

Motion Adaptive Patch-Based Low-Rank Approach for Compressed Sensing Cardiac Cine MRI

Huisu Yoon, Kyung Sang Kim, *Student Member, IEEE*, Daniel Kim, Yoram Bresler, *Fellow, IEEE*, and Jong Chul Ye*, *Senior Member, IEEE*

Abstract—One of the technical challenges in cine magnetic resonance imaging (MRI) is to reduce the acquisition time to enable the high spatio-temporal resolution imaging of a cardiac volume within a short scan time. Recently, compressed sensing approaches have been investigated extensively for highly accelerated cine MRI by exploiting transform domain sparsity using linear transforms such as wavelets, and Fourier. However, in cardiac cine imaging, the cardiac volume changes significantly between frames, and there often exist abrupt pixel value changes along time. In order to effectively sparsify such temporal variations, it is necessary to exploit temporal redundancy along motion trajectories. This paper introduces a novel patch-based reconstruction method to exploit geometric similarities in the spatio-temporal domain. In particular, we use a low rank constraint for similar patches along motion, based on the observation that rank structures are relatively less sensitive to global intensity changes, but make it easier to capture moving edges. A Nash equilibrium formulation with relaxation is employed to guarantee convergence. Experimental results show that the proposed algorithm clearly reconstructs important anatomical structures in cardiac cine image and provides improved image quality compared to existing state-of-the-art methods such as k-t FOCUSS, k-t SLR, and MASTeR.

Index Terms—Compressed sensing dynamic magnetic resonance imaging (MRI), generalized Huber approximation, multiple object functions, Nash equilibrium, overlapped patches, patch-based low-rank, proximal mapping, rank penalty, relaxation.

I. INTRODUCTION

DATA collected by a magnetic resonance imaging (MRI) scanner is a spatial Fourier transform of an object image. Hence, in order to obtain an image without aliasing, k -space samples need to satisfy the Nyquist sampling criterion. However, MRI is an inherently slow imaging modality since it is

designed to acquire 2-D (or 3-D) k -space data through 1-D free induction decay or echo signals. This often limits the use of MRI, especially for high resolution dynamic cardiac cine imaging. While breath-holding can be used to minimize the respiratory motion, the number of acquired image slices during breath-holding is limited due to the short scan window. Furthermore, breath-holding intervals vary between individuals, which introduces difficulties in determining clinical protocols.

To address this issue, many researchers have recently applied compressed sensing (CS) approaches to dynamic imaging applications [1]–[8]. Compressed sensing (CS) tells us that accurate reconstruction is possible as long as the nonzero support is sparse and the sampling basis is incoherent [9], [10]. In k-t SPARSE [1], two types of transforms are used for time-varying cardiac images—a wavelet transform along the spatial dimensions and a Fourier transform along the temporal dimension. k-t FOCUSS [3]–[6], [8] employs a FOCUSS (FOCAL Under-determined System Solver) [11] algorithm to impose a sparsity constraint in the temporal transform domain. A reweighted norm approach for sparse recovery in k-t FOCUSS generalizes k-t BLAST/SENSE [12] so that it indeed outperforms classical k-t BLAST/SENSE. Related methods have been proposed by many authors using group sparsity [13], and data-driven principal component analysis (PCA) or Karhunen–Loève transform (KLT) [4], [6].

In these approaches, a temporal transform is usually applied separably along the temporal dimension at each pixel location so that if the temporal variations are mainly due to the intensity variation rather than object motion (as often observed in perfusion imaging [14] or T_2 decay [6]), the separable temporal transform can still accurately capture the dynamic variations. However, in cardiac cine imaging, the pixel intensity variations are mainly due to the motion of heart, so such a separable linear transform along the temporal dimension is not effective in capturing temporal redundancies. For example, Fig. 1(a) illustrates the motion trajectories of two distinct pixels in a moving heart. Specifically, two different pixel positions A and B are designated such that A is in relatively stationary background and B is in the moving region of the heart. If we just follow the pixel intensity variations at A and B along time, we obtain two independent temporal traces as shown in Fig. 1(b). However, if we plot pixel value changes along the heart wall motion as shown in Fig. 1(c), the variations of the pixel values are quite similar except a constant offset. Hence, if we want to represent the temporal intensity pixel variations using a linear combination of basis functions, the time traces in Fig. 1(b) require two basis functions, whereas the temporal

Manuscript received April 07, 2014; revised May 27, 2014; accepted June 04, 2014. Date of publication June 12, 2014; date of current version October 28, 2014. This work was supported by Korea Science and Engineering Foundation under Grant NRF-2013M3A9B2076548 and NRF-2009-0081089. The work of Y. Bresler was supported in part by the National Science Foundation (NSF) under Grant CCF-1018660 and Grant CCF-1320953. *Asterisk indicates corresponding author.*

H. Yoon and K. S. Kim are with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 291, Korea (e-mail: tea@kaist.ac.kr; kssigari@kaist.ac.kr).

D. Kim is with the Department of Radiology, University of Utah, Salt Lake City, UT 84108 USA (e-mail: daniel.kim@hsc.utah.edu).

Y. Bresler is with the Coordinate Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: ybresler@uiuc.edu).

*J. C. Ye is with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 291, Korea (e-mail: jong.ye@kaist.ac.kr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2014.2330426

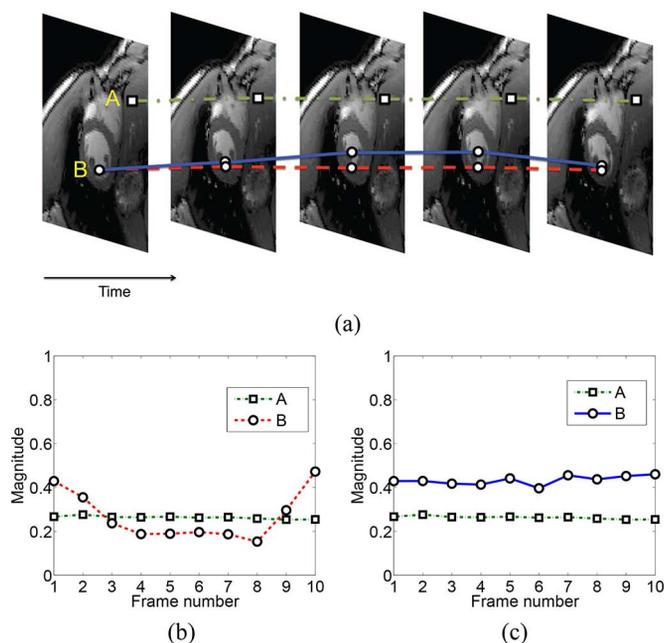


Fig. 1. (a) Dynamic heart motion along the temporal dimension. Pixels A and B are located in a relatively stationary background, and in the fast moving heart wall, respectively. Pixel value variations of A and B (b) along the temporal axis, and (c) along the motion path, respectively.

variations in Fig. 1(c) can be represented using a single basis function; so, the resulting temporal variations can be more effectively sparsified. This implies that to further accelerate data acquisition for cardiac cine MRI, we need a motion adaptive reconstruction framework.

The k-t FOCUSS with motion estimation/compensation (ME/MC) algorithm [4], [8] is one of the earliest works to address this issue. More specifically, when a high resolution reference frame is available, k-t FOCUSS with ME/MC searches for matching blocks from the reference frame in a prediction step, and the remaining motion residual signals are reconstructed using the k-t FOCUSS algorithm. However, one of the limitations of k-t FOCUSS with ME/MC is that after the prediction, the energy of the differences between the original k -space measurement and the synthesized k -space measurement from the predicted image components, is significantly reduced compared to that of the original measurements. Therefore, the resulting sparse recovery problem for the residual tends to be more difficult due to the reduction of the signal-to-noise ratio of the measurements. Hence, instead of using a prediction-residual scheme, Asif *et al.* [15] employed temporal total variation (TV) along motion trajectory as a regularization term. However, as will be explained later, such pixel-by-pixel motion regularization is often sensitive to noise.

Meanwhile, low-rank structures of images have been exploited by many researchers in the medical imaging area. While sparsity promoting methods such as k-t SPARSE or k-t FOCUSS with Fourier basis usually rely on analytic transforms to sparsify the data, the low-rank methods exploit the low-rank structure originated from spatiotemporal correlations of data itself. Liang *et al.* utilized the low rank structure in dynamic

MR image and have developed corresponding reconstruction methods [16]–[23]. Specifically, they proposed the so-called partially separable function (PSF) model that exploits the low-rank property of spatiotemporal volumes. Zhao *et al.* [24] combined ideas from k-t SPARSE/k-t FOCUSS with the previous low-rank modeling ideas from [16]–[23]. Adluru *et al.* used a reordering scheme to reduce the rank of spatiotemporal image [25].

Recently, it has been shown that patch-based signal processing is effective in capturing geometric self-similarity and is robust to noise as demonstrated in the nonlocal means algorithm [26], [27]. Collaborative filtering such as the block matching and 3-D filtering (BM3D) algorithm [28] is another important generalization of patch-based signal processing in which collections of similar patches are grouped and processed together to denoise the image using transform sparsity. Related ideas have been also applied to medical imaging. Specifically, Ravishankar and Bresler applied a dictionary learning algorithm [29] for static MRI reconstruction [30]. To generalize nonlocal regularization schemes, Yang *et al.* proposed a variational framework in compressed sensing MR reconstruction [31]–[33]. Wang *et al.* used patch-based regularization penalty for PET image reconstruction, which was shown superior to pixel-based regularization [34]. Furthermore, there are recent works that exploit the low-rankness of a set of similar patches. Trzasko *et al.* introduced a patch-based low-rank generalization for dynamic image denoising in [35] and for calibration-free parallel imaging applications [36]. Akçakaya *et al.* applied patch-based low-rank filtering for coronary artery imaging problem [37]. Chen *et al.* [38] searched for similar patches along the temporal motion and used the low-rank of a matrix formed from them.

In this paper, we extend the idea of patch-based signal processing further so that the geometric dissimilarities of patches within a spatiotemporal neighborhood are penalized. More specifically, we propose a *motion adaptive* spatio-temporal patch-based low rank penalty to capture geometric similarity along motion trajectory as well as within an image frame. This spatio-temporal low rank patch penalty stems from the observation that a low rank penalty is effective in extracting the major principal components such as mean intensity variations and edge structures, rather than geometrically insignificant features like noise and aliasing artifacts. Therefore, the low rank penalty tends to drive a reconstruction algorithm away from noise and aliasing artifacts.

However, this approach raises several technical difficulties. First, motion adaptive grouping of similar patches depends on the unknown spatio-temporal image, thus it varies with iterative image update. This causes the structure of the cost function to vary as well. Therefore, standard convergence proofs that utilize the convergence of a cost function value cannot be used. Second, the search for similar patches usually produces overlapping patches, in which the number of overlaps varies pixel-by-pixel. Hence, a penalized least squares formulation using a single regularization parameter often produces a pixel-dependent inhomogeneous penalty, which often results in slow convergence. To address these technical issues and guarantee the convergence of the algorithm, we propose a Nash

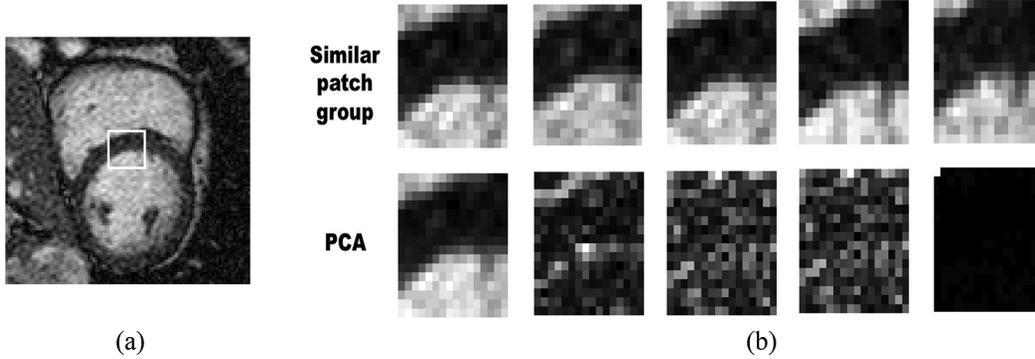


Fig. 2. (a) Reference patch selection. (b) Upper row: similar patch group. Lower row: principal components from the collection of similar patches (from left to right: PC1, PC2, . . . , PC5).

equilibrium (NE) formulation [39], [40] with relaxation (known as Krasnosel’skiĭ–Mann iteration [41] in the optimization literature), in which closed-form solutions for the decoupled subprograms are obtained in an alternating manner, followed by a relaxation step. We also provide a convergence proof for the proposed algorithm. Note that the proof in [39] assumes that the similarity structure does not vary over iterations, and to the best of our knowledge, convergence proofs under our situation were not available before.

Experiments using a numerical phantom and a real data set show that the proposed algorithm accurately reconstructs important anatomic structures in cardiac cine image and provides improved image quality compared to the existing state-of-the-art methods such as k-t SLR [7] and MASTeR [15].

II. THEORY

A. Measurement Model

Let $Y \in \mathbb{C}^{m \times T}$ and $X \in \mathbb{R}^{n \times T}$ denote the spatio-temporal k -space measurement and the unknown image, respectively, whose vectorizations are given by

$$\mathbf{x} = \text{vec}(X), \quad \mathbf{y} = \text{vec}(Y). \quad (1)$$

Here, m, n and T denote the numbers of k -space samples, pixels in each frame, and temporal frames, respectively. Then, a forward model for the dynamic MR imaging problem is given by

$$\mathbf{y} = F\mathbf{x} + \zeta \quad (2)$$

where ζ denotes the measurement noise and F denotes a spatio-temporal sensing matrix given by

$$F = \begin{bmatrix} F(1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & F(2) & & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & F(T) \end{bmatrix} \in \mathbb{C}^{mT \times nT}, \quad m < n \quad (3)$$

where $F(t)$ denotes the Fourier sensing matrix corresponding to the t th temporal undersampled k -space trajectory, which makes F rectangular, and thus makes the inverse problem defined by (2) underdetermined. In Cartesian acquisition, $F(t)$ is obtained such that undersampling is performed along the

phase encoding direction, whereas in a radial sampling pattern, a reduced number of spokes are used. In general, the sampling pattern depends on the k -space acquisition trajectory. In this paper, we assume a coil-by-coil reconstruction, but an extension to parallel imaging would be straightforward.

B. Spatio-Temporal Low-Rank Penalty

It is well known that a natural image has geometric self-similarities, i.e., some parts of the image are similar to other parts of the image. This is especially true for dynamic images, which represent continuously varying objects, since many similar parts are found in adjacent temporal frames. Fig. 2 shows an example. In Fig. 2(a), a reference patch is selected from the region of interest of a dynamic heart image. The upper row of Fig. 2(b) shows a collection of patches consisting of the reference patch as well as its similar patches retrieved in a 3-D spatio-temporal neighborhood. Note that such a search for similar patches in a temporal domain is similar to that used in block-matching algorithms in motion estimation. Furthermore, such self similarity can be effectively exploited to remove noise or aliasing artifacts. Specifically, the lower row of Fig. 2(b) shows principal components extracted from the above group of similar patches. Note that PC 1 and PC 2 (the first two figures) include the major edge information. Other principal components (PC3, PC4, and PC5) contribute little to image quality and normally they contain noise or unnecessary signals. This suggests that a low-rank constraint for the similar patches is effective in capturing the geometric information since the low rank decomposition is related to principal component analysis.

To write a patch-based low rank constraint, we define an operator $\mathcal{R}_{pq} : \mathbf{x} \mapsto \mathcal{R}_{pq}\mathbf{x} \in \mathbb{C}^B$ that extracts a vectorized version of a B -pixel square patch from an image frame in \mathbf{x} . Here, p and q denote the indexes for a reference patch and its corresponding similar patches, respectively, where $p = 1, \dots, P; q = 1, \dots, Q$.

As can be seen in Fig. 3, for a given reference patch $\mathbf{v}_{p1} = \mathcal{R}_{p1}\mathbf{x}$, we search for similar patches $\{\mathbf{v}_{pq}\}_{q=2}^Q$ within the current frame as well as within the adjacent temporal frames. As a similarity metric, we compute the l_2 difference between $\mathcal{R}_{p1}\mathbf{x}$ and a candidate, and choose the $Q-1$ patches whose differences are smaller than others. If patches are normalized, then the similarity search is based on the correlation coefficients between

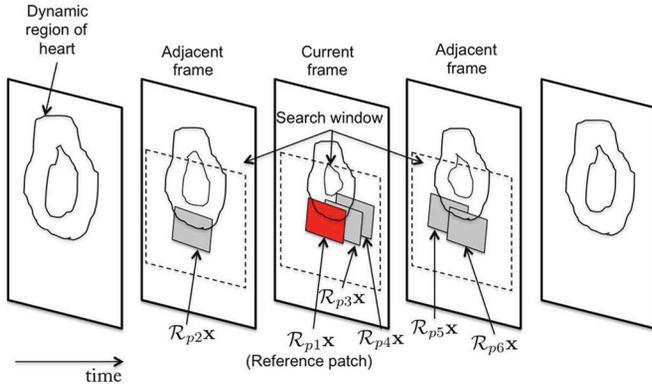


Fig. 3. Proposed similar patch search scheme using a spatio-temporal window. Search window in the current frame is centered at the reference patch. Patch search is also conducted in the same positioned window in multiple temporally adjacent frames. Similar patches that are identified are vectorized and grouped together to construct patch matrix $V_p(\mathbf{x})$, as shown in (4).

patches, and is invariant to scaling of the intensity of patches. To reduce the computational burden of the search, the search range is restricted to within a spatio-temporal window of fixed size around the reference patch $\mathcal{R}_{p1}\mathbf{x}$. We perform the same construction for the whole spatio-temporal image by moving the reference patches and by considering overlapping patches in each frame such that the same pixel can contribute multiple patches. Hence, the total number of patch groups P is about the same as the number of spatio-temporal voxels in \mathbf{x} , i.e., $P \simeq nT$.

Formally, define a mapping $V_p : \mathbf{x} \mapsto V_p(\mathbf{x})$ where

$$V_p(\mathbf{x}) = [\mathcal{R}_{p1}\mathbf{x}, \mathcal{R}_{p2}\mathbf{x}, \dots, \mathcal{R}_{pQ}\mathbf{x}] \in \mathbb{C}^{B \times Q}. \quad (4)$$

Then, its adjoint V_p^* places back the patches at the original positions in \mathbf{x} , and $V_p^*V_p : \mathbb{C}^{nT} \rightarrow \mathbb{C}^{nT}$ is a mapping from an image \mathbf{x} to another image of the same size. Let $(\mathbf{V}_p^*\mathbf{V}_p)$ denote the $nT \times nT$ matrix representing this operator. We call the number of times that the i th pixel appears in patch matrix $V_p(\mathbf{x})$ the *number of references* to the i th pixel by patch matrix $V_p(\mathbf{x})$. Then it is easy to show that

$$\sum_{p=1}^P (\mathbf{V}_p^*\mathbf{V}_p) = \mathbf{Z} \quad (5)$$

where $\mathbf{Z} \in \mathbb{R}^{nT \times nT}$ denotes a diagonal matrix whose (i, i) element is equal to the total number of references to the i th pixel by the patch matrices $\{V_p(\mathbf{x})\}_{p=1}^P$. Furthermore, we have the following identity:

$$\mathbf{x} = \mathbf{Z}^{-1} \left(\sum_{p=1}^P V_p^*V_p(\mathbf{x}) \right). \quad (6)$$

Accordingly, in our model, a patch-based self-similarity constraint is imposed using the following rank penalty:

$$\Psi(\mathbf{x}) = \sum_{p=1}^P \text{Rank}(V_p(\mathbf{x})). \quad (7)$$

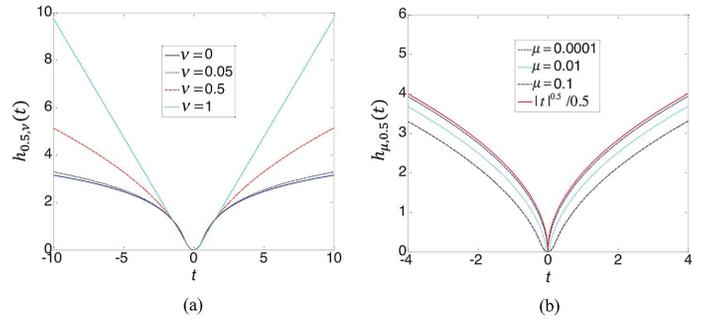


Fig. 4. Generalized ν -Huber function plot. In (a), μ is fixed while ν is varied. In (b), μ is varied while ν is fixed.

There is an important issue in using the penalty in (7). Specifically, the rank operator is not convex, so direct optimization using the rank penalty produces a nonconvex optimization problem. One of the popular convex relaxations for the rank function is the nuclear norm [42]. However, as it has been shown that a nonconvex penalty outperforms the convex nuclear norm, we define instead the following rank penalty [7], [43], [44]

$$\|V_p(\mathbf{x})\|_{h_{\mu, \nu}} = \sum_{k=1}^{\text{Rank}(V_p)} h_{\mu, \nu}(\sigma_k(V_p(\mathbf{x}))), \quad 0 < \nu \leq 1 \quad (8)$$

where $\|\cdot\|_{h_{\mu, \nu}}$ should be regarded as notation rather than a norm, $\sigma_k(V_p(\mathbf{x}))$ denotes the k -largest singular value of $V_p(\mathbf{x})$, and the generalized ν -Huber function $h_{\mu, \nu}(t)$ is defined as

$$h_{\mu, \nu}(t) = \begin{cases} |t|^2/2\mu, & \text{if } |t| < \mu^{1/(2-\nu)} \\ |t|^\nu/\nu - \delta, & \text{if } |t| \geq \mu^{1/(2-\nu)} \end{cases} \quad (9)$$

with $\delta = (1/\nu - 1/2)\mu^{\nu/(2-\nu)}$ to make the function continuous and differentiable. By using the definition (8), the penalty term is defined as

$$\Psi(\mathbf{x}) = \sum_{p=1}^P \|V_p(\mathbf{x})\|_{h_{\mu, \nu}}. \quad (10)$$

Fig. 4 shows the shapes of generalized ν -Huber function $h_{\mu, \nu}(t)$ with various μ and ν parameters. As can be seen in Fig. 4(a), the generalized Huber function is not convex for $0 < \nu < 1$ except in the quadratic region near the origin. The reason for allowing the quadratic region at the origin will be explained later. Fig. 4(b) shows the effect of the μ parameter. As μ increases, the magnitude of the Huber function decreases, and the quadratic region expands.

The generalized ν -Huber function is closely related to Schatten ν -quasi norm that includes the popular nuclear norm as a special case. More specifically, the Schatten ν -norm for the matrix $V_p(\mathbf{x})$ is defined as

$$\|V_p(\mathbf{x})\|_\nu = \sum_{k=1}^{\text{Rank}(V_p)} \sigma_k^\nu(V_p(\mathbf{x})). \quad (11)$$

Note that this corresponds to the nuclear norm when $\nu = 1$. However, as far as we know, there exists no closed form shrinkage rule for Schatten ν -quasi-norm except for the cases $\nu = 1$ (Nuclear norm) and $\nu = 1/2$ [45]. Therefore, existing rank-minimization approaches using Schatten ν -quasi-norm usually resort to a reweighted norm approach [46] or to an approximation using the Huber function. Specifically, the approximation using the Huber function has been used in a low rank matrix completion framework for the entire spatio-temporal image volume in dynamic MRI in [7] and [43].

Compared to the Schatten quasi-norm, the proposed rank proxy has several advantages that allow a closed form shrinkage rule. More specifically, for the generalized ν -Huber function in (9), it is easy to show that $h_{\mu,\nu}(t) - |t|^2/(2\mu)$ is strictly concave for $0 \leq \nu < 1$ as it subtracts the quadratic term at the origin [43], [44]. Hence, there exists a convex conjugate of the form $|s|^2/2\mu + g_{\mu,\nu}(s)$ for some function $g_{\mu,\nu}(s)$ such that the following Legendre–Fenchel transform holds:

$$\mu h_{\mu,\nu}(t) - |t|^2/2 = \min_s \{ |s|^2/2 + \mu g_{\mu,\nu}(s) - ts \} \quad (12)$$

which leads us to the following half-quadratic regularization form:

$$h_{\mu,\nu}(t) = \min_s \{ |s - t|^2/\mu + g_{\mu,\nu}(s) \}. \quad (13)$$

Since an operation on the singular values does not change when applying an orthonormal transformation matrix, the corresponding rank penalty for patch matrix V_p is given by

$$\begin{aligned} \|V_p(\mathbf{x})\|_{h_{\mu,\nu}} &= \sum_{k=1}^{\text{Rank}(V)} h_{\mu,\nu}(\sigma_k(V_p(\mathbf{x}))) \\ &= \sum_{k=1}^{\text{Rank}(V_p)} \left\{ \min_s |s - \sigma_k(V_p(\mathbf{x}))|^2/\mu + g_{\mu,\nu}(s) \right\} \\ &= \min_{\{s_k\}} \sum_k |s_k - \sigma_k(V_p(\mathbf{x}))|^2/\mu + g_{\mu,\nu}(s_k) \\ &= \min_{W \in \mathbb{C}^{B \times Q}} \left\{ \frac{1}{\mu} \|V_p(\mathbf{x}) - W\|_F^2 + \|W\|_{g_{\mu,\nu}} \right\} \end{aligned} \quad (14)$$

where the notation $\|W\|_{g_{\mu,\nu}}$ is defined as

$$\|W\|_{g_{\mu,\nu}} = \sum_{k=1} g_{\mu,\nu}(\sigma_k(W)). \quad (15)$$

Then, equivalent representation of the penalty term is given by

$$\Psi(\mathbf{x}) = \min_{W_p} \sum_{p=1}^P \left\{ \frac{1}{\mu} \|V_p(\mathbf{x}) - W_p\|_F^2 + \|W_p\|_{g_{\mu,\nu}} \right\}. \quad (16)$$

For $\nu = 1$, $g_{\mu,\nu}(s)$ becomes $|s|$ and $\|W_p\|_{g_{\mu,\nu}}$ becomes $\|W_p\|_*$ —the nuclear norm [44], but in general, there exists no closed form expression for $g_{\mu,\nu}(s)$. However, as will be shown later, the expression for $g_{\mu,\nu}$ itself is not important, since (13) and (16) allow closed form shrinkage rules [44].

C. Nash Equilibrium (NE) Problem Formulation

A standard formulation of the image recovery problem is given by a penalized least-squares approach

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \mathcal{C}(\mathbf{x}), \quad \mathcal{C}(\mathbf{x}) \triangleq \|\mathbf{y} - F\mathbf{x}\|_F^2 + \lambda \Psi(\mathbf{x}) \quad (17)$$

where λ denotes the regularization parameter. However, in our patch-based low rank approach, the standard formulation introduces several technical issues. First, the grouping of similar patches depends on the unknown image. Hence, if we use an iterative algorithm that successively updates the image content, the resulting similarity structure and the corresponding rank penalty need to be updated too from one iteration to the next. This causes the structure of the cost function to vary over iterations. Second, the number of pixel references generated by the search for similar patches is not homogeneous across the image. Hence, the rank penalty with a single regularization parameter as in (17) highly penalizes the pixels with a large number of references. As discussed in Section IV, the resulting nonuniform regularization often makes the algorithm highly sensitive to the choice of the regularization parameter λ .

To address these technical issues, this paper converts the reconstruction problem to a multi-objective optimization problem [47] and employs a Nash equilibrium (NE) formulation [39]. More specifically, we are interested in finding a Pareto optimum of the two objective functions corresponding to the rank penalty and data fidelity, respectively

$$\begin{aligned} \{W_p^{(k+1)}\}_{p=1}^P &= \arg \min_{\{W_p\}} \sum_{p=1}^P \frac{1}{\mu} \|V_p(\mathbf{x}^{(k)}) - W_p\|_F^2 \\ &\quad + \|W_p\|_{g_{\mu,\nu}}, \end{aligned} \quad (18)$$

$$\hat{\mathbf{x}}^{(k+1)} = \arg \min_{\mathbf{x}} \|\mathbf{y} - F\mathbf{x}\|^2 + \lambda \|\mathbf{x} - \mathbf{w}^{(k+1)}\|^2 \quad (19)$$

where $\mathbf{w}^{(k+1)} = Z^{-1}(\sum_{p=1}^P V_p^* W_p^{(k+1)})$. As will be discussed later in detail in Section IV, the NE formulation decouples the data fidelity and denoising, and overcomes the aforementioned drawbacks of the standard penalized least square by providing guaranteed convergence and making the algorithm robust to the choice of regularization parameter. Furthermore, each subprogram results in a closed form solution as described in the following.

1) *Minimization With Respect to $\{W_p\}$* : To solve for $\{W_p\}$, we need $\{V_p(\mathbf{x})\}_{p=1}^P$ that are constructed using the similarity relationship $\{\mathcal{R}_{pq}^{(k)}\}_{p,q=1}^{P,Q}$, which again needs to be estimated from the image \mathbf{x} . However, in our algorithm, the image \mathbf{x} is the very object that needs to be estimated, so $\{\mathcal{R}_{pq}^{(k)}\}_{p,q=1}^{P,Q}$ is not known *a priori*. To address this issue, we rely on a refinement scheme: we iteratively estimate a similarity mapping $\{\mathcal{R}_{pq}^{(k)}\}_{p,q=1}^{P,Q}$ using the previous estimate of the image by following the procedure described in Section II-B and Fig. 3.

Then, for a given estimate of similar patch grouping $\{V_p^{(k)}(\mathbf{x}^{(k)})\}_{p=1}^P$, we need to solve the following minimization problem patch-by-patch:

$$\begin{aligned} W_p^{(k+1)} &= \arg \min \left\{ \frac{1}{\mu} \|V_p^{(k)}(\mathbf{x}^{(k)}) - W_p\|_F^2 + \|W_p\|_{g_{\mu,\nu}} \right\}. \end{aligned} \quad (20)$$

Chartrand [44] and Chartrand and Saneva [48] showed that (20) has a closed form solution given as

$$W_p^{(k+1)} = L \text{shrink}_\nu(\Sigma, \mu) U^H \quad (21)$$

where $\text{shrink}_\nu(\Sigma, \mu)$ denotes an element by element singular value shrinkage operator with the following shrinkage rule:

$$\text{shrink}_\nu(t, \mu) := \max\{0, |t| - \mu|t|^{\nu-1}\} t/|t|. \quad (22)$$

For $\nu = 1$, the shrinkage rule is equivalent to soft thresholding. Indeed, (22) generalizes soft-thresholding to the non-convex case. For example, as $\nu \rightarrow 0$, the shrinkage operation approaches hard thresholding, and the result is equivalent to PCA that retains the significant few principal components.

2) *Minimization With Respect to \mathbf{x}* : Next, for a given $\{W_p^{(k+1)}\}$, the minimization in (18) with respect to \mathbf{x} is just a linear least-squares step

$$(F^H F + \lambda I) \mathbf{x} = F^H \mathbf{y} + \lambda \mathbf{w}^{(k+1)} \quad (23)$$

whose closed form solution is given by

$$\hat{\mathbf{x}}^{(k+1)} = (F^H F + \lambda I)^{-1} (F^H \mathbf{y} + \lambda \mathbf{w}^{(k+1)}). \quad (24)$$

In Cartesian sampling, the matrix inversion in (24) can be simplified because F and F^H are implemented on a Cartesian grid and $F^H F$ is circulant. More specifically, multiplying by the full Fourier matrix F_f , (24) can be written as

$$F_f (F^H F F_f^H + \lambda F_f^H) F_f \mathbf{x} = F_f F^H \mathbf{y} + \lambda F_f \mathbf{w}^{(k+1)} \quad (25)$$

where the matrix $F_f F^H F F_f^H$ is diagonal with one only at sampled positions, and $F_f F_f^H = I$. Then, problem (25) can be solved by simple element-wise inversion and FFT. For general sampling patterns such as radial or spiral, the matrix inversion problem in (24) can be easily solved using a conjugate gradient method.

3) *Relaxation for the Update of \mathbf{x}* : The final step of the proposed algorithm is a relaxation step (often called Krasnosel'skii–Mann iteration) [41]. More specifically, rather than using the solution in (24) as the update for the $k + 1$ th estimate of the unknown image, we introduce a relaxation factor $\beta \in (0, 1)$ such that the new update is the weighted average between (24) and the previous update $\mathbf{x}^{(k)}$, i.e.,

$$\begin{aligned} \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \beta (\hat{\mathbf{x}}^{(k+1)} - \mathbf{x}^{(k)}) \\ &= \mathbf{x}^{(k)} + \beta (T_{(k)} \mathbf{x}^{(k)} - \mathbf{x}^{(k)}) \end{aligned} \quad (26)$$

where $T_{(k)}$ denotes the entire operator including shrinkage and least square fitting such that maps $\mathbf{x}^{(k)}$ to $\hat{\mathbf{x}}^{(k+1)}$, i.e., $\hat{\mathbf{x}}^{(k+1)} = T_{(k)} \mathbf{x}^{(k)}$. The main motivation for introducing the relaxation parameter in (26) is to guarantee the convergence of the iteration

Algorithm 1 Motion Adaptive Dynamic MRI using Spatio-Temporal Low Rank Patch Penalty

- 1: Determine $\mathbf{x}^{(0)}$ using k-t FOCUSS with NLMC algorithm (**Algorithm 2**).
 - 2: Set $\eta, \mu, \nu > 0, \beta \in (0, 1)$.
 - 3: Set spatio-temporal search window size for similar patch search.
 - 4: **for** $k = 0, 1, \dots$ **do**
 - 5: **for** $p = 1, \dots, P$ **do**
 - 6: $V_p^{(k)}(\mathbf{x}^{(k)}) = [\mathcal{R}_{p1}^{(k)} \mathbf{x}^{(k)}, \mathcal{R}_{p2}^{(k)} \mathbf{x}^{(k)}, \dots, \mathcal{R}_{pQ}^{(k)} \mathbf{x}^{(k)}]$
 - 7: $[L, \Sigma, U] = \text{svd}(V_p^{(k)}(\mathbf{x}^{(k)}))$
 - 8: $W_p^{(k+1)} = L \text{shrink}_\nu(\Sigma, \mu) U^*$
 - 9: **end for**
 - 10: $\mathbf{w}^{(k+1)} = (\sum V_p^* V_p)^{-1} (\sum V_p^* W_p^{(k+1)})$
 - 11: $\hat{\mathbf{x}}^{(k+1)} = (F^H F + \lambda I)^{-1} (F^H \mathbf{y} + \lambda \mathbf{w}^{(k)})$.
 - 12: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \beta (\hat{\mathbf{x}}^{(k+1)} - \mathbf{x}^{(k)})$.
 - 13: **end for**
-

to a fixed point.¹ Then, we have the following main convergence result Theorem 2.1 for our algorithm. Here, the theorem is based on the assumption that the intersection of the fixed points sets is not empty after enough iterations. This condition is likely to hold in practice because the similar patch selection after enough iterations does not change too much. Here, Algorithm 1 summarizes the overall steps of the proposed method.

Theorem 2.1: For $\nu = 1$, the sequence $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$ generated by Algorithm 1 converges to a fixed point.

Proof: See Appendix. \square

The resulting fixed point from Algorithm 1 is a NE solution for the two objective functions given in (18) and (19) that correspond to the low rank penalty and data fidelity criterion, respectively. Here, an NE solution implies that an individual solution can receive no incremental benefit from changing its value, assuming other solution remain unchanged. Hence, the fixed point produced by the proposed algorithm can be regarded as one of the compromises between the denoising from low rank penalty and data fitting [47].

In Algorithm 1, it is important to note that the similarity functional mapping $\{\mathcal{R}_{pq}^{(k)}\}_{p,q=1}^{P,Q}$ does not need to be stored at all since it is only used for constructing $V_p^{(k)}(\mathbf{x}^{(k)})$, so it can be found patch-by-patch on the fly. This simplifies the implementation of the algorithm.

D. Initialization

Since the search for similar patches depends on the successive image update, we need a reasonable initialization for good performance. Our approach is to use the k-t FOCUSS algorithm with ME/MC [4], [8], where the prediction term is estimated using nonlocal motion compensation (NLMC) [50]. Recall that k-t FOCUSS with ME/MC is formulated as the following reweighted norm approach [4], [8]:

$$\Delta \hat{\mathbf{x}}_l = \arg \min_{\Delta \mathbf{x}} \|\mathbf{y} - F \bar{\mathbf{x}} - F \Delta \mathbf{x}_l\|_2^2 + \eta \|D_l^{-1} \Delta \mathbf{x}_l\|_2^2 \quad (27)$$

¹Note that we only provide a proof for $\nu = 1$ since the proof depends on the firmly nonexpansive shrinkage operator [49]. In general, the shrinkage rule for $\nu < 1$ is not firmly nonexpansive, but our empirical results confirm that the algorithm still converges.

Algorithm 2 k-t FOCUSS with NLMC

- 1: Generate an intermediate quality image \mathbf{x} using k-t FOCUSS.
- 2: Generate a reference frame \mathbf{r} from the k -space data during the diastole phase.
- 3: Set $\bar{\mathbf{x}} = \Delta\mathbf{x}_0 = \mathbf{0}$.
- 4: **for** $p = 0, 1, \dots$ **do**

$$\mathcal{R}_p \bar{\mathbf{x}} \leftarrow \frac{\sum_{i \in \mathcal{N}_p} \mathcal{R}_i \mathbf{r} \exp\left(-\frac{\|\mathcal{R}_p \bar{\mathbf{x}} - \mathcal{R}_i \mathbf{r}\|^2}{h}\right)}{\sum_{i \in \mathcal{N}_p} \exp\left(-\frac{\|\mathcal{R}_p \bar{\mathbf{x}} - \mathcal{R}_i \mathbf{r}\|^2}{h}\right)},$$

- 5: **end for**
- 6: $\bar{\mathbf{x}} = (\sum_p \mathcal{R}_p^* \mathcal{R}_p)^{-1} (\sum_p \mathcal{R}_p^* \mathcal{R}_p(\bar{\mathbf{x}}))$
- 7: Set the unknown by $\mathbf{x} = \bar{\mathbf{x}} + \Delta\mathbf{x}$ and obtain $\Delta\mathbf{x}$ using k-t FOCUSS in (27).
- 8: Output $\mathbf{x}^0 \leftarrow \bar{\mathbf{x}} + \Delta\hat{\mathbf{x}}$

where D_l is a diagonal weighting matrix that is updated using the previous step update \mathbf{x}_{l-1} , and $\bar{\mathbf{x}}$ and $\Delta\mathbf{x}$ denote a prediction term and a residual difference term from the true \mathbf{x} , respectively. Note that the formulation exploits the sparsity of the residual signal. Hence, a better prediction $\bar{\mathbf{x}}$ produces a sparser residual $\Delta\mathbf{x}$ and the compressed sensing recovery improves. To provide a better prediction term $\bar{\mathbf{x}}$, we exploit the idea of non-local means [26]. However, unlike the nonlocal means, the similar patches are identified and taken from a reference frame that is not included in the image frames that are to be reconstructed. In cardiac cine imaging, a reference frame is generated by collecting k -space data from the diastole phase as shown in Fig. 5. Then, starting from an initial k-t FOCUSS [4], [8] reconstruction, the following fixed point equation is used to update the current patch $\mathcal{R}_p \mathbf{x}$

$$\mathcal{R}_p \mathbf{x} \leftarrow \frac{\sum_{i \in \mathcal{N}_p} \mathcal{R}_i \mathbf{r} \exp\left(-\frac{\|\mathcal{R}_p \mathbf{x} - \mathcal{R}_i \mathbf{r}\|^2}{h}\right)}{\sum_{i \in \mathcal{N}_p} \exp\left(-\frac{\|\mathcal{R}_p \mathbf{x} - \mathcal{R}_i \mathbf{r}\|^2}{h}\right)}. \quad (28)$$

In (28), $h > 0$ is a hyper parameter, and \mathbf{r} denotes the reference frame. \mathcal{R}_p denotes an operator to extract the p th patch. The set \mathcal{N}_p is the neighborhood used to search for similar patches. It is defined as the subset of patches within a predefined search window in the reference image such that the l_2 norm difference from the p th patch of \mathbf{x} is less than some predefined threshold value. Because we allow patch overlap, the overlapped pixel values are averaged after each iteration of (28). After obtaining a prediction term by solving (28), we apply the k-t FOCUSS algorithm (27) and solve for the $\Delta\mathbf{x}$. We call this combined scheme k-t FOCUSS algorithm with NLMC (nonlocal motion compensation). This initialization step is performed once at the beginning of the algorithm, starting from the raw data \mathbf{y} , to produce the initial reconstructed image \mathbf{x}^0 . Algorithm 2 summarizes the initial reconstruction process.

III. EXPERIMENTAL RESULTS

A. Method

The proposed algorithm was validated by experiments using numerical dynamic cardiac phantom data and real *in vivo*

dynamic cardiac data sets with different sampling trajectories. Retrospective downsampling from the fully sampled data set was used to compare the results with those of fully sampled reconstruction. During the diastole phase, we used a uniform random sampling pattern, and for the remaining phases, we used a Gaussian random sampling pattern.

The first data set was obtained using a Mathematical Cardiac-Torso (MCAT) numerical phantom that mimics dynamic heart motion. We used the program by Segars *et al.* [51], [52]. It contains left and right ventricle structures and shows systole and diastole phase dynamics. The matrix size was set to 128×128 , and the number of time frames was 50. The data was corrupted by additive Gaussian noise to produce SNR of 33.8 dB and the downsampling rate used for the experiment was 6.1.

A real cardiac cine data set was acquired using a 1.5T Philips scanner at Yonsei University Medical Center, Korea: the acquisition sequence was ECG triggered bSSFP, with a flip angle of 50° . The heart rate of the subject was 75 bpm and retrospective cardiac gating was used for data acquisition. The imaging parameters were as follows: FOV = 345×270 mm², matrix size = 256×256 , TR = 3.17 ms, the number of cardiac phases was 25, slice thickness 10.0 mm and the sampling trajectory was Cartesian. Another cardiac cine data set with radial trajectory was acquired at Yonsei University Medical Center, Korea, using an electrocardiogram-triggered segmented bSSFP pulse sequence from a healthy volunteer on a 3T whole-body MR scanner (Magnetom Trio; Siemens Medical Solutions, Erlangen, Germany). Imaging parameters were as follows: FOV 200×300 mm², TR/TE 3.4/1.7 ms, flip angle 40° , and slice thickness 3 mm. The number of views for each frame was 192: the sample number along the readout direction was 384. Thus, the matrix size was 384×384 . The cardiac parameters were as follows: 32 cardiac phases, eight views per segment, and 24 heartbeats. This provided us with 192 views per frame and temporal resolution of 3.4×8 ms \cong 30 ms. The downsampling rate was 8. Specifically, 48 spokes of radial trajectory were uniformly sampled at each frame, and those sampling locations were randomly rotated at each frame such that there was no overlap in sampling trajectory along the temporal dimension. In the diastole phase, these downsampled spokes were collected and used for reconstruction to generate a reference frame, as shown in Fig. 5. Both *in vivo* data sets were obtained with breath-holding.

We compared k-t FOCUSS with NLMC, k-t SLR, MASTeR, and the proposed algorithm. The k-t SLR algorithm exploits the global low-rank structure of spatiotemporal signals [7]. Specifically, the problem is posed to solve a matrix recovery problem with a low-rank penalty. A sparsity prior is also used to improve the reconstruction quality [7]. For k-t SLR, we used the source codes from <http://www.engineering.uiowa.edu/~jcb/software.html>. MASTeR (motion-adaptive spatio-temporal regularization) is a pixel intensity based regularization scheme [15]. MASTeR models temporal sparsity using motion-adaptive linear transforms between neighboring frames. Specifically, forward- and backward pixel motions are estimated and then the estimated motion operators are used in an l_1 regularization problem to reconstruct a motion compensated

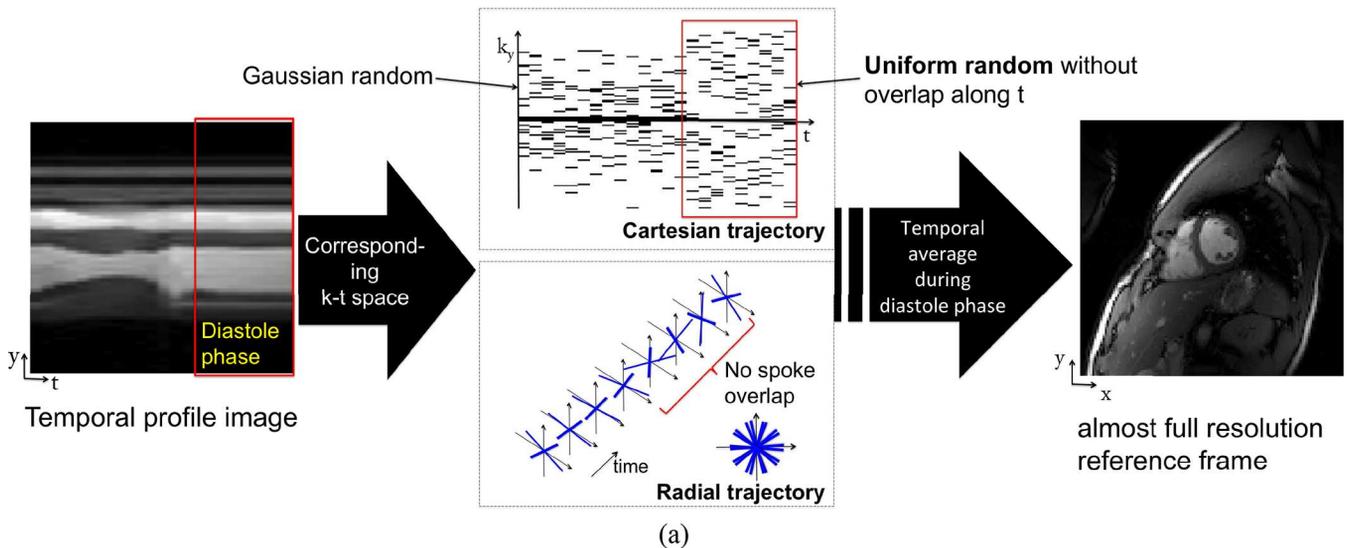


Fig. 5. Scheme for reference frame generation. Reference image with high resolution can be generated by using the properties of the diastole phase: the temporal average of uniformly down-sampled k-space data during diastole phase provides a high-resolution reference image.

image. For MASTeR, we used the open source codes from <http://users.ece.gatech.edu/~sasif/dynamicMRI/index.html>.

The algorithms are compared visually, as well as quantitatively. As numerical measures, we adopted the mean squared error (MSE) and the high frequency normalized error norm (HFEN) [30], [53]. HFEN is computed to complement MSE because sometimes the MSE values do not fully reflect the visual quality. HFEN is computed as

$$\text{HFEN} = \frac{1}{N} \sum_{i=1}^N \frac{\|\text{LoG}(\mathbf{x}_i) - \text{LoG}(\hat{\mathbf{x}}_i)\|_F^2}{\|\text{LoG}(\mathbf{x}_i)\|_F^2} \quad (29)$$

where \mathbf{x} is the ground truth and $\hat{\mathbf{x}}$ is the reconstruction image, LoG stands for Laplacian of Gaussian filter. In this experiment, the filter size and a standard deviation were set to 15×15 and 1.5, respectively, as in [30] and [53].

B. Numerical Phantom Results

Reconstruction results at two cardiac phases with the largest motions are compared in (A) and (B) of Fig. 6. The first rows show image frames at a certain time index, and the second rows show their difference images from the ground truth. The difference images are zoomed in to emphasize the differences between the algorithms. The positions of the selected cardiac phases are indicated in the temporal profile image Fig. 6(C). In Fig. 6(A), k-t FOCUSS with NLMC results show blurring artifacts near edges and it is difficult to observe the fine valve structure. Also, there still remains temporal blurring as shown in Fig. 6(C). While k-t SLR shows results comparable to those of k-t FOCUSS with NLMC, it also performs denoising, so the overall image appears less noisy. However, as shown in Fig. 6(A), the fine structures such as edges and valves were not clearly reconstructed. As indicated by the arrows in Fig. 6(A), MASTeR results also show aliasing artifacts near edge and valve structures. On the other hand, as shown in Fig. 6(A), in the proposed algorithm, edges in the dynamic region and valve structure were clearly reconstructed. In Fig. 6(B), the results

show a similar trend, and the proposed algorithm recovers the fine valve structure well. Fig. 6(C) shows temporal profile images, which clearly show that the proposed method recovers detailed temporal edges. In Figs. 6(D) and (E), the algorithms are compared using quantitative measures, MSE and HFEN. From these visual and numerical comparisons, we can see that the proposed algorithm enjoys several advantages over the other algorithms.

In Fig. 7(a) and (b), the MSE and HFEN values of the algorithms are plotted along the SNR levels by fixing the reduction factor as 6.1. Here, we found that the proposed algorithm, k-t SLR, and MASTeR all have similarly low MSE and HFEN numbers at high SNR levels. However, MASTeR has high MSE for low SNR. This is because it is difficult to obtain accurate pixel-by-pixel motion estimates at low SNR. Fig. 7(c) and (d) shows the performance dependence on the sampling reduction factors by fixing SNR = 33.8 dB. The proposed algorithm shows values lower or comparable to those of other algorithms under various reduction factors.

There are several parameters that need to be determined in the proposed algorithm. These parameters were selected to optimize MSE values of the reconstructions. In Fig. 8, MSE values are plotted with respect to the magnitudes of the regularization parameters on a log scale. We can see that MSEs are relatively unaffected in a wide range of λ, μ, ν values. Patch size, window size, and number of patches Q were fixed as 4×4 , $10 \times 10 \times 5$, and 5, respectively. In Fig. 8(a), MSE value stays low in a range of λ parameter ($10^{-2} \sim 1.0$). In Fig. 8(b), MSE values remain small with respect to μ in a range of $10^{-1} \sim 10^{-3}$. Fig. 8(c) shows the following: when μ is large, the MSE is not affected much by ν , but for a small μ , smaller ν value is better since it gives a hard-thresholding effect. The relaxation parameter $\beta \in (0, 1)$ gives a trade-off between the convergence guarantee and the speed of the algorithm, and a small value for $\beta \simeq 0$ slows convergence. Based on our experiments, we found that a fixed value $\beta = 0.95$ did not reduce the convergence speed, so we used this value throughout the paper.

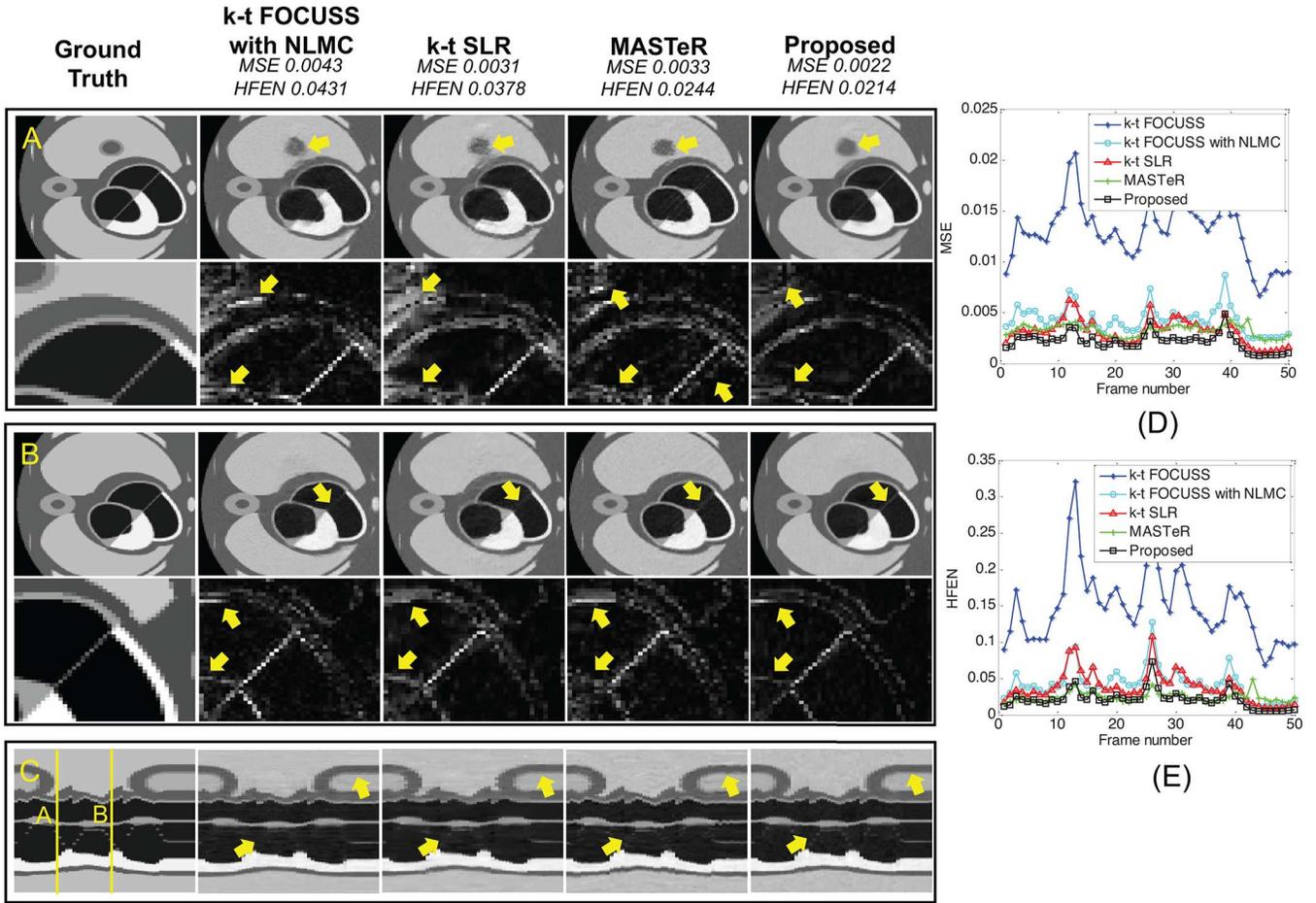


Fig. 6. Reconstruction results of the MCAT dynamic cardiac phantom with 33.8 dB Gaussian noise. The k -space data was obtained in cartesian coordinates, and the acceleration ratio was $\times 6.1$. Box A represents the 13th frame images from the systole phase of the ground-truth, reconstruction results using k-t FOCUSS with NLMC, k-t SLR, MASTeR, and the proposed algorithm, respectively. The first row represents image frames, and the second row represents difference frames. Box B represents the 28th frame images from the diastole phase of the ground-truth, reconstruction results using k-t FOCUSS with NLMC, k-t SLR, MASTeR, and the proposed algorithm, respectively. Box C represents the 80th temporal slice profiles of ground-truth, reconstruction results using k-t FOCUSS with NLMC, k-t SLR, MASTeR, and the proposed method, respectively. In box C, the time instants of frame A and frame B are indicated. Arrows indicate important areas such as cardiac edges and valve structures. Fig. 6(D) and (E) represents MSE and HFEN values along time frames, respectively.

Fig. 9(a)–(c) shows MSE plots with respect to patch and search window sizes. Parameters λ , μ , and ν were fixed as 0.07, 0.005, and 0.003, respectively. From the MSE curves for patch and window sizes, we observed that large size is not needed. Hence, it is beneficial to use a small parameters to reduce the computational cost. Fig. 9(d) also shows that the number Q of similar patches need not be large since after some threshold the performance improvement saturates. These observations show that the search sizes and the number of patches do not have to exceed some extent.

C. Real In Vivo Data

Based on the observations in Section III-B of the effect of parameters, for the case of real Cartesian data with $256 \times 256 \times 25$ matrix size, we used the following parameters. Patch size was 4×4 and the spatio-temporal window size was $10 \times 10 \times 4$. The number of similar patches was fixed as $Q = 5$. Parameters λ , μ and ν were set to 0.001, 0.05, and 0.02, respectively. Fig. 10 shows the reconstruction results of real Cartesian data with matrix size of $256 \times 256 \times 25$. Frames (A) and (B) are in the systole and diastole phases, respectively. Second rows of (A) and

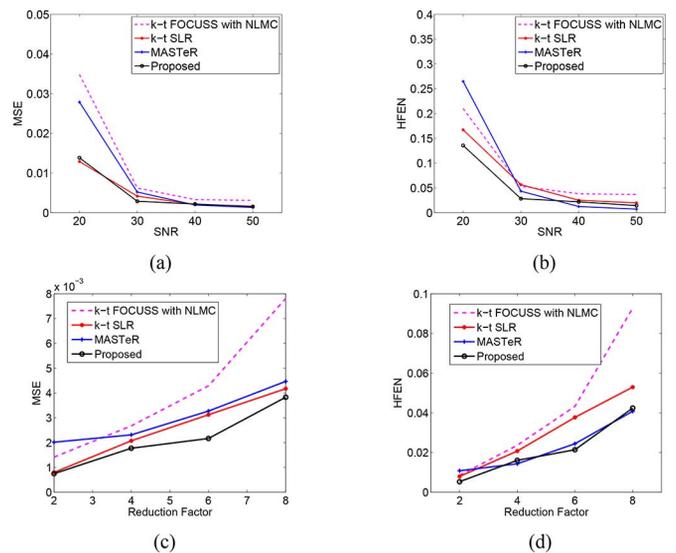


Fig. 7. (a), (b) MSE and HFEN plots for various algorithm for the fixed reduction factor of 6.1. (c), (d) MSE and HFEN plots for various algorithm for the fixed SNR of 33.8 dB.

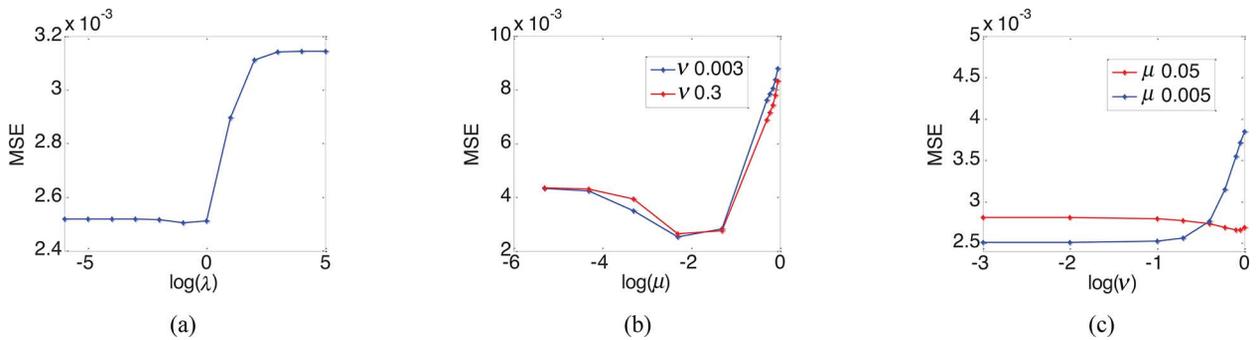


Fig. 8. MSE plots with respect to various hyper-parameters. The x -axis is represented in log scale. Images (a), (b), and (c) represent λ , μ , and ν , respectively. In (a), μ and ν were fixed as 0.005 and 0.003. In (b) and (c), λ was fixed as 0.07.

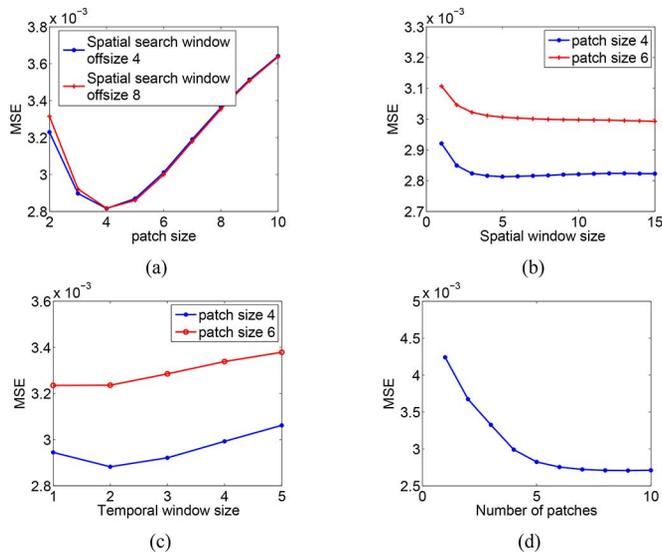


Fig. 9. MSE plots versus (a) patch size, (b) spatial window size variation, (c) temporal window size, and (d) number of similar patches. Patch and window sizes are defined as follows: Patch size 4 represents a 4×4 sized patch. Spatial window size 4 means 12×12 sized spatial window for patch size 4. Temporal window size 2 stands for temporal dimension of the spatio-temporal window equal to 5—the frame including the reference patch and the two adjacent frames before and after the given frame.

(B) show difference images in zoomed regions. In Fig. 10(A), k-t SLR and MASTeR show artifacts near the edges. While k-t FOCUSS with NLMC shows smooth edge reconstruction, the proposed algorithm provides clearer edge structures. In the diastole frame reconstructions (B), all the algorithms show better quality than those in the systole phase (A), but k-t SLR and the proposed algorithm show slightly better reconstruction results. In temporal profile images (C), the proposed algorithm shows the best performance. k-t FOCUSS with NLMC suffers from temporal blurring and noise. k-t SLR shows high denoising quality. However, edges were not well reconstructed. MASTeR suffers from rough motion reconstruction. The proposed algorithm shows the motion most clearly. Averaged MSE and HFEN values and their values along temporal frames [(D) and (E)] show that the proposed algorithm is better than other algorithms.

To confirm that the proposed algorithm provides better reconstruction than the existing ones regardless of sampling trajectory, Fig. 11 shows the reconstruction results for radial data. Parameters were set as follows: patch size was 4×4 , window size

was $12 \times 12 \times 5$, μ , ν , λ , and Q were 0.005, 0.03, 0.1, and 5, respectively. In Fig. 11, we can observe trends similar to the Cartesian case. In Fig. 11(A), k-t FOCUSS with NLMC shows remaining artifacts within the left ventricle. k-t SLR shows cleaner reconstruction, but it does not capture edge structure clearly near rapid heart motion areas. In the MASTeR result, aliasing artifacts remain and some parts of edges were not well reconstructed. The proposed algorithm shows clear distinction between cardiac structures and offers much improved heart boundaries. In the case of the diastole frame reconstructions (B), all the algorithms show better reconstruction than systole frame (A). Specifically, k-t SLR and the proposed algorithm show the best quality. Temporal profile images in Fig. 11(C) again confirm the advantages of the proposed algorithm. The proposed algorithm clearly depicts moving structures better than other algorithms. Quantitative measures in (D) and (E) support the observation that the proposed algorithm is better than the other during systole, while k-t SLR is better during diastole. The advantage of the proposed algorithm over k-t SLR is that it shows low MSE values over all time frames, whereas k-t SLR shows higher deviations along frames.

D. Computation Time

The proposed algorithm was implemented using MATLAB on an Intel i3-3220 processor running at 3.30 GHz with 4 GB RAM. We measured the computation time for the above real data sets. For the case of real Cartesian data with $256 \times 256 \times 25$ matrix size, the alternating minimization steps of singular value thresholding, as well as the matrix inverse in (24) were iterated five times. For k-t FOCUSS with NLMC, the patch size was 4×4 and the window size was 18×18 . Then, the total computation time of the proposed method was about 411 s, which includes the time required for initialization with k-t FOCUSS with NLMC. For k-t SLR, the following parameters were used: μ_1 and μ_2 parameters were chosen from several experiments, 0.001 and 0.0001, respectively; k-t SLR β parameter was obtained from continuation from a small value; 0.1 was used as nonconvex Schatten quasi-norm for singular value shrinkage. A fast implementation of SVD was used. Then, the computation time was about 428 s. For MASTeR results, CWT wavelet transform with level 4 was used and motion type was bi-directional. The computation time was 6334 s. These parameters for k-t SLR and MASTeR were selected that yielded the minimum MSE values among

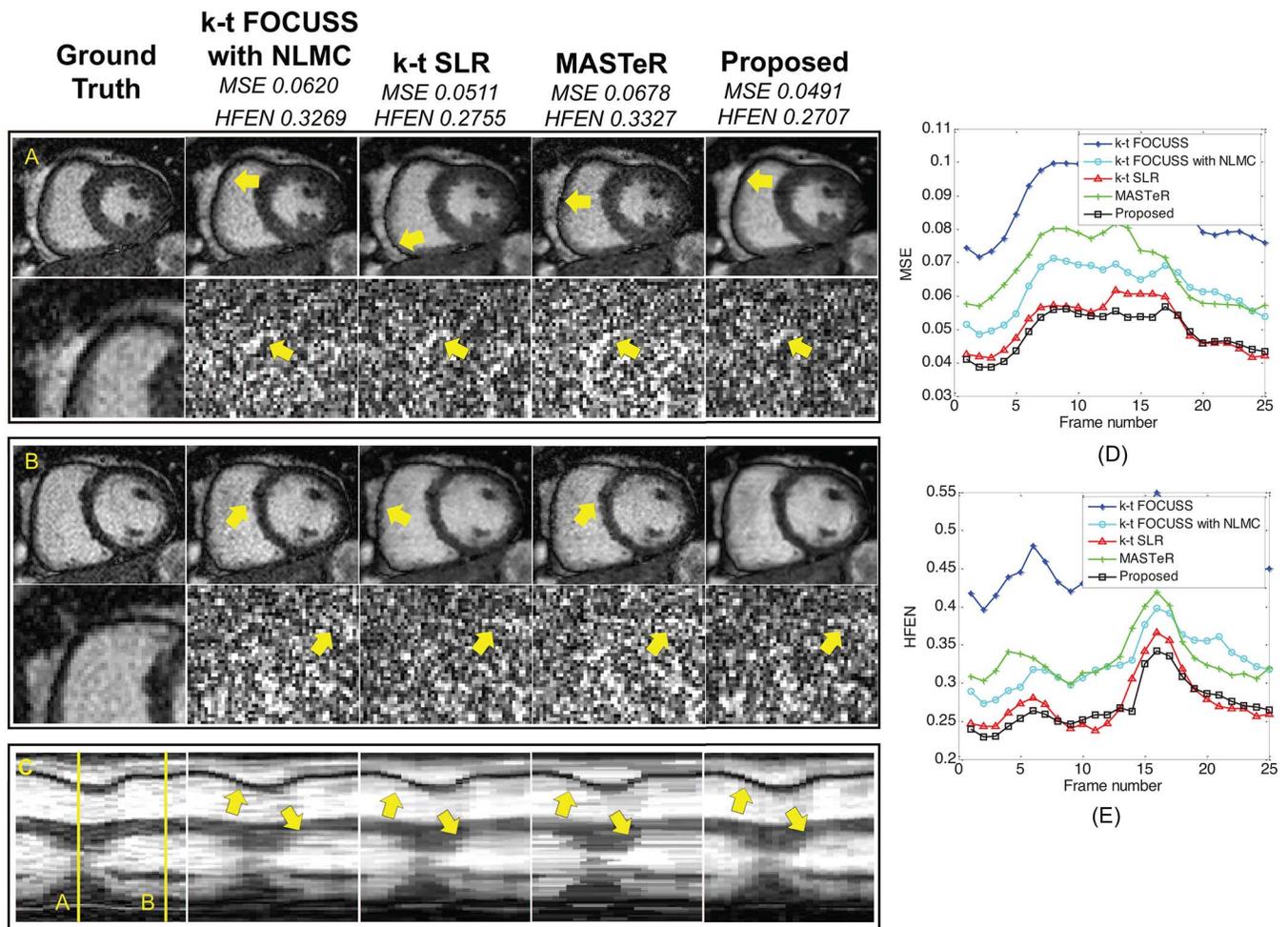


Fig. 10. Reconstructed images at $\times 8$ acceleration: Boxes (A) and (B) represent frames from the systole and diastole phases, respectively. (C) represents temporal slice profile images. Each column represents ground truth, k-t FOCUSS with NLMC, k-t SLR, MASTeR and Proposed algorithm, respectively. Images (D) and (E) represent MSE and HFEN values along time frames.

TABLE I
RECONSTRUCTION TIME FOR VARIOUS ALGORITHMS

Acquisition type	Matrix size (number of pixels)	k-t SLR (sec)	MASTeR (sec)	Proposed (sec)
Real Cartesian	$256 \times 256 \times 25$ (1638400)	428	6334	411
Real radial	$384 \times 384 \times 32$ (4718592)	4458	29047	1605

several experiments. The computational time from the various algorithms are summarized in Table I, which shows that the proposed method has the best run time.

IV. DISCUSSION

A. Initialization

To confirm the importance of initialization for the refinable similarity search scheme, different initial images were tested to assess the convergence rate. Four kinds of initial images were used: ones from k-t FOCUSS with ME/MC, k-t FOCUSS with NLMC, k-t FOCUSS only, k-t SLR, and patch-based reconstruction without ME/MC. In Fig. 12(a) and (b), we plotted MSE along iterations of phantom data and real Cartesian data. As expected, if we use a better initialization, we obtain faster

convergence. Also, the final MSE value is smaller with better initialization.

While the results in Fig. 12(a) and (b) show that k-t SLR initialization appears to provide fastest convergence, as shown in Fig. 13, the systole phase image with k-t SLR initialization still has remaining artifacts near edges, which was observed in the original k-t SLR. The image with k-t FOCUSS initialization has larger error near the cardiac boundary. Because k-t FOCUSS with NLMC initialization has fewer artifacts owing to the nonlocal motion compensation and the convergence rate is also comparable with k-t SLR initialization, we chose k-t FOCUSS with NLMC as our initialization.

Note that there are additional hyper-parameters for NLMC. For the case of NLMC, we used a fixed threshold value to determine the number of similar patches, rather than fixing it. The threshold was empirically set by observing the patch numbers and the geometry of selected patches and the same threshold was used for all patches. Note that the similar patch search is required only once for NLMC since this step is for initialization.

B. Reference Frame Generation

To use NLMC, at least one reference frame is required. One could use an image averaged over the entire acquisi-

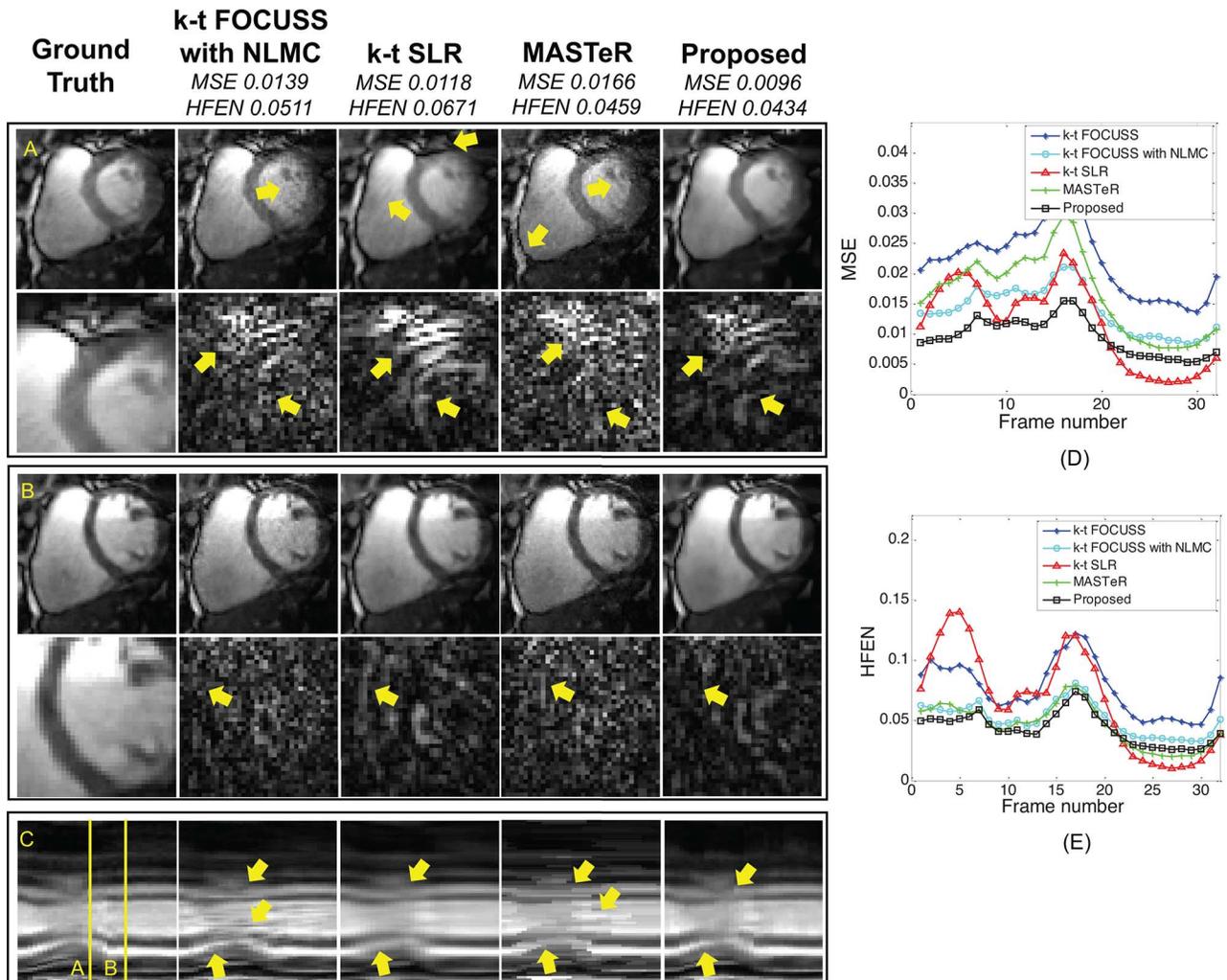


Fig. 11. Image reconstructed from radial data at $\times 8$ acceleration: Boxes (A) and (B) represent frames from the systole and diastole phases, respectively. (C) represents temporal slice profile images. Each column represents ground truth, k-t FOCUSS with NLMC, k-t SLR, MASTeR and the proposed algorithm, respectively. (D) and (E) represent MSE and HFEN values along time frames.

tion time as a reference. However, due to motion artifacts, the temporal average is often blurry, which has a negative impact on the motion compensation step. To address this, in the proposed method the reference frame was generated by merging k -space samples during diastole and applying an inverse Fourier transform, since the heart is relatively stationary during this period. This very simple procedure does not require any variable sampling rate for each cardiac phase, which simplifies the pulse sequence program. However, when designing a sequence using prospective ECG (electrocardiogram) gating, we need to estimate in advance which time points correspond to diastole since the sampling pattern needs to be switched from Gaussian random sampling to nonoverlapped uniform sampling patterns. To estimate the window for generating the reference frame, we can use a technique used in X-ray CT imaging to determine the optimal window for CT angiography [54]–[56]. More specifically, the patient's heart rate is measured and the diastole phase is estimated as a predefined percentage from the whole R-R peak. However, for patients with arrhythmia, such estimation may not be accurate, which may result in loss of image quality.

C. Proposed Formulation versus Penalized Least Squares Formulation

Instead of the proposed formulation, if we use the standard penalized least squares formulation (17), then the minimization with respect to \mathbf{x} is given by

$$(F^H F + \lambda Z) \mathbf{x}^{(k)} = F^H \mathbf{y} + \frac{\lambda}{\mu} \sum_{p=1}^P V_p^{(k)*} (W_p^{(k)}). \quad (30)$$

Compared to (24), there are two main differences in (30): the presence of the nonuniform diagonal matrix Z instead of the identity on the left-hand side, and the unweighted offset on the right-hand side of (30) from overlapped patch contribution, i.e., $\sum_{p=1}^P V_p^{(k)*} (W_p^{(k)})$, instead of the uniform W in (24).

In order to show the shortcoming of the penalized least squares formulation (30), Figs. 14(a) and (b) shows the results from the phantom data reconstructed using (30) and (24), respectively. The results indicate that the standard penalized least squares formulation is not effective in removing artifact from abrupt scene changes and is prone to boundary artifacts. We

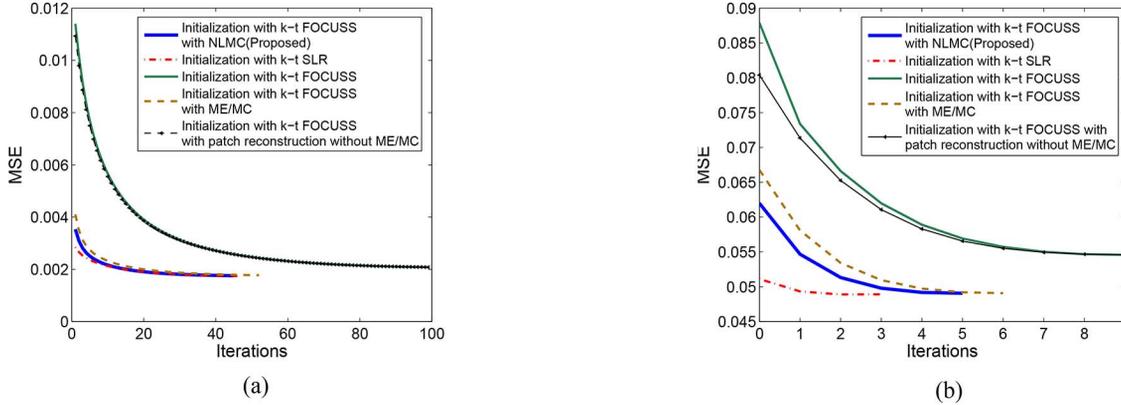


Fig. 12. Convergence plot of the proposed algorithm using various initialization. (a) Phantom data. (b) Real Cartesian data.

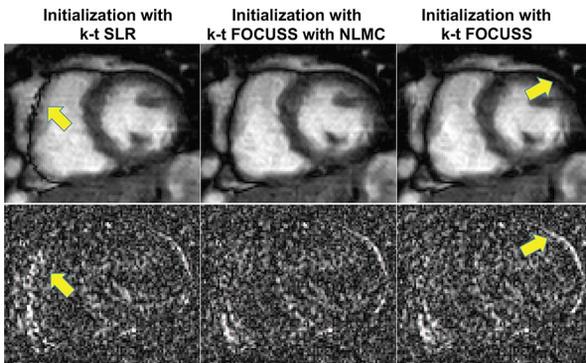


Fig. 13. Comparison of images reconstructed from different initializations. Arrows indicate the positions where the proposed algorithm shows better reconstruction quality than others, which are especially near cardiac edges.

also found that the results by (30) are very sensitive to the regularization parameter. This is because pixels in areas with abrupt changes or near the boundary have far fewer references than pixels in other areas, which reduces the effect of regularization there. In the proposed formulation, the patch contributions are weighted so that the effect of regularization is more uniform, which improves the results and reduces their sensitivity to the choice of the regularization parameter. The MSE plot in Fig. 14(c) also confirms that the proposed formulation is more effective.

D. Comparison With Related Prior Works

As discussed before, globally and locally low-rank structures have been investigated by many researchers, and the proposed algorithm deals with the locally low-rankness using a collection of similar patches. As shown in [35], a similar restriction to locally low-rank operations can potentially overcome some of the challenges faced by low-rank-promoting methods when the row and column dimensions of the Casorati matrix differ significantly. For example, as seen in Figs. 10 and 11, the global low-rank approach (k-t SLR) works as well for relatively motionless areas and in diastole. Our patch-based low-rank method is advantageous in the moving parts and in systole. This is because local processing is better at capturing local motion and fine structures than global methods.

Another advantage of the patch-based approach is robustness to noise. As seen in the SNR plot in Fig. 7(c), the pixel-based motion compensated method (MASTeR) shows degraded performance at low SNR. The patch-based approach produces better quality images than pixel-based approaches in noisy environments. The denoising effect is also verified visually. In Fig. 10, the proposed algorithm shows denoised reconstruction while preserving edges.

The generalized Huber function has been exploited in earlier works such as [24] and [43]. In [43], Hu *et al.* used a generalized Huber function to approximate the Schatten quasi-norm functional for full image series. Zhao *et al.* used a Huber function jointly for partial separability and sparsity constraints [24]. Lingala *et al.* used the Schatten quasi-norm as a surrogate for rank penalty of full image series [7]. Majumdar *et al.* combined Schatten-quasi norm with the FOCUSS algorithm to use low-rankness of difference images obtained during iterations [57]. Our algorithm is, however, general in the sense that it combines the patch-based low-rank structures of a spatiotemporal volume and provides a convergent optimization framework.

In terms of optimization method, Trzasko *et al.* [36] used a fast iterative shrinkage-thresholding algorithm (FISTA) [58] for patch-based rank penalty for calibration-free parallel imaging (CLEAR). To apply FISTA, Trzasko *et al.* applied patch-based SVD for each iteration of the gradient step. Since most of the computational complexity in our algorithm is in singular value shrinkage, the application of such a shrinkage operation for each gradient step is computationally inefficient. Therefore, in our optimization framework, we first solve for \mathbf{x} , after which singular value thresholding is applied. A similar acceleration approach has been taken in the alternating directional method of multiplier (ADMM) or the alternating augmented Lagrangian method (ALM) [59]. Note that in our algorithm formulation, each step can be understood as the alternating application of a BM3D type collaborative filtering step and l_2 penalized least squares step.

In this paper, we adopted spatiotemporal search to exploit the similarity of patches in both spatial and temporal dimensions. During the revision of this paper, we found that Chen *et al.* independently proposed a nonconvex Schatten quasi-norm for patch penalty in their recent work [38], which is similar to our approach. They used a low-rank patch penalty for respiratory

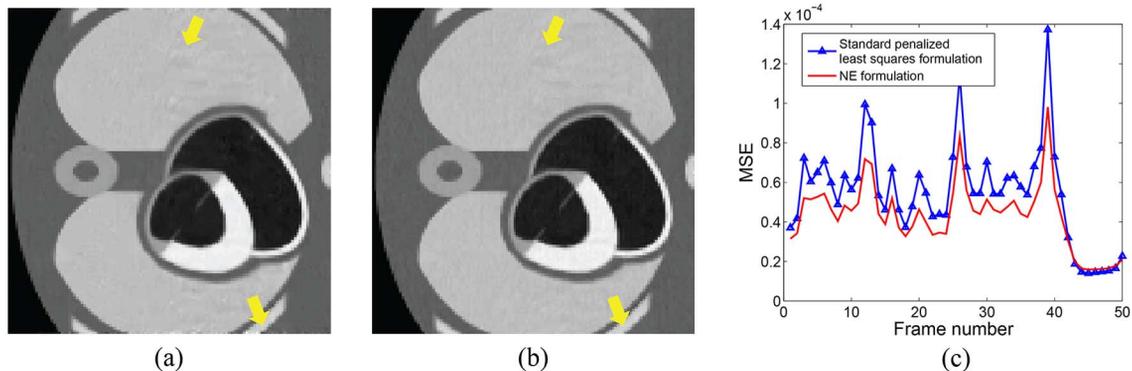


Fig. 14. Comparison of images reconstructed using (a) standard penalized least squares formulation, and (b) NE formulation. (c) Comparison of MSE plots for the reconstruction using standard penalized least squares and NE formulations.

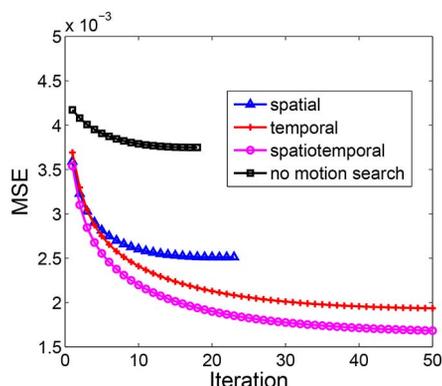


Fig. 15. Search neighborhood comparison. There are four ways of similarity searches in this experiment. 1) Spatial search in the frame itself. 2) Temporal search along motion. 3) Spatiotemporal search in a 3-D neighborhood. 4) Patches at the same position along time frames (no motion search).

motion correction with iterative soft thresholding. In their work, similar patches were searched along the motion trajectory while in the proposed algorithm the patches are searched spatiotemporally. Fig. 15 compares these different search methods using a numerical experiment with the MCAT phantom (see the simulation setup in Section III). The results in Fig. 15 show that the proposed spatiotemporal search gives the lowest MSE values.

Note that there exist three types of similarities in multi-channel spatio-temporal images: 1) adjacent and overlapping patches in an image frame of one channel, 2) patches in adjacent positions in temporally adjacent frames contiguous in time, and 3) patches in different channels due to the smoothly varying coil sensitivity. In this paper, to exploit the spatio-temporal redundancies, we focus on first two types of similarities. In particular, we propose a spatiotemporally motion adaptive low-rank patch penalty to accelerate the cardiac cine MR acquisition. Trzasko *et al.* exploited the low rankedness of multichannel images for calibration-free parallel imaging applications [36]. Combination of the spatio-temporal low rank structure in addition to multi-channel one is an important issue and will be investigated in the future.

V. CONCLUSION

In this paper, a motion-adaptive dynamic MRI method using a low-rank patch penalty was proposed. Geometric similarity within a spatio-temporal window was exploited based on the

observation that rank structures are relatively less sensitive to global intensity changes but make it easier to capture edges and fine structures. The reconstruction problem is formulated as a multi-objective optimization problem for data fidelity and low-rank penalty, which is followed by a relaxation step. We provided a rigorous proof for the convergence of the algorithm even though the similar patch groups keep updating during iteration.

Extensive experimental results show that the proposed patch-based algorithm clearly reconstructs the important cardiac structures and provides competitive image quality and other advantages compared to the existing state-of-the-art compressed sensing dynamic MRI methods.

APPENDIX

A. General Mathematical Preliminaries

Let \mathcal{H} denotes a real Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and the associated norm $\|\cdot\|$. Even though we provide a proof for real Hilbert space, the generalization to complex Hilbert space is straightforward by separating the real and imaginary part of $x \in \mathcal{H}$ and the operator T

$$x \rightarrow \begin{bmatrix} x_r \\ x_i \end{bmatrix}, \quad T \rightarrow \begin{bmatrix} T_r & -T_i \\ T_i & T_r \end{bmatrix}$$

where subscript r and i denote the real and imaginary part, respectively.

For a given mapping $T : D \rightarrow \mathcal{H}$, the set of the *fixed points* of an operator $T : D \rightarrow D$ is denoted by $\text{Fix}T = \{x \in D | Tx = x\}$. Then, we need the following standard results from convex analysis [49] (see also [60, p. 112] for the definition of α -separating operators):

Definition A.1: Let D be a nonempty subset of \mathcal{H} and let $T : D \rightarrow \mathcal{H}$. Then, T is called

1) *firmly nonexpansive* if

$$\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle, \quad \forall x, y \in D. \quad (31)$$

2) *nonexpansive* if

$$\|Tx - Ty\| \leq \|x - y\|, \quad \forall x, y \in D \quad (32)$$

3) α -*separating*, where $\alpha > 0$, if

$$\langle q - x, Tx - x \rangle \geq \alpha \|Tx - x\|^2, \quad \forall x \in \mathcal{H}, q \in \text{Fix}T. \quad (33)$$

Lemma A.1: Let D be nonempty subset of \mathcal{H} and $T : D \rightarrow \mathcal{H}$. Then, T is firmly nonexpansive if and only if $2T - I$ is expansive.

Proof: See Proposition 4.2 in [49]. \square

The following relates nonexpansive operators with α -separating operators (see [60, p. 113]).

Lemma A.2: Let $T : D \rightarrow \mathcal{H}$ have a fixed point. Then, T is α -separating with

- 1) $\alpha = 1$, if T is firmly nonexpansive;
- 2) $\alpha = 1/2$, if T is nonexpansive.

Proof:

- 1) Let T be firmly nonexpansive. Then, for any $x \in D$ and $y \in \text{Fix}T$, we have

$$\begin{aligned} \langle y - x, Tx - x \rangle - \|Tx - x\|^2 \\ = \langle y - Tx, Tx - x \rangle \end{aligned} \quad (34)$$

$$= \langle Ty - Tx, y - x - y + Tx \rangle \quad (35)$$

$$= \langle Ty - Tx, y - x \rangle - \|Ty - Tx\|^2 \geq 0 \quad (36)$$

where the last inequality comes from the firmly nonexpansiveness of T .

- 2) Let T be nonexpansive. Then, thanks to Lemma A.2, there exists firmly nonexpansive operator R such that $T = 2R - I$. Then, from the fact that R is a $\alpha = 1$ separating operator, we can show that T is a $\alpha = 1/2$ separating operator. \square

The following standard result reveals an important class of firmly nonexpansive operators [49].

Lemma A.3: Let f be a proper lower semi-continuous convex function and $\gamma \in \mathbb{R}_{++}$. Then, $\text{Prox}_{\gamma f}$ is firmly nonexpansive, where the proximal mapping is defined as

$$\text{Prox}_f x = \arg \min_{y \in \mathcal{H}} f(y) + \frac{1}{2} \|y - x\|^2. \quad (37)$$

We also need the following lemma regarding the firmly nonexpansiveness of a composite mapping.

Lemma A.4: Let $T : \mathcal{H} \rightarrow \mathcal{H}$ be firmly nonexpansive, and $L : \mathcal{X} \rightarrow \mathcal{X}$ be a bounded linear mapping. Then, $L^* \circ T \circ L$ is firmly nonexpansive.

Proof: See Corollary 4.6 in [49]. \square

B. Proof for Theorem 2.1

The proof is related to the proof in [39], but the result in [39] cannot be used for the iteration-dependent patch similarity update. Specifically, we must consider the iteration-dependent patch extraction operator, which makes the proof more involved than that in [39].

Specifically, let the singular value decomposition of the similar patch matrix $V_p(\mathbf{x})$ be

$$V_p(\mathbf{x}) = \Psi_p \Sigma_p \Phi_p^H$$

where Σ_p denotes the singular value matrix and $\{\Psi_p, \Phi_p\}$ denotes the corresponding left and right singular vector matrices, respectively. By applying a property of the Kronecker product to $\Sigma_p = \Psi_p^H V_p(\mathbf{x}) \Phi_p$, we have $\boldsymbol{\sigma}_p = (\Phi_p^H \otimes \Psi_p^H) \mathbf{v}_p(\mathbf{x})$, where

$\boldsymbol{\sigma}_p$ and $\mathbf{v}_p(\mathbf{x})$ are vectorizations of Σ_p and $V_p(\mathbf{x})$, respectively. This leads to the following analysis transform:

$$\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{\sigma}_1 \\ \vdots \\ \boldsymbol{\sigma}_P \end{bmatrix} = \begin{bmatrix} (\Phi_1^H \otimes \Psi_1^H) \mathbf{v}_1(\mathbf{x}) \\ \vdots \\ (\Phi_P^H \otimes \Psi_P^H) \mathbf{v}_P(\mathbf{x}) \end{bmatrix} = \mathcal{A} \mathbf{x} \quad (38)$$

Alternatively, for a given set of left and right singular vectors $\{\Psi_p, \Phi_p\}_{p=1}^P$ and singular values $\{\boldsymbol{\sigma}_p\}_{p=1}^P$ for similar patch groups $\{W_p\}_{p=1}^P$, an image \mathbf{x} can be synthesized as follows:

$$\mathbf{x} = Z^{-1} \sum_{p=1}^P V_p^* \circ \text{UNVEC}((\Phi_p \otimes \Psi_p) \boldsymbol{\sigma}_p) = \mathcal{S} \boldsymbol{\sigma} \quad (39)$$

where $\text{UNVEC}(\cdot)$ denote the operator to convert vector to the original patch matrix. It is clear that the analysis and synthesis formulation (38) and (39) are both linear mappings and the following properties hold:

$$\mathcal{S} \mathcal{A} = I, \mathcal{S} = Z^{-1} \mathcal{A}^* \quad (40)$$

where Z denotes the diagonal matrix whose elements correspond to the number of pixel references. Then, our formulation can be equivalently represented as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \left\{ \|\boldsymbol{\theta}\|_{g_{\mu, \nu}} + \frac{1}{2\mu} \|\boldsymbol{\theta} - \mathcal{A} \mathbf{x}\|^2 \right\} = \text{Prox}_{\mu g}(\mathcal{A} \mathbf{x}) \quad (41)$$

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\{ \|\mathbf{y} - F \mathbf{x}\|^2 + \lambda \|\mathbf{x} - \mathcal{S} \boldsymbol{\theta}\|^2 \right\} = \text{Prox}_{f/(2\lambda)}(\mathcal{S} \boldsymbol{\theta}) \quad (42)$$

where Prox_h denotes the proximal mapping of the function h and

$$f(\mathbf{x}) = \|\mathbf{y} - F \mathbf{x}\|^2, g(\boldsymbol{\theta}) = \|\boldsymbol{\theta}\|_{g_{\mu, \nu}}. \quad (43)$$

Accordingly, for given $\boldsymbol{\theta}^{(k)}$ and $\mathbf{x}^{(k)}$, the update equations from our formulation are given by

$$\boldsymbol{\theta}^{(k+1)} = \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)} \mathbf{x}^{(k)} \right) = \text{shrink}_{\nu}(\mathcal{A}^{(k)} \mathbf{x}^{(k)}, \mu) \quad (44)$$

$$\begin{aligned} \hat{\mathbf{x}}^{(k+1)} &= \text{Prox}_{f/(2\lambda)} \left(\mathcal{S}^{(k)} \boldsymbol{\theta}^{(k+1)} \right) \\ &= (F^* F + \lambda I)^{-1} \left(F^* \mathbf{y} + \lambda \mathcal{S}^{(k)} \boldsymbol{\theta}^{(k+1)} \right) \end{aligned} \quad (45)$$

where we use the explicit dependence of the analysis and synthesis operator on iteration by using the superscript $^{(k)}$. Then, the corresponding fixed point iteration for $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ is given by

$$\begin{aligned} \mathbf{x}^{(k+1)} &= \text{Prox}_{f/(2\lambda)} \left(\mathcal{S}^{(k)} \circ \text{Prox}_{\mu g}(\mathcal{A}^{(k)} \mathbf{x}^{(k)}) \right) \\ &:= T_{(k)} \mathbf{x}^{(k)}. \end{aligned} \quad (46)$$

Equipped with these, we can show that for each $k \in \mathbb{N}$ the mapping $T_{(k)}$ is a nonexpansive mapping, and therefore, by Lemma 0.2, is $\alpha = 1/2$ separating. The proof will be given in the next section.

Proposition A.5: For each $k \in \mathbb{N}$ and for $\nu = 1$, the operator $T_{(k)}$ in (46) is nonexpansive.

One of the remaining technical difficulties in proving the convergence in our problem is that our operator $T_{(k)}$ itself varies along iteration. Therefore, the standard convergence results for iteration *independent* T operator must be modified using the following result:

Lemma A.6: Let $\{T_{(k)}\}_{k \in \mathbb{N}}$ be sequence of $\alpha = 1/2$ separating operators that depends on each iteration. Consider the following iteration:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \beta(T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)}), \beta \in (0, 1). \quad (47)$$

Suppose that $\bigcap_{k=k_0}^{\infty} \text{Fix}T_{(k)} \neq \emptyset$. Then

$$\lim_{k \rightarrow \infty} \left\| T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right\| = 0 \text{ and } \{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \text{ converges.} \quad (48)$$

Proof: See the following section. \square

Now, we are ready to show the proof of Theorem 2.1.

(Proof of Theorem 2.1): The proof can be easily obtained by combining Proposition A.5 and Lemma A.6, which show that the proposed update strategy (47) using the operator $T_{(k)}$ defined in (46) gives a convergent sequence. This concludes the proof. \square

C. Proof of Subcomponents of Part B

Proof of Proposition A.5: Let $\mathcal{W} := (F^*F/\lambda + I)^{-1}$. Then, it is easy to see that $\|\mathcal{W}\| \leq 1$ for $\lambda > 0$. Now, for given two distinct sequences $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ and $\{\mathbf{x}'^{(k)}\}_{k \in \mathbb{N}}$, using the expressions (44) and (45), we have

$$\begin{aligned} \left\| T_{(k)}\mathbf{x}^{(k)} - T_{(k)}\mathbf{x}'^{(k)} \right\| &= \left\| \mathbf{x}^{(k+1)} - \mathbf{x}'^{(k+1)} \right\| \\ &= \left\| \mathcal{W}\mathcal{S}^{(k)}\boldsymbol{\theta}^{(k+1)} - \mathcal{W}\mathcal{S}^{(k)}\boldsymbol{\theta}'^{(k+1)} \right\| \\ &\leq \left\| \mathcal{S}^{(k)}\boldsymbol{\theta}^{(k+1)} - \mathcal{S}^{(k)}\boldsymbol{\theta}'^{(k+1)} \right\| \\ &= \left\| \mathcal{S}^{(k)} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}^{(k)} \right) - \mathcal{S}^{(k)} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}'^{(k)} \right) \right\| \end{aligned}$$

where the first inequality comes from $\|\mathcal{W}\| \leq 1$.

Now, note that $\text{Prox}_{\mu g}$ is firmly nonexpansive (thus, nonexpansive). Using Lemma A.4, we know that $\mathcal{A}^{(k)*} \circ \text{Prox}_{\mu g} \circ \mathcal{A}^{(k)}$ is firmly nonexpansive. Thus

$$\begin{aligned} \left\| \mathcal{A}^{(k)*} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}^{(k)} \right) - \mathcal{A}^{(k)*} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}'^{(k)} \right) \right\| \\ \leq \left\| \mathbf{x}^{(k)} - \mathbf{x}'^{(k)} \right\|. \end{aligned} \quad (49) \quad (50)$$

Furthermore, $\mathcal{S}^{(k)} = Z^{-(k)}\mathcal{A}^{(k)*}$ and $\|Z^{-(k)}\| \leq 1$. Hence, we have

$$\begin{aligned} \left\| T_{(k)}\mathbf{x}^{(k)} - T_{(k)}\mathbf{x}'^{(k)} \right\| \\ \leq \left\| \mathcal{S}^{(k)} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}^{(k)} \right) - \mathcal{S}^{(k)} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}'^{(k)} \right) \right\| \\ \leq \left\| \mathcal{A}^{(k)*} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}^{(k)} \right) - \mathcal{A}^{(k)*} \circ \text{Prox}_{\mu g} \left(\mathcal{A}^{(k)}\mathbf{x}'^{(k)} \right) \right\| \\ \leq \left\| \mathbf{x}^{(k)} - \mathbf{x}'^{(k)} \right\|. \end{aligned}$$

Therefore, $T_{(k)}$ is nonexpansive. \square

Proof of Lemma A.6: Let $\mathbf{z} \in \bigcap_{k=k_0}^{\infty} \text{Fix}T_{(k)}$. Then, for all $k \geq k_0$, we have

$$\left\| \mathbf{x}^{(k+1)} - \mathbf{z} \right\|^2 = \left\| \mathbf{x}^{(k)} + \beta \left(T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right) - \mathbf{z} \right\|^2 \quad (51)$$

$$\begin{aligned} &= \left\| \mathbf{x}^{(k)} - \mathbf{z} \right\|^2 + \beta^2 \left\| T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right\|^2 \\ &\quad - 2\beta \left\langle \mathbf{z} - \mathbf{x}^{(k)}, T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right\rangle \end{aligned} \quad (52)$$

$$\leq \left\| \mathbf{x}^{(k)} - \mathbf{z} \right\|^2 - \beta(1 - \beta) \left\| T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right\|^2 \quad (53)$$

where we use the property (33) for $\alpha = 1/2$ to arrive at the last inequality. Now, by applying this multiple times, we have

$$\begin{aligned} \left\| \mathbf{x}^{(k+1)} - \mathbf{z} \right\|^2 &\leq \left\| \mathbf{x}^{(k_0)} - \mathbf{z} \right\|^2 \\ &\quad - \beta(1 - \beta) \sum_{l=k_0}^k \left\| T_{(l)}\mathbf{x}^{(l)} - \mathbf{x}^{(l)} \right\|^2. \end{aligned}$$

Since $\|\mathbf{x}^{(k+1)} - \mathbf{z}\|^2 \geq 0$, we have $\sum_{k=k_0}^{\infty} \|T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)}\|^2 < \|\mathbf{x}^{(k_0)} - \mathbf{z}\|^2 / \beta(1 - \beta) < \infty$, and by adding the first k_0 terms, we have $\sum_{k=0}^{\infty} \|T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)}\|^2 < \infty$. Since the sum converges, the sequence of partial sums is Cauchy; hence, for every $\epsilon > 0$ there exists a number N such that for all $m > n > N$, $\sum_{k=n}^{m-1} \|T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)}\| < \epsilon$. Using the triangular inequality, we have

$$\begin{aligned} \left\| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \right\| &\leq \sum_{k=n}^{m-1} \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\| \\ &= \beta \sum_{k=n}^{m-1} \left\| T_{(k)}\mathbf{x}^{(k)} - \mathbf{x}^{(k)} \right\| \leq \beta\epsilon. \end{aligned} \quad (54)$$

Thus, $\{\mathbf{x}^{(k+1)}\}_{k \in \mathbb{N}}$ is a Cauchy sequence, therefore, it converges. \square

REFERENCES

- [1] M. Lustig, J. Santos, D. Donoho, and J. Pauly, "kt SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proc. 13th Annu. Meet. ISMRM*, Seattle, WA, 2006, p. 2420.
- [2] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [3] H. Jung, J. Ye, and E. Kim, "Improved k-t BLAST and k-t SENSE using FOCUS," *Phys. Med. Biol.*, vol. 52, no. 11, p. 3201, 2007.
- [4] H. Jung, K. Sung, K. Nayak, E. Kim, and J. Ye, "k-t FOCUS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, 2009.
- [5] H. Jung, J. Park, J. Yoo, and J. Ye, "Radial k-t FOCUS for high-resolution cardiac cine MRI," *Magn. Reson. Med.*, vol. 63, no. 1, pp. 68–78, 2010.
- [6] L. Feng, R. Otazo, H. Jung, J. H. Jensen, J. C. Ye, D. K. Sodickson, and D. Kim, "Accelerated cardiac T2 mapping using breath-hold multiecho fast spin-echo pulse sequence with k-t FOCUS," *Magn. Reson. Med.*, vol. 65, no. 6, pp. 1661–1669, 2011.
- [7] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: kt SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.
- [8] H. Jung and J. Ye, "Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques," *Int. J. Imag. Syst. Technol.*, vol. 20, no. 2, pp. 81–98, 2010.
- [9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [10] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

- [11] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: Re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.
- [12] J. Tsao, P. Boesiger, and K. P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1031–1042, Oct. 2003.
- [13] M. Usman, C. Prieto, T. Schaeffter, and P. Batchelor, "k-t group sparse: A method for accelerating dynamic MRI," *Magn. Reson. Med.*, vol. 66, no. 4, pp. 1163–1176, 2011.
- [14] R. Otazo, D. Kim, L. Axel, and D. K. Sodickson, "Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI," *Magn. Reson. Med.*, vol. 64, no. 3, pp. 767–776, 2010.
- [15] M. S. Asif, L. Hamilton, M. Brummer, and J. Romberg, "Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI," *Magn. Reson. Med.*, vol. 7, pp. 800–812, 2012.
- [16] Z.-P. Liang, H. Jiang, C. P. Hess, and P. C. Lauterbur, "Dynamic imaging by model estimation," *Int. J. Imag. Syst. Technol.*, vol. 8, no. 6, pp. 551–557, 1997.
- [17] A. S. Gupta and Z. Liang, "Dynamic imaging by temporal modeling with principal component analysis," in *Proc. 9th Annu. Meet. ISMRM*, 2001, p. 10.
- [18] J. Ji and Z.-P. Liang, "High resolution cardiac magnetic resonance imaging: A model-based approach," in *Proc. 23rd Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2001, vol. 3, pp. 2268–2271, IEEE.
- [19] D. Xu, L. Ying, and Z.-P. Liang, "Dynamic MRI using spatiotemporal modeling with phased array coils," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2006, pp. 722–725.
- [20] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in *Proc. Joint Meet. 6th Int. Symp. Noninvasive Funct. Source Imag. Brain Heart Int. Conf. Funct. Biomed. Imag.*, 2007, pp. 181–182.
- [21] C. Brinegar, Y.-J. L. Wu, L. M. Foley, T. K. Hitchens, Q. Ye, C. Ho, and Z.-P. Liang, "Real-time cardiac MRI without triggering, gating, or breath holding," in *Proc. 30th Annu. IEEE Int. Conf. EMBS*, 2008, pp. 3381–3384.
- [22] B. Zhao, J. P. Haldar, C. Brinegar, and Z.-P. Liang, "Low rank matrix recovery for real-time cardiac MRI," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010, pp. 996–999.
- [23] J. P. Haldar and Z.-P. Liang, "Spatiotemporal imaging with partially separable functions: A matrix recovery approach," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010, pp. 716–719.
- [24] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z.-P. Liang, "Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1809–1820, Sep. 2012.
- [25] G. Adluru, L. Chen, D. Feinberg, J. Anderson, and E. V. R. DiBella, "Improving rank constrained reconstructions using prior information with reordering debella," in *Proc. 21th Annu. Meet. ISMRM*, Salt Lake City, UT, 2013, p. 2248.
- [26] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 2, pp. 60–65.
- [27] A. Buades, B. Coll, and J. M. Morel, "Image denoising methods. A new nonlocal principle," *SIAM Rev.*, vol. 52, no. 1, pp. 113–147, 2010.
- [28] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [29] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [30] S. Ravishanker and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, May 2011.
- [31] Z. Yang and M. Jacob, "A unified energy minimization framework for nonlocal regularization," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2011, pp. 1150–1153.
- [32] Z. Yang and M. Jacob, "Robust non-local regularization framework for motion compensated dynamic imaging without explicit motion estimation," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Barcelona, Spain, 2012, pp. 1056–1059.
- [33] Z. Yang and M. Jacob, "Nonlocal regularization of inverse problems: A unified variational framework," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3192–3203, Aug. 2012.
- [34] G. Wang and J. Qi, "Penalized likelihood pet image reconstruction using patch-based edge-preserving regularization," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2194–2204, 2012.
- [35] J. Trzasko, A. Manduca, and E. Borisch, "Local versus global low-rank promotion in dynamic MRI series reconstruction," in *Proc. Annu. Meet. ISMRM*, Salt Lake City, UT, 2011, p. 4371.
- [36] J. Trzasko and A. Manduca, "Calibrationless parallel MRI using CLEAR," in *Proc. IEEE Conf. Rec. 45th Asilomar Conf. Signals, Syst. Comput.*, 2011, pp. 75–79.
- [37] M. Akçakaya, T. Basha, B. Goddu, L. Goepfert, K. Kissinger, V. Tarokh, W. Manning, and R. Nezafat, "Low-dimensional-structure self-learning and thresholding (LOST) : Regularization beyond compressed sensing for MRI reconstruction," *Magn. Reson. Med.*, vol. 66, no. 3, pp. 756–767, 2011.
- [38] X. Chen, M. Salerno, C. H. Meyer, and F. H. Epstein, "Block low-rank sparsity with motion guidance (BLOSM) for accelerated dynamic MRI," in *Proc. 21th Annu. Meet. ISMRM*, Salt Lake City, UT, 2013, p. 4555.
- [39] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, Apr. 2012.
- [40] K. Leyton-Brown and Y. Shoham, "Essentials of game theory: A concise multidisciplinary introduction," *Synthesis Lectures Artif. Intell. Mach. Learn.*, vol. 2, no. 1, pp. 1–88, 2008.
- [41] H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. New York: Springer, 2011, vol. 49.
- [42] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundat. Computat. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
- [43] Y. Hu, S. G. Lingala, and M. Jacob, "A fast majorize-minimize algorithm for the recovery of sparse and low-rank matrices," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 742–753, Feb. 2012.
- [44] R. Chartrand, "Nonconvex splitting for regularized low-rank+ sparse decomposition," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5810–5819, Nov. 2012.
- [45] Z. Xu, X. Chang, F. Xu, and H. Zhang, " $l_{1/2}$ -regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.
- [46] K. Mohan and M. Fazel, "Iterative reweighted least squares for matrix rank minimization," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput.*, 2010, pp. 653–661.
- [47] K. Miettinen, *Nonlinear Multiobjective Optimization*. New York: Springer, 1999, vol. 12.
- [48] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, p. 035020, 2008.
- [49] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer, 2011.
- [50] H. Yoon and J. C. Ye, "MMSE optimal non-local motion compensation for compressed sensing cardiac cine imaging using k-t FOCUSS," in *Proc. 20th Annu. Meet. ISMRM*, 2012, p. 78.
- [51] W. P. Segars, D. S. Lalush, and B. M. Tsui, "A realistic spline-based dynamic heart phantom," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 3, pp. 503–506, Jun. 1999.
- [52] W. P. Segars, D. S. Lalush, and B. M. Tsui, "Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 1, pp. 89–97, Feb. 2001.
- [53] S. Lingala and M. Jacob, "Blind compressive sensing dynamic MRI," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1132–1145, Jun. 2013.
- [54] A. Kopp, S. Schroeder, A. Kuettner, M. Heuschmid, C. Georg, B. Ohnesorge, R. Kuzo, and C. Claussen, "Coronary arteries: Retrospectively ECG-gated multi-detector row CT angiography with selective optimization of the image reconstruction window," *Radiology*, vol. 221, no. 3, pp. 683–688, 2001.
- [55] B. Ohnesorge, T. Flohr, C. Becker, A. Kopp, U. Schoepf, U. Baum, A. Knez, K. Klingenberg-Regn, and M. Reiser, "Cardiac imaging by means of electrocardiographically gated multisection spiral CT: Initial experience," *Radiology*, vol. 217, no. 2, pp. 564–571, 2000.
- [56] M. Vembar, M. Garcia, D. Heuschler, R. Haberl, D. Matthews, G. Böhme, and N. Greenberg, "A dynamic approach to identifying desired physiological phases for cardiac imaging using multislice spiral CT," *Med. Phys.*, vol. 30, p. 1683, 2003.
- [57] A. Majumdar, "FOCUSS based Schatten-p norm minimization for real-time reconstruction of dynamic contrast enhanced MRI," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 315–318, May 2012.
- [58] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2009, pp. 693–696.
- [59] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.
- [60] *Nonlinear Analysis and Optimization: A Conference in Celebration of Alex Ioffe's 70th and Simeon Reich's 60th Birthdays*, A. Leizarowitz, Ed. Providence, RI: AMS, 2008.