Journal of Computers Vol. 32 No. 6, 2021, pp. 144-158
doi:10.53106/199115992021123206012

Design and Construction of High Performance Compressed Sensing Measurement Matrix

Yu-chen Yue[1], Jian-Hua Luo[2], Hua Li[1*]

[1] Department of Armament and control, Army Academy of armored Forces, Feng tai, Beijing
Yuchen.Yue@foxmail.com

[2] Center of maneuver and training, Army Academy of armored Forces, Feng tai, Beijing
jh.luo@VIP.sina.com

Received 1 May 2021; Revised 7 June 2021; Accepted 8 June 2021

Abstract. With the advancing application of compression theory in practical engineering problems, the hardware feasibility of the measurement matrix has gradually become an important factor in considering the comprehensive performance of the compressed sensing system. Although random matrix of independently identically distribution can be used as a universal CS measurement matrix due to its nearly incoherent nature with any signal, the hardware realization of the kind is hard to reach and tends to occupy large storage space. Thus, it is difficult to apply in resource-constrained scenarios. To solve this problem, an innovate Diagonal Toeplitz/Circulant structure block based sparse matrix is proposed in the paper. It then proves that the matrix satisfies the Restricted Isometry Property (RIP) conditions with a probability close to 1. The simulation experiments show that the measurement matric proposed in the paper can not only ensure the accurate reconstruction of sparse signals, but also greatly save the time required for measurement and reconstruction.

Keywords: compressed sensing, Toeplitz matrix, Correlation, measurement matrix, Gerschgorin

1 Introduction

Compressive Sensing (CS) [1-3] is a new theory of data acquisition and sampling. This theory breaks through the traditional Nyquist sampling theorem in terms of signal acquisition method. The Nyquist sampling theorem requires the sampling frequency to be higher than twice the bandwidth of the sampled signal, while compressed sensing requires the sampling signal to be sparse or sparse in the transform domain. In this regard, compressive sensing can sample the signal while compressing the signal at a much lower sampling rate than the Nyquist's. Therefore, problems like the huge amount of sampled data and the serious waste of sampling resources such as sensor element, sampling time and data storage space can be solved by compressed sensing.

Compressed sensing theory mainly includes two parts: signal acquisition and signal reconstruction. At the signal acquisition end, the performance of the measurement matrix is directly related to the number of signal measurements. A measurement matrix with excellent performance can lead to a higher accuracy of signal reconstruction with a small number of measurements. On the signal acquisition end, when reconstructing the raw signal measurement data collected with the same sampling rate, a measurement matrix with excellent performance has higher signal reconstruction accuracy.

Although the random measurement matrix has excellent performance and meets the RIP conditions, this type of matrix is very difficult to implement in hardware, and requires a high storage space, which cannot meet the needs of solving practical problems. Therefore, most scholars worldwide turn their research interest indeterministic measurement matrices and structured measurement matrices. Several promising research results have been achieved so far, mainly on certainty matrices, such as polynomial matrices [4-5], measurement matrices based on chaotic sequence [6], parity-check matrices [7-8] and so

* Corresponding Author

on. Among the results, the polynomial matrices and parity-check matrices perform better in reconstructing the sparse signal. However, the requirements for the number of rows and columns of the matrix are very strict, and the quality of compressible signals is poor. Besides, the measurement matrix based on the chaotic sequence is a dense matrix, which requires more storage space. Structural measurement matrices include the measurement matrix of block compressed sensing [9], sparse random matrix [10], Toeplitz matrix [11], generalized rotation matrix [12]. Among them, the sub-matrices and sparse random matrices that constitute the block compressed sensing measurement matrix are random matrices. Compared with deterministic matrices, hardware implementation is still difficult. The Toeplitz matrix and the generalized rotation matrix have a certain structure, which reduces the storage space, but the reconstruction accuracy is average.

In view of the advantages and disadvantages of the above measurement matrices, this paper proposes and constructs an innovate diagonal Toeplitz / Circulant block based sparse matrix. This matrix introduces the idea of blocking, viewing the sparse Toeplitz matrix as the block matrix. The matrix elements only contain ±1 and 0. The matrix are endowed with the properties of structure based on the structural characteristics of the Toeplitz Matrix or Circulant Matrix.

2 A New Design and Construction of Diagonal Toeplitz/Circulant Structure Block Based Sparse Matrix

Firstly, the paper will give a clear definition of the block matrix:

Definition 1: If $M \times N$ matrix B is constituted by several $m \times n$ dimensional matrices $\mathbf{B}_{i,j}$ ($i = 1,2,...,W$, $j = 1,2,...,L$), and the structure of the matric B is as follows:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,L-1} & \mathbf{B}_{1,L} \\ \mathbf{B}_{2,1} & \mathbf{B}_{2,2} & \cdots & \mathbf{B}_{2,L-1} & \mathbf{B}_{2,L} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{B}_{W,1} & B_{W,2} & \cdots & \mathbf{B}_{W,L-1} & \mathbf{B}_{W,L} \end{bmatrix} \tag{1}$$

Then the matrix $\mathbf{B}$ is defined as a block matrix, and $M = mW$, $N = nL$.

So the basic structure of the $N \times N$ dimensional Toeplitz matrix T is as follows:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & t_{-3} & \cdots & t_{-(N-1)} \\ t_1 & t_0 & t_{-1} & t_{-2} & \ddots & \vdots \\ t_2 & t_1 & \ddots & \ddots & \ddots & t_{-3} \\ t_3 & t_2 & \ddots & \ddots & t_{-1} & t_{-2} \\ \vdots & \ddots & \ddots & t_1 & t_0 & t_{-1} \\ t_{N-1} & \cdots & t_3 & t_2 & t_1 & t_0 \end{bmatrix} \tag{2}$$

Except for the elements in the first row and the first column of the matrix $\mathbf{T}$, the other elements are equal to those on the upper left. A special form of the Toeplitz matrix is called Circulant Matrix (C matrix for short). The basic structure of the $N \times N$ dimensional Circulant matrix $\mathbf{C}$ is as follows:

$$\mathbf{C} = \begin{bmatrix} t_0 & t_1 & t_2 & t_3 & \cdots & t_{N-1} \\ t_{N-1} & t_0 & t_1 & t_2 & \ddots & \vdots \\ t_{N-2} & t_{N-1} & \ddots & \ddots & \ddots & t_3 \\ t_{N-3} & t_{N-2} & \ddots & \ddots & t_1 & t_2 \\ \vdots & \ddots & \ddots & t_{N-1} & t_0 & t_1 \\ t_1 & \cdots & t_{N-3} & t_{N-2} & t_{N-1} & t_0 \end{bmatrix} \tag{3}$$

It can be seen that the $N \times N$ dimensional C Matrix is constructed by cyclic shifting the first-row of $N-1$ times, and ensuring that each cycle will shift the first-row elements to the right by one bit. The

nonzero vector $\mathbf{t}=\begin{bmatrix}t_1 & t_2 & \cdots & t_l\end{bmatrix}\in\mathbb{R}^{l\times 1}$, which is the first-row of the matrix $\mathbf{T}$, is a K-sparse vector, $0<K\ll l$. To guarantee that every column of the $\mathbf{T}$ contains at least one nonzero vector, the paper sets the first element t_1 and the last element t_l as nonzero. Other $K-2$ nonzero elements are randomly distributed in **t**, which means that the random distribution in **T** includes K "nonzero bands". The smaller the K value is, the sparser the matrix **T** is. The nonzero elements in **t** follows the Bernoulli random distribution:

$$t_j \sim \begin{cases} 1 & \text{with prob. } 1/2 \\ -1 & \text{with prob. } 1/2 \end{cases} \tag{4}$$

Above is the sparce vector element. A sparce Toeplitz matrix is constructed according to the methods of building a Toeplitz matrix:

$$\mathbf{A}=\begin{bmatrix} a_1 & a_2 & \cdots & a_l & 0 & \cdots & 0 \\ 0 & a_1 & a_2 & \cdots & a_l & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_1 & a_2 & \cdots & a_l \end{bmatrix} \tag{5}$$

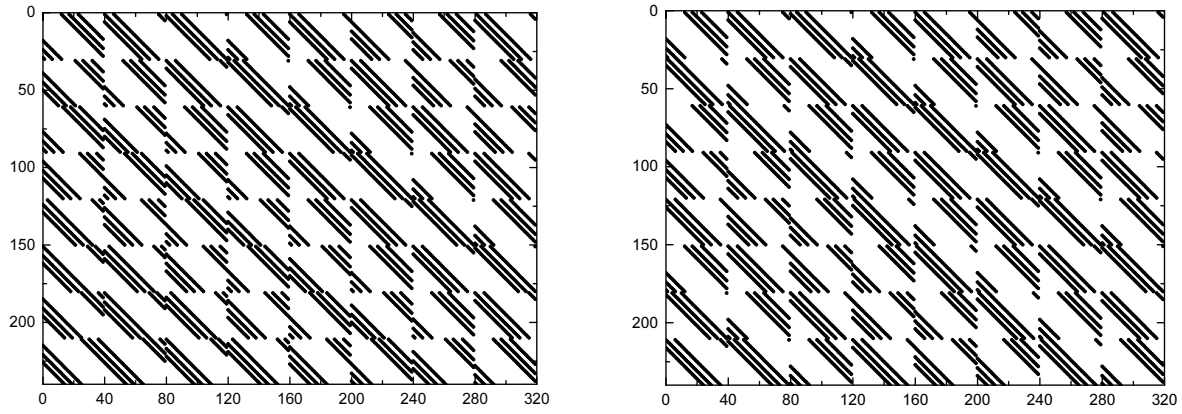
To ensure there is no signal information loss at the perception process, and each column of **A** has at least one nonzero element, the condition $N\le 2M$ should be met. The matrix $\mathbf{A}^c$ viewing as a block matrix is obtained after shifting **A** towards left or right and performing cyclic shifting for c times. In this part, **A** is uniformly set to shift towards right to perform the cyclic shifting. Amon them, c is a positive integer that obeys a random distribution in $(0,L]$, which means that the cyclic shifting value of each matrix is random and is distinguished by subscripts. The sparce Toeplitz block matrix $\mathbf{\Phi}$ is as equation (6) based on the sparse Toeplitz matrix **A**.

$$\mathbf{\Phi}=\begin{bmatrix} \mathbf{A}^{c_1} & \mathbf{A}^{c_{-2}} & \mathbf{A}^{c_{-3}} & \cdots & \cdots & \mathbf{A}^{c_{-L}} \\ \mathbf{A}^{c_2} & \mathbf{A}^{c_1} & \mathbf{A}^{c_{-2}} & \ddots & \ddots & \vdots \\ \mathbf{A}^{c_3} & \mathbf{A}^{c_2} & \mathbf{A}^{c_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{A}^{c_{-2}} & \mathbf{A}^{c_{-3}} \\ \vdots & \ddots & \ddots & \mathbf{A}^{c_2} & \mathbf{A}^{c_1} & \mathbf{A}^{c_{-2}} \\ \mathbf{A}^{c_L} & \cdots & \cdots & \mathbf{A}^{c_3} & \mathbf{A}^{c_2} & \mathbf{A}^{c_1} \end{bmatrix} \tag{6}$$

The matrix $\mathbf{\Phi}$ is a diagonal Toeplitz structure block sparse matrix. The diagonal circulant structure block sparse matrix $\mathbf{\Phi}'$ is as equation (8) constructed according to the sparse Toeplitz matrix **A**.

$$\mathbf{\Phi}'=\begin{bmatrix} \mathbf{A}^{c_1} & \mathbf{A}^{c_2} & \mathbf{A}^{c_3} & \cdots & \cdots & \mathbf{A}^{c_L} \\ \mathbf{A}^{c_L} & \mathbf{A}^{c_1} & \mathbf{A}^{c_2} & \ddots & \ddots & \vdots \\ \mathbf{A}^{c_{L-1}} & \mathbf{A}^{c_2} & \mathbf{A}^{c_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{A}^{c_2} & \mathbf{A}^{c_3} \\ \vdots & \ddots & \ddots & \mathbf{A}^{c_L} & \mathbf{A}^{c_1} & \mathbf{A}^{c_2} \\ \mathbf{A}^{c_2} & \cdots & \cdots & \mathbf{A}^{c_{L-1}} & \mathbf{A}^{c_L} & \mathbf{A}^{c_1} \end{bmatrix} \tag{7}$$

The diagonal Toeplitz structure block sparse matrix $\mathbf{\Phi}$ and diagonal circulant structure block sparse matrix $\mathbf{\Phi}'$ are shown in Fig. 1, the block matrix is a 30×40 sparse Toeplitz matrix.



(a) Distribution diagram of diagonal Toeplitz structure block sparse matrix

(b) Distribution diagram of diagonal circulant structure block sparse matrix

Fig. 1. The distribution diagram of proposed matrices

3 RIP Conditions of the Proposed Matrix

3.1 Proof Method of Matrix RIP Condition Based on Gerschgorin

To begin with, the paper needs to clarify the definition for any matrix $\mathbf{\Phi}$ that satisfies RIP:

Definition 2: In terms of random signal $\mathbf{x} \in \mathbb{R}^n$, there is only a small number of nonzero elements (less than $K \in \mathbb{N}^+$) in $\mathbf{x}$. If $\exists \delta_K \in (0,1)$, and δ_K, $\mathbf{\Phi}$ and $\mathbf{x}$ satisfy the following conditions:

$$(1-\delta_K)\|\mathbf{x}\|_2^2 \le \|\mathbf{\Phi x}\|_2^2 \le (1+\delta_K)\|\mathbf{x}\|_2^2 \tag{8}$$

Then $\mathbf{\Phi}$ satisfies the K-RIP condition, which means that $\mathbf{\Phi}$ satisfies RIP(K,δ_K). However, there is an NP-hard problem to prove whether any matrix can satisfy the RIP conditions according to the definition above. So, this definition cannot be directly used to guide the construction of the measurement matrix. Some researchers have proved that certain reasonably designed Toeplitz matrices can meet the RIP conditions with a high probability. Reference [13] and [14] used the Gerschgorin theorem to prove that two Toeplitz matrices like equation (2) and equation (6) generated by Gauss random variables can satiny the RIP conditions with a high probability. On this basis, the Gerschgorin theorem is introduced to prove whether the matrix satisfies the conditions.

The Gerschgorin method proves whether the singular values of K order sub-matrices in $\mathbf{\Phi}$ are all within the interval $(\sqrt{1-\delta_K}, \sqrt{1+\delta_K})$ to decide whether the matrix $\mathbf{\Phi}$ meets the RIP conditions. It is the equivalent to saying $\mathbf{\Phi}$ satisfy the RIP conditions when the diagonal elements of the Gram matrix are close to 1 with a high probability, and the diagonal elements are highly restricted. The mathematical description is **Lemma 1**:

Lemma 1: $\exists \delta_d, \delta_s, \delta_K \in (0,1)$, if $\delta_K = \delta_d + \delta_s$, the Gram matrix of $\mathbf{\Phi}$ is $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi}$, the element $G_{i,i}$ on the diagonal of $\mathbf{G}$ meets $|G_{i,i} - 1| < \delta_d$, and other elements $G_{i,j} = \phi_i \phi_j$ satisfies $|G_{i,j}| < \delta_s / K$, $i \ne j$, then Φ satisfy RIP(K,δ_K).

The **Lemma 1** shows a basic idea to prove whether the matrix can satisfy the RIP conditions. However, it is still not enough to tackle with the problem. Then the paper introduces three more lemmas [15]:

Lemma 2: $\{x_1, x_2, ..., x_k\}$ is a bounded independent and identically distributed random sequence with an expectation of 0 and a variance of σ^2, the variable x_i satisfies $|x_i| \le \alpha$, and $\mathrm{E}[x_i^2] = \sigma^2$. When $t \ge 0$, there is

$$\Pr\left(\left|\sum_{i=1}^{k} x_i^2 - k\sigma^2\right| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{k\alpha^4}\right) \tag{9}$$

$\Pr(\bullet)$ stands for the probability in this part.

Lemma 3: $\{x_1, x_2, ..., x_k\}$ and $\{y_1, y_2, ..., y_k\}$ are all bounded independent and identically distributed random sequences with mean value of 0. Variables x_i and y_i satisfies $|x_i| \leq \alpha$ and $|y_i| \leq \alpha$. $|x_i y_i| \leq \alpha^2$. If $t \geq 0$, there are

$$\Pr\left(\left|\sum_{i=1}^{k} x_i y_i\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2k\alpha^4}\right) \tag{10}$$

Lemma 4: If $\exists m, n \in \mathbb{Z}$ and $m \geq 1$, $n \geq 3$, then when the value of the element in each column of the $m \times n$-dimensional matrix is unique and different from each other, the non-diagonal element $G_{i,j}$ ($i \neq j$) in the Gram matrix **G** of **Φ** can be expressed at least as the form of the summation of 3 irrelevant elements.

It can decompose any matrix **Φ** corresponding to the non-diagonal element $G_{i,j}$ of the Gram matrix **G** into the form of a summation sequence of incoherent random variables applying **Lemma** 4. Then use **Lemma** 2 and **Lemma** 3 to prove the diagonal elements $G_{i,i}$ is approximately equal to 1 and the maximum probability $G_{i,j}$ is approximately equal to zero. The flow of the matrix based on the Gerschgorin method that satisfies the RIP conditions is shown in Fig. 2:

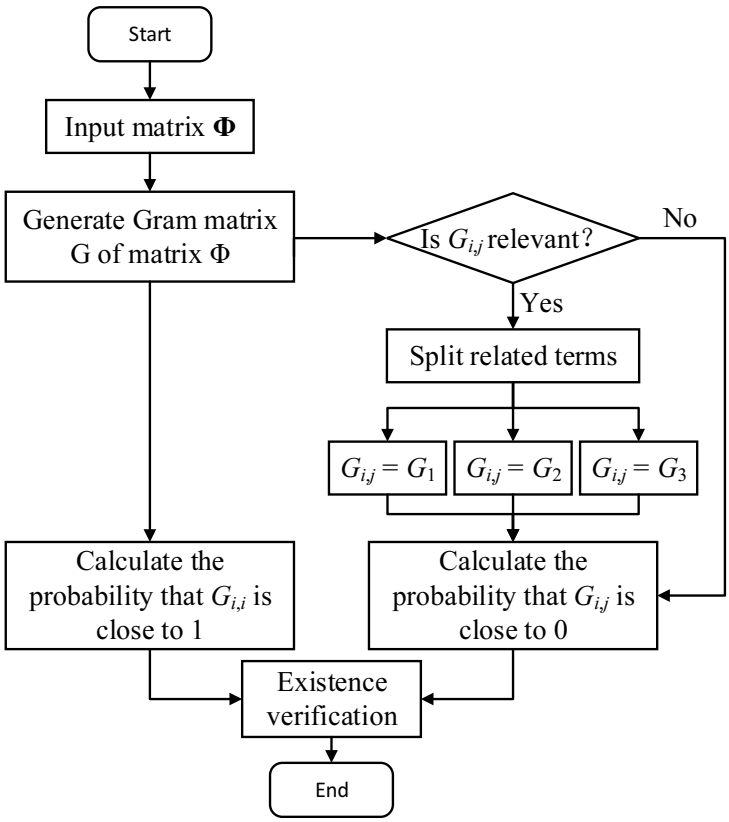


Fig. 2. Flow chart of proof that the matrix satisfies the RIP conditions

3.2 Matrix RIP Proving Conditions

The proposed matrix has an inner and outer double-layer structure. The outer structure is the same as the Toeplitz matrix shown in formula (2). Its characteristic is that except for the first row and first column of the block matrix, the rest of the block matrix is always the same as the upper left square matrix. ; The inner structure is the structure of the block matrix, which is obtained by performing c cyclic shifts on the sparse Toeplitz matrix. The block matrix has an obviously discontinuous sparse non-zero element band structure. The diagonal form of the cyclic structure block sparse matrix $\mathbf{\Phi}'$ is a direct extension of $\mathbf{\Phi}$, so the result of the proof of $\mathbf{\Phi}$ can be directly applied to $\mathbf{\Phi}'$. The proof conclusion that the proposed diagonal Toeplitz structure block sparse matrix satisfies the RIP condition is expressed by **Lemma 5**.

Lemma 5: A random sequence $\{x_1, x_2, ..., x_H\}$ be a random Bernoulli sequence with expectation 0 and variance $1/HW$, where the element x_h in $\{x_1, x_2, ..., x_H\}$ is bounded, $h = 1, 2, ..., H$, and satisfies $|x_h| \le 1/\sqrt{HW}$. With this sequence, a $M \times N$ -dimensional diagonal Toeplitz structured block sparse matrix with $W \times L$ $m \times n$ -dimensional block matrices as shown in equation (10) is generated. So $\exists \delta_K \in (0,1)$, when $K \le C_2 \sqrt{\frac{H}{\ln(NW)}}$, the probability that $\mathbf{\Phi}$ satisfies RIP(K, δ_K) is higher than $1 - \exp(-C_1 H / K^2)$. The constants C_1 and C_2 are only related to δ_K, and $C_1 < \delta_K^2 / 128$, $C_2 \le \sqrt{\delta_K^2 - 128C_1}/16$.

The proof process of **lemma 5** is as follows:

Firstly, it is proved that the diagonal element $G_{i,i}$ of Gram matrix $\mathbf{G} = \mathbf{\Phi}^T \mathbf{\Phi}$ can be close to 1 with high probability, that is $|G_{i,i} - 1| < \delta_d$, $G_{i,i} = \langle \phi_i, \phi_i \rangle$, $i = 1, 2, ..., N$. ϕ_i is the i th column vectors in $\mathbf{\Phi}$.

Let the mean of column weight of $\mathbf{\Phi}$ be l, $W \le l \le M$, and the mean of column weight of block matrix $\mathbf{A}_{p,q}$ be l', $1 \le l' \le H$, $p = 1, 2, ..., W$, $q = 1, 2, ..., L$. The column weight represents the number of non-zero elements in the column vector of the matrix. There are l' independent elements in the block matrix $\mathbf{A}_{p,q}$. However, l non-zero elements in each column of $\mathbf{\Phi}$ can't guarantee their independence. **Lemma 2** and **Lemma 3** require that the variables in the sequence are independent of each other, so the correlation of non-zero elements in the column vector of block matrix is calculated. $G_{i,i} = \sum_{r}^{l} x_h^r$, x_h^r is the r th non-zero element in ϕ_i. From **lemma 2**, it is concluded that:

$$\begin{aligned} \Pr\left(\left| \sum_{r=1}^{l'} \left(x_h^r\right)^2 - \frac{l'}{HW} \right| \ge t \right) &= \Pr\left(\sum_{r=1}^{l'} \left(x_h^r\right)^2 \ge \frac{l'}{HW} + t \text{ or } \sum_{j=1}^{l'} \left(x_h^r\right)^2 \le \frac{l'}{HW} - t \right) \\ &\le 2\exp\left(-\frac{2H^2W^2}{l'} t^2 \right) \end{aligned} \tag{14}$$

Relaxing the $\Pr\left(\left| \sum_{r=1}^{l'} \left(x_h^r\right)^2 - \frac{l'}{HW} \right| \ge t \right)$ in equation (14), then:

$$\Pr\left(\left| \sum_{r=1}^{l'} \left(x_h^r\right)^2 - \frac{1}{W} \right| \ge -\frac{l'}{HW} + \frac{1}{W} + t \right) = \Pr\left(\sum_{r=1}^{l'} \left(x_h^r\right)^2 \ge -\frac{l'}{HW} + \frac{2}{W} + t \text{ or } \sum_{r=1}^{l'} \left(x_h^r\right)^2 \le \frac{l'}{HW} - t \right) \tag{15}$$

According to equation (14) and (15), the size relationship between $\Pr\left(\left| \sum_{r=1}^{l'} \left(x_h^r\right)^2 - \frac{l'}{HW} \right| \ge t \right)$ and $\Pr\left(\left| \sum_{r=1}^{l'} \left(x_h^r\right)^2 - \frac{1}{W} \right| \ge -\frac{l'}{HW} + \frac{1}{W} + t \right)$ can be determined by the range of $\sum_{r=1}^{l'} \left(x_h^r\right)^2$. Because of $1 \le l' \le H$

and $\frac{1}{W} \geq \frac{l'}{HW}$, then:

$$\begin{aligned} \Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2 - \frac{1}{W}\right| \geq -\frac{l'}{HW} + \frac{1}{W} + t\right) &\leq \Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2 - \frac{l'}{HW}\right| \geq t\right) \\ &\leq 2\exp\left(-\frac{2H^2W^2}{l'}t^2\right) \end{aligned} \tag{16}$$

It is obvious that $2\exp\left(-\frac{2t^2H^2W^2}{l'}\right)$ in equation (16) decreases monotonically. Set $-\frac{l'}{HW} + \frac{1}{W} + t = \frac{\delta_d}{W}$, the equation (17) can be scaling to:

$$\begin{aligned} \Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2 - \frac{1}{W}\right| \geq \frac{\delta_d}{W}\right) &\leq 2\exp\left[-\frac{2H^2W^2}{l'}\left(\frac{\delta_d}{W} - \frac{1}{W} + \frac{l'}{HW}\right)^2\right] \\ &\leq 2\exp\left[-\frac{2H^2}{l'}\left(\delta_d - 1\right)^2\right] \end{aligned} \tag{17}$$

Set $\delta_0 \in (0,1)$ and $\delta_d + \delta_0 = 1$, then substituting δ_0 into equation (17):

$$\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2 - \frac{1}{W}\right| \geq \frac{\delta_d}{W}\right) \leq 2\exp\left(-\frac{2H^2\delta_0^2}{l'}\right) \tag{18}$$

According to equation (18), it can be concluded that each diagonal element $G_{i,i}$ satisfies the following conditions:

$$\Pr\left(\left|\sum_{r=1}^{l}\left(x_h^r\right)^2 - 1\right| \geq \delta_d\right) \leq W\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2 - \frac{1}{W}\right| \geq \frac{\delta_d}{W}\right) \leq 2W\exp\left(-\frac{2H^2\delta_0^2}{l'}\right) \tag{19}$$

The upper bound of the joint distribution probability of all diagonal elements $G_{i,i}$ satisfying $\left|G_{i,i} - 1\right| \geq \delta_d$ is obtained:

$$\Pr\left(\bigcup_{i=1}^{N}\left\{\left|G_{i,i} - 1\right| \geq \delta_d\right\}\right) \leq 2N\exp\left(-\frac{2H^2\delta_0^2}{l'}\right) \tag{20}$$

The equation (19) proof that the diagonal elements of Gram matrix $\mathbf{G}$ of proposed matrix $\mathbf{\Phi}$ all close to 1 in very high probability.

Next, proving that the amplitudes of other elements in the Gram matrix $\mathbf{G}$ of $\mathbf{\Phi}$ are limited in a high probability, that is $\left|G_{i,j}\right| < \delta_s / K$, $G_{i,j} = \left\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \right\rangle$, $i, j = 1,2,...,N$, $i \neq j$.

The block matrix $\mathbf{A}_{p,q}$ in $\mathbf{\Phi}$ is a circularly shifted sparse Toeplitz matrix. Due to the shift amount is a random value, the column vectors $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ in $\mathbf{\Phi}$ may have the same element values at the same position. So when calculating $\Pr\left(\left|G_{i,j}\right| \geq \delta_s / K\right)$, three kinds of situations should be considered: ① $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are completely the same; ② The element values of $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ at the same position are the same; ③ $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are completely different. In the above three cases, the off-diagonal elements in the Gram matrix $\mathbf{G}$ of $\mathbf{\Phi}$ are denoted as $G_{i,j}^1$, $G_{i,j}^2$ and $G_{i,j}^3$ respectively.

Situation 1: When $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are completely the same, $G_{i,j}^1 = \left\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \right\rangle = \sum_{r=1}^{l}\left(x_h^r\right)^2$. In this case, it can

apply **lemma 2** to get the following results:

$$\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2-\frac{l'}{HW}\right|\geq t\right)\leq 2\exp\left(-\frac{2H^2W^2}{l'}t^2\right) \tag{21}$$

Since the right term of the inequality in equation (21) is monotonically decreasing and $-\frac{l'}{HW}-t\leq\frac{l'}{HW}-t$, then substituting into equation (21) :

$$\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2\right|\geq\frac{l'}{HW}+t\right)\leq\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2-\frac{l'}{HW}\right|\geq t\right)\leq 2\exp\left(-\frac{2H^2W^2}{l'}t^2\right) \tag{22}$$

Set $\frac{l'}{HW}+t=\frac{\delta_0}{WK}$ and substitute into equation (22):

$$\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2\right|\geq\frac{\delta_0}{WK}\right)\leq 2\exp\left[-\frac{2H^2W^2}{l'}\left(\frac{\delta_0}{WK}-\frac{l'}{HW}\right)^2\right] \tag{23}$$

Set $l'\geq 2\delta_s H$, then $\left(\frac{\delta_2}{DS}\right)^2\leq\left(\frac{\delta_0}{DS}-\frac{l'}{HW}\right)^2$ and substitute into equation (23):

$$\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{WK}\right)\leq 2\exp\left(-\frac{2H^2\delta_s^2}{l'K^2}\right) \tag{24}$$

Therefore, in the first situation, the upper bound of the joint distribution probability of all non-diagonal elements $G_{i,j}$ in $\mathbf{G}$ satisfying $\left|G_{i,j}\right|\leq\delta_s/K$ is as follows:

$$\Pr\left(\left|\mathbf{G}_{i,j}^1\right|\geq\frac{\delta_s}{K}\right)\leq W\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{WK}\right)\leq 2W\exp\left(-\frac{2H^2\delta_s^2}{l'K^2}\right) \tag{25}$$

Situation 2: When the values of elements at the same position in $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are different, then $G_{i,j}=\sum_l x_e x_f$, $1\leq e\leq f\leq H$. However, the elements in $\sum_{l'}x_e x_f$ are not necessarily independent of each other, so it is necessary to divide $\sum_{l'}x_e x_f$ into two independent parts t_1 and t_2. There are only two cases in t_1 and t_2, one is the same quantity, the other is one quantity difference. Let $t_1+t_2=l'$ and $t_1\leq t_2<l'$, by applying **lemma 2**, it can get the following results:

$$\Pr\left(\left|G_{i,j}^2\right|\geq\frac{\delta_s}{K}\right)\leq W\Pr\left(\left|\sum_l x_e x_f\right|\geq\frac{\delta_s}{K}\right)\leq W\Pr\left(\left|\sum_l x_e x_f\right|\geq\frac{\delta_s}{WK}\right) \tag{26}$$

Because of $t_1+t_2=l'$, the $\Pr\left(\left|G_{i,j}^2\right|\geq\frac{\delta_s}{K}\right)$ can switch to:

$$\Pr\left(\left|G_{i,j}^2\right|\geq\frac{\delta_s}{K}\right)\leq 4W\exp\left(-\frac{H^2\delta_s^2}{8K^2l'}\right) \tag{27}$$

Situation 3: When the values of elements in $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are partly the same, then $G_{i,j}=\sum_{l_1}x_e x_f+\sum_{r=1}^{l_2}\left(x_h^r\right)^2$, $1\leq e<f\leq H$ and $l_1+l_2=l$. Set $W_1+W_2=W$ and $W_1\leq W_2<W$, $l_1\in\left[1,mW_1\right)$, $l_2\in\left[1,mW_2\right)$, then:

$$\begin{aligned}\Pr\left(\left|G_{i,j}^3\right|\geq\frac{\delta_s}{K}\right)&=\Pr\left(\left|\sum_{l_1}x_e x_f+\sum_{r=1}^{l_2}x_h^r\right|\geq\frac{\delta_s}{K}\right)\\&\leq 2\max\left\{\Pr\left(\left|\sum_{l_1}x_e x_f\right|\geq\frac{\delta_s}{2K}\right),\Pr\left(\left|\sum_{r=1}^{l_2}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{2K}\right)\right\}\end{aligned}\tag{28}$$

It is similar to the proof method of Situation 1 and Situation 2, then $\Pr\left(\left|\sum_{l_1}x_e x_f\right|\geq\frac{\delta_s}{2K}\right)$ in equation (29) satisfy to:

$$\Pr\left(\left|\sum_{l_1}x_e x_f\right|\geq\frac{\delta_s}{2K}\right)\leq W_1\Pr\left(\left|\sum_{l'}x_e x_f\right|\geq\frac{\delta_s}{2W_1K}\right)\leq 4W_1\exp\left(\frac{H^2W^2\delta_s^2}{32l'W_1^2K^2}\right)\tag{29}$$

$\Pr\left(\left|\sum_{r=1}^{l_2}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{2K}\right)$ satisfy to:

$$\Pr\left(\left|\sum_{r=1}^{l_2}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{2K}\right)\leq W_2\Pr\left(\left|\sum_{r=1}^{l'}\left(x_h^r\right)^2\right|\geq\frac{\delta_s}{2W_2K}\right)\leq W_2\exp\left(-\frac{H^2W^2\delta_s^2}{2W_2^2K^2l'}\right)\tag{30}$$

Then

$$\Pr\left(\left|G_{i,j}^3\right|\geq\frac{\delta_s}{K}\right)<8W\exp\left(-\frac{H^2\delta_s^2}{32l'K^2}\right)\tag{31}$$

Get the sum of $\Pr\left(\left|G_{i,j}^1\right|\geq\frac{\delta_s}{K}\right)$, $\Pr\left(\left|G_{i,j}^2\right|\geq\frac{\delta_s}{K}\right)$ and $\Pr\left(\left|G_{i,j}^3\right|\geq\frac{\delta_s}{K}\right)$, then get $\Pr\left(\left|G_{i,j}\right|\geq\frac{\delta_s}{K}\right)$:

$$\Pr\left(\left|G_{i,j}\right|\geq\frac{\delta_s}{K}\right)\leq 14W\exp\left(-\frac{H^2\delta_s^2}{32l'K^2}\right)\tag{32}$$

The elements in **G** satisfy to $G_{i,j}=G_{j,i}$. Therefore, the number of different non-diagonal elements $G_{i,j}$ is $\left(N^2-N\right)/2<N^2/2$. The upper bound of joint probability of $G_{i,j}$ can be obtained by substituting $\left(N^2-N\right)/2<N^2/2$ into equation (32):

$$\Pr\left(\bigcup_{i=1}^{N}\bigcup_{\substack{j=1\\j\neq i}}^{N}\left\{\left|G_{i,j}\right|\geq\frac{\delta_s}{K}\right\}\right)\leq 7N^2W\exp\left(-\frac{H^2\delta_s^2}{32l'K^2}\right)\tag{33}$$

Equation (34) proves that the high probability of amplitude of elements on non-diagonal lines in **G** is limited. By combining equations (20) and (33), substituting δ_0 to δ_s and $\delta_d=\delta_s=\delta_0=\delta_K/2$ and making reasonable scaling. Then combining with the conditions set in **lemma 4**, it can obtain the probability that the proposed matrix does not meet the rip condition $\Pr(\text{Non}-\text{RIP})$ is:

$$\Pr(\text{Non}-\text{RIP})\leq 8N^2W\exp\left(-\frac{H\delta_K^2}{64K^2}\right)\tag{34}$$

For any $C_1<\frac{\delta_K^2}{64}$, $C_2\leq\sqrt{\delta_K^2-64C_1}/16$, the upper bound of the probability that **Φ** does not satisfy RIP condition is obtained by substituting C_1 and C_2 into equation (34):

$$\Pr\left(\text{Non}-\text{RIP}\right) \le \exp\left(-\frac{C_1 H}{K^2}\right) \tag{35}$$

K has to satisfy that:

$$K \le C_2 \sqrt{\frac{H}{\ln\left(NW\right)}} \tag{36}$$

The proof is complete. In conclusion, the proposed matrix $\mathbf{\Phi}$ satisfy the RIP condition with the probability of $1-\exp\left(-C_1 H/K^2\right)$.

4 Performance Analysis and Simulation Experiments for the Matrix

This Section will verify the performance of matrices through simulation experiments. It starts with the analysis of restricted isometry phenomenon. Fig. 3 shows a Gram matrix elements distribution diagram, in which the matrix $\mathbf{\Phi}$ is a 240×320 dimensional matrix with the size of the middle block matrix is 30×40, and the number of nonzero elements in the row vectors of $\mathbf{\Phi}$ is $w=40$. The axis Z element represents the probability of the element $G_{i,j}$ in $\mathbf{G}$ infinitely close to 1.

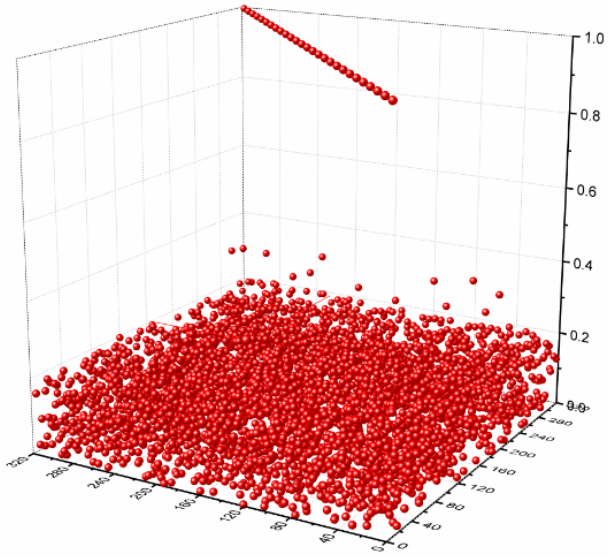


Fig. 3. Gram matrix elemental distribution diagram

It can be seen that the diagonal elements of the Gram matrix that proposed matrix $\mathbf{G}$ converge at 1, and the non-diagonal elements all fluctuate around 1. So it is fair to say that the matrix satisfies the RIP. Then, compressed sampling is used towards one-dimensional sparse signals and image signals to prove the sampling property of the matrix. The measurement matrix used in the comparison experiment is the Gauss random matrix at its best and the scrambling Hadamard matrix that has the best properties in structural matrices. The experiment used a laptop computer with a 64-bit Win7 operation system, Inter Core i7 2.8HZ CPU, and a 16GB RAM.

4.1 One-dimensional Sparse Signal Experiment

A one-dimensional signal with a sparsity of was generated as the experiment signal. The measurement rate is $1/4$, which means that the number of the sampling $M=240$. gradually increased at the step length of 5, increasing from 20 gradually to 70. Its nonzero elements obey the Gaussian distribution $N\left(0,1/2\right)$. The reconstruction algorithm applies Gradient Projection for Sparse Reconstruction [16]. When the reconstruction of Mean Square Error (MSE) satisfies the following condition:

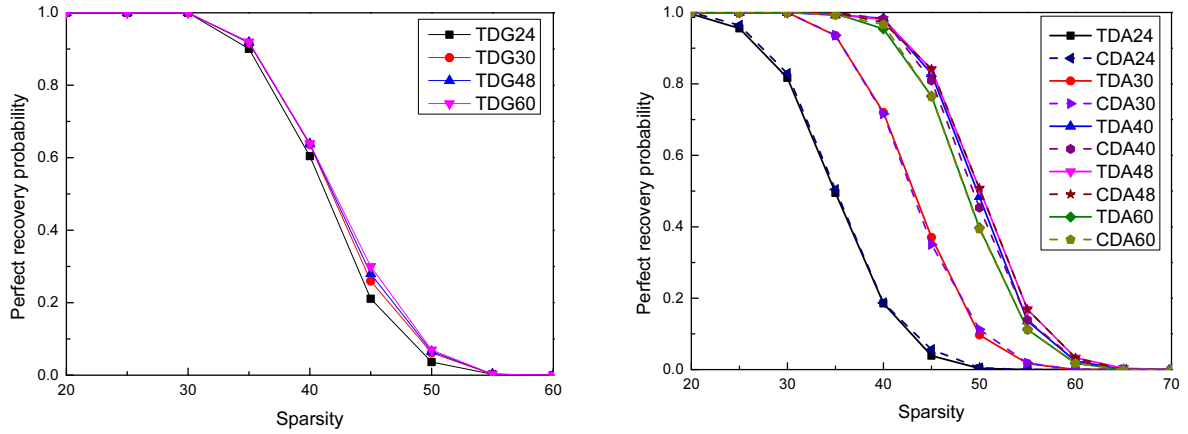
$$\text{MSE} = \frac{\left\|\mathbf{x} - \hat{\mathbf{x}}\right\|_2^2}{\left\|\mathbf{x}\right\|_2^2} \le 0.1 \tag{37}$$

So, the signal can be defined as a perfect recovery, and $\hat{\mathbf{x}}$ is the reconstruction signal. Using Gauss random matrix as the comparison measurement matrix, three types of block structures, including non-structure, Toeplitz structure, and circulant structure were chosen to endure the stability of the experiment result. Each type was conducted for 100 times and then got the average value to have the probability of accurate reconstruction. For convenience, Table 1 aims to show a simplified naming method of the experimented matrix.

Table 1. Matrix naming rule table

Matrix Type / Matrix Structure	Gauss Matrix/G	Sparse Cyclic shifting Toeplitz Matrix/A
Non-structure/N	NG	NA
Diagonal Toeplitz structure /TD	TDG	TDA
Diagonal Circulant structure /CD	CDG	CDA

Where 6 different sizes 16×16, 24×24, 30×30, 40×40, 48×48 and 60×60 were chosen, and block Gauss Random matrix and the proposed matrix were used to reconstructing $\mathbf{x}$, the results are as the Fig. 4(a) and Fig. 3(b).



(a) Correlations between block sizes of Gauss Random matrix and signal sparsity

(b) Correlations between Diagonal Toeplitz/Circulant block structure sizes and sparsity

Fig 4. The reconstruction results of one-dimensional signals by the different matrices

The performance curves of TDG24, TDG30, TDG48 and TDG60 almost coincide. The four curves have the consistent trends and same accurate reconstruction rates. So it proves that the sizes of the block matrices have little effect on Gauss matrix performance. When the sizes of the block matrices are the same, property curves of TDA and CDA coincide, which means that the block matrix properties of Toeplitz structure and Circulant structure are similar. In the real engineering context, a feasible circulant structure can be used and performance will not change. Therefore, the following discussion will only focus on the TD matrix. The size of the block matrix has a significant effect on the properties of TDA. When the size increases, the sparsity increases and the property improves accordingly, From Fig. 3(a) and Fig. 3(b), the property of TDA24 is not as good as that of GAU. TDA30 has better properties not only compared to TDA 24, but also to TDG30. Similarly, the properties of TDA40, TDA48 and TDA 60 significantly outperform TDA30, TDA24, and TDG with the same block matrix size. It concludes that on the one hand, the influence of matrix structure towards the properties is much high than that of matrix sparsity; on the other hand, when the size of the block matrix is huge, it can be seen that although the property of TDA60 is much higher than that of TDA30, it is a little lower compared to those of TDA40 and TDA48. The reason lies in the sparsity of matrix is huge, which makes the property decrease.

When the block size is fixed to 30×30, the block matrix is set as a sparse Toeplitz matrix, and the number of non-zero elements in each row of all structured matrices is set $w=30$. After 100 tests, the probability of accurate reconstruction is calculated as shown in Fig. 5.

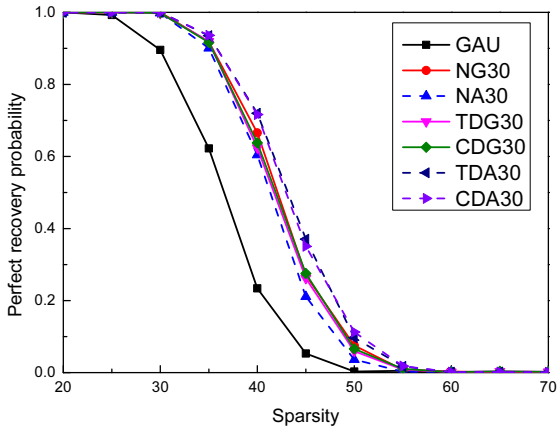


Fig 5. Relationship between signal precision reconstruction rate and sparsity

The results show that the overall reconstruction performance of Gauss matrix is significantly lower than that of other block structures, and the reconstruction performance of Gauss is significantly lower than that of other matrices when the sparsity is 40%, the accuracy of reconstruction rate is nearly 40%. The reason is that the Gauss matrix used in experiment is composed of several rows in the orthogonal Gauss matrix, which has caused some damage to the orthogonality of the original Gauss matrix, However, block Gauss matrix is orthogonal, so it can still keep good column orthogonality after forming block structure matrix. Therefore, ng, TDG and CDG are better than Gauss matrix of Gauss random selection; NG30 curve, tdg40 curve and cdg30 curve almost coincide, which shows that the matrix structure has little influence on the reconstruction performance of the structured matrix with Gauss matrix.

4.2 Image Signal Experiment

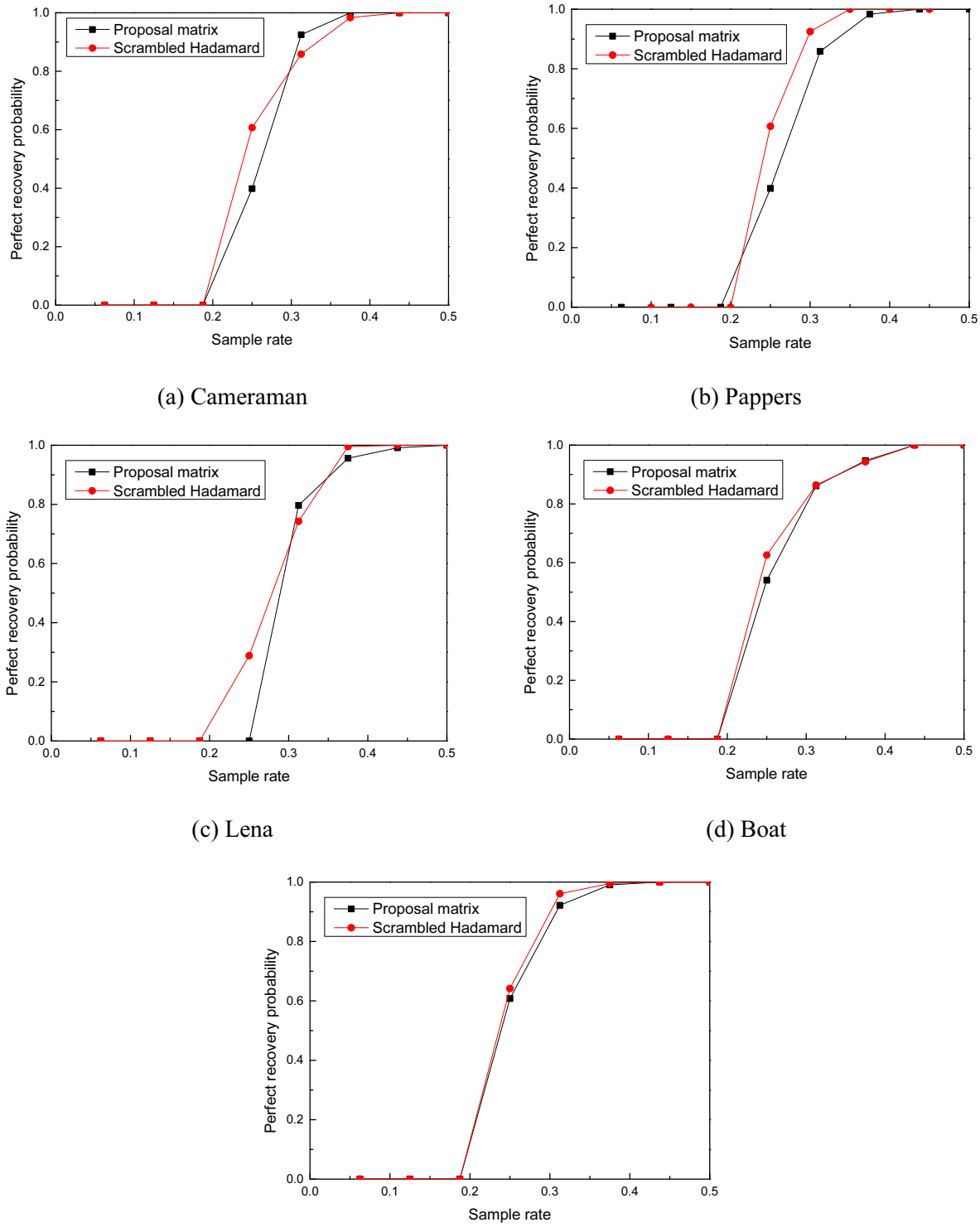
Five 512×512 gray-scale images: Cameraman, Peppers, Lena, Boat and Mandrill are taken as experimental objects to test the performance of measurement matrix, as shown in Fig. 6.

(a) Cameraman (b) Peppers (c) Lena (d) oat (e) Mandril

Fig. 6. Test image

The five images have different characteristics: Peppers contains a lot of smooth region and edge information, but less texture information; Cameraman and Lena also contain a large number of smooth regions, but these two images contain rich edge information and texture information; Boat and Mandrill contain a lot of texture information, but less edge information and smooth area. The minimum total variation (min TV) algorithm based on block compressed sensing is used to reconstruct the image. The measurement rate is set as $\omega=M/N\in\{0.1,0.2,0.3,0.4,0.5\}$ and M as the number of rows of the measurement matrix. The proposed matrix and the scrambled Hadamard matrix proposed in reference [14] are used to reconstruct the experimental image respectively. After 50 times of experiment repetition, the probability of successful reconstruction of the image is calculated. Reference [17] points out that the

change of image block size $B \times B$ will affect the reconstruction quality and computational complexity of the whole image at the same time. With the increase of image block size, the weaker the block effect of the reconstructed image, the higher the reconstruction quality. When $B \geq 32$, the overall reconstruction quality of the image is stable at a high level, but with the continuous increase of image block size, the computational complexity of image reconstruction also increases, the reconstruction efficiency is reduced. Therefore, the image block size is set to 32×32. The experimental results are shown in Fig. 7.



(a) Cameraman

(b) Pappers

(c) Lena

(d) Boat

(e) Mandrill

Fig 7. The accurate reconstruction rate of two measurement matrices for each image under different measurement rates

It can be seen from Fig. 7 that the proposed matrix and scrambled Hadamard matrix are used to measure each image at different measurement rates, and the probability of accurate image reconstruction is basically the same. In most cases, the probability of accurate image reconstruction by proposed matrix is higher than scrambled Hadamard matrix. When the measurement rate is close to 0.4, the image is completely accurately reconstructed. The reconstruction effect of the proposed matrix and the scrambled Hadamard matrix is basically the same, and the performance of the proposed matrix is slightly better than that of the scrambled Hadamard matrix, because the proposed matrix has sparsity and the non-correlation of the matrix column vector is stronger.

The above tests on one-dimensional signals and images show that the proposed diagonal Toeplitz / circulant structure block sparse matrix has obvious advantages in sampling efficiency and recovery effect compared with the classical perceptual matrix, and it is easy to construct and has better application potential.

5 Conclusion

In order to solve the problems of complex process, difficult hardware implementation and high computational complexity of random measurement matrix generation, this paper proposes a diagonal Toeplitz / circulant structure block sparse matrix. The matrix only contains the elements in the set of $\{-1,0,1\}$, in which the non-zero elements are random Bernoulli distribution with probability $1/2$, which greatly simplifies the generation process of Toeplitz matrix and Circulant matrix and makes the matrix very sparse. It is proved theoretically that the proposed matrix can satisfy the rip condition of compressed sensing. Experimental results show that the proposed matrix has lower computational complexity and higher image reconstruction efficiency than the block disordered Hadamard matrix with quasi-Gaussian characteristics under the same measurement rate. In addition, the proposed matrix has obvious advantages in memory overhead, computational complexity and hardware implementation. It is proved that the proposed matrix is a universal and high-performance compressed sensing measurement matrix.

References

[1] E. J. Candes, J. Romberg, and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory, 52(2)(2006) 489-509.

[2] D. L. Donoho, Compressed sensing. IEEE Transactions on Information Theory, 52(4)(2006) 1289-1306.

[3] D. L. Donoho, Y. Tsaig, Extensions of compressed sensing. Signal Processing, 86(3)(2006) 533-548.

[4] Y. Cen, D. Wang, D. Zhang, Measurement Matrix Construtionvia Polynomial Graphs in the Finite Field for Compressed Sensing, Journal of Qiannan Normal University for Nationalities, 37(4)(2017) 1-8.

[5] Y. Guo, C. Wang, S. Du, A compressed sensing measurement matrix construction algorithm based on generalized variable parameter Fibonacci chaotic system, Computer Engineering & Science, 43(03)(2021) 503-510.

[6] F. Du, R. Ye, B. Yan, Application of a PEG-based Construction Matrix in Compressed Sensing, Radio Communications Technology, 46(4)(2020) 465-470.

[7] S. Yin, Research on the construction method and application of the compressed sensing measurement matrix, Yanshan University, 2017

[8] J. Guo, J. Dang, Data compression collecting method for vibration signals based on optimal deterministic measurement matrix, Journal of Vibration and Shock, 38(7)(2019) 195-203.

[9] Y. Zhu, W. Liu, Q. Shen, An Adaptive Block Compressed Sensing Algorithm Based on Saliency, Telecom Technology, 26(12)(2019) 28-33.

[10] W. Du, C. Ma, Q. Wang, Research on Image Reconstruction Algorithm Based on Optimized Hadamard Matrix, Journal of Chongqing University of Science and Technology: Natural Sciences Edition, 22(1)(2020) 85-88.

[11] H. Gan, T. Zhang, Y. Hua, Toeplitz-block sensing matrix based on bipolar chaotic sequence, Acta Physica Sinica, 70(3)(2021) 279-290.

[12] H. Cao, T. Sun, H. Wang, Generalized rotation measurement matrix and its application in underwater echo signals, Technical Acoustics, 2019(6) 623-628.

[13] L. Quan, S. Xiao, X. Xue, Fast sensing method in compressive sensing with low complexity, Journal of Xidian University, 44(1)(2017) 106-111.

[14] W. Du, C. Ma, Q. Wang, Research on Image Reconstruction Algorithm Based on Optimized Hadamard Matrix, Journal of Chongqing University of Science and Technology: Natural Sciences Edition, 22(1)(2020) 85-88.

[15] Holger Rauhut. Circulant and Toeplitz Matrices in Compressed Sensing, Mathematics, 2009.

[16] M.A.T. Figueiredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction[J] · IEEE Journal of Selected Topics in Signal Processing, 1(4)(2007) 586-597.

[17] M. Wang, A design and research on the compressed sensing measurement matrix, Ph.D thesis, Xidian University in Xi`an, 2014.<http://www.cancerresearchuk.org/aboutcancer/statistics/cancerstatsreport/>, 2003 (accessed 13.03.03).