State of the art and prospects of structured sensing matrices in compressed sensing

Kezhi LI[1], **Shuang CONG** (✉)[2]

1   Automatic Complex Communication Networks, Signal and Systems Center, Royal Institute of Technology (KTH), Stockholm 10044, Sweden
2   Department of Automation, University of Science and Technology of China, Hefei 230027, China

Abstract   Compressed sensing (CS) enables people to acquire the compressed measurements directly and recover sparse or compressible signals faithfully even when the sampling rate is much lower than the Nyquist rate. However, the pure random sensing matrices usually require huge memory for storage and high computational cost for signal reconstruction. Many structured sensing matrices have been proposed recently to simplify the sensing scheme and the hardware implementation in practice. Based on the restricted isometry property and coherence, couples of existing structured sensing matrices are reviewed in this paper, which have special structures, high recovery performance, and many advantages such as the simple construction, fast calculation and easy hardware implementation. The number of measurements and the universality of different structure matrices are compared.

Keywords   compressed sensing, structured sensing matrices, RIP, coherence

1   Introduction

In the digital revolution, people are now employing various signal processing techniques and new sensing systems in general electronic products with ever-increasing resolution and fidelity. The conventional manners of sampling signals, images, videos, or other data obey the celebrated Shannon's theorem, that requires to sample a signal at a sampling rate

at least twice the highest frequency present in a signal (so-called Nyquist rate) to retain signal information intact [1, 2]. The Shannon's theorem solves the problem in theory perfectly, yet unfortunately it is not omnipotent. In many applications such as remote surveillance or spectroscopy, sampling in the result with Nyquist rate is expensive, or even physically impossible. So as long as the recovery performance achieves an acceptable level, people want to build devices which are capable of acquiring samples at a necessary rate as low as possible. In some other applications, such as imaging system or video processing, sampling a large number of measurements seems feasible. However, because of the limited storage space and using of advanced compression techniques, people often discard the most received data, and just save a small amount of the compressed data (e.g., JEPG). Apparently it will waste lots of valuable sensing resources since the entire data are sampled at first.

Aiming at solving above problems, the compressed sensing (short for CS) theory [3–7] has become one of the hottest research areas in signal processing since 2006. The research of CS has been growing very fast and it focuses on acquiring and reconstructing sparse or compressible signals. By using CS, compressed measurements can be acquired directly and one may recover the original sparse or compressible signal faithfully even when the sampling rate is much lower than the Nyquist rate. An N-length signal $\mathbf{x}$ is regarded as sparse if $\mathbf{x}$ has K nonzero values and $K \ll N$. Compressible $\mathbf{x}$ means that $\mathbf{x}$ can be well-approximated by another sparse signal $\mathbf{f}$ in certain domain $\mathbf{\Psi}$ by using only K nonzero coefficients:

$\mathbf{x} = \mathbf{\Psi f}$, $|\mathbf{f}|_0 = K$. Normally the traditional compression techniques preserve the values and locations of the largest coefficients, such as JPEG, JPEG2000, MPEG. While CS has more efficient sensing or sampling protocols that capture the essential information content embedded in the original signal and obtain the condensed data straightforwardly. More precisely, these protocols are nonadaptive linear transforms, which can be represented by well-designed matrices, called *sensing matrices* $\mathbf{\Phi}$. These matrices should be incoherent to the sparsifying matrix $\mathbf{\Psi}$ of the compressible signal. With the measurements and the sensing matrix, the process of exact reconstructing signals from a subset of measurements can be implemented by solving a nonlinear optimization problem. The approaches to solving the nonlinear problems are named *reconstruction/recovery algorithms*. From a mathematical point of view, compressed sensing is also deemed as a technique of finding sparse solutions to underdetermined linear systems.

The CS theory is a revolution in both the theory of reliable signal sampling and physical design of sensors. Since the original signal can be sensed from fewer linear projections rather than acquired in its initial domain, the *sensing matrices* play an important role in the CS framework. The property of the sensing matrices affects the number of necessary measurements and the recovery performance directly. Early researchers proved that a random projection is one of the best solutions [5,8]. The projection matrices are generated by orthogonalizing measured vectors uniformly and independently on a unit sphere. In addition, sensing matrices consisting of independent and identically distributed (i.i.d.) entries drawn from a Gaussian or Bernoulli distribution also perform well in both theory and practice [3, 4]. Though the problem has been solved in mathematics, there still exist many obstacles to overcome. One main drawback of the pure random sensing matrices is that they require huge storage-memory, namely $M \times N$ entries to recover a length-N signal, and high computational cost for signal reconstruction. Moreover, the difficulty of hardware implementation also makes them expensive in practice.

To simplify the sensing scheme, many structured sensing matrices have been proposed in recent years. In this paper, after explaining some terminologies such as restricted isometry property and coherence, we give an introduction to couples of existing structured sensing matrices, including subsampled incoherent bases, random Toeplitz matrices, random demodulator matrices, random convolution matrices, structurally random matrices, convolutional matrices using sequences and some other structured sensing matrices. These matrices have special structures which equip them efficiency

in the construction, calculation or hardware implementation. For many of them, corresponding fast recovery algorithms have been developed by exploiting their specific structures. Here we put our emphasis on how the structured matrices are generated and what are their recovery performances. Based on these, the number of measurements and the universality of different structure matrices are compared. The paper will help readers to understand the characteristics of popular sensing matrices well and may inspire them to explore or pursue more efficient sensing schemes in the CS area.

The remaining paper is organized as follows. In Section 2, we describe the core concepts of this paper: sensing matrices, and introduce two prevalent criteria that examine the effectiveness of sensing matrices: restricted isometry property and coherence. In Section 3, couples of structured sensing matrices are analyzed. The overview of applications is discussed in Section 4. Finally, the prospects of structured sensing matrices are discussed in Section 5 followed by the conclusions in Section 6.

2 Sensing matrices

In our real world, normally the useful signals are not random. Images, videos or voices often contain specific structures and strong correlation among pixels, frames or samples. These structures and correlations are the assumptions behind the sparse representation theory. Given an N-dimensional signal $\mathbf{x}$, $\mathbf{\Psi}$ denotes the sparsifying transform basis for $\mathbf{x}$, where throughout this paper we assume that $\mathbf{\Psi}$ is an $N \times N$ normalized unitary matrix satisfying $\mathbf{\Psi}^* \mathbf{\Psi} = N \mathbf{I}_N$. So $\mathbf{x}$ can easily be decomposed by means of a linear superposition of K elementary components:

$$\mathbf{x} = \sum_{k=1}^{K} f_k \psi_k, \tag{1}$$

which can be rewritten in the form of matrix multiplication as

$$\mathbf{x} = \mathbf{\Psi f}, \tag{2}$$

where $\mathbf{f}$ is a length N sparse vertical vector with K nonzero values, $K \ll N$. Typical transforms $\mathbf{\Psi}$ include discrete Fourier transform (DFT), discrete cosine transform (DCT) and discrete wavelet transform (DWT). Sometimes the number of nonzero values in $\mathbf{f}$ is larger than K. In this case people usually encode the most significant K non-zero entries of $\mathbf{f}$ and disregard the rest, which is also the core principle of the image compression standard JPEG (using DCT) and JPEG2000 (using DWT).

Now assuming a length-N signal $\mathbf{x}$ as defined in Eq. (1), the data acquisition process can be described as

$$\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi \Psi f} = \mathbf{\Theta f}, \qquad (3)$$

where the measurement $\mathbf{y}$ represents an $M \times 1$ sampled vector, $\mathbf{\Phi}$ is an $M \times N$ measurement/sensing matrix, $\mathbf{\Theta} = \mathbf{\Phi \Psi}$. (3) is the kernel equation of the sensing process. The sensing process with a random Gaussian measurement matrix $\mathbf{\Phi}$ and a DCT matrix $\mathbf{\Psi}$ is illustrated in Fig. 1, in which there are four columns of $\mathbf{\Psi}$ that correspond to nonzero $\mathbf{f}_i$ coefficients; the measurement vector $\mathbf{y}$ is a linear combination of these columns. The CS theory considers problems based on the fundamental Eq. (3). These problems can be summarized as how to design efficient sensing matrix $\mathbf{\Phi}$, and how to recover $\mathbf{x}$ given $\mathbf{y}$ and $\mathbf{\Phi}$.

Here we focus on the problem of designing proper sensing matrices. Conventionally, if K entries in $\mathbf{y}$ are more important than other entries, people may capture the signal energy roughly and recover the original signal from K measurements, like what we do for recovering natural image from its frequency spectrum. However, in the CS framework the entries of $\mathbf{f}$ are assumed sparse and randomly distributed, which means people do not know where the large entries locate. In this circumstance $\mathbf{x}$ is able to be recovered by exploiting the sparsity from M measurements, $M \ll N$ and $M > K$. Because only partial M measurements are captured to recover the original signal, a good sensing scheme should spread out the information of the non-zero entries into every measurement $\mathbf{y}_k$ evenly, in case of losing significant information. Follow this intuitive idea, people found that random projection is one of the best candidates as a sensing matrix [5, 9]. Besides, if $\mathbf{\Phi}$ represents a Gaussian or Bernoulli random operator, $\mathbf{x}$ can also be faithfully recovered from $\mathbf{y}$ using nonlinear optimization approaches provided that $M \geqslant O(K \log(N/K))$ [3, 4, 10]. These early works made by D. Donoho, E. Candès, T. Tao and Romberg established the foundation of the CS theory.

From Eq. (3) an essential question might be raised instinctively: apart from the general random operators, what kinds of sensing matrices $\mathbf{\Phi}$ are capable to recover $\mathbf{x}$ uniquely from measurements $\mathbf{y}$? Fortunately, two important criteria for evaluating proper operators were created to provide fundamental insights into the geometry of sensing matrices. The most well-known one is often referred as the Restricted Isometry Property (RIP):

Definition 1 (RIP [5])   An $M \times N$ matrix $\mathbf{\Theta} = \mathbf{\Phi \Psi}$ is said to satisfy the RIP with parameters (K, δ) $(\delta \in (0, 1))$ if

$$(1 - \delta)\|\mathbf{f}\|^2 \leqslant \|\mathbf{\Theta f}\|^2 \leqslant (1 + \delta)\|\mathbf{f}\|^2, \text{ for all } \mathbf{f} \in \Gamma, \qquad (4)$$

where Γ represents the set of all length-N vectors with K nonzero coefficients.

Generally speaking, RIP requires the sensing matrix to act as a near isometry on the set of all K-sparse signals. It is consistent with the thought of spreading energy behind random sensing matrices. So measurement $\mathbf{y}$ preserves the energy that does not shrink or expand too much comparing with the original signal $\mathbf{x}$. If $\mathbf{\Theta}$ satisfies the RIP, many reconstruction algorithms can be used to recover any K-sparse signal $\mathbf{f}$ from M measurements $\mathbf{\Theta f}$, such as Basis Pursuit (BP) or Matching Pursuit recovery algorithms [11, 12]. In addition, RIP guarantees the uniqueness of the reconstruction result $\mathbf{f}$, which does not hold automatically for some other RIP-related property, such as the weaker Statistical Restricted Isometry Property (StRIP) [13].

Because there is no existing algorithm for efficiently verifying whether a matrix satisfies RIP, people also need the coherence property to examine the "quality" of $\mathbf{\Theta}$:

Definition 2 (Coherence [14, 15])   The coherence $\mu(\mathbf{\Theta})$ is the largest absolute inner product between any two
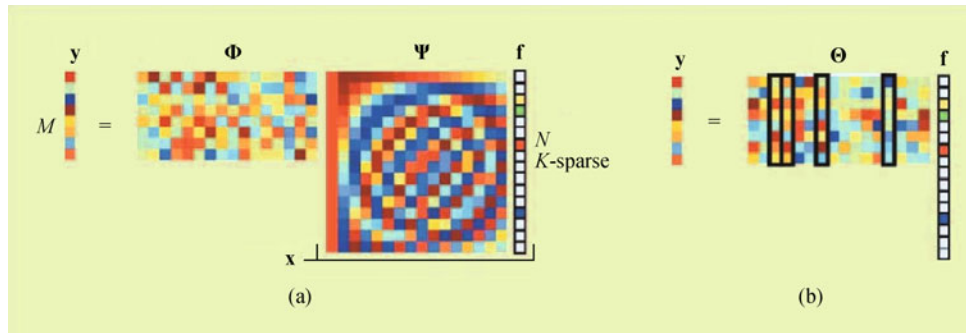
Fig. 1   (a) Compressive sensing measurement process with a random Gaussian measurement matrix $\mathbf{\Phi}$ and DCT matrix $\mathbf{\Psi}$ as sparsifying matrix; (b) measurement process with $\mathbf{\Theta} = \mathbf{\Phi \Psi}$. The original scheme figure is from [7]. We use $\mathbf{f}$ to denote the K sparse vector

normalized columns of $\mathbf{\Theta}$

$$\mu(\mathbf{\Theta}) = \max_{1 \leqslant i \neq j \leqslant N} |\langle \Theta_i, \Theta_j \rangle|, \qquad (5)$$

where Θ_i, Θ_j represent two columns of $\mathbf{\Theta}$.

If $\mathbf{\Theta} = \mathbf{\Phi\Psi}$, the coherence can also be quantified by calculating the maximal correlation among all rows of $\mathbf{\Phi}$ and all columns of $\mathbf{\Psi}$

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \max_{1 \leqslant i, j \leqslant N} |\langle \Phi_i, \Psi_j \rangle| = \max_{1 \leqslant i, j \leqslant N} |\mathbf{\Theta}(i, j)|. \qquad (6)$$

Note that for a unitary matrix $\mathbf{\Phi}$ with $\mathbf{\Phi}^*\mathbf{\Phi} = N\mathbf{I}_N$, the mutual coherence coefficient μ is bounded by $1 \leqslant \mu(\mathbf{\Theta}) \leqslant \sqrt{N}$ [6]. When $\mathbf{\Phi}$ is chosen as the DFT or the Walsh-Hadamard transform and $\mathbf{\Psi}$ is an identity matrix, $\mu(\mathbf{\Theta}) = 1$. If $\mathbf{\Phi}$ is a matrix of random basis vectors or a matrix of i.i.d. Gaussian entries $\mathcal{N}(0, 1)$, the mutual coherence between $\mathbf{\Phi}$ and any orthonormal matrix $\mathbf{\Psi}$ is on the order of $O\left(\sqrt{2 \log N}\right)$ with very high probability, far from the lower bound [16]. Coherence μ is a core concept in constructing CS matrix, and it will be frequently used in the sensing matrix analysis.

3 Structured sensing matrices

The initial work of CS focused on randomized sensing matrices, in which the entries of matrices are independently generated from standard probability distributions. For instance, with overwhelming probability, all matrices satisfying random Gaussian/Bernoulli distribution obey the RIP could be uniquely recovered from number of measurements M and

$$M \geqslant C \cdot K \log(N/K) \qquad (7)$$

where C is some constant depending on each instance [4]. As mentioned in the introduction, pure random matrices are not easily applicable to real implementations due to its large storage and heavy computation. In recent years some structured sensing matrices have been proposed. Unlike pure random matrices, special constructions make structured sensing matrices suitable for various applications, and we will introduce them chronologically and analyze their performances respectively.

3.1 Subsampled incoherent bases

For subsampled incoherent base matrices, the most famous examples are random subsampled Fourier and Walsh-Hadamard matrices. An $M \times N$ sensing matrix is constructed by random selecting rows from an $N \times N$ square DFT (or FFT)

matrix $\mathbf{F}$ or a Walsh-Hadamard transform (WHT) matrix $\mathbf{H}$, respectively. Specifically,

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{bmatrix}, \qquad (8)$$

where $\omega = e^{-\frac{2\pi i}{N}}$ is a primitive Nth root of unity in which $i = \sqrt{-1}$ and

$$\mathbf{H}_N = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{bmatrix}, \qquad (9)$$

with an initial matrix $\mathbf{H}_1 = [+1]$. The oversampling factor for partial DFT matrix was proved as $(\log N)^6$ at first in [5], then was improved to $(\log N)^4$ in [17]. Generally speaking, the RIP property of sampled unitary matrix is summarized as following theorems.

Theorem 1 (RIP for randomly subsampled unitary matrix [17, 18])   Suppose that the $M \times N$ matrix $\mathbf{\Theta}$ is a randomly subsampled unitary matrix, i.e., it can be written as $\mathbf{\Theta} = \frac{1}{\sqrt{M}}\mathbf{R}_{\Omega}\mathbf{U}$, where $\frac{1}{\sqrt{M}}$ is a normalizing coefficient, $\mathbf{R}_{\Omega}$ is a random sampling operator which selects M samples out of N ones uniformly at random, and $\mathbf{U}$ is an $N \times N$ unitary matrix satisfying $\mathbf{U}^*\mathbf{U} = N\mathbf{I}_N$. Then $\mathbf{\Theta}$ satisfies the RIP with high probability provided that

$$M \geqslant O\left(\delta^{-2}\mu^2(\mathbf{U})K \log^4 N\right), \qquad (10)$$

where δ denotes the restricted isometry constant in Definition 1.

Theorem 1 implies that the RIP bound of a randomly subsampled unitary matrix depends on $\mu(\mathbf{U})$. Note that for a unitary matrix $\mathbf{U}$ with $\mathbf{U}^*\mathbf{U} = N\mathbf{I}_N$, $1 \leqslant \mu(\mathbf{U}) \leqslant \sqrt{N}$. When $\mathbf{U}$ is chosen as the FFT or the Walsh-Hadamard transform, $\mu(\mathbf{U}) = 1$ and by Eq. (10), one has

$$M \geqslant O\left(\delta^{-2}K \log^4 N\right). \qquad (11)$$

All above bounds are for the uniform reconstruction, which means that once the sampling operator $\mathbf{\Phi}$ is constructed, all sparse signals in a certain basis $\mathbf{\Psi}$ can be recovered as long as M is sufficiently large. If one fixes $\mathbf{x}$ and wants to recover it specifically, the problem turns to a non-uniform one and this weaker assumption leads to less measurements. In detail:

Theorem 2 (Non-uniform recovery [19])   Assume that $\mathbf{\Theta}$ is a randomly subsampled unitary matrix that follows the same

definition as in Theorem 1. Let $\mathbf{f}$ in (3) be a fixed arbitrary K-sparse signal. Then $\mathbf{f}$ can be faithfully recovered from $\mathbf{y}$ using l_1 norm optimization, if M satisfies

$$M \geqslant O(\mu^2(\mathbf{U})K \log N). \tag{12}$$

In addition, if we fix $\mathbf{f} \in \mathbf{R}^N$ and suppose that the coefficient sequence $\mathbf{f}$ of $\mathbf{x}$ is K-sparse in the basis $\mathbf{\Psi}$; select M measurements in the $\mathbf{\Phi}$ domain uniformly at random, then if

$$M \geqslant C \cdot \mu^2(\mathbf{\Phi}, \mathbf{\Psi})K \log N, \tag{13}$$

for some positive constant C, its l_1 norm minimization solution is exact with overwhelming probability [6, 20]. For the cases of DFT and WHT matrices, $\mathbf{\Phi} = \mathbf{F}$ or $\mathbf{H}$, the bound of M holds for $O(K \log N)$. The theorems listed here are also very useful to prove the feasibility of other structured sensing matrices.

Although partial FFT (or WHT) has near-optimal theoretical guarantee, easy hardware implementation and fast-computable recovery, its major shortcoming is the lack of the *universality* property. A universal sensing matrix means that the matrix can handle signals that are sparse in any domain. If $\mathbf{\Phi}$ is a Gaussian random matrix, the matrix $\mathbf{\Phi}\mathbf{\Psi}$ will remain Gaussian for any unitary transform $\mathbf{\Psi}$. However, if $\mathbf{\Phi}$ is randomly sampled from a FFT, it will not be universal, as $\mu(\mathbf{F}\mathbf{\Psi})$ can not be $O(1)$ for all bases $\mathbf{\Psi}$, eg. when $\mathbf{\Psi} = \mathbf{F}^*$, $\mu(\mathbf{F}\mathbf{\Psi})$ will be large.

3.2   Random toeplitz matrices

Because all elements in random matrices are required to satisfy the i.i.d. random distribution, it becomes natural to raise a question one step further: can we reduce the randomness a little and achieve a similar reconstruction performance? Bajwa et al. first followed this thought to propose random Toeplitz matrices (RTM) in 2007 [21, 22]. In RTM, the entries are independence distributed in one row, while reserve certain structure among other rows. Specifically, if a probability distribution $P(a)$ yields an i.i.d. CS matrix (having unit-norm columns in expectation) then an $M \times N$ (partial) Toeplitz matrix $\mathbf{A}$ (also having unit-norm columns in expectation) of the form

$$\mathbf{A} = \begin{bmatrix} a_{N-1} & a_{N-2} & \cdots & a_0 \\ a_N & a_{N-1} & \cdots & a_1 \\ \vdots & \vdots & & \vdots \\ a_{N+M-2} & a_{N+M-3} & \cdots & a_{M-1} \end{bmatrix}, \tag{14}$$

where the entries $\{a_i\}_{i=1}^{N+M-2}$ have been drawn independently from $P(a)$, is also a CS matrix in the sense that it satisfies

RIP of order $3K$ with high probability for every $\delta \in (0, 1/3)$ provided $M > C \cdot K^3 \log(N/K)$, where C is a constant [21].

In the technical aspect, the proof of RIP of RTM used the celebrated Hajnal-Szemeredi theorem on equitable coloring of graphs to partition an $M \times 3K$ Toeplitz-structured submatrix $\mathbf{A}_T$ into roughly $O(K^2)$ i.i.d. submatrices having dimensions approximately equal to $O(M/K^2) \times 3K$. By using random Toeplitz matrices, only $O(N)$ independent random variables are required to generate. Multiplication with Toeplitz matrices can be more efficiently implemented using fast Fourier transform, resulting in faster acquisition and reconstruction algorithms. In addition, Toeplitz-structured matrices meet the naturally requirement for certain application areas such as system identification. Later Haupt et. al. and Rauhut improved the bound of M to $O(K^2 \log N)$ [22] and $O(K \log^2(N))$ [23], respectively.

Meanwhile, random Toeplitz matrices also have disadvantages. For example, RTM are proved to be able to recover signals sparse only in the time domain. Their strong structures make them not suitable for processing signals sparse in other bases, such as DCT domain.

3.3   Random demodulator

The random demodulation (RD) matrix was proposed by Tropp et al. in 2010 [24]. Pseudorandom binary sequence are often used to modulate the input signal. Similar implementations include Bernoulli or Rademacher random variables. The random demodulator is a sampling system that can be used to acquire sparse, bandlimited signals in an analog model. Figure 2 displays a block diagram for the RD system [24]. It is for a continuous-time signal f whose highest frequency is less than $W/2$ hertz. Tropp et al. modulated the signal by multiplying the signal with a high-rate pseudonoise sequence, which smeared the tones across the entire spectrum. Then a low-pass anti-aliasing filter was applied to capture the signal $\mathbf{x}$ by sampling $\mathbf{x}$ at a relatively low rate. Simulations suggested that the RD requires just $O(K \log(W/K))$ samples per second to stably reconstruct the original signal.
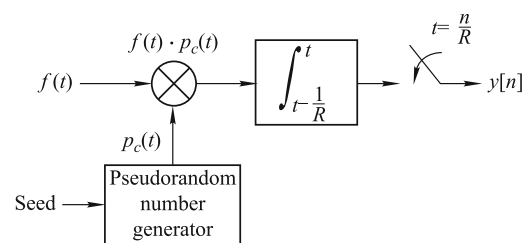
Fig. 2   Block diagram for the random demodulator. The components include a random number generator, a mixer, an accumulator, and a sampler (taken from [24])

In mathematics, the random demodulator can be seen as a linear system that maps a continuous-time signal to a discrete sequence of samples. To express the system in matrix form, let $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{W-1}$ be the chipping sequence in a diagonal matrix $\mathbf{D}$, $\mathbf{H}$ is an $R \times W$ accumulate-and-dump sampler matrix, where R is the sampling rate. Assume that W is divisible by R, the overall action of the system is

$$\mathbf{\Theta} = \mathbf{\Phi\Psi} = \mathbf{HD} \cdot \hat{\mathbf{F}}, \qquad (15)$$

where

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & \cdots & & & \\ & & & 1 & 1 & \cdots \\ & & & & & \ddots \\ & & & & & 1 & 1 & \cdots \end{bmatrix}, \qquad (16)$$

there are W/R 1 in each row of $\mathbf{H}$, and

$$\mathbf{D} = \begin{bmatrix} \varepsilon_0 & & & \\ & \varepsilon_1 & & \\ & & \ddots & \\ & & & \varepsilon_{W-1} \end{bmatrix}, \qquad (17)$$

and $\hat{\mathbf{F}}$ is a $W \times W$ permuted DFT matrix with

$$\hat{\mathbf{F}} = \frac{1}{\sqrt{W}} \left[e^{-2\pi i \cdot nw/W} \right]_{n,w}, \qquad (18)$$

where $n = 0, 1, \ldots, W-1$ and $w = 0, \pm 1, \ldots, \pm(W/2-1), W/2$.

The main advantage of the RD system is it bypasses the need for a high-rate analog-to-digital converter (ADC). It is typically much easier to implement demodulation rather than sampling, thus a low-rate ADC is allowed to use and a more robust system with low-power can be achieved. In theory, the RD guaranteed the recovery of random general signals with the sampling rate of $R \sim O(K \log W + K \log^3 W)$ in the noiseless case and $R \sim O(K \log^6 W)$ in the noisy case, where C is a positive constant.

3.4 Random convolution

The random convolution (RC) model was first proposed by Romberg in 2007 [25, 26]. In the RC, the construction has two steps. The signal $\mathbf{x} \in \mathbb{R}^N$ was circularly convolved with a "pulse" $\mathbf{h} \in \mathbb{R}^N$, then subsampled. The pulse is supposed to be random and its energy spreads uniformly across the discrete spectrum. If one writes the convolution of $\mathbf{x}$ and $\mathbf{h}$ into the matrix form as $\mathbf{Hx}$, where [25]

$$\mathbf{H} = N^{-1/2} \mathbf{F}^* \mathbf{\Sigma F}, \qquad (19)$$

with $\mathbf{F}$ as the discrete Fourier matrix and $\mathbf{\Sigma}$ as a diagonal matrix whose non-zero elements are the Fourier transform of $\mathbf{h}$.

The matrix $\mathbf{\Sigma}$ can be generated by

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_0 & & & \\ & \sigma_1 & & \\ & & \ddots & \\ & & & \sigma_{N-1} \end{bmatrix}, \qquad (20)$$

where the diagonal entries σ_w are unit magnitude complex numbers with random phases as follows:

$$\begin{aligned} &w = 0 : \sigma \sim \pm 1 \text{ with equal probability,} \\ &1 \leqslant w < N/2 : \sigma_w = e^{j\theta_w}, \text{ where } \theta_w \sim \text{Uniform}([0, 2\pi]) \\ &w = N/2 : \sigma_{N/2+1} \sim \pm 1 \text{ with equal probability} \\ &N/2 + 1 \leqslant w \leqslant N - 1 : \sigma_w = \sigma_{N-w}^*, \text{the conjugate of } \sigma_{N-w}. \end{aligned} \qquad (21)$$

From Eq. (21) one can see that the action of $\mathbf{H}$ on a signal $\mathbf{x}$ can be broken down into a DFT followed by a randomization of the phases with symmetric constraints, followed by an inverse DFT. Fourier optics imaging architecture implementing random convolution followed by randomly premodulated summation (RPMS) is shown in Fig. 3. Alternatively, the random sampling process can also be substituted with randomly pre-modulated summation, which means to break them into blocks of size N/M, and summarize each block with a single number. This action will influence the bound of sufficient recovery measurements with a factor of $\log N$.
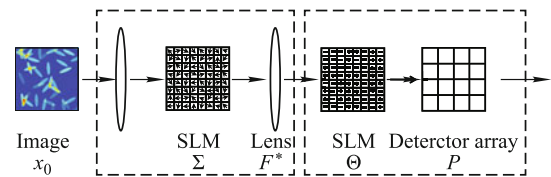
Fig. 3 Fourier optics imaging architecture implementing random convolution followed by RPMS [25]. SLM represents the spatial light modulator

Random convolution is significant since it is deemed as an efficient data acquisition strategy that can recover noiseless N-length signals in any fixed representation from $O(K \log N)$ measurements, which is relatively small for structured CS matrices. The randomness exists in both sampling process and entries generation, making RC universal (or uniform) towards the choice of signal representation. It is specially important for signals sparse in unknown bases.

3.5 Structurally random matrices

Structurally random matrix (SRM) is a novel framework of fast and efficient CS introduced by Do et al. [27, 28]. In the SRM, the sensing signal is prerandomized by scrambling its
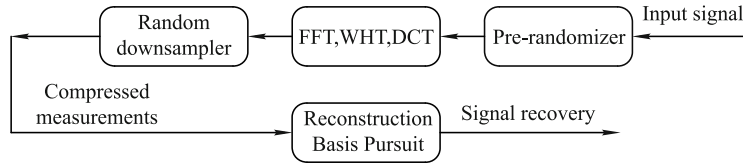
Fig. 4   Block diagram for sampling scheme of SRM [27]

sample locations for flipping its sample signs and then fast-transforming the randomized samples. The sensing measurements are obtained by subsampling the resulting transform coefficients finally. The sampling algorithm contains three steps. The diagram is illustrated in Fig. 4.

As shown in Fig. 4, the sampling procedure is (i) pre-randomizing a signal; (ii) applying some fast transform to the randomized signal; (iii) randomly subsampling the transform coefficients to get compressed measurements. If decomposing the algorithm mathematically as a product of 3 matrices, then the SRM can be represented as [27],

$$\mathbf{A} = \mathbf{DFR}, \tag{22}$$

where

- **R**, the randomizer, is a random permutation matrix (denoted as the global randomizer) or a random diagonal matrix of Bernoulli i.i.d. entries (denoted as the local randomizer).

- **F** is some computable transform such as the FFT, the DCT, the WHT, etc.

- **D**, the random downsampler, is a matrix composed of nonzero rows of a random diagonal matrix whose diagonal entries D_{ii} are i.i.d. binary random variables with $P(D_{ii} = 1) = M/N$, where M is the number of measurements.

The reconstruction algorithm can be any l_1 norm minimization or greedy pursuit algorithm. SRMs are highly relevant to large-scale, real-time compressed sensing applications as they have fast computation and support block-based processing. Meanwhile, SRMs have theoretical sensing performance of $O(K \log N)$ measurements for exact recovery, which is comparable to that of completely random sensing matrices. In the construction of sensing operator, SRMs use the random downsampler, fast transform and random diagonal matrix, like random convolution. SRMs provide the properties of universality and hardware implementation friendliness for reconstructing sparse signals.

3.6   Structured matrices using sequences

In the most previous work, random sequences have been ex-ploited to generate sensing matrices. [22, 29] use Bernoulli random sequence. An alternative way is to obtain matrices from diagonal unimodular sequences σ with random phases [25], i.e., $\sigma_k = e^{j\theta_k}$, where θ_k is a random variable that is uniformly distributed in $[0, 2\pi)$. In [30, 31], σ can be perfect or nearly perfect sequences.

Different from random sequences, recently many researchers adopt deterministic sequences to construct sensing matrices. These sequences are generated delicately and many of them have been widely implemented in communication and coding theory. Because the sequences have determined the formulation, the sensing matrices based on sequences often have less randomness, and many of them are even deterministic [13, 32–35]. Here we only introduce one of them named convolutional matrices using deterministic filter [34] as an example to have a look how to construct sensing matrices employing sequences.

The sampling operator $\mathbf{\Phi}$ can be represented as a partial circulant matrix with the following form [34]

$$\mathbf{\Phi} = \frac{1}{\sqrt{M}}\mathbf{R}_\Omega \mathbf{A}, \tag{23}$$

where $\mathbf{A}$ is a circulant matrix that can be expressed as

$$\mathbf{A} = \begin{bmatrix} a_0 & a_{N-1} & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_2 \\ \vdots & \vdots & & \vdots \\ a_{N-1} & a_{N-2} & \cdots & a_0 \end{bmatrix}. \tag{24}$$

For $\mathbf{\Phi}$ given in (23), the measurement process can be realized by circularly convolving $\mathbf{x}$ with a filter $\mathbf{a} = [a_0\ a_1\ \cdots\ a_{N-1}]^{\mathrm{T}}$ and then downsample the output at locations indexed by Ω. As known the circulant matrix $\mathbf{A}$ can be diagonalized using FFT. This property enables the convolutional matrix with fast computation. It is easy to see that the filter vector $\mathbf{a}$ (i.e., the first column of $\mathbf{A}$) can be obtained by taking the inverse FFT of sequence $\sigma = [\sigma_0\ \sigma_1\ \cdots\ \sigma_{N-1}]^{\mathrm{T}}$, i.e.,

$$\mathbf{a} = \frac{1}{\sqrt{N}}\mathbf{F}^*\sigma. \tag{25}$$

σ may adopt various unimodular sequences. The coherence bounds for different sequences are given in Table 1. For real

Table 1 Coherence parameter $\mu(\mathbf{A})$ for different diagonal sequences σ

	σ	N	$\mu(\mathbf{A})$
Complex matrices	FZC	Arbitrary	1
	m-sequence	$2^k - 1, k \in \mathbb{N}$	$\sqrt{1 + \frac{1}{N}}$
	Legendre sequence	$N \equiv 3 \pmod 4$ and N prime	$\sqrt{1 + \frac{1}{N}}$
		$N \equiv 1 \pmod 4$ and N prime	$1 + \frac{1}{\sqrt{N}}$
	Golay sequence	$2^{\kappa_1} 10^{\kappa_2} 26^{\kappa_3}, \kappa_1, \kappa_2, \kappa_3 \in \mathbb{N}$	$\sqrt{2}$
Real matrices	Extended FZC	Even N	$4 + \frac{4}{\sqrt{N}}$
		Odd N	$2.69 + \frac{8.15}{\sqrt{N}}$
	Extended Golay	Even N, $N = 2^{\kappa_1} 10^{\kappa_2} 26^{\kappa_3}, \kappa_1, \kappa_2, \kappa_3 \in \mathbb{N}$	$2\left(1 + \frac{1}{\sqrt{N}}\right)$
		Odd N, $N = 2^{\kappa_1} 10^{\kappa_2} 26^{\kappa_3} - 1, \kappa_1, \kappa_2, \kappa_3 \in \mathbb{N}$	$1 + \frac{2}{\sqrt{N}}$

sensing matrices $\mathbf{A}$, the diagonal sequence needs to be conjugate symmetric, shown as extended sequences in Table 1.

Using the uniform and non-uniform theorems, the coherence bounds reveal that $M \geqslant O(\delta^{-2} \mu K \log^4 N)$ measurements are enough for uniform recovery and $M \geqslant O(\delta^{-2} \mu K \log N)$ for non-uniform recovery, where δ denotes the restricted isometry constant. These convolutional matrices are not universal, while they show the effectiveness for signals sparse in both the time and frequency domain. When σ is the Frank-Zadoff-Chu (FZC) sequence, the corresponding sensing matrices are also capable for recovering signals sparse in the DCT domain. The number of measurements in regard to different sequences can be calculated easily from Table 1 and Theorems 1 and 2.

3.7 Other sensing matrices

Many other sensing matrices were developed in recent years. To accelerate the computational speed for large data, block structures were introduced for Gaussian matrix [36], Toeplitz matrix [37], Hadamard matrix [38] and SRM [27], etc. The block structure means the sensing matrices have the following form with structured matrices as blocks $\mathbf{A}_i, i = 1, 2, \ldots, l$.

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & & & \\ & \mathbf{A}_2 & & \\ & & \ddots & \\ & & & \mathbf{A}_l \end{bmatrix}. \tag{26}$$

Block-based sensing has more advantageous for realtime applications since the encoder does not need to send the sampled data until the whole signal is measured. Besides structure sensing matrices, the sensing matrices can even be deterministic. Various deterministic matrices have been introduced in [13, 32, 39–42]. Comparing to structured sensing matrices, deterministic sensing matrices has fixed forms and there is no randomness in the construction. Specifically, sec-

ond order Reed-Muller codes are used in [13, 43] and dual of extended binary BCH codes are exploited in [13, 44]. Many other sequences are also employed in the deterministic matrix design, such as discrete chirp sequences [13, 32, 34], Kerdock and Delsarte-Goethals codes [45], Sidelnikov sequences [46] and Alltop sequences [13, 47, 48], etc. Deterministic sensing matrices have fixed constructions, and thus normally they can not guarantee to recover all signals with high probability. They are able to recover most signals but an exponential fraction with high probability. Some papers focus on the problems of designing sensing matrices that lead to good (expected-case) mean squared error (MSE) performance rather than the worst case [49]. For more information regarding sensing matrices the readers may refer to references [18, 50] and CS website [51].

3.8 Relations between structured sensing matrices

The sensing matrices introduced in this section are not developed independently. They are associated with each other. Subsampled Fourier and Hadamard matrices were firstly proved as the qualified structured sensing matrices. They belong to the subsampled incoherent bases group of matrices. Random Toeplitz matrices are very famous and significant to many applications, such as channel estimation [22] and system identification [40]. The randomness exists in each row while between rows they have strong structure. In real time signal processing the modulation idea has been widely implemented, which is also used in random demodulator. The celebrated random convolution is actually a special modulation of signals in the Fourier domain. Moreover, SRM are a group of structured matrices generated from an approach based on random convolution but with Bernoulli diagonal phase modulation for signals in more flexible domains. Finally, structured and deterministic sensing matrices using sequences are analyzed as a new sub-area in sensing matrix design.

Practically people may utilize different structured sens-

ing matrices according to the sensing models and hardware constraints. For instance, if one needs structured matrices to model the 1-dimensional convolution in sensing processing, random Toeplitz or Circulant matrices are employed due to the natural of the convolution calculation. In addition, in the same model if the objective signal is sparse in the Fourier domain and the phases of the modulated signal can be symmetrical, random convolution are suitable to solve this problem accelerated by fast algorithms. When the phases can only be modulated as ± 1 and should be determined in advance for hardware reasons, convolutional CS matrices using deterministic sequences are recommended with a price of larger number of measurements M comparing with that of random convolution. If fixing the entire sensing scheme, the sensing matrices will be deterministic and there is no randomness in construction. In this case deterministic matrices constructed from coding theory are the only candidates, and usually they are with strict size constraints. In general, there is a tradeoff between randomness and number of measurements M. Less randomness facilitates the sensing scheme, however it often leads to more measurements and consequently longer sensing time.

4   Applications of structured sensing matrices

Essentially CS theory can be recognized as a data processing technique that recovers sparse data from under-determined equations. The advantage of CS is to process sparse signals that can not be processed appropriately before, or obtain the compressed data using proper physical instruments directly. Fortunately most of the signals in the real world that people are interest in belong to sparse signals or can be approximated in certain domain. So from its emergence CS has been implemented in numerous applications including communications, machine learning, imaging, geophysical data analysis, radar, remote sensing, data streaming, quantum state tomography, and so on. For instance the matrices mentioned previously, the Toeplitz matrices are quite suitable for communication channel estimation [22]; random demodulators are designed for sampling of sparse wideband analog signals [24, 52]; random convolution matrices can be exploited in radar imaging [25]; also, the validity of SRMs has been verified in image processing [27, 28]; convolutional matrices using sequences have widely applications in communication and signal processing [34, 35, 53, 54]. Apart from these works, here we simply introduce two celebrated applications of CS, in medical imaging and single pixel camera.

A promising application for compressed sensing is in reducing the sampling rate in magnetic resonance imaging (MRI) [55–57]. The main motivation of CS MRI is that, MRI scanners sequentially sample the human's body in the 2-D continuous Fourier domain, and sensed coefficients satisfy the sparse property which is also the prerequisite of the theory of CS. Moreover, MRI is very time consuming. In order to obtain a clearer image, one often needs a long time to collect the data. However, the speed of data collection is limited by physical and physiological constraints. Applying the CS technique may accelerate the scanning process with the same accuracy due to fewer CS measurements being required. The schematic diagram of MRI using CS is shown in Fig. 5(a).

Another category of the application involves the design of new acquisition hardware that is able to acquire projections of a signal against a class of vectors. In this case, the sensing process is accomplished by physical optical instruments, and the research normally focuses on the problem of how to design sensing matrices whose entries belong to some patterns/bases that can be easily implemented on the hardware. One example is the framework of recovering an image based on optical modulators, known as the single pixel camera shown in Fig. 5(b) [58]. The digital micromirror device (DMD) is a reflective spatial light modulator that selectively redirects parts of the light beam [59]. The DMD is comprised of an array of bacterium-sized, electrostatically actuated micro-mirrors, and each mirror rotates about a hinge and can swing between two stages $+10°$ or $-10°$. The state of each mirror depends on the bit loaded in the corresponding position of the programmable sensing matrix, and many structured sensing matrices may be implemented in this scenario. People have tested that the system works well when matrix entries are drawn randomly from a fast transform such as a Walsh Hadamard transform [60]. Many advanced imaging hardware architectures based on the single pixel camera model have been developed after these techniques mature, e.g., in terahertz imaging [61, 62]. With regard to real applications, actually it is not trivial to decide which strategy or structured sensing matrices we shall use. Because different matrices have their own features and performances, we have to investigate the practical scenarios and make the tradeoff between number of measurements M, universality or not, hardware constraints, computation and so on.

5   Prospects and future works

As a key research area of the encoding part of the com-
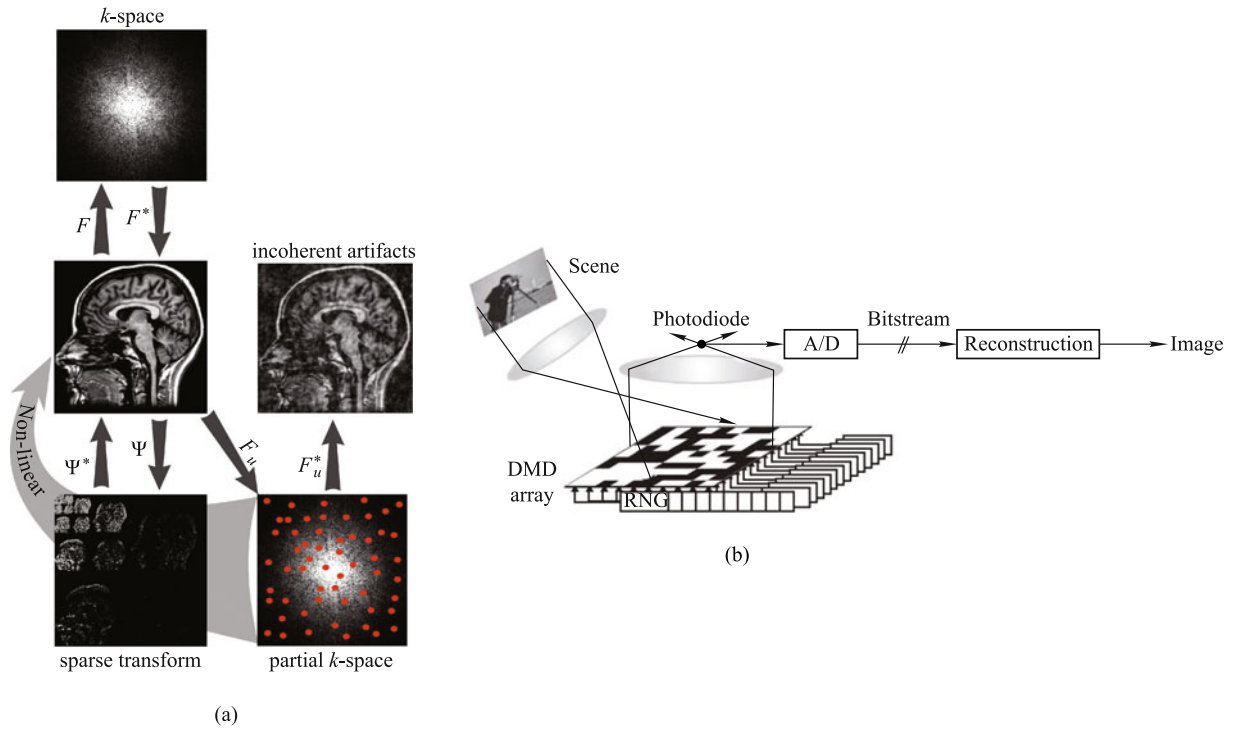
Fig. 5 (a) Illustration of the domains and operators used in [57] as the requirements of CS: sparsity in the transform domain, incoherence of the undersampling artifacts and the need for nonlinear reconstruction that enforces sparsity; (b) diagram of the single pixel camera. The image **x** is reflected off a digital micro-mirror device (DMD) array whose mirror orientations are modulated in the pseudorandom pattern supplied by the random number generator (RNG) [58]

pressed sensing theory, the research of structured sensing matrices is really important and has attracted more and more attention in recent several years. Although in literatures people have proposed many structured sensing matrices, matrices with special structures are deadly needed with regard to special settings or hardware requirements.

Generally speaking, the future development of sensing matrices will focus on two aspects. The first one is to use less randomness and less memory storage. For instance, comparing with full random matrices, more sparse sensing matrices with certain structure have and will be exploited to reduce the calculation in CS [63]. The structure of a network also can be embodied in a matrix revealed by a one-to-one correspondence with an expander graph [64]. The second aspect is to design sensing matrices satisfying certain structure in reality. This will be the main motivation for developing more structured sensing matrices. Take several examples to illustrate it. In the communication system the convolution process is equivalent to a Toeplitz matrix consisting of the transmitting signals multiplying the system impulse response function. That's the reason why Toeplitz CS matrices could be utilized in sparse channel estimation [22]. In [52] the authors proposed a practical sampling system called modulated wideband converter (MWC) by adopting periodic waveforms, a low-pass filter and a low rate sampler.

They proved that perfect recovery of multi-bandlimited signals from the proposed samples can be achieved under certain necessary and sufficient conditions. In mathematics, the sampling process can be reformed as a structured sensing matrices $\mathbf{y} = \mathbf{S\bar{F}D}$, where $\mathbf{S\bar{F}D}$ represent the sign pattern matrix, reorder Fourier matrix and diagonal matrix, respectively. This matrix structure comes from the hardware design, and it performs well in practice [52]. In addition, the structured sensing matrices were also implemented in statistical physics, such as the seeding matrix with coupling block diagonal structure. This work was proposed in [65] for a framework named seeded compressed sensing. Krzakala et al. proved that in their model the experimental recovery results approached the theoretical limit for large systems. To sum up, people will continue to work on pursuing various structured sensing matrices with less randomness/measurements, better performances and hardware friendly property incorporating knowledge from other fields such as coding theory, communication, random matrix theory etc., subject to specific requirements based on real settings.

6 Conclusion

After explaining the fundamental knowledge of sensing ma-

trices, RIP and coherence, we reviewed couples of existing structured sensing matrices, including subsampled incoherent bases, random Toeplitz matrices, random demodulator matrices, random convolution matrices, structurally random matrices and other structured sensing matrices. For each of them, we concentrated on the structure of the matrix, the measurement bounds and its scope of application. Though it is difficult to cover all of the developments in structured sensing matrices area, here we aim to explain the main idea and demonstrate a few well-known examples that are representatives of a wider class of the CS problem.

References

1. Shannon C. Communication in the presence of noise. Proceedings of the Institute of Radio Engineers (IRE), 1949, 37(1): 10–21

2. Nyquist H. Certain topics in telegraph transmission theory. Transactions of the American Institute of Electrical Engineers, 1928, 47(2): 617–644

3. Donoho D L. Compressed sensing. IEEE Transactions on Information Theory, 2006, 52(7): 1289–1306

4. Candès E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory, 2006, 52(2): 489–509

5. Candès E, Tao T. Near optimal signal recovery from random projections: universal encoding strategies. IEEE Transactions on Information Theory, 2006, 52: 5406–5425

6. Candès E, Wakin M. An introduction to compressive sampling. IEEE Signal Processing Magazine, 2008, 25(2): 21–30

7. Baraniuk R. Compressive sensing. IEEE Signal Processing Magazine, 2007, 24(4): 118–121

8. Candès E, Romberg J. Practical signal recovery from random projections. Proceedings of Society of Photographic Instrumentation Engineers, 2005, 5674: 76–86

9. Haupt J, Nowak R. Signal reconstruction from noisy random projections. IEEE Transactions on Information Theory, 2006, 52(9): 4036–4048

10. Candès E, Romberg J. Quantitative robust uncertainty principles and optimally sparse decompositions. Foundations of Computational Mathematics, 2006, 6(8): 227–254

11. Donoho D, Elad M, Temlyakov V. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Transactions on Information Theory, 2006, 52(1): 6–18

12. Tropp J, Gilbert A. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Transactions on Information Theory, 2007, 53(12): 4655–4666

13. Calderbank R, Howard S, Jafarpour S. Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property. IEEE Journal of Selected Topics in Signal Processing, 2010, 4(2):

358–374

14. Donoho D L, Elad M. Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization. Proceedings of National Academy of Sciences, 2003, 100(5): 2197–2202

15. Tropp J. Greed is good: algorithmic results for sparse approximation. IEEE Transactions on Information Theory, 2004, 50(10): 2231–2242

16. Candès E, Plan Y. Near-ideal model selection by l_1 minimization. Annals of Statistics, 2009, 37(5): 2145–2177

17. Rudelson M, Vershynin R. On sparse reconstruction from Fourier and Gaussian measurements. Communications on Pure and Applied Mathematics, 2008, 61(8): 1025–1045

18. Duarte M, Eldar Y. Structured compressed sensing: From theory to applications. IEEE Transactions on Signal Processing, 2011, 59(9): 4053–4085

19. Candès E, Plan Y. A probabilistic and RIPless theory of com-pressed sensing. IEEE Transaction on Information Theory, 2011, 57(11): 7235–7254

20. Candès E, Romberg J. Sparsity and incoherence in compressive sampling. Inverse Problems, 2007, 23(6): 969–985

21. Bajwa W U, Haupt J D, Raz G M, Wright S J, Nowak R D. Toeplitz-structured compressed sensing matrices. In: Proceedings of IEEE/SP the 14th Workshop on Statistical Signal Processing. 2007, 8: 294–298

22. Haupt J, Bajwa W, Raz G, Nowak R. Toeplitz compressed sensing matrices with applications to sparse channel estimation. IEEE Transactions on Information Theory, 2010, 56(11): 5862–5875

23. Rauhut H. Circulant and Toeplitz matrices in compressed sensing. In: Proceedings of Signal Processing with Adaptive Sparse Structured Representations. 2009, 1–6

24. Tropp J, Laska J, Duarte M, Romberg J, Baraniuk R. Beyond Nyquist: Effcient sampling of sparse bandlimited signals. IEEE Transactions on Information Theory, 2010, 56(1): 520–544

25. Romberg J. Compressive sensing by random convolution. SIAM Journal on Imaging Scicences, 2009, 2(4): 1098–1128

26. Romberg J. Sensing by random convolution. In: Proceedings of the 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing. 2007, 137–140

27. Do T, Gan L, Nguyen N, Tran T. Fast and efficient compressive sensing using structurally random matrices. IEEE Transactions on Signal Processing, 2012, 60(1): 139–154

28. Do T, Tran T and Gan L. Fast compressive sampling with structurally random matrices. In: Proceedings of IEEE International Conference on Acoustics Speech, Signal Process. 2008, 3369–3372

29. Rauhut H, Romberg J, Tropp J. Restricted isometries for partial random circulate matrices. Applied and Computational Harmonic Analysis, 2012, 32(2): 242–254

30. Zepernick H J, Finger A. Pseudo Random Signal Processing: Theory and Application. West Sussex: Wiley, 2005

31. Fan P Z, Darnell M. The synthesis of perfect sequences. Lec-ture notes in Computer Science, 1995, 1025: 63–73

32. Applebaum L, Howard S, Searle S, Calderbank R. Chirp sensing codes: deterministic compressed sensing measurements for fast recovery. Applied and Computational Harmonic Analysis, 2009, 26(2): 283–290

33. Gan L, Li K, Ling C. Golay meets hadamard: golay-paired hadamard matrices for fast compressed sensing. In: Proceedings of Information Theory Workshop. 2012, 637–641

34. Li K, Gan L, Ling C. Convolutional compressed sensing using de-

terministic sequences. IEEE Transaction on Signal Processing, 2013, 61(3): 740–752

35. Gan L, Li K, Ling C. Novel Toeplitz sensing matrices for compressing radar imaging. In: Proceedings of International Workshop on Compressed Sensing Applied to Radar Imaging. 2012, 1–6

36. Gan L. Block compressed sensing of natural images. In: Proceedings of International Conference on Digital Signal Processing. 2007, 403–406

37. Sebert F, Zou Y M, Ying L. Toeplitz block matrices in compressed sensing and their applications in imaging. In: Proceedings of International Conference on Information Technology and Applications. 2008, 47–50

38. Gan L, Do T, Tran T. Fast compressive imaging using scram-bled block Hadamard ensemble. In: Proceedings of European Signal Processing Conference. 2008, 1–6

39. Devore R A. Deterministic construction of compressed sensing matrices. Journal of Complexity, 2007, 23: 918–925

40. Saligrama V. Deterministic designs with deterministic guarantees: Toeplitz compressed sensing matrices, sequence design and system identfication. IEEE Transactions on Information Theory, 2012, 99

41. Iwen M A. Simple deterministically constructible RIP matrices with sub-linear Fourier sampling requirements. In: Proceedings of Conference in Information Sciences and Systems. 2008, 870–875

42. Monajemi H, Jafarpour S, Gavish M, D. Donoho L, Ambikasaran S, Bacallado S, Bharadia D, Chen Y, Choi Y, Chowdhury M. Deterministic matrices matching the compressed sensing phase transitions of Gaussian random matrices. In: Proceedings of the National Academy of Sciences, 2013, 110(4): 1181–1186

43. Howard S, Calderbank A, and Searle S, A fast reconstruction algorithm for deterministic compressive sensing using second order reed muller codes. In: Proceedings of the 42nd Annual Conference on Information Sciences and Systems. 2008, 11–15

44. Ailon N, Liberty E. Fast dimension reduction using rademacher series on dual BCH codes. In: Proceedings of Annual ACM-SIAM Symposium on Discrete Algorithms. 2008, 1–6

45. Calderbank R, Howard S, Jafarpour S. A sublinear algorithm for sparse reconstruction with l_2/l_2 recovery guarantees. In: Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing. 2009, 209–212

46. Yu N Y. Deterministic compressed sensing matrices from multiplicative character sequences. In: Proceedings of the 45th Annual Conference on Information Sciences and Systems. 2011, 1–5

47. Alltop W. Complex sequences with low periodic correlations. IEEE Transaction on Information Theory, 1980, 26(3): 350–354

48. Strohmer T, Heath R. Grassmanian frames with applications to coding and communication. Applied and Computational Harmonic Analysis, 2003, 14: 257–275

49. Chen W, Rodrigues M, Wassell I. Projection design for statistical compressive sensing: a tight frame based approach. IEEE Transactions on Signal Processing, 2013, 61(8): 2016–2029

50. Rauhut H. Compressive sensing and structured random matrices. In: Radon Series Computational and Applied Mathematics, Theoretical Foundations and Numerical Methods for Sparse Recovery. New York: DeGruyter, 2010, 9: 1–92

51. Compressive sensing resources. http://www-dsp.rice.edu/cs

52. Mishali M and Eldar Y, From theory to practice: sub-nyquist sampling of sparse wideband analog signals. IEEE Journal of Selected Topics in Signal Processing, 2010, 4(2): 375–391

53. Li, K, Ling C, Gan L. Deterministic compressed sensing matrices: where Toeplitz meets Golay. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. 2011, 3748–3751

54. Li K, Gao S, Ling C, Gan L. Wynerziv coding for distributed compressive sensing. In: Proceedings of Workshop of Signal Processing with Adaptive Sparse Structured Representations, 2011, 1–6

55. Lustig M, Donoho D, Santos J, Pauly J. Compressed sensing MRI. IEEE Signal Processing Magazine, 2008, 25(2): 72–82

56. Gorodnitsky I F, George J, Rao B. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. Electrocephalography and Clinical Neurophysiology, 1995, 95: 231–251

57. Lustig M, Donoho D, Pauly J M. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine, 2007, 58(6): 1182–1195

58. Duarte M F, Davenport M A, Takhar D, Laska J N, Sun T, Kelly K F, Baraniuk R G. Single pixel imaging via compressive sampling. IEEE Signal Processing Magazine, 2008, 83–91

59. Sampsell J. An overview of the digital micromirror device (DMD) and its application to projection displays. In: Proceedings of International Symposium Digest Technical Papers, 1993, 24: 1012–1015

60. Takhar D, Laska J, Wakin M, Duarte M, Baron D, Sarvotham S, Kelly K, Baraniuk R. A new compressive imaging camera architecture using optical-domain compression. Computational Imaging IV , 2006, 606: 43–52

61. Chan W L, Charan K, Takhar D, Kelly K F, Baraniuk R G, Mittleman D M. A single-pixel terahertz imaging system based on compressive sensing. Applied Physics Letters, 2008, 93(12)

62. Shen H, Newman N, Gan L, Zhong S, Huang Y, Shen Y. Com-pressed terahertz imaging system using a spin disk. In: Proceedings of the 35th International Conference on Infrared, Millimeter and Terahertz Waves. 2010, 1–2

63. Gilbert A, Indyk P. Sparse recovery using sparse matrices. Proceedings of the IEEE, 2010, 98(6): 937–947

64. Xu W, Hassibi B. Efficient compressive sensing with deterministic guarantees using expander graphs. In: Proceeding of IEEE Information Theory Workshop. 2007, 414–419

65. Krzakala F, Mezard M, Sausset F, Sun Y F, Zdeborova L. Statistical-physics-based reconstruction in compressed sensing. Physical Review X2, 021005, 2012
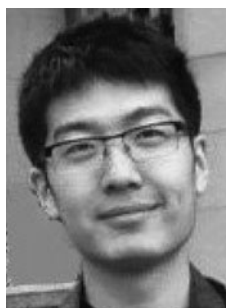
Kezhi Li is now a research assistant in the Automatic Complex Communication Networks, Signal and Systems Centre, School of Electrical Engineering at the Royal Institute of Technology (KTH), Sweden. He received his PhD in the Imperial College London in 2013. His research interests include signal processing, compressed sensing and its applications in communication, system identification and tomography.

Shuang Cong is a professor in the Department of Automation at the University of Science and Technology of China. She received her PhD in system engineering from the University of Rome "La Sapienza", Rome, Italy, in 1995. Her research interests include advanced control strategies for motion control, fuzzy logic control, neural networks design and applications, robotic coordination control, and quantum system control.