33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Adaptive Sampling Design for Compressed Sensing MRI

Saiprasad Ravishankar and Yoram Bresler

Abstract— **Compressed Sensing (CS) takes advantage of the sparsity of MR images in certain bases or dictionaries to obtain accurate reconstructions from undersampled k-space data. The (pseudo) random sampling schemes used most often for CS may have good theoretical asymptotic properties; however, with limited data they may be far from optimal. In this paper, we propose a novel framework for improved adaptive sampling schemes for highly undersampled CS MRI. While the proposed framework is general, we apply it with a recently proposed MRI reconstruction algorithm employing adaptive image-patch based sparsifying dictionaries. Numerical experiments demonstrate up to 7 dB improvements in reconstruction PSNR using the adapted sampling scheme, on top of the large improvements reported in our previous work for the adaptive patch-based reconstruction scheme over analytical sparsifying transforms.**

I. Introduction

Magnetic Resonance imaging (MRI) is a non-invasive and non-ionizing imaging technique that offers a variety of contrast mechanisms and enables excellent visualization of both anatomical structure and physiological function. However, a major drawback of MRI that affects clinical throughput and image quality, especially in dynamic imaging applications, is that it is a relatively slow imaging modality. This is because the measurements in MRI, samples in k-space of the spatial Fourier transform of the object, are acquired sequentially in time. Therefore, many techniques try to reduce the amount of data required for accurate reconstruction.

Compressed Sensing enables accurate reconstruction of images from significantly fewer measurements than the number of unknowns, or than mandated by traditional Nyquist sampling, provided the underlying image has a sparse representation in some transform domain, and the acquisition is appropriately incoherent with the transform. The typical formulation of the CSMRI reconstruction problem enforces both sparsity and k-space data consistency using the following Lagrangian setup [1]

$$\min_x \|F_u x - y\|_2^2 + \lambda \|\Psi x\|_1 \qquad (1)$$

Here, $x \in \mathbb{C}^P$ represents as a vector, the P-pixel 2D complex image to be reconstructed, and $y \in \mathbb{C}^m$ represents the k-space measurements. The two are related (in the absence of noise) as $F_u x = y$, where $F_u \in \mathbb{C}^{m \times P}$ is the undersampled Fourier encoding matrix. Matrix Ψ represents a global, typically orthonormal transform such as wavelets. The quality of image reconstruction in CSMRI depends on the choice of the sampling pattern (F_u) and sparsifying transform or *dictionary*.

Various sampling schemes have been proposed for CSMRI such as Cartesian sampling with random phase encodes and pseudo random 2D sampling [1]. Wang et al. [2] approximate pseudo random 2D sampling (for a single 2D image) with bidirectional Cartesian sampling that is realized with a pulse sequence program that switches the directions of phase encoding and frequency encoding during data acquisition. To account for the unequal distribution of signal energy across k-space, Lustig et al. [1] perform variable density random sampling of k-space by drawing sample positions according to a probability density function (pdf). Out of many such candidate patterns, they choose the one that has the lowest mutual coherence with the sparsifying transform. However, the procedure uses an ad-hoc model for the pdf and is nonadaptive. Moreover, finding the optimal sampling scheme that maximizes the incoherence for a given number of samples is a combinatorial problem that is intractable.

Although random sampling has good asymptotic properties and there are theoretical performance guarantees for CS based on mutual coherence, these results are inapplicable for the small sample (matrix) sizes in CSMRI – especially with high subsampling. Existing methods for sampling pattern selection in CSMRI may therefore have room for improvement.

In this paper, we focus on the design of adaptive sampling patterns using training image scans. Such sampling patterns capture the underlying structure in k-space to provide superior reconstructions for CSMRI in test scans. Our framework for sampling design also involves image reconstruction as one of its components. While the proposed sampling design algorithm is general, we choose here a recently proposed [3] adaptive reconstruction framework described in the next section.

II. Adaptive Sparsifying Dictionaries for MRI

Numerous transforms, either separately or in combination, have been tested for CSMRI such as wavelets, finite differences [1], and contourlets [4]. However, CSMRI reconstructions obtained with non-adaptive sparsifying transforms, typically suffer from many artifacts at higher undersampling factors ($>$ 3 fold) [5]. We recently proposed the idea (DLMRI) of simultaneously learning an image patch-based dictionary and reconstructing the image using only the undersampled k-space data [3]. The optimization problem

This work was supported in part by the National Science Foundation (NSF) under grants CCF 06-35234 and CCF 10-18660.

S. Ravishankar and Y. Bresler are with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, IL 61801, USA (e-mail: ravisha3, ybresler@illinois.edu).

proposed there is

$$(P0)\ \min_{x,D,\mathscr{A}} \sum_{ij} \left\|R_{ij}x - D\alpha_{ij}\right\|_2^2 + \nu \left\|F_u x - y\right\|_2^2 \qquad (2)$$
$$s.t.\ \ \left\|\alpha_{ij}\right\|_0 \leq T_0\ \forall\, i,j.$$

Operator $R_{ij} \in \mathbb{C}^{n\times P}$ extracts (as a vector) the 2D image patch x_{ij} of size $\sqrt{n}\times\sqrt{n}$ pixels from x as $x_{ij} = R_{ij}x$, where (i,j) indexes the top-left corner of the patch in the 2D image; $D \in \mathbb{C}^{n\times K}$ is the image patch-based complex dictionary and $\alpha_{ij} \in \mathbb{C}^K$ is the sparse representation of the patch in D with at most T_0 non-zeros; $\mathscr{A}$ denotes the set $\{\alpha_{ij}\}_{ij}$; and ν is a weight that depends on the measurement noise level (σ) as $\nu = \frac{\lambda}{\sigma}$ with λ being a positive constant. The formulation enforces both k-space data fidelity and sparsity of the overlapping patches of the reconstructed image.

An iterative alternating algorithm is used to solve (P0). Significant improvements in reconstruction quality were observed [3] with such an adaptive framework as compared to non-adaptive CSMRI methods such as those of Lustig et al. [1]. The adaptive reconstruction framework also significantly improved the undersampling limit achievable with CSMRI (almost 2.5-4 times higher undersampling factors than existing CSMRI methods at comparable reconstruction errors [3]). However, the issue of optimal sampling was not discussed in that work. This is the topic of this paper.

III. Design of the Sampling Pattern

We work with a fully sampled training image scan(s) for this section. The goal is to choose a sampling pattern in k-space that gives the best reconstruction of the training image(s) at a given undersampling factor (M). The entire k-space of the training image(s) is partitioned into J *cells*. Examples of such cells for Cartesian and 2D pseudo random sampling are shown in Figure 1. The total number of sample points is kept fixed, but they can be re-distributed by moving a sample point from one cell to another. The cost function that we optimize is

$$(P1)\ \min_{S_M} \max_j \sum_{i=1}^{N} \left\| \left(y_j^i - H_j\left(S_M y^i\right)\right) \odot W_j \right\|_2^2 \qquad (3)$$

where y^i represents the k-space values of the i^{th} reference image (N references are assumed), S_M is the undersampling mask in k-space at the undersampling factor M (i.e. $S_M y^i$ represents the undersampled k-space measurements), H represents the reconstruction method for obtaining the full k-space from the undersampled one, and W is a weighting function for k-space. The subscript j is used to index the values in the j^{th} k-space cell C'_j, so that $y_j^i \in \mathbb{C}^{|C'_j|}$, where $|C'_j|$ denotes the number of pixels in C'_j. The problem (P1) thus minimizes the maximum (weighted) reconstruction error in the k-space cells for the reference image(s).

Recently, Seeger et al. [6] also proposed optimization of k-space sampling for CSMRI. The optimization there was done sequentially on a single sagittal brain slice using information gain as the criterion, and the resulting sampling pattern was tested on other test data using the reconstruction strategy

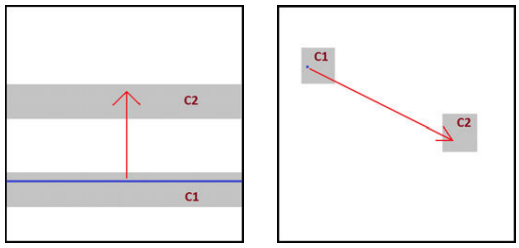

Fig. 1. Sampling Design: Two cells, C1 and C2, are shown in k-space for Cartesian sampling (left) and 2D pseudo random sampling (right). The arrows show the phase encode/sample point in C1 being moved to another location in C2.

of Lustig et al. [1]. However, as opposed to their work, we optimize directly the errors in k-space and our cost function is adapted to both the training data and the reconstruction strategy. We also work at higher undersampling factors than [6] and do not apriori dedicate samples for the k-space center.

Our problem formulation (P1) of finding the optimal undersampling pattern S_M is combinatorial and NP-hard. We propose an approximate algorithm for its solution. In order to optimize the cost function in (P1), we start with an initial pattern, and iteratively modify it based on the quality of the reconstruction it provides in k-space. A key idea that distinguishes the proposed approach, is that the reconstruction results in each step provide not only a measure of the effectiveness of the sampling pattern used, but are also employed to prescribe *how* the pattern should be modified to improve the results.

A. Algorithm

Our algorithm alternates between two steps - reconstruction of reference image(s) with a fixed sampling pattern (reconstruction update step), and update of the sampling pattern given the reconstructions (sampling update step). In the following, we will first work with the case $N = 1$ (single reference) for simplicity, and then generalize to the multiple reference/training images case.

The training image is first reconstructed [3] from the initial undersampling pattern. This reconstruction is then transformed to k-space using the discrete Fourier transform (DFT), producing the reconstructed k-space data, which is then partitioned into cells similarly to the training k-space. In each of the cells, we compute the l_2 norm of the difference between the reconstructed and training k-space data. This value is normalized (weighted) by the p-th power of the peak magnitude of the training k-space data in the cell (i.e., $W_j = \frac{1}{\|y_j\|_\infty^p}$ is constant over the k-space locations within the cell), producing the *total cell error*.

The total errors for the various cells are then sorted in increasing order and the cells corresponding to the top L error values are chosen and modified (improved) during the sample re-distribution process (sampling update step). These are the "bad" cells that require more samples to reduce their errors (choosing $L > 1$ allows more than 1 cell to be improved, while also reducing the cost in (3)). Denote the cells sorted in ascending order according to their total errors

by $\{C_j\}_{j=1}^{J}$, with the corresponding total errors $\{E_j\}_{j=1}^{J}$. Let $e \triangleq sE_{J-L+1}$, where s is a fixed parameter. Thus, e is a measure of the total error in the "best" of the L "bad" cells.

The k-space samples are re-distributed by moving samples from low-error cells to the top L cells. The top L cells are handled sequentially beginning with cell C_J. The non-sampled point (at location P_1) in cell C_J of the reconstructed k-space that has the highest point-wise error is chosen as the candidate to be sampled. The sampled point P_2 from C_1 of the reconstructed k-space that has the lowest point-wise error (note that this error need not be zero as data fidelity may not be exactly enforced during reconstruction) is moved to P_1. The value of the reconstructed k-space at P_1 is then set equal to the training value at that point, and at P_2 the value is set to zero. Sample points are sequentially added to C_J at non-sampled locations preferenced by the point-wise reconstruction error (higher errors first). This is done until the total error of C_J falls below e. Furthermore, points to be added to C_J are taken from sampled locations of C_1 preferenced by their point-wise reconstruction error (lower errors first). This is done as long as the total error of the reconstructed k-space in C_1 remains below e. Once that condition is violated, sample points are instead chosen for removal from the next cell, i.e. C_2 and so on. Thus, sample points are sequentially moved from the cells with lower total errors to the top L cells (starting with C_J, then C_{J-1} and so on) until the latter have total errors less than e.

It is possible that the total error of a cell (among the top L cells) may fail to diminish below e. This can happen either if we run out of sampled points due to saturation of errors (near e) in all the low error cells or if there are no more unsampled points left in the high error cell in which case the error there is due to the reconstruction method not enforcing 'exact' data consistency. Once the sample re-distribution process (sampling update step) is complete, the image is reconstructed using the new sampling pattern.

The sampling design iterates over the two steps of reconstruction and sample/phase encode re-distribution. Different values of the power factor p can result in different types of re-distribution (corresponding to different weighting functions W). The value $p = 0$ implies constant weighting (of 1 for all cells) on the k-space reconstruction error, which generally implies that cells (with less samples) near the center of k-space will get higher preference (more likely to be in the top L bad cells) due to higher signal energy concentration near the k-space center. Thus, the predominant movement of sampled points in this case would be towards the center of k-space. Higher values of p result in more general re-distribution of sampled points. While the algorithm is outlined for a single training scan, it can be easily extended to the case of multiple training images (of same size and scan parameters) by working with cumulative (over the training set) errors of k-space cells in the sampling update step.

IV. Numerical Experiments

We performed simulations to test the performance of our sampling framework. The training and test images (512×512 complex MRI scans kindly provided by Prof. Michael Lustig, UC Berkeley) used in our simulations are shown in Fig. 2. The training image was chosen as one slice (with rich features) of a multi-slice data acquisition. In the numerical experiments, we work with simulated k-space data that are obtained by 2D DFT of the complex MR images. The undersampling patterns designed with the training image were tested on other slices of the multi-slice acquisition as well as on a test image from a different scan (Top right of Fig. 2). 2D pseudo random sampling (Fig. 3) and Cartesian sampling (Fig. 7) schemes are used. The parameters of the DLMRI algorithm [3] were set as $n = 36, K = n, T_0 = 7, \lambda = 140$. The k-space of the training slice is shown in Fig. 7.

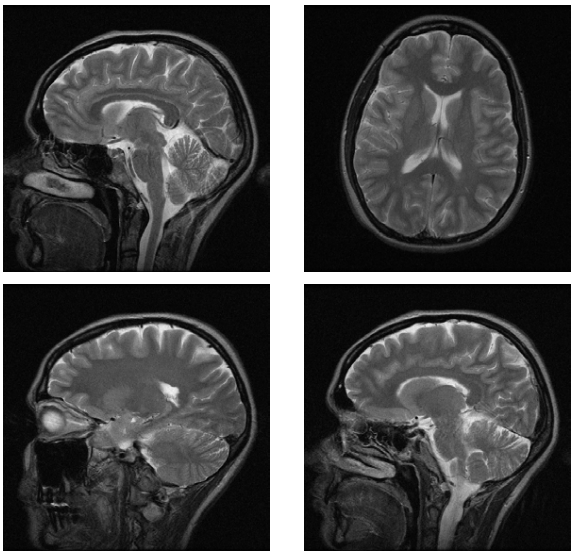

Fig. 2. Training image (Top left) and test images

In Fig. 3, 5.3 fold undersampling is employed on the k-space of the training slice. The parameters for the sampling design are set as $J = 16384, p = 0.25, s = 0.74, L = 52$, and cell size of 4×4. Five iterations of the algorithm are executed (with no sampling update step in the last iteration) and the initial (variable density random pattern) and final sampling patterns are shown. Based on the locations at which samples were added/removed (Fig. 3), it can be inferred that the algorithm captures the underlying k-space structure. The peak signal to noise ratio (PSNR) in dB computed as the ratio of the peak intensity value of the original image to the root mean square (rms) reconstruction (from the undersampling pattern) error for that image (rms error computed between image magnitudes) is plotted over algorithm iterations for the training image reconstruction. The PSNR improves by 7 dB with iterations indicating that our sampling design algorithm leads to better sampling patterns. The PSNR also converges quickly indicating fast algorithmic convergence.

The training image reconstructions with the initial and final undersampling patterns shown in Fig. 4 depict the improvement in reconstruction error. The reconstruction error magnitudes (computed as $\left| |\widehat{I}| - |I| \right|$, where I is the original image and $\widehat{I}$ is the reconstruction) also show errors of

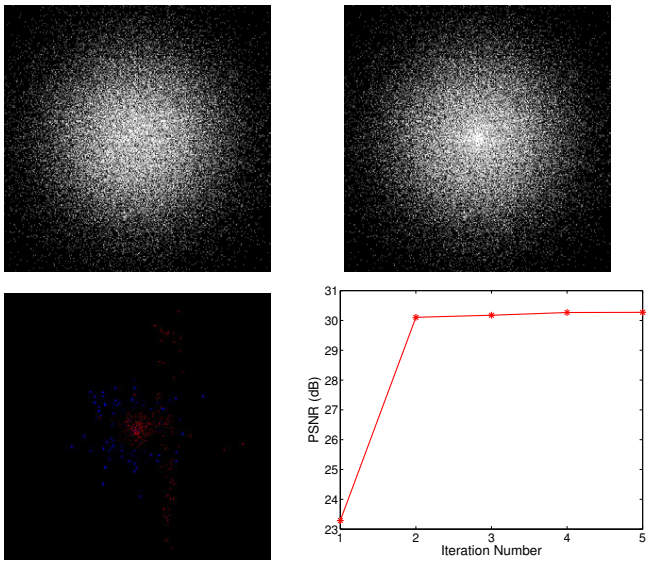

Fig. 3. 2D random sampling at 5.3 fold undersampling. Top: Initial sampling pattern (left), final sampling pattern (right). Bottom: k-space locations (compare final and initial patterns) where samples were added (in red) and removed (in blue) after 5 iterations (left), plot of training image reconstruction PSNR over iterations (right) beginning with initial reconstruction.

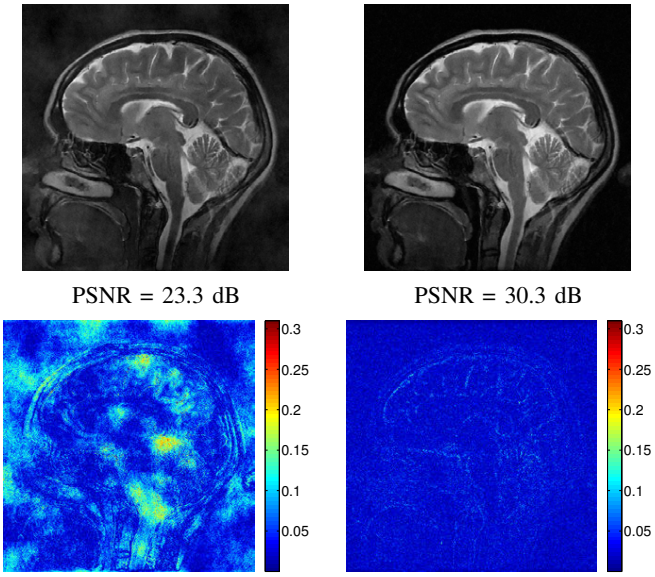

Fig. 4. Top: Training image reconstruction with initial sampling pattern (left) and final sampling pattern (right). Bottom: Reconstruction error magnitudes for Top row images.

much smaller magnitude and structure for the final sampling pattern compared to the initial one. The initial and final undersampling patterns are also tested on the test images. The test reconstructions and error maps shown in Figs. 5 and 6 show upto 5.5 dB improvement in reconstruction PSNR with the adapted sampling pattern compared to the initial pattern. The significant improvements on a variety of test images indicate the promise of our adaptive sampling design.

Fig. 7 employs Cartesian sampling with 4.3 fold undersampling of k-space. The parameters for sampling design

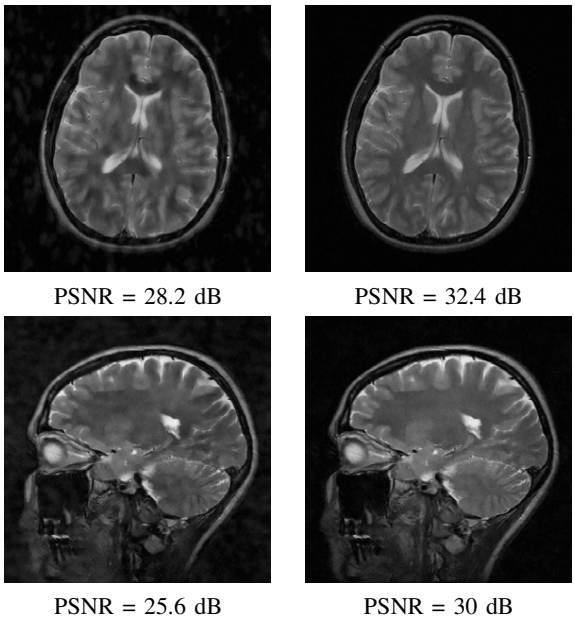

Fig. 5. Test image reconstructions with initial sampling pattern (left) and final pattern (right).

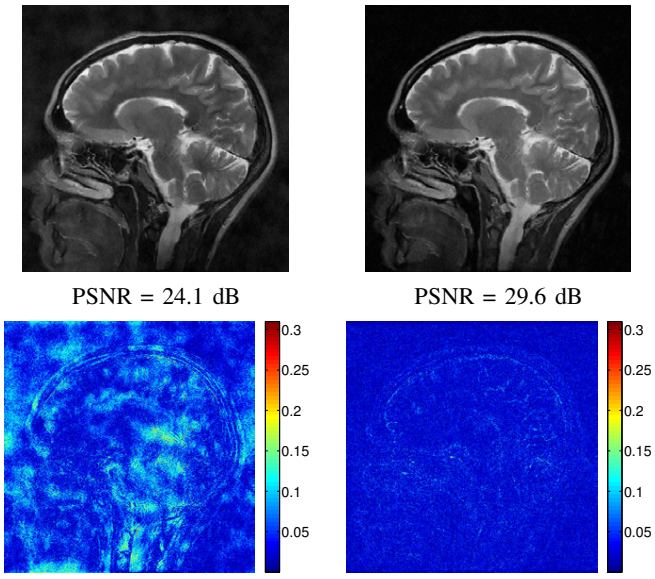

Fig. 6. Top: Test image reconstructions with initial sampling pattern (left) and final pattern (right). Bottom: Reconstruction error magnitudes for top row images.

are set as $J = 51, p = 0.04, s = 0.74, L = 3$, and cell size of 10 in the phase-encoding direction. Five iterations of the algorithm are executed. The final undersampling pattern is shown along with the initial pattern. Since $p = 0.04$ is small, sample points move more towards the center of k-space. The PSNR of the training slice reconstruction is plotted over algorithm iterations. It increases by nearly 6.3 dB indicating good improvements on training data. The initial and adapted undersampling patterns are also tested on the test images. The test reconstructions and error maps shown in Fig. 8 show upto 6.6 dB improvement in reconstruction PSNR with

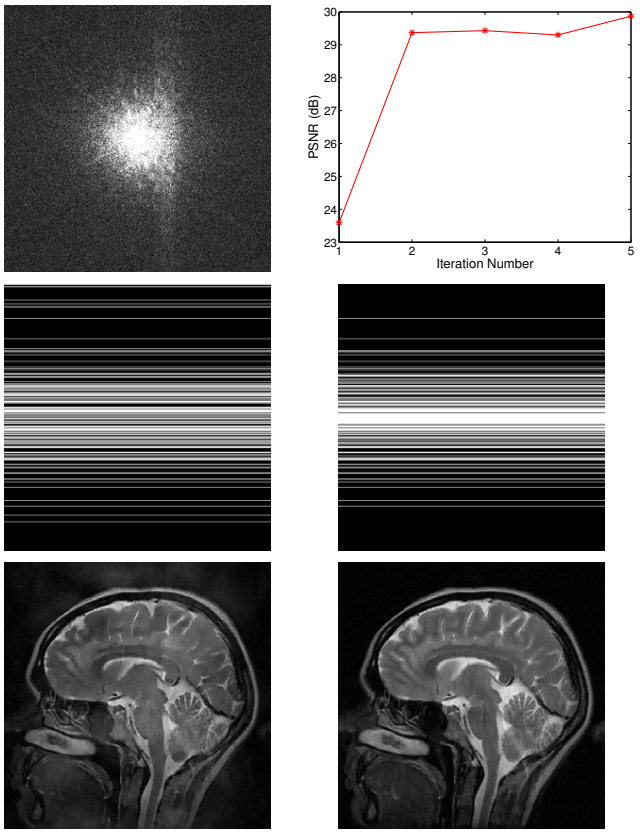

Fig. 7. Cartesian sampling at 4.3 fold undersampling. Top: Training image k-space (left), plot of training image reconstruction PSNR over iterations beginning with initial reconstruction (right). Middle: Initial sampling pattern (left), final pattern (right). Bottom: Training image reconstruction with initial sampling pattern (left) and final sampling pattern (right).

the adapted sampling pattern compared to the initial one. The promising improvements in reconstruction performance shown on both training and test data indicate the superior performance of adaptive sampling design.

V. Conclusions

In this paper, we introduced an adaptive sampling framework for CSMRI. This framework was also combined with an adaptive reconstruction framework that learns the sparsifying dictionary directly from the sampled k-space data. The iterative algorithm for sampling design utilizes fully sampled training image scans to adapt an initial undersampling pattern. The k-space errors of the image reconstructed from the undersampled k-space data are reduced in each iteration. Significant improvements in reconstruction PSNR were observed in both training and test images when using the adapted sampling pattern compared to the initial pattern. The proposed framework for sampling design is generic and can be combined with any reconstruction strategy. A more detailed study of the parameters involved in sampling design and a comparison to the work of Seeger et al. [6] will be presented elsewhere. We also plan to study the performance of alternative choices for the weighting function (W).

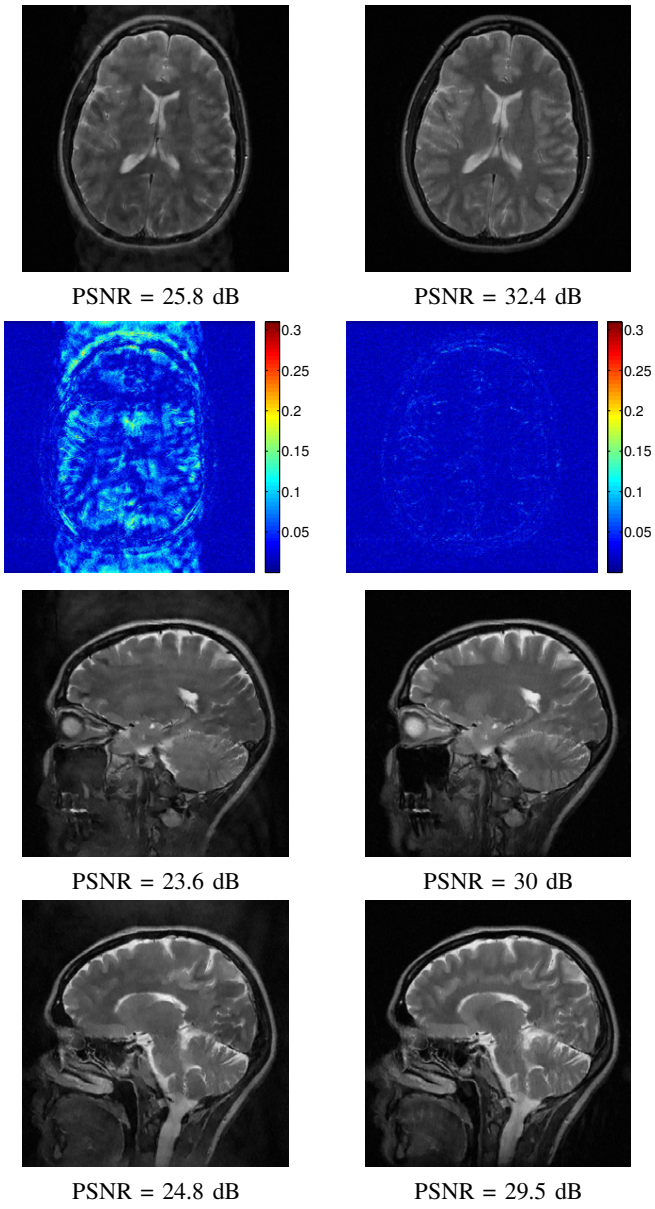

Fig. 8. Test image reconstructions with initial sampling pattern (left) and final pattern (right). Second row: Reconstruction error magnitudes for first row images.

References

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

[2] H. Wang, D. Liang, and L. Ying, "Pseudo 2D random sampling for compressed sensing mri," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2009, pp. 2672–2675.

[3] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, 2011.

[4] X. Qu, X. Cao, D. Guo, C. Hu, and Z. Chen, "Combined sparsifying transforms for compressed sensing MRI," *Electronics Letters*, vol. 46, no. 2, pp. 121–123, 2010.

[5] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Proc. IEEE Int. Conf. Comp. Vis. and Pat. Rec.*, 2008, pp. 1–8.

[6] M. Seeger, H. Nickisch, R. Pohmann, and B. Schlkopf, "Optimization of k-space trajectories for compressed sensing by bayesian experimental design," *Magn. Reson. Med.*, vol. 63, no. 1, pp. 116–126, 2010.