Accelerated Whole-Brain Multi-parameter Mapping Using Blind Compressed Sensing

Sampada Bhave,[1] Sajan Goud Lingala,[2] Casey P. Johnson,[3] Vincent A. Magnotta,[3] and Mathews Jacob[1]*

Purpose: To introduce a blind compressed sensing (BCS) framework to accelerate multi-parameter MR mapping, and demonstrate its feasibility in high-resolution, whole-brain $T_{1\rho}$ and T_2 mapping.

Methods: BCS models the evolution of magnetization at every pixel as a sparse linear combination of bases in a dictionary. Unlike compressed sensing, the dictionary and the sparse coefficients are jointly estimated from undersampled data. Large number of non-orthogonal bases in BCS accounts for more complex signals than low rank representations. The low degree of freedom of BCS, attributed to sparse coefficients, translates to fewer artifacts at high acceleration factors (R).

Results: From 2D retrospective undersampling experiments, the mean square errors in $T_{1\rho}$ and T_2 maps were observed to be within 0.1% up to $R = 10$. BCS was observed to be more robust to patient-specific motion as compared to other compressed sensing schemes and resulted in minimal degradation of parameter maps in the presence of motion. Our results suggested that BCS can provide an acceleration factor of 8 in prospective 3D imaging with reasonable reconstructions.

Conclusion: BCS considerably reduces scan time for multi-parameter mapping of the whole brain with minimal artifacts, and is more robust to motion-induced signal changes compared to current compressed sensing and principal component analysis-based techniques. **Magn Reson Med 000:000–000, 2015. © 2015 Wiley Periodicals, Inc.**

Key words: $T_{1\rho}$ imaging; T_2 imaging; dictionary learning; blind compressed sensing (BCS); 3D multi-parameter mapping

INTRODUCTION

The quantification of multiple tissue parameters from MRI datasets is emerging as a powerful tool for tissue characterization (1–8). Parameters such as proton density, longitudinal and transverse relaxation times (denoted by T_1 and T_2), relaxation times in the rotating frame ($T_{1\rho}$ and $T_{2\rho}$), as well as diffusion have been

shown to be useful in diagnosis of various diseases including cerebral ischemia (9), Parkinson's disease (2–4), Alzheimer's disease (2,5,7), epilepsy (7), multiple sclerosis (6,7), edema (8), necrosis (8), liver fibrosis (10), and intervertebral disc and cartilage degeneration (11–13). Although a single parameter may be sensitive to a number of tissue properties of interest, it may not be specific. Acquiring additional parameters can improve the specificity. The main bottleneck in the routine clinical use of multi-parameter mapping is the long scan time associated with the acquisition of MR images with multiple weightings or contrast values. In addition, long scan times are likely to result in motion-induced artifacts in the data.

A common approach to reduce the scan time is to limit the number of weighted images from which the parameters are estimated. However, this approach precludes the use of multi-exponential fitting methods, limits the accuracy of fits, and restricts the dynamic range of estimated tissue parameters. Several researchers have proposed to accelerate the acquisition of the weighted images using parallel imaging, model-based compressed sensing, and low-rank signal modeling (14–22). The use of parallel imaging alone can only provide moderate acceleration factors (14). Model-based compressed sensing methods rely on large dictionaries generated by Bloch equation simulations of all possible parameter combinations (23). A challenge associated with this scheme is its vulnerability to patient motion, mainly because the dictionary basis functions cannot account for motion-induced signal changes. Another problem with the direct application of this scheme to multiparameter imaging is the rapid growth in the size of the dictionaries with the number of parameters, which also results in increased complexity of the nonlinear recovery algorithm. In this context, methods such as k–t principal component analysis (PCA) and partial separable function (PSF) models that estimate the basis functions from the measured data itself are more desirable; the basis functions can model motion-induced signal changes and thus provide improved recovery of weighted images (22,24).

The main contribution of this article is to optimize the blind compressed sensing (BCS) scheme, which was originally introduced for dynamic imaging (25), to accelerate multi-parameter mapping. The BCS scheme represents the evolution of the magnetization of the pixels as a sparse linear combination of basis functions in a finite dictionary $\mathbf{V}$. Specifically, the Casorati matrix of the data $\mathbf{X}$ is modeled as $\mathbf{X} = \mathbf{UV}$. This model is ideally suited for multi-parameter mapping since there are finite number of distinct tissue types in the specimen with unique

parameter values. The proposed algorithm learns the dictionary basis functions, as well as their sparse coefficients $\mathbf{U}$, from the undersampled data by solving a constrained optimization problem. The criterion is a linear combination of the data fidelity term and a sparsity promoting ℓ_1 norm on the coefficient matrix $\mathbf{U}$, subject to the Frobenius norm constraint on the dictionary $\mathbf{V}$. Unlike methods such as (26) that pre-estimate the dictionary, this approach provides robustness to patient-specific motion. When the data is truly low-rank, the k–t PCA and PSF schemes (22,24) requires very few basis functions to represent it. However, in many cases (e.g., in the presence of motion, multiple tissue types, and simultaneous mapping of multiple parameters), the rank of the dataset can be considerably higher; the larger degrees of freedom will translate to a tradeoff between accuracy and artifacts, especially at high acceleration factors. The BCS scheme uses a considerably larger dictionary of non-orthogonal basis functions, which provides a richer representation of the data compared to the smaller dictionary of orthogonal basis functions used in the k–t PCA and PSF schemes. The sparsity of the coefficients ensures that the number of active basis functions at each voxel are considerably lower than the rank of the dataset. Since the basis functions used at different spatial locations are different, the BCS scheme can be viewed as a locally low-rank scheme; the appropriate basis functions (subspace) at each voxel are selected independently. Since the number of basis functions required at each voxel is considerably lower than the global rank, the BCS scheme can provide a richer representation with lower degrees of freedom; this translates to better trade-offs between accuracy and achievable acceleration, especially in multi-parametric datasets with inter-frame motion.

The BCS algorithm (25) was inspired by the theoretical work on BCS by Gleichman et al. (27). The work by Gleichman et al. considers the same sensing matrix for all time frames, for simplicity of the derivations. The proposed scheme uses different sensing matrices for different frames. The experiments in (25,28) clearly demonstrate the benefit of higher spatial and temporal incoherency offered by this sampling strategy. In addition, the algorithm used in (27) is fundamentally different from our setting. The proposed scheme is also motivated by and has similarities to the PSF model introduced by Liang et al. (16,24,29). However, there are several key differences between the PSF implementations and the proposed scheme. For example, (24) uses the power factorization method to exploit the low-rank structure of $\mathbf{X}$. They jointly estimate $\mathbf{U}$ and $\mathbf{V}$ by alternating between two quadratic optimization schemes involving data consistency terms. Our previous work shows that the BCS scheme provides improved reconstructions than low-rank methods, including power-factorization (24,25), mainly because of the richer dictionary and the lower degrees of freedom. Zhao et al. (29), assumes the data to be low-rank and pre-estimates the orthogonal basis set $\mathbf{V}$ from low resolution data; they then estimate the coefficients using a sparsity penalty on U. This approach can be seen as the first step of our iterative algorithm to jointly estimate $\mathbf{U}$ and $\mathbf{V}$. Specifically, the joint estima-

tion of U and V will provide a richer dictionary with non-orthogonal basis functions, which provide sparser coefficients than the orthogonal basis functions in (29). This is not unexpected since extensive work in image processing have shown that over-complete and non-orthogonal dictionaries/frames offer more compact representations than orthogonal basis sets. The comparisons in Figure 4 demonstrate the performance improvement offered by the proposed joint estimation scheme.

We study the utility of the proposed BCS scheme to simultaneously recover $T_1\rho$ and T_2 maps from undersampled weighted images. We rely on Cartesian subsampling schemes. The proposed scheme yields reasonable estimates from the whole-brain for eightfold under-sampling over the fully-sampled setup, thereby reducing the scan time to 20 min.

METHODS

In a multi-parameter imaging, the k-space data corresponding to different image contrasts is often sequentially acquired by manipulating the sequence parameters (e.g., echo time [TE], spin lock duration/amplitude, and flip angle). We denote the parametric dimension by p. The multi-coil undersampled acquisition of such an experiment is modeled as:

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

where $\mathbf{b}(\mathbf{k}, p)$ represents the concatenated vector of the $\mathbf{k} - p$ measurements from all the coils. $\gamma(\mathbf{x}, p)(\mathbf{x} = (x_1, y_1))$ denotes the underlying images pertaining to different contrasts; and $\mathbf{n}$ is additive noise. $\mathcal{A}$ is the operator that models coil sensitivity and Fourier encoding on a specified $\mathbf{k} - p$ sampling trajectory.

BCS Formulation

The BCS model relies on the assumption that there exists a finite number of distinct tissue types with unique relaxation parameter values within the specimen of interest; the evolution of the magnetization of the tissue types as a function of p can be represented efficiently as a linear combination of basis functions in a dictionary $\mathbf{V}_{R \times N}$. Here R denotes the total number of basis functions in the dictionary and N is the total number of contrast weighted images in the dataset. The signal evolution at the pixel specified by $\mathbf{x}$ is modeled as a sparse linear combination of basis functions $v_i(p); i = 1, ..R$ in $\mathbf{V}$ (also see Fig. 1):

$$\gamma(\mathbf{x}, p) = \sum_{i=1}^{R} \underbrace{u_i(\mathbf{x})}_{\text{sparse spatial weights}} \underbrace{v_i(p)}_{\text{learned bases}} . \qquad [2]$$

Using the Casorati matrix notation (16), the above equation can be rewritten as

$$\Gamma_{M \times N} = \begin{pmatrix} \gamma(\mathbf{x}_1, p_1) & . & \gamma(\mathbf{x}_1, p_N) \\ . & . & . \\ . & . & . \\ \gamma(\mathbf{x}_M, p_1) & . & \gamma(\mathbf{x}_M, p_N) \end{pmatrix} = \mathbf{U}_{M \times R} \, \mathbf{V}_{R \times N}, \quad [3]$$
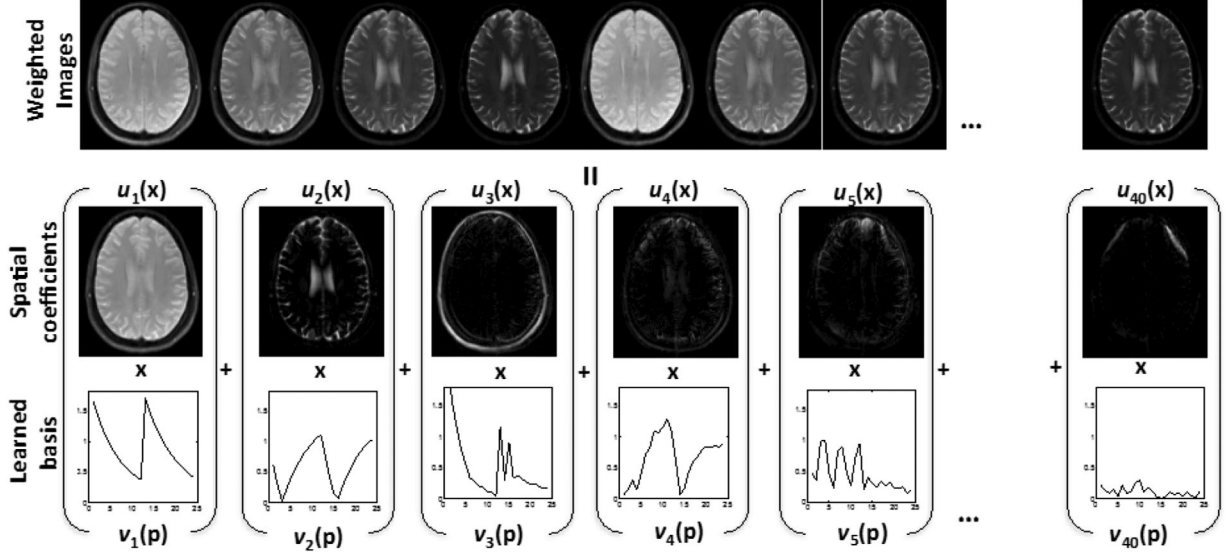
FIG. 1. BCS model representation: The model representation of the multi-parameter signal of a single brain slice with 24 parametric measurements (12 TEs and 12 TSLs) is shown above. The signal Γ is decomposed as a linear combination of spatial weights $u_i(x)$ [$\mathbf{x}$ are the spatial locations (pixels)] and temporal basis functions in $v_i(p)$ (p are the parametric measurements). We observe that only 3–4 coefficients per pixel are sufficient to represent the data. The Frobenius norm attenuates the insignificant basis functions.

where M is the total number of pixels in the image, $u_i(\mathbf{x})$ and $v_i(p)$ in Eq. [2] are respectively the ith column and row entries of $\mathbf{U}, \mathbf{V}$. We formulate the joint recovery of $\mathbf{U}$, $\mathbf{V}$ from under-sampled multi-coil $k - p$ measurements as the following constrained minimization problem:

$$\{\mathbf{U}^*, \mathbf{V}^*\}$$
$$= \underset{\mathbf{U}, \mathbf{V}}{\arg \min} \underbrace{||\mathcal{A}(\mathbf{UV}) - \mathbf{b}||_F^2 + \lambda ||\mathbf{U}||_{l_1}}_{\mathcal{C}(\mathbf{U}, \mathbf{V})} \text{ such that } ||\mathbf{V}||_F^2 < 1.$$

[4]

The first term in Eq. [4] ensures data consistency. The second term promotes sparsity on the spatial coefficients $u_i(\mathbf{x})$ using a convex ℓ_1 norm prior on $\mathbf{U}$, which is given by $||\mathbf{U}||_{l_1} = \left(\sum_{i=1}^{M} \sum_{j=1}^{R} |u(i,j)| \right)$, and λ is the regularization parameter. The optimization problem is constrained by imposing unit Frobenius norm on the over-complete dictionary $\mathbf{V}$, making the recovery problem well posed. Note that we are jointly estimating the sparse coefficients $\mathbf{U}$ and the subject-specific dictionary $\mathbf{V}$ directly from the under-sampled $k - p$ data. Since the dictionary is subject-specific, this approach ensures that any deviations from the true parametric encoding, such as subject motion, field inhomogeneity, and chemical shift artifacts, are learned by the basis functions. The number of active bases at a specified voxel depends on several factors that include partial volume effects, motion, and magnetization disturbances due to inhomogeneity artifacts. The spatial weights $u_i(\mathbf{x})$ are encouraged to be sparse since we expect only a few tissue types to be active at any specified voxel. The main difference of the proposed implementation from (25) is the use of an efficient algorithm and the extension to multi-coil formulation which enables better recovery at high acceleration rates.

Optimization Algorithm

We majorize an approximation of the ℓ_1 penalty on $\mathbf{U}$ in (4) as $||\mathbf{U}||_{\ell_1} \approx \min_{\mathbf{L}} \frac{\beta}{2} ||\mathbf{U} - \mathbf{L}||^2 + ||\mathbf{L}||_{\ell_1}$, where $\mathbf{L}$ is an auxiliary variable. This approximation becomes exact as $\beta \to \infty$. When β is small, the majorization is equivalent to the Frobenius norm on $\mathbf{U}$ (30). We use a variable splitting and augmented Lagrangian optimization scheme to enforce the constraint in Eq. [4]. Thus, the optimization problem corresponds to

$$\{\mathbf{U}^*, \mathbf{V}^*\} = \arg \min_{\mathbf{U}, \mathbf{V}, \mathbf{Q}, \mathbf{L}} ||\mathcal{A}(\mathbf{UV}) - \mathbf{b}||_F^2$$
$$+ \frac{\beta}{2} ||\mathbf{U} - \mathbf{L}||^2 + \lambda ||\mathbf{L}||_{l_1} \text{ such that } ||\mathbf{Q}||_F^2 < 1, \mathbf{V} = \mathbf{Q}$$

[5]

Here, $\mathbf{Q}$ is the auxiliary variable for $\mathbf{V}$. The constraint $\mathbf{V} = \mathbf{Q}$ is enforced by adding the augmented Lagrangian term $\frac{\alpha}{2} ||\mathbf{V} - \mathbf{Q}||^2 + \langle \mathbf{\Lambda}, (\mathbf{V} - \mathbf{Q}) \rangle$ to the above cost function. Here, $\mathbf{\Lambda}$ is the Lagrange multiplier term that will enforce the constraint. These simplifications enable us to decouple the optimization problem in (4) into different subproblems. We use an alternating strategy to solve for the variables $\mathbf{U}, \mathbf{V}, \mathbf{Q}$, and $\mathbf{L}$. All of these subproblems are solved independently in an efficient fashion as described below, assuming the other variables to be fixed.

Update on $\mathbf{L}$: The subproblem can be solved analytically as

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

[6]

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

Update on $\mathbf{U}$: The subproblem on $\mathbf{U}$, assuming the other variables to be fixed, can be written as

$$U_{n+1} = \arg\min_{U} ||\mathcal{A}(U_n V_n) - b||_2^2 + \frac{\lambda\beta}{2}||U - L_{n+1}||_2^2 \quad [7]$$

Since it is a quadratic problem, we solve it using a conjugate gradient (CG) algorithm. Here, $U_n, V_n,$ and L_n are the variables at the nth iteration.

Update on Q: This subproblem, assuming the other variables to be fixed, is solved using a projection scheme as specified in Eq. [8].

$$Q_{n+1} = \begin{cases} V_n & \text{when} \quad ||V_n||_F^2 \leq 1 \\ \dfrac{1}{||V_n||_F} V_n & \text{else} \end{cases} \quad [8]$$

Note that Q_n is obtained by scaling V_n so that the Frobenius norm is unity.

Update on V: Minimizing the cost function with respect to V, assuming other variables to be constant yields

$$V_{n+1} = ||\mathcal{A}(U_{n+1}V_n) - b||_F^2 + < \Lambda_n, V_n - Q_{n+1} \\ > + \frac{\alpha}{2}||V_n - Q_{n+1}||_F^2. \quad [9]$$

The quadratic problem is solved using a CG algorithm. This usually takes a few steps to converge. We use the steepest ascent method to update the Lagrange multiplier at each iteration

$$\Lambda_{n+1} = \Lambda_n + \alpha \ (V_{n+1} - Q_{n+1}) \quad [10]$$

The convergence of the algorithm depends on α and β parameters. Since we use the augmented Lagrangian framework for enforcing constraint on the dictionary, it is not necessary for α to tend to ∞ for the constraint to hold, allowing faster convergence. However, α is progressively updated every iteration to improve the convergence. The inner loop is terminated once the constraint is satisfied, meaning the difference between V and Q is less than a threshold of 10^{-5}. In contrast, for the majorization to well approximate the ℓ_1 penalty, β needs to be a high value. As discussed earlier, the majorization is only exact when $\beta \to \infty$. Since the condition number of the U subproblem is dependent on β, convergence of the algorithm will be slow at high values of β. In addition, the algorithm may converge to a local minimum if it is directly initialized with a large β value. Hence, we use a continuation on β, where we initialize it with a small value and increment it gradually when the cost in Eq. [4] stagnates to a threshold level of 10^{-3}. Our previous work shows that this strategy significantly minimizes the convergence of the algorithm to local minima (25). The outer loop is terminated when constraints for sparse approximation are achieved; in other words, when the cost function given in Eq. [4] converges to a threshold value of 10^{-6}.

The pseudocode of the algorithm is shown below.

Algorithm: BCS(A, b, λ)
Input: $b \to k$ − space measurements
Initialize $\beta > 0$
while $| C_n - C_{n-1} | > 10^{-5} C_n$

do $\begin{cases} \text{Initialize } \alpha > 0 \\ \textbf{while} ||V - Q||^2 > 10^{-5} \\ \quad \textbf{do} \begin{cases} \text{Update } L : \text{Shrinkage Step} \leftarrow \text{Eq. [6]} \\ \text{Update } U : \text{Quadratic sub-problem} \leftarrow \text{Eq. [7]} \\ \text{Update } Q : \text{Projection step} \leftarrow \text{Eq. [8]} \\ \text{Update } V : \text{Quadratic sub-problem} \leftarrow \text{Eq. [9]} \\ \text{Update } \Lambda : \text{Lagrange multiplier update} \leftarrow \text{Eq. [10]} \\ \alpha = 5 * \alpha : \text{Continuation parameter update.} \end{cases} \\ \textbf{if} |C_n - C_{n-1}| < 10^{-3} C_n \\ \quad \textbf{then} \{\beta = 10 * \beta : \text{Continuation parameter update.} \end{cases}$

return (U,V)

Data Acquisition

To demonstrate the utility of the proposed BCS scheme in recovering $T_{1\rho}$, T_2, and S_0 parameters, healthy volunteers were scanned on a Siemens 3T Trio scanner (Siemens Healthcare, Erlangen, Germany) using a vendor provided 12-channel phased array coil. Written informed consent was obtained and the study was approved by the Institutional Review Board. The coil sensitivity maps were obtained using the Walsh method for coil map estimation (31).

To test the feasibility of the algorithm and to optimize the parameters, we first acquired a single-slice fully-sampled axial 2D dataset using a turbo spin echo sequence, combined with $T_{1\rho}$ preparatory pulses (32) and T_2 preparatory pulses (33). Scan parameters were turbo factor of 8, matrix size $= 128 \times 128$, field of view (FOV) $= 22 \times 22$ cm^2, pulse repetition time (TR) $= 2500$ ms, slice thickness $= 5$ mm, B_1 spin lock frequency $= 330$ Hz, and bandwidth $= 130$ Hz/pixel. $T_{1\rho}$- and T_2-weighted images were obtained by changing the duration of the $T_{1\rho}$ (referred as spin lock time [TSL]) and duration of the T_2 preparation pulses (referred as TE), respectively. The data was collected for 12 equispaced TSL and 12 equispaced TE values, both ranging from 10 ms to 120 ms. This provided a total of 24 parametric measurements. The scan time for this dataset was 16 min. Note that five or six TSLs are sufficient for $T_{1\rho}$ estimation using a single exponential fit. However, our main motivation is the future use of this scheme for multiparameter mapping (e.g., joint imaging of $T_{1\rho}$, T_2, $T_{1\rho}$ dispersion imaging, as well as time-resolved parametric mapping). The proposed scheme will prove very useful in these settings. Moreover, larger number of parametric images are essential for more sophisticated models such as multi-exponential model to account for partial volume issues.

To demonstrate the utility of the approach in accelerated 3D imaging, we acquired a prospective 3D dataset using a segmented 3D gradient echo sequence based on the 3D MAPSS approach (34). Scan parameters were FOV $= 22 \times 22 \times 22$ cm^3, matrix size $= 128 \times 128 \times 128$, 64 lines/segments, TR/TE $= 5.6/2.53$ ms, recovery
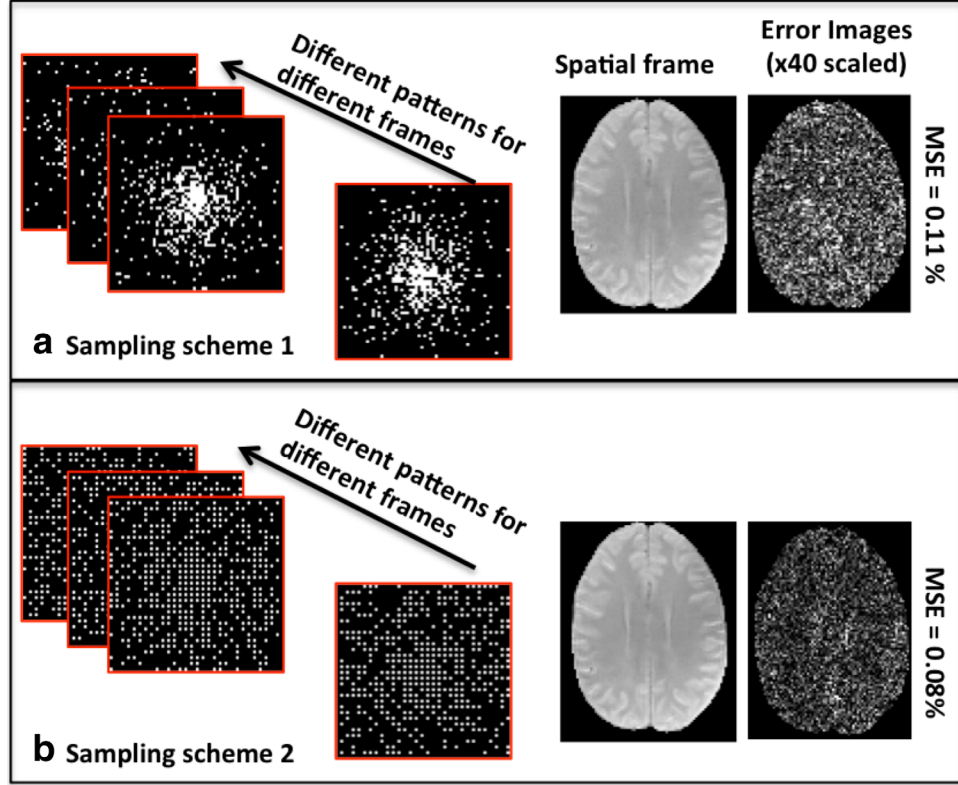
FIG. 2. Choice of sampling trajectories: The sampling patterns for a specific frame for the two choices of sampling schemes are shown in (**a**) and (**b**), respectively. The results are shown for an under-sampling factor of 8. The first sampling scheme [shown in (a)] is a pseudo-random variable density pattern, while the second sampling scheme [shown in (b)] is a combination of a uniform 2 × 2 under-sampling pattern and a pseudo-random variable density pattern. The second column shows one of the weighted images of the reconstructed data using BCS. As seen from the error images in third column, sampling scheme 2 yields better performance. Note that the sampling patterns are randomized over different parameter values to increase incoherency.

time = 1500 ms, resolution 1.7 mm isotropic, bandwidth = 260 Hz/pixel, B_1 spin lock frequency = 330 Hz and constant flip angle = 10°. The readout (frequency encode) direction was (k_x), which enabled us to choose an arbitrary sampling pattern. TEs and TSLs of the T_2 and $T_{1\rho}$ preparation pulses were varied uniformly from 10 to 100 ms providing 10 measurements of each. Scan time of the prospective 3D dataset was 20 min. To be consistent with the 2D dataset the phase encoding plane (phase encode, slice encode) was oriented along the axial ($k_y - k_z$) plane. We perform the recovery of each $y - z$ slice independently.

Optimization and Validation of the Algorithm Using Fully Sampled 2D Acquisition

We used the fully sampled 2D dataset to determine an optimal sampling pattern, optimize the parameters, and compare with other algorithms.

Determination of a Sampling Scheme

To choose an optimal sampling scheme that will work well with the multichannel BCS scheme, we retrospectively under-sample the 2D dataset using two different under-sampling schemes shown in Figure 2a,b. Both patterns correspond to an 8-fold under-sampling. Figure 2a shows the pseudo-random variable density trajectory which oversamples the center of k-space. The sampling scheme 2 as shown in Figure 2b is a combination of a 2 × 2 uniform Cartesian under-sampling pattern and a pseudo-random variable density pattern as in Figure 2a. Acceleration factors of 6, 8, 10, and 12 were achieved as 4-fold uniform under-sampling and 1.5-, 2-, 2.5-, and 3-fold random variable density under-sampling, respectively. The 2 × 2 uniform sampling pattern for different frames is randomly integer shifted in the range $[x, y] = [-1, 1] \times [-1, 1]$ as done in (35) to achieve more incoherency. This sampling scheme may be replaced with Poisson disc sampling (36). We compare the reconstructions provided by the proposed algorithm from the dataset under-sampled using both schemes.

Details of Algorithms and Determination of Their Parameters

We compare the BCS algorithm against compressed sensing (CS);[26] and *k–t* principal component analysis (PCA) (22) methods. A training dataset of 10,000 exponentials is generated assuming the exponential model in Eq. [12] for the CS scheme. A dictionary of 1000 atoms is learned from the training dataset using k-SVD algorithm (37). Specifically, we vary the T_2 and $T_{1\rho}$ values from 1 ms to 300 ms in steps of 3. The learned dictionary is then optimized for signal approximation with at most K

atoms. The sparsity value K is chosen as 7 based on the model fit with respect to fully sampled dataset. The dictionary learned from the training phase is used in the reconstruction. The data is reconstructed using an iterative procedure, which iterates between obtaining the K-term estimate of the signal using orthogonal matching pursuit algorithm and minimizing data consistency as described in (26). k–t PCA is implemented as a two-step approach, where the first step is to estimate the orthogonal basis functions from the training data. The basis functions are estimated from the center 9×9 grid of the fully sampled k-space data using PCA. In the second step, the estimated basis functions are used in reconstruction of the data. We also compare the BCS algorithm with the k–t PCA method with ℓ_1 sparsity constraint enforced on the coefficients. The algorithms are implemented in MATLAB on a quad core linux machine with a NVDIA Tesla graphical processing unit. The regularization parameters of all the algorithms were chosen such that the error between reconstructions and the fully sampled data specified by

$$\text{MSE} = \left(\frac{||\Gamma_{\text{recon}} - \Gamma_{\text{orig}}||_F^2}{||\Gamma_{\text{orig}}||_F^2} \right) \qquad [11]$$

is minimized. We iterate all algorithms until convergence (until the change in the criterion/cost function is less than a threshold which is $10e^{-6}$). With this setting, k–t PCA takes about 10–15 iterations, k–t PCA with ℓ_1 constraint takes 7–8 iterations, BCS takes 60–70 iterations while CS takes around 100 iterations to converge.

We also compare the BCS and k–t PCA methods for their compression capabilities. The 2D dataset with and without motion is represented using different number of basis functions in case of k–t PCA and different regularization parameters (and equivalently different sparsities) in case of BCS. For BCS model, we considered dictionary $\mathbf{V}_{\text{us}}$ estimated from 6-fold undersampled data. To determine the model representation at different compression factors, we solved for the model coefficients $\mathbf{U}$ using the following equation:

$$\hat{\mathbf{U}}_\lambda = \arg\min_{\mathbf{U}} ||\Gamma - \mathbf{U}\mathbf{V}_{\text{us}}||_2^2 + \lambda||\mathbf{U}||_{\ell_1} \qquad [12]$$

We varied the range of λ and minimized the above problem to control the sparsity levels of $\hat{\mathbf{U}}_\lambda$, and hence the compression capabilities. A threshold of 0.1% was applied on $\hat{\mathbf{U}}_\lambda$ to shrink the coefficients that were very small and were not fully decayed to zero during the above ℓ_1 minimization problem. The model approximation error is given by $||\Gamma - \mathbf{U}_\lambda \mathbf{V}_{\text{us}}||_F^2$.

Comparison of the Algorithms

We estimate the parameters S_0, $T_1\rho$, and T_2 by fitting the mono-exponential model

$$M(p) = S_0 . \exp\left(\frac{-\text{TE}(p)}{T_2} \right) . \exp\left(\frac{-\text{TSL}(p)}{T_1\rho} \right) \qquad [13]$$

to the reconstructed images on a pixel by pixel basis using a linear least-squares algorithm. The mean square

error of the parameter maps obtained from the BCS, CS, and k–t PCA algorithms are compared to the ones obtained from the fully sampled data. We mask the reconstructed images before computing the parameter maps to limit our evaluation of $T_1\rho$, T_2, and S_0 to the brain tissue.

The performance of the reconstruction scheme at higher acceleration was assessed by retrospectively under-sampling the dataset at acceleration factors of 6, 8, 10, 12, and 15 using the sampling scheme shown in Figure 2b. To determine the robustness of the proposed scheme to motion, we constructed a simulated dataset with inter-frame motion by adding translational motion resulting in 1 pixel shift and rotational motion of $1°$ to frames 16–21 of the 2D dataset, out of 24 frames. The reconstructed images are aligned to compensate for inter-frame motion, prior to fitting. To demonstrate the advantage of acquiring multiple parameters over single parameter, we compared the $T_1\rho$ maps obtained by applying BCS, k–t PCA, and CS schemes on the combined dataset ($T_1\rho + T_2$) and the $T_1\rho$ only dataset.

Validation of the BCS Algorithm Using Prospective 3D Acquisition

The prospectively under-sampled 3D dataset is recovered using the BCS scheme. The dataset was under-sampled on a Cartesian grid with an acceleration factor of $R = 8$ using the under-sampling scheme 2. Each of the 128 slices in the dataset are recovered independently using BCS. The parameter maps are estimated from the pixels by fitting the mono-exponential model to the data. The mean square error metric could not be used for the 3D experiments as the fully sampled ground truth was not available. Hence, we determine the regularization parameter λ using the L-curve strategy (38).

RESULTS

Fully Sampled 2D Acquisition

The comparisons of the two undersampling patterns at acceleration factor of 8 is shown in Figure 2. The mean square error values and the error images in third column show that sampling scheme 2 (shown in Fig. 2b) provides better reconstructions. Sampling scheme 2 samples outer k-space more than sampling scheme 1 (shown in Fig. 2a), which reduces blurring of the high frequency edges. In other words, the sampling scheme 2 is both randomly and uniformly distributed in k-space making it suitable for multi-channel CS applications. The aliasing introduced by the 2×2 uniform grid in sampling scheme 2 is resolved using information from multiple coils. Using different sampling patterns for different frames increases incoherency and thus helps in better reconstructions. We use sampling scheme 2 for all the subsequent experiments.

We demonstrate the choice of the parameters in BCS and k–t PCA schemes in Figure 3 using 8-fold retrospectively under-sampled data. The comparisons were done in two regimes: one where the subject was still, and one with head motion during part of the scan. In Figure 3a,b, we show the model approximation error as a function of
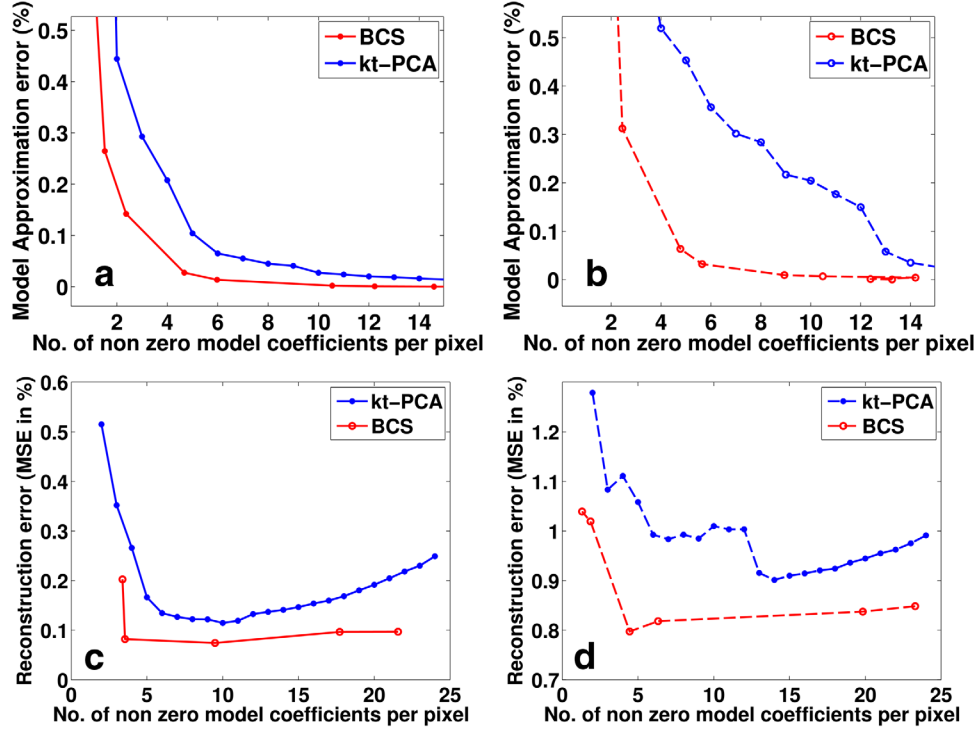
FIG. 3. Comparison of BCS and *k–t* PCA model representation: **a** and **b**: show the model approximation error against the number of non-zero coefficients per pixel of BCS and *k–t* PCA without and with motion, respectively. **c** and **d**: show the reconstruction error against the average number of non-zero model coefficients per pixel of BCS and *k–t* PCA models on the 2D dataset without and with motion respectively. We observe that BCS gives better reconstructions with less number of non-zero model coefficients than *k–t* PCA both in case of with and without motion. In other words, the degree of freedom of BCS is less than that of *k–t* PCA. BCS model gives better compression than *k–t* PCA model as seen from (a) and (b). Note: For (a) and (b), the basis functions in case of BCS were estimated from 6 fold undersampled data and the basis functions of *k–t* PCA were estimated from center of k-space of the fully sampled data.

number of non-zero coefficients per pixel while representing the 2D dataset without and with motion for BCS and *k–t* PCA using learned basis functions, respectively. In case of BCS scheme, the basis functions learned from BCS reconstruction of 6-fold under-sampled data were used, whereas in case of *k–t* PCA, basis functions estimated from center k-space of the fully sampled data were used. We observe that BCS provides better compression capabilities than *k–t* PCA. In other words, the model fitting error in BCS is lower with less number of non-zero coefficients per pixel as compared to *k–t* PCA. We observe from Figure 3c,d that the better signal representation offered by BCS translates to better reconstruction. Specifically, the optimal number of non-zero coefficients that yield minimum reconstruction errors in the *k–t* PCA model (10 in case without motion and 14 for case with motion) is considerably higher than that of BCS model (≈4 in case without motion and ≈5 in case with motion).

In Figure 4, we compare the performance of BCS against *k–t* PCA scheme with and without sparsity constraint and CS schemes for different acceleration factors without motion (right) and in the presence of motion (left). We observe that BCS is capable of providing reconstructions with lower errors, compared with CS and *k–t* PCA schemes with and without sparsity constraint. The better performance of BCS in cases without and with

motion can be attributed to the richer dictionary and lower degrees of freedom over other methods. The ℓ_1 norm on the coefficients and Frobenius norm constraint on the dictionary attenuates the insignificant basis functions which model the artifacts and noise as shown in Figure 5a and thereby minimize noise amplification. In contrast, since the model order (number of non-zero coefficients) in *k–t* PCA without sparsity constraint is fixed a priori, basis functions modeling noise are also learned, especially in the case with motion. This is demonstrated in Figure 5c. Imposing a sparsity constraint on **U** in *k-t* PCA method improves the results over *k–t* PCA without regularization. This scheme can be seen as the first iteration of the BCS scheme. The results in the article clearly demonstrate the benefit in re-estimating the basis functions. Specifically, the BCS scheme enables the learning of non-orthogonal basis functions, which provide sparser coefficients. The CS method, conversely, exhibited motion artifacts as the dictionary is learned from the data model, which does not contain signal prototypes that account for patient-specific motion fluctuations. The comparison of $T_{1\rho}$ and T_2 parameter maps at acceleration factor of 8 are shown in Figure 6. We observe that BCS provides superior reconstructions which translate into better parameter maps as compared to other two schemes in both with and without motion datasets. Figure 7 shows the parameter maps for different
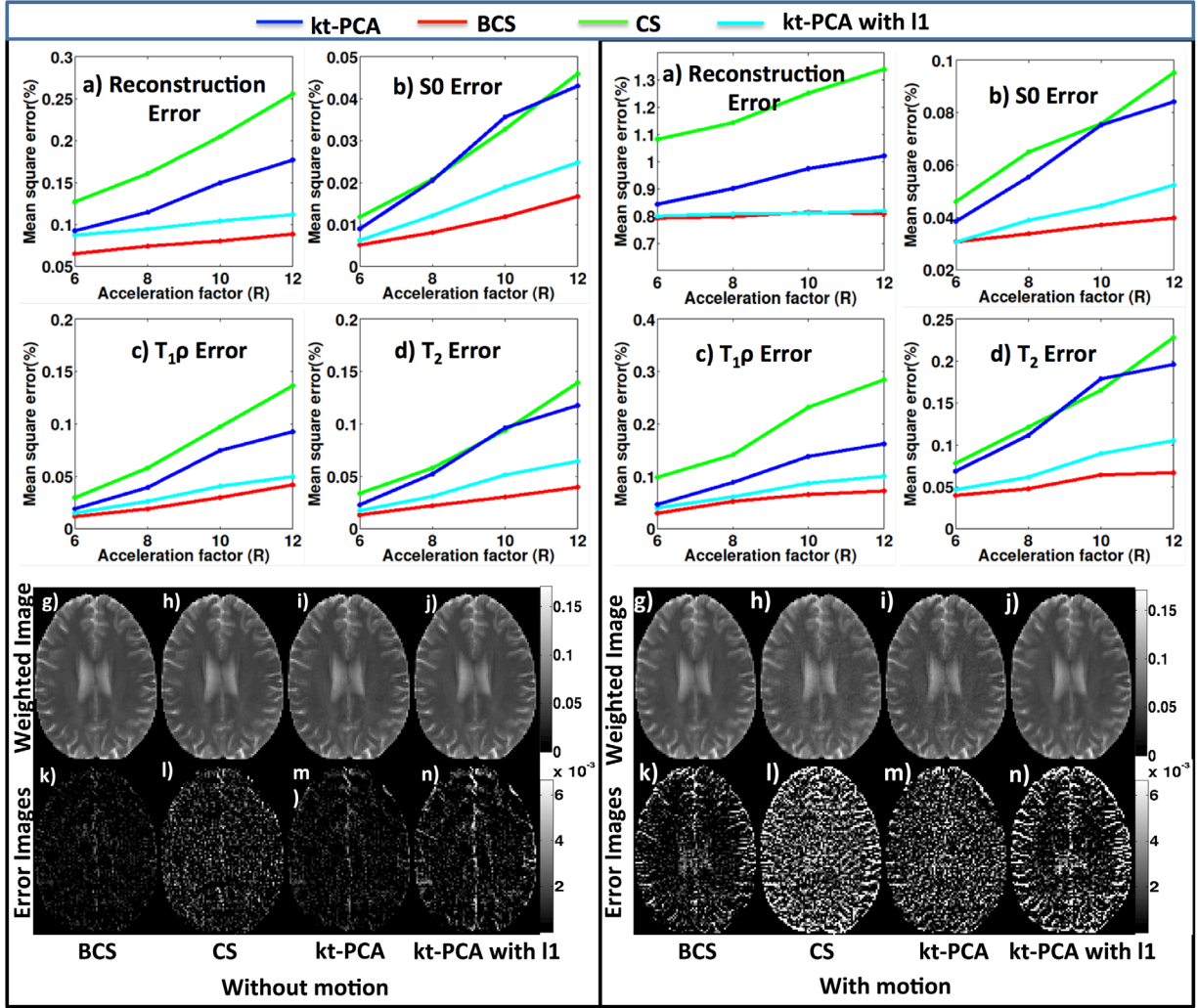
FIG. 4. Comparison of the proposed BCS scheme with different reconstruction schemes on retrospectively under-sampled 2D dataset: The results for dataset without and with motion are shown in (**i**) and (**ii**), respectively. The plots for reconstruction error, S_0 map error, $T_1\rho$ map error, and T_2 map error for BCS, CS, k–t PCA, and k–t CPA with ℓ_1 sparsity schemes are shown in (**a–d**). It is observed that the BCS scheme provides better recovery in both cases. The images in (**g–j**) show one weighted image of the reconstructed dataset at acceleration factor of 8 using the 4 different schemes. We observe that the CS and k–t PCA schemes were sensitive to motion and resulted in spatial blurring as seen in (**ii**)-(**h–j**), which is also evident from the error images.

acceleration factors. Acceleration factors up to 15 were achieved with minimal degradation. All the schemes yield better $T_1\rho$ maps in case of the combined ($T_1\rho + T_2$) dataset as compared to the only $T_1\rho$ dataset as seen in Figure 8. In addition, we observe that BCS gives better performance than other schemes, thus confirming that combining $T_1\rho$ and T_2 datasets does not affect the reconstructions, instead it enables to achieve higher acceleration and improves the specificity of $T_1\rho$.

Prospective 3D Acquisition

The optimal regularization parameter is chosen using the L-curve method as shown in Figure 9. The λ value of 0.07 is then used to recover all the slices. The parameter maps for the prospectively under-sampled 3D dataset recovered using the BCS scheme are shown in Figure 10. These results demonstrate that the BCS scheme yields good parameter maps with reasonable image quality. The

acceleration factor of $R = 8$ enables us to obtain reliable $T_1\rho, T_2$, and S_0 estimates from the entire brain within a reasonable scan time (20 min).

DISCUSSION

We have introduced a BCS framework to accelerate multiparameter mapping of the brain. The fundamental difference between CS schemes and the proposed framework is that BCS learns a dictionary to represent the signal, along with the sparse coefficients from the under-sampled data. This approach enables the proposed scheme to account for motion-induced signal variations. Since the number of different tissue types in the specimen is finite, this approach also enables use of smaller dictionaries, resulting in a computationally efficient algorithm. The main difference of the proposed scheme versus k–t PCA scheme is the non-orthogonality of the basis functions and the sparsity of the coefficients. The
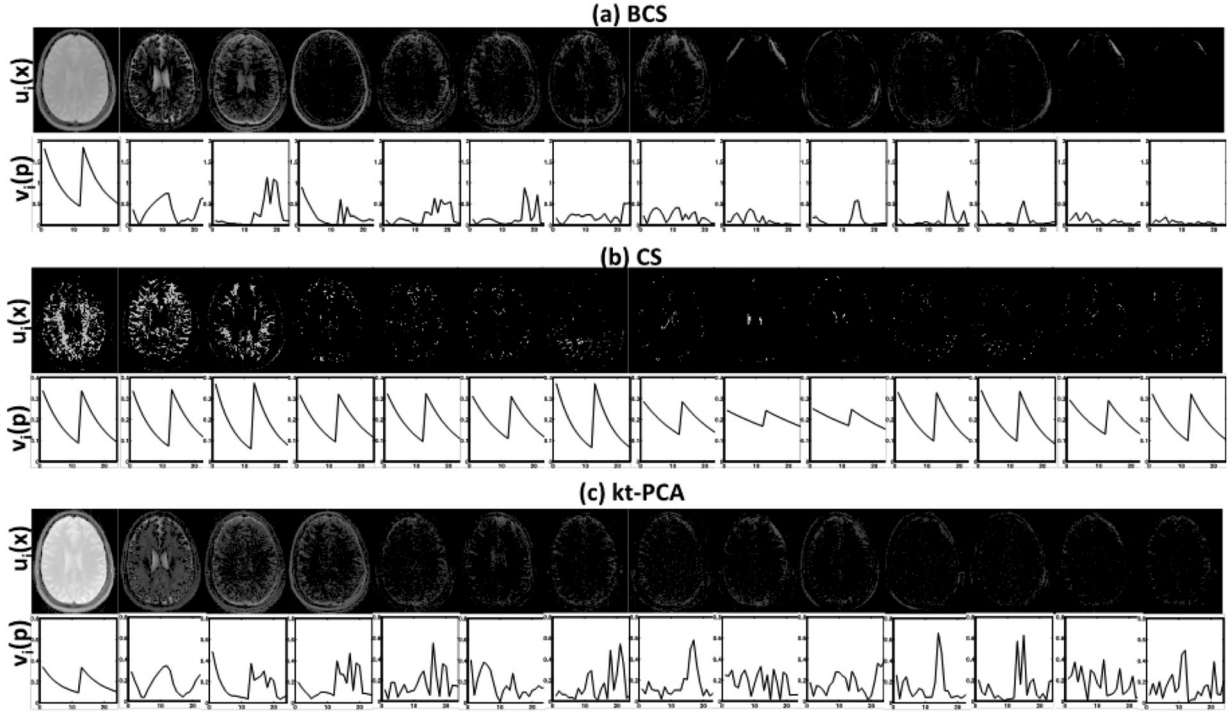
FIG. 5. Model Coefficients and dictionary basis functions for the 2D data with motion: Few spatial coefficients $\mathbf{u}_i(x)$ and their corresponding basis functions $\mathbf{v}_i(p)$ for BCS, CS, and k–t PCA schemes are shown in (**a–c**), respectively. The product entries $\mathbf{u}_i(x)\mathbf{v}_i(p)$ are sorted according to Frobenius norm and first 14 entries are shown here. Since the Frobenius norm constraint attenuates the insignificant basis functions BCS reconstructions have less noise amplification, whereas the basis functions estimated using k–t PCA scheme are noisy.
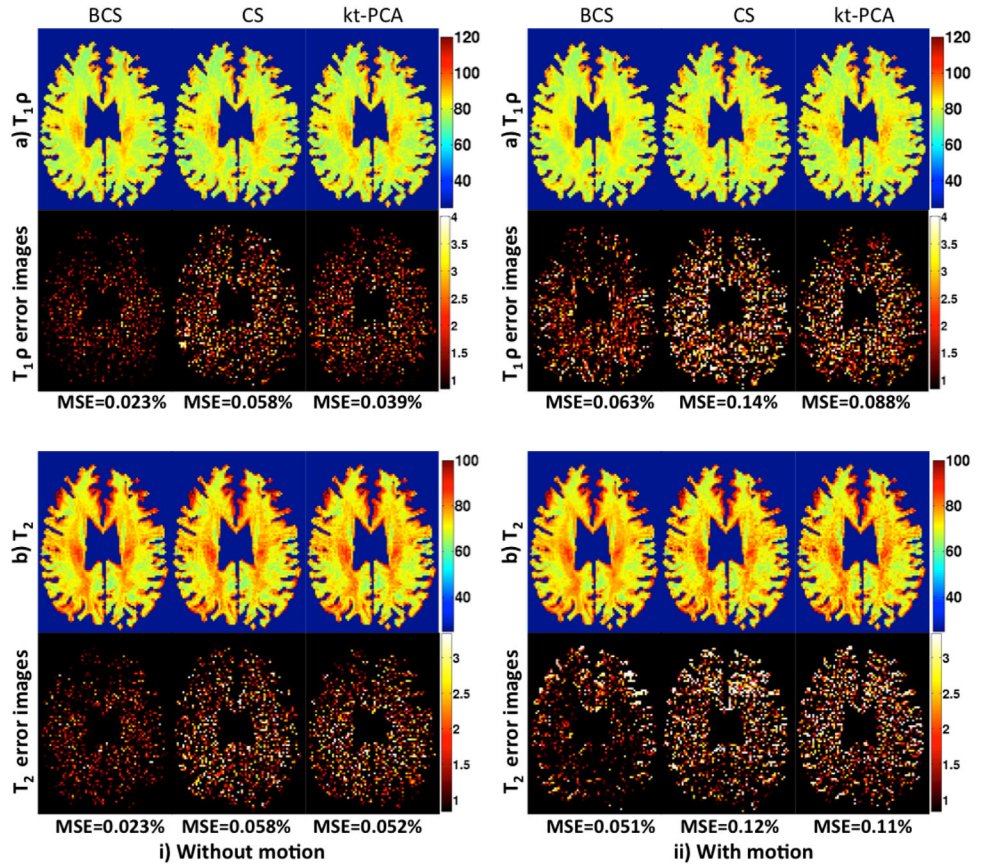
FIG. 6. $T_1\rho$ and T_2 parameter maps for retrospectively undersampled 2D dataset: The $T_1\rho$ and T_2 parameter maps obtained using BCS, CS, and k–t PCA schemes on the 2D dataset with and without motion are shown in (**i**) and (**ii**), respectively. The maps are obtained at acceleration factor of 8. We observe that BCS scheme performs better than CS and k–t PCA schemes in both cases with and without motion. The noise in reconstructions using the k–t PCA and CS schemes propagates to the parameter maps and hence the degradation is higher in case of k–t PCA and CS as compared to BCS.
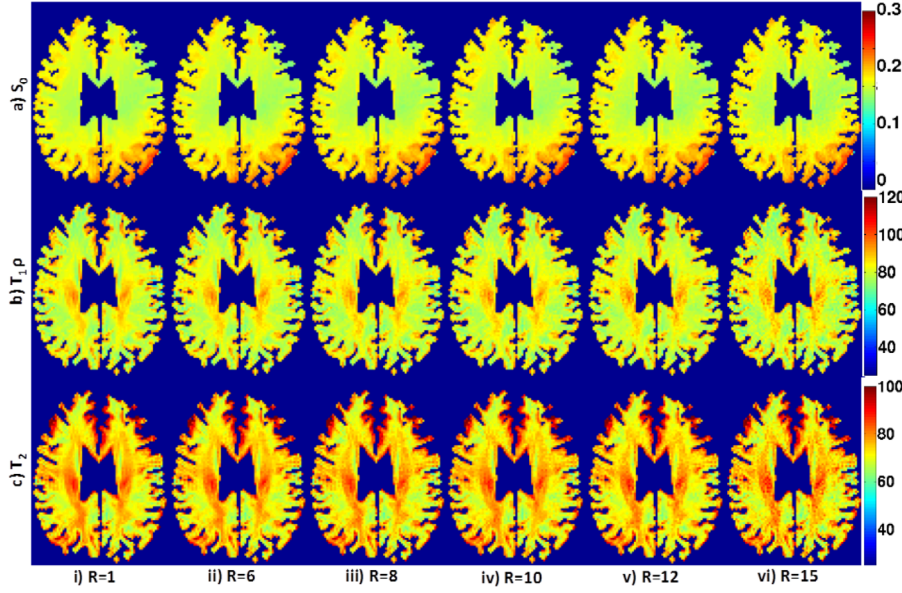
FIG. 7. Parameter maps of a retrospectively under-sampled 2D dataset at different acceleration factors: S_0, $T_1\rho$, and T_2 parameter maps (**a–c**) at acceleration factors $R = 1$, 6, 8, 10, 12, and, 15 are shown in (**i–vi**). We observe reasonable reconstructions for acceleration factors up to 15 with minimal degradation in contrast.

richer model and the fewer degrees of freedom due to the sparsity of the coefficients translate to lower artifacts at high acceleration factors.

Since the k–t PCA basis functions are estimated from the center 9×9 k-space of the fully sampled data, it does not exploit the redundancy due to parallel MRI. The k–t PCA performance may be further improved by a pre-reconstruction step, where the missing k-space data is interpolated from the known samples using GRAPPA (39) or SPIRiT (40), prior to estimating the basis functions. However, no such pre-reconstruction is necessary in BCS since the dictionary is updated itera-

tively with the coefficients in the reconstruction process. The k–t PCA reconstructions, specially in presence of motion can be improved using the model consistency condition (MOCCO) technique (41) introduced recently. Such a model consistency relaxation could also be realized with the BCS model, which is yet to be explored.

Our comparisons with k–t PCA and CS schemes in the case of subjects experiencing head motion show that BCS is more robust to motion. This behavior can be attributed to the ability of the BCS scheme to learn complex basis functions that capture the motion-induced
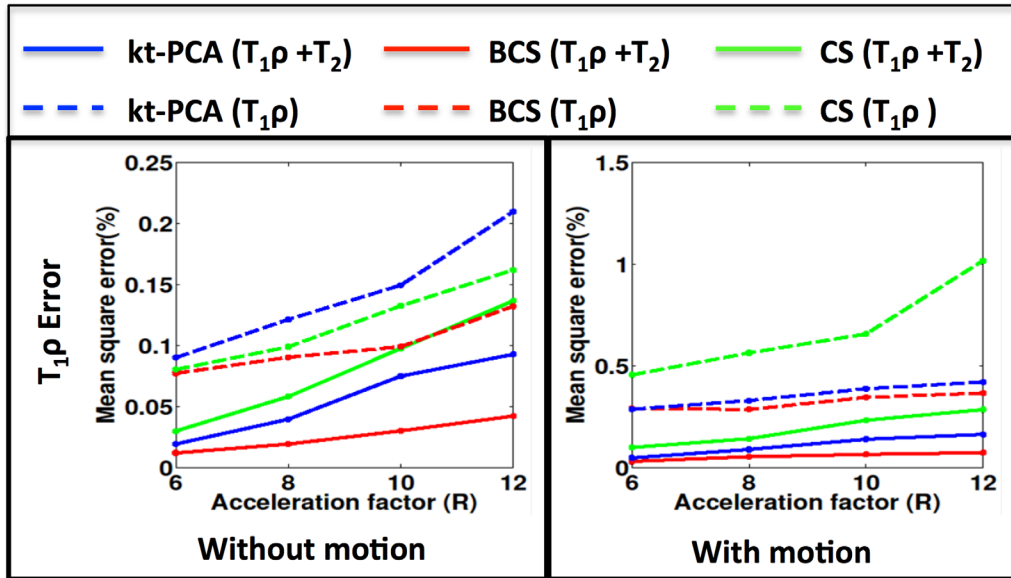

FIG. 8. Comparison of $T_1\rho$ maps errors obtained from reconstructions of combined ($T_1\rho + T_2$) dataset and the $T_1\rho$ only dataset: the $T_1\rho$ maps errors at different accelerations for all the schemes on the combined dataset (solid lines) and only $T_1\rho$ dataset (dotted lines) are shown. The plot on left shows comparisons for the datasets without any motion and the plot on the right shows comparisons for datasets with motion. We observe in both cases that BCS performs better than CS and k–t PCA schemes. In other words, combining the datasets improves the reconstructions.

signal variations. The ability to be robust to motion induced signal variations is especially important in high-resolution whole-brain multi-parameter mapping experiments, where the acquisition time can be significant.

Based on our work that combined low-rank and spatial smoothness priors (28), we observed that the use of spatial smoothness priors along with low-rank priors as in Zhao et al., ISMRM, 2012 can provide better reconstructions. While spatial smoothness priors can be additionally included with BCS to improve performance, this is beyond the scope of this article.

The proposed method can only compensate for inter-frame motion. We correct for the motion using registration of the images in the time series, prior to estimation of the parameter maps. An alternative to this approach is the joint estimation of motion and the low-rank dataset as in (42). The improvement in the results comes from superior reconstruction of the image series, which translates into good quality parameter maps.

The quality of the reconstructions depends on the regularization parameter λ. We used the *L*-curve method to optimize λ. We observed that the value of λ did not vary much across different datasets acquired with the same protocol. Therefore, in the practical setting, once the λ is tuned for one dataset, it could be used to recover other datasets that are acquired using the same protocol. In order for the majorize-minimize algorithm to converge, β should tend to infinity, and convergence of the algorithm is slow at higher values of β. Thus the continuation method plays a significant role in providing faster convergence. Currently, the reconstruction time for one slice is about 40 min on the GPU. We observed that the CG
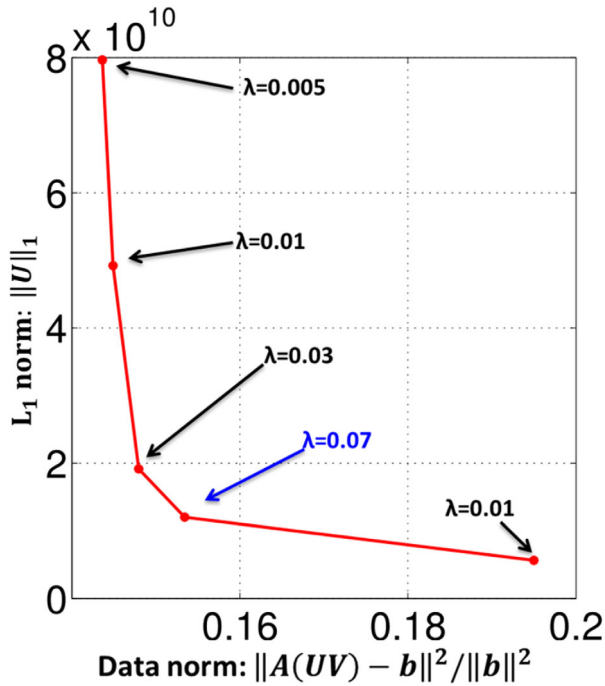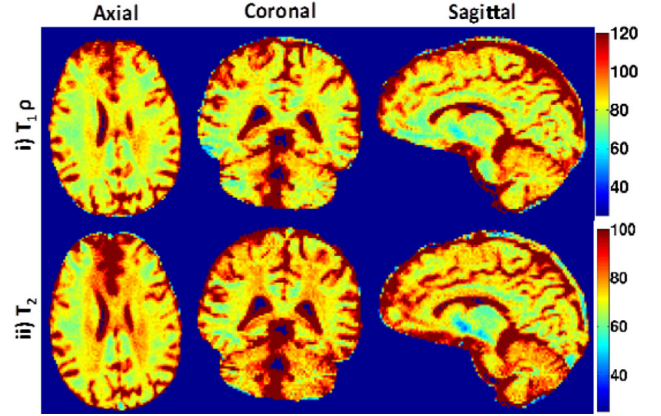

FIG. 10. Parameter maps for 3D prospective undersampled data at $R = 8$: Axial, Coronal, and Sagittal $T_1\rho$ and T_2 parameter maps are shown in (**i**)–(**ii**). With the acceleration of $R = 8$, the scan time was reduced to 20 min. Note: All 128 slices were processed slice by slice to reconstruct the 3D parameter maps.

steps required to solve the quadratic sub-problems are time consuming. These CG steps can be avoided by additional variable splitting in the data consistency term as shown in (43,44), which is a subject of further investigation.

The proposed scheme can be extended in several directions. First, in the current setting, we reconstruct the 3D data slice by slice, but the algorithm can be further modified to reconstruct the entire 3D data at once, thus exploiting the redundancies across slices. However, this will be computationally expensive. Second, additional constraints such as total variation penalty on the coefficients and sparsity of the basis functions (45) can be added to further improve the results. Third, spatial patches can be used to construct dictionaries to exploit the redundancies in the spatial domain (46,47). Lastly, we use a single exponential model to estimate the parameter maps. However, several other models like multi-exponential model (48) which will accommodate for partial volume effects or a Bloch equation simulation-based approach can be used for parameter fitting. Since, these extensions are beyond the scope of this article, we plan to investigate these in future.

CONCLUSION

We introduced a BCS framework, which learns an over-complete dictionary and sparse coefficients from under-sampled data, to accelerate MR multi-parameter brain mapping. The proposed scheme yields reasonable parameter estimates at high acceleration factors, thereby considerably reducing scan time. The robustness of the BCS scheme to motion makes it well suited for multiparameter mapping in a setting with high probability of patient-specific motion or in a dynamic setting like in cardiac applications.

FIG. 9. Choice of regularization parameter λ: The λ parameter was optimized using the *L*-curve strategy (38). We change λ and plot the data consistency error against the smoothness penalty. λ value of 0.07 was chosen as the regularizing parameter for the 3D dataset.

REFERENCES

1. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, Griswold MA. Magnetic resonance fingerprinting. Nature 2013;495: 187–192.

2. Haris M, Singh A, Cai K, Davatzikos C, Trojanowski JQ, Melhem ER, Clark CM, Borthakur A. T1rho (T1ρ) MR imaging in Alzheimer's disease and Parkinson's disease with and without dementia. J Neurol 2011;258:380–385.

3. Nestrasil I, Michaeli S, Liimatainen T, Rydeen C, Kotz C, Nixon J, Hanson T, Tuite PJ. T1ρ and T2ρ MRI in the evaluation of Parkinson's disease. J Neurol 2010;257:964–968.

4. Michaeli S, Sorce DJ, Garwood M, Ugurbil K, Majestic S, Tuite P. Assessment of brain iron and neuronal integrity in patients with Parkinson's disease using novel MRI contrasts. Mov Disord 2007;22:334–340.

5. Borthakur A, Sochor M, Davatzikos C, Trojanowski JQ, Clark CM. T1rho MRI of Alzheimer's disease. Neuroimage 2008;41:1199–1205.

6. Zipp F. A new window in multiple sclerosis pathology: nonconventional quantitative magnetic resonance imaging outcomes. J Neurol Sci 2009;287:S24–S29.

7. Deoni SC. Quantitative relaxometry of the brain. Top Magn Reson Imag 2010;21:101.

8. Alexander AL, Hurley SA, Samsonov AA, Adluru N, Hosseinbor AP, Mossahebi P, Tromp DP, Zakszewski E, Field AS. Characterization of cerebral white matter properties using quantitative magnetic resonance imaging stains. Brain Connect 2011;1:423–446.

9. Jokivarsi KT, Hiltunen Y, Gröhn H, Tuunanen P, Gröhn OH, Kauppinen RA. Estimation of the onset time of cerebral ischemia using T1ρ and T2 MRI in rats. Stroke 2010;41:2335–2340.

10. Deng M, Zhao F, Yuan J, Ahuja A, Wang YJ. Liver T1ρ MRI measurement in healthy human subjects at 3 T: a preliminary study with a two-dimensional fast-field echo sequence. Liver 2012;85.

11. Wang YXJ, Zhao F, Griffith JF, Mok GS, Leung JC, Ahuja AT, Yuan J. T1rho and T2 relaxation times for lumbar disc degeneration: an in vivo comparative study at 3.0-Tesla MRI. Eur Radiol 2013;23:228–234.

12. Li X, Kuo D, Theologis A, Carballido-Gamio J, Stehling C, Link TM, Ma CB, Majumdar S. Cartilage in anterior cruciate ligament–reconstructed knees: MR imaging T1ρ and T2—Initial Experience with 1-year Follow-up. Radiology 2011;258:505.

13. Li X, Pai A, Blumenkrantz G, Carballido-Gamio J, Link T, Ma B, Ries M, Majumdar S. Spatial distribution and relationship of T1ρ and T2 relaxation times in knee cartilage with osteoarthritis. Magn Reson Med 2009;61:1310–1318.

14. Robson PM, Grant AK, Madhuranthakam AJ, Lattanzi R, Sodickson DK, McKenzie CA. Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions. Magn Reson Med 2008;60:895–907.

15. Huang C, Graff CG, Clarkson EW, Bilgin A, Altbach MI. T2 mapping from highly undersampled data by reconstruction of principal component coefficient maps using compressed sensing. Magn Reson Med 2012;67:1355–1366.

16. Liang ZP. Spatiotemporal imaging with partially separable functions. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, Washington DC, USA, 2007. pp. 988–991.

17. Haldar JP, Liang ZP. Spatiotemporal imaging with partially separable functions: a matrix recovery approach. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, Rotterdam, Netherlands, 2010. pp. 716–719.

18. Velikina JV, Alexander AL, Samsonov A. Accelerating MR parameter mapping using sparsity-promoting regularization in parametric dimension. Magn Reson Med 2013;70:1263–1273.

19. 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.

20. Feng L, Otazo R, Jung H, Jensen JH, Ye JC, Sodickson DK, Kim D. Accelerated cardiac T2 mapping using breath-hold multiecho fast spin-echo pulse sequence with k-t FOCUSS. Magn Reson Med 2011; 65:1661–1669.

21. Zhao B, Lam F, Liang Z. Model-based MR parameter mapping with sparsity constraints: parameter estimation and performance bounds. IEEE Trans Med Imaging 2014;33:1832–1344.

22. Petzschner FH, Ponce IP, Blaimer M, Jakob PM, Breuer FA. Fast MR parameter mapping using k-t principal component analysis. Magn Reson Med 2011;66:706–716.

23. Li W, Griswold M, Yu X. Fast cardiac T1 mapping in mice using a model-based compressed sensing method. Magn Reson Med 2012;68: 1127–1134.

24. Zhao B, Haldar JP, Brinegar C, Liang ZP. Low rank matrix recovery for real-time cardiac MRI. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, Rotterdam, Netherlands, 2010. pp. 996–999.

25. Lingala SG, Jacob M. Blind compressive sensing dynamic MRI. IEEE Trans Med Imaging 2013;32:1132.

26. Doneva M, Börnert P, Eggers H, Stehning C, Sénégas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. Magn Reson Med 2010;64:1114–1120.

27. Gleichman S, Eldar YC. Blind compressed sensing. IEEE Trans Inf Theory 2011;57:6958–6975.

28. Lingala SG, Hu Y, DiBella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: kt SLR. IEEE Trans Med Imaging 2011;30:1042–1054.

29. Zhao B, Lu W, Liang Z. Highly accelerated parameter mapping with joint partial separability and sparsity constraints, vol. 2233. In Proc Int Symp Magn Reson Med, Melbourne, Australia, 2012.

30. Hu Y, Lingala SG, Jacob M. A fast majorize–minimize algorithm for the recovery of sparse and low-rank matrices. IEEE Trans Image Proc 2012;21:742–753.

31. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array MR imagery. Magn Reson Med 2000;43:682–690.

32. Charagundla SR, Borthakur A, Leigh JS, Reddy R. Artifacts in T1ρ-weighted imaging: correction with a self-compensating spin-locking pulse. J Magn Reson 2003;162:113–121.

33. Brittain JH, Hu BS, Wright GA, Meyer CH, Macovski A, Nishimura DG. Coronary angiography with magnetization-prepared T2 contrast. Magn Reson Med 1995;33:689–696.

34. Li X, Han ET, Busse RF, Majumdar S. In vivo T1ρ mapping in cartilage using 3D magnetization-prepared angle-modulated partitioned k-space spoiled gradient echo snapshots (3D MAPSS). Magn Reson Med 2008;59:298–307.

35. Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI. Magn Reson Med 2008;59:365–373.

36. Vasanawala S, Murphy M, Alley MT, Lai P, Keutzer K, Pauly JM, Lustig M. Practical parallel imaging compressed sensing MRI: summary of two years of experience in accelerating body MRI of pediatric patients. In IEEE International Symposium Biomedical Imaging: From Nano to Macro. IEEE, Chicago, USA, 2011. pp. 1039–1043.

37. Aharon M, Elad M, Bruckstein A. k-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Sign Proc 2006;54:4311–4322.

38. Hansen PC, O'Leary DP. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J Sci Comput 1993;14:1487–1503.

39. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 2002;47:1202–1210.

40. Lustig M, Pauly JM. SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space. Magn Reson Med 2010;64:457–471.

41. Velikina JV, Samsonov AA. Reconstruction of dynamic image series from undersampled MRI data using data-driven model consistency condition (MOCCO). Magn Reson Med in press.

42. Lingala SG, DiBella E, Jacob M. Deformation corrected compressed sensing (DC-CS): a novel framework for accelerated dynamic MRI. IEEE Trans Med Imaging arXiv preprint arXiv:1405.7718, 2014.

43. Ramani S, Fessler JA. Parallel MR image reconstruction using augmented lagrangian methods. IEEE Trans Med Imaging 2011;30:694–706.

44. Ravishankar S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans Med Imaging 2011;30:1028–1041.

45. Lingala SG, Jacob M. Blind compressed sensing with sparse dictionaries for accelerated dynamic MRI. In 10th International Symposium on Biomedical Imaging (ISBI), IEEE, 2013. pp. 5–8.

46. Wang Y, Zhou Y, Ying L. Undersampled dynamic magnetic resonance imaging using patch-based spatiotemporal dictionaries. In 10th International Symposium on Biomedical Imaging, ISBI, IEEE, San Francisco, USA, 2013. pp. 294–297.

47. Protter M, Elad M. Image sequence denoising via sparse and redundant representations. IEEE Trans Image Proc 2009;18:27–35.

48. Kroeker RM, Mark Henkelman R. Analysis of biological NMR relaxation data with continuous distributions of relaxation times. J Magn Reson 1986;69:218–235.