

Rethinking Sampling in Parallel MRI: A Data-Driven Approach

Baran Gözcü, Thomas Sanchez, Volkan Cevher

Laboratory for Information and Inference Systems (LIONS), EPFL, Switzerland

{baran.goezcue, thomas.sanchez, volkan.cevher}@epfl.ch

Abstract—In the last decade, Compressive Sensing (CS) has emerged as the most promising, model-driven approach to accelerate MRI scans. CS relies on the key sparsity assumption and proposes random sampling for data acquisition. The practical CS approaches in MRI employ variable-density (VD) sampling, where samples are drawn at random based on a parametric probability model which focuses on the center of the Fourier domain. In stark contrast to this model-driven sampling approaches, we propose a data-driven framework for optimizing sampling in parallel (multi-coil) MRI. Our approach does not assume any structure in the data, and instead optimizes a performance metric (e.g. PSNR) for any given reconstruction algorithm, based on our earlier learning-based sampling framework previously applied to 2D MRI which we also extend to 3D MRI setting in this work by employing lazy evaluations in the greedy algorithm. We show boosted performance for the parallel MRI based on this sampling approach and highlight the inefficiency of variable density approaches. This suggests that data-driven sampling methods could be the key to unlocking the full power of CS applied to MRI.

Index Terms—Parallel MRI, compressive sensing, learning-based subsampling, greedy algorithm

I. INTRODUCTION

Imaging speed has historically been an important criterion in many MRI applications. Due to many physical and physiological limitations in the acquisition, much effort has been devoted to the design of reconstruction methods that use less Fourier data without sacrificing image quality. This has recently allowed the rapid rise of Compressed Sensing (CS) [1], which has been successfully combined with parallel MRI to produce better quality images at lower scan times [2].

Traditional compressive MRI has been *model-driven*, in the sense that it assumes models such as sparsity or structured sparsity in some basis and proposes reconstruction algorithms such as basis pursuit exploiting these models along with theoretical performance guarantees [1], [3]. As for practical sampling, variable density approaches [4], [5] have dominated the arena where one tunes the parameters of the model. We will show that this itself is a limiting factor for a better reconstruction performance.

In the area of machine learning, *data-driven* approaches outperform the classical *model-driven* methods for various applications. In the light of this paradigm shift and based

on our prior work, we propose a data-driven framework for sampling in the parallel MRI setting which is more relevant for clinical applications [6]. In our approach, we employ a combinatorial algorithm that lets the data decide sampling masks matched to specific parallel MRI decoders in use. Our masks are structured, with a certain amount of low frequency content as well as a more uniform spread on higher frequencies which in turn manages to significantly reduce the artifacts seen in the variable density approach. Moreover, the generalization of our masks is supported by statistical learning theory.

Our contributions in this work are two-fold: We extend our learning-based framework [7] to the multicoil MRI setting and show that this provides reliable and boosted performance compared to variable density approaches. As a second contribution, we also consider 3D MRI setting for which we propose using lazy greedy algorithm which successfully tackles the combinatorial problem that is harder since one can subsample in both phase and frequency encoding directions in 3D MRI.

II. MODEL-DRIVEN COMPRESSED SENSING

In the *compressive sensing* (CS) problem [1], [8], one seeks to recover a sparse vector via a small number of linear measurements. In the special case of compressive parallel MRI, these measurements take the specific form of subsampled Fourier measurements and are described as follows:

$$\mathbf{b}_j = \mathbf{P}_\Omega \Psi \mathbf{S}_j \mathbf{x} + \mathbf{w}_j, \quad (1)$$

where $\Psi \in \mathbb{C}^{p \times p}$ is the Fourier transform operator applied to the vectorized image $\mathbf{x} \in \mathbb{C}^p$, $\mathbf{P}_\Omega : \mathbb{C}^p \rightarrow \mathbb{C}^n$ is a subsampling operator that selects the rows of Ψ indexed by the set Ω , with $|\Omega| = n$. We refer to Ω as the *sampling pattern* or *mask*. j ranges from 1 to C and denotes the measurement on each coil with its corresponding sensitivity $\mathbf{S}_j \in \mathbb{R}^{p \times p}$. Here, $\mathbf{w}_j \in \mathbb{C}^n$ is additive noise, independent on each coil. This setting describes parallel measurements acquired on different coils for the same underlying image \mathbf{x} .

Given the vectorized measurements $\mathbf{b} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_C]^T$, $\mathbf{b} \in \mathbb{C}^{nC}$ (along with knowledge of Ω), a *reconstruction algorithm* (also referred to as a *decoder*) forms an estimate $\hat{\mathbf{x}}$ of \mathbf{x} . This algorithm is treated as a general function, and is written as $\hat{\mathbf{x}} = g(\Omega, \mathbf{b})$. A wide variety of decoding techniques have been proposed for compressive MRI; hereafter, we present a few of the most widely-used and best-performing techniques, for both the single-coil and multi-coil settings.

This work is supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement number 725594 - time-data), by Hasler Foundation Program: Cyber Human Systems (project number 16066) and by the Department of the Navy, Office of Naval Research (ONR) under a grant number N62909-17-1-2111.

A. Single-coil and multi-coil reconstruction algorithms

In the general CS problem, decoders based on convex optimization have received considerable attention, both due to their theoretical guarantees and practical performance. The formulation of a single-coil problem is same as Equation (1), the only difference being $C = 1$, and \mathbf{S}_1 equals $p \times p$ identity matrix. In the noiseless setting (i.e., $\mathbf{w} = 0$), a notable choice is *basis pursuit* (BP) [1] which minimizes ℓ_1 norm of the image in a sparse domain. A similar type of convex optimization formulation is *total variation* (TV) minimization.

SENSE is a method that requires the knowledge of coil sensitivities to retrieve an unaliased image from parallel acquisitions [9]. We use its implementation in the BART toolbox [10], based on FISTA [11] to iteratively solve the following:

$$\min_{\mathbf{z}} \|\mathbf{b} - \mathbf{P}_\Omega \Psi \mathbf{S} \mathbf{z}\|_2^2 + \lambda \|\Phi \mathbf{z}\|_1 \quad (2)$$

where $\mathbf{b} \in \mathbb{C}^{nC}$ is the coil-wise subsampled Fourier measurements, $\mathbf{S} = [\mathbf{S}_1 \dots \mathbf{S}_C]^T \in \mathbb{R}^{Cp \times p}$ are the stacked coil sensitivities, $\mathbf{P}_\Omega \Psi$ is the subsampled Fourier transform applied coil-wise to the product of the image with the stacked coil sensitivities and Φ is a sparsifying operator. Recently, annihilating filter based low-rank Hankel matrix approach (ALOHA) has been also shown to be effective as a parallel MRI algorithm that exploits low-rankness property [2]. Other multicoil algorithms include [12]–[15].

B. Sampling mask design

Generally speaking, most popular approaches for designing Ω for compressive MRI make use of *random variable-density sampling* according to a non-uniform probability distribution in order to minimize the coherence [4]. Such approaches have notable limitations: They typically require parameters to be tuned and it is generally unclear which particular sampling distribution will be most effective for a given decoding rule and anatomy. We focus on Cartesian masks in this work, however we also refer to recent non-Cartesian sampling schemes proposed in [16]–[18].

Recently, alternative design methods have used fully sampled training images for sampling optimization. In [19]–[21], the training data is used to construct a sampling distribution, from which the samples are then drawn randomly. In [22], [23], a single training image is used at a time to choose a phase encode to sample. Much like the above-mentioned variable-density sampling approaches, these existing adaptive algorithms assume sampling models or contains parameters for their mask selection whose tuning is non-trivial. Moreover, to our knowledge, none of these works have provided theoretical justifications of the mask selection method.

III. DATA-DRIVEN COMPRESSED SENSING

We follow our learning-based framework of [7], and we summarize our main ideas hereafter. We first assume that we have a set of training samples $x_1, \dots, x_m \in \mathbb{C}^p$ representative of an unknown test signal of interest x coming from the same underlying probability distribution. We wish to find the sampling pattern Ω that will perform best on x , according to

a performance metric $\eta_\Omega(x, \hat{x})$, and given constraints on the design of the mask. It is desirable to constrain Ω to contain for instance only horizontal and/or vertical lines in k -space. To formalize this, we define a set \mathcal{S} of subsets of $\{1, \dots, p\}$ which typically contains sets of phase encodes in the k -space. The final mask then takes the form $\Omega = \bigcup_{j=1}^{\ell} S_j$, $S_j \in \mathcal{S}$, for some $\ell > 0$. We assume that there exists a *cost function* $c(\Omega) > 0$ associated with each sampling pattern, and that the final cost must satisfy $c(\Omega) \leq \Gamma$ for some $\Gamma > 0$, and we consider here the case where the cost is the total number of indices in Ω . This then leads to the following problem

$$\hat{\Omega} = \arg \max_{\Omega \in \mathcal{A}} \frac{1}{m} \sum_{j=1}^m \eta_\Omega(x_j, \hat{x}_j). \quad (3)$$

where \mathcal{A} is the set of feasible Ω 's according to \mathcal{S} . Given that enough training samples are considered, statistical learning theory guarantees that the result obtained using the empirical average in Equation 3 will perform well on x [7].

A. Learning-based variable density

An intuitive approach given in [7] is to find an approximate solution to the combinatorial problem (3) by generating a candidate masks $\Omega_1, \dots, \Omega_L$ using parametric variable-density methods, and choose the one with the best empirical performance on the training set (Algorithm 1). While similar ideas have already been used when performing parameter sweeps in existing works (e.g., see [19]), our framework justifies why the empirical performance is the correct quantity to optimize.

Algorithm 1 Learning-Based variable density (LB-VD)

Input: Training data $\mathbf{x}_1, \dots, \mathbf{x}_m$, decoder g , candidate masks $\Omega_1, \dots, \Omega_L$

Output: Sampling pattern Ω

- 1: **for** $\ell = 1, \dots, L$ **do**
 - 2: For each j , set $\mathbf{b}_j \leftarrow \mathbf{P}_{\Omega_\ell} \Psi \mathbf{x}_j$, $\hat{\mathbf{x}}_j \leftarrow g(\Omega_\ell, \mathbf{b}_j)$
 - 3: $\eta_\ell \leftarrow \frac{1}{m} \sum_{j=1}^m \eta(\mathbf{x}_j, \hat{\mathbf{x}}_j)$
 - 4: $\Omega \leftarrow \Omega_{\ell^*}$, where $\ell^* = \arg \max_{\ell=1, \dots, L} \eta_\ell$
 - 5: **return** Ω
-

B. Learning-based greedy mask optimization

In [7], a greedy algorithm is used to find an approximate solution to (3) for single-coil reconstructions. We start from a base mask Ω , and at each iteration, we create candidate masks $\Omega' = \Omega \cup S$ with each subset $S \in \mathcal{S}$ not yet included in the mask Ω . After reconstructing the image \hat{x} in the training dataset using the decoder g for all the Ω' , the subset S that yields the larger increase in the performance metric η is permanently added to the mask Ω and this procedure is repeated while the cost constraint Γ is not violated. As it will be seen in Section IV, this algorithm can seamlessly optimize the sampling in the case of parallel MRI as well.

C. Learning-based lazy greedy algorithm

Algorithm 2 is infeasible in terms running time in the case 3D MRI where one is allowed to sample in both phase and frequency encoding directions. A faster stochastic approach was proposed in [24] for dynamic MRI, but cannot be readily applied to the 3D setting, as it would require unreasonably large

Algorithm 2 Learning-based greedy (LB)

Input: Training data $\mathbf{x}_1, \dots, \mathbf{x}_m$, decoder g , sampling subsets \mathcal{S} , cost function c , maximum cost Γ

Output: Sampling pattern Ω

- 1: $\Omega \leftarrow \emptyset$
 - 2: **while** $c(\Omega) \leq \Gamma$ **do**
 - 3: **for** $S \in \mathcal{S}$ such that $c(\Omega \cup S) \leq \Gamma$ **do**
 - 4: $\Omega' = \Omega \cup S$
 - 5: For each j , set $\mathbf{b}_j \leftarrow \mathbf{P}_{\Omega'} \Psi \mathbf{x}_j$, $\hat{\mathbf{x}}_j \leftarrow g(\Omega', \mathbf{b}_j)$
 - 6: $\eta(\Omega') \leftarrow \frac{1}{m} \sum_{j=1}^m \eta(\mathbf{x}_j, \hat{\mathbf{x}}_j)$
 - 7: $\Omega \leftarrow \Omega \cup S^*$, $S^* = \arg \max_{S: c(\Omega \cup S) \leq \Gamma} \frac{\eta(\Omega \cup S) - \eta(\Omega)}{c(\Omega \cup S) - c(\Omega)}$
 - 8: **return** Ω
-

batchs. Another way of speeding up is to use lazy evaluations [25] which is equivalent to but faster than the greedy algorithm for submodular functions (functions with diminishing return property) which possess optimization guarantees [26]. While in our present setting, verifying submodularity is difficult due to the fact that the performance of reconstructions generally cannot be written in a closed form, we were motivated by the fact that Algorithm 2 works well in practice even if the objective function is not submodular. Indeed, as we will observe in Section IV, lazy evaluations as well perform well for mask optimization.

The lazy greedy Algorithm 3 relies on the fact that if at the i -th iteration of the greedy algorithm, an element S is expected to bring a marginal benefit $\Delta(S|\Omega_i) = \frac{\eta(\Omega_i \cup S) - \eta(\Omega_i)}{c(\Omega_i \cup S) - c(\Omega_i)}$, then it holds that $\Delta(S|\Omega_j) \leq \Delta(S|\Omega_i)$ for all $j \geq i$. Exploiting this fact, one can keep a list of upper bounds on the marginal benefits of each element S , called $\rho(S)$, initialized at $+\infty$. Then, at each iteration, the element S with largest value $\rho(S)$ is picked and updated as $\rho(S) = \Delta(S|\Omega)$. If $\rho(S) \geq \rho(S') \forall S' \in \mathcal{S}$, the marginal benefit of this element S is larger than the upper bounds of each other marginal contribution, and is consequently added permanently to Ω . This speeds up the algorithm, since $\rho(S) \geq \rho(S')$ can be satisfied after a few trials, instead of trying each available S at every iteration [27].

Algorithm 3 Learning-based lazy greedy (LB-L)

Input: Training data $\mathbf{x}_1, \dots, \mathbf{x}_m$, decoder g , sampling subsets \mathcal{S} , cost function c , maximum cost Γ

Output: Sampling pattern Ω

- 1: $\Omega \leftarrow \emptyset$
 - 2: $\rho(S) \leftarrow +\infty \forall S \in \mathcal{S}$ s.t. $c(\Omega \cup S) \leq \Gamma$
 - 3: **while** $c(\Omega) \leq \Gamma$ **do**
 - 4: $\Omega' \leftarrow \Omega \cup S$, where $S = \arg \max_{S' \in \mathcal{S}: c(\Omega \cup S') \leq \Gamma} \rho(S')$
 - 5: For each j , set $\mathbf{b}_j \leftarrow \mathbf{P}_{\Omega'} \Psi \mathbf{x}_j$, $\hat{\mathbf{x}}_j \leftarrow g(\Omega', \mathbf{b}_j)$
 - 6: $\eta(\Omega') \leftarrow \frac{1}{m} \sum_{j=1}^m \eta(\mathbf{x}_j, \hat{\mathbf{x}}_j)$
 - 7: $\rho(S) \leftarrow \frac{\eta(\Omega') - \eta(\Omega)}{c(\Omega') - c(\Omega)}$
 - 8: **if** $\rho(S) \geq \rho(S') \forall S' \in \mathcal{S}$ s.t. $c(\Omega \cup S') \leq \Gamma$ **then**
 - 9: $\Omega = \Omega \cup S$
 - 10: **return** Ω
-

IV. NUMERICAL EXPERIMENTS

A. Implementation Details

We consider the decoders described in Section II: BP, TV, SENSE, and ALOHA. For BP, we let the sparsifying transform to be the shearlet transform [28]. For SENSE, we used ESPIRiT method to estimate coil sensitivities [13]. The MRI data used in the following subsection was acquired on a 3T MRI system (SIEMENS). The protocols were approved by the local ethics committee. The data set consists of 2D T1-weighted brain scans of seven healthy subjects, which were scanned with a FLASH pulse sequence and a 12-channel receive-only head coil. In our experiments, we use 4 different slices of sizes 256×256 from two subjects as training data and 100 slices from 5 subjects as test data. To speed up the computations, we compressed our 12-channel data into 4-channel. The 3D MRI dataset of knee images that is used in the last subsection is acquired by a 3T system (GE Healthcare) using FSE CUBE sequence with the proton density weighting. The number of channels is 8 and the matrix size is 320×320 . We used the data of 8 subjects as training data and remaining 12 subjects as test data. This dataset is publicly available at <http://mridata.org>. Note that once we learn a mask for a particular anatomy and scan parameters, it can be used on all subsequent scans in a clinical setting.

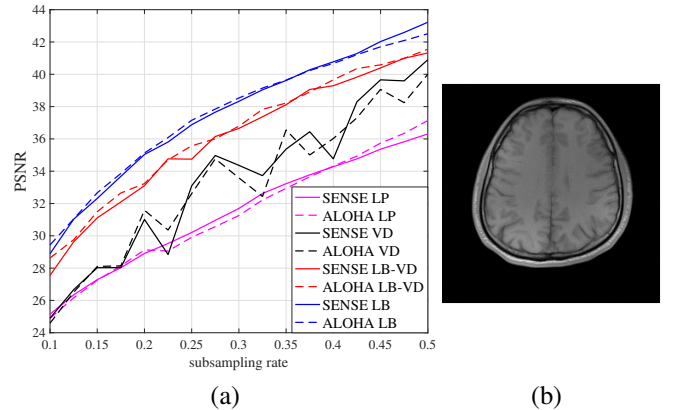


Fig. 1. (a) PSNR performances of various masks averaged over 100 test images: LP (low-pass mask), VD (variable density), LB-VD and LB are masks produced by Algorithm 1 and 2 respectively. (b) Ground truth image for the reconstructions given in Figure 2.

B. Optimal masks for multi-coil 2D MRI setting

In this section we compare the performances of various mask selection methods. As it can be seen in Figure 1(a) for a range of subsampling rates, the coherence-based VD masks can perform better than low-pass (LP) mask. However, they have weaker performance compared to the parameter sweep with PSNR as the objective function (Algorithm 1 - LB-VD). This is expected as the LB-VD masks exploit the knowledge of the training set, whereas the coherence-based VD masks do not. The best performance is provided by the greedy Algorithm 2 (LB). Indeed, the greedy LB algorithm is free from the constraint of a parametric distribution, which allows the data to decide which mask is best suited to a given decoder. Also, the VD-based masks result in artifacts especially visible in the ALOHA reconstructions that are suppressed in the LB-case.

TABLE I
CROSS PERFORMANCE OF OPTIMAL SINGLE AND MULTI-COIL MASKS AT
20% SUBSAMPLING RATE AVERAGED OVER 100 TEST SLICES.

Decoder		Single-coil		Multi-coil	
Mask	Metric	TV	BP	SENSE	ALOHA
	TV LB		32.82	34.23	34.56
BP LB		32.90	34.27	35.02	35.07
SENSE LB		31.51	31.82	36.88	36.61
ALOHA LB		29.00	26.95	30.73	37.15

In Figure 2, we have also included optimal single coil masks (TV and BP) which result in inferior performance when used with the multi-coil SENSE and ALOHA reconstruction algorithms. Compared to these single coil masks, the optimal multi-coils masks are found to be more spread in k-space. As expected, one obtains the best performances in terms of visual quality and performance metrics when the LB-mask is paired with the reconstruction algorithm for which it was trained. In Table I, we provide an average cross-performance of single/multi coils masks and reconstruction algorithms. We see that it is important to optimize single and multi-coil masks separately as they are quite different from each other, which was also observed by [29] recently.

C. Optimal masks for multi-coil 3D MRI setting

In this section we demonstrate the effectiveness of Algorithm 3 in providing masks for 3D MRI by comparing it to common 3D MRI subsampling masks: Controlled Aliasing in Parallel Imaging Results in Higher Acceleration (CAPIR-INHA) [30], Poisson disc (PD) sampling and its variable density variant (VD-PD) [31], and adaptive random sampling method [19]. For brevity we used only SENSE algorithm as the multi-coil reconstruction algorithm and for all masks we took the central 24×24 region as calibration region to estimate the coil sensitivities. As seen in Figure 3, the mask (LB-L) given by Algorithm 3 provides superior image quality compared to the benchmarks and comparable result to more recent VD-PD sampling whose polynomial decay parameter is optimized on the training data using Algorithm 1 (LB-VD).

V. CONCLUSION

With this work, we contend that the conventional wisdom of variable density-based methods have been limiting the performance of CS applied to parallel MRI. We presented a data-driven approach to sampling optimization, where, for any given reconstruction algorithm, we let the training data decide on the sampling pattern without any assumption on the underlying structure of the signal. This approach was shown to outperform low-pass sampling, coherence-based and learning-based variable-density samplings across all rates considered. We also extended our framework to 3D MRI setting by employing the lazy greedy algorithm.

Following this work, we anticipate that pairing our learning-based method with recent neural network reconstruction algorithms for parallel imaging will provide an even more powerful data-driven approach to compressive MRI. We also plan to apply our framework to non-Cartesian sampling schemes.

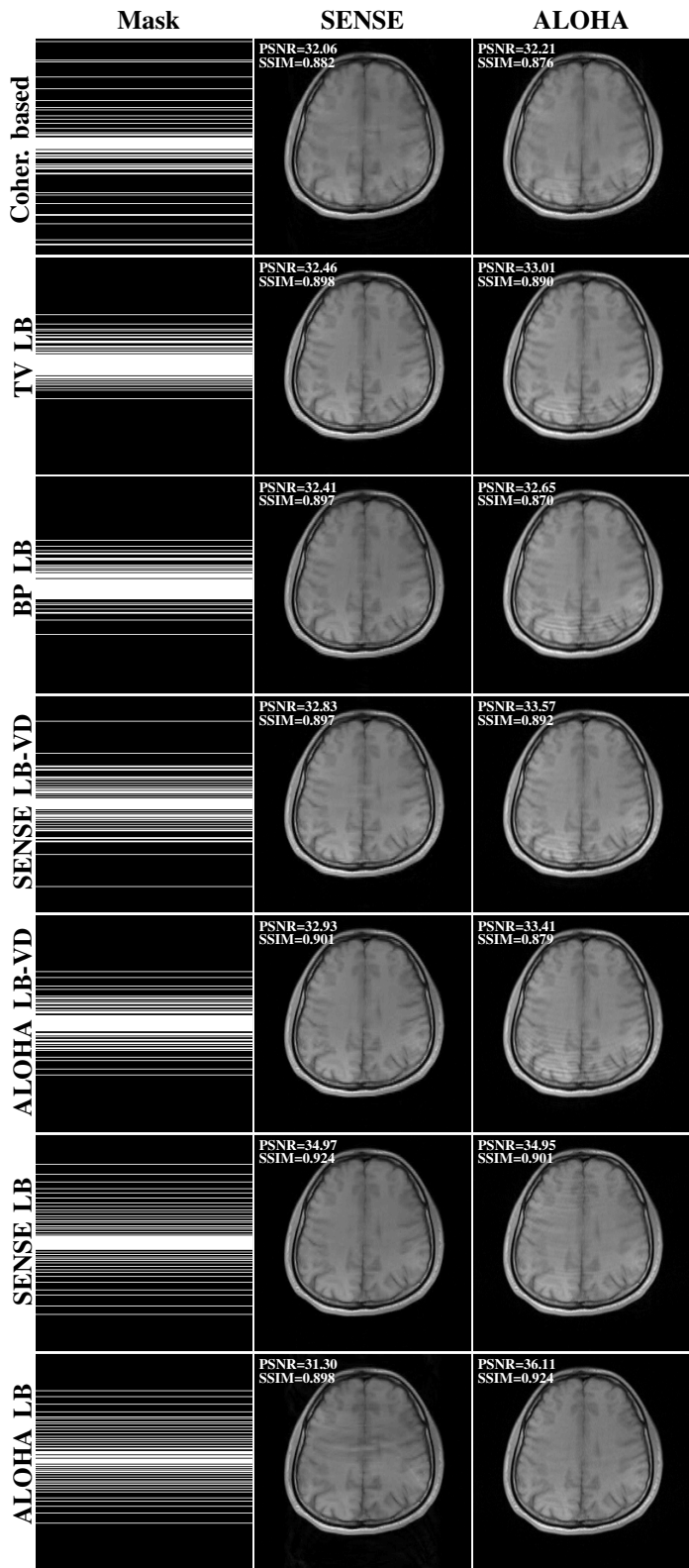


Fig. 2. Optimized masks and example reconstructions under SENSE and ALOHA decoding at 20% sampling rate. The corresponding ground truth is shown in Figure 1.

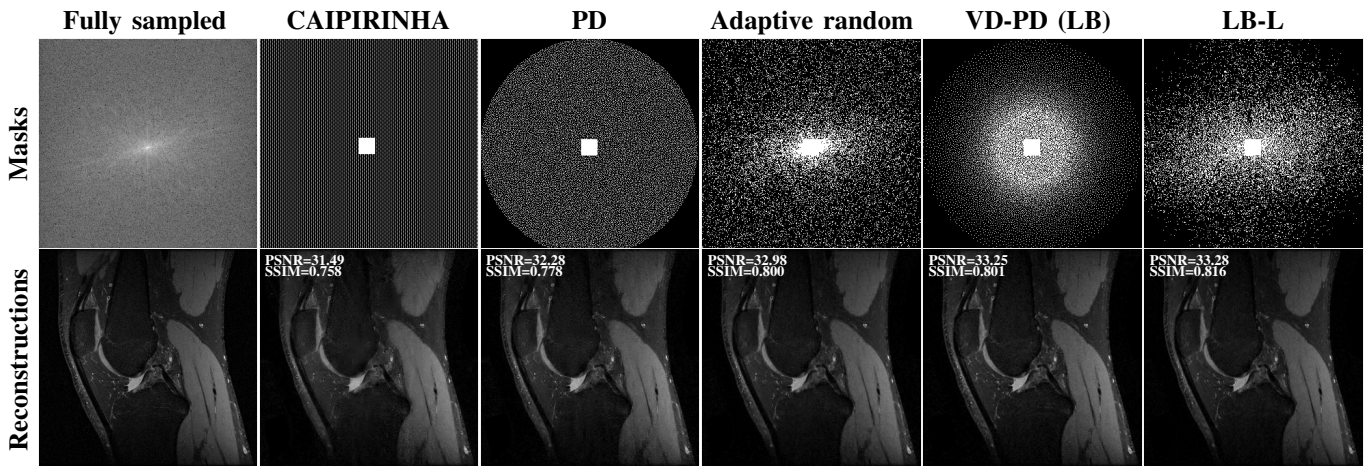


Fig. 3. Optimized masks and example reconstructions of knee images under SENSE decoding at 6-fold acceleration.

REFERENCES

- [1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.
- [2] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank hankel matrix," *IEEE Trans. Comput. Imaging*, vol. 2, no. 4, pp. 480–495, 2016.
- [3] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [4] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [5] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing," in *Forum of Mathematics, Sigma*, vol. 5, Cambridge University Press, 2017.
- [6] J. A. Fessler, "Limitations & caveats of deep learning," *ISMRM course on Deep Learning*, 2018.
- [7] B. Gözcü, R. K. Mahabadi, Y.-H. Li, E. Ilıcak, T. Cukur, J. Scarlett, and V. Cevher, "Learning-based compressive MRI," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1394–1406, 2018.
- [8] E. J. Candès *et al.*, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3, pp. 1433–1452, Madrid, Spain, 2006.
- [9] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.
- [10] M. Uecker, F. Ong, J. I. Tamir, D. Bahri, P. Virtue, J. Y. Cheng, T. Zhang, and M. Lustig, "Berkeley advanced reconstruction toolbox," in *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 23, p. 2486, 2015.
- [11] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [12] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, 2002.
- [13] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT: an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 990–1001, 2014.
- [14] P. J. Shin, P. E. Larson, M. A. Ohliger, M. Elad, J. M. Pauly, D. B. Vigneron, and M. Lustig, "Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion," *Magn. Reson. Med.*, vol. 72, no. 4, pp. 959–970, 2014.
- [15] J. P. Haldar and J. Zhuo, "P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data," *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1499–1514, 2016.
- [16] C. Boyer, N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss, "On the generation of sampling schemes for magnetic resonance imaging," *SIAM Journal on Imaging Sciences*, vol. 9, no. 4, pp. 2039–2072, 2016.
- [17] C. Lazarus, P. Weiss, N. Chauffert, F. Mauconduit, M. Bottlaender, A. Vignaud, and P. Ciuciu, "SPARKLING: Novel non-Cartesian sampling schemes for accelerated 2D anatomical imaging at 7T using compressed sensing," in *25th annual meeting of the International Society for Magnetic Resonance Imaging*, 2017.
- [18] C. Lazarus, P. Weiss, N. Chauffert, F. Mauconduit, L. El Gueddari, C. Destrieux, I. Zemmoura, A. Vignaud, and P. Ciuciu, "Variable-density k-space filling curves for accelerated magnetic resonance imaging," 2018.
- [19] F. Knoll, C. Clason, C. Diwok, and R. Stollberger, "Adapted random sampling patterns for accelerated MRI," *Magn. Reson. Mat. Phys. Biol. Med.*, vol. 24, no. 1, pp. 43–50, 2011.
- [20] Y. Zhang, B. S. Peterson, G. Ji, and Z. Dong, "Energy preserved sampling for compressed sensing MRI," *Comput. Math. Methods Med.*, vol. 2014, 2014.
- [21] J. Vellagoundar and R. R. Machireddy, "A robust adaptive sampling method for faster acquisition of MR images," *Magn. Reson. Imaging*, vol. 33, no. 5, pp. 635–643, 2015.
- [22] D.-d. Liu, D. Liang, X. Liu, and Y.-t. Zhang, "Under-sampling trajectory design for compressed sensing MRI," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 73–76, IEEE, 2012.
- [23] S. Ravishanker and Y. Bresler, "Adaptive sampling design for compressed sensing MRI," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 3751–3755, IEEE, 2011.
- [24] T. Sanchez, B. Gözcü, R. B. van Heeswijk, E. Ilıcak, T. Çukur, *et al.*, "Scalable learning-based sampling optimization for compressive dynamic mri," *arXiv preprint arXiv:1902.00386*, 2019.
- [25] M. Minoux, "Accelerated greedy algorithms for maximizing submodular set functions," in *Optimization techniques*, pp. 234–243, Springer, 1978.
- [26] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions I," *Mathematical programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [27] A. Krause and D. Golovin, "Submodular function maximization," 2014.
- [28] G. Kutyniok, W.-Q. Lim, and R. Reisenhofer, "Shearlab 3D: Faithful digital shearlet transforms based on compactly supported shearlets," *ACM Trans. Math. Softw.*, vol. 42, no. 1, p. 5, 2016.
- [29] J. P. Haldar and D. Kim, "OEDIPUS: An experiment design framework for sparsity-constrained MRI," *IEEE transactions on medical imaging*, 2019.
- [30] F. A. Breuer, M. Blaimer, M. F. Mueller, N. Seiberlich, R. M. Heidemann, M. A. Griswold, and P. M. Jakob, "Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA)," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 55, no. 3, pp. 549–556, 2006.
- [31] S. Vasanawala, M. Murphy, M. T. Alley, P. Lai, K. Keutzer, J. M. Pauly, and M. Lustig, "Practical parallel imaging compressed sensing mri: Summary of two years of experience in accelerating body MRI of pediatric patients," in *2011 IEEE international symposium on biomedical imaging: From nano to macro*, pp. 1039–1043, IEEE, 2011.