Published in final edited form as:
IEEE Trans Med Imaging. 2012 September ; 31(9): 1809–1820. doi:10.1109/TMI.2012.2203921.

Image Reconstruction from Highly Undersampled (k, *t*)-Space Data with Joint Partial Separability and Sparsity Constraints

Bo Zhao [Student Member, IEEE],
Department of Electrical and Computer Engineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Justin P. Haldar [Member, IEEE],
Ming Hsieh Department of Electrical Engineering, Signal and Image Processing Institute, Dana and David Dornsife Cognitive Neuroscience Imaging Center, and Brain and Creativity Institute, University of Southern California, Los Angeles, CA, USA

Anthony G. Christodoulou [Student Member, IEEE], and
Department of Electrical and Computer Engineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Zhi-Pei Liang [Fellow, IEEE]
Department of Electrical and Computer Engineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Bo Zhao: bozhao1@illinois.edu; Zhi-Pei Liang: z-liang@illinois.edu

Abstract

Partial separability (PS) and sparsity have been previously used to enable reconstruction of dynamic images from undersampled (k, *t*)-space data. This paper presents a new method to use PS and sparsity constraints jointly for enhanced performance in this context. The proposed method combines the complementary advantages of PS and sparsity constraints using a unified formulation, achieving significantly better reconstruction performance than using either of these constraints individually. A globally convergent computational algorithm is described to efficiently solve the underlying optimization problem. Reconstruction results from simulated and *in vivo* cardiac MRI data are also shown to illustrate the performance of the proposed method.

Index Terms

constrained reconstruction; partial separability modeling; low-rank matrices; sparsity; dynamic imaging; real-time cardiac MRI

I. Introduction

This paper addresses an image reconstruction problem that often arises in dynamic imaging, such as cardiac imaging [1], dynamic contrast-enhanced cancer imaging [2], time-resolved angiographic imaging [3], and functional neuroimaging [4]. We focus on dynamic Fourier imaging, in which the measured data can be modeled as

$$s(\mathbf{k}, t)=\int\rho(\mathbf{x}, t)\exp(-i2\pi\mathbf{k}\cdot\mathbf{x})d\mathbf{x}+\eta(\mathbf{k}, t), \quad (1)$$

where $\rho(\mathbf{x}, t)$ is the desired image function and $\eta(\mathbf{k}, t)$ is the measurement noise. Conventional imaging methods often assume $\rho(\mathbf{x}, t)$ to be support-limited in $(\mathbf{x}, t)$-space and reconstruct $\rho(\mathbf{x}, t)$ from samples of $s(\mathbf{k}, t)$ acquired at the Nyquist rate. This paper addresses the problem of reconstructing $\rho(\mathbf{x}, t)$ from highly undersampled data, which frequently arises in accelerated dynamic imaging. Although conventional reconstruction methods cannot handle sub-Nyquist data, it was demonstrated very early that high-quality images can be reconstructed from undersampled data with appropriate constraints [5], [6]. Over the last few years, significant advances have been made in theory and algorithms for solving the undersampled data reconstruction problem, resulting in a large number of new methods (e.g., [6]–[34] and references therein). These methods can be roughly grouped into two classes: (a) sparsity-constrained methods, including the popular compressed sensing methods and its many variants [7]–[15], and (b) methods utilizing spatiotemporal correlation in various forms, such as limited spatiotemporal support [16]–[21] and spatiotemporal partial separability (PS) [6], [22]–[34].

Extending our early conference papers [28], [35], this paper proposes a new method to use PS and sparsity jointly for image reconstruction from highly undersampled data. PS and sparsity constraints each have different strengths and limitations for dynamic image reconstruction. The proposed method combines the complementary advantages of both constraints by using spatial-spectral sparsity to regularize PS model-based reconstruction, achieving much better performance than using either of these constraints individually.

Several other methods have also been proposed recently to use joint PS and sparsity constraints for dynamic imaging. For example, Gao et al. proposed a robust principal component analysis-based model for dynamic computed tomography [31], which represents dynamic images as a linear combination of a PS component and a sparse component. This image model is distinct from our proposed formulation, in which the dynamic image is assumed to have both PS and sparsity characteristics simultaneously. $\mathbf{k}$-t SLR [32], [33] is similar to our proposed method except that it enforces the PS constraints implicitly through the nuclear norm or Schatten p-norm. We will demonstrate later in the paper that using the PS constraints explicitly in the proposed method has both imaging and computational advantages.

For convenience, we summarize here the key notation and symbols used throughout the paper. We use $\mathbb{C}$ to denote the field of complex numbers. We denote vectors by bold lower-case letters (e.g., $\mathbf{d}$), and matrices by bold uppercase letters (e.g., $\mathbf{X}$). $\mathbf{x}_i$ and $\mathbf{X}_{i,j}$ denote the ith and (i, j)th entry of $\mathbf{x}$ and $\mathbf{X}$, respectively. $\mathbf{X}^H$ is the Hermitian transpose of $\mathbf{X}$, and $\mathcal{A}^*$ is the adjoint operator of a linear operator $\mathcal{A}$. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. $\mathcal{I}$ represents the identity operator. vec($\mathbf{X}$) denotes the vector constructed by concatenating the columns of $\mathbf{X}$. We use the following vector norms (or quasinorms): a) ℓ_0 quasi-norm $\|\mathbf{d}\|_0 \triangleq \Sigma_i \mathbf{1}(\mathbf{d}_i \neq = 0)$, where $\mathbf{1}(\cdot)$ is the indicator function, b) ℓ_1 norm $\|\mathbf{d}\|_1 \triangleq \Sigma_i |\mathbf{d}_i|$, and c) ℓ_2 norm $\|\mathbf{d}\|_2 \triangleq \sqrt{\mathbf{d}^H\mathbf{d}}$. Denoting the ith singular value of $\mathbf{X}$ as σ_i, we also use the following matrix norms: a) Frobenius norm $\|\mathbf{X}\|_{\mathrm{F}} \triangleq (\sum_i \sigma_i^2)^{1/2}$, b) nuclear norm $\|\mathbf{X}\|_* \triangleq \Sigma_i \sigma_i$, and c) Schatten p-norm $\|\mathbf{X}\|_p \triangleq (\sum_i \sigma_i^p)^{1/p}$.

II. The proposed method

A. Formulation

We use a discrete image model in which $\rho(\mathbf{x}, t)$ is represented by its samples on a grid of N spatial locations $\{\mathbf{x}_n\}_{n=1}^{N}$ and M time instants $\{t_m\}_{m=1}^{M}$. Let

$$\mathbf{C}=\begin{bmatrix} \rho(\mathbf{x}_1,t_1) & \dots & \rho(\mathbf{x}_1,t_M) \\ \vdots & \ddots & \vdots \\ \rho(\mathbf{x}_N,t_1) & \dots & \rho(\mathbf{x}_N,t_M) \end{bmatrix} \in \mathbb{C}^{N\times M} \tag{2}$$

(the Casorati matrix [24]) and $\boldsymbol{\rho} = \mathrm{vec}(\mathbf{C}) \in \mathbb{C}^{NM\times 1}$ be the matrix and vector representations of $\rho(\mathbf{x}, t)$. We further assume that D samples are acquired from $s(\mathbf{k}, t)$. Then, (1) can be written as

$$\mathbf{d}=\boldsymbol{\Phi}\boldsymbol{\rho}+\boldsymbol{\eta}, \tag{3}$$

where $\mathbf{d} \in \mathbb{C}^{D\times 1}$ contains the measured data, $\boldsymbol{\Phi} \in \mathbb{C}^{D\times NM}$ is the measurement matrix that models the operations of the spatial Fourier transform and $(\mathbf{k}, t)$-space sparse sampling, and $\boldsymbol{\eta} \in \mathbb{C}^{D\times 1}$ is the noise vector.

In accelerated dynamic imaging, $D \ll NM$ and (3) is highly underdetermined. Here we propose a new method to solve this underdetermined system of equations, by imposing the PS and sparsity constraints jointly using the following formulation:

$$\widehat{\boldsymbol{\rho}}=\arg\min_{\boldsymbol{\rho}\in\mathbb{C}^{NM\times 1}} \left\|\mathbf{d}-\boldsymbol{\Phi}\boldsymbol{\rho}\right\|_2^2+R_{\mathrm{r}}(\boldsymbol{\rho})+R_{\mathrm{s}}(\boldsymbol{\rho}). \tag{4}$$

In this formulation (called "PS-Sparse"), the penalty functions $R_{\mathrm{r}}(\cdot)$ and $R_{\mathrm{s}}(\cdot)$ are used to incorporate the PS and sparsity constraints respectively, as explained below.

A1. PS constraints—The PS constraint assumes that $\rho(\mathbf{x}, t)$ is spatiotemporally partially separable in the following sense [24]:

$$\rho(\mathbf{x},t)=\sum_{\ell=1}^{L} u_\ell(\mathbf{x})v_\ell(t), \tag{5}$$

where L is the order of the PS model, and $\{u_\ell(\mathbf{x})\}_{\ell=1}^{L}$ and $\{v_\ell(t)\}_{\ell=1}^{L}$ are sets of spatial and temporal functions, respectively. The PS model was proposed to capture spatiotemporal correlation often observed in dynamic image sequences [24]. It is easy to show that with (5), any set of functions of the form $\{\rho(\mathbf{x}_\ell, t)\}_{\ell=1}^{\widehat{L}}$ are linearly dependent if $\hat{L} > L$. The PS constraint defined in (5) also implies that the rank of $\mathbf{C}$ is upper bounded by L [24]. As a consequence, $\mathbf{C}$ has at most $2L(M + N - L)$ real degrees of freedom [27], which is often much smaller than the total number of elements in $\mathbf{C}$. Selection of L for (5) often needs to balance the representation capability of the model and the numerical condition of the resulting model fitting problem. When L is too low, the model may fail to capture some temporal features, although the corresponding model fitting problem is often well-conditioned. When L is too high (with respect to the number of measurements available), the model fitting problem is often ill-conditioned, which can amplify modeling errors and

measurement noise. A key contribution of this paper is a new algorithm that enables high-order PS model to be used without the concomitant ill-conditioning problem.

The PS constraint can be enforced in several ways. One way is to enforce it "implicitly" using the nuclear norm or Schatten p-norm ($p < 1$). In this case, we have $R_r(\boldsymbol{\rho}) = \lambda_r \|\mathbf{C}\|_*$ or $R_r(\boldsymbol{\rho}) = \lambda_r \|\mathbf{C}\|_p^p$, which are respectively convex and non-convex surrogates for rank($\mathbf{C}$). These implicit PS constraints have been investigated for dynamic imaging in [32], [33].

Another way is to enforce the PS constraint explicitly with a known L [24], [29], by using

$$R_r(\boldsymbol{\rho}) = \begin{cases} 0, & \text{if rank}(\mathbf{C}) \le L, \\ \infty, & \text{else.} \end{cases} \tag{6}$$

Using (6) is equivalent to forcing $\mathbf{C}$ to have the following decomposition,[1]

$$\mathbf{C} = \mathbf{U}_s \mathbf{V}_t, \tag{7}$$

where $\mathbf{U}_s \in \mathbb{C}^{N \times L}$ represents a basis for the spatial (or column) subspace of $\mathbf{C}$, and $\mathbf{V}_t \in \mathbb{C}^{L \times M}$ contains a basis for the temporal (or row) subspace of $\mathbf{C}$.

An even stronger way to enforce the explicit PS constraints is to assume that $\mathbf{V}_t$ in (7) is known, or can be accurately estimated prior to full image reconstruction. This assumption is generally valid when specialized data acquisition schemes are used to sample ($\mathbf{k}$, t)-space [8], [22]–[27]. For example, if $s(\mathbf{k}, t)$ is fully sampled at the temporal Nyquist rate for a number of $\mathbf{k}$-space locations, then the dominant (rank-L) temporal subspace of these measurements can be extracted using singular value decomposition (SVD). Assuming that this subspace is representative of the dominant subspace of $\mathbf{C}$, $\mathbf{V}_t$ can be obtained from the L dominant right singular vectors from the above SVD. This approach was originally described in [22]–[24], and similar approaches have also recently been proposed in [8], [25].[2]

In this paper, we focus on imposing the explicit PS constraints with a known $\mathbf{V}_t$ (or estimated as described above). We further assume, without loss of generality, that $\mathbf{V}_t$ has L orthonormal rows. It will be demonstrated later in the paper that this approach has advantages over using implicit PS constraints.

A2. Sparsity constraints—Dynamic image sequences often have spatiotemporal features that lead to approximately sparse representations [8], [9]. Here we assume that $\Psi\,\boldsymbol{\rho}$ is sparse (or has a small ℓ_0 norm) under a certain sparsifying transform Ψ. Generally, the choice of Ψ is application-dependent. In this paper, we use the temporal Fourier transform as an example to derive our proposed algorithm. Directly enforcing the sparsity constraints through the ℓ_0 norm is not desirable, although greedy algorithms for ℓ_0 minimization exist [36]. An important development in compressed sensing is the establishment of a theoretical foundation for the use of surrogate cost functions that are more tractable to optimize. Specifically, the convex ℓ_1 norm is much easier to optimize and has proven optimality for sparsity-constrained inverse problems under certain conditions [37], [38], although

[1]Note that $\mathbf{U}_s$ and $\mathbf{V}_t$ in (7) are not unique, but the columns of $\mathbf{U}_s$ span the same column subspace and rows of $\mathbf{V}_t$ span the same row subspace.

[2]The method in [8] uses a full-rank $\mathbf{V}_t$ (i.e., rank($\mathbf{V}_t$) = M) in contrast to a low-rank $\mathbf{V}_t$ here. With a full-rank $\mathbf{V}_t$, the rank constraint is not imposed and $\mathbf{V}_t$ serves as a spatiotemporal transform in [8].

noncovex surrogate functions, such as the ℓ_q quasinorm ($0 < q < 1$) [39], can also be used. In this paper, we use the ℓ_1 norm to impose the sparsity constraint and the corresponding penalty function is $R_s(\boldsymbol{\rho}) = \lambda_s \|\Psi\boldsymbol{\rho}\|_1$. This constraint and its variants have been widely used for dynamic imaging in previous work (e.g., [8]).

A3. Reconstruction with joint PS and sparsity constraints—Under the assumptions described above, the solution to (4) is given by $\hat{\mathbf{C}} = \hat{\mathbf{U}}_s\mathbf{V}_t$, where $\hat{\mathbf{U}}_s$ is given by

$$\widehat{\mathbf{U}}_s = \arg\min_{\mathbf{U}_s \in \mathbb{C}^{N\times L}} \left\|\mathbf{d} - \Omega(\mathbf{F}_s\mathbf{U}_s\mathbf{V}_t)\right\|_2^2 + \lambda\left\|\mathrm{vec}(\mathbf{U}_s\mathbf{V}_f)\right\|_1, \tag{8}$$

where $\Omega : \mathbb{C}^{N\times M} \rightarrow \mathbb{C}^{D\times 1}$ represents the ($\mathbf{k}$, t)-space sampling operator, $\mathbf{F}_s \in \mathbb{C}^{N\times N}$ is the spatial Fourier matrix, $\mathbf{V}_f = \mathbf{V}_t\mathbf{F}_t$, and $\mathbf{F}_t \in \mathbb{C}^{M\times M}$ is an orthonormal temporal Fourier matrix. The solution to (8) reduces to the basic PS-constrained reconstruction [24] (referred to as Basic-PS) when $\lambda = 0$, i.e.,

$$\widehat{\mathbf{U}}_s = \arg\min_{\mathbf{U}_s \in \mathbb{C}^{N\times L}} \left\|\mathbf{d} - \Omega(\mathbf{F}_s\mathbf{U}_s\mathbf{V}_t)\right\|_2^2, \tag{9}$$

and to a basic ($\mathbf{x}$, f)-domain sparsity constrained reconstruction [8] (referred to as Basic-Sparse) if $L = M$ and $\mathrm{rank}(\mathbf{V}_t) = M$, i.e.,

$$\widehat{\mathbf{C}} = \arg\min_{\mathbf{C} \in \mathbb{C}^{N\times M}} \left\|\mathbf{d} - \Omega(\mathbf{F}_s\mathbf{C})\right\|_2^2 + \lambda\left\|\mathrm{vec}(\mathbf{C}\mathbf{F}_t)\right\|_1. \tag{10}$$

The proposed method incorporates both the PS and sparsity constraints in a single formulation, providing some desirable advantages over (9) and (10) summarized here (to be demonstrated in Section III): a) In Basic-PS, the model order L cannot be high because it is limited by the number of measurements available. In principle, for the solution of the Lth order Basic-PS to be well-defined, each $\mathbf{k}$-space location has to have at least L temporal samples. In practice, many more measurements are often needed to avoid an ill-conditioned model fitting problem. The proposed method overcomes the limitations with the sparsity constraint serving as an effective regularizer. b) A major limitation of Basic-Sparse is that spatiotemporal blurring often appears when ($\mathbf{k}$, t)-space is highly undersampled. This is because many different sparse solutions can closely match the highly undersampled ($\mathbf{k}$, t)-space data. By imposing a PS constraint with $L \ll M$, the proposed method effectively exploits spatiotemporal correlations in the data to significantly reduce the set of candidate sparse solutions. This frequently yields better reconstruction results, removing blurring in the sparsity-constrained reconstructions.

B. Algorithm

In this section, we present a globally convergent, efficient solution algorithm to solve (8). The cost function of (8) is convex and coercive over $\mathbb{C}^{N\times L}$; therefore, there exists at least one optimal solution [40]. In most cases of interest, strict convexity also holds, which guarantees uniqueness of the optimal solution to (8). We focus here on finding any optimal solution. Note that (8) is a large-scale, non-smooth convex optimization problem for which a number of algorithms can be adapted (e.g., [41]–[45]). We develop an algorithm based on additive half-quadratic regularization [42], [46], [47] with continuation [42], [45], [48]. The proposed algorithm is efficient and simple to implement, although computational efficiency could potentially be improved by considering other algorithms. In the following, we first present an overview of the proposed algorithm in Section II-B1, then describe each step of the

algorithm in detail from Section II-B2 to II-B5, and finally present a convergence analysis in Section II-B6.

B1. Summary of Algorithm—Because the regularization term in (8) is non-differentiable, solving (8) directly can be challenging. We approximate it with the following differentiable cost function,

$$\widehat{\mathbf{U}}_{\mathrm{s}}=\arg\min_{\mathbf{U}_{\mathrm{s}}\in\mathbb{C}^{N\times L}}\left\|\mathbf{d}-\Omega(\mathbf{F}_{\mathrm{s}}\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{t}})\right\|_2^2+\lambda\sum_{n=1}^{N}\sum_{m=1}^{M}\phi((\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{f}})_{n,m}). \tag{11}$$

where $\phi(t):\mathbb{C}\rightarrow\mathbb{C}$ is the Huber function defined as

$$\phi(t)=\begin{cases}|t|^2/2\alpha, & \text{if } |t|\le\alpha,\\ |t|-\alpha/2, & \text{if } |t|>\alpha.\end{cases} \tag{12}$$

The parameter α controls the accuracy of the approximation of (8) by (11); Note that as $\alpha\rightarrow 0$, (11) approaches (8).

In Section II-B2, we will show that for a fixed α, solving (11) is equivalent to solving the following half-quadratic minimization problem:

$$\min_{\{\mathbf{U}_{\mathrm{s}},\mathbf{G}\}}\left\|\mathbf{d}-\Omega(\mathbf{F}_{\mathrm{s}}\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{t}})\right\|_2^2+\frac{\lambda}{2\alpha}\left\|\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{f}}-\mathbf{G}\right\|_{\mathrm{F}}^2+\lambda\left\|\mathrm{vec}(\mathbf{G})\right\|_1, \tag{13}$$

where $\mathbf{G}$ is an auxiliary matrix. Note that (13) is non-quadratic in $\mathbf{G}$ but quadratic in $\mathbf{U}_{\mathrm{s}}$.

Equation (13) can be solved by a simple alternating minimization procedure. Specifically, at the $(\ell+1)$th iteration, we first optimize (13) over $\mathbf{G}$ with a fixed $\mathbf{U}_{\mathrm{s}}^{(\ell)}$, i.e.,

$$\mathbf{G}^{(\ell+1)}=\arg\min_{\mathbf{G}\in\mathbb{C}^{N\times M}}\frac{1}{2\alpha}\|\mathbf{U}_{\mathrm{s}}^{(\ell)}\mathbf{V}_{\mathrm{f}}-\mathbf{G}\|_{\mathrm{F}}^2+\left\|\mathrm{vec}(\mathbf{G})\right\|_1, \tag{14}$$

and then minimize (13) with respect to $\mathbf{U}_{\mathrm{s}}$ with a fixed $\mathbf{G}^{(\ell+1)}$, i.e.,

$$\mathbf{U}_{\mathrm{s}}^{(\ell+1)}=\arg\min_{\mathbf{U}_{\mathrm{s}}\in\mathbb{C}^{N\times L}}\left\|\mathbf{d}-\Omega(\mathbf{F}_{\mathrm{s}}\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{t}})\right\|_2^2+\frac{\lambda}{2\alpha}\|\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{f}}-\mathbf{G}^{(\ell+1)}\|_{\mathrm{F}}^2. \tag{15}$$

General solutions to these two problems are described in Sections II-B3 and II-B4, respectively, and a faster computational algorithm is described in Section II-B5 for the case of Cartesian Fourier sampling. The above alternating minimization procedure is repeated until $\mathbf{G}^{(\ell+1)}$ and $\mathbf{U}_{\mathrm{s}}^{(\ell+1)}$ converge. This iterative procedure can be viewed as a majorize-minimize algorithm for solving (11).

For a fixed non-zero α, the solution to (11) only yields an approximation to the solution of (8). To obtain a better solution, we apply a continuation scheme in which we gradually reduce α to zero. The continuation procedure starts with a large value of α, for which the optimization problem is approximately quadratic. In this case, the alternating minimization typically converges rapidly. For each subsequent continuation step, we decrease the value of α and use the previous solution as an initialization for the next round of optimization. This

procedure is repeated until the cost function of (11) well approximates that in (8). Convergence of this procedure is described in Section II-B6.

B2. Equivalence of (11) and (13)—Note that the Huber function $\phi(t)$ can be equivalently expressed as [47]

$$\phi(t)=\min_{g\in\mathbb{C}}\frac{1}{2\alpha}|t-g|^2+|g|. \tag{16}$$

where g is an auxiliary variable. With (16), the regularization term in (11) can be written as

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\phi(|\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{f}}|_{n,m})=\min_{\mathbf{G}\in\mathbb{C}^{N\times M}}\frac{1}{2\alpha}\left\|\mathbf{U}_{\mathrm{s}}\mathbf{V}_{\mathrm{f}}-\mathbf{G}\right\|_{\mathrm{F}}^{2}+\left\|\mathrm{vec}(\mathbf{G})\right\|_{1}. \tag{17}$$

Therefore, minimizing (11) is equivalent to minimizing (13).

B3. Solution of (14)—Equation (14) is separable with respect to each entry of **G**, and $\mathbf{G}^{(\ell+1)}$ can be obtained analytically through [49]

$$\mathbf{G}^{(\ell+1)}=\mathscr{S}(\mathbf{U}_{\mathrm{s}}^{(\ell)}\mathbf{V}_{\mathrm{f}}), \tag{18}$$

where $\mathbb{C}^{M\times M}:\mathbb{C}^{N\times M}\rightarrow\mathbb{C}^{N\times M}$ is a soft-thresholding operator defined as

$$\mathscr{S}(\mathbf{Q})_{n,m}=\begin{cases}0, & \text{if } |\mathbf{Q}_{n,m}|<\alpha,\\ \frac{\mathbf{Q}_{n,m}}{|\mathbf{Q}_{n,m}|}(|\mathbf{Q}_{n,m}|-\alpha), & \text{if } |\mathbf{Q}_{n,m}|\geq\alpha,\end{cases} \tag{19}$$

for any $\mathbf{Q}\in\mathbb{C}^{N\times M}$.

B4. Solution of (15)—Equation (15) is a standard linear least squares problem, and its optimal solution is given by the following normal equations:

$$\left(\mathscr{A}^{*}\mathscr{A}+\frac{\lambda}{2\alpha}\mathscr{B}^{*}\mathscr{B}\right)\mathbf{U}_{\mathrm{s}}^{(\ell+1)}=\mathscr{A}^{*}(\mathbf{d})+\frac{\lambda}{2\alpha}\mathbf{G}^{(\ell+1)}\mathbf{V}_{\mathrm{f}}^{H}, \tag{20}$$

where $\mathscr{A}(\mathbf{U})\perp\!\!\!\perp\Omega(\mathbf{F}_{\mathrm{s}}\mathbf{U}\mathbf{V}_{\mathrm{t}})$ and $\mathscr{B}(\mathbf{U})\perp\!\!\!\perp\mathbf{U}\mathbf{V}_{\mathrm{f}}$. It can be shown that $\mathscr{B}^{*}\mathscr{B}$ is positive definite when $\mathbf{V}_{\mathrm{t}}$ has full row rank. Since $\mathscr{A}^{*}\mathscr{A}$ is also positive semidefinite, the coefficient matrix corresponding to $\mathscr{A}^{*}\mathscr{A}+\frac{\lambda}{2\alpha}\mathscr{B}^{*}\mathscr{B}$ is positive definite, guaranteeing a unique solution to (20). Since we assumed that the rows of $\mathbf{V}_{\mathrm{t}}$ are orthonormal, it can be shown that the rows of $\mathbf{V}_{\mathrm{f}}$ are also orthonormal, and thus (20) can be rewritten as

$$\left(\mathscr{A}^{*}\mathscr{A}+\frac{\lambda}{2\alpha}\mathscr{I}\right)\mathbf{U}_{\mathrm{s}}^{(\ell+1)}=\mathscr{A}^{*}(\mathbf{d})+\frac{\lambda}{2\alpha}\mathbf{G}^{(\ell+1)}\mathbf{V}_{\mathrm{f}}^{H}. \tag{21}$$

Equation (21) can be solved efficiently using iterative matrix solvers. Our implementation uses the conjugate gradient (CG) algorithm, initialized with $\mathbf{U}_{\mathrm{s}}^{(\ell)}$ to improve computational speed.

B5. Efficient computation of (15) for Cartesian Fourier sampling—If $\mathbf{F}_s$ is orthonormal, as is often the case with Cartesian Fourier sampling, (15) can be decoupled to enable much more efficient computation. Specifically, in this case, (15) can be converted into

$$\mathbf{U}_k^{(\ell+1)} = \arg\min_{\mathbf{U}_k \in \mathbb{C}^{N\times L}} \left\| \Omega(\mathbf{U}_k \mathbf{V}_t) - \mathbf{d} \right\|_2^2 + \frac{\lambda}{2\alpha} \| \mathbf{U}_k \mathbf{V}_f - \mathbf{B}^{(\ell+1)} \|_F^2, \tag{22}$$

where $\mathbf{U}_s^{(\ell+1)} = \mathbf{F}_s^H \mathbf{U}_k^{(\ell+1)}$ and $\mathbf{G}^{(\ell+1)} = \mathbf{F}_s^H \mathbf{B}^{(\ell+1)}$. Equation (22) has a decoupled structure, which can be understood by rewriting

$$\Omega(\mathbf{U}_k \mathbf{V}_t) = \begin{bmatrix} \Omega_1(\mathbf{u}_1 \mathbf{V}_t) \\ \Omega_2(\mathbf{u}_2 \mathbf{V}_t) \\ \vdots \\ \Omega_N(\mathbf{u}_N \mathbf{V}_t) \end{bmatrix}, \tag{23}$$

where $\mathbf{u}_n \in \mathbb{C}^{1\times L}$ represents the nth row of $\mathbf{U}_k$, and Ω_n is the sampling operator that takes samples from the nth row of $\mathbf{U}_k \mathbf{V}_t$. The solution to (22) can be shown to be equivalent to

$$\mathbf{u}_n^{(\ell+1)} = \arg\min_{\mathbf{u}_n \in \mathbb{C}^{1\times L}} \left\| \Omega_n(\mathbf{u}_n \mathbf{V}_t) - \mathbf{d}_n \right\|_2^2 + \frac{\lambda}{2\alpha} \| \mathbf{u}_n \mathbf{V}_f - \mathbf{b}_n^{(\ell+1)} \|_F^2, \tag{24}$$

for $n = 1, \ldots, N$, where $\mathbf{d}_n$ is the measured data from the nth row of $\Omega^*(\mathbf{d})$ and $\mathbf{b}_n^{(\ell)}$ is the nth row of $\mathbf{B}^{(\ell)}$. To obtain the solution to (24), we solve

$$\left(\mathcal{A}_n^* \mathcal{A}_n + \frac{\lambda}{2\alpha} \mathcal{I} \right) \mathbf{u}_n^{(\ell+1)} = \mathcal{A}_n^*(\mathrm{d}_n) + \frac{\lambda}{2\alpha} \mathbf{b}_n^{(\ell+1)} \mathbf{V}_f^H, \tag{25}$$

for $n = 1, \ldots, N$, where $\mathcal{A}_n(\mathbf{u}) \perp\!\!\!\perp \Omega_n(\mathbf{u}\mathbf{V}_t)$. Note that each of the normal equations in (25) has only L unknowns. Equation (25) implies that the coefficient matrix of the linear least squares problem (22) can be permuted to have a block diagonal structure. There are a total of N blocks, each of which has $L \times L$ elements. Such a block diagonal matrix guarantees that CG has fast convergence (CG converges to the optimal solution within L iterations [50], assuming exact arithmetic).

B6. Convergence analysis—The proposed algorithm has the following convergence properties. First, as $\alpha \to 0$, the optimal solution of (13) converges to that of (8). Second, for a fixed α, the sequence $\{(\mathbf{G}^{(\ell)}, \mathbf{U}_s^{(\ell)})\}$ generated by the alternating minimization algorithm converges to an optimal solution of (13). Third, some entries of $\mathbf{G}$ have finite convergence, and the other entries of $\mathbf{G}$ and $\mathbf{U}_s$ have q-linear (quotient-linear) convergence rates. The first property follows from the fact that the continuation procedure belongs to a special case of the quadratic penalty method and thus has guaranteed global convergence (see Proposition 4.2.1 in [40]). The derivation of the remaining two properties is more involved, and is given in the Appendix.

III. Application example

In this section, we show some representative results from simulation and *in vivo* real-time cardiac MRI experiments to illustrate the performance of the proposed method.

A. Simulation results

A complex-valued numerical cardiac phantom was used for our simulation studies. The simulation was designed to emulate single-channel, real-time cardiac MR experiments with variable heart rate and variable respiratory rate. The phantom was created from real human cardiac MR data which were collected using retrospective ECG-gating during a single breath-hold and used to generate a time series of images representing a single prototype cardiac cycle. We applied multiple time-warps to this prototype to simulate heart rate variability, and the individual time-warped heart-beats were concatenated together to form a long image sequence spanning multiple cardiac cycles. Subsequently, respiration was modeled by applying an additional quasi-periodic spatial deformation [51] to this image sequence. The simulated cardiac and respiratory variability were consistent with statistics from the PhysioBank archive [52]. Acquisition parameters for our simulations included: acquisition matrix size = 200 × 256, field of view (FOV) = 273 mm × 350 mm, effective spatial resolution = 1.36 mm × 1.36 mm, slice thickness = 6 mm, and T_R = 3 ms.

A Cartesian Fourier-based sampling scheme was used for our simulations, part of which is shown in Fig. 1(a). We acquire one **k**-space line per T_R in an interleaved fashion, alternating between sampling one line from central **k**-space and sampling one line from outer **k**-space. The lines from central **k**-space are acquired in a sequential order, while the lines acquired from outer **k**-space follow a random order. Let N_c and N_p respectively denote the number of lines for central and complete **k**-space. With this (**k**, t)-space sampling pattern, one full set of **k**-space lines (called a "data frame") is acquired in a time interval of $2(N_p - N_c)T_R$. The total acquisition time T_a is $2N_{df}(N_p - N_c)T_R$, with N_{df} being the total number of data frames.

In image reconstruction, we assume that the entire data acquisition interval can be divided into subintervals of length T_m, as illustrated in Fig. 1(b), during which temporal signal variations are negligible. Under this assumption, the measured data within each T_m interval can be treated as an instantaneous snapshot. This kind of temporal modeling is commonly used in real-time imaging, although other types of temporal modeling, such as the bandlimited model in [30], could also be used.

Using the above sampling and temporal modeling schemes, the measured data can be partitioned into two sets. One set contains the fully sampled central **k**-space data with high temporal resolution, and the other contains the sparsely sampled data with high spatial resolution. We determine $\mathbf{V}_t$ from all the data in the fully sampled region of **k**-space as described in Section II-A [8], [22]–[27]. Note that with this scheme for sampling and $\mathbf{V}_t$ estimation, $\mathbf{V}_t$ is automatically synchronized temporally with all the data used for image reconstruction. Also $\mathbf{V}_t$ can represent general temporal variations (e.g., periodic or aperiodic) in the data.

Throughout all the simulations, we used $N_c = 8$ and $N_p = 200$ for data acquisition. We used $T_m = 2N_cT_R = 16T_R$ for the temporal modeling. As a consequence, 16 readout lines (i.e., 8 percent of the full k-space data) were measured at each reconstructed time point, and the reconstructions were performed at a frame rate of 21 images per second (i.e., an effective temporal resolution of 48 ms).

A1. General evaluation—We compared PS-Sparse with Basic-PS and Basic-Sparse, using simulated data for T_a = 35.3 s (i.e., N_{df} = 32). We performed Basic-PS reconstruction for $1 \le L \le 32$. With the proposed sampling scheme, the Basic-PS solution is unique for these values of L, since N_{df} (equal to the minimum number of times each **k**-space line is sampled) is always greater than or equal to L. We also performed PS-Sparse reconstructions using the same data set for $1 \le L \le 800$. Note that in our simulation setup, $L = M = T_a/T_m =$

800 is the full model order, for which PS-Sparse corresponds to Basic-Sparse. We manually chose the regularization parameter λ for PS-Sparse at every 10 model orders (including the full model order) to minimize the following reconstruction error:

$$\text{error}=\frac{\|\mathbf{C}_g-\widehat{\mathbf{U}}_s\mathbf{V}_t\|_F}{\|\mathbf{C}_g\|_F}, \tag{26}$$

where $\mathbf{C}_g \in \mathbb{C}^{N\times M}$ is the gold standard.

As a comparison, Eckart-Young (EY) approximation [50], defined as

$$\widehat{\mathbf{E}}_L = \arg\min_{\text{rank}(\mathbf{E})\le L}\|\mathbf{E}-\mathbf{C}_g\|_F^2, \tag{27}$$

with L ranging from 1 to 800 was also computed from the gold standard using SVD. The rank-L EY approximation represents the best possible performance that any order-L PS model can achieve. The error of the EY approximations reflects the underlying model mismatch between the PS model and the ground truth.

The reconstruction errors for these simulations are shown in Fig. 2. Note that when L is small, both Basic-PS and PS-Sparse have large reconstruction errors, due to the limited capability of the low-order PS models. As L increases, the reconstruction accuracy of Basic-PS and PS-Sparse improves, as expected. However, as we continue to increase L (e.g., beyond 16), the Basic-PS reconstructions quickly deteriorate because the inverse problem becomes increasingly ill-conditioned. This ill-conditioning problem is nicely overcome by PS-Sparse, with sparsity as a regularizer, as shown in Fig. 2(b). As L approaches the full model, PS-Sparse reduces to Basic-Sparse and the reconstruction error becomes worse.

To illustrate the reconstruction improvement of PS-Sparse over Basic-PS and Basic-Sparse, a set of mid-systolic cardiac snapshots are shown in Fig. 3; the corresponding temporal variations from a vertical line through the left ventricle are shown in Fig. 4. At very low model orders (such as $L = 8$), the PS constraint is too restrictive and both Basic-PS and PS-Sparse reconstructions suffer from severe blurring (e.g., in the endocardial border and papillary muscles). Increasing the model order to 16 benefits both reconstruction methods. The advantages of PS-Sparse over Basic-PS become clearer as the model order further increases. For example, at $L = 32$, the Basic-PS reconstruction has severe artifacts due to ill-conditioning, which are removed in the PS-Sparse reconstruction. The PS-Sparse reconstruction remains stable for much higher model orders (e.g., $L = 60$). When L equals the full model order, the PS-Sparse reconstruction reduces to the Basic-Sparse reconstruction, and image artifacts, such as spatiotemporal blurring, appear.

To complement the results in Figs. 3 and 4, reconstructions of four different cardiac phases, ranging from end-systole to end-diastole in a cardiac cycle, are shown in Fig. 5. As can be seen, the PS-Sparse reconstructions are consistently better than the other reconstructions across all cardiac phases.

A2. Influence of T_a—We have also evaluated PS-Sparse with respect to different data acquisition window lengths, which provides useful insight into the performance of PS-Sparse in accelerating imaging experiments. The simulation study was done with $T_a = 26.5$, 35.3, 44.2, and 53.0 s (i.e., $N_{df} = 24$, 32, 40 and 48). Image reconstruction was done using Basic-PS, Basic-Sparse, and PS-Sparse, with $L = 8$, 16, and 24 for Basic-PS and PS-Sparse. As before, we manually selected the regularization parameter of each method for each data

acquisition window length to optimize their performance. Fig. 6 shows the reconstruction errors as a function of N_{df}. The reconstruction errors of Basic-PS of all three model orders improve as N_{df} increases, as expected, because longer data acquisition windows resulted in better conditioning of the PS model fitting. This effect becomes more pronounced for higher-order PS models (e.g., $L = 24$). For Basic-Sparse, since the ratio of the number of measurements to the number of reconstructed spatiotemporal voxels remains the same for different acquisition window lengths, the reconstruction error of Basic-Sparse stays at about the same level for different values of N_{df}. For PS-Sparse, the reconstruction error is relatively less sensitive to N_{df} for both high and low model orders, and is lower than that of either Basic-PS (with matched model order) or Basic-Sparse for a given N_{df}.

Fig. 7 shows the reconstructions of one mid-diastolic phase using $N_{df} = 24$ and 48. When $L = 8$, the PS constraint causes significant blurring in both Basic-PS and PS-Sparse reconstructions, which did not improve as the length of data acquisition window increases. For $L = 24$, Basic-PS reconstructed a higher quality image, though this only occurred with a relatively long data acquisition ($N_{df} = 48$). With shorter data acquisition windows, the Basic-PS reconstructions have serious artifacts due to ill-conditioning problem. In contrast, PS-Sparse with the same model order reconstructs higher-quality images at both $N_{df} = 24$ and 48. Although the Basic-Sparse reconstructions are relatively stable for $N_{df} = 24$ and 48, they both suffer from significant blurring artifacts. To better illustrate the image artifacts, we show the error images of these reconstructions in Fig. 8, which were obtained by subtracting the reconstructions from the gold standard.

A3. Influence of noise—We have also investigated the noise sensitivity of PS-Sparse using simulations. We simulated data acquisition with $T_a = 35.3$ s (i.e., $N_{df} = 32$), and contrast-to-noise ratios (between the blood pool and background) of 20 and 10. Reconstructions were performed using Basic-PS, Basic-Sparse, and PS-Sparse with $L = 24$ and manually-selected regularization parameters. We include a video clip of these reconstruction results in the supplementary material. As can be seen, the Basic-PS reconstructions suffer from noticeable artifacts at both noise levels. Basic-Sparse reconstructions are also corrupted by significant spatiotemporal blurrings. These problems were nicely overcome by PS-Sparse.

A4. Comparison with implicit PS reconstruction—Lingala et al. recently proposed to use implicit PS constraints based on the Schatten p-norm along with a sparsity constraint for dynamic image reconstruction [32], [33]. The reconstruction problem was formulated as

$$\overline{\mathbf{C}} = \arg \min_{\mathbf{C} \in \mathbb{C}^{N \times M}} \left\| \mathbf{d} - \Omega(\mathbf{F}_s \mathbf{C}) \right\|_2^2 + \lambda_s \left\| \mathrm{vec}(\mathbf{C}\mathbf{F}_t) \right\|_1 + \lambda_r \left\| \mathbf{C} \right\|_p^p, \tag{28}$$

where $p = 1$ was used in [33], and $p < 1$ in [32].

We compared PS-Sparse with the above two methods, denoted as NN-Sparse and Sch-Sparse, respectively. We used ($\mathbf{x}$, f)-domain sparsity constraints for our comparisons for the sake of consistency, noting however that [32] used an alternative sparsity constraint based on spatiotemporal smoothness. The data was simulated with $T_a = 26.5$ s (i.e., $N_{df} = 24$). For PS-Sparse, we used $L = 32$ and manually selected λ. For NN-Sparse and Sch-Sparse, we manually chose the regularization parameters λ_s and λ_r to optimize their performance. We chose $p = 0.1$ for Sch-Sparse. Note that Sch-Sparse is nonconvex and subject to local minima, but that our results were insensitive to the initialization in this case. Fig. 9 shows reconstructions from the end-diastolic and end-systolic cardiac phases. Although Sch-Sparse has slightly better performance over NN-Sparse, its reconstructed spatial and temporal

images still suffer from significant blurring and artifacts. In contrast, the corresponding PS-Sparse reconstructions show higher spatial and temporal fidelity than both the NN-Sparse and Sch-Sparse reconstructions.

The superior performance of PS-Sparse can be attributed to the use of the explicit PS constraints. Since the explicit PS constraints have fewer degrees of freedom than the implicit PS constraints, the low-rank structure was enforced more strongly. It has been observed empirically that low-rank matrix recovery using explicit PS constraints can be more effective than the nuclear norm constraint when the number of measurements is small [53]–[55].

Computationally, PS-Sparse is more efficient than NN-Sparse and Sch-Sparse, especially in the case of Cartesian Fourier sampling. In addition to using explicit PS, PS-Sparse also employs a pre-determined temporal subspace, resulting in a much smaller number of degrees of freedom and a significantly simplified computational problem. In addition, the proposed solution algorithm avoids computation-intensive SVD evaluations used in the implementations of NN-Sparse [33] and Sch-Sparse [32]. We implemented these methods on a 2×Intel® Quad-Core Xeon® 3.16GHz workstation with 48GB RAM running MATLAB® R2011b on a Linux platform. In the experiments described above, PS-Sparse took 11 minutes while NN-Sparse and Sch-Sparse each took more than 2 hours based on the implementations described in [32], [33]. However, it should be noted that a faster augmented Lagrangian algorithm has been implemented on graphics processing units [56], leading to significant reductions in computation time compared to the implementations in [32], [33]. Our proposed method can also benefit from similar implementation, although in-depth evaluation of computational efficiency are beyond the scope of this paper.

B. In vivo results

In this section, we show some representative results of applying PS-Sparse to real-time cardiac imaging. Experimental data were collected from rats on a Bruker (Billerica, MA) Avance AV1 4.7 T magnet, equipped with a 40 G/cm shielded gradient set and a 4-channel array coil. A customized FLASH pulse sequence was used for data acquisition. Acquisition parameters for *in vivo* experiments included: T_R = 7.5 ms, T_E = 2.4 ms, acquisition matrix size = 128 × 128, FOV = 40 mm × 40 mm, effective spatial resolution = 0.31 mm × 0.31 mm, and slice thickness = 1.5 mm. A similar sampling pattern to Fig. 1(a) was used with N_c = 1. All experiments were performed in compliance with federal and institutional regulations and guidelines.

Image reconstruction was done using a sliding window method [57], Basic-PS, Basic-Sparse, and PS-Sparse. For Basic-PS, Basic-Sparse and PS-Sparse, we used $T_m = 2T_R$, (i.e., two readout lines for each reconstructed time point). The reconstructions were performed at a frame rate of 67 images per second. We manually chose L = 16 for both Basic-PS and PS-Sparse based on empirical evaluation of the reconstruction quality. Specifically, an initial estimate of the model order was selected using the method described in [58]. We performed image reconstruction for a range of model orders about this initial estimate, and chose the final model order based on visual evaluation of reconstruction quality. The regularization parameters for Basic-Sparse and PS-Sparse were chosen based on the discrepancy principle [59]. The noise variance for the discrepancy principle was estimated from the variance of the signal in high-frequency regions of **k**-space where measurement noise is dominant [60].

Some representative reconstruction results are shown in Fig. 10. The sliding window (SW) reconstruction is included as a reference to illustrate the level of undersampling in (**k**, t)-space as well as the noise level of the measured data. For Basic-PS and Basic-Sparse, the reconstructions suffer from poor conditioning and blurring artifacts respectively, which is

consistent with our simulation results. As can be seen from the last row of Fig. 10, PS-Sparse produced much improved reconstructions in both magnitude and phase. A video clip of these reconstructions is included in the supplementary material.

IV. Discussion

PS-Sparse makes use of specialized data acquisition and processing to pre-estimate the temporal subspace of **C**. It results in a simplified problem that can be solved efficiently. Pre-estimating the temporal subspace from a subset of the measured data (e.g., a number of fully sampled central **k**-space points) is based on the assumption that this set of data and **C** share the same temporal subspace. In addition to the results shown in this paper, this strategy has been successfully used in a number of other dynamic MRI contexts [24], [26], [61]. However, it is worth noting that the use of the PS model itself does not require pre-estimation of $\mathbf{V}_t$, and reconstruction accuracy can be further improved by using all, rather than just a subset, of the measured data to determine the temporal subspace [29], [30], [62]. The proposed method can be extended to simultaneously estimate $\mathbf{U}_s$ and $\mathbf{V}_t$ from all of the measured data. For example, the PowerFactorization algorithm can be used to perform this extension [29], [30], [53], [62], in which $\mathbf{U}_s$ and $\mathbf{V}_t$ are estimated in an alternating manner. Some preliminary investigations indicate that this approach can further improve robustness to PS modeling error and measurement noise, although the solution is computationally much more challenging.

The examples shown in this paper made use of (**x**, *f*)-domain sparsity, which is a useful constraint for real-time cardiac imaging. Extending the proposed method to other types of sparsity constraints is mathematically straightforward. For example, we can reformulate the reconstruction problem with joint PS and spatial total variation (TV) constraints. The resulting optimization problem can be solved using the algorithm in [35]. Although the TV constraints did not necessarily improve reconstruction results in that case, further improvement could potentially be obtained with other sparsifying transforms.

Both PS-Sparse and Basic-PS use an explicit model order for image reconstruction. Generally, choice of the model order L must be made to balance representational power, the required number of measurements, and noise sensitivity. Because of the additional sparsity constraints, PS-Sparse demonstrates improved robustness to the choice of the model order compared to Basic-PS. In practice, this makes model-order selection easier for PS-Sparse than for Basic-PS. However, in the absence of ground truth, defining meaningful quantitative metrics to evaluate the performance of PS-Sparse under different model orders is nontrivial and is an interesting open problem. Preliminary investigation on the use of information-theoretic model-selection methods for this purpose has been done in [29], [62], but further study is necessary to gain deeper insight.

PS-Sparse also requires the choice of the regularization parameter λ. In our simulation studies, the regularization parameters were manually tuned for optimal performance of the algorithms, but this was only possible because we knew the ground truth. Although automatic selection of optimal regularization parameters is still an open problem, a number of methods have been proposed [59], including: a) discrepancy principle-based method, b) L-curve method, and c) generalized cross-validation method. For processing our *in vivo* experimental data, we used the discrepancy principle-based method to select λ, which yielded good empirical results. It is also worth reiterating that in PS-Sparse, the sparsity constraint serves as a regularizer to stabilize the PS model. So a large range of λ values would produce similar reconstruction results, as long as stability is achieved. However, if λ is chosen too big, PS-Sparse would overemphasize the sparsity constraint, resulting in blurring artifacts.

The simulations and experiments presented in this paper used a random sampling pattern that can lead to large **k**-space jumps between subsequent phase encodings. This kind of sampling is reasonable for certain types of pulse sequences, such as the FLASH sequence used in our *in vivo* experiment. However, abrupt changes in the phase encoding gradient pattern could cause problems like eddy currents, which can lead to undesirable image artifacts for pulse sequences like balanced SSFP [63]. Previous work has addressed such problems by slightly reducing the randomness of the sampling pattern to obtain a trajectory with reduced eddy current artifacts [63], [64]. To enable PS-Sparse with balanced SSFP, we can use the same type of approaches to modify the sampling pattern. Preliminary simulations (not shown in the paper) indicate that PS-Sparse has similar performance with these modified trajectories.

This paper is focused on presenting a general image reconstruction method that can be applied to a variety of dynamic imaging applications. Although improved performance was demonstrated in real-time cardiac imaging, systematic assessment of the clinical utility of PS-Sparse still requires further investigation with respect to specific clinical applications (e.g., cardiac functional analysis and wall motion abnormality detection). PS-Sparse can be subject to many of the limitations shared by other constrained reconstruction methods. In particular, accelerated data acquisition can come at the risk of the potential loss of image features, and the usefulness of the proposed method for specific clinical applications will be dependent on application-specific data characteristics and the image features of clinical interest. Some preliminary results have demonstrated the potential utility of PS-Sparse for a range of dynamic imaging applications, including first-pass myocardial perfusion imaging [61], [65], dynamic speech imaging [66], and parameter mapping [67].

V. Conclusion

This paper presented a new method for image reconstruction from highly undersampled (**k**, *t*)-space data using joint partial separability and sparsity constraints. The proposed method incorporates the complementary advantages of both constraints into a single formulation by using sparsity constraint to regularize PS model-based reconstruction. An efficient, globally convergent algorithm based on half-quadratic regularization with continuation was presented to solve the resulting optimization problem. The performance of the proposed method was evaluated using simulated and *in vivo* real-time cardiac imaging data. The results demonstrated that the proposed method yields significantly better performance than using partial separability and sparsity constraints individually. The proposed method should prove useful for various dynamic imaging applications.

Acknowledgments

This work was supported in part by research grants: NIH-P41-EB001977, NIH-P41-EB015904, NIH-R01-EB013695 and NIH-R21-EB009768.

References

1. Finn JP, Nael K, Deshpande V, Ratib O, Laub G. Cardiac MR imaging: State of the technology. Radiology. 2006; 241:338–354. [PubMed: 17057063]
2. Choyke PL, Dwyer AJ, Knopp MV. Functional tumor imaging with dynamic contrast-enhanced magnetic resonance imaging. J Magn Reson Imaging. 2003; 17:509–520. [PubMed: 12720260]
3. Prince, MR.; Grist, TM.; Debatin, JF. 3D Contrast MR Angiography. New York: Springer-Verlag; 1999.
4. Buxton, RB. Introduction to Functional Magnetic Resonance Imaging: Principles & Techniques. Cambridge: Cambridge University Press; 2002.

5. Willis NP, Bresler Y. Lattice-theoretic analysis of time-sequential sampling of spatiotemporal signals–Part I. IEEE Trans Info Theory. 1997; 43:190–207.
6. Liang ZP, Jiang H, Hess CP, Lauterbur PC. Dynamic imaging by model estimation. Int J Imaging Syst Tech. 1997; 8:551–557.
7. Lustig M, Santos JM, Donoho DL, Pauly JM. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. Proc Int Symp Magn Reson Med. 2006:2420.
8. Jung H, Ye JC, Kim EY. Improved k-t BLAST and k-t SENSE using FOCUSS. Phys Med Biol. 2007; 52:3201–3226. [PubMed: 17505098]
9. Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI. Magn Reson Med. 2008; 59:365–373. [PubMed: 18228595]
10. Trzasko J, Haider C, Borisch E, Campeau N, Glockner J, Riederer S, Manduca A. Sparse-CAPR: Highly accelerated 4D CE-MRA with parallel imaging and nonconvex compressive sensing. Magn Reson Med. 2011:1019–1032. [PubMed: 21608028]
11. Velikina JV, Johnson KM, Block WF, Samsonov AA. Design of temporally constrained compressed sensing methods for accelerated dynamic MRI. Proc Int Symp Magn Reson Med. 2010:4865.
12. Adluru G, McGann C, Speier P, Kholmovski EG, Shaaban A, Dibella E. Acquisition and reconstruction of undersampled radial data for myocardial perfusion magnetic resonance imaging. J Magn Reson. 2009; 29:466–473.
13. Lu W, Vaswani N. Modified compressive sensing for real-time dynamic MR imaging. Proc IEEE Int Conf Image Process. 2009:3045–3048.
14. Liang D, Ying L. Compressed-sensing dynamic MR imaging using partially known support. Proc IEEE Eng Med Biol Conf. 2010:2829–2832.
15. Usman M, Prieto C, Schaeffter T, Batchelor PG. k-t group sparse: A method for accelerating dynamic MRI. Magn Reson Med. 2011; 66:1163–1176. [PubMed: 21394781]
16. Madore B, Glover GH, Pelc NJ. Unaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. Magn Reson Med. 1999; 42:813–828. [PubMed: 10542340]
17. Tsao J, Boesiger P, Pruessmann KP. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. Magn Reson Med. 2003; 50:1031–1042. [PubMed: 14587014]
18. Xu D, King KF, Liang ZP. Improved k-t SENSE by adaptive regularization. Magn Reson Med. 2007; 57:918–930. [PubMed: 17457871]
19. Aggarwal N, Bresler Y. Patient-adapted reconstruction and acquisition dynamic imaging method (PARADIGM) for MRI. Inverse Probl. 2008; 24:045015.
20. Mistretta CA. Undersampled radial MR acquisition and highly constrained back projection (HYPR) reconstruction: Potential medical imaging applications in the post-Nyquist era. J Magn Reson Imaging. 2009; 29:501–516. [PubMed: 19243031]
21. Sharif B, Derbyshire JA, Faranesh AZ, Bresler Y. Patient-adaptive reconstruction and acquisition in dynamic imaging with sensitivity encoding (PARADISE). Magn Reson Med. 2010; 64:501–513. [PubMed: 20665794]
22. Sen Gupta A, Liang Z-P. Dynamic imaging by temporal modeling with principle component analysis. Proc Int Symp Magn Reson Med. 2001:10.
23. Ji J, Liang Z-P. High resolution cardiac magnetic resonance imaging: A model-based approach. Proc IEEE Eng Med Biol Conf. 2001:2268–2271.
24. Liang Z-P. Spatiotemporal imaging with partially separable functions. Proc IEEE Int Symp Biomed Imaging. 2007:988–991.
25. Pedersen H, Kozerke S, Ringgaard S, Nehrke K, Kim WY. k-t PCA: Temporally constrained k-t BLAST reconstruction using principle component analysis. Magn Reson Med. 2009; 62:706–716. [PubMed: 19585603]
26. Brinegar C, Schmitter SS, Mistry NN, Johnson GA, Liang ZP. Improving temporal resolution of pulmonary perfusion imaging in rats using the partially separable functions model. Magn Reson Med. 2010; 64:1162–1170. [PubMed: 20564601]

27. Haldar JP, Liang Z-P. Low-rank approximations for dynamic imaging. Proc IEEE Int Symp Biomed Imaging. 2011:1052–1055.
28. Zhao B, Haldar JP, Liang Z-P. PSF model-based reconstruction with sparsity constraint: Algorithm and application to real-time cardiac MRI. Proc IEEE Eng Med Biol Conf. 2010:3390–3393.
29. Haldar JP, Liang Z-P. Spatiotemporal imaging with partially separable functions: A matrix recovery approach. Proc IEEE Int Symp Biomed Imaging. 2010:716–719.
30. Zhao B, Haldar JP, Brinegar C, Liang Z-P. Low rank matrix recovery for real-time cardiac MRI. Proc IEEE Int Symp Biomed Imaging. 2010:996–999.
31. Gao H, Cai JF, Shen Z, Zhao H. Robust principal component analysis-based four-dimensional computed tomography. Phys Med Biol. 2011; 56:3181–3198. [PubMed: 21540490]
32. Lingala SG, Hu Y, Dibella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR. IEEE Trans Med Imaging. 2011; 30:1042–1054. [PubMed: 21292593]
33. Lingala SG, Hu Y, Jacob M. Real-time cardiac MRI using low-rank and sparsity penalties. Proc IEEE Int Symp Biomed Imaging. 2010:988–991.
34. Trzasko J, Manduca A, Borisch E. Local versus global low-rank promotion in dynamic MRI series reconstruction. Proc Int Symp Magn Reson Med. 2011:4371.
35. Zhao B, Haldar JP, Christodoulou AG, Liang Z-P. Further development of image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints. Proc IEEE Int Symp Biomed Imaging. 2011:1593–1596.
36. Tropp JA, Wright SJ. Computational methods for sparse solution of linear inverse problems. Proc IEEE. 2010; 98:948–958.
37. Candès E, Tao T. Decoding by linear programming. IEEE Trans Info Theory. 2005; 51:4203–4215.
38. Donoho DL. Compressed sensing. IEEE Trans Info Theory. 2006; 52:1289–1306.
39. Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process Lett. 2007; 14:707–710.
40. Bertsekas, DP. Nonlinear Programming. 2. Belmont: Athena Scientific; 1999.
41. Yin W, Osher S, Goldfarb D, Darbon J. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. SIAM J Imaging Sciences. 2008; 1:143–168.
42. Wang Y, Yang J, Yin W, Zhang Y. A new alternating minimization algorithm for total variation image reconstruction. SIAM J Imaging Sciences. 2008; 1:248–272.
43. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sciences. 2009; 2:183–202.
44. Yang J, Zhang Y. Alternating direction algorithm for ℓ_1 problems in compressive sensing. SIAM J Sci Comput. 2011; 33:250–278.
45. Becker S, Bobin J, Candès EJ. NESTA: A fast and accurate first-order method for sparse recovery. SIAM J Imaging Sciences. 2011; 4:1–39.
46. Geman D, Yang C. Nonlinear image recovery with half-quadratic regularization. IEEE Trans Image Process. 1995; 4:932–946. [PubMed: 18290044]
47. Nikolova M, Ng M. Analysis of half-quadratic minimization methods for signal and image recovery. SIAM J Sci Comput. 2005:937–965.
48. Madsen K, Bruun H. A finite smoothing algorithm for linear ℓ_1 estimation. SIAM J Optimization. 1993; 3:223–225.
49. Donoho DL. De-noising by soft-thresholding. IEEE Trans Info Theory. 1995; 41:613–627.
50. Trefethen, LN.; DB. Numerical linear algrbra. Philadelphia: SIAM; 1997.
51. Tsai Y-C, Lin H-D, Hu Y-C, Yu1 C-L, Lin K-P. Thin-plate spline technique for medical image deformation. J Med Biol Eng. 2000; 20(4):203–210.
52. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation. 2000; 101:215–220.
53. Haldar JP, Hernando D. Rank-constrained solutions to linear matrix equations using PowerFactorization. IEEE Signal Process Lett. 2009; 16:584–587. [PubMed: 22389578]

54. Lee K, Bresler Y. ADMiRA: Atomic decomposition for minimum rank approximation. IEEE Trans Info Theory. 2010:4402–4416.
55. Dai W, Milenkovic O. SET: An algorithm for consistent matrix completion. Proc IEEE Int Conf Acoust, Speech, Signal Process. 2010
56. Lingala SG, Hu Y, Dibella E, Jacob M. Accelerated first pass cardiac perfusion MRI using improved k-t SLR. Proc IEEE Int Symp Biomed Imaging. 2011:1280–1283.
57. d' Arcy JA, Collins DJ, Rowland IJ, Padhanni AR, Leach MO. Applications of sliding window reconstruction with cartesian sampling for dynamic contrast enhanced MRI. NMR Biomed. 2002:174–183. [PubMed: 11870913]
58. Ulfarsson MO, Solo V. Dimension estimation in noisy PCA with SURE and random matrix theory. IEEE Trans Signal Process. 2008; 56:5804–5816.
59. Vogel, CR. Computational methods for Inverse Problems. Philadelphia: SIAM; 2002.
60. Sümbül U, Santos JM, Pauly JM. Improved time series reconstruction for dynamic magnetic resonance imaging. IEEE Trans Med Imaging. 2009; 28:1093–1104. [PubMed: 19150785]
61. Christodoulou AG, Zhao B, Zhang H, Ho C, Liang Z-P. Four-dimensional MR cardiovascular imaging: Method and applications. Proc IEEE Eng Med Biol Conf. 2011:3732–3735.
62. Haldar, JP. PhD dissertation. Univ. of Illinois; Urbana-Champaign, Urbana, IL: 2011. Constrained imaging: Denoising and sparse sampling.
63. Bieri O, Markl M, Scheffler K. Analysis and compensation of eddy currents in balanced SSFP. Magn Reson Med. 2005; 54:129–137. [PubMed: 15968648]
64. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. Magn Reson Med. 2009; 61:103–116. [PubMed: 19097216]
65. Christodoulou AG, Zhang H, Zhao B, Ye Q, Hitchens TK, Ho C, Liang Z-P. High resolution 3D first pass myocardial perfusion imaging. Proc Int Symp Magn Reson Med. 2012:1151.
66. Fu M, Christodoulou AG, Naber AT, Kuehn DP, Liang Z-P, Sutton BP. High-frame-rate multislice speech imaging with sparse samping of (k, *t*)-space. Proc Int Symp Magn Reson Med. 2012:12.
67. Zhao B, Lu W, Liang Z-P. Highly accelerated parameter mapping with joint partial separability and sparsity constraints. Proc Int Symp Magn Reson Med. 2012:2233.

Appendix

The convergence of the alternating minimization procedure (i.e., (18) and (20)) for solving (13) is established here. We first convert (18) and (20) into a similar form to that used in [42], and then establish the nonexpansiveness of the linear operators in the alternating minimization. Based on these nonexpansiveness properties, we apply the theorems in [42] to obtain the finite and *q*-linear convergence results in Section II-B6.

For convenience, the proposed alternating minimization for solving (13) is first summarized below:

$$\mathbf{G}^{(\ell+1)}=S(\mathbf{U}_{\mathrm{s}}^{(\ell)}\mathbf{V}_{\mathrm{f}}), \tag{29}$$

$$\mathbf{U}_{\mathrm{s}}^{(\ell+1)}=\left(\mathscr{A}^{*}\mathscr{A}+\frac{\lambda}{2\alpha}\mathscr{I}\right)^{-1}\left(\mathscr{A}^{*}(\mathbf{d})+\frac{\lambda}{2\alpha}\mathbf{G}^{(\ell+1)}\mathbf{V}_{\mathrm{f}}^{H}\right). \tag{30}$$

We next convert (29) and (30) to a form similar to (3.4) and (3.5) in [42]. First, the matrix representation of $\mathscr{A}$ is denoted by $\mathbf{K}\in\mathbb{C}^{P\times NL}$, i.e.,

$$\mathrm{vec}(\mathscr{A}(\cdot))=\mathbf{K}\mathrm{vec}(\cdot). \tag{31}$$

Second, we have

$$\mathrm{vec}\left(\left(\mathscr{A}^*\mathscr{A}+\frac{\lambda}{2\alpha}\mathscr{I}\right)(\cdot)\right)=\left(\mathbf{K}^H\mathbf{K}+\frac{\lambda}{2\alpha}\mathbf{I}_{NL}\right)\mathrm{vec}(\cdot), \tag{32}$$

Let $\mathbf{D}=(\mathbf{V}_{\mathrm{f}}^T\otimes\mathbf{I}_N)\in\mathbb{C}^{MN\times NL}$, where $\otimes$ is the Kronecker product of two matrices. Based on the property of the Kronecker product and the fact that $\mathbf{V}_{\mathrm{f}}\mathbf{V}_{\mathrm{f}}^H=\mathbf{I}_L$, it can be shown that the columns of $\mathbf{D}$ are orthonormal, i.e.,

$$\mathbf{D}^H\mathbf{D}=\mathbf{I}_{NL}. \tag{33}$$

Finally, we introduce the following two positive definite matrices,

$$\mathbf{M}=\mathbf{K}^H\mathbf{K}+\frac{\lambda}{2\alpha}\mathbf{D}^H\mathbf{D}=\mathbf{K}^H\mathbf{K}+\frac{\lambda}{2\alpha}\mathbf{I}_{NL} \tag{34}$$

and

$$\mathbf{T}=\frac{\lambda}{2\alpha}\mathbf{D}\mathbf{M}^{-1}\mathbf{D}^H. \tag{35}$$

It can be easily shown that the maximum singular value of $\mathbf{T}$ is less than 1.

Based on the above notation, (29) and (30) can be rewritten as follows,

$$\mathrm{vec}(\mathbf{G}^{(\ell+1)})=s(\mathrm{vec}(\mathbf{U}_{\mathrm{s}}^{(\ell)}\mathbf{V}_{\mathrm{f}}))=s\circ h(\mathrm{vec}(\mathbf{G}^{(\ell)})), \tag{36}$$

$$\mathrm{vec}(\mathbf{U}_{\mathrm{s}}^{(\ell+1)})=\mathbf{M}^{-1}\left(\frac{\lambda}{2\alpha}\mathbf{D}^H\mathrm{vec}(\mathbf{G}^{(\ell)})+\mathbf{K}^H\mathbf{d}\right), \tag{37}$$

where s performs element-by-element soft-thresholding of a vector with the threshold value of α, $h(\mathbf{v})=\mathbf{D}\mathbf{M}^{-1}(\frac{\lambda}{2\alpha}\mathbf{D}^H\mathbf{v}+\mathbf{K}^H\mathbf{d})$ for $\forall\ \mathbf{v}\in\mathbb{C}^{NM\times 1}$, and $\circ$ represents the composition of two operators.

The above alternating minimization procedures (36) and (37) have the same structure as (3.4) and (3.5) in [42]. Furthermore, we can establish that s and h in (36) have the same nonexpansiveness properties as those in Proposition 3.1 and 3.2 in [42], which were used to prove the following theorems for convergence. Specifically, for s, we have that $|s(\mathbf{a})_i-s(\mathbf{b})_i|\le|\mathbf{a}_i-\mathbf{b}_i|$, for $i=1,\ldots,NM$ and $\forall\ \mathbf{a},\mathbf{b}\in\mathbb{C}^{NM\times 1}$, and $|s(\mathbf{a})_i-s(\mathbf{b})_i|=|\mathbf{a}_i-\mathbf{b}_i|$ if and only if $s(\mathbf{a})_i-s(\mathbf{b})_i=\mathbf{a}_i-\mathbf{b}_i$; for h, $\|h(\mathbf{w})-h(\tilde{\mathbf{w}})\|_2\le\|\mathbf{w}-\tilde{\mathbf{w}}\|_2$, $\forall\ \mathbf{w},\tilde{\mathbf{w}}\in\mathbb{C}^{NM\times 1}$, and $\|h(\mathbf{w})-h(\tilde{\mathbf{w}})\|_2=\|\mathbf{w}-\tilde{\mathbf{w}}\|_2$ if and only if $h(\mathbf{w})-h(\tilde{\mathbf{w}})=\mathbf{w}-\tilde{\mathbf{w}}$.

As a result, the global convergence property for the proposed method can be established based on Theorem 3.4 of [42]. Let $(\mathbf{G}^\infty,\mathbf{U}_{\mathrm{s}}^\infty)$ denote an optimal solution of (13).

Theorem 1 (Theorem 3.4 [42])

For any fixed $\alpha > 0$, the sequence $\{(\mathbf{G}^{(\ell)}, \mathbf{U}_s^{(\ell)})\}$ generated by (36) and (37) from any starting point $(\mathbf{G}^{(0)}, \mathbf{U}_s^{(0)})$ converges to $(\mathbf{G}^{\infty}, \mathbf{U}_s^{\infty})$ of (13).

Furthermore, we can apply Theorems 3.5 and 3.6 in [42] to characterize the convergence rate of $\mathbf{G}$ and $\mathbf{U}_s$. The following theorems will make use of two new index sets defined as:

$I=\{i:|\mathrm{vec}(\mathbf{U}_s^{\infty}\mathbf{V}_f)_i|=|h(\mathrm{vec}(\mathbf{G}^{\infty}))_i|<\alpha\}$ and $E=\{1, \ldots, NL\}\backslash I$, where \ represents a set difference operation.

Theorem 2 (Theorem 3.5 [42])

For the sequence $\{(\mathbf{U}_s^{(\ell)}, \mathbf{G}^{(\ell)})\}$ generated by (36) and (37) from any starting point $(\mathbf{U}_s^{(0)}, \mathbf{G}^{(0)})$, there exists $\zeta_0 \leq \left\|\mathbf{G}^{(0)}-\mathbf{G}^{\infty}\right\|_F^2/\omega^2<+\infty$ such that $\forall i \in I$, $\mathrm{vec}(\mathbf{G}^{(\ell)})_i = \mathrm{vec}(\mathbf{G}^{\infty})_i$ as long as $\ell \geq \zeta_0$, where

$$\omega \triangleq \min_{i\in I}\{\alpha-|h(\mathrm{vec}(\mathbf{G}^{\infty}))_i|\}>0.$$

Theorem 3 (Theorem 3.6 [42])

Let $\mathbf{T}_{EE}=[\mathbf{T}_{i,j}]_{i,j\in E}$ be a $|E|\times|E|$ submatrix of $\mathbf{T}$ whose entries are taken from $\mathbf{T}$ according to E. The sequence $\{(\mathbf{U}_s^{(\ell)}, \mathbf{G}^{(\ell)})\}$ generated by (36) and (37) satisfies

$$\begin{aligned}\left\|\mathrm{vec}(\mathbf{G}^{(\ell+1)}-\mathbf{G}^{\infty})_E\right\|_2 &\leq c_1\left\|\mathrm{vec}(\mathbf{G}^{(\ell)}-\mathbf{G}^{\infty})_E\right\|_2,\\ \left\|\mathbf{M}^{\frac{1}{2}}\mathrm{vec}(\mathbf{U}_s^{(\ell+1)}-\mathbf{U}_s^{\infty})\right\|_2 &\leq c_2\left\|\mathbf{M}^{\frac{1}{2}}\mathrm{vec}(\mathbf{U}_s^{(\ell)}-\mathbf{U}_s^{\infty})\right\|_2,\end{aligned}$$

for $\ell \geq \zeta_0$, where $c_1=\sqrt{\sigma_{\max}((\mathbf{T}^2)_{EE})}$, $c_2=\sqrt{\sigma_{\max}((\mathbf{T})_{EE})}$, $\sigma_{\max}(\cdot)$ represents the maximum singular value of a matrix, and $\mathbf{M}^{\frac{1}{2}}$ is any positive semidefinite Hermitian square root of the Hermitian matrix $\mathbf{M}$, i.e., $\mathbf{M}=\mathbf{M}^{\frac{1}{2}}(\mathbf{M}^{\frac{1}{2}})^H=(\mathbf{M}^{\frac{1}{2}})^2$.

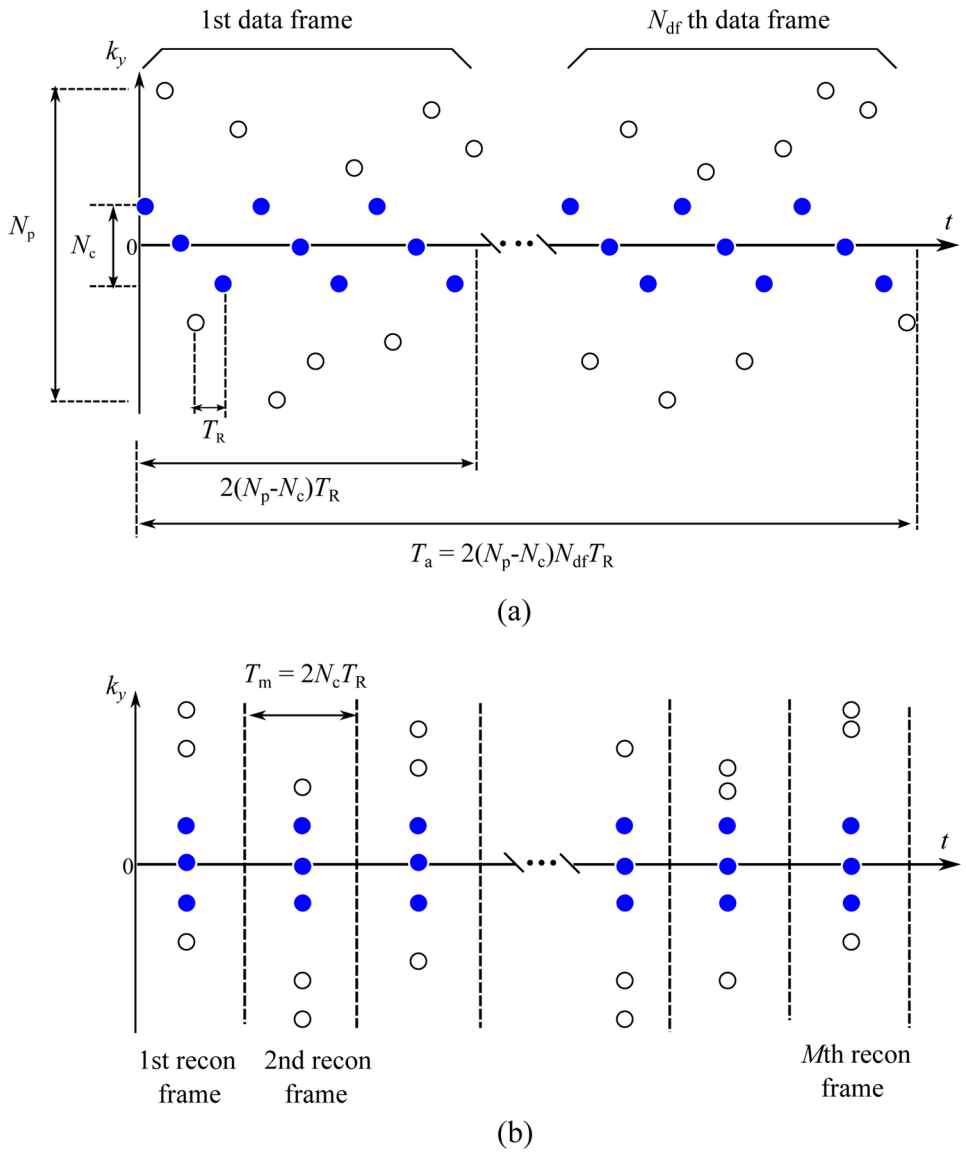

Fig. 1.

(a) (**k**, *t*)-space sampling pattern and (b) temporal modeling used for image reconstruction. As shown in (a), we acquire one k-space line per T_R in an interleaved fashion, alternating between sampling one line from central **k**-space and one line from outer **k**-space. The lines acquired from central **k**-space follow a sequential order, while the lines acquired from outer **k**-space follow a random order. Using this sampling pattern, it takes $2(N_p - N_c)T_R$ to collect a full set of **k**-space lines, denoted as a data frame. The total data acquisition time T_a is $2(N_p - N_c)T_R N_{df}$ with a total number of N_{df} data frames. In image reconstruction, we assume that the entire data acquisition interval can be subdivided into intervals of length T_m, shown by dashed lines in (b), during which temporal signal variations are negligible.

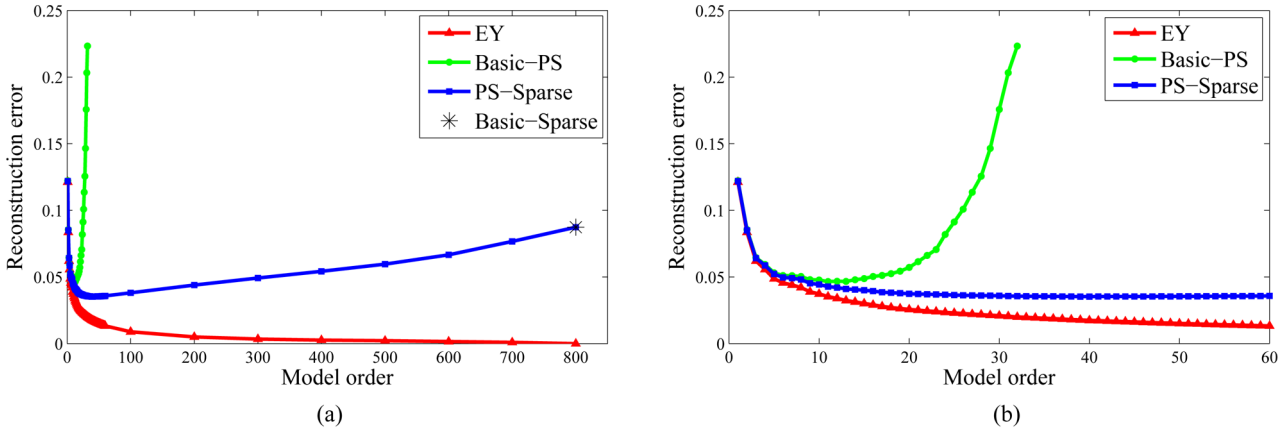

Fig. 2.

(a) Reconstruction error versus L (ranging from 1 to the full model order) with $N_{df} = 32$ as calculated using (26), and (b) zoomed-in plot showing model orders from 1 to 60. Note that PS-Sparse is significantly more robust to changes in model order than Basic-PS. PS-Sparse also has smaller reconstruction error than Basic-Sparse except when the full model order is adopted (when PS-Sparse becomes Basic-Sparse).

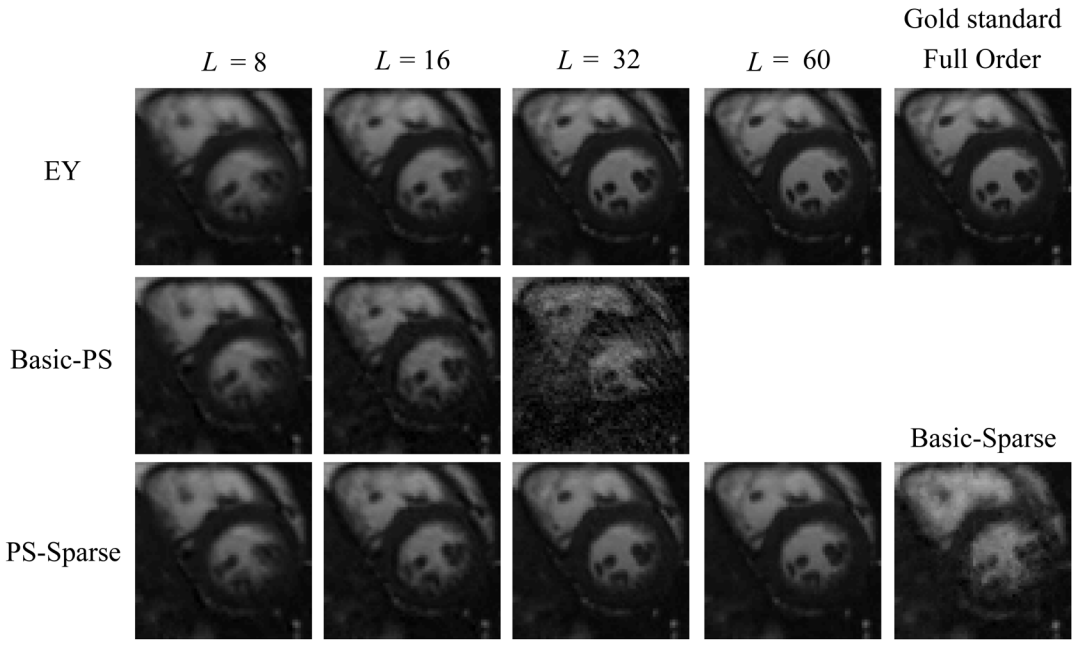

Fig. 3.

Reconstructions from the mid-systolic cardiac phase (zoomed in on the heart). Note that rank-*L* EY provides the best performance that *L*th order Basic-PS and PS-Sparse could achieve, and that full-order EY is equivalent to the gold standard. Note the improvements of PS-Sparse reconstructions over Basic-PS reconstructions and the robustness of PS-Sparse to different model orders.

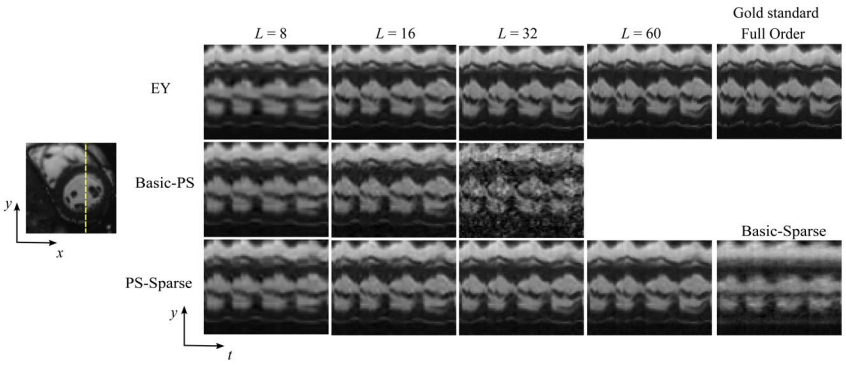

Fig. 4.

The temporal variations of a vertical line passing through the left ventricle (zoomed in on the heart), corresponding to reconstructions in Fig. 3. Note the ill-conditioning problem with Basic-PS with $L = 32$, and the blurring artifacts in the Basic-Sparse reconstruction.

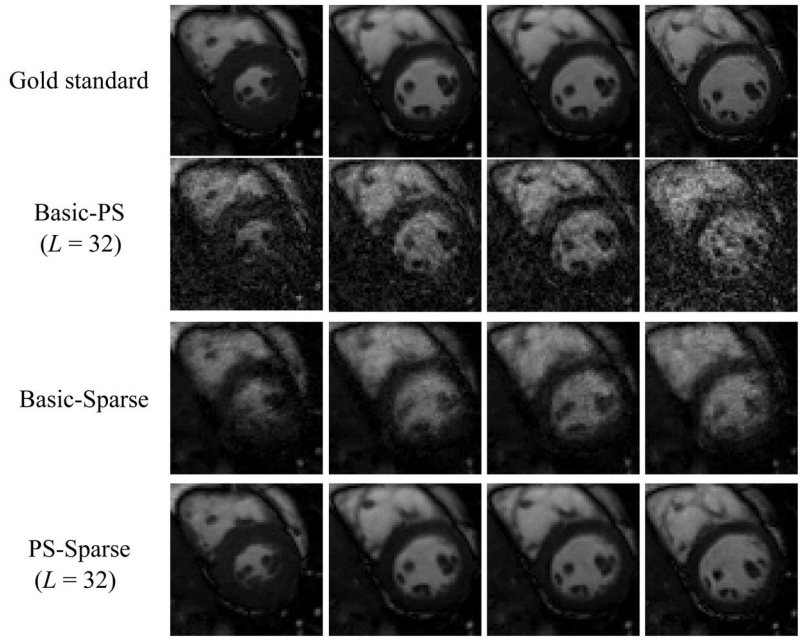

Fig. 5.

Reconstructions in four cardiac phases ranging from end-systole to end-diastole based on the same measurements used in Fig. 3. To avoid repetition, four time points shown here are different from the single time point shown in Fig. 3. Note the improved reconstruction quality of PS-Sparse over Basic-PS and Basic-Sparse.

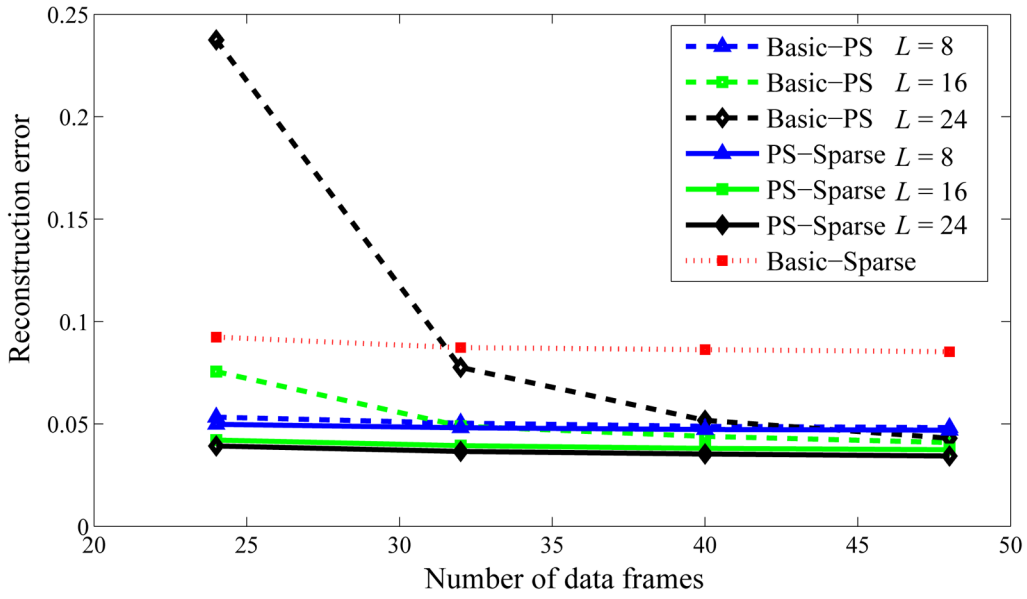

Fig. 6.

Reconstruction error versus the number of data frames for Basic-PS (dashed lines), Basic-Sparse (red dotted line), and PS-Sparse (solid lines). Note that PS-Sparse has lower reconstruction error than either Basic-PS (with the same model order) or Basic-Sparse at each data acquisition window length. In addition, the reconstruction error of PS-Sparse is relatively robust to N_{df} for both high and low model orders.

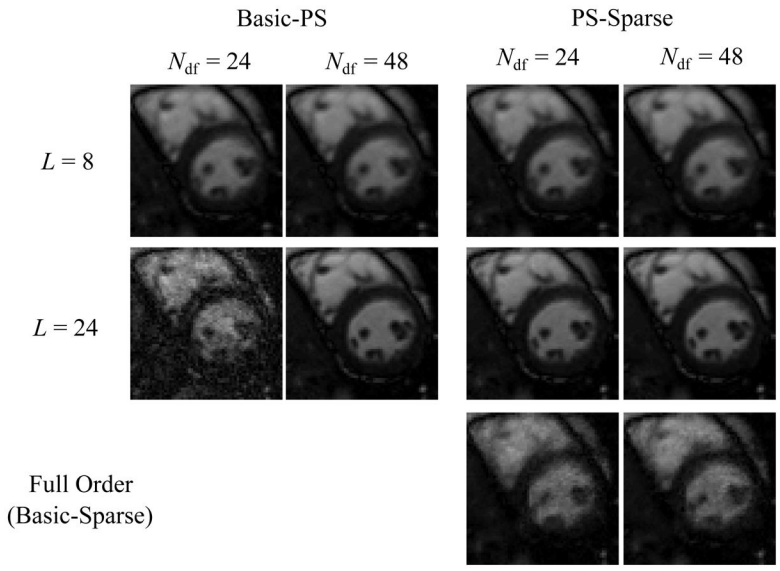

Fig. 7.

Reconstructions of the mid-diastolic cardiac phase (zoomed in on the heart) using N_{df} = 24 and 48. Note the superior performance of PS-Sparse in the presence of limited measurements and also its robustness to the number of measurements.

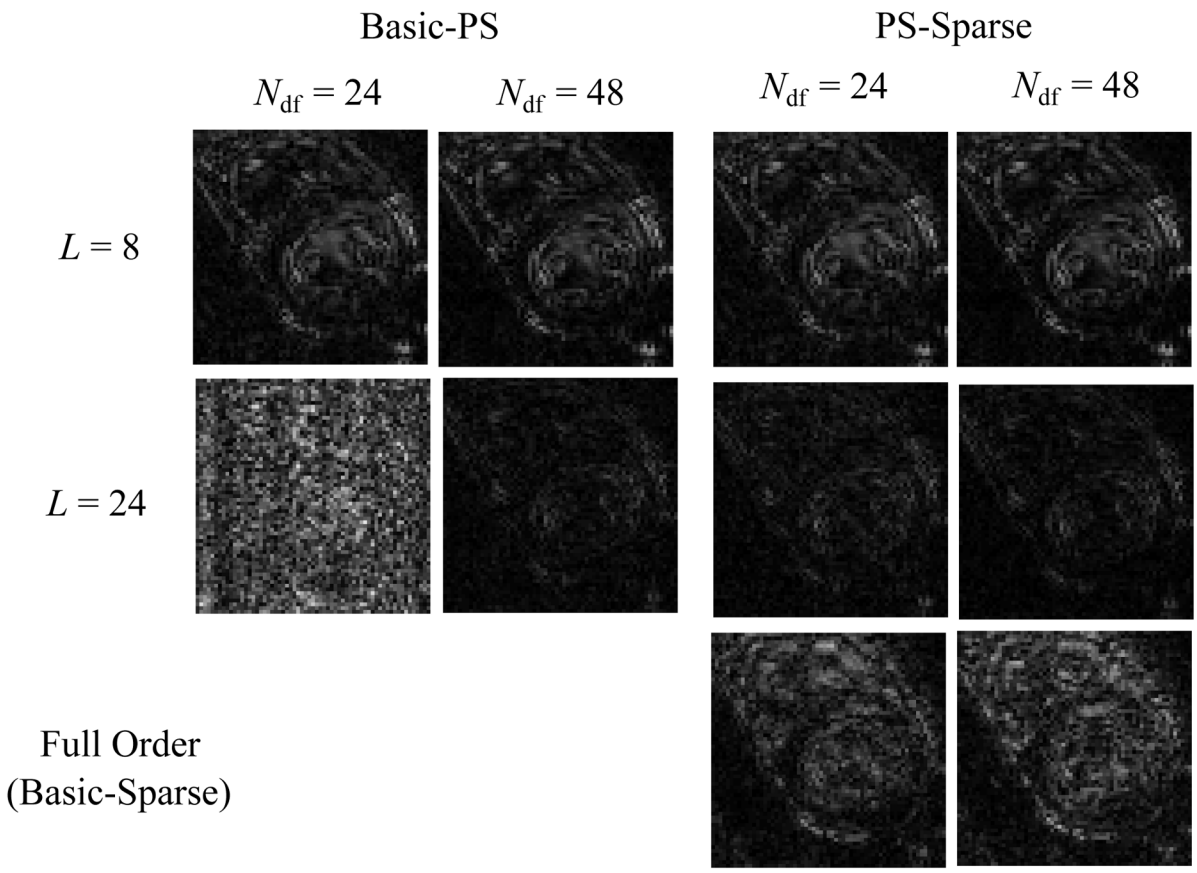

Fig. 8.
Error images for the reconstructions in Fig. 7, which were obtained by subtracting the reconstructions from the gold standard. Note that the error images have been scaled up for visualization purpose (the maximum value of the error images is one seventh of that in Fig. 7.).

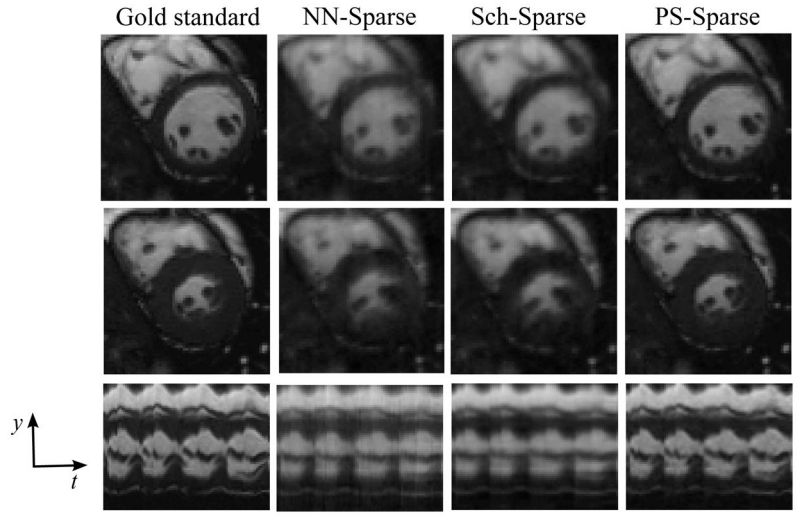

Fig. 9.

Comparison of NN-Sparse, Sch-Sparse, and PS-Sparse. The first two rows show end-diastolic and end-systolic cardiac images, and the last row shows the reconstructed temporal variations of a line passing through the left ventricle. Note that the PS-Sparse reconstructions have higher spatial and temporal fidelity than the NN-Sparse and Sch-Sparse reconstructions (e.g., the myocardial border and the motion of the papillary muscles are much better defined in the PS-Sparse reconstructions than in the NN-Sparse and Sch-Sparse counterparts.).

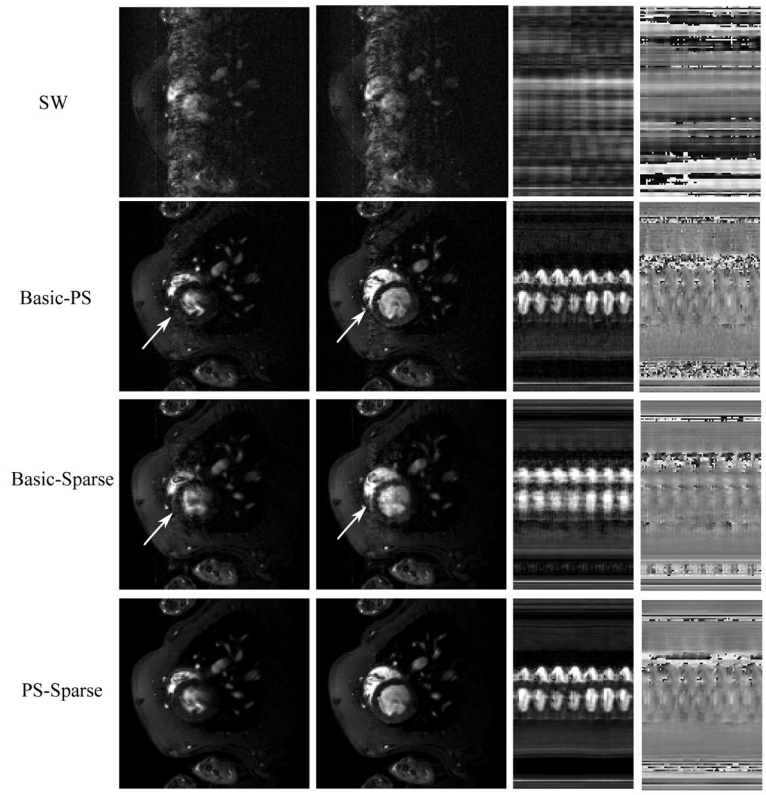

Fig. 10.

Reconstructions from an *in vivo* real-time cardiac MR experiment on rats. The first two columns show the reconstructions in two cardiac phases, and the last two columns show the reconstructed magnitude and phase of temporal variations for a line passing through the left ventricle. Consistent with the simulation results, PS-Sparse shows improved performance over Basic-PS and Basic-Sparse (e.g., regions marked by arrows). Furthermore, the PS-Sparse reconstructions are also capable of capturing the phase changes in dynamic image sequences.