determined, so that the largest eigenvalue can be calculated. However, the matrices $U$ and $G$ are too large to calculate the eigenvalue conveniently, for example, as for a 4-coil $256 \times 256$ image, the size of corresponding $U$ and $G$ reaches the size of $262144 \times 262144$. Therefore, we relax the bound so as to could efficiently calculate it.

Notice that matrix $\left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)$ is Hermitian matrix and with the linearity and triangle inequality of matrix norm [24], so that

$$e_i \left( \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right)
= \left\| \tilde{U}^T \tilde{U} + \lambda_1 (G - I)^H (G - I) \right\|_2
\leq \left\| \tilde{U}^T \tilde{U} \right\|_2 + \lambda_1 \left\| (G - I)^H (G - I) \right\|_2$$ (30)

$$\leq \left\| \tilde{U}^T \tilde{U} \right\|_2 + \lambda_1 \left\| (G - I)^2 \right\|_2.
\leq \left\| \tilde{U}^T \tilde{U} \right\|_2 + \lambda_1 \left\| (G - I) \right\|_2^2 + \lambda_1 \left\| (G - I) \right\|_2^2.$$ (31)

As mentioned, the matrix $G$ is a block circulant matrix as shown below

$$G = \begin{bmatrix} G_{1,1} & G_{1,2} & \cdots & G_{1,J} \\ G_{2,1} & G_{2,2} & \cdots & G_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ G_{J,1} & G_{J,2} & \cdots & G_{J,J} \end{bmatrix},$$

where $G_{m,n} \ (m, n = 1, 2, \cdots, J)$ is circulant matrix. Thus, now, the eigenvalue can easily be obtained by linear combination weights.
Therefore,

\[
e_i \left( \hat{U}^T \hat{U} + \lambda_1 (G - I)^H (G - I) \right) \\
\leq \left\| \hat{U}^T \hat{U} \right\|_2 + \lambda_1 (\|G\|_2 + \|I\|_2)^2 \\
= \left( \sum_{j=1}^{J} \sum_{n=1}^{J} \|G_{m,n}\|_2 + \|I\|_2 \right)^2
\]

Let \( c = 1 + \lambda_1 \left( \sum_{m=1}^{M} \sum_{n=1}^{N} \|G_{m,n}\|_2 + 1 \right)^2 \) we have

\[
L(\gamma) = \max_i \left\{ \frac{1}{\gamma}, \left| e_i \left( \hat{U}^T \hat{U} + \lambda_1 (G - I)^H (G - I) \right) \right| \right\} \\
= \frac{1}{\gamma}, \quad 0 < \gamma \leq \frac{1}{c}, \\
L(\gamma) = \max_i \left\{ \frac{1}{\gamma}, \left| e_i \left( \hat{U}^T \hat{U} + \lambda_1 (G - I)^H (G - I) \right) \right| \right\} \\
= c, \quad \gamma > \frac{1}{c}.
\]

The Eq. (32) means that, when \( 0 < \gamma \leq 1/c \), one has \( L(\gamma) = 1/\gamma \), which satisfies the convergence condition of pFISTA; whereas when \( \gamma > 1/c \), then \( L(\gamma) = c > 1/\gamma \), which does not satisfy the convergence condition of pFISTA. In summary, when \( 0 < \gamma \leq 1/c \), pFISTA-SPiRiT is guaranteed to converge.

V. EXPERIMENTAL RESULTS

In this section, we first conducted experiments on multi-coils MRI brain images to assess the validity of the convergence criteria we derived. And then we compared the reconstructions of pFISTA-parallel and the widely adopted algorithm - ADMM [14]. The ADMM softwares to solve SENSE and SPIRiT analysis reconstruction models were implemented by ourselves. Last, we discussed the convergence and results under other tight frames with different \( \gamma \).

Three datasets are used in experiments. The first brain dataset shown in Fig. 3 (a) was acquired from a healthy volunteer on a 1.5T Philips MRI scanner equipped with an 8-coil head coil using the 2D T1-weighted fast-field-echo sequence (matrix size = 256 x 256, TR/TE = 1700 ms/390
Fig. 5. Reconstructions of three different brain images by pFISTA-SPIRiT with different step size $\gamma$. The 8, 12 and 32-coil data shown in Fig. 3 were used in (a), (b) and (c), respectively. (d) is the reconstructions of pFISTA-SPIRiT at different iteration in 8-coil data. All experiments used the same sampling pattern depicted in Fig. 3.

Relative $\ell_2$ norm error (RLNE) is adopted as objective criteria to quantify the reconstruction performance. The RLNE is defined as

$$RLNE = \frac{\|x_{ref} - x_{rec}\|_2}{\|x_{ref}\|_2},$$

where $x_{ref}$ denotes the vectorized reference image that is a square root of sum of squares (SSOS) of the fully sampled image and $x_{rec}$ the vectorized reconstructed image that is the SSOS image of pFISTA-SPIRiT reconstructed image and modular image of pFISTA-SENSE reconstructed image. We should point out that a lower RLNE, a higher consistency between the reference image and the reconstructed image.

For SENSE, the fully sampled $256 \times 64$ area of the k-space center is used to calculate sensitivity map, and for SPIRiT, the fully sampled $256 \times 22$ area is used to estimate linear combination weights. The shift-invariant discrete wavelets transform (SIDWT) [7], [25], [26], if not mentioned otherwise, is adopted as the tight frame in experiments. In all experiments involving SIDWT, Daubechies wavelets with 4 decomposition levels are utilized. For pFISTA-SENSE, $\lambda = 10^{-3}$ is set and for pFISTA-SPIRiT, we set $\lambda = 10^{-3}$ and $\lambda_1 = 1$, and $5 \times 5$ SPIRiT kernel is used. All computation procedures run on a CentOS 7 computation server with two Intel Xeon CPUs of 3.5 GHz and 112 GB RAM.

A. Main Results

As mentioned above, once if the parameter meets the condition $0 < \gamma \leq 1/c$, both the pFISTA-SENSE and pFISTA-SPIRiT converge. Thus, here we perform reconstructions by
pFISTA-SENSE and pFISTA-SPIRiT with various $\gamma$ in the recommended range, respectively, to verify if the recommended $\gamma$ could enable the convergence of the algorithm.

As shown in Fig. 4, for three tested brain images, pFISTA-SENSE converges when using the $\gamma$ ranged from $0.01/c$ to $1/c$. And all $\gamma$ used eventually allow the RLNEs decreasing to a comparatively low level. More importantly, the larger the $\gamma$, the faster the algorithm converges, this observation is consistent with the Eq. (16). The intermediate reconstructed images manifest the convergence speeds of pFISTA-SENSE with various $\gamma$. The undersampling artifacts were quickly removed within 100 iterations when with parameter $\gamma = 1/c$ and the algorithm produced a nice image (Fig. 4 (d)). As $\gamma$ decreased, the algorithm took more time to converge to the final stage yielding satisfying results, for instance, when $\gamma = 0.01/c$, the program cost about 800 iterations to eventually eliminate the undersampling artifacts (Fig. 4 (d)). All recommended $\gamma$ enable promising results but with different convergence rate. Importantly, the algorithm reaches the fastest convergence speed when $\gamma = 1/c$, thus we recommend $\gamma = 1/c$ when carrying out pFISTA-SENSE experiments. It is also worth to point out that the number of coils of parallel imaging makes no influence on the convergence of pFISTA-SENSE with offered range $0 < \gamma \leq 1/c$. In a word, the convergence criteria we provided can escort pFISTA-SENSE to achieve satisfying results of different coils parallel imaging experiments.

In addition, we observe similar phenomenon on pFISTA-SPIRiT experiments (Fig. 5). With $0 < \gamma \leq 1/c$, pFISTA-SPIRiT empirically convergences to a level of promising low RLNE. The larger the $\gamma$, the faster the algorithm converges, and the fastest convergence speed is achieved when $\gamma = 1/c$, thus we also recommend $\gamma = 1/c$ for pFISTA-SPIRiT experiments. The intermediate results of pFISTA-SPIRiT with monotonically decreasing $\gamma$ also reveal the increasing convergence rate as $\gamma$ rises (Fig. 5 (d)). The pFISTA-SPIRiT could be applied in multi-coils imaging experiments with guaranteed convergence if $0 < \gamma \leq 1/c$.

We observed in our experiments that pFISTA-parallel allows as fast convergence speed as the fastest ADMM offers and also close reconstruction error of ADMM (Fig. 6). This indicates that the relaxation of the convergence criterion of pFISTA-SPIRiT is reasonable as it enables comparable results as the ADMM provides with the best hand-crafted optimal penalty parameter $\beta$. As shown in Fig. 6, the reconstruction of the 8-coil T1-weighted brain image, the convergence of ADMM is sensitive to the parameter $\beta$ selection, relatively larger or smaller $\beta$ would result in noticeable discrepancy (Figs. 6 (a-b)). And $\beta = 0.01$ yields the fastest convergence speed of ADMM. Importantly, pFISTA-parallel with the recommended parameter $\gamma = 1/c$ very close reconstruction error as well as reconstructed images (Figs. 6 (c-j)).

### B. Discussion on other tight frames

Tight frame is crucial for sparse MRI reconstruction. In this section, we conduct experiments using pFISTA-SENSE and pFISTA-SPIRiT with four other tight frames, contourlet [15], [27], shearlet [28], patch based directional wavelets (PBDW) [9], and PBDW in SIDWT domain (PBDWS) [29]. By exploring the image geometry, contourlet provides a sparse expansion for images that have smooth contours [27], and therefore provides improvements compared to the contourlet transform [28]. PBDW trains the geometric directions on the pixels of image patches and provides an adaptively sparse representation for image [9]. PBDWS extend the PBDW into SIDWT domain to enhance the ability of sparsifying [29]. Here, the filters used in contourlet are ladder structure filters and the decomposition levels are $[5,4,4,3]$, the filters used in shearlet are Meyer filters [30] and the decomposition level used in shearlet is 4, and the filters used in PBDW and PBDWS are the Haar wavelets and the decomposition level is 3.

The results shown in Fig. 7 reveal that the RLNEs of pFISTA-SENSE with PBDW and PBDWS tight frames are smaller than that under SIDWT and shearlet tight frames. The result is reasonable since PBDW and PBDWS are tight frames trained using pre-reconstructed image, admitting better sparse representation of the image. Notably, parameter $\gamma = 1/c$ enables the fastest convergence speed though the tight frame varies. Those results indicates that the tight frames used will not affect the convergence of both pFISTA-SENSE and pFISTA-SPIRiT.
VI. CONCLUSION

As a simple and fast algorithm to solve sparse reconstruction model, pFISTA has been successfully extended to solve parallel imaging problems, but its convergence criterion needs to be proved to help quickly and conveniently choose a satisfying parameter. In this work, we prove the sufficient conditions for the convergence of parallel imaging version pFISTA for solving sparse reconstruction models. More explicitly, we offer a bound served as guidance about how to choose the pFISTA parameter for solving both SENSE and SPIRiT. Experimental results evince the validity and effectiveness of the convergence criterion. This work is useful to help user quickly choose a proper parameter to obtain faithful results and fast convergence speed, and to promote the application of sparse reconstruction.

ACKNOWLEDGMENTS

Authors appreciate the help of Yunsong Liu for revising the manuscript. Xiaobo Qu is grateful to Prof. Chun Yuan for hosting his visit at University of Washington.

REFERENCES

[1] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[2] T. T. E. J. Cand'es, J. Romberg, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
[3] J. M. P. M. Lustig, D. Donoho, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
[4] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, “An efficient algorithm for compressed MR imaging using total variation and wavelets,” in 2008 *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Conference Proceedings, pp. 1–8.
[5] J. Huang, S. Zhang, and D. Metaxas, “Efficient MR image reconstruction for compressed MR imaging,” *Medical Image Analysis*, vol. 15, no. 5, pp. 670–679, 2011.
[6] M. Guerquin-Kern, M. Habelin, K. P. Puessmann, and M. Unser, “A fast wavelet-based reconstruction method for magnetic resonance imaging.” *IEEE Transactions on Medical Imaging*, vol. 30, no. 9, pp. 1649–1660, 2011.
[7] C. A. Baker, K. King, D. Liang, and L. Ying, “Translation-invariant dictionaries for compressed sensing in magnetic resonance imaging,” in 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, Conference Proceedings, pp. 1602–1605.
[8] S. Ravishankar and Y. Bresler, “MR image reconstruction from highly undersampled k-space data by dictionary learning,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1028–1041, 2011.
[9] X. Qu, D. Guo, B. Ning, Y. Hou, Y. Lin, S. Cai, and Z. Chen, “Undersampled MRI reconstruction with patch-based directional wavelets,” *Magnetic Resonance Imaging*, vol. 30, no. 7, pp. 964–977, 2012.
[10] Z. Zhan, J. Cai, D. Guo, Y. Liu, Z. Chen, and X. Qu, “Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1850–1861, 2016.
[11] X. Xu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, “Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator,” *Medical Image Analysis*, vol. 18, no. 6, pp. 843–856, 2014.
[12] Z. Lai, X. Yu, Y. Liu, D. Gao, J. Ye, Z. Zhan, and Z. Chen, “Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform,” *Medical Image Analysis*, vol. 27, pp. 93–104, 2016.
[13] M. Vetterli, J. Kovacevi, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge University Press, 2014.
[14] Y. Liu, Z. Zhan, J.-F. Cai, D. Guo, Z. Chen, and X. Qu, “Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging.” *IEEE Transactions on Medical Imaging*, vol. 35, no. 9, pp. 2130–2140, 2016.
[15] X. Qu, W. Zhang, D. Guo, C. Cai, S. Cai, and Z. Chen, “Iterative thresholding compressed sensing MRI based on contourlet transform, inverse problems in science and engineering.” *Inverse Problems in Science and Engineering*, vol. 18, no. 6, pp. 737–758, 2010.
[16] Y. Liu, J.-F. Cai, Z. Zhan, D. Guo, J. Ye, Z. Chen, and X. Qu, “Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging.” *PloS One*, vol. 10, no. 4, p. e0119584, 2015.
[17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
[18] S. Ramani and J. A. Fessler, “Parallel MR image reconstruction using augmented Lagrangian methods,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 3, pp. 694–706, 2010.
[19] I. Daubechies, M. Defrise, and C. De Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,” *Commp...

Fig. 7. Empirical convergence using four other tight frames. (a–d) are the RLNEs of pFISTA-SENSE with different γ using contourlet, shearlet, PBDW, and PBDWS, respectively; and (e–h) are the RLNEs of pFISTA-SPIRiT with different γ using contourlet, shearlet, PBDW, and PBDWS, respectively. Note: 8-coil image in Fig. 3 (a) and the Cartesian sampling pattern with sampling rate of 0.25 are adopted in all experiments.
[20] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
[21] S. T. Ting, R. Ahmad, N. Jin, J. Craft, J. Serafim da Silveira, H. Xue, and O. P. Simonetti, “Fast implementation for compressive recovery of highly accelerated cardiac cine MRI using the balanced sparse model,” *Magnetic Resonance in Medicine*, vol. 77, no. 4, pp. 1505–1515, 2017.
[22] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: Sensitivity encoding for fast MRI,” *Magnetic Resonance in Medicine*, vol. 42, pp. 952–962, 1999.
[23] M. Lustig and J. M. Pauly, “SPIRIT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space,” *Magnetic Resonance in Medicine*, vol. 64, no. 2, pp. 457–71, 2010.
[24] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Siam, 2000, vol. 71.
[25] R. R. Coifman and D. L. Donoho, *Translation-invariant de-noising*. Springer, 1995, pp. 125–150.
[26] M. H. Kayvanrad, A. J. McLeod, J. S. Baxter, C. A. McKenzie, and T. M. Peters, “Stationary wavelet transform for under-sampled MRI reconstruction,” *Magnetic Resonance Imaging*, vol. 32, no. 10, pp. 1353–1364, 2014.
[27] M. N. Do and M. Vetterli, “The contourlet transform: an efficient directional multiresolution image representation,” *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.
[28] G. Easley, D. Labate, and W.-Q. Lim, “Sparse directional image representations using the discrete shearlet transform,” *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 25–46, 2008.
[29] B. Ning, X. Qu, D. Guo, C. Hu, and Z. Chen, “Magnetic resonance image reconstruction using trained geometric directions in 2D redundant wavelets domain and non-convex optimization,” *Magnetic Resonance Imaging*, vol. 31, no. 9, pp. 1611–1622, 2013.
[30] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. American Mathematical Soc., 2001, vol. 22.