VARIABLE DENSITY COMPRESSED SENSING IN MRI. THEORETICAL VS HEURISTIC SAMPLING STRATEGIES.

Nicolas Chauffert[(1,2)], *Philippe Ciuciu*[(1,2)], *Pierre Weiss*[(3,4)]

[(1)]CEA/DSV/I²BM NeuroSpin center, Bât. 145, F-91191 Gif-sur-Yvette, France
[(2)] INRIA Saclay Ile-de-France, Parietal team, 91893 Orsay, France.
[(3)] Institut des Technologies Avances du Vivant (CNRS UMS 3039), 31106 Toulouse, France.
[(4)] Institut de Mathmatiques de Toulouse (CNRS UMR 5219), 31062 Toulouse Cedex 9, France.
nicolas.chauffert@inria.fr,philippe.ciuciu@cea.fr,pierre.armand.weiss@gmail.com

ABSTRACT

The structure of Magnetic Resonance Images (MRI) and especially their compressibility in an appropriate representation basis enables the application of the compressive sensing theory, which guarantees exact image recovery from incomplete measurements. According to recent theoretical results on the reconstruction guarantees, a near optimal strategy is to downsample the k-space using an independent drawing of the acquisition basis entries. Here, we first bring a novel answer to the synthesis problem, which amounts to deriving the optimal distribution (according to a given criterion) from which the data should be sampled. Then, given that the sparsity hypothesis is not fulfilled in the low frequency band in MRI, we extend this approach by densely sampling this center and drawing the remaining samples from the optimal distribution. We compare this theoretical approach to heuristic strategies, and show that the proposed two-stage process drastically improves reconstruction results on anatomical MRI.

Index Terms— MRI, compressive sensing, wavelets, synthesis problem, variable density random undersampling.

1. INTRODUCTION

Decreasing scanning time is a crucial issue in MRI since it could increase patient comfort, make the exam cost cheaper and improve image quality by limiting the patient's movement geometric distortions. A simple way to reduce acquisition time consists of acquiring less data samples by downsampling the k-space. Compressed Sensing (CS) theory [1, 2] gives guarantees of recovering a sparse signal from incomplete measurements in an acquisition basis. These guarantees depend on the sparsity of the image in a representation basis (e.g., wavelet basis) and on the mutual properties of the acquisition and representation bases.

Recent results [3, 4] show that a near optimal sampling strategy consists of measuring a set of coefficients drawn independently according to a distribution that depends on the sensing basis. Using this approach, the number of measurements required to guarantee exact recovery of a given sparse signal with high probability is $O(s \ln(n))$, where s is the sparsity level. Uniform reconstruction guarantees involve $O(s \ln^\alpha(n))$ bounds (with $\alpha > 1$) on the number of measurements needed [3].

In this paper, we propose an answer to the synthesis problem in the MRI framework, which amounts to finding the optimal downsampling pattern of the k-space. Following [3], we derive the optimal distribution from which the data samples are independently drawn (Section 3). Then, given that MRI images are not sparse in the low frequencies, we propose a two-stage sampling process: first, a given area of the k-space center is fully sampled in order to recover the low frequencies, and second the existing theory is applied to the remaining high frequency content of MRI images for which the sparsity assumption is more tenable (Section 4). In Section 5, we compare our method with the state-of-the-art [5], and recover very close reconstruction results, while the proposed approach more likely meets the required sparsity assumption for designing CS sampling schemes. Our results show that the proposed two-stage strategy drastically improves reconstruction results.

2. NOTATION

In this paper, we consider a discrete 2D k-space composed of n pixels, in which the MRI signal is acquired. The acquisition of the complete k-space gives access to the Fourier transform of a reference image, to which we will compare our reconstruction results. We assume that the image is sparse in a representations basis (e.g., a wavelet basis) as pointed out in [5], and we denote by x this sparse representation. Let $\Psi_1 \ldots \Psi_n \in \mathbb{C}^n$ be the wavelet atoms of an orthogonal wavelet tranform and $\boldsymbol{\Psi} = [\Psi_1, \ldots, \Psi_n] \in \mathbb{C}^{n \times n}$. The MRI image then reads $\boldsymbol{\Psi}x$. Let us introduce $\boldsymbol{F}^*$ the Fourier transform and $\boldsymbol{A}_0 = \boldsymbol{F}^* \boldsymbol{\Psi}$ the orthogonal transform between the *representation* (the wavelet basis) and the *acquisition* (the k-space) domains. $\boldsymbol{A}_0 \in \mathbb{C}^{n \times n}$ is then an orthogonal matrix. Finally, we will denote by $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ a matrix composed of m lines of $\boldsymbol{A}_0$. $\boldsymbol{A}$ is called the acquisition matrix, and m represents the number of Fourier coefficients actually measured. We define the ℓ_1 problem as the convex optimization problem that consists of finding the vector of minimal ℓ_1-norm subject to the constraint of matching the observed data $y = \boldsymbol{A}x$:

$$\underset{\boldsymbol{A}w=y}{\operatorname{argmin}} \|w\|_1. \tag{1}$$

3. OPTIMAL SAMPLING IN AN ORTHOGONAL SYSTEM

3.1. Optimal sampling distribution

First, we summarize a result recently introduced by Rauhut [3]. Let $P = (P_1, \ldots, P_n)$ be a discrete probability measure. Let us de-

fine the scalar product $\langle .,.\rangle_P$ by $\langle x,y\rangle_P = \sum_{i=1}^n x_i \bar{y}_i P_i$. Matrix $\tilde{\boldsymbol{A}}_0$ is defined by $(\tilde{\boldsymbol{A}}_0)_{ij} = (\boldsymbol{A}_0)_{ij}/P_i^{1/2}$, i.e. the lines of $\boldsymbol{A}_0$ are normalized by $P^{1/2}$. Then, for all possible distribution P, $\tilde{\boldsymbol{A}}_0$ has orthogonal columns with repect to $\langle .,.\rangle_P$. Following [3], the infinite norm of $\tilde{\boldsymbol{A}}_0$ is given by: $K(P) = \sup_{1\leqslant i,j\leqslant n} |(\tilde{\boldsymbol{A}}_0)_{ij}|$. Rauhut's result [3, Theorem 4.4] links the number of required measurements m to the sparsity level s of *any* unknown signal in order to guarantee its exact recovery from an independent sampling of its Fourier coefficients:

Theorem 1. *[3, Th. 4.4] Consider a sequence of m i.i.d. indexes drawn from the law P, and denote $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ the matrix composed of lines of $\tilde{\boldsymbol{A}}_0$ corresponding to these indexes. Assume that,*

$$\frac{m}{\ln(m)} \geqslant CK(P)^2 s \ln^2(s)\ln(n) \qquad (2)$$

$$m \geqslant DK(P)^2 s \ln(\epsilon^{-1}) \qquad (3)$$

Then, with probability $1-\epsilon$, every s-sparse vector $x \in \mathbb{C}^n$ can be recovered from observations $y = \boldsymbol{A}x$, by solving the ℓ_1 minimization problem in Eq. (1). The values C and D are universal constants.

Let us denote a_i^* the i-th line of $\boldsymbol{A}_0$. We show the following result:

Proposition 1. *Since $K(P) = \sup_{1\leqslant i,j\leqslant n} |(\tilde{\boldsymbol{A}}_0)_{ij}| = \sup_{1\leqslant i\leqslant n} |\frac{\|a_i\|_\infty}{P_i^{1/2}}|$,*

(i) *the optimal distribution π that minimizes the upper bound in (2)–(3) is: $\pi_i = \|a_i\|_\infty^2/L$ where $L = \sum_{i=1}^n \|a_i\|_\infty^2$.*

(ii) $K(\pi)^2 = L$.

Proof.

(i) Taking $P = \pi$, we get $K(\pi) = \sqrt{L}$. Now assume that $q \neq \pi$, since $\sum_{k=1}^n q_k = \sum_{k=1}^n \pi_k = 1$, $\exists j \in [1,n]$ such that $q_j < \pi_j$. Then $K(q) \geqslant \|a_j\|_\infty/q_j^{1/2} > \|a_j\|_\infty/\pi_j^{1/2} = \sqrt{L} = K(\pi)$. So, π is the distribution that minimizes $K(P)$.

(ii) is a consequence of π's definition. $\qquad\square$

In the next part, we assess the upper bound in inequality (2).

3.2. Discussion on the upper bound

In [6, 4], or even in [3, Th. 4.2], an $O(s\log(n))$ upper bound has been derived. Nevertheless, these results only consider the probability to recover *a given* sparse signal. The result introduced in Theorem 1 gives a *uniform* result. It is thus more general since it enables to apply the CS theory to *all* sparse signals and for our concern, to all MRI images. Moreover, it is more general than the one proposed in [6], since no assumption on the sign of the sparse signal entries is needed. Roughly speaking, the upper bound in Eq. (2) shows that the number of measurements needed to perfectly recover an s-sparse signal is $O(s\log^4(n))$ (since $m \leqslant n$ and $s \leqslant n$). According to [3], this bound is the best known result for uniform recovery.

Nevertheless, this bound is not usable in practice to determine the number of measurements. Indeed, for a 2D 256×256 image, $\log^4(n) = 15128$, $K^2(\pi) = L$ only depends on the choice of the wavelet representation. In our experiments, we used Symmlets with 10 vanishing moments and got $L = 8.34$. Rauhut [3] suggests that $C \gg 1$, which actually makes the upper bound unusable in Eq. (2).

Recent results [7] give $O(s^2)$ bounds for the number of measurements needed to guarantee exact reconstruction. Nevertheless, the

constants are lower and guarantee the reconstruction of very sparse signals. Unfortunately, the $O(s^2)$ bound called quadratic bottleneck is a strong limit for applicability in large scale scenarii.

3.3. Rejection of samples

An important issue of this theoretical approach is that k-space positions that appear more likely according to π are drawn more than once. To select a given position at most once, we introduce an intuitive alternative to the solution proposed in [6], which consists of rejecting samples associated with k-space positions that have already been visited. Let us show that this strategy is an improvement over the one proposed in Theorem 1.

Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ denote the matrix composed of m lines of $\tilde{\boldsymbol{A}}_0$, corresponding to m independent drawings. $\boldsymbol{A}$ has m_1 different lines ($m_1 \leqslant m$), and $\boldsymbol{A}_1 \in \mathbb{C}^{m_1 \times n}$ denotes the corresponding matrix. Let $\boldsymbol{A}_2 \in \mathbb{C}^{m\times n}$ denote the matrix obtained after m_2 additional independent drawings, where m_2 is the smallest integer such that the actual number of different samples matches m exactly. Then:

Proposition 2. *If x is the unique solution of the following ℓ_1 problem:* $\underset{\boldsymbol{A}w=\boldsymbol{A}x}{\operatorname{argmin}}\|w\|_1$, *it is also the unique solution of* $\underset{\boldsymbol{A}_2 w=\boldsymbol{A}_2 x}{\operatorname{argmin}} \|w\|_1$

Proof. Assume that x^* fulfills $\boldsymbol{A}_2 x^* = \boldsymbol{A}_2 x$ and $\|x^*\|_1 \leqslant \|x\|_1$, then since all lines of $\boldsymbol{A}$ are lines of $\boldsymbol{A}_2$, we get $\boldsymbol{A}x^* = \boldsymbol{A}x$. Because x is the unique solution of $\underset{\boldsymbol{A}w=\boldsymbol{A}x}{\operatorname{argmin}}\|w\|_1$ and $\|x^*\|_1 \leqslant \|x\|_1$, then $x^* = x$ and x is also the unique solution of: $\underset{\boldsymbol{A}_2 w=\boldsymbol{A}_2 x}{\operatorname{argmin}} \|w\|_1$. $\qquad\square$

In Theorem 1, m is the number of drawings. Prop. 2 proves that the result still holds if m is the number of *different samples* drawn according to law P.

3.4. Preliminary results

The proposed sampling strategy was tested on the Fourier tranform of a reference image (Fig. 1(a)) using several sampling patterns. We compare the above mentioned independent drawing from π distribution shown in Fig. 1(b)) to polynomial distributions $P(x,y) = (1 - (\sqrt{2}/n)\sqrt{k_x^2 + k_y^2})^p$ for variable power of decay $p = 1:6$ and $-n/2 < k_x, k_y \leqslant n/2$ [5]. In our experiments, the k-space is downsampled at rate 5 meaning that only 20% of the Fourier coeffients are measured; see Fig. 1(c)-(d).

Since our sampling schemes involve randomness, we performed a Monte-Carlo study and generated 10 sampling schemes from wich we perform reconstruction by solving the ℓ_1 minimization problem. For comparison purposes, we computed method-specific average value of peak Signal-to-Noise Ratio (PSNR) over the 10 reconstruction results as well as the corresponding standard deviations. The results are summarized in Tab. 1 where it is shown that an independent drawing from π gives worse results than those obtained with the empirical polynomial distribution with a power $p = 4:6$. Optimality of exponents $p = 5, 6$ are in agreement with previous work [8]. This poor reconstruction performance can be justified by the fact that MRI images are not really sparse in the wavelet domain since a lot of low-frequency wavelet coefficients are non-zero. This lack of sparsity justifies the weakness of this theoretical approach in comparison with polynomial or more generally variable density samplings.

The non-sparsity of the low frequency image content in the wavelet domain means that a lot of image information is contained around the k-space center. This explains why high order polynomial
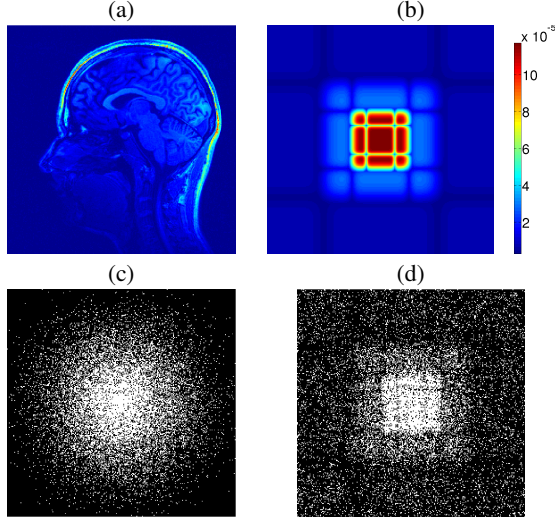
Fig. 1. **Example of k-space sampling schemes: selected samples appear in white color.** (a): Reference 256×256 used in our experiments. (b): Optimal distribution π. (c): Sampling pattern based on an independent drawing from a 4th order polynomial density [5]. (d): Sampling pattern based on an independent drawing from π.

distributions provide better reconstruction results. In what follows, we propose a novel two-step sampling process to overcome this limitation.

Table 1. Comparison of the reconstruction results in terms of PSNR for various k-space downsampling methods. **Bold font** indicate the best performance with respect to (wrt) the PSNR and its Std. Dev.

Sampling density		Mean PSNR (dB)	Std. dev.
Polynomial decay. Exponent:	1	23.55	1.40
	2	29.40	2.48
	3	32.00	3.01
	4	35.52	0.57
	5	**36.09**	0.14
	6	**35.94**	0.04
π		33.38	2.26

4. IMPROVING IMAGE SPARSITY

Most CS theories are based on sparsity of the signal of interest in a transform basis. Fig. 2 shows that low frequency wavelet coefficients contain a lot of information and do not fulfill the sparsity hypothesis. The following method tends to decompose the wavelet representation in two parts, one dedicated to low frequencies and the other to high frequencies in which the MRI image is more sparse. Since low frequency wavelets impact the k-space center, their recovery is performed by fully sampling this center. High frequencies are then reconstructed using CS theory.

4.1. Acquiring the k-space center: a sparsifying technique.

Let n_1 be the number of low frequency wavelets. Without loss of generality, assume that $\Psi_1, \ldots, \Psi_{n_1}$ are the low frequency wavelet atoms. Let $\Omega = \bigcup_{1 \leqslant i \leqslant n_1} \mathrm{supp}(\boldsymbol{F}^*\Psi_i)$. By definition of Ω, let us

Fig. 2. Representation of the typical MRI image shown in Fig. 1(a) in a wavelet basis (Symmlets with 10 vanishing moments).

introduce vector $x_\Omega = \boldsymbol{\Psi}^*\boldsymbol{F}y_\Omega$ where $y := \boldsymbol{A}_0 x$ and for $1 \leqslant i \leqslant n$, $(y_\Omega)_i = y_i 1_\Omega(i)$. Then $x_i = (x_\Omega)_i$ for $1 \leqslant i \leqslant n_1$ and vector $x_{\Omega^\perp} = x - x_\Omega$ is sparse since it contains no low frequency wavelet coefficient.

4.2. A two-stage reconstruction

The signal of interest $x_{\Omega^\perp}$ is now more sparse. We adopt the same strategy as in Section 3: we draw samples according to π and perform rejection if the sample drawn is located within Ω. Then, we recover the signal $x_{\Omega^\perp}$ by solving $\underset{\boldsymbol{A}w=y-\boldsymbol{A}x_\Omega}{\operatorname{argmin}} \|w\|_1$. We notice that:

(i) Since we reject samples already drawn, we can sample from the law $\pi^\star$ defined by $\pi_i^\star = \|a_i\|_\infty^2/L^\star$ if $i \notin \Omega$, 0 otherwise, with $L^\star = \sum_{i \notin \Omega} \|a_i\|_\infty^2$. The profile of $\pi^\star$ is shown in Fig. 3(b). Amongst the remaining frequencies, the more likely k-space positions are the lower frequencies. High frequencies remains unlikely and will be rarely visited by these schemes.

(ii) For the particular case of Shannon wavelets, $\Omega^\perp = \bigcup_{n_1 < j \leqslant n} \mathrm{supp}(\boldsymbol{F}^*\Psi_i)$. Then, matrix $\boldsymbol{A}_{\Omega^\perp}$ (composed of lines of $\tilde{\boldsymbol{A}}_0$ corresponding to frequencies included in $\Omega^\perp$) has orthogonal columns. Then, Theorem. 1 can be applied with the optimal distribution $\pi^\star$ and $\Omega = \sqrt{L^\star}$. Bounds (2)–(3) are thus improved since $L^\star < L$.
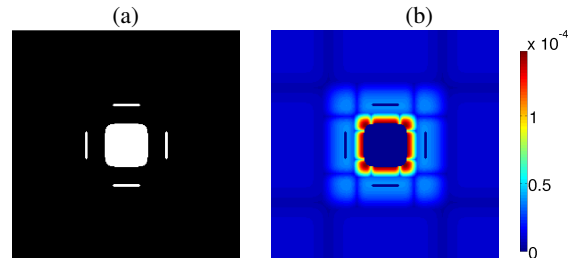

Fig. 3. Representation of Ω for a Symmlet-10 transform (a) and distribution $\pi^\star$ (b).

5. RESULTS

Here, we compare several classical sampling schemes used in MRI with our two-stage approach either combined with high frequency sampling from $\pi^\star$-distribution or from polynomial densities, as illustrated in Fig. 4. For each scheme, the number of samples acquired represents only 20% of the full k-space.

As shown in Tab. 2, sampling the whole k-space center drastically improves the reconstruction performance in comparison with Tab. 1 irrespective of the inital approach. Moreover, in contrast to what we observed in Section 3.4 with π, our two-stage approach
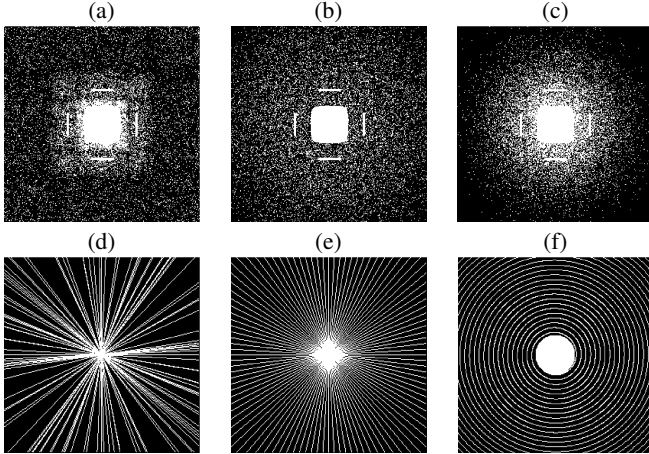
Fig. 4. **Various k-space sampling schemes.** (a)-(c): *Two-stage sampling schemes*. High frequencies sampled according to $\pi^\star$ (a) and to polynomial densities [5] with $p = 1$ (b) and $p = 4$ (c). (d)-(f): *Classical sampling schemes*. (d): Radial sampling with independent drawing of angles between lines. (e): Uniform radial sampling. (f): Spiral sampling, with a complete acquisition of low frequencies.

Table 2. Reconstruction results in terms of mean PSNR (and Std. dev. for random schemes) for the different patterns introduced in Fig. 4. **Bold font** indicate the best performance wrt the PSNR.

Sampling density		Mean PSNR (dB)	Std. dev.
Two-stage strategies:	$\pi^\star$	35.87	0.08
	Polynomial (1)	35.03	0.07
	Polynomial (2)	35.94	0.08
	Polynomial (3)	**36.30**	0.08
	Polynomial (4)	**36.32**	0.15
	Polynomial (5)	**36.24**	0.27
	Polynomial (6)	35.86	0.05
Random radial		31.97	0.52
Radial		34.22	
Spiral		31.15	

based on $\pi^\star$ achieves very close reconstruction results (while less accurate by 0.45dB) to those based on a similar two-step procedure relying on high order polynomial densities [5]. As expected, the gain g_p in dB for low order polynomial densities is larger ($g_p \in (4.3, 11.48)$ dB) for $p = 1 : 3$) since increasing the exponent in such densities makes the sampling denser around the k-space center. For the highest order density ($p = 6$), we even observe a loss of PSNR indicating that our approach fails in this context. On the other hand, the gain we obtained for $\pi^\star$ is more striking since $g_\pi = 2.49$ dB. Also, the standard deviation measured for all random schemes strongly decreases using our two-step approach. Besides, the two-stage sampling schemes perform better than spiral and radial patterns irrespective of the strategy with respect to angles (regular or random). Nevertheless, continuity is required for practical implementation of MRI sequences and we are currently investigating the design of optimal continuous trajectories. Our aim is to derive continuous sampling patterns from conciliating optimal drawing and continuity using Markov chains.

The reasons for which the reconstruction results based on the optimal distribution $\pi^\star$ are less accurate than those based on polynomial density sampling are still unclear. Two main hypotheses for-

mulate as follows: i.) even after the first step (subtracting low frequencies), the signal would not be sparse enough preventing us from meeting the conditions of Theorem. 1. ii.) the bounds given in Theorem. 1 might be sub-optimal and minimizing these bounds would not actually give the optimal sampling distribution. In particular, real images have a level of sparsity that increases among wavelet subbands, and this property should be taken into account in order to derive theoretical reconstruction results.

Although this theoretical approach does not seem optimal in terms of reconstruction performance, it encourages the use of compressed sensing for MRI. Indeed, this theory tends to sample the low frequencies more densely than the high frequencies, from considerations on the transform basis. The heuristics which have the best practical results [5] propose similar sampling density profiles. Theories [3, 4] suit with the fact that MRI images mainly contain low frequency information, though theoretical framework does not take this observation into consideration.

6. CONCLUSION AND FUTURE WORK

We proposed a two-stage method to design k-space sampling schemes. First, the signal is sparsified by acquiring the whole center of the k-space in order to recover the wavelet low frequency coefficients. Second, an independent random sampling of high frequencies is performed according to an optimal distribution for a given criterion so as to recover the remaining wavelet coefficients. Our results are comparable to the state-of-the-art. Also, we have shown that sampling the whole low frequencies drastically improves reconstruction quality. Our method is general enough to be used for any acquisition system measuring a sparse signal, which does not perfectly fulfill the sparsity hypothesis. In MRI, it seems that sparsity in the wavelet domain depends on the subbands. An interesting outlook of this work would be to include this remark in the theoretical framework.

7. REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[3] H. Rauhut, "Compressive Sensing and Structured Random Matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, M. Fornasier, Ed., vol. 9 of *Radon Series Comp. Appl. Math.*, pp. 1–92. deGruyter, 2010.

[4] E. J. Candès and Y. Plan, "A probabilistic and ripless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.

[5] M. Lustig, D.L 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, Dec. 2007.

[6] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 595–598, 2011.

[7] A. Juditsky, F.K. Karzan, and A. Nemirovski, "On low rank matrix approximations with applications to synthesis problem in compressed sensing," *SIAM J. on Matrix Analysis and Applications*, vol. 32, no. 3, pp. 1019–1029, 2011.

[8] F. Knoll, Ch. Clason, C. Diwoky, and R. Stollberger, "Adapted random sampling patterns for accelerated MRI," *Magn. Reson. Med.*, vol. 24, pp. 43–50, 2011.