A General Framework for Compressed Sensing and Parallel MRI Using Annihilating Filter Based Low-Rank Hankel Matrix

Kyong Hwan Jin, Dongwook Lee, and Jong Chul Ye, *Senior Member, IEEE*

***Abstract*—Parallel MRI (pMRI) and compressed sensing MRI (CS-MRI) have been considered as two distinct reconstruction problems. Inspired by recent k-space interpolation methods, an annihilating filter-based low-rank Hankel matrix approach is proposed as a general framework for sparsity-driven k-space interpolation method which unifies pMRI and CS-MRI. Specifically, our framework is based on a novel observation that the transform domain sparsity in the primary space implies the low-rankness of weighted Hankel matrix in the reciprocal space. This converts pMRI and CS-MRI to a k-space interpolation problem using a structured matrix completion. Experimental results using *in vivo* data for single/multicoil imaging as well as dynamic imaging confirmed that the proposed method outperforms the state-of-the-art pMRI and CS-MRI.**

***Index Terms*—Annihilating filter, cardinal spline, compressed sensing, parallel MRI, pyramidal representation, structured low rank block Hankel matrix completion, wavelets.**

I. Introduction

MAGNETIC resonance imaging (MRI) is an imaging system that sequentially acquires k-space data corresponding to the Fourier transform of an object. This enables us to apply various advanced signal processing techniques. Recently, compressed sensing theory [1], [2] has been extensively employed in accelerated MRI [3]–[5]. Compressed sensing algorithms can restore original signals from much less k-space data by exploiting the sparsity of an unknown image in the total variation (TV) or wavelet transform domains, and incoherent sampling schemes such as Gaussian random or Poisson disc are usually required. Accurate MRI reconstruction from less data makes compressed sensing a hot topic in the research community; thus, it has been applied across many different application areas such as in pediatric imaging [6], dynamic cardiac MRI [7]–[9], perfusion imaging [10], angiography [11], and so on.

Manuscript received December 18, 2015; revised June 9, 2016; accepted July 21, 2016. Date of publication August 18, 2016; date of current version November 4, 2016.This work was supported by Korea Science and Engineering Foundation under Grant NRF-2014R1A2A1A11052491 and NRF-2016R1A2B3008104. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ali Bilgin.

K. H. Jin was with the Department of Bio and Brain Engineering, KAIST, Daejeon 305-701, South Korea. He is now with the Biomedical Imaging Group, Ecole Polytechnique Federale de Lausanne, Lausanne 1015, Switzerland (e-mail: kyonghwan.jin@gmail.com).

D. Lee and J. C. Ye are with the Department of Bio and Brain Engineering, KAIST, Daejeon 305-701, South Korea (e-mail: jong.ye@kaist.ac.kr; zeronimo@kaist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2016.2601296

On the other hand, parallel MRI (pMRI) [12]–[14] 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. Representative parallel imaging algorithms such as SENSE (sensitivity encoding) [12] or GRAPPA (generalized autocalibrating partially parallel acquisitions) [13] require regularly sampled k-space data for computationally efficient reconstruction. Moreover, additional k-space data, the so-called auto calibration (ACS) lines, are often required to estimate the coil sensitive maps or GRAPPA kernels [13].

Because the aim of the two approaches is an accelerated acquisition by reducing the k-space data, extensive research efforts have been made to synergistically combine the two for further acceleration. 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:

$$\min_{\mathbf{f}} \|W\mathbf{f}\|_1 \quad \text{subject to} \quad \mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_r \end{bmatrix} = \begin{bmatrix} A[S_1] \\ \vdots \\ A[S_r] \end{bmatrix} \mathbf{f} \tag{1}$$

where $\mathbf{f}$ and $\mathbf{g}_i$ denote the unknown image and the k-space measurements from the i-th coil, respectively; A is a subsampled Fourier matrix; W is a sparsifying transform, and $[S_i]$ denotes a diagonal matrix whose diagonal elements come from the i-th coil sensitivity map. The multichannel version of k-t FOCUSS [7] is one of the typical examples of such approaches. On the other hand, l_1-SPIRiT (l_1- iTerative Self-consistent Parallel Imaging Reconstruction) [15] utilizes the GRAPPA type constraint as an additional constraint for a compressed sensing problem:

$$\min_{F} \quad \|\Psi F\|_{1,2}$$

$$\text{subject to} \quad G = AF, \text{VEC}(F) = M \cdot \text{VEC}(F)$$

where $\|\cdot\|_{1,2}$ denotes the $(1,2)$-mixed norm of a matrix, $F = [\,\mathbf{f}_1\ \mathbf{f}_2\ \ldots\ \mathbf{f}_r\,]$, $G = [\,\mathbf{g}_1\ \mathbf{g}_2\ \ldots\ \mathbf{g}_r\,]$ denote the unknown images and their k-space measurements for the given set of coils, and Ψ denote a discrete wavelet transform matrix, and M is an image domain GRAPPA operator, and $\text{VEC}(\cdot)$ is the vectorization operator. In both approaches, an accurate estimation of coil sensitivity maps or GRAPPA kernel is essential to fully exploit the coil sensitivity diversity.

In order to overcome these difficulties, calibration-less parallel imaging methods have been extensively investigated, among

2333-9403 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

which SAKE (simultaneous autocalibrating and k-space estimation) [16] represents one of the first steps. In SAKE, the missing k-space elements are reconstructed by imposing the data consistency and the structural maintenance constraints of the block Hankel structure matrix. However, the origin of the low rankness in the Hankel structured matrix for the case of a single coil measurement was not addressed, so it was not clear whether SAKE could outperform the compressed sensing approach when it is applied to single coil data. Haldar [17], [18] later discovered that a Hankel structured matrix constructed by a single coil k-space measurement is low-ranked when an unknown image has finite support or a slow-varying phase. Based on this observation, he developed the so called LORAKS (Low-rank modeling of local k-space neighborhoods) algorithm [17] and its parallel imaging version, P-LORAKS (Low-rank modelling of local k-space neighborhoods with parallel imaging data) [18]. However, it was not clear how the existing theory can deal with large classes of image models that are not finite supported but can be sparsified using various transforms such as wavelet transforms or total variations (TV), etc.

Therefore, one of the main goals of this paper is to develop a theory that unifies and generalizes k-space low-rank methods to also allow for transform sparsity models which are critical for practical MRI applications. Toward this goal, we show that the transform domain sparsity in the signal space can be directly related to the existence of *annihilating filters* [19]–[21] in the *weighted* k-space. Interestingly, the commutative relation between an annihilating filter and weighted k-space measurements provides a rank-deficient Hankel structured matrix, whose rank is determined by the sparsity level of the underlying signal in the transform domain. Therefore, by performing a low-rank matrix completion approach, the missing weighted k-space data in the Hankel structured matrix can be recovered, after which the original k-space data can be recovered by removing the weights.

Interestingly, our new framework is so general that it can generalize the existing compressed sensing MR approaches in very unique ways. For example, even though the original authors did not explicitly mention, C-based LORAKS in [17] indeed utilizes a special case of annihilating filter that exploits the image domain sparsity in term of finite support condition. Moreover, if an image can be sparsified with wavelet transforms, the low rank structured matrix completion problem can be solved using a pyramidal decomposition after applying scale dependent k-space weightings. In addition, we show that there exist additional inter-coil annihilating filter relationships that are unique in pMRI, which can be utilized to construct a concatenated Hankel matrix that is low ranked.

Another important advantage of the proposed algorithm is that, compared to the existing CS-MRI, the reconstruction errors are usually scattered throughout the entire images rather than exhibiting systematic distortion along edges because the annihilating filter relationships are specifically designed for estimating the edge signals. Given that many diagnostic errors are caused by the systematic distortion of images, we believe that our annihilating filter-based low rank Hankel matrix approach (ALOHA) framework may have a great potential in clinical applications.

The remainder of this paper is as follows. Section II-B discusses the relationship between the transform domain sparsity and the low-rankness of Hankel structured matrix in weighted k-space. In Section III, pyramidal decomposition and parallel imaging version of the proposed method will be provided. Section IV then explains the implementation detail. Experimental results are provided in Section V, which is followed by the discussion in Section VI and conclusion in Section VII.

II. Theory

A. Notations

A $(n-d+1)\times d$ Hankel structured matrix generated from an n-dimensional vector $\mathbf{x}=[x[0],\ldots,x[n-1]]^T\in\mathbb{C}^n$ has the following structure:

$$\mathscr{H}(\mathbf{x})=\begin{bmatrix} x[0] & x[1] & \cdots & x[d-1] \\ x[1] & x[2] & \cdots & x[d] \\ \vdots & \vdots & \ddots & \vdots \\ x[n-d] & [n-d+1] & \cdots & x[n-1] \end{bmatrix}. \quad (2)$$

where d is called the *matrix pencil parameter.* We denote the space of this type of Hankel structure matrices as $\mathcal{H}(n,d)$. An $n\times d$ *wrap-around* Hankel matrix generated from an n-dimensional vector $\mathbf{u}=[u[0],\ldots,u[n-1]]^T\in\mathbb{C}^n$ is defined as:

$$\mathscr{H}_c(\mathbf{u})=\left[\begin{array}{cccc} u[0] & u[1] & \cdots & u[d-1] \\ u[1] & u[2] & \cdots & u[d] \\ \vdots & \vdots & \ddots & \vdots \\ u[n-d] & u[n-d+1] & \cdots & u[n-1] \\ \hline u[n-d+1] & u[n-d+2] & \cdots & u[0] \\ \vdots & \vdots & \ddots & \vdots \\ u[n-1] & u[0] & \cdots & u[d-2] \end{array}\right].$$

Note that $n\times d$ wrap-around Hankel matrix is equivalent to the standard Hankel matrix with respect to an augumented vector with the periodic boundary expansion:

$$\tilde{\mathbf{u}}=\left[\mathbf{u}^T\ \underbrace{u[0]\,u[1]\,\ldots\,u[d-2]}_{(d-1)}\right]^T\in\mathbb{C}^{n+d-1}.$$

Therefore, the two terms - wrap-around Hankel matrix and Hankel matrix with periodic boundary conditions - will be used interchangeably.

B. Transform Domain Sparsity and Low-Rankness in Weighted k-Space

Here, we describe the relationship between transform domain sparsity and low rankness of *weighted* Hankel matrix, which is the key idea of the proposed algorithm. For better readability, the theory here is outlined by assuming 1-D signals, but the principle can be extended for multidimensional signals [22].

Note that typical signals of our interest may not be sparse in the image domain, but can be sparsified in a transform domain. For example, consider a L-spline signal model [23], [24].

Specifically, the signal f of our interest is assumed to satisfy the following partial differential equation:

$$\mathrm{L}f = w \tag{3}$$

where L denotes a constant coefficient linear differential equation (or whitening operator in [23], [24]):

$$\mathrm{L} := a_K \partial^K + a_{K-1}\partial^{K-1} + \cdots + a_1 \partial + a_0 \tag{4}$$

and w is a driving continuous domain sparse signal or *sparse innovation* given by

$$w(x) = \sum_{j=0}^{r-1} c_j \delta\left(x - x_j\right) \qquad x_j \in [0, \tau]. \tag{5}$$

Here, without loss of generality, we set $\tau = n$ for a positive integer $n \in \mathbb{Z}$. This model includes many class of signals with the finite rate of innovations [19]–[21]. For example, if the underlying signal is piecewise constant, we can set L as the first differentiation. In this case, f corresponds to the total variation (TV) signal model, and this TV signal model will be extensively used throughout the paper.

Now, by taking the Fourier transform of (3), we have

$$\hat{y}(\omega) := \mathcal{F}\{\mathrm{L}f(x)\} = \hat{l}(\omega)\hat{f}(\omega) = \sum_{j=0}^{r-1} a_j e^{-i\omega x_j} \tag{6}$$

where

$$\hat{l}(\omega) = a_K (i\omega)^K + a_{K-1}(i\omega)^{K-1} + \ldots + a_1 (i\omega) + a_0 \ .$$

In the standard Nyquist sampling, we should measure discrete set of Fourier samples from a deterministic grid, whose grid size should be set to the Nyquist limit $\Delta = 2\pi/n$ to avoid aliasing artifacts; so the discrete spectrum can be represented as

$$\hat{y}[k] := \left.\hat{y}(\omega)\right|_{\omega = k\Delta} = \hat{l}[k\Delta]\hat{f}[k\Delta] = \sum_{j=0}^{r-1} c_j e^{-i2\pi k x_j / n}, \tag{7}$$

for $k \in [0, \ldots, n-1]$. The discrete spectral sampling model in Eq. (7) implies that the unknown signal in the image domain is a periodic streams of Diracs with a period n, which is indeed a signal with the finite rate of innovation (FRI) with rate $\rho = 2k/n$ [19]–[21]. Therefore, theoretical results from the FRI sampling theory can be used [19]–[21], which tells us that we can find a minimum length annihilating filter $\hat{h}[k]$ such that

$$(\hat{h} * \hat{y})[k] = \sum_{l=0}^{r} \hat{h}[l]\hat{y}[k-l] = 0, \quad \forall k. \tag{8}$$

The specific form of the *minimum length* annihilating filter $\hat{h}[k]$ for the case of (7) can be found in [19], which has the following z-transform

$$\hat{h}(z) = \sum_{l=0}^{k} \hat{h}[l] z^{-l} = \prod_{j=0}^{r-1} (1 - e^{-i2\pi x_j / n} z^{-1}) \ , \tag{9}$$

whose filter length is $r+1$ [19].

Now, let $\mathbf{y} = [\,\hat{y}[0] \ \ldots \ \hat{y}[n-1]\,]$ and $\min\{n-d+1, d\} > r$. Suppose, futhermore, that a Hankel structure matrix $\mathscr{H}(\hat{\mathbf{y}}) \in \mathcal{H}(n,d)$ is constructed by

$$\mathscr{H}(\hat{\mathbf{y}}) = \begin{bmatrix} \hat{y}[0] & \hat{y}[1] & \cdots & \hat{y}[d-1] \\ \hat{y}[1] & \hat{y}[2] & \cdots & \hat{y}[d] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}[n-d] & \hat{y}[n-d+1] & \cdots & \hat{y}[n-1] \end{bmatrix} \tag{10}$$

then we can show the following key result:

Theorem 2.1: Let $r+1$ denote the length of the minimum size of annihilating filters that annihilates discrete Fourier data $\hat{y}[k]$. Assume that $\min\{n-d+1, d\} > r$. Then, for a given Hankel structured matrix in (10), we have

$$\mathrm{RANK}\mathscr{H}(\hat{\mathbf{y}}) = r, \tag{11}$$

where $\mathrm{RANK}(\cdot)$ denotes a matrix rank.

Proof: See our companion paper [25].

Indeed, the proof of Theorem 2.1 informs us that the explicit form of the Fourier samples $\hat{y}[k]$ given by

$$\hat{y}[k] := \sum_{s=0}^{p-1} \sum_{l=0}^{m_s - 1} c_{s,l} k^l {d_s}^k \text{ for } 0 \le k \le n-1,$$

with

$$r = \sum_{s=0}^{p-1} m_s$$

is necessary and sufficient condition to have a rank-r Hankel matrix. Accordingly, we concluded that the signals with the finite rate of innovation (FRI) correspond to this class signals [25].

We further showed that for the case of the cardinal spline cases (where the knots $\{x_j\}$ are located on the uniform grid), the k-space data is also periodic; accordingly, a wrap-around Hankel matrix can be equivalently obtained from the periodic boundary condition [25]. Moreover, the spectral weighting comes from the DFT of the discrete whitening filter that has attenuating behaviour at high frequency regions, which makes the algorithm more robust to noise boosting. See [25] for more details. This cardinal spline model will be used throughout the paper for MR specific applications, where the unknown images should be reconstructed on a fixed grid.

Now, it is important to emphasize that Theorem 2.1 implies the following relationship:

$$\text{FRI signals} \quad \stackrel{\mathcal{F}}{\Longrightarrow} \quad \text{low-ranked weighted Hankel structured matrix,}$$

where $\mathcal{F}$ denotes the Fourier transform. Therefore, if some of k-space data of a FRI signals are missing, we can construct an appropriate weighted Hankel matrix with missing elements, which are recovered based on low rank matrix completion [26]–[30]:

$$\begin{aligned} (P) \quad & \min_{\mathbf{m} \in \mathbb{C}^n} \ \mathrm{RANK}\mathscr{H}(\mathbf{m}) \\ & \text{subject to} \ \ P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{y}}) \ , \end{aligned} \tag{12}$$

where

$$\hat{\mathbf{y}} = \mathbf{l} \odot \hat{\mathbf{f}}$$

and $\odot$ denotes the Hadamard product, and $\hat{\mathbf{l}}$ and $\hat{\mathbf{f}}$ denotes the vectors composed of discrete samples of $\hat{l}(\omega)$ and $\hat{f}(\omega)$, respectively. By solving (P) we can obtain the missing data $m(\omega) = \hat{l}(\omega)\hat{f}(\omega)$ in the Fourier domain. Then, the missing spectral data $\hat{f}(\omega)$ can be obtained by dividing by the weight, i.e. $\hat{f}(\omega) = m(\omega)/\hat{l}(\omega)$, when $\hat{l}(\omega) \neq 0$. As for the signal $\hat{f}(\omega)$ at the spectral null of the filter $\hat{l}(\omega)$, the corresponding elements should be specifically obtained as sampled measurements, which can be easily done in MR acquisition.

C. Sampling Rate, Stability and Compressibility

Among various algorithm to solve matrix completion problem (P), one of the most well-characterised approaches is a convex relaxation approach using the nuclear norm [26], [27], [29], [30]. More specifically, the missing k-space elements can be found by solving the following nuclear norm minimization problem:

$$(P1) \quad \min_{\mathbf{m}\in\mathbb{C}^n} \ \|\mathscr{H}(\mathbf{m})\|_* \quad \text{subject to} \ \ P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{y}}) \tag{13}$$

where $\|\cdot\|_*$ denotes the matrix nuclear norm. Therefore, the remaining question is to verify whether the low-rank matrix completion approach $(P1)$ does not compromise any optimality compared to the standard Fourier CS with the l_1 sparsity penalty, which is the main interest in this section.

Toward this goal, we first derive an algebraic bound. Considering that $\min\{n-d+1, d\} \geq r+1$ to allow rank deficient Hankel matrix, we have $n \geq r+d \geq 2r+1$. In fact, by taking all n samples as measurements, we can obtain the bunched sampling pattern used in the classical sampling theory of FRI signals [19]–[21], where the number of required samples are equal to

$$m \geq 2r+1. \tag{14}$$

Note that the minimum sampling rate in (14) is indeed equal to the algebraic bound of the standard compressed sensing approach [1], [2]. Second, we are interested in deriving the performance guarantee with random Fourier samples. Here, the notion of the incoherence plays a crucial role. We recall the definitions using our notations. Suppose that $M \in \mathbb{C}^{n_1\times n_2}$ is a rank-r matrix whose SVD is $U\Sigma V^*$. M is said to satisfy the *standard incoherence* condition with parameter μ if

$$\max_{1\leq i\leq n_1} \|U^*\mathbf{e}_i\|_2 \leq \sqrt{\frac{\mu r}{n_1}}, \ \max_{1\leq j\leq n_2} \|V^*\mathbf{e}_j\|_2 \leq \sqrt{\frac{\mu r}{n_2}}, \tag{15}$$

where $\mathbf{e}_i$ denotes the standard unit coordinate vector with 1 at the i-th elements, and zeros in all other locations. Then, by extending the result by Chen and Chi [31], we obtained a general performance guarantee [25]:

Theorem 2.2 ([25]): Let $\Omega = \{j_1, \ldots, j_m\}$ be a multi-set consisting of random indices where j_k's are i.i.d. following the uniform distribution on $\{0, \ldots, n-1\}$. Suppose that a Hankel matrix is constructed from $\hat{\mathbf{y}} \in \mathbb{C}^n$ is of rank-r and satisfies the standard incoherence condition in (15) with parameter μ. Then, there exists an absolute constant c_1 such that $\hat{\mathbf{y}}$ is the unique minimizer to (13) with probability $1 - 1/n^2$, provided

$$m \geq c_1 \mu c_s r \log^\alpha n, \tag{16}$$

where $\alpha = 2$ and $c_s := n/d$ if the Hankel matrix has the wrap-around property; $\alpha = 4$ and $c_s := \max\{n/(n-d+1), n/d\}$, otherwise.

Proof: See our companion paper [25].

In compressed sensing MRI, the reconstruction image grid is always fixed; therefore, cardinal spline model that has knot locations on a fixed grid is more appropriate. Therefore, the results in (16) implies that the required number of samples are proportional to the sparsity level up to $\log^2(n)$ factor. Considering that the standard compressed sensing analysis showed that the required number of Fourier samples in the l_1 minimization approach is proportional to the sparsity level up to $\log^q(n)$ factor for some integer q [1], [2], the result in (16) is comparable to the standard CS-MRI approach.

In practice, we observed that the number of k-space samples are only a few times the sparsity level, which are not far from the algebraic bound (14). This will be again confirmed by the empirical verification in our paper. Moreover, controlled numerical experiments in [25] confirmed that the proposed low-rank matrix completion approach is consistently better than the standard CS in various signal models. This suggests an important observation: by reformulating the compressed sensing problem as a low rank Hankel structured matrix completion problem in the measurement domain, we may not expect any performance loss.

In Theorem II.3 of our companion paper [25], we also considered the recovery of $\hat{\mathbf{y}}$ from its partial entries with noise, and derived a more improved version of the stability results than that of Chen and Chi [31]. Note that the stability result in [25] can be used to show the *compressibility* of the proposed structured low-rank matrix completion method. In the compressed sensing literatures, the compressibility implies that the dominant r-coefficients can be *stably* recovered by compressed sensing algorithms even though the underlying signal is not perfectly r-sparse [1], [2]. More specifically, let $y(x)$ be s-sparse FRI signal. Suppose, furthermore, that the sampling rate in (16) is only sufficient to recover r-dominant sparse coefficients with $r < s$. Then, the contribution of the remaining $s-r$ non-zero coefficients works as noises in the Fourier domain with an appropriate noise bound δ. Therefore, the stability result in [25] guarantees the stable recovery of r-dominant sparse coefficients, implying that the proposed method can be used to recovery signals which is not perfectly r-sparse. Recall that this compressibility is one of the most basic and fundamental components of compressed sensing, making CS approaches very effective for reconstructing MR image that is not perfectly sparse [1], [2].

III. ALOHA for Accelerated MRI

Inspired by the theoretical finding in the previous section, this section will explain two realizations of the low-rank matrix completion approaches that are useful for MR applications. The first one is wavelet-based pyramidal decomposition approach and the second one is a generalization for multichannel

parallel MRI. Note that these are particular instances of the ALOHA algorithm and other variations of ALOHA may be possible for various scenarios as demonstrated in recent applications [32]–[34].

A. ALOHA for Wavelet Sparse Signals

One of the technical issues associated with the wavelet analysis is that the standard wavelet multi-resolution analysis are usually conducted on fixed grid; so, the signal model in (3) should be modified. Toward this goal, we found that a cardinal spline model [23], [24], [35] is very useful.

A cardinal spline is a special case of a non-uniform spline where the knots are located on the integer grid. More specifically, a function $f(x)$ is called a *cardinal* L-*spline* if and only if

$$\mathrm{L}f(x) = w(x), \tag{17}$$

where the operator L is continuous domain whitening operator and the continuous domain *sparse innovation* signal $w(x)$ is given by

$$w(x) := \sum_{p\in\mathbb{Z}} a[p]\delta(x-p)\ , \tag{18}$$

whose singularities are located on integer grid. Here, we assume that the number of nonzero coefficients $\{a[n]\}$ on the integer grid is r.

One of the main advantages of using cardinal setup over the non-uniform splines is that we can recover signals by exploiting the sparseness of *sampled* signal rather than exploiting off the grid singularity. More specifically, consider "L-compatible" generalized wavelets which, at a given resolution level s, are such that[1]

$$\psi_s(x) = \mathrm{L}^*\phi_s(x). \tag{19}$$

Here, L^* is the adjoint operator of L and $\phi_s(x)$ is some smoothing kernel with good localization properties. Then, the wavelet analysis gives us

$$\begin{aligned} u_s(x) &= \langle f, \psi_s(\cdot - x)\rangle \\ &= \langle f(\cdot), \mathrm{L}^*\phi_s(\cdot - x)\rangle \\ &= \langle \mathrm{L}f(\cdot), \phi_s(\cdot - x)\rangle \\ &= \langle w(\cdot), \phi_s(\cdot - x)\rangle = (\bar{\phi}_s * w)(x) \\ &= \sum_{i\in\mathbb{Z}} a[i]\bar{\phi}_s(x-i) \end{aligned} \tag{20}$$

where $\bar{\phi}_s(x) = \phi_s(-x)$. Note that $u_s(x)$ is indeed a *smoothed* version of continuous domain innovation $w(x)$ in (18), because all the sparsity information of the innovation $w(x)$ is encoded in its coefficients $\{a[i]\}$, and aside from the interpolant $\bar{\phi}_s(\cdot)$, $u_s(x)$ in (20) still retains the same coefficients of the original innovation (18).

In particular, if the underlying L-cardinal spline signal is a TV signal, i.e. $\mathrm{L} = \frac{d}{dx}$, then we can define the s-scale wavelet as

$$\psi_s(x) = 2^{-s/2}\psi_{\mathrm{Haar}}\left(\frac{x}{2^s}\right)\ , \tag{21}$$

where the centered Haar wavelet is given by

$$\psi_{\mathrm{Haar}}(x) = \begin{cases} 1, & -\frac{1}{2} \le x < 0 \\ -1, & 0 \le x < \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

For the given Haar wavelet, the corresponding smoothing kernel $\phi(x)$ that satisfes

$$\psi_{\mathrm{Haar}}(x) = \mathrm{L}^*\phi(x)\ , \tag{23}$$

is given by [23], [24]:

$$\begin{aligned} \phi(x) &= -\wedge(2x), \\ \text{where} \quad \wedge(x) &:= \begin{cases} x-1, & -1 \le x < 0 \\ -x+1, & 0 \le x < 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{24}$$

because $\mathrm{L}^* = -\mathrm{L} = -\frac{d}{dt}$ for TV signal model. Note that $\wedge(2x)$ in (24) is the triangular function with the support size of 1 as shown in Fig. 1(b), which is associated with the Haar wavelet basis in Fig. 1(a) thanks to Eq. (23).

One of the most fundamental and novel observations made in this paper is that the basis composed of triangular functions allow us to analyze the following sampled signal $w_s(x)$ rather than directly analysing $u_s(x)$:

$$\begin{aligned} w_s(x) &= u_s(x)\sum_{p=-\infty}^{\infty}\delta\left(x-2^s p\right) \\ &= \sum_{p=-\infty}^{\infty} u_s[p]\delta(x-2^s p) \end{aligned} \tag{25}$$

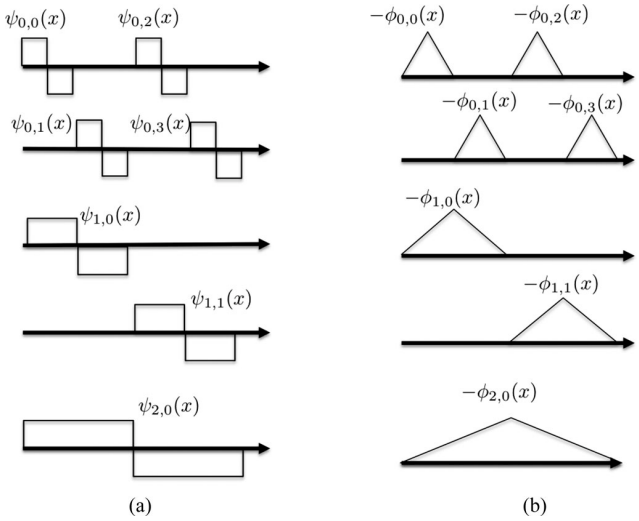

Fig. 1. (a) Haar wavelet basis, and the (b) associated hierarchical triangular basis [35]. Note that absolute scale is ignored for simplicity.

[1]The analysis in the following is significantly influenced by the theory of sparse stochastic processes [35], so we follow the original authors' notation.

where

$$u_s[p] := u_s(t)|_{t=2^s p}\,.$$

In the following, we will discuss the property of $w_s(x)$ for $s=0$ and $s>0$, respectively.

1) When $s=0$: In this case, we have

$$\begin{aligned} w_0(x) &= \sum_{p=-\infty}^{\infty} u_0[p]\delta(x-p) \\ &= \sum_{p=-\infty}^{\infty} (\phi_0 * a)[p]\delta(x-p) \\ &= \sum_{p=-\infty}^{\infty} a[p]\delta(x-p) \end{aligned}$$

where we use the fact that $\phi_0(x)=\phi_0(-x)$ in (24) and its support size is 1. This implies that $w_0(x)$ has the same sparsity level as the underlying innovation $w(x)$, implying that $w_0(x)$ is a stream of r-Diracs. Therefore, by combining the relationship in [20], for the finite supported signal $w(x)$ with the support size n, the resulting Fourier spectrum becomes the discrete Fourier transform (DFT) such that

$$\hat{w}_0[k] = \hat{w}_0(e^{i\omega})\big|_{\omega=2\pi k/n} = \hat{\psi}_0^*[k]\hat{f}[k]$$

where $\hat{\psi}_0^*[k]$ and $\hat{f}[k]$ denote the DFT spectrum of $\psi_0(-x)$ and $f(x)$, respectively. Because $\hat{\psi}_0^*[k]$ can be computed deterministically, we can construct a Hankel matrix $\mathscr{H}(\hat{\mathbf{w}}_0) = \mathscr{H}(\hat{\boldsymbol{\psi}}_0^* \odot \hat{\mathbf{f}})$. Moreover, thanks to the sparsity preservation, we have

$$\text{RANK}\mathscr{H}(\hat{\mathbf{w}}_0) = \text{RANK}\mathscr{H}(\hat{\boldsymbol{\psi}}_0^* \odot \hat{\mathbf{f}}) = r. \tag{26}$$

In addition, due to the periodicity of DFT spectrum, we can use the following wrap-around Hankel matrix:

$$\mathscr{H}_c(\hat{\mathbf{w}}_0) = \left[\begin{array}{cccc} \hat{w}_0[0] & \hat{w}_0[1] & \cdots & \hat{w}_0[d-1] \\ \hat{w}_0[1] & \hat{w}_0[2] & \cdots & \hat{w}_0[d] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_0[n-d] & \hat{w}_0[n-d+1] & \cdots & \hat{w}_0[n-1] \\ \hline \hat{w}_0[n-d+1] & \hat{w}_0[n-d+2] & \cdots & \hat{w}_0[0] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_0[n-1] & \hat{w}_0[0] & \cdots & \hat{w}_0[d-2] \end{array}\right], \tag{27}$$

where the bottom block is an augmented block. As discussed before, this can be easily constructed as the standard Hankel matrix by imposing the periodic boundary conditions. Since the bottom block can be also annihilated using the same annihilating filter, we can see the rank of the wrap-around Hankel expansion is the same as the original Hankel structured matrix:

$$\text{RANK}\mathscr{H}_c(\hat{\mathbf{w}}_0) = \text{RANK}\mathscr{H}(\hat{\mathbf{w}}_0) = r.$$

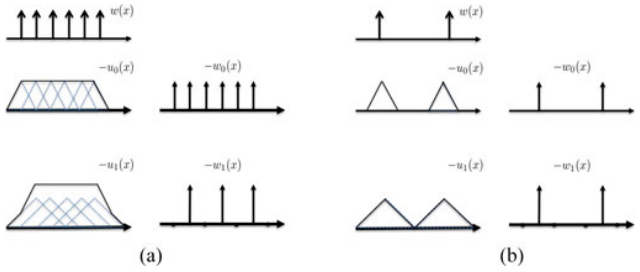

Fig. 2. (a) Best case where the sparsity is reduced by half across scale, and (b) the worst case where the sparsity is persevered across scale.

Therefore, the missing DFT coefficients can be interpolated using the following low-rank matrix completion:

$$\begin{aligned} &\min_{\mathbf{g}\in\mathbb{C}^n} \quad \|\mathscr{H}_c(\mathbf{g})\|_* \\ &\text{subject to} \quad P_\Omega(\mathbf{g}) = P_\Omega(\hat{\mathbf{l}} \odot \hat{\mathbf{f}})\,, \end{aligned} \tag{28}$$

where $\hat{\mathbf{l}} = \hat{\boldsymbol{\psi}}_0^*$.

2) When $s>0$: First, we consider $s=1$. In this case, we have

$$\begin{aligned} w_1(x) &= u_1(x) \sum_{p=-\infty}^{\infty} \delta(x-2p) \\ &= \sum_{p=-\infty}^{\infty} u_1(2p)\delta(x-2p) \\ &= \sum_{p=-\infty}^{\infty} (\phi_1 * a)[2p]\delta(x-2p) \end{aligned}$$

Then, as shown in Figs. 2(a) and 2(b), we can easily see that the sparsity level of the sequence $(\phi_1 * a)[2p]$ does not increase and, furthermore, in many cases, it is reduced nearly by half.

On the other hand, the spectrum of $w_1(x)$ has aliasing components due to the downsampling by factor 2, so we have

$$\hat{w}_1[k] = \frac{1}{2}\left(\hat{\psi}_1^*[k]\hat{f}[k] + \hat{\psi}_1^*[k+n/2]\hat{f}[k+n/2]\right) \tag{29}$$

whose period is reduced by $n/2$. Because of the reduction of the period, in constructing the Hankel matrix (27), we only need to consider the low frequency spectrum up to $n/2$, which can reduce the complexity of the structured matrix completion. However, one key technical challenge is due to the aliasing component in (29) which also aliases into the low frequency part.

To deal with this problem, we propose the following approach. Specifically, let $f_{\text{low}}(x)$ denote the ideal low-pass part of the signal $f(x)$ by zeroing out the DFT spectrum above $n/2$, i.e.

$$f_{\text{low}}(x) = h_{\text{low}}(x) * f(x),$$

where $h_{\text{low}}(x)$ denotes the ideal low-pass filter. Then, for $\mathrm{L} = \frac{d}{dx}$, it is easy to show the following:

$$\mathrm{L}f_{\text{low}}(x) = h_{\text{low}}(x) * \mathrm{L}f(x) = h_{\text{low}}(x) * w(x),$$

because the differentiation and the linear operator L commutes. Therefore, for given $\phi(x)$ in (24), we have

$$\begin{aligned} u_s^{\rm low}(x) &= \langle f_{\rm low}(\cdot), \psi_s(\cdot - x)\rangle \\ &= \langle f_{\rm low}(\cdot), {\rm L}^*\phi_s(\cdot - x)\rangle \\ &= \langle (h_{\rm low} * w)(\cdot), \phi_s(\cdot - x)\rangle \\ &= \sum_{i\in\mathbb{Z}} a[i](h_{\rm low} * \phi_s)(x - i) \end{aligned}$$

and the corresponding sampled signal becomes

$$\begin{aligned} w_1^{\rm low}(x) &= u_1^{\rm low}(x) \sum_{p=-\infty}^{\infty} \delta(x - 2p) \\ &= \sum_{p=-\infty}^{\infty} (\phi_1 * a * h_{\rm low})[2p]\delta(x - 2p) \end{aligned} \tag{30}$$

The role of ideal low pass filter is doubling the resolution, so the non-zero coefficient of $(\phi_1 * a * h_{\rm low})[p]$ increases approximately twice than $(\phi_1 * a)[p]$. However, in (30), the discrete signals are down-sampled by factor of two with zero-padding, so we can assume that the net sparsity level may not increase from using only ideal low-pass signals. On the other hand, one of the important benefits of using this trick is that we no more have aliasing component in its DFT spectrum:

$$\begin{aligned} \hat{w}_1^{\rm low}[k] &= \frac{1}{2}\left(\hat{\psi}_1^*[k]\hat{f}^{\rm low}[k] + \hat{\psi}_1^*[k + n/2]\hat{f}^{\rm low}[k + n/2]\right) \\ &= \frac{1}{2}\hat{\psi}_1^*[k]\hat{f}^{\rm low}[k], \qquad k = 0, \ldots, n/2 - 1, \end{aligned} \tag{31}$$

because the frequency content of $f^{\rm low}$ above $n/2$ is zero. Therefore, the missing DFT coefficients at $s = 1$ can be formulated using the following reduced size low-rank matrix completion problem:

$$\begin{aligned} &\min_{\mathbf{g}\in\mathbb{C}^{n/2}} \ \|\mathscr{H}_c(\mathbf{g})\|_* \\ &\text{subject to} \quad P_\Omega(\mathbf{g}) = P_\Omega(\hat{\mathbf{l}} \odot \hat{\mathbf{f}}^{\rm low}) \,, \end{aligned} \tag{32}$$

where $\hat{\mathbf{l}} = \hat{\boldsymbol{\psi}}_0^*$ and $\hat{\mathbf{f}}^{\rm low}$ denotes the Fourier samples at only below $n/2$. Note the computational complexity reduction due to the reduced Hankel matrix size. For the scale $s > 2$, we can perform similar procedure.

Accordingly, the resulting algorithm is a pyramidal decomposition of a low rank Hankel matrix completion algorithm as shown in Fig. 3 to decouple the aliasing components in wavelet analysis by slightly compromising sparsity level. Here, the low-rank matrix completion is solved from the lowest scale, i.e. $s = 0$, up to the highest scale (see caption of Fig. 3 for more discussion of the figures). Here, a care needs to be taken because the low frequency k-space data corresponding to the scaling function coefficients should be acquired additionally during MR data acquisition. This information as well as the annihilating filter size then determines the the depth of the pyramidal decomposition as will be discussed in detail later in Discussion (see Supplement Material).

There are several advantages of using wavelet approaches compared to the direct operator weighting in ALOHA framework. First, the pyramidal decomposition structure for low rank matrix completion can significantly reduce the overall computational complexity. Moreover, it has been observed in [35] that the wavelet approach is more robust for noise and the model mismatch, which is also consistently observed in ALOHA framework as will be discussed later. Finally, in solving the s-scale of structured matrix completion, rather than starting from zero-initial guess for the missing k-space data, they can be initialised using the values estimated at the lower scale, i.e. $s - 1$. This significantly accelerates the convergence of the algorithm.

Before we finish this section, it is worth to mention that a similar fine-to-coarse scale wavelet coefficients reconstruction was recently proposed in Fourier compressed sensing problem [36].

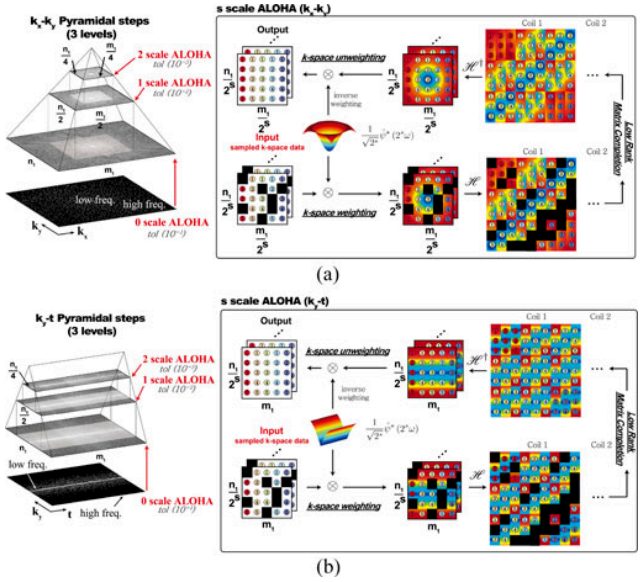

Fig. 3. ALOHA implementation using pyramidal decomposition. Construction of Hankel matrices from (a) k_x-k_y data by assuming that 2-D dydadic wavelet transform of images is sparse, and (b) $k - t$ subsampled data by assuming that dynamic images can be sparsified using spatial wavelet and temporal Fourier transform. In the box, each reconstruction unit at the s-scale (the s-scale ALOHA) is illustrated, which consists of five steps: (1) the k-space region extraction, (2) k-space element by element weighting using s-scale wavelet spectrum $\frac{1}{\sqrt{2^s}}\hat{\psi}^*(2^s\omega)$, (3) low-rank Hankel matrix completion, (4) k-space unweighting by dividing the interpolated k-space data using the s-scale wavelet spectrum, and (5) the k-space data replacement using the interpolated data. Note that for the case of dynamic MRI, one dimensional weigting is required along the phase encoding direction, whereas 2-D weighting is necessary for the case of static imaging. In the figure, $\mathscr{H}^\dagger$ corresponds to the pseudo-inverse operation that takes the average value from the Hankel matrix and putting it back to the original k-space domain. The color coding in the Hankel structure matrix indicates the values of weighting.

B. Generalization to Parallel MRI

Beside the low-rank property originating from sparsity in the transform domain, there exists an additional low-rank relationship that is unique in parallel MRI. The relationship we described here has similarity to SAKE and P-LORAKS when the image itself is sparse, and our contribution is its generalization to transform domain sparse signals and theoretical verification.

Here, to allow seamless integration with wavelet analysis, the following theory is derived based on the cardinal spline models and wavelet analysis in the previous section.

In pMRI, the unknown image $g_i(x)$ from the i-th coil can be represented as

$$g_i(x) = s_i(x)f(x), \quad i = 1, \ldots, N_c, \tag{33}$$

where $s_i(x)$ denotes the i-th coil sensitivity map, $f(x)$ is an unknown image, and N_c denotes the number of coils. To make the algorithm general, we impose the most basic assumption: the coil sensitivity map is bounded, i.e. $|s_i(x)| < \infty$, $\forall i$. The main goal of parallel imaging is, then, to exploit the common signal $f(x)$ that is measured through multiple channels. There are two distinct scenarios: 1) $g_i(x)$ is sparse by itself, and 2) $g_i(x)$ can be sparsified by applying transform such as total variations. In the following we investigate the two scenario separately.

1) Sparse Signals: Under the cardinal setup, to model the image domain sparse signal on a fixed grid, we again assume that $g_i(x)$ is a stream of Diracs given by

$$g_i(x) = \sum_{l \in I_i} g_i[l]\delta(x - l), \quad i = 1, \ldots, N_c,$$

where I_i is a non-zero index set on the integer grid. Now, define the union of index set

$$I_f = \bigcup_{i=1}^{N_c} I_i, \quad r = |I_f|,$$

where $|\cdot|$ denotes the cardinality of a set. Because the coil sensitivity map is bounded, Eq. (33) implies that $f(x)$ is a stream of r-Diracs whose non-zero coefficients exist only on the index set I_f:

$$f(x) = \sum_{l \in I_f} f[l]\delta(x - l),$$

and, accordingly, we have

$$g_i(x) = \sum_{l \in I_f} s_i[l]f[l]\delta(x - l), \quad \text{where} \quad s_i[l] = s_i(x)|_{x=l}\,.$$

Then, under the usual assumption of the finite supported signal with the support size n, the resulting spectrum becomes DFT spectrum represented by

$$\hat{g}_i[k] = \hat{f}[l] \circledast \hat{s}_i[l] = \sum_{l=0}^{n-1} \hat{f}[l]\hat{s}[(k-l)_n], \tag{34}$$

where $(\cdot)_n$ denotes the mod operation, because the multiplication in the discrete-time domain becomes periodic convolution in the DFT domain. Therefore, a Hankel matrix $\mathscr{H}_c(\hat{\mathbf{g}}_i)$ with the matrix pencil size d and the periodic boundary condition can have the following decomposition:

$$\mathscr{H}_c(\hat{\mathbf{g}}_i) = \mathscr{H}_c(\hat{\mathbf{f}})\mathcal{C}(\hat{\mathbf{s}}_i) \quad \in \mathbb{C}^{n\times d} \tag{35}$$

where $\mathscr{H}_c(\hat{\mathbf{f}})$ is $n \times n$ wrap around Hankel matrix and $\mathcal{C}(\hat{\mathbf{s}}_i)$ are given by

$$\mathscr{H}_c(\hat{\mathbf{w}}) = \begin{bmatrix} \hat{f}[0] & \hat{f}[1] & \cdots & \hat{f}[n-1] \\ \hat{f}[1] & \hat{f}[2] & \cdots & \hat{f}[0] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{f}[n-1] & \hat{f}[0] & \cdots & \hat{f}[n-2] \end{bmatrix} \in \mathbb{C}^{n\times n}\,.$$

and

$$\mathcal{C}(\hat{\mathbf{s}}_i) = \begin{bmatrix} \hat{s}_i[0] & \hat{s}_i[1] & \cdots & \hat{s}_i[d-1] \\ \hat{s}_i[n-1] & \hat{s}_i[0] & \cdots & \hat{s}_i[d-2] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{s}_i[1] & \hat{s}_i[2] & \cdots & \hat{s}_i[d] \end{bmatrix} \in \mathbb{C}^{n\times d}\,.$$

Then, for a given horizontally augmented matrix $\mathcal{Y}_h$:

$$\mathcal{Y}_h = \begin{bmatrix} \mathscr{H}_c(\hat{\mathbf{g}}_1) & \ldots & \mathscr{H}_c(\hat{\mathbf{g}}_{N_c}) \end{bmatrix} \in \mathbb{C}^{n\times(N_c d)}\,. \tag{36}$$

with $\mathscr{H}(\hat{\mathbf{g}}_i) \in \mathbb{C}^{n\times d}$, we have the following decomposition:

$$\mathcal{Y}_h = \mathscr{H}_c(\hat{\mathbf{f}}) \begin{bmatrix} \mathcal{C}(\hat{\mathbf{s}}_1) & \ldots & \mathcal{C}(\hat{\mathbf{s}}_{N_c}) \end{bmatrix}. \tag{37}$$

Due to the rank inequality $\text{RANK}(AB) \leq \min\{\text{RANK}(A), \text{RANK}(B)\}$, we therefore have the following rank bound:

$$\text{RANK}\,\mathcal{Y}_h \leq \text{RANK}\,\mathscr{H}_c(\hat{\mathbf{f}}) = |I_f| = r \quad , \tag{38}$$

where the equality comes from Theorem 2.1.[2] Therefore, if the matrix pencil size d is chosen such that $N_c d > r$, then the concatenated matrix becomes low-ranked.

Note that the decomposition structure in (37) implies the following inter-coil annihilating filter relationship:

$$\hat{g}_i[k] \circledast \hat{s}_j[k] - \hat{g}_j[k] \circledast \hat{s}_i[k] = 0, \quad i \neq j,$$

which are quite often used in multi-channel deconvolution problems [37]. This is because $\binom{N_c}{2}$ combination of the vertically augmented sensitivity filter vectors with the appropriate $+1, -1$ weighting live in the null space of $[\mathcal{C}(\hat{\mathbf{s}}_1) \ \ldots \ \mathcal{C}(\hat{\mathbf{s}}_{N_c})]$ (accordingly, in the null space of $\mathcal{Y}_h$) [37].

2) TV Signals: Similar to the previous section, we also assume that $g_i(x)$ is a cardinal L-spline model (17) where $\mathrm{L} = \frac{d}{dx}$. Because $s_i(x)$ is bounded and the following chain rule holds:

$$\mathrm{L}g_i(x) = s_i(x)\mathrm{L}f(x) + f(x)\mathrm{L}s_i(x)\,, \quad i = 1, \ldots, N_c,$$

the cardinal L-spline model assumption implies that both $\mathrm{L}f(x)$ and $\mathrm{L}s_i(x)$ are also cardinal L-splines:

$$\mathrm{L}f(x) = w(x) := \sum_{l \in I_a} a[l]\delta(x - l) \tag{39}$$

$$\mathrm{L}s_i(x) = b_i(x) := \sum_{l \in I_{b_i}} b_i[l]\delta(x - l) \tag{40}$$

[2]In this special instance, we do not even need Theorem 2.1. This is because $n \times n$-Hankel matrix $\mathscr{H}_c(\hat{\mathbf{f}})$ becomes circulant, so the rank of the circulant matrix is equal to the non-zero elements after taking DFT.

where I_a and I_{b_i} denote the index sets where $a[l]$ and $b_i[l]$ are non-zero, respectively. We further define the union set

$$I_b = \bigcup_{i=1}^{N_c} I_{b_i}. \tag{41}$$

Then, the wavelet analysis at the 0-th scale using centered Haar wavelet gives us

$$\begin{aligned} u_0^i(x) &= \langle g_i(\cdot), \psi_0(\cdot - x)\rangle \\ &= \langle g_i(\cdot), \mathrm{L}^*\phi_0(\cdot - x)\rangle \\ &= \langle \mathrm{L}g_i(\cdot), \phi_0(\cdot - x)\rangle \\ &= \langle s_i(\cdot)\mathrm{L}f(\cdot), \phi_0(\cdot - x)\rangle + \langle f(\cdot)\mathrm{L}s_i(\cdot), \phi_0(\cdot - x)\rangle \\ &= \sum_{l\in I_a} a[l]s_i(l)\phi_0(x-l) + \sum_{l\in I_{b_i}} b_i[l]f(l)\phi_0(x-l) \end{aligned} \tag{42}$$

where we use $\bar{\phi}_0(x) = \phi_0(x)$ for the case of centred triangular basis. Recall that one of the main advantages of cardinal spline model is that the sparsity level is still preserved after discretization, so we can analyze the following signal:

$$w_0^i(x) = u_0^i(x)\sum_{l\in\mathbb{Z}}\delta(x-l) \tag{43}$$

$$= \sum_{l\in\mathcal{F}} a[l]s_i[i]\delta(x-l) + \sum_{l\in\mathcal{S}_i} b_i[l]f[i]\delta(x-l) \tag{44}$$

where $s_i[i] := s_i(x)|_{x=i}$ and $f[i] := f(x)|_{x=i}$. We further define a truncated sequences $f_{\mathrm{tr}}[l]$:

$$f_{\mathrm{tr}}[l] = \begin{cases} f[l], & l \in I_b \\ 0, & \text{otherwise} \end{cases}, \tag{45}$$

where I_b is defined in (41). Then, for the finite supported signal with the support size n, the resulting spectrum becomes DFT spectrum such that

$$\begin{aligned} \hat{w}_0^i[k] &= \hat{\psi}_0^*[k]\hat{g}_i[k] \\ &= \mathrm{DFT}\left\{a[l]s_i[l] + b_i[l]f_{\mathrm{tr}}[l]\right\} \qquad (46) \\ &= \hat{w}[k] \circledast \hat{s}_i[k] + \hat{f}_{\mathrm{tr}}[k] \circledast \hat{b}_i[k] \qquad (47) \end{aligned}$$

where $\hat{w}[k]$ denotes the DFT spectrum of (39), and $\hat{f}_{\mathrm{tr}}[k]$ is the DFT spectrum of the sequence in (45). Then, the $n \times d$ Hankel matrix with the periodic boundary condition constructed from $\hat{w}_0^i[k]$ can be decomposed as

$$\mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_i) = \mathscr{H}_c(\hat{\mathbf{w}})\mathcal{C}(\hat{\mathbf{s}}_i) + \mathscr{H}_c(\hat{\mathbf{f}}_{\mathrm{tr}})\mathcal{C}(\hat{\mathbf{b}}_i) \tag{48}$$

where $\hat{\mathbf{l}} = \hat{\boldsymbol{\psi}}_0^*$ and $\mathscr{H}_c(\hat{\mathbf{w}}) \in \mathbb{C}^{n\times n}$ and $\mathcal{C}(\hat{\mathbf{s}}_i) \in \mathbb{C}^{n\times d}$ are given by

$$\mathscr{H}_c(\hat{\mathbf{w}}) = \begin{bmatrix} \hat{w}[0] & \hat{w}[1] & \cdots & \hat{w}[n-1] \\ \hat{w}[1] & \hat{w}[2] & \cdots & \hat{w}[0] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}[n-1] & \hat{w}[0] & \cdots & \hat{w}[n-2] \end{bmatrix}.$$

and

$$\mathcal{C}(\hat{\mathbf{s}}_i) = \begin{bmatrix} \hat{s}_i[0] & \hat{s}_i[1] & \cdots & \hat{s}_i[d-1] \\ \hat{s}_i[n-1] & \hat{s}_i[0] & \cdots & \hat{s}_i[d-2] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{s}_i[1] & \hat{s}_i[2] & \cdots & \hat{s}_i[d] \end{bmatrix} \in \mathbb{C}^{n\times d}.$$

Similarly, we can define $\mathscr{H}_c(\hat{\mathbf{f}}_{\mathrm{tr}}) \in \mathbb{C}^{n\times n}$ and $\mathcal{C}(\hat{\mathbf{b}}_i) \in \mathbb{C}^{n\times d}$. Then, a horizontally augmented matrix $\mathcal{Y}_h$ can be decomposed as following:

$$\begin{aligned} \mathcal{Y}_h &= \left[\mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_1) \cdots \mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_{N_c})\right] \qquad (49) \\ &= \mathscr{H}_c(\hat{\mathbf{w}})\left[\mathcal{C}(\hat{\mathbf{s}}_1) \cdots \mathcal{C}(\hat{\mathbf{s}}_{N_c})\right] \\ &\quad + \mathscr{H}_c(\hat{\mathbf{f}}_{\mathrm{tr}})\left[\mathcal{C}(\hat{\mathbf{b}}_1) \cdots \mathcal{C}(\hat{\mathbf{b}}_{N_c})\right] \qquad (50) \end{aligned}$$

where $\hat{\mathbf{l}} = \hat{\boldsymbol{\psi}}_0^*$. Due to the rank inequality $\mathrm{RANK}(AB) \le \min\{\mathrm{RANK}(A), \mathrm{RANK}(B)\}$ as well as $\mathrm{RANK}(A+B) \le \mathrm{RANK}(A) + \mathrm{RANK}(B)$, the rank of the concatenated matrix is, therefore, given by

$$\begin{aligned} \mathrm{RANK}\,\mathcal{Y}_h &\le \mathrm{RANK}\,\mathscr{H}_c(\hat{\mathbf{w}}) + \mathrm{RANK}\,\mathscr{H}_c(\hat{\mathbf{f}}_{\mathrm{tr}}) \\ &= |I_a| + |I_b| \end{aligned} \tag{51}$$

where $|I_a|$ and $|I_B|$ denote the cardinality of the sets I_a and I_b defined in (39) and (41), respectively. Therefore, if $\mathrm{L}f(x)$ and $\mathrm{L}s_i(x), i = 1, \ldots, N_c$ are sufficiently sparse such that $|I_a| + |I_b| < N_c d$, then the resulting concatenated Hankel matrix becomes low-ranked. The analysis can be also extended for the scale $s \ge 1$ by combining the discussion in the previous section.

3) Low-Rank Matrix Completion: Due to the aforementioned low-rankness of the concatenated matrix, the multichannel version of the ALOHA can be formulated as

$$\begin{aligned} &\min_{\{\mathbf{m}_i\}_{i=1}^r} \ \mathrm{RANK}\left[\mathscr{H}_c(\mathbf{m}_1) \cdots \mathscr{H}_c(\mathbf{m}_r)\right] \qquad (52) \\ &\text{subject to} \quad P_\Omega(\mathbf{m}_i) = P_\Omega(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_i), \quad i = 1, \ldots, N_c. \end{aligned}$$

Before we finish this section, we like to investigate the feasibility of the different concatenation order. More specifically, one could construct a vertically augmented matrix $\mathcal{Y}_v$:

$$\mathcal{Y}_v = \begin{bmatrix} \mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_1) \\ \vdots \\ \mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_{N_c}) \end{bmatrix} \in \mathbb{C}^{N_c n\times d}. \tag{53}$$

Due to the wrap-around structure of the individual Hankel matrix, note that $\mathcal{Y}_v$ is not a transpose of $\mathcal{Y}_h$. Accordingly, the analysis in the previous section does not hold and we cannot utilize the low-rankness from the inter-coil relationship.

However, if we stack Hankel matrix without wrap-around property (i.e. $\mathscr{H}(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_i) \in \mathbb{C}^{(n-d+1)\times d}$ are used in (53) instead of $\mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_i)$), then due to the special structure of the Hankel matrix, $\mathcal{Y}_v$ becomes the transpose of $\mathcal{Y}_h$; accordingly, by interchanging the role of $n-d+1$ and d, one can make the rank of $\mathcal{Y}_v$ equal to that of $\mathcal{Y}_h$. However, the theoretical analysis of the concatenated Hankel matrices without wrap around property turns out to be quite involved due to the boundary condition

and data truncation; and even more, it is even not necessary in accelerated MR except for the super-resolution imaging (see [22]), because the image should be recovered on a fixed grid rather than on a continuum which makes the cardinal spline model more appropriate. Under this condition, the performance of the resulting rank minimization for the vertical stacking approach becomes deteriorated as will be shown in Discussion (see Supplement Material).

IV. Implementation Details

A. 2D Hankel Structured Matrix Construction

In 3D imaging or dynamic acquisition of MR data, the readout direction is usually fully sampled and the other two encoding directions are under sampled, so the problem becomes a 2-D imaging problem. Thus, this section presents an explicit way of constructing a 2D Hankel structured matrix. Specifically, if $\hat{h}[n,m]$ is a $p_1 \times q_1$ size 2D annihilating filter, then the corresponding annihilating filter relation is given by

$$(\hat{h} * \hat{f})[n,m] = \sum_{i=0}^{p_1-1} \sum_{j=0}^{q_1-1} \hat{h}[i,j]\hat{f}[n-i, m-j] = 0, \quad (54)$$

for all $n, m \in \Omega$. Let $n_1 \times m_1$ k-space data matrix be defined by

$$\hat{F} := \begin{bmatrix} \hat{f}[0,0] & \cdots & \hat{f}[0, m_1-1] \\ \vdots & \ddots & \vdots \\ \hat{f}[n_1-1, 0] & \cdots & \hat{f}[n_1-1, m_1-1] \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_0 & \cdots & \hat{\mathbf{f}}_{m_1-1} \end{bmatrix}.$$

Similarly, we define $p_1 \times q_1$ annihilating filter matrix $\hat{H}$. Then, (54) can be equivalently represented as

$$\mathscr{H}(\hat{F})\hat{\mathfrak{h}} = \mathbf{0}, \quad (55)$$

where a 2-D Hankel structured matrix $\mathscr{H}(\hat{F})$ is constructed as

$$\begin{bmatrix} \mathscr{H}(\hat{\mathbf{f}}_0) & \mathscr{H}(\hat{\mathbf{f}}_1) & \cdots & \mathscr{H}(\hat{\mathbf{f}}_{q_1-1}) \\ \mathscr{H}(\hat{\mathbf{f}}_1) & \mathscr{H}(\hat{\mathbf{f}}_2) & \cdots & \mathscr{H}(\hat{\mathbf{f}}_{q_1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathscr{H}(\hat{\mathbf{f}}_{m_1-q_1}) & \mathscr{H}(\hat{\mathbf{f}}_{m_1-q_1+1}) & \cdots & \mathscr{H}(\hat{\mathbf{f}}_{m_1-1}) \end{bmatrix} \quad (56)$$

with $\mathscr{H}(\hat{\mathbf{f}}_j) \in \mathbb{C}^{(n_1-p_1+1)\times p_1}$ given by

$$\begin{bmatrix} \hat{f}[0,j] & \hat{f}[1,j] & \cdots & \hat{f}[p_1-1,j] \\ \hat{f}[1,j] & \hat{f}[2,j] & \cdots & \hat{f}[p_1,j] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{f}[n_1-p_1,j] & \hat{f}[n_1-p_1+1,j] & \cdots & \hat{f}[n_1-1,j] \end{bmatrix},$$

and the annihilating filter vector is given by

$$\hat{\mathfrak{h}} = \overline{\text{VEC}(\hat{H})}, \quad (57)$$

where the overline denotes an operator that reserves the order of a vector. Similar to 1-D cases, we again impose a periodic boundary condition to construct a wrap-around Hankel matrix.

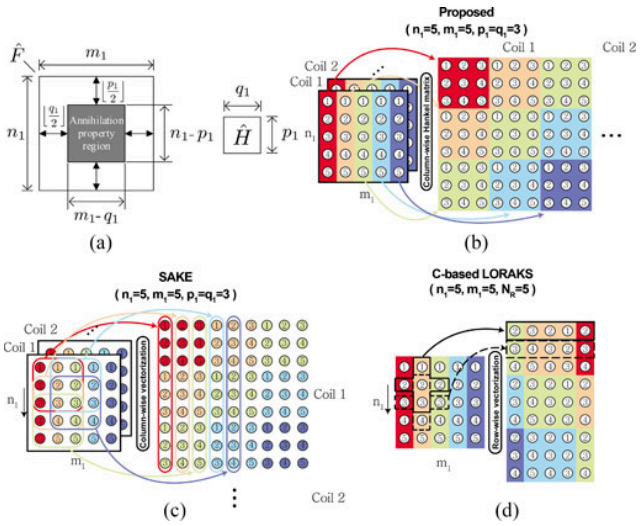

Fig. 4. (a) Area where annihilation property holds. Various ways of constructing block Hankel matrices: (b) ALOHA, (c) SAKE, and (d) LORAKS. In (d), N_R denotes the number of neighborhood pixels.

Using this, we can construct an augmented matrix $\mathcal{Y}$ in (49) from N_c-channels.

The augmented matrix structure $\mathscr{H}(\hat{F})$ illustrated in Fig. 4(b) is similar to those of SAKE and C-based LORAKS/P-LORAKS in Fig. 4(c) and (d), respectively, with the following differences. Compared to SAKE, ALOHA stacks the multi-coil Hankel matrices side by side. Second, unlike the SAKE and ALOHA, C-based LORAKS uses an adjustable parameter R as the radius of non-separable neighborhoods according to the software manual provided by the original authors. Moreover, the most important novelty of the proposed method is the k-space weighting to construct a *weighted* Hankel structured matrix.

B. Hankel Structured Matrix Completion Algorithm

In order to solve Eqs. (12) and (52), we employ an SVD-free structured rank minimization algorithm [38] with an initialization using the low-rank factorization model (LMaFit) algorithm [39]. This algorithm does not use the singular value decomposition (SVD), so the computational complexity can be significantly reduced. Specifically, the algorithm is based on the following observation [40]:

$$\|A\|_* = \min_{U,V: A = UV^H} \|U\|_F^2 + \|V\|_F^2 . \quad (58)$$

Hence, (12) can be reformulated as the nuclear norm minimization problem under the matrix factorization constraint:

$$\min_{U,V: \mathscr{H}(\mathbf{m}) = UV^H} \|U\|_F^2 + \|V\|_F^2 \quad \text{subject to} \quad P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{f}}). \quad (59)$$

By combining the two constraints, we have the following cost function for an alternating direction method of multiplier

(ADMM) step [41]:

$$L(U,V,\mathbf{m},\Lambda) := \iota(\mathbf{m}) + \frac{1}{2}\left(\|U\|_F^2 + \|V\|_F^2\right) + \frac{\mu}{2}\|\mathscr{H}(\mathbf{m}) - UV^H + \Lambda\|_F^2 \quad (60)$$

where $\iota(\mathbf{m})$ denotes an indicator function:

$$\iota(\mathbf{m}) = \begin{cases} 0, & \text{if } P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{f}}) \\ \infty, & \text{otherwise} \end{cases}.$$

One of the advantages of the ADMM formulation is that each subproblem is simply obtained from (60). More specifically, $\mathbf{m}^{(n+1)}, U^{(n+1)}$ and $V^{(n+1)}$ can be obtained, respectively, by applying the following optimization problems sequentially:

$$\begin{aligned} &\min_{\mathbf{m}} \iota(\mathbf{m}) + \frac{\mu}{2}\left\| \mathscr{H}(\mathbf{m}) - U^{(n)}V^{(n)H} + \Lambda^{(n)}\right\|_F^2 \\ &\min_{U} \frac{1}{2}\|U\|_F^2 + \frac{\mu}{2}\left\| \mathscr{H}(\mathbf{m}^{(n+1)}) - UV^{(n)H} + \Lambda^{(n)}\right\|_F^2 \\ &\min_{V} \frac{1}{2}\|V\|_F^2 + \frac{\mu}{2}\left\| \mathscr{H}(\mathbf{m}^{(n+1)}) - U^{(n+1)}V^{H} + \Lambda^{(n)}\right\|_F^2 \end{aligned} \quad (61)$$

and the Lagrangian update is given by

$$\Lambda^{(n+1)} = \mathcal{Y}^{(n+1)} - U^{(n+1)}V^{(n+1)H} + \Lambda^{(n)}.$$

It is easy to show that the first step in (61) can be reduced to

$$\mathbf{m}^{(n+1)} = P_{\Omega^c}\mathscr{H}^\dagger\left\{U^{(n)}V^{(n)H} - \Lambda^{(n)}\right\} + P_\Omega(\hat{\mathbf{f}}), \quad (62)$$

where P_{Ω^c} is a projection mapping on the set Ω^c and $\mathscr{H}^\dagger$ corresponds to the Penrose-Moore pseudo-inverse mapping from our block Hankel structure to a vector. Hence, the role of the pseudo-inverse is taking the average value and putting it back to the original coordinate. Next, the subproblem for U and V can be easily calculated by taking the derivative with respect to each matrix, and we have

$$\begin{aligned} U^{(n+1)} &= \mu\left(\mathcal{Y}^{(n+1)} + \Lambda^{(n)}\right)V^{(n)}\left(I + \mu V^{(n)H}V^{(n)}\right)^{-1} \\ V^{(n+1)} &= \mu\left(\mathcal{Y}^{(n+1)} + \Lambda^{(n)}\right)^H U^{(n+1)} \\ &\quad \left(I + \mu U^{(n+1)H}U^{(n+1)}\right)^{-1} \end{aligned} \quad (63)$$

Note that the computational complexity of our ADMM algorithm is dependent on the matrix inversion in (63), whose complexity is determined by the estimated rank of the Hankel matrix. Therefore, even though the Hankel matrix has large size, the estimated rank is much smaller, which significantly reduces overall complexity.

Now, for faster convergence, the remaining issue is how to initialize U and V. For this, we employ an algorithm called the low-rank factorization model (LMaFit) [39]. More specifically, for a low-rank matrix Z, LMaFit solves the following optimization problem:

$$\min_{U,V,Z} \frac{1}{2}\|UV^H - Z\|_F^2 \text{ subject to } P_I(Z) = P_I(\mathscr{H}(\hat{\mathbf{f}})) \quad (64)$$

and Z is initialized with $\mathscr{H}(\hat{\mathbf{f}})$ and the index set I denotes the positions where the elements of $\mathscr{H}(\hat{\mathbf{f}})$ are known. LMaFit solves a linear equation with respect to U and V to find their updates and relaxes the updates by taking the average between the previous iteration and the current iteration. Moreover, the rank estimation can be done automatically. LMaFit uses QR factorization instead of SVD, so it is also computationally efficient. Even though the problem (64) is non-convex due to the multiplication of U and V, the convergence of LMaFit to a stationary point was analyzed in detail [39]. However, the LMaFit alone cannot recover the block Hankel structure, which is the reason we use an ADMM step afterward to impose the structure.

Note that the automatic rank estimation is another important advantages over SAKE and LORAKS that require time consuming full search or manual tuning for the rank estimation.

C. Reconstruction Flow

As shown in Fig. 3, the ALOHA framework in the present work is comprised with several major steps: pyramidal decomposition, k-space weighting, Hankel matrix formation, rank estimation, SVD-free low rank matrix completion, and k-space unweighting. Here, we will explain these in more detail.

The pyramidal decomposition is performed as follows. First, in static MR data acquisition illustrated in Fig. 3 a, the $k_x - k_y$ corresponds to the two phase encoding directions that are downsampled. Thus, the Hankel matrix is constructed from $k_x - k_y$ data. After a k-space interpolation from a finer scale, the data at the current scale is defined to contain one-fourth of data around zero frequency from that of the previous scale. Second, in the case of dynamic MR imaging shown in Fig. 3(b), k_x samples from the readout direction are fully acquired, whereas the k_y directional phase encoding are downsampled along the temporal direction t. Therefore, the data in $k_y - t$ space (or simply, $k - t$ space) are downsampled, from which we construct a Hankel structure matrix. In pyramidal decomposition, after a k-space interpolation from a finer scale, the $k_y - t$ data in the current scale contains a half of the data from that of the previous scale. Note that the wavelet decomposition is performed only along the spatial domain, so the pyramidal decomposition is only performed along k_y direction. This construction of Hankel matrix is due to the observation that the dynamic signal is sparse in spatial wavelet and temporal Fourier transform domain [32]. See more details in our recent work [32].

In both cases, the estimated k-space data at the lower scale are used to initialize the low rank matrix completion algorithm at the current scale. This accelerates the convergence speed. Moreover, due to the additional chance of refining the estimates, more important k-space samples at the low frequency regions are refined furthermore compared to the high frequency k-space samples. Consequently, the overall computational burden of the low rank matrix competition algorithm is significantly reduced while the overall quality is still maintained.

The k-space weighting is performed using wavelets. Specifically, we use a Haar wavelet expansion whose spectrum is

given by

$$\hat{\psi}_0(\omega) = \frac{i\omega}{2}\left(\frac{\sin\omega/4}{\omega/4}\right)^2. \tag{65}$$

The corresponding k-space weighting at the s-scale is given by

$$\hat{\psi}_s^*(\omega) = \frac{-1}{\sqrt{2^s}}\frac{i2^s\omega}{2}\left(\frac{\sin 2^s\omega/4}{2^s\omega/4}\right)^2. \tag{66}$$

For the case of static MRI in Fig. 3(a), we use 2-D weighting by assuming that the image is sparse in 2-D dyadic wavelet transform domain. Care needs to be taken when applying the weighting to 2D Fourier domain because there are two frequency variables (ω_x, ω_y). One could use a separable weighting $\hat{l}(\omega_x, \omega_y) = \hat{l}(\omega_x)\hat{l}(\omega_y)$; however, the resulting problem is that the missing k-space components along the frequency axis $\omega_x = 0$ or $\omega_y = 0$ cannot be recovered. Consequently, we applied the weighting sequentially along each axis, i.e. we solve (12) by applying $\hat{l}(\omega_x)$ first, which is followed by solving (12) with $\hat{l}(\omega_y)$. However, simultaneous weighting would be possible as demonstrated in a recent work for off-the-grid recovery of piecewise constant image [22]. For the case of dynamic imaging in Fig. 3(b), one dimensional weighting along the phase encoding direction was applied as explained in detail in [32]. Finally, after the k-space interpolation, the k-space unweighting is done in k-space pixel-by-pixel by dividing the reconstructed value with (66). Note that (66) has zero value at the DC frequency. However, because we acquire the DC value as well as some of the low frequency k-space data, the problem of dividing by zero never happened.

We used TITAN GTX graphic card for graphic processor unit (GPU) and i7-4770k CPU and the codes were written in MATLAB 2015a (Mathwork, Natick). To accelerate the algorithm, most part of the MATLAB codes except the LMaFit were implemented using Compute Unified Device Architecture (CUDA) for GPU. LMaFit step was implemented using the original authors' CPU version code.

D. MR Acquisition and Reconstruction Parameters

To assess the performance of ALOHA for single coil compressed sensing imaging, k-space raw data from an MR headscan was obtained with Siemens Tim Trio 3T scanner using balanced steady-state free precession (bSSFP) sequence. The acquisition parameters were as follows: TR/TE $= 10.68/5.34$ ms, 208×256 acquisition matrix (partial Fourier factor 7/8, oversampling factor 50%), and number of slices is 104 with 2 mm slice thickness. The field-of-view (FOV) was $178 \times 220\,\text{mm}^2$. We used the central coronal slice.

A retrospective down-sampling mask was generated according to a two dimensional Gaussian distribution and the data at the central 7×7 region around zero frequency were obtained additionally. This is equivalent to assume a 3D imaging scenario where the readout direction is fully sampled and the downsampling is done in the remaining 2-D phase encoding direction. Downsampling factors of four was used to generate sampling masks. However, data was obtained as partial Fourier measurements, so effective downsampling ratio was 4.13. The 2-D k-space weighting using (66) was used. The ALOHA reconstruction was conducted using the following parameters: three levels of pyramidal decomposition, and decreasing LMaFit tolerance values $(5 \times 10^{-2}, 5 \times 10^{-3}, 5 \times 10^{-4})$ at each level of the pyramid. In addition, an initial rank estimate for LMaFit started with one and was refined automatically in an increasing sequence, the annihilating filter size was 23×23, and the ADMM parameter was $\mu = 10^3$.

For the compressed sensing approach, we used two approaches with the same data and the same sampling masks: (1) the sparsity in wavelet domain (which we denote l_1-wavelet) [3], and (2) the split Bregman method for the total variation [42]. In case of l_1-wavelet approach, we implemented an ADMM algorithm using wavelet domain sparsity. In addition, for comparison with the existing state-of-the art approach using Hankel structured matrix completion algorithm, C-based LORAKS was compared because it exploits the image domain sparsity. There are other types of LORAKS such as G-based LORAKS and S-based LORAKS, which may improve the quality of image better than C-based LORAKS depending on situtation. However, because C based-LORAKS (from now on, referred as 'C-LORAKS') is similar to the ALOHA without weighting, we use C-LORAKS as a reference to contrast why the proposed ALOHA framework has many advantages. The implementation of C-LORAKS was based on the source code available in the original author's homepage, which requires manual setting of estimated ranks. We chose the rank for LORAKS that gave the best reconstruction quality. In Discussion (see Supplementary Material), we also provided reconstruction results by ALOHA without weighting to control the confounds and confirm the importance of the k-space weighting in constructing Hankel matrix.

The parameters for the l_1-wavelet and TV approaches were optimized to have the best performance in terms of the normalized mean square error (NMSE), which is defined by $\text{NMSE}(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\|_2^2/\|\mathbf{y}\|_2^2$, where $\mathbf{x}$ and $\mathbf{y}$ denote the reconstructed and the ground-truth images, respectively. C-LORAKS parameters were also chosen manually to give the best reconstruction quality.

To evaluate the performance of ALOHA in static parallel imaging, k-space raw data from an MR headscan was obtained with Tim Trio 3T scanner using 2D GRE sequence. The acquisition parameters were as follows: TR/TE $= 8.6/4$ ms, 231×512 acquisition matrix, and six z-slices with 7mm slice thickness. The field-of-view (FOV) was $250 \times 250\,\text{mm}^2$, and the number of coils was four. Retrospectively undersampled 2-D k-space data at the acceleration factor of four were obtained according to two dimensional Gaussian distribution in addition to the 7×7 central region around zero frequency. The data from 4 receiver coils were used. For comparison, we used the identical data and sampling masks for SAKE with ESPIRiT [16], [43]. Note that GRAPPA [13] requires ACS lines, so with the additional 50 samples along ACS, the effective downsampling ratio was 3.1, which is not good as the four time acceleration in other algorithms. SAKE with ESPIRiT are combined algorithm with both low rank matrix completion algorithms for Hankel structured matrix collected from the ACS

k-space data and SENSE algorithm for filling high-frequency k-space [43]. SAKE with ESPIRiT was recently proposed to reduce the computational burden of the original SAKE without noticeable reconstruction quality loss by performing low rank matrix completion only for the 65×65 central region, and after that coil sensitivities are estimated using the reconstruction data. The estimated coil sensitivities are used to estimate the remaining k-space missing data through ESPIRiT [43]. Accordingly, the image quality from SAKE+ESPIRiT was quite similar to that of SAKE with a slight NMSE loss, but the computational time of SAKE was significantly higher than SAKE+ESPIRiT. Therefore, reconstruction reconstruction SAKE with ESPIRiT were only illustrated. The parameters for SAKE with ESPIRiT were chosen such that they provided the best reconstruction results. The parameters for the ALOHA are as follows: three levels of pyramidal decomposition with decreasing LMaFit tolerances $(10^{-2}, 10^{-3}, 10^{-4})$, and 5×5 annihilating filter. The same LMaFit rank estimation strategy and ADMM parameter used for single coil experiments were employed. We generated the square root of sum of squares (SSoS) image from multi-coil reconstructions.

We also validated the performance of ALOHA for accelerated dynamic cardiac data in the $k-t$ domain. A cardiac cine data set was acquired using a 3T whole-body MRI scanner (Siemens; Tim Trio) equipped with a 32-element cardiac coil array. The acquisition sequence was bSSFP and prospective cardiac gating was used. The imaging parameters were as follows: FOV= $300 \times 300\,\text{mm}^2$, acquisition matrix size= 128×128, TE/TR $= 1.37/2.7$ ms, receiver bandwidth $= 1184$ Hz/pixel, and flip angle $= 40^\circ$. The number of cardiac phases was 23 and the temporal resolution was 43.2ms. The k-t space samples including four lines around zero frequency were retrospectively obtained at the reduction factor of eight according to a Gaussian distribution. For comparison, k-t FOCUSS [7], C-LORAKS [17], and SAKE [16] was used. In case of C-LORAKS (single coil) and SAKE (multi coil), these algorithms were applied to k-t domain for dynamic reconstruction. The parameters in k-t FOCUSS, C-LORAKS, and SAKE, were selected to give the best NMSE values. For the ALOHA in single coil data, the following parameters were used: three level of pyramidal decomposition only along the phase encoding direction, decreasing LMaFit tolerances $(10^{-1}, 10^{-2}, 10^{-3})$ at each scale, and 13×7 annihilating filter. The same LMaFit rank estimation strategy and ADMM parameter was $\mu = 10$. The k-space weighting in Eq. (66) was applied only along the phase encoding direction.

Next, we investigated the synergetic improvement of dynamic imaging from multi-channel acquisition. Four representative coils out of 32 were used. The reason we chose only four coils was to verify that the generalised ALOHA can maximally exploit the multi-channel diversity even with the small number of coils. The four coils were chosen such that it covers every area of images. The same four coils were used for all algorithms for fair comparison. In the ALOHA, the annihilating filter size was 5×5. The same LMaFit rank estimation strategy and ADMM parameter used before were employed. After the reconstructions of k-space samples, the inverse Fourier transform was applied, and the SSoS images were obtained by combining the reconstructed images.

V. Results

A. Static MR Experiments

Reconstructed results from a single coil brain data are shown in Fig. 5 with the NMSE values. From the NMSE values, we observed that the performance ALOHA was quantitatively superior to the performance of l_1-wavelet and TV based compressed sensing approach. The reconstruction results by ALOHA has less perceivable distortion compared to those of l_1-wavelet and TV approaches. This can be easily observed from the magnified images in the third and the fourth rows of Fig. 5. In the case of TV, structural distortion around the image edges was easily recognizable. The reconstruction results of l_1-wavelets exhibits Gibb's ringing artifacts and distorted grey matter structure at the cerebellum area (see the fourth row), whereas the C-LORAKS resulted in a significant noise boost in the white matter area (see the third row). On the other hand, ALOHA reconstruction provided best edge structures without boosting noises.

Next, we compared our parallel imaging results with those of the existing approaches for additional multichannel saggital brain data set. The reconstruction results in Fig. 6 are from Gaussian random sampling patterns at the acceleration factor of 4. The NMSE results in Fig. 6 showed that ALOHA was most accurate. From the magnified images at the second row of Fig. 6, we observed that proposed method provided reconstruction results more accurately than other algorithms. More specifically, in Fig. 6, we observed that the reconstruction by GRAPPA and SAKE with ESPIRiT are much noisier compared to the ALOHA reconstruction. Moreover, the overall reconstruction errors from SAKE with ESPIRiT were higher than those from ALOHA. The reconstruction time was 7.6 sec with our preliminary GPU implementation of ALOHA. Note that the current GPU implementation is only preliminary because we still use the original CPU version of LMaFit algorithm from the original authors. Accordingly, with an optimized GPU implementation of LMaFit may further reduce the overall computational time, which will be reported later.

On the other hand, the reconstruction time for GRAPPA, SAKE and SAKE+ESPIRIT for this example were 9.8 sec, 118.4 sec, and 17.6 sec, respectively. For the case of SAKE, unlike the SAKE+ESPIRiT, the Hankel structured matrix is constructed using all k-space data similar to ALOHA, but the computational time was significantly higher than ALOHA. Accordingly, we found that our computation time was competitive to the existing approaches even with much improved reconstruction quality.

B. Dynamic MR Experiments

Using the sub-sampled k-space data at the acceleration factor of eight, the average NMSE values of k-t FOCUSS, C-LORAKS, and ALOHA, were 1.616×10^{-2}, 1.363×10^{-2}, and, 1.224×10^{-2}, respectively (see Fig. 7(a)). The sub-sampled data was collected according to a Gaussian distri-

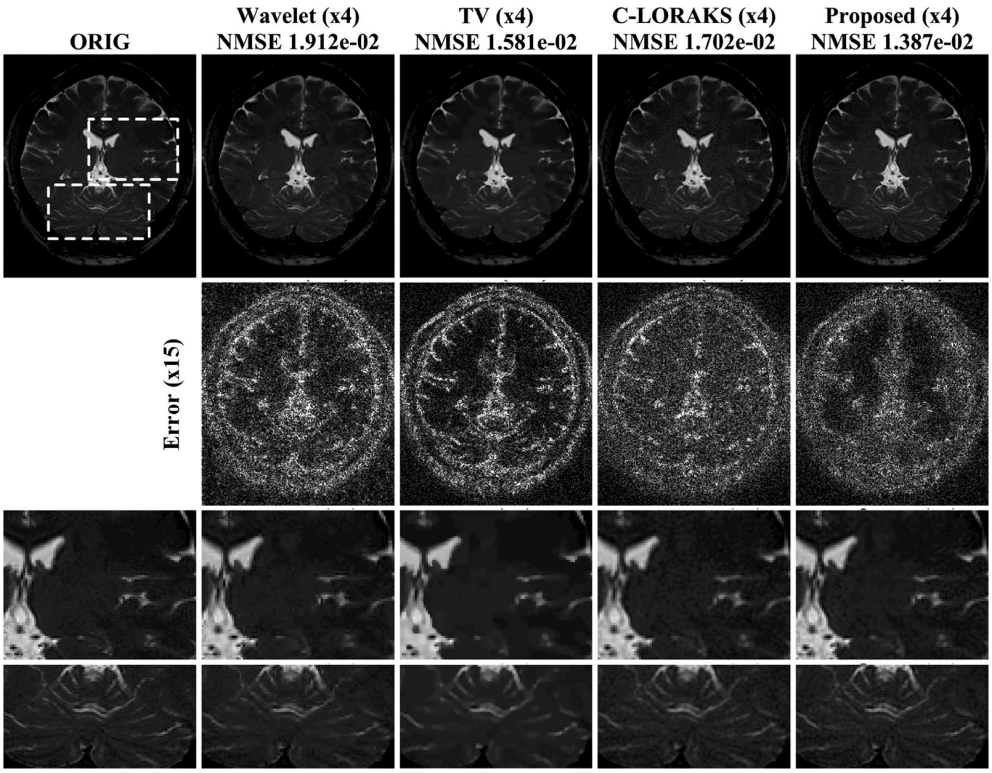

Fig. 5. Comparison with l_1-wavelet compressed sensing, TV compressed sensing, C-LORAKS, and the proposed method at the four fold acceleration factors. The data was acquired from a single channel coil. The first row shows reconstructed images, and the second row shows difference images between the ground-truth and the reconstructions, and the third and last row show the magnified views of distorted regions in the reconstructed images.

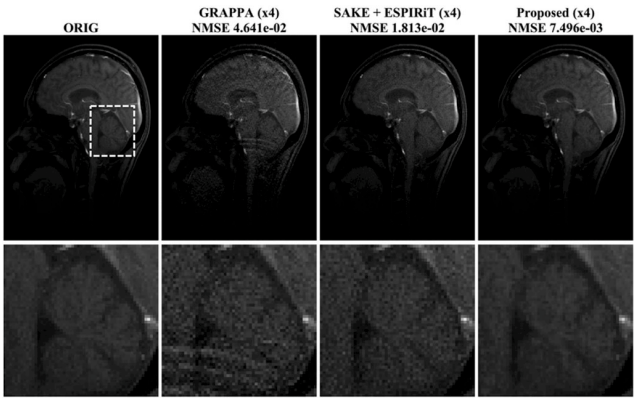

Fig. 6. Parallel imaging results using GRAPPA, SAKE with ESPIRiT and the proposed method at the four fold acceleration of Gaussian random sampling pattern. The second row shows the magnified views corresponding to a white box with a broken line. Note that GRAPPA requires ACS lines, so with the additional 50 samples along ACS region, the effective downsampling ratio was 3.1.

bution and included the four center lines around zero frequency. The average NMSE values were calculated using all temporal frames. The temporal NMSE plot in Fig. 7(a) also confirmed that the proposed method outperformed k-t FOCUSS and C-LORAKS across all temporal frames. Moreover, as shown in Fig. 7(a), the temporal profile of an heart area (indicated as a broken purple line) by the proposed reconstruction was the most accurate, which showed sharp transition to systole phase that were comparable to the true one, whereas the temporal variation in the k-t FOCUSS and C-LORAKS reconstruction became smoother and more blurry. Moreover, as indicated by the yellow arrow in the magnified heart area, false muscle structures were observed in the conventional approaches.

The NMSE values of the parallel dynamic imaging results from k-t FOCUSS, SAKE, and the proposed method using four coil k-space data were 8.75×10^{-3}, 4.983×10^{-3} and 3.571×10^{-3}, respectively, which quantitatively showed that the proposed method outperformed k-t FOCUSS and SAKE (see Fig. 7(b)). Reduced residual artifacts compared to the conventional approaches were perceivable in the ALOHA difference images in Fig. 7(b). Moreover, the temporal profiles of the proposed reconstruction showed much sharp transition to the systole phase which were comparable to the true one, whereas the dynamic slice profile from k-t FOCUSS and SAKE showed smoother and more blurry transition as shown in Fig. 7(b). Finally, in the magnified view of the heart area in Fig. 7(b), we can observed the broken heart muscle structures in the conventional reconstruction, which was well-preserved in ALOHA reconstruction.

VI. Discussion

Discussion can be found in the Supplementary Material.

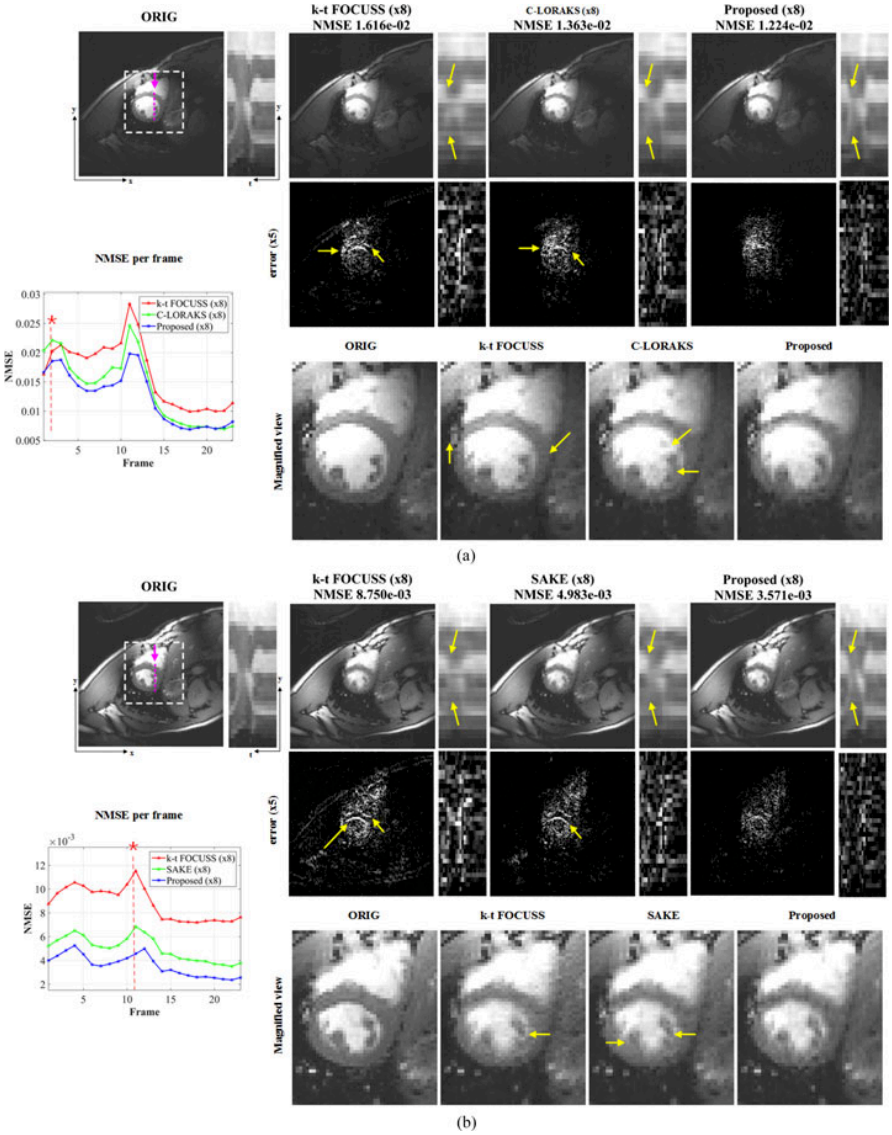

Fig. 7. Reconstruction results from 8 fold accelerated dynamic MR data using (a) single coil and (b) four coils data set. Purple lines denote the regions corresponding to y-t cross sections that are magnified along temporal axis. The second rows in both (a) and (b) show the difference images between the ground-truth and the reconstructions. Yellow arrows indicate noticeable differences between results.

VII. Conclusion

In this paper, we proposed a general framework for annihilating filter based low-rank Hankel matrix approach (ALOHA) for static and dynamic MRI inspired by recent calibration-free k-space methods such as SAKE and C-LORAKS. Because natural images can be much more effectively sparsified in the transform domains, we generalized the idea to include signals that can be sparsfied in the transform domains. Our analysis showed that the transform domain sparsity can be equivalently represented as a low-rank Hankel structured matrix in the weighted k-space domain, whose weighting function is determined solely by the transform, not by the data. When signals are effectively sparsified in dyadic wavelet transform, we showed that the corresponding low rank Hankel matrix completion problem can be implemented using a pyramidal decomposition, which significantly reduces the overall computational complexity and improves the noise robustness. For parallel imaging data, we verified that by stacking Hankel matrix from each coil side by side, we may exploit the coil sensitivity diversity.

Reconstruction results from single coil static MR imaging confirmed that the proposed method outperformed the existing compressed sensing frameworks. We further demonstrated superior performance of the proposed method in static parallel MR imaging even without calibration data. Furthermore, the algorithm was successfully extended to dynamic accelerated MRI along k-t domain with both single coil and multi coil dynamic MR data. Therefore, we concluded that the proposed algorithm was very effective in unifying the compressed sensing and parallel MRI.

ACKNOWLEDGEMENT

The authors would like to thank Prof. Michael Unser at EPFL for the many insightful discussions. The authors also like to thank Prof. Sung-Hong Park at KAIST for providing the single coil brain MR k-space data set.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[2] E. J. 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, Feb. 2006.
[3] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.
[4] D. Liang, B. Liu, J. Wang, and L. Ying, "Accelerating SENSE using compressed sensing," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1574–1584, 2009.
[5] J. P. Haldar, D. Hernando, and Z.-P. Liang, "Compressed-sensing MRI with random encoding," *IEEE Trans. Med. Imag.*, vol. 30, no. 4, pp. 893–903, Apr. 2011.
[6] S. S. Vasanawala, M. T. Alley, B. A. Hargreaves, R. A. Barth, J. M. Pauly, and M. Lustig, "Improved pediatric MR imaging with compressed sensing 1," *Radiology*, vol. 256, no. 2, pp. 607–616, 2010.
[7] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, 2009.
[8] H. Jung, J. Park, J. Yoo, and J. C. Ye, "Radial k-t FOCUSS for high-resolution cardiac cine MRI," *Magn. Reson. Med.*, vol. 63, no. 1, pp. 68–78, 2010.
[9] H. Yoon, K. Kim, D. Kim, Y. Bresler, and J. Ye, "Motion adaptive patch-based low-rank approach for compressed sensing cardiac cine MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 11, pp. 2069–2085, Nov. 2014.
[10] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: kt SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.
[11] T. Çukur, M. Lustig, and D. G. Nishimura, "Improving non-contrast-enhanced steady-state free precession angiography with compressed sensing," *Magn. Reson. Med.*, vol. 61, no. 5, pp. 1122–1131, 2009.
[12] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.
[13] M. A. Griswold *et al.*, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, 2002.
[14] K. P. Pruessmann, "Encoding and reconstruction in parallel MRI," *NMR Biomed.*, vol. 19, no. 3, pp. 288–299, 2006.
[15] M. Lustig and J. M. Pauly, "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space," *Magn. Reson. Med.*, vol. 64, no. 2, pp. 457–471, 2010.
[16] P. J. Shin *et al.*, "Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion," *Magn. Reson. Med.*, vol. 72, no. 4, pp. 959–970, 2014.
[17] J. Haldar, "Low-rank modeling of local k-space neighborhoods (LORAKS) for constrained MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 3, pp. 668–681, Mar. 2014.
[18] J. P. Haldar and J. Zhuo, "P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data," *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1499–1514, 2015.
[19] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
[20] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–Fix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.
[21] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2788–2805, Aug. 2005.
[22] G. Ongie and M. Jacob, "Off-the-grid recovery of piecewise constant images from few Fourier samples," *J. Imaging Sci.*, Vol. 9, no. 3, pp. 1004–1041, 2016.
[23] M. Unser, P. D. Tafti, and Q. Sun, "A unified formulation of Gaussian versus sparse stochastic processes–Part I: Continuous-domain theory," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1945–1962, May 2014.
[24] M. Unser, P. Tafti, A. Amini, and H. Kirshner, "A unified formulation of Gaussian versus sparse stochastic processes–Part II: Discrete-domain theory," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 3036–3051, May 2014.
[25] J. C. Ye, J. M. Kim, K. H. Jin, and K. Lee, "Compressive sampling using annihilating filter-based low-rank interpolation," arXiv preprint arXiv:1511.08975, Nov. 2015.
[26] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
[27] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
[28] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.
[29] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1548–1566, Mar. 2011.
[30] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, Jun. 2010.
[31] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6576–6601, Oct. 2014.
[32] D. Lee, K. H. Jin, E. Y. Kim, S.-H. Park, and J. C. Ye, "Acceleration of MR parameter mapping using annihilating filter-based low rank Hankel matrix (ALOHA)," *Magn. Reson. Med.*, Jan. 2016, DOI: 10.1002/mrm.26081.
[33] J. Lee, K. H. Jin, and J. C. Ye, "Reference-free single-pass EPI Nyquist ghost correction using annihilating filter-based low rank Hankel matrix (ALOHA)," *Magn. Reson. Med.*, Jan. 2016, DOI 10.1002/mrm.26077.
[34] M. Mani, M. Jacob, D. Kelley, and V. Magnotta, "Multi-shot multi-channel diffusion data recovery using structured low-rank matrix completion," Aug. 2016, doi: 10.1002/mrm.26382.
[35] M. Unser and P. D. Tafti, *An Introduction to Sparse Stochastic Processes*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[36] Y. Zhang and P. L. Dragotti, "On the reconstruction of wavelet-sparse signals from partial Fourier information," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1234–1238, Sep. 2015.
[37] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 202–219, Feb. 1999.
[38] M. Signoretto, V. Cevher, and J. A. Suykens, "An SVD-free approach to a class of structured low rank matrix optimization problems with application to system identification," in *IEEE Conference on Decision and Control*, no. EPFL-CONF-184990.
[39] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Program. Comput.*, vol. 4, no. 4, pp. 333–361, 2012.
[40] N. Srebro, "Learning with matrix factorizations," Ph.D. dissertation, Dept. of Elect. Eng., Comput. Sci., Massachusetts Inst. of Technol., Cambridge, MA, USA, 2004.
[41] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[42] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
[43] M. Uecker *et al.*, "ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 990–1001, 2014.

Kyong Hwan Jin, photograph and biography not available at the time of publication.

Dongwook Lee, photograph and biography not available at the time of publication.

Jong Chul Ye, photograph and biography not available at the time of publication.