Semi-Universal Convolutional Compressed Sensing Using (Nearly) Perfect Sequences

Kezhi Li, *Student Member, IEEE,* Lu Gan, *Member, IEEE,* and Cong Ling, *Member, IEEE*

***Abstract*—In this paper, a new class of circulant matrices for compressed sensing are proposed by introducing nearly perfect sequences into the random convolution framework. Measurements are obtained by circularly convolving the signal not with a random sequence, but with a well-designed deterministic sequence, followed by random subsampling. Both uniform recovery and non-uniform recovery of sparse signals are investigated, based on the coherence parameter of the proposed sensing matrices. Many examples of the sequences are investigated, particularly the Frank-Zadoff-Chu (FZC) sequence and the Golay sequence. A salient feature of the proposed sensing matrices is their semi-universality, i.e., they can not only handle sparse signals in the time domain, but also those in the frequency or discrete-cosine transform (DCT) domain.**

***Index Terms*—Compressed sensing, Frank-Zadoff-Chu sequence, Golay sequence, nearly perfect sequences, random convolution.**

I. Introduction

COMPRESSED sensing (CS) is a growing theory in signal processing aiming at simultaneous sampling and compression of a signal [1], [2]. Consider a length-N signal $\mathbf{x}$ and suppose that the basis $\mathbf{\Psi}$ provides a K sparse representation of $\mathbf{x}$. That is, $\mathbf{x} = \mathbf{\Psi f}$, where $\mathbf{f}$ can be approximated using only $K \ll N$ non-zero entries and $\mathbf{\Psi}$ is referred to as the sparsifying transform. Throughout this paper, we assume that $\mathbf{\Psi}$ is an $N \times N$ normalized unitary matrix satisfying $\mathbf{\Psi}^*\mathbf{\Psi} = \mathbf{I}_N$. The data acquisition process in CS can be described as

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{e} = \mathbf{\Phi \Psi f} + \mathbf{e}, \tag{1}$$

where $\mathbf{y}$ represents an $M \times 1$ sampled vector, $\mathbf{\Phi}$ is an $M \times N$ measurement/sensing matrix and $\mathbf{e}$ is a noise vector. It was shown in [1], [2] that if $\mathbf{\Phi}$ is a Gaussian or Bernoulli random operator, $\mathbf{x}$ can be faithfully recovered from $\mathbf{y}$ using nonlinear optimization provided that $M \geq \mathcal{O}(K \log(N/K))$.

Although Gaussian or Bernoulli operators offer optimal theoretical bounds, they require huge memory for storage and high computational cost for signal reconstruction. Besides, fully random matrices are often difficult or expensive in hardware implementation. Taking these facts into account, many researchers have investigated structured random or deterministic operators [3]–[10]. Among them, one class of fast CS sampling operators is through convolving the signal of interest with a random filter and then subsampling [4], [5], [8], [9], [11]. Such operators are memory efficient, fast computable and hardware friendly in implementation [11]. They hold great potential in applications such as sparse channel estimation [4], Fourier optics [8], Radar imaging [8], [11] and coded aperture imaging [12].

Note that for convolution-based CS, most existing works focus on filters with independent and identically distributed (i.i.d.) random coefficients. In this paper, we propose a new framework by convolution with a *deterministic* filter followed by random sampling of the outputs. A deterministic filter has the advantage of being more convenient to implement. The filter coefficients are obtained by taking the inverse fast Fourier transform (IFFT) of a unimodular (nearly) perfect sequence (e.g., the Frank-Zadoff-Chu (FZC) [13] and the Golay sequences [14]) or its extended version (a precise definition will be given in Section III). We show that such a scheme is *semi-universal*, i.e., it can efficiently sample a sparse signal in the time ($\mathbf{\Psi} = \mathbf{I}_N$) or spectral ($\mathbf{\Psi} = \frac{1}{\sqrt{N}}\text{IFFT}$) domain. Specifically, for *all* K-sparse signals of length N in the time or spectral domain, robust reconstruction can be achieved when the number of measurements satisfies $M \geq \mathcal{O}\left(K \log^4 N\right)$, while for *any* given K-sparse signal, it can be recovered from only $M \geq \mathcal{O}(K \log N)$ measurements. To our best knowledge, these bounds offer the strongest theoretical guarantee for convolution-based CS. In addition, when the filter is constructed from the FZC sequence, these results also hold for sparse signals in the DCT domain.

The rest of this paper is organized as follows. Section II gives a brief review of CS theory and in particular, random convolution-based CS. Section III introduces the framework of the proposed system and summarizes the main results of this paper. Sections IV and V are devoted to the analysis of complex and real coefficient sensing matrices, respectively. Simulation results are given in Section VI, followed by conclusions in Section VII.

Notation: In this paper, bold letters are used to denote a vector or a matrix. For an $M \times N$ matrix $\mathbf{A}$, $\mathbf{A}(p, q)$ ($0 \leq p \leq M-1$, $0 \leq q \leq N-1$) represents the element on its p-th row and q-th column. $\mathbf{A}^T$ and $\mathbf{A}^*$ denote the transpose and Hermitian transpose of $\mathbf{A}$, respectively. For an $N \times N$ given matrix $\mathbf{A}$, we use $\mu(\mathbf{A})$ to denote its coherence parameter, i.e., the maximum magnitude,

$$\mu(\mathbf{A}) = \max_{0 \leq p,q \leq N-1} |\mathbf{A}(p, q)|.$$

The material in this paper was partially presented at IEEE Information Theory Workshop (ITW'09), Oct. 2009, Taormina, Italy, and at International Conference on Acoustics, Speech and Signal Processing (ICASSP'11), May 2011, Prague, Czech. This work was supported in part by the UK MOD University Defence Research Centre (UDRC) in Signal Processing and by the UK EPSRC under Grant EP/I038853/1.

K. Li and C. Ling are with the Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK (e-mail: k.li08@imperial.ac.uk, cling@ieee.org).

L. Gan is with the School of Engineering and Design, Brunel University, London UB8 3PH, UK (e-mail: lu.gan@brunel.ac.uk).

For two $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$, their mutual coherence $\mu(\mathbf{A}, \mathbf{B})$ is defined as

$$\mu(\mathbf{A}, \mathbf{B}) = \mu(\mathbf{A}\mathbf{B}) = \max_{0 \leq p,q \leq N-1} |\mathbf{A}(p,:)\mathbf{B}(:,q)|,$$

where $\mathbf{A}(p,:)$ and $\mathbf{B}(:,q)$ correspond to the p-th row of $\mathbf{A}$ and q-th column of $\mathbf{B}$, respectively. $\mathbf{F}$ is the $N \times N$ fast Fourier transform (FFT) matrix where $\mathbf{F}(p,q) = e^{-\frac{2\pi jpq}{N}}$ $(0 \leq p, q \leq N-1)$. If a is a function of N, $a \doteq b$ means $\lim_{N \to \infty} a = b$. We use the standard asymptotic notation $f(x) = \mathcal{O}(g(x))$ when $\limsup_{x \to \infty} |f(x)/g(x)| < \infty$.

II. Review of Compressed Sensing

In this section, we first review uniform and non-uniform recovery in compressed sensing, and in particular theoretical performance bounds for randomly subsampled unitary matrices. We then highlight existing works of random convolution-based CS.

A. Uniform vs. Non-Uniform Recovery

Let $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$. Then (1) can be re-written as

$$\mathbf{y} = \mathbf{\Theta}\mathbf{f} + \mathbf{e}. \tag{2}$$

Hence, reconstruction of $\mathbf{x} = \mathbf{\Psi}\mathbf{f}$ is equivalent to recovery of a K-sparse vector $\mathbf{f}$. Note that as $M < N$, eq. (2) is in general under-determined. To recover $\mathbf{f}$ (or equivalently, $\mathbf{x}$) from $\mathbf{y}$, nonlinear optimization is required. In the noiseless case (i.e, $\mathbf{e} = \mathbf{0}$, exact recovery can be achieved by a standard l_1 minimization program [15]

$$\min \|\mathbf{f}\|_{l_1} \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Theta}\mathbf{f}. \tag{3}$$

In the noisy case, $\mathbf{f}$ can be reconstructed using the unconstrained LASSO [16] that solves the l_1 regularized square problem

$$\min \lambda \|\mathbf{f}\|_{l_1} + \frac{1}{2}\|\mathbf{y} - \mathbf{\Theta}\mathbf{f}\|^2, \tag{4}$$

where λ is the Lagrangian constant. In addition to l_1-based algorithms, many greedy algorithms, such as orthogonal matching pursuit [17], subspace pursuit [18], CoSaMP [19] and their variants have been proposed for sparse signal reconstruction. These algorithms require lower computational complexity than l_1-based optimization with somewhat weaker theoretical guarantees.

In CS reconstruction, *uniform recovery* [2] 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. To achieve uniform recovery, many existing sparse recovery algorithms such as l_1 optimization [15], LASSO [16], subspace pursuit [18] and CoSAMP [19] require that the restricted isometry property (RIP) is satisfied.

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

$$(1-\delta)\|\mathbf{f}\|^2 \leq \|\mathbf{\Theta}\mathbf{f}\|^2 \leq (1+\delta)\|\mathbf{f}\|^2, \text{ for all } \mathbf{f} \in \Gamma, \tag{5}$$

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

For RIP constant δ required in different sparse recovery algorithms, please refer to [7] for details. Note that the RIP is a very strong restriction. Among existing sampling operators, it is known that the i.i.d. Gaussian and Bernoulli matrices satisfy the RIP when $M \geq \mathcal{O}(\delta^{-2} K \log N)$. However, as we have mentioned before, these full random operators are impractical for large-scale CS data acqusition. Another subclass of operators satisfying the RIP is random sampled unitary matrix, as summarized in the following theorem [7], [20].

Theorem 1 (RIP for randomly subsampled unitary matrix): Suppose that the $M \times N$ matrix $\mathbf{\Theta}$ is a randomly subsampled unitary matrix, i.e., it can be written as $\mathbf{\Theta} = \sqrt{\frac{N}{M}}\mathbf{R}_\Omega \mathbf{U}$, where $\sqrt{N/M}$ is a normalizing coefficient, $\mathbf{R}_\Omega$ is a random sampling opertor 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 [7], [20]

$$M \geq \mathcal{O}\left(\delta^{-2}\mu^2(\mathbf{U}) K \log^4 N\right). \tag{6}$$

The above theorem 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 $\mathbf{U}^*\mathbf{U} = N\mathbf{I}_N$, we have $1 \leq \mu(\mathbf{U}) \leq \sqrt{N}$. In case when $\mathbf{U}$ is chosen as the FFT or the Walsh-Hadamard transform, $\mu(\mathbf{U}) = 1$ and by Eq. (6), we have

$$M \geq \mathcal{O}\left(\delta^{-2} K \log^4 N\right). \tag{7}$$

One can also observe that compared with the optimal bound provided by full random matrix, there is an extra $\log^3 N$ factor in (7). To address this issue, several researchers have relaxed the conditions and investigated the case of non-uniform recovery, where one just needs to reconstruct a given sparse signal. Theorem 2 presents the results for non-uniform recovery of a randomly subsmapled unitary matrix using l_1-based optimization.

Theorem 2 (Non-uniform recovery): Assume that $\mathbf{\Theta}$ is a randomly subsampled unitary matrix that follows the same definition as in Theorem 1. Let $\mathbf{f}$ in (2) be a fixed arbitary K-sparse signal. Then $\mathbf{f}$ can be faithfully recovered from $\mathbf{y}$ using l_1-based optimization (i.e., (3) in the noiseless case and (4) in the noisy case) if M satisfies [21]

$$M \geq \mathcal{O}(\mu^2(\mathbf{U}) K \log N). \tag{8}$$

Theorem 2 implies that using randomly subsampled FFT or Walsh-Hadamard transform (WHT), the number of samples required for non-uniform reconstruction is nearly optimal. This is because that non-uniform recovery is much weaker than uniform recovery. It should also be noted that the above non-uniform recovery only holds for l_1 optimization. It is still unknown whether we can get similar bounds for fast greedy recovery algorithms such as the subspace pursuit [18] and CoSaMP [19].

At this point, it is worth pointing out that although partial FFT (or WHT) has near-optimal theoretical guarantee, easy hardware implementation and fast-computable recovery, its major shortcoming is the lack of *universality*. A universal sensing matrix means that it can handle signals that are sparse on any domain. If $\mathbf{\Phi}$ is a Gaussian random matrix, the matrix $\mathbf{\Phi\Psi}$ 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 $\mu(\mathbf{F\Psi})$ can not be $\mathcal{O}(1)$ for all bases $\mathbf{\Psi}$. In this paper, we will propose a new class of randomly subsampled circulant matrices that are *semi-universal*, which can be used to efficiently sample sparse signals on both time and frequency domains.

B. Random Convolution for CS

Tropp *et al.* [5], [22] first proposed the idea of CS using convolution with an i.i.d. sequence followed by fixed regular sampling. The effectiveness of such an approach has been demonstrated through numerical simulations. Later, many people have investigated cyclic convolution with an N-tap random filter [4], [8], [9], in which the sampling operator $\mathbf{\Phi}$ can be represented as a partial circulant matrix with the following form

$$\mathbf{\Phi} = \sqrt{\frac{N}{M}}\mathbf{R}_\Omega \mathbf{A} \tag{9}$$

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

$$\mathbf{A} = \begin{bmatrix} a_0 & a_{N-1} & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & a_{N-2} & \cdots & a_0 \end{bmatrix}. \tag{10}$$

For $\mathbf{\Phi}$ given in (9), the measurement process can be realized by circularly convolving $\mathbf{x}$ with a filter $\mathbf{a} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{N-1}\end{bmatrix}^T$ and then downsample the output at locations indexed by Ω. It is also well known that a circulant matrix $\mathbf{A}$ can be diagonalized using FFT as follows

$$\mathbf{A} = \frac{1}{\sqrt{N}}\mathbf{F}^*\Sigma\mathbf{F}, \tag{11}$$

where $\Sigma = \text{diag}(\sigma) = \text{diag}(\sigma_0, \sigma_1, \cdots, \sigma_{N-1})$. Eq. (11) suggests that a circulant operator is fast computable. 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 IFFT[1] of sequence $\sigma = \begin{bmatrix}\sigma_0 & \sigma_1 & \cdots & \sigma_{N-1}\end{bmatrix}^T$, i.e.,

$$\mathbf{a} = \frac{1}{\sqrt{N}}\mathbf{F}^*\sigma. \tag{12}$$

In other words, σ is the normalized FFT of $\mathbf{a}$. It is clear that when σ is a unimodular sequence, i.e., $|\sigma_k| = 1$ ($0 \le k \le N-1$), $\mathbf{A}$ is a unitary matrix satisfying $\mathbf{A}^*\mathbf{A} = N\mathbf{I}_N$.

In most existing works, the coefficient vector $\mathbf{a}$ is constructed randomly. In [4], [9], $\mathbf{a}$ is a binary random sequence, where each a_i takes the values of $+1$ and -1 with equal probability. An alternative way is to obtain $\mathbf{a}$ from σ, which is a binary random sequence [9] or a unimodular sequence with random phases [8], i.e., $\sigma_k = e^{j\theta_k}$, where θ_k is a random variable that is uniformly distributed in $[0, 2\pi)$.

The sampling operator $\mathbf{R}_\Omega$ can be either deterministic or random, as summarized below.

1) Deterministic subsampling: In deterministic sampling, Ω is chosen as any arbitrary subset of $\{1, 2, \cdots, N\}$ with cardinality $|\Omega| = M$. It was shown in [9] that $\mathbf{\Phi}$ given by (9) satisfies RIP with parameters (K, δ) provided that $M \ge \mathcal{O}\left((K\log N)^{\frac{3}{2}}\right)$. Note that there is an extra term $\sqrt{K\log N}$ factor in this bound. On the other hand, non-uniform recovery results have been investigated in [23], where the author considered the recovery of a given K-sparse signal whose nonzero components have random signs. Under this model, it was established in [23] that exact recovery can be achieved using l_1 optimization when $M \ge \mathcal{O}(K\log^2 N)$. However, unlike Theorem 2, this bound only holds for noiseless measurement and hence the guarantee for stable recovery under noisy measurements is unclear. Besides, as we will show later, when $\mathbf{R}_\Omega$ is a deterministic operator, $\mathbf{\Phi}$ given in (9) works poorly for a spectrally sparse signal, which implies that it cannot be used directly to sample a natural image (which is often sparse in the DCT or the wavelet domain).

2) Random subsampling: To achieve a universal convolution-based CS, Romberg [8] proposed to use a random sampling operator $\mathbf{R}_\Omega$. Under such a setting, it can be shown that for any orthonormal basis $\mathbf{\Psi}$, $\mathbf{A}$ given by (11) satisfies

$$\mu(\mathbf{A\Psi}) = \mathcal{O}(\sqrt{\log N}). \tag{13}$$

By Theorem 1, such a universal operator satisfies the RIP when $M \ge \mathcal{O}(\delta^{-2}K\log^5 N)$ and by Theorem 2, $M \ge \mathcal{O}(\delta^{-2}K\log^2 N)$ measurements are required for non-uniform recovery. Note that compared with the optimal bounds offered by a randomly subsampled unitary matrix, there is an extra $\log N$ factor in random convolution. It is thus natural to ponder: Can we get better bounds for convolution-based CS systems?

III. Deterministic filter for convolutional CS

In this Section, we propose a new framework which answers the afore-posed question affirmatively.

A. Our Proposal

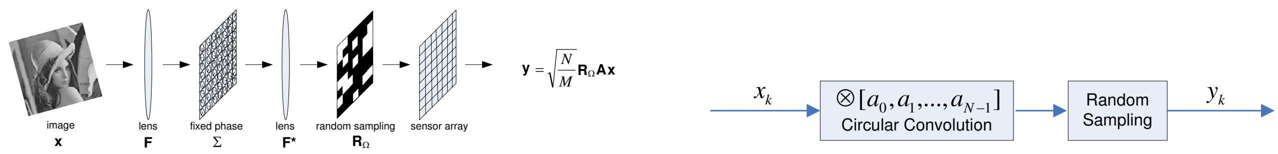
Unlike previous work, in this paper, we propose the use of a *deterministic filter* followed by random sampling for convolution-based CS. Specifically, for $\mathbf{\Phi}$ given in (9), $\mathbf{R}_\Omega$ is a random sampling operator and $\mathbf{A}$ is constructed from (11) by choosing σ as a deterministic unimodular sequence. As mentioned before, under this constraint, $\mathbf{A}$ is a unitary matrix satisfying $\mathbf{A}^*\mathbf{A} = N\mathbf{I}_N$. When $\mathbf{\Psi} = \mathbf{I}_N$, by Theorem 1 and Theorem 2, the performance bounds depend on $\mu(\mathbf{A})$, which is given by

$$\mu(\mathbf{A}) = \max(|a_0|, |a_1|, \cdots, |a_N|).$$

Hence, we need to find a unimodular sequence σ so that the maximum magnitude of $\mathbf{a} = \frac{1}{\sqrt{N}}\mathbf{F}^*\sigma$ is minimized. To this end, we consider the construction of σ using a perfect or nearly perfect sequence [24], [25], whose definition is given below:

Definition 2 (Perfect and nearly perfect sequences): A sequence $\mathbf{s}$ of period N is called a perfect sequence if it has

[1] For convenience, the definition here differs from the standard one IFFT $= \frac{1}{N}\mathbf{F}^*$ by a factor of $1/\sqrt{N}$.



(a) (b)

Fig. 1. Implementation schemes of convolutional CS: (a) Implementation in frequency domain: Fourier optics imaging. The lenses transform the signal to Fourier domain and back. (b) Implementation in time domain: convolution with a deterministic filter followed by random sampling.

perfect periodic autocorrelation

$$R_{\mathbf{s}}(l) = \sum_{k=0}^{N-1} s_k \cdot s^*_{\text{mod }(k+l,N)} = \begin{cases} N & l = iN \\ 0 & l \neq iN \end{cases} \tag{14}$$

where $l = 0, 1, 2, \cdots$ is an integer. A nearly perfect sequence is a sequence with the off-peak autocorrelation magnitude bounded by a small ϵ, i.e.,

$$|R_{\mathbf{s}}(l)| < \epsilon, \quad l \neq iN. \tag{15}$$

Note that when σ is a unimodular sequence, $\mathbf{\Phi}$ in general is complex valued. Yet real-valued filters are desired in many applications, such as image processing. To generate real sensing matrices, we need conjugate symmetry of the diagonal sequence σ. More precisely, we define an extended sequence

$$\sigma_k = \begin{cases} \pm 1 & k = 0 \\ e^{-j\theta_k} & 1 \leq k \leq N/2 - 1 \\ \pm 1 & k = N/2 \\ e^{j\theta_{N-k}} & N/2 + 1 \leq k \leq N - 1; \end{cases} \tag{16}$$

when N is even, and

$$\sigma_k = \begin{cases} \pm 1 & k = 0 \\ e^{-j\theta_k} & 1 \leq k \leq (N-1)/2 \\ e^{j\theta_{N-k}} & (N+1)/2 \leq k \leq N - 1, \end{cases} \tag{17}$$

when N is odd, for some phases θ_k. Thus, to get a real-valued filter, we just need to use a length-N_0 unimodular sequence $\mathbf{s} = [s_0, s_1, \cdots, s_{N_0-1}]$, where

$$N_0 = \begin{cases} \frac{N}{2}, & N \text{ even}, \\ \frac{N-1}{2}, & N \text{ odd}; \end{cases} \tag{18}$$

then σ can be obtained by extending $\mathbf{s}$ using (16) and (17) for even N and odd N, respectively.

The scheme of convolutional CS is illustrated in Fig. 1. In particular, Fig. 1(a) shows the frequency domain implementation: the FFT of the signal is multiplied with a deterministic diagonal matrix $\mathbf{\Sigma}$, fed to IFFT (up to a scaling factor $N^{-1/2}$), then randomly subsampled. It is similar to the imaging architecture in [8] but with fixed coefficients. Fig. 1(b) shows the alternative implementation in the time domain: the signal passes through a deterministic filter (i.e., convolved with the IFFT of sequence σ), then randomly subsampled. By using a deterministic construction of σ (or equivalently, filter coefficient vector $\mathbf{a}$), not only can we simplify the design and implementation, it also offers better theoretical guarantee, as explained in the next subsection.

B. Main Results

Lemma 1 (Bound on the coherence parameter): Let the complex-valued matrix $\mathbf{A}$ be defined by (11) where $\sigma = \mathbf{s}$. If $\mathbf{s}$ is a unimodular perfect sequence, then $\mu(\mathbf{A}) = 1$. If $\mathbf{s}$ is a unimodular nearly perfect sequence satisfying (15), then

$$\mu(\mathbf{A}) \leq \sqrt{1 + \epsilon}. \tag{19}$$

Proof: First, we examine the FFT $\hat{\mathbf{s}} = \mathbf{F}\mathbf{s}$ of sequence $\mathbf{s}$. By the Wiener-Khinchin theorem, the power spectrum $|\hat{\mathbf{s}}|^2$ is given by the FFT of the periodic autocorrelation function $R_{\mathbf{s}}$. Thus, we have (for $0 \leq k \leq N-1$)

$$\begin{aligned} |\hat{s}_k|^2 &= \sum_{l=0}^{N-1} R_{\mathbf{s}}(l) e^{-\frac{j2\pi kl}{N}} \\ &\leq N + \left| \sum_{l=1}^{N-1} R_{\mathbf{s}}(l) e^{-\frac{j2\pi kl}{N}} \right| \\ &\leq N + (N-1)\epsilon. \end{aligned} \tag{20}$$

Now consider the sequence $\mathbf{a} = \frac{1}{\sqrt{N}}\mathbf{F}^*\mathbf{s}$. It is easy to show that $\mathbf{F}^*\mathbf{s}$ is a reversed version of $\hat{\mathbf{s}}$ (with respect to index k), hence the same magnitudes. From (20), the coherence parameter, i.e., the peak magnitude of $\mathbf{a}$, is bounded by

$$\mu(\mathbf{A}) \leq \frac{1}{\sqrt{N}}\sqrt{N + (N-1)\epsilon} \leq \sqrt{1 + \epsilon}.$$

When $\epsilon = 0$, we get the ideal bound $\mu(\mathbf{A}) = 1$ for perfect sequences. ■

Lemma 1 forms the motivation of the method to be developed in this paper. It shows that $\mu(\mathbf{A})$ will be small for perfect or nearly-perfect sequences. In particular, we require $\mu(\mathbf{A}) = \mathcal{O}(1)$ in this paper, which means that ϵ is bounded by a constant. The only known binary perfect sequence is $[1, 1, -1, 1]$. So in general we have to use polyphase perfect sequences. We refer to [26] for a unified construction of polyphase perfect sequences. A survey of binary and quadriphase nearly-perfect sequences is given in [27]. Examples of binary sequences with $\epsilon = 1$ are m-sequences, Legendre sequences, and twin-prime sequences. An example of quadriphase sequences with $\epsilon = 1$ is the complementary-based sequence. Sequences with other values of ϵ such as 2, 3, and 4, can be found in [27].

Yet, Lemma 1 has some limitations: firstly, it is difficult, if not impossible to extend to real sensing matrices; secondly, the bound (19) is pessimistic; thirdly, an extension to other domains (e.g., DCT) seems difficult. In the following, we derive, in a case-by-case manner, the bound on $\mu(\mathbf{A})$ for both complex and real sensing matrices, which is often better than (19).

TABLE I
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 chirp	Even N	$4+\frac{4}{\sqrt{N}}$
		Odd N	$2.69+\frac{8.15}{\sqrt{N}}$
	Extended Golay	Even N, $N = 2^{\kappa_1}10^{\kappa_2}26^{\kappa_3}, \kappa_1, \kappa_2, \kappa_3 \in \mathbb{N}$	$2+\frac{2}{\sqrt{N}}$
		Odd N, $N = 2^{\kappa_1}10^{\kappa_2}26^{\kappa_3} \pm 1, \kappa_1, \kappa_2, \kappa_3 \in \mathbb{N}$	$2+\frac{1}{\sqrt{N}}$

Table I lists the different unimodular sequences σ used in this paper, along with the corresponding N and $\mu(\mathbf{A})$. Derivations of $\mu(\mathbf{A})$ are heavily based on Gaussian sums given in Appendix A. Details will be explained in Section IV and Section V for complex and real-coefficient filters, respectively. Based on Lemma 1, Table I, Theorem 1 and Theorem 2, we arrive at the following theorem:

Theorem 3: Consider a CS sampling operator $\mathbf{\Phi}$ given in (9), where $\mathbf{R}_\Omega$ is a random sampling operator and the unitary circulant matrix $\mathbf{A}$ is generated from (11), with $\sigma = [\sigma_0, \sigma_1, \cdots, \sigma_{N-1}]^T$, $|\sigma_k| = 1$, $k = 0, 1 \cdots N-1$, being a unimodular sequence as listed in Table I. Then, for all K-sparse signals in the time ($\mathbf{\Psi} = \mathbf{I}_N$) or spectral domain ($\mathbf{\Psi} = \frac{1}{\sqrt{N}}\mathbf{F}^*$), $M \geq \mathcal{O}(K \log^4 N)$ measurements are required for uniform recovery; for any given K-sparse signal in the time or spectral domain, $M \geq \mathcal{O}(K \log N)$ measurements are needed using l_1-based reconstruction.

Proof: As one can see from Table I, in our proposed systems,

$$\mu(\mathbf{A}) = \mathcal{O}(1). \tag{21}$$

Hence, it can be easily derived from Theorem 1 and Theorem 2 that Theorem 3 holds for time-domain K-sparse signals (i.e., $\mathbf{\Psi} = \mathbf{I}_N$). To see this is the case in frequency domain (i.e., $\mathbf{\Psi} = \frac{1}{\sqrt{N}}\mathbf{F}^*$), let us examine the coherence parameter $\mu(\mathbf{A\Psi})$. Note that

$$\mathbf{A\Psi} = \frac{1}{\sqrt{N}}\mathbf{F}^*\Sigma\mathbf{F}\frac{1}{\sqrt{N}}\mathbf{F}^* = \mathbf{F}^*\Sigma.$$

Obviously that the square matrix $\mathbf{F}^*\Sigma$ is unitary, and all the entries are unimodular, which implies an ideal coherence parameter $\mu\left(\frac{1}{\sqrt{N}}\mathbf{A}\mathbf{F}^*\right) = 1$. Therefore, Theorem 3 also holds for spectrally sparse signals. ∎

In addition to the above Theorem, we will show in Theorem 4 that if σ is chosen as the FZC sequence, similar bounds hold for sparse signals in the DCT domain. Details will be given in the next Section. It is well known that most natural images are sparse in the DCT domain. Thus, this represents a major advantage of the proposed scheme over the partial FFT sensing matrix, which does not work for the DCT domain, because the FFT and DCT matrices are mutually coherent [6].

C. Connections With Existing Work

The comparison between the proposed scheme and existing random convolution-based operators is shown in Table II. It can be seen that the proposed scheme offers the strongest theoretical performance guarantee for both uniform and non-uniform CS recovery, thanks to its deterministic construction of σ. Recall that for a random filter, the coherence parameter is bounded by $\mathcal{O}(\sqrt{N})$ [8], [9]. The $\mathcal{O}(1)$ coherence parameter associated with deterministic σ enables us to remove the extra $(\log N)$ factor in existing random convolution. Although the proposed scheme can not offer universality, it works for time and frequency domains (and the DCT domain for FZC sequences). This implies that the proposed scheme can be used as a hardware friendly, memory efficient and fast computable solution for large scale CS applications, e.g., hyperspectral imaging.

Chirp sequences were applied to radio interferometry in [28], where the sensing matrix was constructed in a different way, namely, it was the product of a rectangular binary matrix $\mathbf{M}$, Fourier matrix $\mathbf{F}$, diagonal matrix $\mathbf{C}$ implementing chirp modulation and diagonal matrix $\mathbf{D}$ implementing the primary beam. The coherence was analyzed when $\mathbf{\Psi}$ is formed by Gaussian waveforms. Chirp sequences were also used to construct deterministic sensing matrices in [29], which cannot be implemented through convolution. Besides, the sizes of the sampling operators in [29] are restricted to be $M \times M^2$. Upon completion of this work, we learned that perfect sequences (including the FZC sequence) were used as the entries of Toeplitz sensing matrices in radio spectrum estimation [30]. Yet the analysis of the coherence parameter or RIP were limited to the cases $\mathbf{\Psi} = \mathbf{I}_N$ and $\mathbf{\Psi} = \mathbf{F}$ in [30]. Taking a step forward, we generalize the analysis to the case of DCT in this paper. Besides, we extend it to the case of real-valued sensing matrices in Section V (the sensing matrices in [30] are complex-valued.).

IV. Unitary Matrix A

In this Section, we deal with unitary sensing matrices $\mathbf{A}$ ($\mathbf{A}^*\mathbf{A} = N\mathbf{I}_N$) which are generated from polyphase and binary sequences σ.

A. Polyphase Sequences

When the diagonal sequence σ is a unimodular perfect sequence whose elements have unit modulus $|\sigma_k| = 1$, we have the ideal coherence parameter $\mu(\mathbf{A}) = 1$. Here, we use a well known chirp-like sequence, the FZC sequence with

TABLE II
COMPARISON OF DIFFERENT CIRCULANT OPERATORS FOR COMPRESSED SENSING

Measurement Operator $\mathbf{\Phi}$	Random Convolution [8]	Partial Circulant Operator [9]	This work
Filter Coefficients	Random	Random	Deterministic
Sub-sampling Operator	Random	Deterministic	Random
Universality	Yes	No	Semi-Universal
Restricted Isometry Property	$M \geq \mathcal{O}(K \log^5 N)$	$M \geq \mathcal{O}((K \log N)^{\frac{3}{2}})$	$M \geq \mathcal{O}(K \log^4 N)$
Non-Uniform Recovery	$M \geq \mathcal{O}(K \log N)$	$M \geq \mathcal{O}(K \log^2 N)$	$M \geq \mathcal{O}(K \log N)$

perfect auto-correlation [13]. The mth sequence within the FZC family is given by [31][2]

$$s_k = \begin{cases} e^{-\frac{j\pi m k^2}{N}}, & \text{for even } N \\ e^{-\frac{j\pi m k(k+1)}{N}}, & \text{for odd } N \end{cases} \tag{22}$$

for $k = 0, 1, \cdots, N-1$.

In what follows, we show the matrix $\mathbf{A}$ generated from the FZC sequence can also recover the sparse signals in the DCT domain.

Theorem 4: Let IDCT represent the inverse DCT matrix, and let $\mathbf{A}$ be generated from the FZC sequence with $m = 1$. The matrix

$$\mathbf{U} = \mathbf{A} \cdot \text{IDCT} = N^{-1/2}\mathbf{F}^*\Sigma\mathbf{F} \cdot \text{IDCT} \tag{23}$$

has coherence parameter $\mu(\mathbf{U}) \leq 6\sqrt{2}$.

Proof: For clarity we only give the proof for even N here. The case of odd N is similar and is omitted. We apply a result from [31] that the Fourier dual of a unimodular perfect sequence yields another unimodular perfect sequence. Specifically, the elements of $\mathbf{A}$ are given by [31]

$$\begin{aligned} \mathbf{A}(p,q) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{j2\pi kp}{N} \cdot \frac{-j\pi k^2}{N} \cdot \frac{-j2\pi kq}{N}} \\ &= e^{\frac{j\pi(p-q)^2}{N} - \frac{j\pi}{4}}. \end{aligned} \tag{24}$$

Obviously, we may ignore the phase $e^{-\frac{j\pi}{4}}$ when calculating its magnitude. From the definition of IDCT, we have

$$\text{IDCT}(p,q) = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}(p+\frac{1}{2})q\right), & 1 \leq q \leq N-1. \end{cases} \tag{25}$$

Thus, when $q = 0$ and $0 \leq p \leq N-1$,

$$\begin{aligned} |\mathbf{U}(p,0)| &= \left| \sum_{k=0}^{N-1} \mathbf{A}(p,k) \cdot \frac{1}{\sqrt{N}} \right| \\ &= \frac{1}{\sqrt{N}} \left| \sum_{k=0}^{N-1} e^{\frac{j\pi}{N}(p-k)^2 - \frac{j\pi}{4}} \right| \\ &= \frac{1}{\sqrt{N}} \left| \sum_{k=0}^{N-1} e^{\frac{j\pi}{N}k^2} \right| \\ &= \frac{1}{\sqrt{N}} |G_N(N)| \leq \sqrt{2}. \end{aligned} \tag{26}$$

The last step is due to a property of the complete Gauss sum $G_N(N)$ given in Appendix A.

[2]This definition gives a sequence which is the complex conjugate of the standard one [13]. They obviously have the same autocorrelation magnitudes.

When $1 \leq q \leq N-1$ and $0 \leq p \leq N-1$,

$$\begin{aligned} \mathbf{U}(p,q) &= \sum_{k=0}^{N-1} \mathbf{A}(p,k) \cos\left(\frac{\pi}{N}(k+\frac{1}{2})q\right) \cdot \frac{\sqrt{2}}{\sqrt{N}} \\ &= \frac{\sqrt{2}}{2\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{j\pi}{N}(p-k)^2 - \frac{j\pi}{4}} \left(e^{-j\frac{\pi}{N}(k+\frac{1}{2})q} + e^{j\frac{\pi}{N}(k+\frac{1}{2})q}\right). \end{aligned} \tag{27}$$

Rewriting the two terms in the summand as,

$$\begin{aligned} e^{j\frac{\pi}{N}(p-k)^2 - j\frac{\pi}{4} - j\frac{\pi}{N}(k+\frac{1}{2})q} &= e^{j\frac{\pi}{N}\left[(k-p-\frac{1}{2}q)^2 - \frac{1}{4}q^2 - pq - \frac{1}{2}q - \frac{N}{4}\right]}, \\ e^{j\frac{\pi}{N}(p-k)^2 - j\frac{\pi}{4} + j\frac{\pi}{N}(k+\frac{1}{2})q} &= e^{j\frac{\pi}{N}\left[(k-p+\frac{1}{2}q)^2 - \frac{1}{4}q^2 + pq + \frac{1}{2}q - \frac{N}{4}\right]}, \end{aligned} \tag{28}$$

we obtain the bound

$$|\mathbf{U}(p,q)| \leq \frac{\sqrt{2}}{2\sqrt{N}} \left(\left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{N}(k-p-\frac{1}{2}q)^2} \right| + \left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{N}(k-p+\frac{1}{2}q)^2} \right| \right). \tag{29}$$

Now let us examine the first sum. When $q = 2q_0$ is an even number, let $p + q_0 = l, 0 \leq l \leq \frac{3N}{2} - 1$. If $0 \leq l \leq N-1$,

$$\begin{aligned} \left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{N}(k-l)^2} \right| &= \left| \sum_{k=-l}^{-1} e^{j\frac{\pi}{N}k^2} + \sum_{k=0}^{N-1-l} e^{j\frac{\pi}{N}k^2} \right| \\ &= \left| \sum_{k=0}^{l} e^{j\frac{\pi}{N}k^2} + \sum_{k=0}^{N-1-l} e^{j\frac{\pi}{N}k^2} - 1 \right| \\ &= |G_{2N}(l+1) + G_{2N}(N-l) - 1| \\ &\leq 2\sqrt{N} + 1 \end{aligned} \tag{30}$$

where the last step is due to a property of the incomplete Gauss sum $|G_{2N}(l)| \leq \sqrt{N}$ for $l \leq N$. If $N \leq l \leq \frac{3N}{2} - 1$,

$$\begin{aligned} \left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{N}(k-l)^2} \right| &\leq \left| \sum_{k=-l}^{N-1-l} e^{\frac{j\pi}{N}k^2} \right| = \left| \sum_{k=-l+N}^{2N-1-l} e^{\frac{j\pi}{N}k^2} \right| \\ &= |G_{2N}(2N-l) + G_{2N}(l-N+1) - 1| \\ &\leq 2\sqrt{N} + 1 \end{aligned} \tag{31}$$

where again the property $|G_{2N}(l)| \leq \sqrt{N}$ for $l \leq N$ is applied.

For the second sum, the calculation is similar and the bound is exactly the same. So for even q, $0 \leq p, q \leq N-1$, the bound of $\mathbf{U}$ will be

$$\begin{aligned} |\mathbf{U}(p,q)| &\leq \sqrt{\frac{2}{N}} \left(2\sqrt{N} + 1\right) \\ &= 2\sqrt{2} + \frac{\sqrt{2}}{\sqrt{N}}. \end{aligned} \tag{32}$$

When $q = 2q_0 + 1$ is an odd number, denote by

$$Q_N(m) = \sum_{k=0}^{m-1} e^{j\frac{\pi}{N}(k+\frac{1}{2})^2} = \sum_{k=0}^{m-1} e^{j\frac{\pi}{4N}(2k+1)^2}, 0 \le m \le N \tag{33}$$

the modified Gauss sum. Again, let $l = p + q_0$. After some tedious calculation, we may break up (29) into

$$\begin{aligned} |\mathbf{U}(p,q)| =& \frac{\sqrt{2}}{2\sqrt{N}} \left(\left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{4N}(2k-2p-2q_0-1)^2} \right| \right. \\ & \left. + \left| \sum_{k=0}^{N-1} e^{j\frac{\pi}{4N}(2k-2p+2q_0+1)^2} \right| \right) \\ \le & \frac{\sqrt{2}}{2\sqrt{N}} \left(|Q_N(\mathrm{mod}(p+q_0,N)+1)| \right. \\ & + |Q_N(N-1-\mathrm{mod}(p+q_0,N))| \\ & + |Q_N(|p-q_0|)| \\ & \left. + |Q_N(N-|p-q_0|)| \right). \end{aligned} \tag{34}$$

Meanwhile, since the Gauss sum $G_{8N}(2m)$ can be written as

$$\begin{aligned} G_{8N}(2m) &= \sum_{k=0}^{2m-1} e^{\frac{2\pi j}{8N}k^2} \\ &= \sum_{k \text{ odd}} e^{\frac{2\pi j}{8N}k^2} + \sum_{k \text{ even}} e^{\frac{2\pi j}{8N}k^2} \\ &= \sum_{v=0}^{m-1} e^{\frac{2\pi j}{8N}(2v+1)^2} + \sum_{v=0}^{m-1} e^{\frac{2\pi j}{8N}(2v)^2} \\ &= \sum_{v=0}^{m-1} e^{\frac{\pi j}{4N}(2v+1)^2} + \sum_{v=0}^{m-1} e^{\frac{\pi j}{N}(v)^2} \\ &= Q_N(m) + G_{2N}(m), \end{aligned} \tag{35}$$

we have

$$\begin{aligned} |Q_N(m)| &= |G_{8N}(2m) - G_{2N}(m)| \\ &\le |G_{8N}(2m)| + |G_{2N}(m)| \\ &\le 2\sqrt{N} + \sqrt{N} = 3\sqrt{N} \end{aligned} \tag{36}$$

for $0 \le m \le N$, where the inequality $|G_{2N}(l)| \le \sqrt{N}$ for $l \le N$ is applied as before. As a result,

$$|\mathbf{U}(p,q)| \le 6\sqrt{2}, \tag{37}$$

when q is odd.

In summary, the bound of $\mathbf{U}(p,q)$ for $0 \le p, q \le N-1$ is

$$\max\left\{ 2\sqrt{2} + \frac{\sqrt{2}}{\sqrt{N}}, 6\sqrt{2} \right\} = 6\sqrt{2}, \tag{38}$$

which completes the proof. ∎

B. Binary Sequences

In the binary case, the sequences are generally not perfect. So we use nearly perfect sequences. There are various binary sequences with nearly perfect and good autocorrelation property [27]. Specifically, we consider the m-sequence and Legendre sequence.

The m-sequence is a binary signal of length $N = 2^k - 1$, where k is an positive integer. The autocorrelation of an m-sequence $\mathbf{s}$ is given by [24]

$$R_{\mathbf{s}}(l) = \begin{cases} N, & l \equiv 0 \mod N; \\ -1, & \text{otherwise.} \end{cases} \tag{39}$$

Accordingly, from the Wiener-Khinchin relation we have

$$|a_k| = \begin{cases} 1, & k = 0; \\ \sqrt{N+1}, & \text{otherwise.} \end{cases} \tag{40}$$

Therefore, the coherence parameter is

$$\mu_m(\mathbf{A}) = \sqrt{(N+1)/N} \doteq 1. \tag{41}$$

The Legendre sequence can provide a similar bound. A Legendre sequence has a length equal to an odd prime N, and is defined as [32],

$$s_k = \begin{cases} 1, & \text{if } k \text{ is a square or } 0 \pmod N \\ -1, & \text{if } k \text{ is a nonsquare } \pmod N. \end{cases} \tag{42}$$

The FFT of the Lengendre sequence is known as [32]

$$\hat{s}_k = \begin{cases} 1 + s_k\sqrt{N}, & \text{if } N \equiv 1 \pmod 4 \\ 1 + js_k\sqrt{N}, & \text{if } N \equiv 3 \pmod 4. \end{cases} \tag{43}$$

Therefore, we have

$$\mu_{\text{Legendre}}(\mathbf{A}) = \begin{cases} 1 + \frac{1}{\sqrt{N}} \doteq 1 & \text{if } N \equiv 1 \pmod 4 \\ \sqrt{1 + \frac{1}{N}} \doteq 1 & \text{if } N \equiv 3 \pmod 4. \end{cases} \tag{44}$$

The proof of the bound $\mu_{\text{Golay}}(\mathbf{A}) \le \sqrt{2}$ for Golay sequences (for complex sensing matrices) is deferred to the next section, which will simply follow as a corollary (Corollary 1).

V. Orthogonal Matrix A

When σ satisfies the conjugate symmetry, the matrix $\mathbf{A}$ will be an orthogonal matrix satisfying $\mathbf{A}^T\mathbf{A} = N\mathbf{I}_N$.

A. Polyphase Sequences

We define an extended chirp sequence, which resembles but is not the same as the FZC sequence, in the following manner. When N is an even number, the diagonal sequence is given by

$$\sigma_k = \begin{cases} 1 & k = 0 \\ e^{-j\frac{\pi}{N}k^2} & 1 \le k \le \frac{N}{2} - 1 \\ 1 & k = \frac{N}{2} \\ e^{j\frac{\pi}{N}k^2} & \frac{N}{2} + 1 \le k \le N - 1; \end{cases} \tag{45}$$

when N is an odd number, the diagonal sequence is given by

$$\sigma_k = \begin{cases} 1 & k = 0 \\ e^{-j\frac{\pi}{N}k^2} & 1 \le k \le \frac{N-1}{2} \\ -e^{j\frac{\pi}{N}k^2} & \frac{N+1}{2} \le k \le N - 1. \end{cases} \tag{46}$$

Theorem 5: Let σ be defined as (45) or (46), for even and odd N, respectively. Then the coherence parameter satisfies

$$\mu(\mathbf{A}) \doteq \begin{cases} 4, & N \text{ even;} \\ 2.69, & N \text{ odd.} \end{cases} \tag{47}$$

We also adopt the incomplete Gauss sum to prove the result. The case of even N is given below, while the case of odd N is similar and the details are given in Appendix B.

Proof: Since $\sigma_{N-k}=\sigma_k^*, 1\le k\le \frac{N}{2}-1$, we have

$$\begin{aligned}a_l &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}lk}\cdot\sigma_k\\ &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N/2-1} e^{j\frac{2\pi}{N}lk}e^{-j\frac{\pi}{N}k^2}+\frac{(-1)^l}{\sqrt{N}}+\frac{1}{\sqrt{N}}\sum_{k=\frac{N}{2}+1}^{N-1} e^{j\frac{2\pi}{N}lk}e^{j\frac{\pi}{N}k^2}\\ &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N/2-1} e^{\frac{j\pi}{N}\left(-k^2+2lk-l^2\right)}\cdot e^{\frac{j\pi}{N}l^2}+\\ &\quad \frac{1}{\sqrt{N}}\sum_{k=0}^{N/2-1} e^{\frac{j\pi}{N}\left(k^2-2lk+l^2\right)}\cdot e^{-\frac{j\pi}{N}l^2}-\frac{1}{\sqrt{N}}+\frac{(-1)^l}{\sqrt{N}}.\end{aligned} \tag{48}$$

As a result,

$$|a_l|\le\frac{2}{\sqrt{N}}\left|\sum_{k=0}^{N/2-1} e^{\frac{j\pi}{N}(k-l)^2}\right|+\frac{2}{\sqrt{N}},\quad 0\le l\le N-1. \tag{49}$$

Similar to the proof of *Theorem* 4, we analyze the absolute value in different cases. When $0\le l\le\frac{N}{2}-1$,

$$\begin{aligned}|a_l| &\le \frac{2}{\sqrt{N}}\left|\sum_{k=-l}^{-1} e^{j\frac{\pi}{N}k^2}+\sum_{k=0}^{N/2-1-l} e^{j\frac{\pi}{N}k^2}\right|+\frac{2}{\sqrt{N}}\\ &= \frac{2}{\sqrt{N}}\left|\sum_{k=0}^{l} e^{j\frac{\pi}{N}k^2}+\sum_{k=0}^{N/2-1-l} e^{j\frac{\pi}{N}k^2}-1\right|+\frac{2}{\sqrt{N}}\\ &= \frac{2}{\sqrt{N}}\left|G_{2N}(l+1)+G_{2N}\left(\frac{N}{2}-l\right)-1\right|+\frac{2}{\sqrt{N}}\\ &\le \frac{2}{\sqrt{N}}\left(2\sqrt{N}+1\right)+\frac{2}{\sqrt{N}}=4+\frac{4}{\sqrt{N}}.\end{aligned} \tag{50}$$

To deal with the other half, we consider $|a_{N-l}|$ for $0\le l\le\frac{N}{2}-1$. Because

$$\begin{aligned}|a_{N-l}| &\le \frac{2}{\sqrt{N}}\left|\sum_{k=0}^{N/2-1} e^{\frac{j\pi}{N}(k+l-N)^2}\right|+\frac{2}{\sqrt{N}}\\ &= \frac{2}{\sqrt{N}}\left|\sum_{k=0}^{N/2-1} e^{\frac{j\pi}{N}(k+l)^2}\right|+\frac{2}{\sqrt{N}}\\ &= \frac{2}{\sqrt{N}}\left|G_{2N}\left(\frac{N}{2}+l\right)-G_{2N}(l)\right|+\frac{2}{\sqrt{N}},\end{aligned} \tag{51}$$

we have the same bound

$$\begin{aligned}|a_{N-l}| &\le \frac{2}{\sqrt{N}}\left(\left|G_{2N}\left(\frac{N}{2}+l\right)\right|+|G_{2N}(l)|\right)+\frac{2}{\sqrt{N}}\\ &\le \frac{2}{\sqrt{N}}\cdot 2\sqrt{N}+\frac{2}{\sqrt{N}}=4+\frac{2}{\sqrt{N}}.\end{aligned} \tag{52}$$

So for $0\le l\le N-1$,

$$|a_l|\le 4+4/\sqrt{N}. \tag{53}$$

Thus the coherence parameter

$$\mu(\mathbf{A})=\max\{|a_l|\}=4+4/\sqrt{N}\doteq 4. \tag{54}$$

■

B. Binary Sequences

In this subsection, we construct real sensing matrices from binary sequences with low peak-to-mean envelope power ratio (PMEPR). PMEPR has been extensively studied in the area of orthogonal frequency-division multiplexing (OFDM) communications, where the peak value of the IFFT of a data sequence is to be reduced. Using this connection, many of the low-PMEPR sequences can be applied to generate sensing matrices with small $\mu(\mathbf{A})$.

Definition 3 (PMEPR [33]): Let $\mathbf{c}=(c_0,\ldots,c_{N-1})$ be a codeword drawn from a given constellation. The complex envelope of a multicarrier signal with N subcarriers may be represented as

$$S_{\mathbf{c}}(\omega)=\sum_{k=0}^{N-1} c_k e^{j\omega k},\quad 0\le\omega<2\pi. \tag{55}$$

Then the PMEPR of codeword $\mathbf{c}$ is defined as[3]

$$\mathrm{PMEPR}_{\mathbf{c}}=\max_{0\le\omega<2\pi}\frac{|S_{\mathbf{c}}(\omega)|^2}{E\left[\sum_{i=0}^{N-1}|c_i|^2\right]} \tag{56}$$

Theorem 6 ($\mu(\mathbf{A})$ for low PMEPR sequences): Let $\mathbf{s}$ be a binary sequence with a constant PMEPR $C_{\mathbf{s}}$. Construct sequence σ as $\sigma=[s_0,\cdots,s_{N/2-1},s_0,s_{N/2-1},\cdots,s_1]$ for even N, and $\sigma=[s_0,\cdots,s_{(N-1)/2},s_{(N-1)/2},\cdots,s_1]$ for odd N. Then the coherence parameter of $\mathbf{A}$ is bounded by

$$\mu(\mathbf{A})\le\begin{cases}\sqrt{2C_{\mathbf{s}}}+\frac{2}{\sqrt{N}}\doteq\sqrt{2C_{\mathbf{s}}}, & N \text{ even};\\ \sqrt{2C_{\mathbf{s}}}+\frac{1}{\sqrt{N}}\doteq\sqrt{2C_{\mathbf{s}}}, & N \text{ odd}.\end{cases} \tag{57}$$

Proof: We show the proof for even N, while the case of odd N is omitted since it is very similar. Since the PMEPR of $\mathbf{s}$ is $C_{\mathbf{s}}$, we have

$$\frac{|S_{\mathbf{s}}(\omega)|^2}{N/2}\le C_{\mathbf{s}}, \tag{58}$$

where $S_{\mathbf{s}}(\omega)=\sum_{i=0}^{N/2-1}s_ie^{j\omega i}$, for any $0\le\omega<2\pi$.

On the other hand, we have

$$\mu(\mathbf{A})=\frac{1}{\sqrt{N}}\max_k\left|\sum_{i=0}^{N-1}\sigma_ie^{\frac{2\pi j}{N}ki}\right| \tag{59}$$

Due to the symmetry of the sequence, the second half (except the $i=N/2$ term) is a complex conjugate of the first half. Thus, the sum in (59) can bounded by

$$\begin{aligned}\mu(\mathbf{A})\le\frac{1}{\sqrt{N}}\max_k\Bigg\{&\left|\sum_{i=0}^{N/2-1}s_ie^{\frac{2\pi j}{N}ki}\right|+\\ &\left|\sum_{i=0}^{N/2-1}s_ie^{-\frac{2\pi j}{N}ki}+s_0e^{j\pi k}-s_0\right|\Bigg\}.\end{aligned} \tag{60}$$

[3]Sometimes it is also referred to as the peak-to-average power ratio (PAPR) in literature.

From (58), we obtain

$$\begin{aligned}\mu(\mathbf{A}) &\leq \frac{2}{\sqrt{N}} \max_k \left\{ \left| \sum_{i=0}^{N/2-1} s_i e^{\frac{2\pi j}{N} ki} \right| + 1 \right\} \\ &\leq \frac{2}{\sqrt{N}} \left(\sqrt{\frac{N}{2} C_{\mathbf{s}}} + 1 \right) \\ &\leq \sqrt{2C_{\mathbf{s}}} + \frac{2}{\sqrt{N}},\end{aligned} \tag{61}$$

where the second inequality is because of (58). As $N \to \infty$, $\mu(\mathbf{A})$ will approach $\sqrt{2C_{\mathbf{s}}}$. ■

The binary Golay sequences, introduced by Golay in 1961 [14], have found numerous applications in communications and signal processing [33], [34]. They are known to exist for all lengths $2^{\kappa_1}10^{\kappa_2}26^{\kappa_3}$, $\kappa_1, \kappa_2, \kappa_3$ non-negative integers. For the constructions of Golay sequences, please refer to Appendix C.

Definition 4 (Golay sequences): The aperiodic autocorrelation function of a sequence $\mathbf{s}$ is defined by

$$r_{\mathbf{s}}(l) = \sum_{i=0}^{N-l-1} s_i s_{i+l}^*, \quad l = 0, \cdots, N-1. \tag{62}$$

Let $\mathbf{a} = (a_0, a_1, \cdots, a_{N-1})$ and $\mathbf{b} = (b_0, b_1, \cdots, b_{N-1})$ be a pair of binary sequences with values 1 or -1. Then $\mathbf{a}$ and $\mathbf{b}$ form a Golay complementary pair if

$$r_{\mathbf{a}}(l) + r_{\mathbf{b}}(l) = \begin{cases} 2N, & l = 0 \\ 0, & l = 1, \cdots, N-1. \end{cases} \tag{63}$$

A sequence in any complementary pairs is called a *Golay sequence*.

Taking the Fourier transform of (63), the corresponding power spectrum $S_{\mathbf{a}}$ and $S_{\mathbf{b}}$ satisfy the following relation

$$|S_{\mathbf{a}}(\omega)|^2 + |S_{\mathbf{b}}(\omega)|^2 = 2N. \tag{64}$$

It simply follows that

$$|S_{\mathbf{s}}(\omega)|^2 \leq 2N, \tag{65}$$

for any Golay sequence $\mathbf{s}$ with length N. Besides, $|s_i|^2 = 1, i = 0, \cdots, N-1$ holds for all binary Golay sequences. From (56), the PMEPR of a Golay sequence is [33]

$$\mathrm{PMEPR}_{\mathrm{Golay}} = \max_{0 \leq \omega < 2\pi} \frac{|S_{\mathbf{s}}(\omega)|^2}{N} \leq 2. \tag{66}$$

Substituting it into (57), we obtain the $\mu(\mathbf{A})$ based on Golay sequence for real $\mathbf{A}$:

$$\mu_{\mathrm{Golay}}(\mathbf{A}) \leq \begin{cases} 2 + \frac{2}{\sqrt{N}}, & N \text{ even;} \\ 2 + \frac{1}{\sqrt{N}}, & N \text{ odd.} \end{cases} \tag{67}$$

and when $N \to \infty$, $\mu_{\mathrm{Golay}}(\mathbf{A})$ will approach 2.

Remark 1 (Orthogonal symmetric Toeplitz matrices (OSTM)): The authors introduced OSTM [35] as sensing matrices in [36], and proposed to use the Golay sequence as the diagonal sequence in [37]. In general, OSTM may be viewed as a special case of convolutional CS, where the sequence $\mathbf{s} \in \{-1, 1\}^{N/2} \backslash \{(-1, -1, \cdots, -1), (1, 1, \cdots, 1)\}$. However, if the binary sequence $\mathbf{s}$ is randomly selected from this set, the bound on $\mu(\mathbf{A})$ will be poor. In fact, it was shown in [38] that the PMEPR of a random codeword of length N is asymptotically $\log N$ with probability 1. Therefore, our judicious selection of a deterministic sequence with constant $C_{\mathbf{s}}$ leads to a stronger theoretic guarantee of $\mu(\mathbf{A})$.

Corollary 1: Let the diagonal sequence σ be simply a Golay sequence $\mathbf{s}$ of length N such that $\mathbf{A}$ is a *complex* matrix. Then the coherence parameter $\mu(\mathbf{A}) \leq \sqrt{2}$.

Proof: Let $\omega = 2\pi k/N$ in (65), we have the following bound on the FFT of a Golay sequence $\mathbf{s}$:

$$|\hat{s}_k|^2 \leq 2N. \tag{68}$$

Thus, we obtain

$$\mu_{\mathrm{Golay}}(\mathbf{A}) \leq \sqrt{\frac{2N}{N}} = \sqrt{2}. \tag{69}$$

■

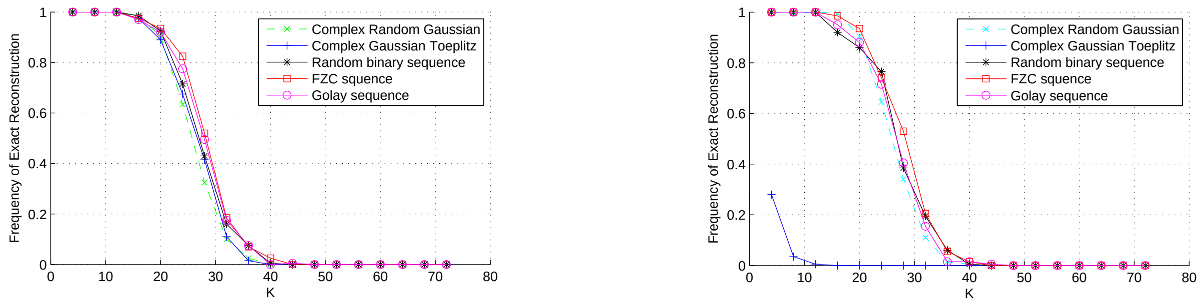
Remark 2: Besides Golay sequences, other sequences with low PMEPR can also be applied to obtain a constant coherence parameter $\mu(\mathbf{A}) \sim \mathcal{O}(1)$. In [39], near-complementary sequences with PMEPR < 4 were proposed[4]. Using these sequences, $\mu(\mathbf{A})$ will be about $2\sqrt{2}$ for a real matrix $\mathbf{A}$(and 2 for a complex matrix $\mathbf{A}$).

VI. Simulation Results

Extensive simulations have been carried out. For illustration purposes, we first present some results for complex sampling matrices of sizes 128×1024. The recovery performance is compared with that of Gaussian random/Toeplitz matrices and random sequence-based matrices (i.e., σ is a random binary sequence). The reconstruction algorithms are based on the fixed-point continuation and active set algorithm (FPC-AS) for solving l_1-regularized least-squares problem [40]. We consider K-sparse signals in the time and DCT domains. For the time domain, the K non-zero elements of signal $\mathbf{x}$ are selected uniformly at random, and its non-zero coefficients x_i ($i = 1, \cdots K$) obey the Gaussian distribution. For the DCT domain, the input signal is generated in the same way but sparse in DCT domain. Fig. 2 depicts the empirical frequencies of exact reconstruction for different sensing matrices with 500 trials run for each sparsity level K. We assume that the exact reconstruction is achieved if the signal to noise ratio (SNR) is greater than 50 dB. From these figures, one can observe that the performance of proposed sensing matrices are quite close to those of random sensing matrices in time domain, while in DCT domain, proposed sensing matrices substantially outperform complex Gaussian Toeplitz matrices.

Then we show simulation results on images using real-valued sensing matrices in Fig. 3 for 10% sampling rate. The Shepp-Logan phantom image (128×128) is used as a test image and the results are compared with random convolution. The fast reconstruction algorithm for Toeplitz matrices in [11] is applied. The figure reveals that random sampling in combination with both random sequences and deterministic sequences have good reconstruction performance, while deterministic sampling yields poor performance. Specifically,

[4] These are q-phase sequences for even integer q.



(a) (b)

Fig. 2. Simulation results for 128×1024 complex sensing matrices for K-sparse signals: (a) in time domain with FPC-AS; (b) in DCT domain with FPC-AS.

the performances of the proposed sensing matrices based on extended chirp and Golay sequences are close to that of random convolution, while the extended Golay sequence results in better performance than the extended chirp sequence.

VII. Conclusions

In this paper, a new class of circulant sensing matrices are proposed. We show that these convolutional sensing matrices have small coherence parameter $\mu(\mathbf{A})$, so that the original signal could be faithfully recovered. The underlying diagonal sequence is fixed as a (nearly) perfect sequence, such as the FZC sequence and Golay sequence. Our proposed matrices have the semi-universal property, in the sense that they are able to reconstruct sparse signals in more than one domain. Experimental results show that these sensing matrices compare favorably with existing structured random matrices.

Appendix A
Gauss Sums

Definition 5: Let N be a positive integer. The exponential sum [41]

$$G_N(m) = \sum_{k=0}^{m-1} e^{2\pi j k^2/N} \tag{70}$$

is an *incomplete Gauss sum* when $m < N$.

When $m = N$, the complete Gauss sum $G_N(N)$ is well known [41]

$$G_N(N) = \begin{cases} (1+j)\sqrt{N}, & \text{if} \quad N \equiv 0 \pmod 4 \\ \sqrt{N}, & \text{if} \quad N \equiv 1 \pmod 4 \\ 0, & \text{if} \quad N \equiv 2 \pmod 4 \\ j\sqrt{N}, & \text{if} \quad N \equiv 3 \pmod 4. \end{cases} \tag{71}$$

Moreover, when $m \le (N+1)/2$,

$$G_N(m) + G_N(N-m+1) = 1 + G_N(N). \tag{72}$$

Define a normalized version $g_N(m)$ as

$$g_N(m) = 2N^{-\frac{1}{2}} \sum_{k=0}^{m} e^{2\pi j k^2/N}, \tag{73}$$

then we have [41]

$$|g_N(m)| \le \begin{cases} \sqrt{2}, & \text{if} \quad N = 4k, m \le N/2 \\ 1.07 + \mathcal{O}(N^{-\frac{1}{2}}), & \text{if} \quad N = 4k+1, m < N/2 \\ 0.95 + \frac{101}{40} N^{-\frac{1}{2}}, & \text{if} \quad N = 4k+2, m \le N \\ \sqrt{1+N^{-1}}, & \text{if} \quad N = 4k+3, m < N/2, \end{cases} \tag{74}$$

where k is a positive integer.

Now we let N be an even integer, $N = 2N_0$ and $L = 2N$, so that

$$\begin{aligned} G_L(m) &= \frac{\sqrt{L}}{2} g_L(m) \\ G_{2N}(m) &= \sum_{k=0}^{m-1} e^{j\frac{2\pi}{2N}k^2} = \sum_{k=0}^{m-1} e^{j\frac{\pi}{N}k^2}. \end{aligned} \tag{75}$$

When $0 \le m \le N$,

$$\begin{aligned} G_{2N}(m) &= \frac{\sqrt{L}}{2} g_L(m) = \frac{\sqrt{2N}}{2} g_{4N_0}(m), \\ |G_{2N}(m)| &\le \frac{\sqrt{2N}}{2} \cdot \sqrt{2} = \sqrt{N}, \end{aligned} \tag{76}$$

where the second equality is because $\max|g_{2N}(m)| = \sqrt{2}$ when $m \le N$ [41]. When $N+1 \le m \le 2N-1$,

$$\begin{aligned} G_{2N}(m) &= (1+j)\sqrt{2N} + 1 - G_{2N}(2N-1-m), \\ |G_{2N}(m)| &\le \sqrt{2} \cdot \sqrt{2N} + 1 + \sqrt{N} = 3\sqrt{N} + 1. \end{aligned} \tag{77}$$

Moreover, when N is an odd integer, let $N = 2N_0 + 1$ and $L = 2N = 4N_0 + 2$. If $m \le N$, then [41]

$$\begin{aligned} G_L(m) &= \frac{\sqrt{L}}{2} g_L(m) = \sum_{k=0}^{m-1} e^{j\frac{\pi}{N}k^2}. \\ |G_{2N}(m)| &\le \frac{\sqrt{2N}}{2} \cdot \left(0.95 + \frac{101/40}{\sqrt{N}}\right). \end{aligned} \tag{78}$$

(a) (b) (c) (d)

Fig. 3. Simulation results for different real-valued sensing matrices to recover the 128×128 Shepp-Logan phantom image: (a) random binary sequence, deterministic sampling; (b) random binary sequence, random sampling; (c) extended chirp sequence, random sampling; (d) extended Golay sequence, random sampling.

APPENDIX B
PROOF OF THEOREM 5 FOR ODD N

Since $\sigma_{N-k} = \sigma_k^*, 1 \le k \le \frac{N-1}{2}$, we have

$$\begin{aligned} a_l &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}lk} \cdot \sigma_k \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{(N-1)/2} e^{j\frac{2\pi}{N}lk} e^{-\frac{\pi}{N}jk^2} - \frac{1}{\sqrt{N}} \sum_{k=(N+1)/2}^{N-1} e^{j\frac{2\pi}{N}lk} e^{\frac{j\pi}{N}k^2} \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{(N-1)/2} e^{\frac{j\pi}{N}\left(-k^2+2lk-l^2\right)} \cdot e^{\frac{j\pi}{N}l^2} + \\ &\quad \frac{1}{\sqrt{N}} \sum_{k=0}^{(N-1)/2} e^{\frac{j\pi}{N}\left(k^2-2lk+l^2\right)} \cdot e^{-\frac{j\pi}{N}l^2} - \frac{1}{\sqrt{N}}. \end{aligned} \tag{79}$$

As a result,

$$|a_l| \le \frac{2}{\sqrt{N}} \left| \sum_{k=0}^{(N-1)/2} e^{\frac{j\pi}{N}(k-l)^2} \right| + \frac{1}{\sqrt{N}}, \quad 0 \le l \le N-1. \tag{80}$$

So

$$\begin{aligned} |a_{N-l}| &\le \frac{2}{\sqrt{N}} \left| \sum_{k=0}^{(N-1)/2} e^{\frac{j\pi}{N}(k+l-N)^2} \right| + \frac{1}{\sqrt{N}} \\ &= \frac{2}{\sqrt{N}} \left| \sum_{k=0}^{(N-1)/2} e^{\frac{j\pi}{N}(k+l)^2} \right| + \frac{1}{\sqrt{N}} \\ &= \frac{2}{\sqrt{N}} \left| G_{2N}\left(\frac{N+1}{2} + l\right) - G_{2N}(l) \right| + \frac{1}{\sqrt{N}}. \end{aligned} \tag{81}$$

In this case let $2N = 4N_0 + 2$ (as $N = 2N_0 + 1$ is odd), where N_0 is an integer. Because for any $m < 4N_0 + 2$, we have [41]

$$\begin{aligned} |g_{4N_0+2}(m)| &< 0.95 + \frac{101/40}{\sqrt{N}}, \\ |G_{2N}(m)| &\le \frac{\sqrt{2N}}{2} \cdot \left(0.95 + \frac{101/40}{\sqrt{N}}\right). \end{aligned} \tag{82}$$

Therefore, we arrive at

$$\begin{aligned} \mu(\mathbf{A}) &= \max\{|a_l|\} \\ &\le \frac{4}{\sqrt{N}} \frac{\sqrt{2N}}{2} \cdot \left(0.95 + \frac{101/40}{\sqrt{N}}\right) + \frac{1}{\sqrt{N}} \\ &\le 2.69 + \frac{8.15}{\sqrt{N}}. \end{aligned} \tag{83}$$

APPENDIX C
CONSTRUCTIONS OF GOLAY SEQUENCES

There are two approaches to constructing binary Golay sequences: direct construction and recursive construction.

A. Direct construction

Consider a *Boolean function* f from $\mathbb{Z}_2^l = \{(x_0, x_1 \cdots, x_{l-1}) | x_i \in \{0,1\}\}$ to $\mathbb{Z}_2$. Any f can be uniquely expressed as a linear combination of the 2^l monomials:

$$1, x_0, x_1, \cdots, x_{l-1}, x_0x_1, x_0x_2, \cdots, x_{l-2}x_{l-1}, \cdots, x_0x_1\cdots x_{l-1}. \tag{84}$$

The resulting expression is known as the *algebraic normal form* of f [34]. With the Boolean function f, we associate a length-2^l sequence $\mathbf{f}$, where the ith element of $\mathbf{f}$ is $f(i_0, i_1, \cdots, i_{l-1})$ and $(i_0, i_1, \cdots, i_{l-1})$ is the binary representation of the integer $i = \sum_{k=0}^{l-1} i_k 2^{l-k-1}$ (i_0 is the most significant bit). Taking $l = 3$ as an example,

$$\begin{aligned} \mathbf{f} = &(f(0,0,0), f(0,0,1), f(0,1,0), f(0,1,1), \\ &f(1,0,0), f(1,0,1), f(1,1,0), f(1,1,1)). \end{aligned} \tag{85}$$

Now we can use this notion to describe the construction of binary Golay sequences.

Theorem 7 ([34]): For any permutation π of $\{0, 1, \cdots, l-1\}$ and any choice of constants $c_k, c \in \mathbb{Z}_2$, let

$$f(x_0, \cdots, x_{l-1}) = \sum_{k=0}^{l-2} x_{\pi(k)}x_{\pi(k+1)} + \sum_{k=0}^{l-1} c_kx_k, \tag{86}$$

then

$$\mathbf{a}(x_0, x_1, \cdots, x_{l-1}) = f(x_0, x_1, \cdots, x_{l-1}) + c \tag{87}$$

generates a binary Golay sequence of length 2^l under the mapping $0 \mapsto 1, 1 \mapsto -1$.

Using this construction, one obtains a set of $2^{(l+1)}l!/2$ Golay sequences of length 2^l.

B. Recursive construction

In this situation, it is helpful to rewrite a sequence $\mathbf{a}$ in the polynomial form,

$$a(z) = a_{N-1}z^{N-1} + a_{N-2}z^{N-2} + \cdots + a_1 z + a_0. \tag{88}$$

Then the corresponding polynomials $(a(z), b(z))$ of the Golay complementary pair $(\mathbf{a}, \mathbf{b})$ satisfy

$$a(z)a(z^{-1}) + b(z)b(z^{-1}) = 2N. \tag{89}$$

This equation is derived from (63). Simple calculation shows that $a(z) + z^N b(z)$ and $a(z) - z^N b(z)$ are also a Golay complementary pair satisfying (89) with length $2N$. So a length-$2N$ Golay pair could be generated from a length-N pair.

In addition, if $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ are Golay complementary pairs of length N_1 and N_2, respectively, then

$$\begin{aligned} &a(z^{N_2})(c(z) + d(z))/2 + z^{N_2(N_1-1)}b(z^{-N_2})(c(z) - d(z))/2, \\ &b(z^{N_2})(c(z) + d(z))/2 - z^{N_2(N_1-1)}a(z^{-N_2})(c(z) - d(z))/2 \end{aligned} \tag{90}$$

also form a Golay complementary pair of length $N_1 N_2$.

Moreover, letting $\tilde{b}$ be the reversal of b, Golay gave the following two constructions:
Concatenation:

$$a(z^{N_2})c(z) + b(z^{N_2})d(z)z^{N_1N_2}, \tilde{b}(z^{N_2})c(z) - \tilde{a}(z^{N_2})d(z)z^{N_1N_2} \tag{91}$$

is a Golay complementary pair of length $2N_1N_2$.
Interleaving:

$$a(z^{2N_2})c(z^2) + b(z^{2N_2})d(z^2)z, \tilde{b}(z^{2N_2})c(z^2) - \tilde{a}(z^{2N_2})d(z^2)z \tag{92}$$

is also a Golay complementary pair of length $2N_1N_2$.

So Golay complementary pairs of lengths $N = 2^{\kappa_1}10^{\kappa_2}26^{\kappa_3}, \kappa_1, \kappa_2, \kappa_3 \geq 0$ can be constructed in foregoing ways from several primitive pairs of lengths $2, 10, 26$. More details could be found in [42].

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Jul. 2006.

[2] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[3] L. Gan, T. Do, and T. D. Tran, "Fast compressive imaging using scrambled block Hadamard ensemble," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Lausanne, Switzerland, Aug. 2008.

[4] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.

[5] J. A. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstrution," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Toulouse, France, May 2006, pp. 872–875.

[6] T. Do, L. Gan, N. Nguyen, and T. Tran, "Fast and efficient compressive sensing using structurally random matrices," *IEEE Trans. Signal Process.*, vol. 60, pp. 139–154, Jan. 2012.

[7] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[8] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Imag. Sci.*, vol. 2, no. 4, pp. 1098–1128, Dec. 2009.

[9] H. Rauhut, J. Romberg, and J. Tropp, "Restricted isometries for partial random circulant matrices," *Appl. Comput. Harmon. Anal.*, May 2011, to be published.

[10] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE J. Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[11] W. Yin, S. P. Morgan, J. Yang, and Y. Zhang, "Practical compressive sensing with Toeplitz and circulant matrices," Rice University CAAM Technical Report TR10-01, 2010.

[12] R. Marcia, Z. Harmany, and R. Willett, "Compressive coded aperture imaging," in *SPIE electronic imaging*, 2009.

[13] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.

[14] M. Golay, "Complementary series," *IEEE Trans. Inf. Theory*, vol. 7, no. 2, pp. 82 –87, Apr. 1961.

[15] l_1-magic. [Online]. Available: http://www.l1-magic.org

[16] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[17] J. A. Tropp, "Greed is good: Algorithm results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, 2004.

[18] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2230–2249, May 2009.

[19] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 301–321, 2009.

[20] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Commun. on Pure and Applied Math.*, vol. 61, pp. 1025–1045, Dec. 2008.

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

[22] J. Tropp, "Random filters for compressive sampling," in *Proc. 40th Ann. Conf. Information Sciences and Systems (CISS)*, Mar. 2006, pp. 216–217.

[23] H. Rauhut, "Compressive sensing and structured random matrices," *Radon Series Comp. Appl. Math.*, Theoretical Foundations and Numercial Methods for Sparse Recovery, Publisher: DeGruyter, vol. 9, pp. 1-92, 2010.

[24] Hans-Jurgen, Zepernick, and A. Finger, *Pseudo Random Signal Processing: Theory and Application*. Wiley, 2005.

[25] P. Z. Fan and M. Darnell, "The synthesis of perfect sequences," *Lecture notes in Computer Science, Cryptography and Codingm Springer*, 1995.

[26] W. H. Mow, "A new unified construction of perfect root-of-unity sequences," in *Proc. of IEEE 4th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, vol. 3, Sep. 1996, pp. 955–959.

[27] H. Luke, H. Schotten, and H. Hadinejad-Mahram, "Binary and quadriphase sequences with optimal autocorrelation properties: A survey," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3271–3282, Dec. 2003.

[28] Y. Wiaux, G. Puy, Y. Boursier, and P. Vandergheynst, "Spread spectrum for imaging techniques in radio interferometry," *Mon. Not. R. Astron. Soc.*, vol. 400, no. 99, pp. 1029–1038, 2009.

[29] L. Applebaum, S. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast

recovery," *Appl. Comput. Harmonic Anal.*, vol. 26, no. 2, pp. 283–290, Mar. 2009.

[30] J. P. Elsner, M. Braun, H. Jakel, and F. K. Jondral, "Compressed spectrum estimation for cognitive radios," in *Proc. of 19th Virginia Tech Symposium on Wireless Communications*, 2009.

[31] C. Li and W. Huang, "A constructive representation for the Fourier dual of the Zadoff-Chu sequences," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4221–4224, Nov. 2007.

[32] J. Jensen, H. Jensen, and T. Hoholdt, "The merit factor of binary sequences related to difference sets," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 617–626, May 1991.

[33] B. Popovic, "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE Trans. on Commun.*, vol. 39, no. 7, pp. 1031 –1033, Jul. 1991.

[34] J. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed-Muller codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2397–2417, Nov. 1999.

[35] A. Bottcher, "Orthogonal symmetric Toeplitz matrices," *Complex Analysis and Operator Theory*, vol. 2, no. 2, pp. 285–298, 2008.

[36] K. Li, C. Ling, and L. Gan, "Statistical restricted isometry property of orthogonal symmetric Toeplitz matrices," in *Proc. IEEE Information Theory Workshop*, Oct. 2009, pp. 183–187.

[37] ——, "Deterministic compressed sensing matrices: where Toeplitz meets Golay," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2011, pp. 3748–3751.

[38] M. Sharif and B. Hassibi, "On multicarrier signals where the PMEPR of a random codeword is asymptotically log$n$," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 895–903, May 2004.

[39] N. Y. Yu and G. Gong, "Near-complementary sequences with low PMEPR for peak power control in multicarrier communications," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 505–513, Jan. 2011.

[40] Z. Wen, W. Yin, H. Zhang, and D. Goldfarb, "On the convergence of an active set method for l_1-minimization," *Optimization Methods and Software*, 2011, to be published.

[41] D. Lehmer, "Incomplete Gauss sums," *Mathematika*, vol. 11, pp. 125–135, 1976.

[42] M. G. Parker, K. G. Paterson, C. Tellambura, and J. G. Proakis, "Golay complementary sequences," *Wiley Encyclopedia of Telecommunications*, New York: Wiley, 2003.