

REVIEW

Open Access

# Compressed sensing MRI: a review from signal processing perspective



Jong Chul Ye

## Abstract

Magnetic resonance imaging (MRI) is an inherently slow imaging modality, since it acquires multi-dimensional k-space data through 1-D free induction decay or echo signals. This often limits the use of MRI, especially for high resolution or dynamic imaging. Accordingly, many investigators has developed various acceleration techniques to allow fast MR imaging. For the last two decades, one of the most important breakthroughs in this direction is the introduction of compressed sensing (CS) that allows accurate reconstruction from sparsely sampled k-space data. The recent FDA approval of compressed sensing products for clinical scans clearly reflect the maturity of this technology. Therefore, this paper reviews the basic idea of CS and how this technology have been evolved for various MR imaging problems.

**Keywords:** MRI; compressed sensing; k-space

## Background

Magnetic resonance imaging (MRI) exploits the nuclear magnetic resonance (NMR) phenomena to enable high contrast imaging. Since the first observation of NMR absorption in a molecular beam in 1938, the main research thrust was to understand and utilize the NMR phenomena for spectroscopy applications. Then, Lauterbur introduced in 1973 the gradient field to encode the spatial origin of the radio waves emitted from the nuclei of the object. This breakthrough allowed for multi-dimensional imaging by NMR physics.

The idea of Lauterbur can be easily understood using k-space interpretation that describes MR sampling as Fourier encoding in 2D or 3D spaces. Specifically, data collected by an MRI scanner are samples of the spatial Fourier transform of an object image. Hence, in order to obtain an image without aliasing artifacts, k-space samples need to satisfy the Nyquist sampling criterion.

Despite this close link to signal sampling theory, until the first demonstration of the sensitivity encoding (SENSE) technique by Prussemann [1], MR imaging was not considered as an important research topic for signal processing. Specifically, Prussemann et al. [1] showed that spatial diversity information from coil sensitivity

maps have additional information that can be exploited for fast signal acquisition. Furthermore, Sodickson et al. [2] proposed the simultaneous acquisition of spatial harmonics (SMASH). These works gave a birth of parallel imaging and iterative reconstruction methods, and has resulted in a flurry of novel ideas and algorithms, including Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA) by Griswold [3] and k-t space method for cardiac imaging such as [4–9].

The common theme in these approaches is that the data redundancy can be exploited to reduce the required sampling rate. Because redundant data can be compactly represented in some transform domains, it is also closely related to the concept of “sparsity”. Originally investigated by Bresler and his students for on-the-fly Fourier imaging [10, 11], the sparsity regularization has become the main workhorse in modern accelerated MRI researches thanks to the introduction of the compressed sensing theory [12, 13]. Ever since the first demonstration of compressed sensing MRI by Lustig et al. [14, 15], the compressed sensing MRI has become the essential tools in modern MR imaging researches. In this paper, we will review these ideas in more detail.

Correspondence: [jong.ye@kaist.ac.kr](mailto:jong.ye@kaist.ac.kr)

Department of Bio and Brain Engineering, Korea Adv. Inst. of Science & Technology (KAIST), 291 Daehak-ro, Daejeon, Korea



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

**Main text**

**MR forward models**

Before we introduce the concept of compressed sensing MR, we begin with the discussion of MR forward model. Mathematically, the forward model for k-space measurement can be described by.

$$b^i(\mathbf{k}) = \int \gamma^i(r) e^{-j2\pi\mathbf{k}^T r} dr, \quad i = 1, \dots, C, \quad (1)$$

where  $\mathbf{r} \in R^d$  and  $\mathbf{k} \in R^d$ ,  $d = 2, 3$  denote the image domain and k-space coordinates, respectively,  $C$  is the number of the coil and the  $i$ -th coil image  $\gamma^i$  is given by

$$\gamma^i(\mathbf{r}) = \mathbf{v}(\mathbf{r}) s^i(\mathbf{r}) \quad (2)$$

Here,  $\mathbf{v}(\mathbf{r})$  denotes the contrast weighted bulk magnetization distribution, and  $s^i(\mathbf{r})$  is the corresponding coil sensitivity of the  $i$ -th coil. Although the expression may give an impression that the measurement is two- or three-dimensional, this is indeed originated from a 1-D measurement since the k-space trajectory is a function of time, i.e.  $\mathbf{k} := \mathbf{k}(t)$ , and we acquire one k-space sample at each time point  $t$ . This makes the MR imaging inherently slow, since we need to scan through a 3-D object via 1-D trajectories.

Dynamic MRI is another important MR technique to monitor dynamic processes such as brain hemodynamics and cardiac motion. Among the various forms of dynamic MR modeling, here we mainly focus on the k-t formulation. Specifically, consider a discrete imaging equation for cartesian trajectory for simplicity. Because the samples along the readout direction are fully sampled, most of the dynamic MR formulation is applied separably after taking the Fourier transform along the read-out direction. More specifically, let  $\gamma(s, t)$  denote the unknown image content (for example, proton density, T1/T2 weighted image, etc.) on the spatial coordinate  $s$  along the phase encoding line at time instance  $t$ . Then, the k-t space measurement  $b(k, t)$  at time  $t$  is given by

$$b(k, t) = \int \gamma(s, t) e^{-j2\pi ks} ds \quad (3)$$

where  $\gamma(s, t)$  is the spatio-temporal image which may be weighted by coil sensitivity map for the case of parallel imaging.

Throughout the paper, for simplicity we often use the operator notation for (1):

$$b^i = F[S^i]v, \quad i = 1, \dots, C,$$

where  $[S^i]$  is a diagonal operator comprised of the  $i$ -th coil sensitivity. Similarly, we use operator notation for (3).

$$b = F\gamma \quad (4)$$

or

$$b^i = F\gamma^i, \quad i = 1, \dots, C \quad (5)$$

for the case of parallel imaging.

**Compressed sensing theory**

**Performance guarantees**

Compressed sensing (CS) theory [12, 16, 17] addresses the accurate recovery of unknown sparse signals from underdetermined linear measurements and has become one of the main research topics in the signal processing area for the last two decades [18–23]. Most of the compressed sensing theories have been developed to address the so-called single measurement vector (SMV) problems [12, 16, 17]. More specifically, let  $m$  and  $n$  be positive integers such that  $m < n$ . Then, the SMV compressive sensing problem is given by

$$(P0) : \begin{array}{l} \text{minimize } \|x\|_0 \\ \text{subject to } \|y - Ax\| < \epsilon, \end{array} \quad (6)$$

where  $y \in R^m$ ,  $A \in R^{m \times n}$ ,  $x \in R^n$ , and  $\epsilon$  denotes the noise level. (P0) implies that the solution favours the sparsest solution.

Since (P0) requires a computationally expensive combinatorial optimization, greedy methods [24], reweighted norm algorithms [25, 26], or convex relaxation using the  $l_1$  norm [12, 27] have been widely investigated as alternatives. In particular, the convex relaxation approach addresses the following  $l_1$  minimization problem:

$$(P1) : \begin{array}{l} \text{minimize } \|x\|_1 \\ \text{subject to } \|y - Ax\| < \epsilon, \end{array} \quad (7)$$

One of the important theoretical tools of CS is the so-called restricted isometry property (RIP), which enables us to guarantee the robust recovery of certain input signals [17]. More specifically, a sensing matrix  $A \in R^{m \times n}$  is said to have a  $k$ -restricted isometry property (RIP) if there is a constant  $0 \leq \delta_k < 1$  such that

$$(1 - \delta_k) \|x\|_2 \leq \|Ax\|_2 \leq (1 + \delta_k) \|x\|_2.$$

for all  $x \in R^n$  with  $\|x\|_0 \leq k$ . It has been demonstrated that  $\delta_{2k} < \sqrt{2} - 1$  is sufficient for  $l_1/l_0$  equivalence [12]. For many classes of random matrices, the RIP condition is satisfied with high probability if the sampling pattern is incoherent and the number of measurements satisfies  $m \geq ck \log(n/k)$  for some constant  $c > 0$  [17].

**Optimization approaches**

In practice, the following unconstrained form of optimization problem is often used:

$$\min_x \frac{1}{2} \|y - Ax\|^2 + \lambda \|\Psi x\|_1 \tag{8}$$

where  $y$  is noisy measurement,  $\lambda$  is the regularization parameter, and  $\Psi$  refers to an analysis transform such that  $\Psi x$  becomes sparse. Note that for the special case of  $\Psi = I$ , (8) is reduced to (P1).

One of the technical issues in solving (8) is that the cost function is not smooth at  $\Psi x = 0$ , which makes the corresponding gradient ill-defined. This leads to the development of various techniques, which are culminated as the new type of convex optimization theory called *proximal optimization* [28]. For example, the popular optimization methods such as forward-backward splitting (FBS) [29], split Bregman iteration [30], alternating directional method of multiplier (ADMM) [31], Douglas-Rachford splitting algorithm [32], and primal-dual algorithm [33] have been developed to solve compressed sensing problems. However, the comprehensive coverages of these techniques deserves another review paper, so in this section we mainly review the ADMM algorithms which have been extensively used for compressed sensing MRI ever since its first introduction to compressed sensing MRI [34].

More specifically, we convert the problem (8) to the following constraint problem:

$$\min_{y,u} \frac{1}{2} \|y - Ay\|^2 + \lambda \|u\|_1 \tag{9}$$

subject to

$$u = \Psi \gamma. \tag{10}$$

Then, the associated ADMM is given by

$$\gamma^{(k+1)} = \arg \min_x \frac{1}{2} \|y - Ay\|^2 + \frac{\mu}{2} \|\zeta^{(k)} + \Psi \gamma - u^{(k)}\|^2$$

$$u^{(k+1)} = \arg \min_u \lambda \|u\|_1 + \frac{\mu}{2} \|\zeta^{(k)} + \Psi \gamma^{(k+1)} - u\|^2$$

$$\zeta^{(k+1)} = \zeta^{(k)} + x^{(k+1)} - u^{(k+1)},$$

Now, each step of ADMM has a closed-form solution. Algorithm 1 summarises the resulting ADMM iteration, where  $\text{shrink}_1$  denotes the soft-tresholding:

$$\text{shrink}_1(x, \tau) = \text{sgn}(x) \max\{0, |x| - \tau\}, \tag{11}$$

for the threshold value  $\tau > 0$ .

In addition to the analysis prior in (8), the total variation (TV) penalty has been also extensively used for imaging applications, because the finite difference operator

**Algorithm 1**  $l_1$ -ADMM

---

```

1: Set  $u^{(0)}, \zeta^{(0)}$ .
2: for  $k = 0, 1, \dots$  do
3:    $\gamma^{(k+1)} = (A^*A + \mu I)^{-1} (A^*y + \mu (u^{(k)} - \zeta^{(k)}))$ 
4:    $u^{(k+1)} = \text{shrink}_1 \{ \zeta^{(k)} + \Psi \gamma^{(k+1)}, \frac{\lambda}{\mu} \}$ 
5:    $\zeta^{(k+1)} = \zeta^{(k)} + \gamma^{(k+1)} - u^{(k+1)}$ 
6: end for

```

---

can sparsify smooth images. Specifically, the TV minimisation problem assumes the following form:

$$\min_x \frac{1}{2} \|y - Ax\|^2 + \lambda \text{TV}(x), \tag{12}$$

where  $\text{TV}(x)$  is given by

$$\text{TV}(x) = \|\nabla x\|_{1,p}, \tag{13}$$

where  $\|\cdot\|_{1,p}$ ,  $p = 1, 2$  denotes the  $l_1/l_p$ -mixed norm. In  $d$ -dimensional space (e.g.  $d = 2$  for images), the discretized implementation can be defined as

$$\|\nabla x\|_{1,p} = \sum_{i=1}^n \|\nabla x(i)\|_p$$

where  $\nabla x(i) \in \mathbb{R}^d$  denotes the gradient of  $x$  at the  $i$ -th coordinate and  $n$  denotes the number of discretized samples.

In order to apply ADMM for (12), we need to focus on the primal formulation of the total variation penalty:

$$\text{TV}(x) = \|\nabla x\|_{1,p} = \sum_{i=1}^n \|\nabla x(i)\|_p \tag{14}$$

Now, we define a splitting variable  $u(i) = \nabla x(i) \in \mathbb{R}^d$ . Then, the constraint optimization formulation is given by

$$\min_{x, \{u(i)\}_{i=1}^n} \frac{1}{2} \|y - Ax\|^2 + \lambda \sum_{i=1}^n \|u(i)\|_p \tag{15}$$

$$\text{Subject to } u(i) = \nabla x(i), i = 1, \dots, n \tag{16}$$

Then, the associated ADMM is given by

$$x^{(k+1)} = \arg \min_x \frac{1}{2} \|y - Ax\|^2 + \frac{\eta}{2} \sum_i \|\zeta^{(k)}(i) + \nabla x(i) - u^{(k)}(i)\|^2$$

$$u^{(k+1)}(i) = \arg \min_{u(i)} \lambda \|u(i)\|_p + \frac{\eta}{2} \|\zeta^{(k)}(i) + \nabla x^{(k+1)}(i) - u(i)\|^2$$

$$\zeta^{(k+1)}(i) = \zeta^{(k)}(i) + \nabla x^{(k+1)}(i) - u^{(k+1)}(i)$$

Each step has the closed form expression. Algorithm 2 summarises the TV-ADMM algorithm, where  $\text{shrink}_{\text{vec},2}(x)$  for  $x \in \mathbb{R}^d$  denotes the vector shrinkage [34]:

$$\text{shrink}_{\text{vec},2}(x, \tau) = \frac{x}{|x|} \max\{0, |x| - \tau\}, x \in \mathbb{R}^d \quad (17)$$

for the threshold value  $\tau > 0$ .

---

**Algorithm 2** TV-ADMM
 

---

```

1: Set  $\eta > 0$  and  $\zeta^{(0)} = u^{(0)} = 0 \in \mathbb{R}^{n \times d}$ 
2: for  $k = 0, 1, \dots$  do
3:    $x^{(k+1)} = (A^*A - \eta\Delta)^{-1}(A^*y + \eta\text{div}(\zeta^{(k)} - u^{(k)}))$ 
4:   for  $i = 1, \dots, n$  do
5:      $u^{(k+1)}(i) = \text{shrink}_p\{\zeta^{(k)}(i) + \nabla x^{(k+1)}(i), \lambda/\eta\}, \quad \text{if } p = 1$ 
        $\text{shrink}_{\text{vec},p}\{\zeta^{(k)}(i) + \nabla x^{(k+1)}(i), \lambda/\eta\}, \text{ if } p = 2$ 
6:      $\zeta^{(k+1)}(i) = \zeta^{(k)}(i) + \nabla x^{(k+1)}(i) - u^{(k+1)}(i)$ 
7:   end for
8: end for

```

---

### Basic MR ingredients for compressed sensing

In order to apply compressed sensing theory for specific imaging applications, the unknown signal should be sparse in some transform domain, and the sensing matrix should be sufficiently incoherent. This section shows why these conditions can be readily satisfied in MR imaging, which is one of the main reasons that allows for successful applications of CS theory to MR imaging.

#### Sparsity of MR images

An MR image is rarely sparse in its own. However, one of the important observations that led to the successful applications of CS to MR is that the sparsity is closely related to signal redundancies. This is because redundant signals can be easily converted to sparse signals using some transforms.

Basically, there are three major directions that have been investigated in compressed sensing MRI: 1) the spatial domain redundancy, 2) the temporal domain redundancy, and 3) the coil domain redundancy. For example, as shown in Fig. 1(a), natural images can be sparsely represented in finite difference or wavelet transform domain, although the image is not sparse in its own. This observation is the main idea that allows for total variation (TV) and wavelet transform approaches for image denoising, and reconstruction. Accordingly, TV and wavelets have been the main transforms that have been extensively used in most of the CS MRI researches. On the other hand, dynamic MR images such

as cardiac cine, functional MRI, and MR parameter mapping have significant redundancy along the temporal dimension as shown in Fig. 1(b). For example, if the image is perfectly periodic, then temporal Fourier transform may be the optimal transform to sparsify the signal. However, in many dynamic MR problems, the temporal variations are dependent on the MR physics as well as specific motion of organs, so the analytic transform such as Fourier transform may not be an optimal solution, but more data-driven approaches such as PCA or dictionary learning are better options. Indeed, these observation leads to dictionary learning approaches that will be discussed later.

On the other hand, the coil redundancy is somewhat distinct compared to the spatial- and temporal- domain redundancy. As shown in Fig. 1(c), the redundancy of the coil image stems from the underlying common images, which results in cross- channel redundancies:

$$s^j(r)\gamma^i(r) = s^i(r)\gamma^j(r), i \neq j, \forall r \quad (18)$$

where  $s^i$  denotes the  $i$ -th coil sensitivity map and  $\gamma^i$  is the coil image. The relationship in (18) is easy to show since the coil image is given by (2). The main idea of the classical parallel MRI is to exploit this relationship. Specifically, the sensitivity encoding (SENSE) [1] exploits the image domain redundancy described in (18), whereas the  $k$ -space domain approaches such as GRAPPA [3] exploits its dual relationship in the  $k$ -space:

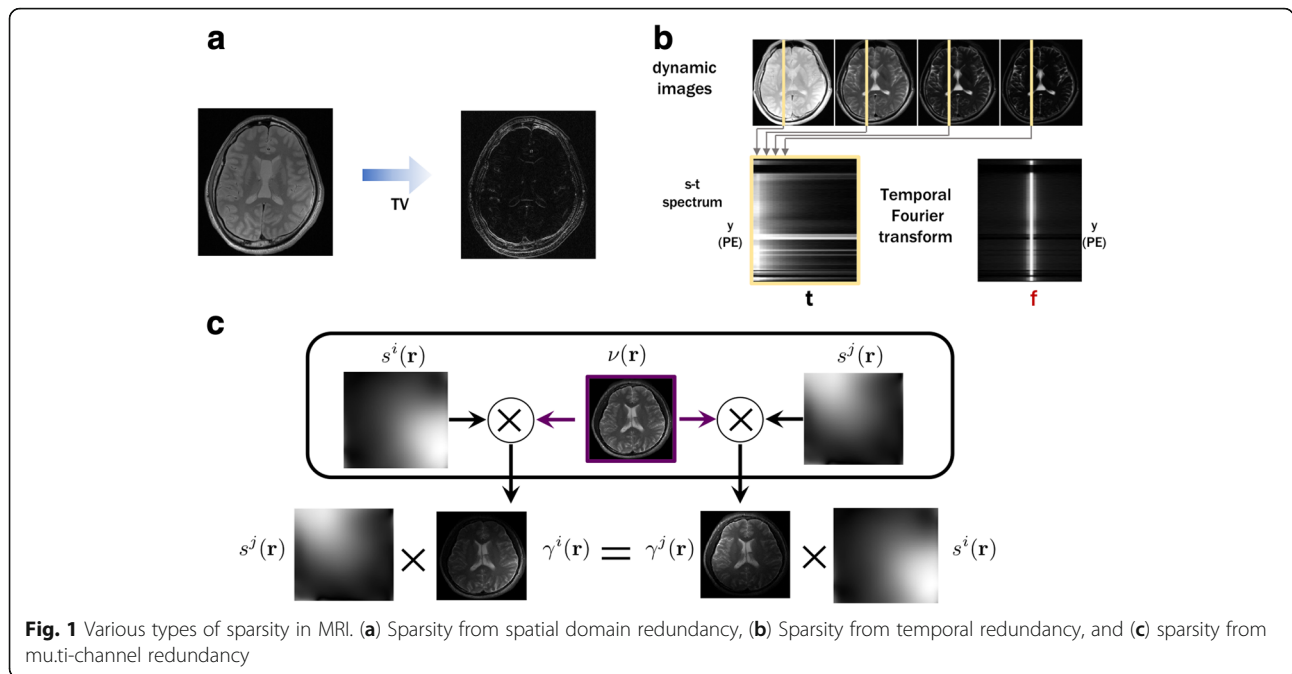
$$\hat{s}^j(r) * \hat{\gamma}^i(r) = \hat{s}^i(r) * \hat{\gamma}^j(r), i \neq j, \forall r \quad (19)$$

where  $*$  denotes the convolution, and  $\hat{s}^j, \hat{\gamma}^j$  denote the Fourier transform of  $s^j$  and  $\gamma^j$  respectively. Later we will describe how these two expressions of the coil dimensional redundancies have been exploited in compressed sensing MRI.

#### Incoherent sampling pattern

In compressed sensing MRI, downsampling pattern is very important to impose the incoherence, but its realization is limited by MR physics. For example, in the 2-D acquisition, the readout direction should be fully sampled, so there is only freedom along the phase encoding direction in designing incoherence sampling patterns. Figure 2(a)-(c) shows the examples of realizable sampling patterns that have been exploited in the literature: a) cartesian undersampling, b) radial trajectory, and c) spiral trajectory. In particular, the radial and spiral trajectories, which have been studied even before the advanced of the compressed sensing [35], have more incoherent radial sampling patterns compared to the cartesian undersampling.





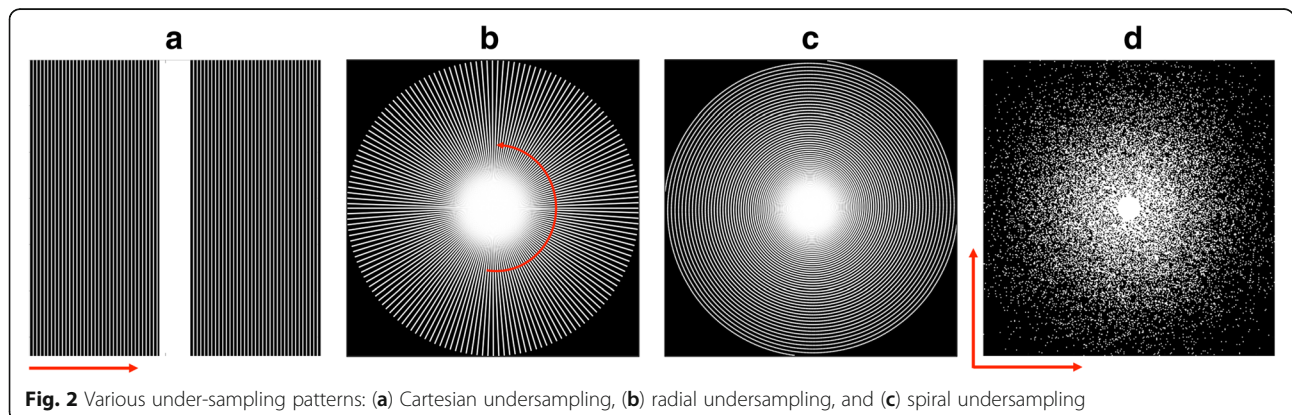
On the other hand, if we deal with 3-D imaging or dynamic imaging, there are more rooms for sampling pattern design, since there are two dimensional degree of freedom. Figure 2(d) shows an example of 2-D random sampling pattern. For the case of wavelet transform-based sparsity imposing prior, Lustig et al. [14] showed that the incoherence can be optimized by generalizing the notion of PSF to Transform Point Spread Function (TPSF). Specifically, TPSF measures how a single transform coefficient of the underlying object ends up influencing other transform coefficients of the measured undersampled object. To calculate the TPSE, a single point in the wavelet transform space at the  $i$ -th location is transformed to the image space and then to the Fourier space. Once the corresponding Fourier space data is subsampled by the given downsampling pattern, its influence in the wavelet transform

domain can be calculated by taking inverse transform followed by the inverse wavelet transform. To have the best incoherence property, the side of TPSF should be as small as possible, such that the side lobe contribution can be removed by shrinkage operation. This was used as the main criterion for sampling pattern design.

**Historical milestones**

The earliest application of the compressed sensing for MRI was done by Lustig and his collaborators [14]. The application of CS to dynamic MRI was pioneered by Jung et al. [18] and Gamper et al. [36], which was later significantly improved by Jung et al. [20]. The positive results of these earlier works have resulted in flurry of new ideas in static and dynamic MRI.

In terms of combining CS with parallel imaging, the first application for static imaging was done by Block et



al. [37] by combining total variation penalty and parallel imaging; and by Liang et al. [38] in more general form using SENSE. In dynamic imaging, the works by Jung et al. [18] and Feng et al. [39] are among the first that combine the SENSE type parallel imaging with the k-t domain compressed sensing. Combination with CS and parallel imaging using GRAPPA type constraints was pioneered by Lustig et al. [40]. Coil compression techniques have been also investigated to reduce the number of coils for CS reconstruction [41]. Feng et al. [42] later combined compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI, which has been approved for clinical use.

Aside from the standard  $l_1$  and TV penalty, several innovative sparsity inducing penalty have been used for compressed sensing MRI. For example, Trzasko et al. [43, 44] proposed a direct  $l_0$  minimization approaches, whereas Knoll et al. [45, 46] proposed a generalized total variation approaches, and Sung et al. [47] combined the approximate message passing algorithm with parallel imaging. This idea was then extended to the dictionary learning [20, 48, 49] and motion compensation [20, 50]. Then, the idea of low-rank regularization was soon introduced [49, 51–53] thanks to the theoretical advances by Candes et al. [54]. The low-rank idea was further extended to structured low-rank approach for parallel imaging [55] and single coil imaging with the finite support [56]. The duality between the compressed sensing and low-rank Hankel matrix approaches were discovered by Jin and his colleagues [57–60] and Ongie et al. [61]. In particular, the unified framework for compressed sensing and parallel MRI in terms of low-rank Hankel matrix approaches was presented by Jin et al. [57] and its theoretical performance guarantees was also given in [62].

In the following, we provide more detailed reviews of these historical milestones in compressed sensing MRI.

### Basic formulation of compressed sensing MRI

Although some MR images such as angiograms are already sparse in the pixel representation, more complicated images are rarely sparse, but only have a sparse representation in some transform domain, for example, in terms of spatial finite-differences or their wavelet coefficients. Based on this observation, Lustig et al. [14] proposed the first compressed sensing MRI using spatial domain wavelet transform as a sparsifying transform. More specifically, the problem is formulated as

$$\begin{aligned} \min_x \quad & \|\Psi\gamma\|_1 \tag{20} \\ \text{subject to} \quad & \|y-DF\gamma\|^2 < \epsilon \end{aligned}$$

where  $\|\cdot\|_1$  denotes the  $l_1$  norm,  $F$  is the Fourier transform,  $\gamma$  denotes the 2-D complex image,  $\Psi$  is a either

finite difference or spatial wavelet transform,  $D$  is the downsampling pattern, and  $y$  is the downsampled k-space measurement. Eq. (20) was solved using a non-linear conjugate gradient method.

The resulting optimization algorithms are, however, computational expensive. For cartesian sampling trajectory, this problem can be overcome as follows. Specifically, the image update step for (20) in ADMM implementation can be summarized as

$$(F^*D^*DF + \mu I)\gamma = F^*D^*y - \mu\Psi^*(\zeta^{(k)} - u^{(k)}),$$

which is computationally expensive due to the matrix inverse to obtain  $\gamma$ . Instead of using a direct matrix inverse, by multiplying the Fourier transform  $F$  to both sides, we have

$$(D^*D + \mu I)\hat{\gamma} = D^*y + \mu F\Psi^*(u^{(k)} - \zeta^{(k)})$$

where  $\hat{\gamma} = F\gamma$ . Note that a diagonal matrix  $D^*D$  consisting of ones and zeros. The ones are at those diagonal entries that corresponds to the sampled locations in the k-space. Let  $\Omega$  denote the sampled location. Then,

$$\hat{\gamma}^{(k+1)} = P_{\Omega} \left( \frac{D^*y + \mu F\Psi^*(u^{(k)} - \zeta^{(k)})}{1 + \mu} \right) + P_{\Omega^c} (F\Psi^*(u^{(k)} - \zeta^{(k)})), \tag{21}$$

where  $P_{\Omega}$  and  $P_{\Omega^c}$  denotes the projection on the sampling location  $\Omega$  and its complement  $\Omega^c$ , respectively. Then, the image update  $\gamma^{(k+1)}$  can be simply done by taking the fast Fourier transform (FFT).

The first experimental demonstration by Lustig et al. [14] clearly confirmed the efficiency of the compressed sensing algorithm, which has led to many other CS approaches using various sparsifying transform, and optimization algorithms, etc. For example, to deal with several hyperparameters to trade off between sparsity and data fidelity terms. Many research efforts have been made to quantify such trade-off, provide different forms of reconstruction, even to propose hyperparameter free reconstruction method [63–70].

### Advanced formulation for compressed sensing MRI Non-Cartesian compressed sensing MRI

Compressed sensing with non Cartesian sampling has been also extensively studied, since they are really a great combination given the sampling behavior of non Cartesian sampling schemes and the incoherence requirement in compressed sensing reconstruction. Aside from the radial and spiral sampling patterns discussed before, the works in [71–83] have focused on designing better sampling trajectories with good incoherent properties. For example, Haldar et al. [82] and

Puy et al. [83] proposed random phase encoding scheme to maximize the incoherency of the sensing matrix. However, one of the main technical issues associated with non-cartesian compressed sensing MRI is that the fast reconstruction trick shown in (21) cannot be used, which increases the overall computational time.

**Combination of parallel imaging with CS**

Recall that parallel MRI (pMRI) [1, 3] exploits the diversity in the receiver coil sensitivity maps that are multiplied by an unknown image. This provides additional spatial information for the unknown image, resulting in accelerated MR data acquisition through k-space sample reduction. Because the aim of the parallel imaging and CS approaches is similar, extensive research efforts have been made to synergistically combine the two for further acceleration [18, 20, 38, 40, 84].

One of the most simplest approaches can be a SENSE type approach that explicitly utilizes the estimated coil maps to obtain an augmented compressed sensing problem [18, 20, 38, 84]. More specifically, if the coil sensitivity is known and given by the sensitivity  $[S^i]$ ,  $i = 1, \dots, C$ , then the SENSE type compressed sensing MRI problem can be formulated as

$$\min_x \frac{1}{2} \sum_{i=1}^C \|b^i - DF[S^i]x\|^2 + \lambda \|\Psi x\|_1 \tag{22}$$

The optimization framework is the standard optimization framework under sparsity constraint, so proximal optimization algorithms can be used to solve this problem.

On the other hand,  $l_1$ -SPIRiT ( $l_1$ -iterative Self-consistent Parallel Imaging Reconstruction) [40] utilizes the GRAPPA type constraint as an additional constraint for a compressed sensing problem:

$$\begin{aligned} &\min_x \|\Psi x\|_1 \\ &\text{subject to } \|b^i - DF[S^i]x\|^2 < \epsilon, \quad i = 1, \dots, C \\ &x = Mx \end{aligned} \tag{23}$$

where  $M$  is an image domain GRAPPA operator. In both approaches, an accurate estimation of coil sensitivity maps or GRAPPA kernel is essential to fully exploit the coil sensitivity diversity. To address this problem, Uecker et al. [85] developed a novel eigen space method to extract the coil sensitivity maps directly from the k-space data, which is one of the most popular methods widely used by MR researchers.

**Blind compressed sensing MR using dictionary learning**

Blind compressed sensing approaches attempted to simultaneously reconstruct the underlying image as well as

the sparsifying transform from highly undersampled measurements. Ravishankar and his colleagues pioneered two distinct approaches - synthesis dictionary learning [48] and analysis transform learning [86] - when the underlying sparsifying transform is unknown a priori.

**Synthesis dictionary-based BCS**

More specifically, let  $P_j$ ,  $j = 1 \dots, N$  represents the operator that extracts a  $m$ -dimensional patch as a vector  $P_j x \in C^m$  from the image  $x$ , where  $N$  denotes the number of patches. Then, dictionary learning is to find the unknown dictionary  $D \in R^{m \times Q}$  and the corresponding sparse coefficient matrix  $C \in C^{m \times n}$  such that  $Y = DC$ , where  $Y = [P_1 x \ P_2 x \dots \ P_N x]$

The synthesis model allows each patch  $P_j x$  to be approximated by a linear combination  $Dc_j$  of a small number of columns from a dictionary  $D \in C^{m \times K}$ , where  $c_j \in C^K$  is sparse. The columns of the learnt dictionary (represented by  $d_k$ ,  $1 \leq k \leq K$ ) in (P0) are additionally constrained to be of unit norm in order to avoid the scaling ambiguity. The dictionary, and the image patch, are assumed to be much smaller than the image. This model can be used as a signal model, and Ravishankar et al. [48] proposed the following patch-based dictionary learning regularizer:

$$\begin{aligned} R(x) &= \min_{D,C} \sum_{j=1}^N \|P_j x - Dc_j\|^2, \text{ s.t. } \|d_i\| \\ &= 1, \forall i, \|C_j\|_0 \leq k \forall j \end{aligned} \tag{24}$$

where  $C_j$  and  $d_j$  denotes the  $j$ -th column of  $C$  and  $D$ , respectively. Then, the associated BCS formulation is given by

$$\begin{aligned} (P0) : &\min_{x,D,C} v \|Ax - b\|_2^2 \\ &+ \sum_{j=1}^N \|P_j x - Dc_j\|^2, \text{ s.t. } \|d_i\| \\ &= 1 \forall i, \|C_j\|_0 \leq s \forall j \end{aligned} \tag{25}$$

where  $A = DF$  denotes the downsampled Fourier transform.

To address the optimization problem (P0), Ravishankar et al. [48] employed the following two-step alternating minimization algorithm. First, the following minimization problem is solved by fixing  $x$ :

$$\begin{aligned} &\min_{D,C} \sum_{j=1}^N \|P_j x - Dc_j\|^2, \text{ s.t. } \|d_i\| \\ &= 1 \forall i, \|C_j\|_0 \leq K \forall j \end{aligned} \tag{26}$$

The K-SVD algorithm [87] was used to learn the dictionary. For a given dictionary  $D$ , the image update can be done by

$$\min_x \left\{ \nu \|Ax - b\|_2^2 + \sum_{j=1}^N \|P_j x - D_{C_j}\|_2^2 \right\} \quad (27)$$

These steps are alternated until convergence. The dictionary learning MRI have shown superior image reconstructions for MRI, as compared to non-adaptive compressed sensing schemes.

**Sparsifying transform-based BCS**

However, the BCS Problem (P0) is both non-convex and NP-hard. Approximate iterative algorithms for (P0) typically solve the synthesis sparse coding problem repeatedly, which makes them computationally expensive. In order to overcome some of the aforementioned drawbacks of synthesis dictionary-based BCS, Ravishankar et al. [86] proposed to use the sparsifying transform model in this work. Sparsifying transform learning has been shown to be effective and efficient in applications, while also enjoying good convergence guarantees. Specifically, they used the following transform learning regularizer:

$$R(x) = \min_{W, C} \sum_{j=1}^N \|WP_j x - C_j\|_2^2 + \lambda Q(W), \text{ s.t. } \|B\|_0 \leq s \quad (28)$$

where  $W \in C^{\wedge}\{m \times m\}$  denotes the unknown transform, the function  $Q(W)$  is a regularizer for the transform given by

$$Q(W) = -\log |\det W| + 0.5 \|W\|_F^2$$

The  $-\log |\det W|$  penalty eliminates degenerate solutions such as those with repeated rows. The  $\|W\|_F$  penalty helps remove a scale ambiguity in the solution.

Then, with additional constraint  $\|x\| \leq E$ , the overall optimization problem is given by

$$(P1) : \min_{x, W, C} \nu \|Ax - b\|_2^2 + \sum_{j=1}^N \|WP_j x - C_j\|_2^2 + \lambda Q(W), \text{ s.t. } \|C\|_0 \leq s, \|x\| \leq E, \quad (29)$$

where  $A = DF$  again denotes the downsampled Fourier transform. One of the important advantages of (P1) is that there exists a closed-form update for  $W$ , so the computationally expensive dictionary learning step can be avoided.

**K-t methods for compressed sensing dynamic MRI**

Dynamic MRI is a technique to acquire temporally varying MR sequences such as cardiac cine, perfusion, time-resolved angiography, functional MRI, etc. In dynamic MRI, there exists significant redundancies along the temporal directions, which can be extensively studied in various compressed sensing approaches.

**k-t SPARSE**

The k-t SPARSE by Lustig et al. [88] is an earliest version of compressed sensing dynamic MRI. Recall that the k-space measurement  $b(k, t)$  at time  $t$  is given by

$$b(k, t) = \int \gamma(s, t) e^{-j2\pi ks} ds \quad (30)$$

Another application of the Fourier transform along the temporal direction

$$\gamma(s, t) = \int p(s, f) e^{-j2\pi ft} df, \quad (31)$$

results in the following 2-D Fourier relationship:

$$b(k, t) = \iint p(s, f) e^{-j2\pi(kl+ft)} dsdf, \quad (32)$$

where  $\rho(s, f)$  denotes the temporal Fourier transform of  $\gamma(s, t)$ .

Note that  $p(s, f)$  is usually sparse because the periodic motions from heart or slow varying motions from fMRI can be easily sparsified using the temporal Fourier transform. Accordingly, the k-t data can be represented as mapping from the spatial-temporal image:

$$b = Ap \quad (33)$$

where  $A = DF$  and  $D$  is a k-t downsampling pattern, and  $F$  now becomes a 2D Fourier transform. In addition to the temporal Fourier transform, the authors in [88] used the wavelet transform in the spatial dimension to exploit the spatial redundancy. Then, k-t SPARSE is formulated based on the following optimization problem:

$$\min_y \|b - DFy\|_2^2 + \lambda \|Wy\|_1 \quad (34)$$

where  $W$  denotes the spatial wavelet transform, respectively.

**K-t FOCUSS**

Preliminary CS dynamic MRI approaches [36, 88] were seemingly different from the classical k-t approach such as k-t BLAST/SENSE [5]. One of the most important contributions of the k-t FOCUSS by Jung et al. [18, 20] was to reveal that the compressed sensing dynamic MRI is not very different from the classical k-t approaches, but rather it can be obtained by a very simple modification of the classical k-t BLAST/SENSE to ensure significant performance improvement.



More specifically, rather than using wavelet transform, k-t FOCUSS exploited the x-f domain sparsity. Then, a standard compressed sensing formulation would be:

$$\min_{\rho} \|b - D F \rho\|_2^2 + \lambda \|\rho\|_1 \tag{35}$$

to enforce the sparseness in x-f image  $\rho$ . However, there exists two additional novel ties in the k-t FOCUSS. First, rather than directly enforcing the sparseness of the x-f image, the k-t FOCUSS further sparsifies the x-f image using the initial estimate.

Specifically, let  $\rho_0$  be the predictable initial estimate of  $\rho$ . Then, the residual

$$x = \rho - \rho_0 \tag{36}$$

should be much more sparse than the original spatio-temporal image. Second, rather than directly using the  $l_1$  minimization, k-t FOCUSS employed the reweighted norm approaches. This results in the following equivalent minimization problem:

$$\min_{x,v} \|b - A \rho_0 - A x\|_2^2 + \lambda \frac{1}{2} \sum_{i=1}^n \left[ \frac{x_i^2}{v_i} + v_i \right], \tag{37}$$

Then, the normal equations with respect to  $x$  and  $v$  are given by

$$-\frac{x_i^2}{v_i^2} + 1 = 0$$

$$-F^H (b - A \rho_0 - A x) + \lambda V^{-1} x = 0$$

where  $V$  is a diagonal matrix whose  $i$ -th diagonal element is  $v_i$ . This result in the following FOCUSS iteration:

$$\rho^{(n+1)} = \rho_0 + W^{(n)} A^H (A^H W^{(n)} A + \lambda I)^{-1} (v - A \rho_0) \tag{38}$$

where the weighting matrix is given by

$$W^{(n+1)} = \text{diag} \left( |x^{(n+1)}|^{\frac{1}{2}} \right).$$

One of the most powerful observations in [18, 20] was that the first iteration of (38) has very similar form to the classical k-t BLAST/SENSE algorithm, except the power factor of weighting matrix. This observation led to an innovative idea to convert k-t BLAST/SENSE to compressed sensing approach. More specifically, by using incoherence sampling patterns, multiple iterations and correct weighting factor for the diagonal matrix, the authors of k-t FOCUSS [18, 20] clearly demonstrated the performance improvement. This observation

suggested that the improvement by the classical k-t algorithms such as k-t BLAST/SENSE [5] was not from the Bayesian perspective as the original authors of [5] had claimed, but indeed is originated from exploiting the sparsity in the spatio-temporal domain [18, 20]. Furthermore, by simply modifying the weight factor and sampling patterns, several additional iterations can significantly improve the performance of k-t BLAST/SENSE.

Another powerful aspect of k-t FOCUSS was that the idea can be easily extended to exploit the sparsity in other transform domains. For example, the residual step in (36) can be interpreted as sparsity promoting step by subtracting the temporal mean images. Thus, Jung et al. [20] proposes motion estimated and compensated modification of k-t FOCUSS to make the residual signal much sparser. More specifically, rather than subtracting the temporal mean values, they subtracted the motion estimated frame. Note that motion estimation and compensation (ME/MC) is an essential step in video coding that uses motion vectors to exploit the temporal redundancies between frames [89, 90]. In order to employ ME/MC within dynamic MRI, there are several technical issues to address. First, at least one reference frame is required. This issue can be easily resolved if we acquire fully sampled data in one frame as often done in dynamic MRI. The main technical difficulty, however, comes from the existence of the low quality current frame. Fortunately, this issue can be addressed using an additional reconstruction step before the ME/MC [20].

In addition, the spatio-temporal signals can be further sparsified using data-driven transform:

$$\rho = D c,$$

where  $D$  denotes the learned temporal dictionary based from the images, whereas  $c$  denotes the coefficients. Then, the imaging problem can be formulated as

$$\min_{c,v,D} \|b - A \rho_0 - A D c\|_2^2 + \lambda \frac{1}{2} \sum_{i=1}^n \left[ \frac{C_i^2}{v_i} + v_i \right], \tag{39}$$

For example, in order to find the temporal basis that can sparsify  $\rho$ , Jung et al. [18, 20] performed the singular value decomposition (SVD) after the image reconstruction, then the new dictionary is used to estimate the new coefficients. This procedure is closely related to the partial separable function (PSF) [91] and k-t SLR (k-t sparse and low-rank decomposition) [49], which will be reviewed soon.

**Partially separable function (PSF) approach**

In the PSF model by Liang et al. [91], the k-t samples are assumed to be decomposed in the following form

$$b(k, t) = \sum_{l=1}^L \Psi_l(k) \varnothing_l(t)$$

for some data dependent spectral and temporal basis function  $\{\Psi_l(k)\}_{l=1}^L$  and  $\{\varnothing_l(t)\}_{l=1}^L$ . Thanks to the partially separable assumption, the so-called Casorati Matrix B for the fully sampled k-t data given by

$$B = \begin{bmatrix} b(k_1, t_1) & \cdots & b(k_1, t_n) \\ \vdots & \ddots & \vdots \\ b(k_m, t_1) & \cdots & b(k_m, t_n) \end{bmatrix}$$

has at most rank L.

In dynamic CS MRI, many of the k-t samples are missing and we are interested in finding the missing components. Hence, by utilizing the low-rankness of the Ca-sorati matrix, the missing k-t samples can be estimated using a low rank matrix completion algorithm. In particular, the authors in [92, 93] proposed the following matrix factorization approach:

$$\{\hat{U}, \hat{V}\} = \arg \min_{U, V} \|A(UV^H) - B\|^2 \quad (40)$$

where A denotes the k-t sampling operator that indicates the missing samples by 0, and  $U \in C^{m \times L}$ ,  $V \in C^{m \times L}$  is used for the low rank matrix factorization  $B = UV^H$ , and the optimization is performed for U and V alternately by fixing the other matrix using the previous estimate. Because the exact rank of B is not known, the authors proposed the incremented powerFactorization (IRFP) algorithm [92], where (40) starts with  $L = 1$  by increasing order with the initialization of the power factorization of  $L + 1$  from that of  $L$ . To avoid an overfitting, the algorithm is terminated as soon as the data fidelity is below some threshold values.

### K-t SLR: K-t sparse and low rank approach

k-t Sparse and Low Rank Approach (k-t SLR) model by Lingala et al. [49] is a more systematic way of learning both basis and sparse coefficients. In this approach, the spatio-temporal signal  $\rho(x, t)$  is first rearranged in a matrix form

$$\tau = \begin{bmatrix} \rho(x_1, t_1) & \cdots & \rho(x_1, t_n) \\ \vdots & \ddots & \vdots \\ \rho(x_m, t_1) & \cdots & \rho(x_m, t_n) \end{bmatrix} \quad (41)$$

Then, using the low rank prior, the optimisation problem can be formulated as

$$\min_{\Gamma} \|b - A(\Gamma)\|^2 + \lambda \phi(\Gamma),$$

where  $A := DF$  is now a downsampled Fourier transform. Here, the rank prior is approximated using the general class of Schatten p-functionals, specified by

$$\phi(\Gamma) = \|\Gamma\|_p^p = \sum_i \sigma_i^p,$$

where  $\{\sigma_i\}$  denotes the singular values of  $\Gamma$ .

In dynamic imaging applications, the images in the time series may have sparse wavelet coefficients or sparse gradients. In addition, if the intensity profiles of the voxels are periodic (e.g., cardiac cine), they may be sparse in the Fourier domain. Based on this observation, Lingala et al. [49] proposed additional sparsity inducing penalty in specified basis sets along with the low-rank property to further improve the recovery rate. Specifically, they chose the 2-D wavelet transform to sparsify each of the images in the time series, while can be a 1-D Fourier transform to exploit the pseudo-periodic nature of motion. Then, the resulting composite minimization problem can be formulated as

$$\min_{\Gamma} \|b - \mathcal{A}(\Gamma)\|^2 + \lambda_1 \|\Gamma\|_p^p + \lambda_2 \left\| \sqrt{\sum_{i=0}^{q-1} |\Phi_i^H \Gamma \Psi_i|^2} \right\|_1 \quad (42)$$

### K-space structured low-rank approaches

#### Basic theory

Compared to the standard compressed sensing approaches, k-space structured low-rank approaches such as SAKE [55], LORAKS [56], ALOHA [57, 59, 60] and GIRAF [61] are relatively new, but has significant potentials in MRI imaging. These approaches are all derived by the k-space convolution relationship and can be used for both static and dynamic imaging. So we first discuss a matrix representation of the convolution. For simplicity, we will consider the 1-D notation.

Specifically, consider a fully sampled k-space measurement from the multi-channel coils:

$$y^i = F \gamma^i, i = 1, \dots, C$$

where  $\gamma^i$  denotes the unknown i-th coil images and  $y^i$  corresponds to its k-space data. Note that the matrix representation of a k-space convolution of  $y^i$  with a d-tap filter h is given by

$$z^i = C(y^i) \bar{h} \quad (43)$$

where  $C(y^i)$  denotes the convolution matrix constructed by the vector  $y^i$ :

$$C(y^i) = \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ y^i[-1] & y^i[0] & \cdots & y^i[d-2] \\ y^i[0] & y^i[1] & \cdots & y^i[d-1] \\ y^i[1] & y^i[2] & \cdots & y^i[d] \\ \vdots & \vdots & \ddots & \vdots \\ y^i[n-d] & y^i[n-d+1] & \cdots & y^i[n-1] \\ y^i[n-d+1] & y^i[n-d+2] & \cdots & y^i[n] \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} \tag{44}$$

and  $h^\top$  denotes a vector that reverses the order of the elements. If we extract the  $n - d$ -rows of the convolution matrix with  $n - d > d$ , we can obtain the following Hankel structured matrix:

$$\mathcal{H}_d(y^i) = \begin{bmatrix} y^i[0] & y^i[1] & \cdots & y^i[d-1] \\ y^i[1] & y^i[2] & \cdots & y^i[d] \\ \vdots & \vdots & \ddots & \vdots \\ y^i[n-d] & y^i[n-d+1] & \cdots & y^i[n-1] \end{bmatrix} \tag{45}$$

By defining  $Y = [y^1, \dots, y^C]$ , we can further defined the extended Hankel matrix

$$\mathcal{H}_{d|C}(Y)h_{(i,k)} = [\mathcal{H}_d(y^1) \cdots \mathcal{H}_d(y^C)] \tag{46}$$

In the following, we will explain how these Hankel structured matrix has been utilised for accelerated MRI.

**SAKE**

A calibrationless parallel imaging reconstruction method, termed simultaneous au- tocalibrating and k-space estimation (SAKE), is a data-driven, coil-by-coil recon- struction method that does not require a separate calibration step for estimating coil sensitivity information [55]. SAKE is based on the following observation in GRAPPA:

$$\sum_{k \neq i}^C y^k * w_{(k,i)} = y^i \tag{47}$$

which implies that the  $i$ -th  $k$ -space measurement can be represented as the linear combination of the filtered  $k$ -space data from other coils. In matrix form, this recursive relationship implies the existence of the null space of the Hankel matrix.

$$\mathcal{H}_{d|C}(Y)h_{(i,k)} = [\mathcal{H}_d(y^1) \cdots \mathcal{H}_d(y^C)]h_{(i,k)} = 0 \tag{48}$$

In other word,  $H_{d|C}(Y)$  is low-ranked. Therefore, SAKE solves the following low- rank matrix completion problem to interpolate the missing  $k$ -space data:

$$\min_M \text{rank}(\mathcal{H}(M))$$

$$\text{subject to } P_\Omega(m^i) = P_\Omega(y^i), i = 1, \dots, C,$$

where  $P_\Omega(\cdot)$  denotes the projection on the measured  $k$ -space samples on the index set  $\Omega$ . The problem was solved using the iterative singular value shrinkage method [94].

**LORAKS**

The low-rank modeling of local-space neighborhoods (LORAKS) [56] was inspired by the finite support condition. More specifically, if the object  $\gamma$  has finite support, we can easily find the function  $w$  such that

$$\gamma w = 0,$$

This results in a convolution relationship in  $k$ -space

$$y * h = 0 \Rightarrow \mathcal{H}(y)h = 0,$$

where  $y$  and  $h$  denote the Fourier spectrum of  $\gamma$  and  $w$ , respectively. Therefore, this gives a single channel version of the low-rank condition, which results in the following rank minimization problem:

$$\min_m \text{rank}(\mathcal{H}(m))$$

$$\text{subject to } P_\Omega(m) = P_\Omega(\hat{y}),$$

**ALPHA and GIRAF**

Annihilating filter-based low rank Hankel matrix (ALPHA) approach [57–60, 62] and GIRAF (Generic Iterative Reweighted Annihilating Filter) [61] can be considered as the full generalization of SAKE and LORAKS for general class of signals with the finite rate of innovations (FRI) for MR measurements. Moreover, the approaches have unified the parallel imaging and compressed sensing as a  $k$ - space interpolation with performance guarantees [62], and can be used for artifact correction [58]. This section describes the fundamental dual relationship between transform domain sparsity and low rankness in reciprocal domain, which is the key ingredient. For better readability, we provide here a high level description by assuming 1-D signals.

The Fourier CS problem of our interest is to recover the unknown signal  $x(t)$  from the Fourier measurement:

$$\hat{x}(\omega) = \mathcal{F}\{x(t)\} = \int x(t)e^{-i\omega t} dt. \tag{49}$$

In classical Nyquist sampling, to avoid aliasing artefacts, the grid size should be at most:

$$\Delta = 2\pi/\tau$$

when the support of the time domain signal  $x(t)$  is  $\tau$ . Then, discrete Fourier data at the Nyquist rate is defined by:

$$\hat{x}[k] = \hat{x}(\omega)|_{\omega=\frac{2\pi k}{\tau}}. \tag{50}$$

We also define a length  $(r + 1)$ -annihilating filter  $\hat{h}[k]$  for  $\hat{x}[k]$  that satisfies

$$(\hat{h} * \hat{x})[k] = \sum_{p=0}^r \hat{h}[p]\hat{x}[k-p] = 0, \quad \forall k. \tag{51}$$

The existence of the minimum length finite length annihilating filter has been extensively studied for FRI signals [95–97]. Let  $r + 1$  denotes the minimum size of annihilating filters that annihilates discrete Fourier data  $\hat{x}[k]$ . Then, a  $d$ -tap annihilating filter  $h$  with  $d > r + 1$  can be easily obtained by convolving an appropriate size FIR filter with the minimum length annihilating filter. In matrix form, this is equivalent to

$$\mathcal{H}_d(\hat{x})\bar{h} = 0 \tag{52}$$

where the Hankel structure matrix  $\mathcal{H}_d(\hat{x})$  is constructed as

$$\mathcal{H}_d(\hat{x}) = \begin{bmatrix} \hat{x}[0] & \hat{x}[1] & \cdots & \hat{x}[d-1] \\ \hat{x}[1] & \hat{x}[2] & \cdots & \hat{x}[d] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}[n-d] & \hat{x}[n-d+1] & \cdots & \hat{x}[n-1] \end{bmatrix} \tag{53}$$

Assume that  $\min(n - d + 1, d) > r$ . Then, we can show the following low rank property [62]:

$$\text{RANK}\mathcal{H}(\hat{x}) = r, \tag{54}$$

Thanks to the low-rankness of the associated Hankel matrix, the missing  $k$ -space data can be easily interpolated using the following low-rank matrix completion [62]:

$$\begin{aligned} &\underset{m \in \mathbb{C}^n}{\text{minimize}} \text{RANK } \mathcal{H}(m) \\ &\text{subject to } P_{\Omega}(m) = P_{\Omega}(\hat{x}), \end{aligned} \tag{55}$$

where  $P_{\Omega}$  is the projection operator on the sampling location  $\Omega$ . Moreover, as shown in [62], the low-rank matrix completion approach (55) does not compromise any optimality compared to the standard Fourier CS.

Note that signals may not be sparse in the image domain, but can be sparsified in a transform domain. In fact, this was the main idea of the compressed sensing.

Specifically, the signal  $x$  of our interest is a non-uniform spline that can be represented by:

$$Lx = w \tag{56}$$

where  $L$  denotes a constant coefficient linear differential equation that is often called the continuous domain whitening operator in [98, 99]:

$$L := a_K \partial^K + a_{K-1} \partial^{K-1} + \dots + a_1 \partial + a_0 \tag{57}$$

and  $w$  is a continuous sparse innovation:

$$w(t) = \sum_{j=0}^{r-1} c_j \delta(t - t_j). \tag{58}$$

For example, if the underlying signal is piecewise constant, we can set  $L$  as the first differentiation. In this case,  $x$  corresponds to the total variation signal model. Then, by taking the Fourier transform of (56), we have

$$\hat{z}(\omega) = \hat{l}(\omega)\hat{x}(\omega) = \sum_{j=0}^{r-1} a_j e^{-i\omega x_j} \tag{59}$$

where

$$l(\omega) = a_K (i\omega)^K + a_{K-1} (i\omega)^{K-1} + \dots + a_1 (i\omega) + a_0 \tag{60}$$

Accordingly, the Hankel matrix  $\mathcal{H}(\hat{z})$  from the weighted spectrum  $\hat{z}(\omega)$  satisfies the following rank condition:

$$\text{RANK}\mathcal{H}(\hat{z}) = r.$$

Thanks to the low-rankness, the missing Fourier data can be interpolated using the following matrix completion problem:

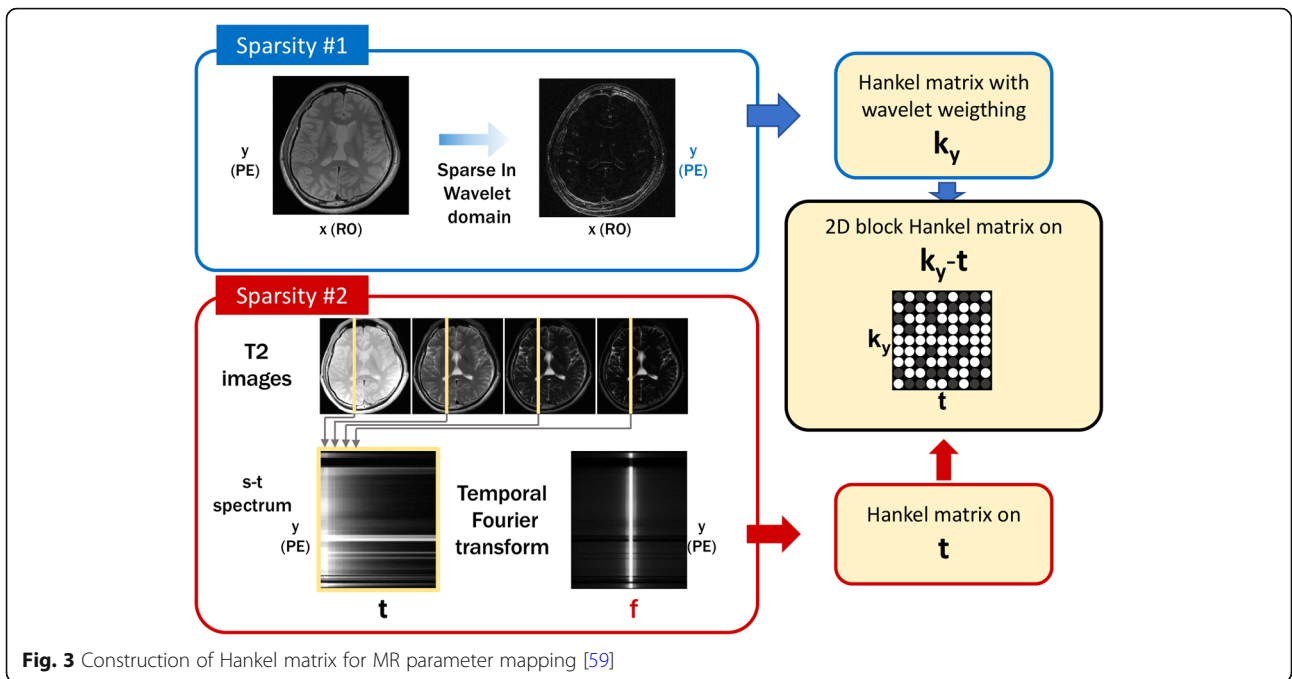
$$(P_w) \min_{m \in \mathbb{C}^n} \|\mathcal{H}(m)\|_*, \text{ subject to } P_{\Omega}(m) = P_{\Omega}(\hat{l} \odot \hat{y}), \tag{61}$$

or, for noisy Fourier measurements  $\hat{y}$ ,

$$(P'_w) \min_{m \in \mathbb{C}^n} \|\mathcal{H}(m)\|_*, \text{ subject to } \|P_{\Omega}(m) - P_{\Omega}(\hat{l} \odot \hat{y})\| \leq \delta \tag{62}$$

where  $\odot$  denotes the Hadamard product,  $\hat{l}$  and  $\hat{x}$  denotes the vectors composed of full samples of  $\hat{l}(\omega)$  and  $\hat{x}(\omega)$ , respectively. After solving  $(P_w)$ , the missing spectral data  $\hat{x}(\omega)$  can be obtained by dividing by the weight, i.e.  $\hat{x}(\omega) = m(\omega)/\hat{l}(\omega)$  assuming that  $\hat{l}(\omega) \neq 0$ .



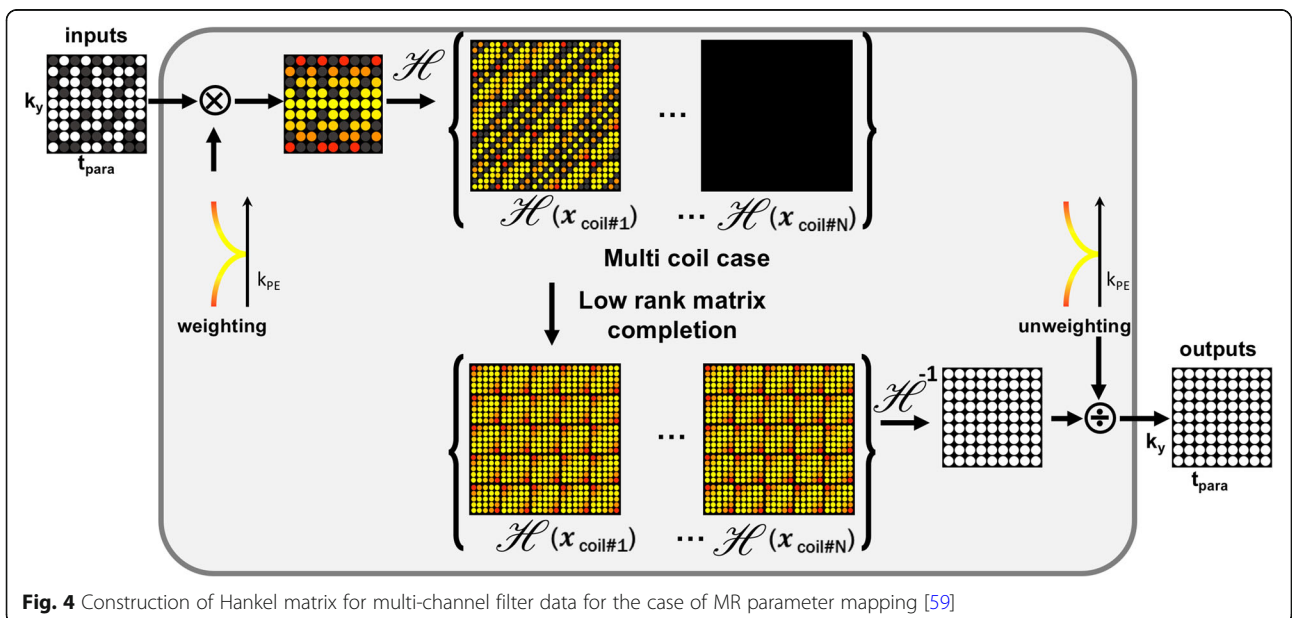


**Fig. 3** Construction of Hankel matrix for MR parameter mapping [59]

The idea can be generalized for any transform domain sparse signals as long as the transform can be represented using shift-invariant filters. Wavelet domain sparse signal belongs to this class. In this case, the weight kernel in the Fourier domain is obtained as the spectrum of the subband filters [57–60, 62]. For example, Fig. 3 showed the construction of the Hankel matrix for the MR parameter mapping, where the

k-space weighting from wavelet weighting is applied only along the phase encoding direction, whereas no weighting is applied along the t-domain since the correspond temporal spectrum is already sparse [59].

The idea can be also easily generalised to the parallel imaging by exploiting (19). Specifically, (19) can be equivalently represented using the matrix representation:



**Fig. 4** Construction of Hankel matrix for multi-channel filter data for the case of MR parameter mapping [59]

$$\mathcal{H}_a(\hat{\gamma}^i)\hat{s}^j = \mathcal{H}_a(\hat{\gamma}^j)\hat{s}^i, \forall i \neq j, \quad (63)$$

This implies that an extended Hankel matrix  $\mathcal{H}_{a|C}([\hat{\gamma}^1 \dots \hat{\gamma}^C])$  in (46) is low ranked.

For example, the multi-channel construction of Hankel matrix for MR parameter mapping [59] is shown in Fig. 4.

One of the most important advantages of the Hankel matrix formulation is that the coil diversity can be readily exploited in addition to the image domain redundancy in a unified framework. This makes the separate coil sensitivity estimation unnecessary.

### Clinical applications

Since the compressed sensing MRI allows significant acceleration of MR acquisition, it has been extensively applied for various clinical applications such as fast cardiac MRI, whole heart MRI, dynamic contrast enhanced (DCE)-MRI, diffusion MRI, spectroscopic, etc., that usually require significant acquisition time using standard methods.

For example, Otazo et al. [84] applied the CS method to the first pass cardiac perfusion MRI and demonstrated feasibility of 8-fold acceleration in vivo imaging using standard coil arrays. They showed that CS method results in similar temporal fidelity and image quality to GRAPPA with 2-fold acceleration [84]. Hsiao et al. [100] applied combined parallel imaging and compressed sensing to achieve 4D phase contrast for the quantification of cardiac flow and ventricular volumes pediatric patients during congenital heart MRI examinations. Vincent et al. [101] employed CS to evaluate LV function and volumes and found that CS strategy with the single breath hold provided similar results to multi breathhold imaging protocols.

For free-breathing contrast-enhanced multiphase liver MRI, Chandarana et al. [102] showed that a combination of compressed sensing, parallel imaging, and radial k-space sampling demonstrated the feasibility of breath-hold cartesian T1 weighted imaging. Espagnet et al. [103] employed golden-angle radial sparse parallel technique for DCE-MRI to evaluate the permeability characteristics of the pituitary gland.

For diffusion MRI, Landman et al. [104] showed that CS reconstruction using standard data can resolve crossing fibers similar to a standard q-ball approach using much richer data with longer acquisition time. Kuhnt et al. [105] showed that High Angular Resolution Diffusion Imaging (HARDI) + CS is a promising approach for fiber tractography in clinical practice.

Finally, for the spectroscopic imaging, Larson et al. [106] developed a CS method for acquiring hyperpolarized <sup>13</sup>C data using multiband excitation pulses and achieved 2 s temporal resolution with full volumetric coverage of a mouse. Geethanath et al. [107] demonstrated

a potential reduction in acquisition time by up to 80% or more for hydrogen 1 MR spectroscopic imaging using CS, with negligible loss of clinical information.

With the commercially available CS reconstruction methods, we expect to see more clinical applications of CS in the near future.

### Conclusions

Nowdays, compressed sensing has become a mature technology, as reflected by recent approval by FDA. Major vendors have started to sell the compressed sensing reconstruction softwares, and many clinical researchers have been evaluating its clinical usefulness.

Despite of this maturity, some of the main technical issues of the compressed sensing are 1) the computational complexity of the algorithm is relative high, and at high acceleration, image quality degradation is still reported. Although the recent state-of-the art CS techniques such as structured Hankel matrix approach can address the quality degradation problems, it also increases the computational complexity, which may interfere the clinical workflow.

Fortunately, for the last two years, the MR image reconstruction field have been rapidly changed thanks to the successful demonstration of of the deep learning-based MR reconstruction technologies [108–113]. The sudden popularity of deep learning approaches can be attributed to the real-time reconstruction in spite of the significant improvement of the image quality. Thus, when originally presented, these techniques were regarded as totally different technology that is nothing to do with the compressed sensing. However, recent theoretical analysis [114] showed that the deep convolutional neural network is closely related to the Hankel matrix decomposition. Therefore, we can still argue that the compressed sensing MRI has renewed interests in the form of deep learning, and it will be interesting to see how this exciting and rapidly evolving field will develop for the coming decades.

### Abbreviations

ALoHA: Annihilating filter-based low-rank Hankel matrix approach; CS: Compressed sensing; FOCUSS: FOCal Underdetermined System Solver; GRAPPA: Generalized Autocalibrating Partially Parallel Acquisitions; /1-SPIRiT: /1-Iterative Self-consistent Parallel Imaging Reconstruction; LORAKS: Low-rank modeling of local-space neighborhoods; MRI: Magnetic resonance imaging; SAKE: Simultaneous autocalibrating and k-space estimation; SENSE: Sensitivity encoding

### Acknowledgements

The author would like to thank his students, Yoseob Han and Dongwook Lee, for providing figures. The author also like to thank two anonymous reviewers who comments have significantly improve the quality of the paper.

### Funding

This work is supported by National Research Foundation of Korea, Grant number NRF-2016R1A2B3008104.

**Availability of data and materials**

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Author's contributions**

The author has read and approved this article.

**Ethics approval**

Not applicable.

**Consent for publication**

Not applicable

**Competing interests**

The author declares that he has no competing interests.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 3 November 2018 Accepted: 4 February 2019

Published online: 29 March 2019

**References**

1. Pruessmann KP, Weigher M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. *Magn. Reson. Med.* 1999;42(5):952–62.
2. Sodickson DK, Manning WJ. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn Reson Med.* 1997;38(4):591–603.
3. 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(6):1202–10.
4. Madore B, Glover GH, Pelc NJ. Unaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. *Magn. Reson. Med.* 1999;42(5):813–28.
5. Tsao J, Boesiger P, Pruessmann KP. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magn Reson Med.* 2003;50(5):1031–42.
6. Kozerke S, Tsao J, Razavi R, Boesiger P. Accelerating cardiac cine 3D imaging using k-t BLAST. *Magn. Reson. Med.* 2004;52:19–26.
7. Sharif B, Derbyshire JA, Faranesh AZ, Bresler Y. Patient-adaptive reconstruction and acquisition in dynamic imaging with sensitivity encoding (paradise). *Magn Reson Med.* 2010;64(2):501–13.
8. Tsao J, Kozerke S, Boesiger P, Pruessmann KP. Optimizing spatiotemporal sampling for k-t BLAST and k-t SENSE: Application to high-resolution real-time cardiac steady-state free precession. *Magn. Reson. Med.* 2005;53:1372–82.
9. Hansen MS, Kozerke S, Pruessman KP, Boesiger P, Pedersen EM, Tsao J. One the influence of training data quality in k-t BLAST reconstruction. *Magn. Reson. Med.* 2004;52:1175–83.
10. Bresler, Y., Gastpar, M., Venkataramani, R.: Image compression on-the-fly by universal sampling in fourier imaging systems. In: Proc. 1999 IEEE Information Theory Workshop on Detection, Estimation, Classification, and Imaging, pp. 48 (1999)
11. Ye, J.C., Bresler, Y., Moulin, P.: A self-refencing level-set method for image reconstruction from sparse Fourier samples. In: to Appear in Proc. IEEE Workshop in Variational and Level Set Methods in Computer Vision, Vancouver, Canada (2001)
12. Candes E, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory.* 2006;52(2):489–509.
13. Donoho DL. Compressed sensing. *IEEE Trans. Inf. Theory.* 2006;52(4):1289–306.
14. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med.* 2007;58(6):1182–95.
15. Lustig M, Donoho DL, Santos JM, Pauly JM. Compressed sensing MRI. *IEEE Signal Process Mag.* 2008;25(2):72–82.
16. Candes EJ, Romberg JK, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Commun Pure Appl Math.* 2006;59(8):1207–23.
17. Candes E, Tao T. Decoding by linear programming. *IEEE Trans. Info. Theory.* 2005;51(12):4203–15.
18. Jung H, Ye JC, Kim EY. Improved k-t BLAST and k-t SENSE using FOCUSS. *Phys. Med. Biol.* 2007;52:3201–26.
19. Ye JC, Tak S, Han Y, Park HW. Projection reconstruction MR imaging using FOCUSS. *Magn Reson Med.* 2007;57(4):764–75.
20. 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(1):103–16.
21. Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med. Phys.* 2008;35:660.
22. Cotter SF, Rao BD. Sparse channel estimation via matching pursuit with application to equalization. *IEEE Trans. Commun.* 2002;50:374–7.
23. Wagadarikar A, John R, Willett R, Brady D. Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics.* 2008;47(10):44–51.
24. Tropp JA. Just Relax: convex programming methods for identifying sparse signals in noise. *IEEE Trans. Inf Theory.* 2006;52(3):1030–51.
25. Gorodnitsky IF, Rao BD. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Trans Signal Process.* 1997;45(3):600–16.
26. Candes EJ, Wakin MB, Boyd SP. Enhancing sparsity by reweighted  $\ell_1$  minimization. *J Fourier Anal Appl.* 2008;14(5):877–905.
27. Chen SS, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit. *SIAM J Sci Comput.* 1999;20(1):33–61.
28. Bauschke HH, Combettes PL. Convex analysis and monotone operator theory in Hilbert spaces. Vol. 408. New York: Springer; 2011.
29. Combettes PL, Wajs VR, et al. Signal recovery by proximal forward-backward splitting. *Multiscale Model Simul.* 2006;4(4):1168–200.
30. Goldstein T, Osher S. The split Bregman method for  $L_1$  regularized problems. *SIAM J Imag Sci.* 2009;2(2):323–43.
31. Wang Y, Yang J, Yin W, Zhang Y. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J Imag Sci.* 2008;1(3):248–72.
32. Combettes PL, Pesquet JC. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Sel. Top. Sign. Proces.* 2007;1(4):564–74.
33. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision.* 2011;40(1):120–45.
34. Ramani S, Fessler JA. Parallel MR image reconstruction using augmented Lagrangian methods. *IEEE Trans. Med. Imaging.* 2011;30(3):694–706.
35. Meyer CH, Hu BS, Nishimura DG, Macovski A. Fast spiral coronary artery imaging. *Magn Reson Med.* 1992;28(2):202–13.
36. Gamper U, Boesiger P, Kozerke S. Compressed sensing in dynamic MRI. *Magn Reson Med.* 2008;59(2):365–73.
37. Block KT, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med.* 2007;57(6):1086–98.
38. Liang D, Liu B, Wang J, Ying L. Accelerating SENSE using compressed sensing. *Magn. Reson. Med.* 2009;62(6):1574–84.
39. Feng L, Srichai MB, Lim RP, Harrison A, King W, Adluru G, Dibella EV, Sodickson DK, Otazo R, Kim D. Highly accelerated real-time cardiac cine MRI using k-t SPARSE-SENSE. *Magn Reson Med.* 2013;70(1):64–74.
40. Lustig M, Pauly JM. SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magn Reson Med.* 2010;64(2):457–71.
41. Zhang T, Pauly JM, Vasanawala SS, Lustig M. Coil compression for accelerated imaging with Cartesian sampling. *Magn Reson Med.* 2013;69(2):571–82.
42. Feng L, Grimm R, Block KT, Chandarana H, Kim S, Xu J, Axel L, Sodickson DK, Otazo R. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn Reson Med.* 2014;72(3):707–17.
43. Trzasko J, Manduca A. Fast spiral coronary artery imaging. *IEEE Trans. Med. Imaging.* 2009;28(1):106–21.
44. Trzasko JD, Haider CR, Borisch EA, Campeau NG, Glockner JF, Riederer SJ, Manduca A. Sparse-CAPR: highly accelerated 4D CE-MRA with parallel imaging and nonconvex compressive sensing. *Magn Reson Med.* 2011;66(4):1019–32.
45. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. *Magn Reson Med.* 2011;65(2):480–91.
46. Knoll F, Clason C, Bredies K, Uecker M, Stollberger R. Parallel imaging with nonlinear reconstruction using variational penalties. *Magn Reson Med.* 2012;67(1):34–41.
47. Sung K, Daniel BL, Hargreaves BA. Location constrained approximate message passing for compressed sensing MRI. *Magn Reson Med.* 2013;70(2):370–81.

48. Ravishanker S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imaging*. 2011;30(5):1028–41.
49. 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(5):1042–54.
50. Asif MS, Hamilton L, Brummer M, Romberg J. Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI. *Magn Reson Med*. 2012;7:800–12.
51. Ak Cakaya M, Basha TA, Goddu B, Goepfert LA, Kissinger KV, Tarokh V, Manning WJ, Nezafat R. Low-dimensional-structure self-learning and thresholding: Regularization beyond compressed sensing for MRI Reconstruction. *Magn Reson Med*. 2011;66(3):756–67.
52. Trzasko JD, Manduca A. Calibrationless parallel MRI using CLEAR. In: *IEEE Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*; 2011. p. 75–9.
53. Yoon H, Kim KS, Kim D, Bresler Y, Ye JC. Motion adaptive patch-based low-rank approach for compressed sensing cardiac cine MRI. *IEEE Trans. Med. Imaging*. 2014;33(11):2069–85.
54. Candès EJ, Recht B. Exact matrix completion via convex optimization. *Found. Comput. Math*. 2009;9(6):717–72.
55. Shin PJ, Larson PE, Ohliger MA, Elad M, Pauly JM, Vigneron DB, Lustig M. Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. *Magn Reson Med*. 2014;72(4):959–70.
56. Haldar JP. Low-rank modeling of local-space neighborhoods (LORAKS) for constrained MRI. *IEEE Trans. Med. Imaging*. 2014;33(3):668–81.
57. Jin KH, Lee D, Ye JC. A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix. *IEEE Trans Computational Imag*. 2016;2(4):480–95.
58. Jin KH, Um J-Y, Lee D, Lee J, Park S-H, Ye JC. MRI artifact correction using sparse+ low-rank decomposition of annihilating filter-based Hankel matrix. *Magnetic Reson Med*. 2017;78(1):327–40.
59. Lee D, Jin KH, Kim EY, Park S-H, Ye JC. Acceleration of MR parameter mapping using annihilating filter-based low rank hankel matrix (ALOHA). *Magn Reson Med*. 2016;76(6):1848–68.
60. Lee J, Jin KH, Ye JC. Reference-free single-pass EPI Nyquist ghost correction using annihilating filter-based low rank Hankel matrix (ALOHA). *Magn Reson Med*. 2016;76(8):1775–89.
61. Ongie G, Jacob M. Off-the-grid recovery of piecewise constant images from few fourier samples. *SIAM J Imaging Sci*. 2016;9(3):1004–41.
62. Ye JC, Kim JM, Jin KH, Lee K. Compressive sampling using annihilating filter-based low-rank interpolation. *IEEE Trans Inf Theory*. 2017;63(2):777–801.
63. Bauer S, Markl M, Honal M, Jung B. The effect of reconstruction and acquisition parameters for GRAPPA-based parallel imaging on the image quality. *Magn Reson Med*. 2011;66(2):402–9.
64. Huang F, Lin W, Duensing GR, Reykowski A. A hybrid method for more efficient channel-by-channel reconstruction with many channels. *Magn Reson Med*. 2012;67(3):835–43.
65. Trzasko JD, Bao Z, Manduca A, McGee KP, Bernstein MA. Sparsity and low-contrast object detectability. *Magn Reson Med*. 2012;67(4):1022–32.
66. Johnson KM, Block WF, Reeder SB, Samsonov A. Improved least squares MR image reconstruction using estimates of k-space data consistency. *Magn Reson Med*. 2012;67(6):1600–8.
67. Park S, Park J. Adaptive self-calibrating iterative GRAPPA reconstruction. *Magn Reson Med*. 2012;67(6):1721–9.
68. Liang D, DiBella EV, Chen R-R, Ying L. k-t ISD: dynamic cardiac MR imaging using compressed sensing with iterative support detection. *Magn Reson Med*. 2012;68(1):41–53.
69. Chang Y, Liang D, Ying L. Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction. *Magn Reson Med*. 2012;68(3):730–40.
70. Huang F, Lin W, Duensing GR, Reykowski A. k-t sparse GROWL: Sequential combination of partially parallel imaging and compressed sensing in k-t space using flexible virtual coil. *Magn Reson Med*. 2012;68(3):772–82.
71. Jung B, Stalder AF, Bauer S, Markl M. On the undersampling strategies to accelerate time-resolved 3D imaging using k-t-GRAPPA. *Magn Reson Med*. 2011;66(4):966–75.
72. Piccini D, Littmann A, NIELLES-Vallespin S, Zenge MO. Spiral phyllotaxis: the natural way to construct a 3D radial trajectory in MRI. *Magn Reson Med*. 2011;66(4):1049–56.
73. Lin W, B'ornert P, Huang F, Duensing GR, Reykowski A. Generalized GRAPPA operators for wider spiral bands: Rapid self-calibrated parallel reconstruction for variable density spiral MRI. *Magn Reson Med*. 2011;66(4):1067–78.
74. Pipe JG, Zwart NR, Aboussouan EA, Robison RK, Devaraj A, Johnson KO. A new design and rationale for 3d orthogonally oversampled k-space trajectories. *Magn Reson Med*. 2011;66(5):1303–11.
75. Pang Y, Vigneron DB, Zhang X. Parallel traveling-wave MRI: A feasibility study. *Magn Reson Med*. 2012;67(4):965–78.
76. Wang H, Liang D, King KF, Nagarsekar G, Chang Y, Ying L. Improving GRAPPA using cross-sampled autocalibration data. *Magn Reson Med*. 2012;67(4):1042–53.
77. Witschey WR, Cocosco CA, Gallichan D, Schultz G, Weber H, Welz A, Hennig J, Zaitsev M. Localization by nonlinear phase preparation and k-space trajectory design. *Magn Reson Med*. 2012;67(6):1620–32.
78. Addy NO, Wu HH, Nishimura DG. Simple method for MR gradient system characterization and k-space trajectory estimation. *Magn Reson Med*. 2012;68(1):120–9.
79. Zahneisen B, Hugger T, Lee KJ, LeVan P, Reisert M, Lee H-L, Assl'ander J, Zaitsev M, Hennig J. Single shot concentric shells trajectories for ultra fast fMRI. *Magn Reson Med* 68(2), 484–494 (2012)
80. Turley DC, Pipe JG. Distributed spirals: a new class of three-dimensional k-space trajectories. *Magn Reson Med*. 2013;70(2):413–9.
81. Layton KJ, Gallichan D, Testud F, Cocosco CA, Welz AM, Barmet C, Pruessmann KP, Hennig J, Zaitsev M. Single shot trajectory design for region-specific imaging using linear and nonlinear magnetic encoding fields. *Magn Reson Med*. 2013;70(3):684–96.
82. Haldar JP, Hernando D, Liang Z-P. Compressed-sensing MRI with random encoding. *IEEE Trans Med Imag*. 2011;30(4):893–903.
83. Puy G, Marques JP, Gruetter R, Thiran J-P, Ville DVD, Vanderheyneyst P, Wiaux Y. Spread spectrum magnetic resonance imaging. *IEEE Trans Med Imag*. 2012;31(3):586–98.
84. Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magn Reson Med*. 2010;64(3):767–76.
85. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M. ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magn Reson Med*. 2014;71(3):990–1001.
86. Ravishanker S, Bresler Y. Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging. *SIAM J Imag Sci*. 2015;8(4):2519–57.
87. Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*. 2006;54(11):4311–22.
88. Lustig M, Santos JM, Donoho DL, Pauly JM. kt SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. In: *Proceedings of the 13th Annual Meeting of ISMRM, Seattle, vol. 2420*; 2006.
89. Le Gall D. MPEG: A video compression standard for multimedia applications. *Commun ACM*. 1991;34(4):46–58.
90. Jung H, Ye JC. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques. *Int J Imag Syst Technol*. 2010;20(2):81–98.
91. Haldar, J.P., Liang, Z.-P.: Spatiotemporal imaging with partially separable functions: a matrix recovery approach. In: *IEEE International Symposium on Biomedical Imaging*, pp. 716–719 (2010). IEEE
92. Haldar JP, Hernando D. Rank-constrained solutions to linear matrix equations using powerfactorization. *IEEE Signal Process Letters*. 2009;16(7):584–7.
93. Zhao B, Haldar JP, Christodoulou AG, Liang Z-P. Image reconstruction from highly undersampled-space data with joint partial separability and sparsity constraints. *IEEE Trans Med Imag*. 2012;31(9):1809–20.
94. Cai J-F, Candès EJ, Shen Z. A singular value thresholding algorithm for matrix completion. *SIAM J Optimization*. 2010;20(4):1956–82.
95. Vetterli M, Marziliano P, Blu T. Sampling signals with finite rate of innovation. *IEEE Trans Signal Process*. 2002;50(6):1417–28.
96. Dragotti PL, Vetterli M, Blu T. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Trans Signal Process*. 2007;55(5):1741–57.
97. Maravic I, Vetterli M. Sampling and reconstruction of signals with finite rate of innovation in the presence of noise. *IEEE Trans Signal Process*. 2005;53(8):2788–805.



98. Unser M, Tafti PD, Sun Q. A unified formulation of Gaussian versus sparse stochastic processes—Part I: Continuous-domain theory. *IEEE Trans Inf Theory*. 2014;60(3):1945–62.
99. Unser M, Tafti PD, Amini A, Kirshner H. A unified formulation of Gaussian versus sparse stochastic processes—Part II: Discrete-domain theory. *IEEE Trans Inf Theory*. 2014;60(5):3036–51.
100. Hsiao A, Lustig M, Alley MT, Murphy M, Chan FP, Herfkens RJ, Vasanawala SS. Rapid pediatric cardiac assessment of flow and ventricular volume with compressed sensing parallel imaging volumetric cine phase-contrast MRI. *Am J Roentgenol*. 2012;198(3):250–9.
101. Vincenti G, Monney P, Chaptinel J, Rutz T, Coppo S, Zenge MO, Schmidt M, Nadar MS, Piccini D, Ch'evre P, et al. Compressed sensing single-breath-hold CMR for fast quantification of LV function, volumes, and mass. *JACC: Cardiovasc Imag*. 2014;7(9):882–92.
102. Chandarana H, Feng L, Block TK, Rosenkrantz AB, Lim RP, Babb JS, Sodickson DK, Otazo R. Free-breathing contrast-enhanced multiphase MRI of the liver using a combination of compressed sensing, parallel imaging, and golden-angle radial sampling. *Investigative Radiol*. 2013;48(1):10–16.
103. Espagnet MR, Bangiyev L, Haber M, Block K, Babb J, Ruggiero V, Boada F, Gonen O, Fatterpekar G. High-resolution DCE-MRI of the pituitary gland using radial k-space acquisition with compressed sensing reconstruction. *Am J Neuroradiol*. 2015;36(8):1444–9.
104. Landman BA, Bogovic JA, Wan H, ElShahaby FEZ, Bazin P-L, Prince JL. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI. *NeuroImage*. 2012;59(3):2175–86.
105. Kuhnt D, Bauer MH, Egger J, Richter M, Kapur T, Sommer J, Merhof D, Nimsky C. Fiber tractography based on diffusion tensor imaging compared with high-angular-resolution diffusion imaging with compressed sensing: initial experience. *Neurosurgery*. 2013;72(0 1):165.
106. Larson PE, Hu S, Lustig M, Kerr AB, Nelson SJ, Kurhanewicz J, Pauly JM, Vigneron DB. Fast dynamic 3D MR spectroscopic imaging with compressed sensing and multiband excitation pulses for hyperpolarized <sup>13</sup>C studies. *Magn Reson Med*. 2011;65(3):610–9.
107. Geethanath S, Baek H-M, Ganji SK, Ding Y, Maher EA, Sims RD, Choi C, Lewis MA, Kodibagkar VD. Compressive sensing could accelerate 1H MR metabolic imaging in the clinic. *Radiology*. 2012;262(3):985–94.
108. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, Knoll F. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med*. 2018;79(6):3055–71.
109. Han Y, Yoo J, Kim HH, Shin HJ, Sung K, Ye JC. Deep learning with domain adaptation for accelerated projection-reconstruction mr. *Magn Reson Med*. 2018;80(3):1189–205.
110. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature*. 2018;555(7697):487.
111. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE Trans Med Imag*. 2018;37(2):491–503.
112. Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., Liang, D.: Accelerating magnetic resonance imaging via deep learning. In: *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium On*, pp. 514–517 (2016). IEEE
113. Kwon K, Kim D, Park H. A parallel MR imaging method using multilayer perceptron. *Med Phys*. 2017;44(12):6209–24.
114. Ye JC, Han Y, Cha E. Deep convolutional framelets: A general deep learning framework for inverse problems. *SIAM J Imag Sci*. 2018;11(2):991–1048.