Compressive Sensing with Chaotic Sequence

Lei YU, Jean Pierre BARBOT, Gang ZHENG, and Hong SUN,

Abstract—Compressive sensing is a new methodology to capture signals at sub-Nyquist rate. To guarantee exact recovery from compressed measurements, one should choose specific matrix, which satisfies the Restricted Isometry Property (RIP), to implement the sensing procedure. In this letter, we propose to construct the sensing matrix with chaotic sequence following a trivial method and prove that with overwhelming probability, the RIP of this kind of matrix is guaranteed. Meanwhile, its experimental comparisons with Gaussian random matrix, Bernoulli random matrix and sparse matrix are carried out and show that the performances among these sensing matrix are almost equal.

Index Terms—Compressive Sensing, Chaos, Logistic Map.

I. INTRODUCTION

Over the recent years, a new sampling theory, called Compressive Sensing [9], [10], [11] (CS for short), has attracted lots of researchers. The central goal of CS is to capture attributes of a signal using very few measurements: for any $N$-dimensional signal $\mathbf{v}$ (w.l.g. $\mathbf{v}$ is $s$-sparse vector), the measurement $\mathbf{y} \in \mathbb{R}^M$ is captured through $\Phi \mathbf{v}$, where $s < M < N$ and $\Phi \in \mathbb{R}^{M \times N}$ is a well chosen matrix satisfying the Restricted Isometry Property (RIP)[8].

Definition 1.1: Matrix $\Phi \in \mathbb{R}^{m \times n}$ satisfies the Restricted Isometry Property of order $s$ if there exists a constant $\delta \in [0, 1]$ such that

$$
(1 - \delta)\|\mathbf{v}\|_2^2 \leq \|\Phi \mathbf{v}\|_2^2 \leq (1 + \delta)\|\mathbf{v}\|_2^2
$$

(1)

for all $s$-sparse vectors $\mathbf{v}$.

In CS framework, finding a proper sensing matrix $\Phi$ satisfying RIP is one of the central problems. Candès and Tao have proposed that matrix with elements drawn by Gaussian distribution or Bernoulli distribution satisfies RIP with overwhelming probability, providing that sparsity $s \leq O(M/\log N)$[10]. And the randomly selected Fourier basis also retains RIP with overwhelming probability with sparsity $s \leq O(M/(\log N)^6)$[10]. On the other hand, many researchers have employed some other techniques to construct deterministic sensing matrix: one group satisfying Statistical Isometry Property (SIRIP) [7], such as Chirp Sensing Codes, second order Reed-Muller code, BCH code by R. Calderbank et. al [2], [14], [7]; one group satisfying RIP-1 [5], such as sparse random matrix by P. Indyk et. al [6] and LDPC by D. Baron et. al [4]; one group satisfying deterministic RIP, such as finite fields by R. A. Devore [12]; etc.

In this paper, we employ chaotic sequence to construct the sensing matrix, called chaotic matrix. Comparing to other techniques, chaotic system generates the “pseudo-random” matrix in deterministic approach and hence verifies the RIP similar to Gaussian or Bernoulli matrix. Moreover, it is easy to be implemented in physical electric circuit and only one initial state is necessary to be memorized. Based on the statistical property of chaotic sequence, it is shown that chaotic matrix satisfies RIP with overwhelming probability, providing that $s \leq O(M/\log N)$.

The main contribution of this paper is to make a connection between chaotic sequence and CS. It is shown by the experiments that the performance of chaotic matrix is somewhat equal to the famous Gaussian random matrix and sparse random matrix. The paper is organized as below. In section II, one chaotic system is recalled and its statistical property is presented. Section III shows the construction of chaotic matrix and proves its RIP. In section IV, experiments are carried out to simulate the performance of chaotic matrix. At the end, the conclusion is given.

II. CHAOTIC SEQUENCE AND ITS STATISTICAL PROPERTY

Let us consider the following quadratic recurrence equation

$$
z_{n+1} = rz_n(1 - z_n)
$$

(2)

where $r$ is a positive constant sometimes known as the “biotic potential” giving the so-called Logistic map. For the special case $r = 4$, the solution for system (2) can be written as below [17]:

$$
z_n = \frac{1}{2}[1 - \cos(2\pi \theta 2^n)]
$$

(3)

where $\theta \in [0, \pi]$ satisfying $z_0 = \frac{1}{2}[1 - \cos(2\pi \theta)]$ with $z_0$ the initial condition of (2). It is well known that chaotic system (2) can produce very complex sequences. Even more, it is often used as the random number generator in practice since (2) takes a very simple dynamics [17]. In this section, we will analyze its statistical properties, the distribution, the correlations and the sampling distance which guarantees the statistical independence.

Denote

$$
x_n = \cos(2\pi \theta 2^n)
$$

(4)

obviously, $z_n$ takes the similar statistical property with $x_n$ since the linear transformation, for instance the fact that $x_n$ and $x_m$ are statistically independent, would result in $z_n$ and $z_m$ are statistically independent.
A. Distribution

Function (4) possesses the following features: zero mean, values bounded within the interval \(-1 \leq x_n \leq 1\), invariant density given by 

\[ \rho(x) = \frac{1}{2} (1 - x^2)^{-1/2} \]

and \( x_0 = \cos(2\pi\theta) \).

B. Correlations

It can be checked that the \( m \)-th moment of \( x_n \) satisfies

\[ E(x_n^m) = 0 \text{ if } m \text{ is odd and} \]

\[ E(x_n^m) = 2^{-m} \left( \frac{m}{2} \right) \]

if \( m \) is even.

C. Statistical independence

In [19], it has been proved that sequence generated by (4) is not independent. However, we can measure its independence through the high order correlations, which is determined by the sampling distance. We have the following lemma.

Lemma 2.1: Denote \( X = \{x_n, x_{n+1}, \ldots, x_{n+k}, \ldots\} \) the sequence generated by (4) with initial state \( x_0 = \cos(2\pi\theta) \), and integer \( d \) the sampling distance, then for any positive integer \( m_0, m_1 < 2^n \), it has

\[ E(x_{n+m_0}^m x_{n+d+m_1}^m) = E(x_{n+m_0}^m)E(x_{n+d+m_1}^m) \]  

(6)

Proof: If there exists at least one odd number in \( m_0, m_1 \), the right side of (6) is equal to 0. For the left side, we have

\[ E(x_{n+m_0}^m x_{n+d+m_1}^m) = \int_{-1}^{1} \rho(x_0) x_{n+m_0}^m x_{n+m_1}^m dx_0 \]

\[ = \int_{0}^{1} \cos^{m_0} (2\pi \theta^n) \cos^{m_1} (2\pi \theta^{n+d}) d\theta \]

\[ = \frac{1}{2^{m_0+m_1} n} \sum \sigma \left( 2^n \sum_{i=1}^{m_0} \sigma_i + 2^{n+d} \sum_{i=1}^{m_1} \sigma_{i+n+d} \right) \]

where the last equation uses the fact that \( \cos \theta = (e^{i\theta} + e^{-i\theta})/2 \) and \( \int_{0}^{1} e^{2\pi i k \theta} d\theta = \delta(k) \), with \( \delta(k) = 0 \) if \( k \neq 0 \) otherwise equals to 1. \( \sum \sigma \) is the summation over all possible configurations, where \( \sigma_i = \pm 1 \) and \( \sigma_{i+n+d} = \pm 1 \).

All possible cases for \( m_0 \) and \( m_1 \) are as below:

1) \( m_1 \) is odd: \( \sum_{i=1}^{m_0} \sigma_i \leq m_0 \) and \( \sum_{i=1}^{m_1} \sigma_{i+n+d} \geq 1 \), hence \( 2^n \sum_{i=1}^{m_0} \sigma_i + 2^{n+d} \sum_{i=1}^{m_1} \sigma_{i+n+d} \neq 0 \);

2) \( m_1 \) is even: it is possible that \( \sum_{i=1}^{m_1} \sigma_{i+n+d} = 0 \), while since \( m_0 \) is odd (the assumption at the beginning of this part of proof), \( \sum_{i=1}^{m_0} \sigma_i \neq 0 \), hence \( 2^n \sum_{i=1}^{m_0} \sigma_i + 2^{n+d} \sum_{i=1}^{m_1} \sigma_{i+n+d} \neq 0 \).

Then we can conclude that the left side of (6) is also equal to 0.

If both \( m_0 \) and \( m_1 \) are even numbers, after a trivial combinatorial analysis, we get

\[ E(x_{n+m_0}^m x_{n+d+m_1}^m) = 2^{-m_0-m_1} \left( \frac{m_0}{2} \right) \left( \frac{m_1}{2} \right) \]

Compare it with equation (5), we have (6).

Remark 2.2: Lemma 2.1 implies that \( x_n \) and \( x_{n+d} \) are statistically independent when \( d \to \infty \), and this result corresponds to that given in [18]. Approximately, if the sampling distance is chosen large enough, for instance \( d = 15 \),

\[ E(x_{n+m_0}^m x_{n+d+m_1}^m) = E(x_{n+m_0}^m)E(x_{n+d+m_1}^m) \]

for all \( m_0, m_1 < 32768 \), hence \( x_n \) and \( x_{n+d} \) can be considered approximately independent, as illustrated in Fig. 1.

III. Chaotic Sensing Matrix

Let \( Z(d, k, z_0) = \{z_n, z_{n+d}, \ldots, z_{n+kd}\} \) be the chaotic sequence sampled from the output sequence produced by Logistic map (2) with sampling distance \( d \) and initial condition \( z_0 \), and let \( x_k \in X(d, k; x_0) \) denote the regularization of \( Z(d, k, z_0) \) as below

\[ x_k = 1 - 2z_{n+kd} \]

(7)

where \( X(d, k, x_0) \) just corresponds to equation (4) and hence fulfills the statistical properties discussed in the previous section.

To construct the sensing matrix \( \Phi \in \mathbb{R}^{M \times N} \), generate sampled chaotic sequence \( X(d, k, x_0) \) with length \( k = M \times N \), then create a matrix \( \Phi \) column by column with this sequence, written as

\[ \Phi = \sqrt{\frac{2}{M}} \begin{pmatrix} x_0 & \ldots & x_{M(N-1)} \\ x_1 & \ldots & x_{M(N-1)+1} \\ \vdots & & \vdots \\ x_{M-1} & \ldots & x_{MN-1} \end{pmatrix} \]

(8)

where the scaler \( \sqrt{\frac{2}{M}} \) is for normalization. By chosen sampling distance \( d = 15 \), then elements of sequence \( X(d, k, x_0) \) are approximately independent and satisfy identical distribution \( \rho(x) \), i.e. a.i.i.d, and hence elements of matrix \( \Phi \) are a.i.i.d.
**Theorem 3.1:** Chaotic matrix $\Phi \in \mathbb{R}^{M \times N}$ constructed following (8) satisfies RIP for constant $\delta > 0$ with overwhelming probability, providing that $s \leq O(M / \log(N/s))$.

**Remark 3.2:** Inherently, this matrix $\Phi$ is sub-gaussian with $a.i.i.d$ elements. In [16], A. Pajor et. al have proved that all sub-gaussian matrix verify the RIP from geometrical point of view. In what follows, a brief proof following R. Baraniuk’s idea [3] connecting Johnson-Lindenstrauss property [15], [1] and RIP, is presented. Moreover, we can see what Lemma 2.1 implies for RIP.

Before giving the proof, let us recall a lemma stated in [1].

**Lemma 3.3:** For $h \in [0, 1/2]$,

$$E[\exp(hQ^2)] \approx \frac{1}{\sqrt{1 - 2h}}, \quad E[Q^4] \approx 3$$

where $Q = \langle x, u \rangle$ with $x$ being any row vector of $\Phi$ and $u$ being any unit vector.

**Remark 3.4:** In Lemma 3.3, $\approx$ represents approximately less, which goes to be strictly $\leq$ when sampling distance $d \to \infty$.

**Proof for Theorem 3.1:** The proof contains two parts: first prove the J-L property for any sub-matrix of $\Phi$, then conclude the RIP using permutation theory.

1) **J-L property:**

Denote $\Phi_T$ the arbitrary column sub matrix of $\Phi$, with index set $|T| = s$. For any unit vector $u \in \mathbb{R}^s$, from Chernoff’s inequality, given some positive value $h$, it has

$$\Pr[\|\Phi_T u\|^2 \geq 1 + \delta] \leq \exp(-h(1+\delta))E[\exp(h\|\Phi_T u\|^2)]$$

$$\approx \exp(-h(1+\delta))E[\exp(hQ^2)]^M$$

$$\approx \exp(-h(1+\delta))\left(\frac{1}{\sqrt{1 - 2h}}\right)^M$$

$$\approx \exp\left(-\frac{M}{2} \left(\delta^2/2 - \delta^3/3\right)\right)$$

$$\approx \exp(-c_1(\delta)M)$$

where the last inequality is obtained by Taylor expansion and setting $h = \frac{1}{2(1+\delta)}$, which is the extremum point, and $c_1(\delta) = \delta^2/4 - \delta^3/6$.

Similarly, we can calculate the lower bound of its probability as follows

$$\Pr[\|\Phi_T u\|^2 \leq 1 - \delta]$$

$$\leq \exp(hM(1+\delta))E[\exp(-hM\|\Phi_T u\|^2)]$$

$$\approx \exp(hM(1+\delta))(E[\exp(-hQ^2)])^M$$

$$\approx \exp(hM(1-\delta))\left(1 - h + \frac{3}{2}h^2\right)^M$$

$$\approx \exp(-c_2(\delta)M)$$

where the last inequality is obtained by Taylor expansion and setting $h = h_{opt} = \frac{-2 - \delta + \sqrt{4 + 8\delta - \delta^2}}{4(1-\delta)}$, which is the extremum point, and $c_2(\delta) = h_{opt}(1-\delta)(1 - h_{opt} + 3h_{opt}^2)/2$.

Choose $c(\delta) = \min\{c_1(\delta), c_2(\delta)\}$, then one finally gets

$$\Pr[\|\Phi_T u\|^2 - 1 \geq \delta] \leq 2\exp(-c(\delta)M) \quad (9)$$

2) **RIP:**

For any $s$-sparse vector $v$, denote $T$ the set of locations where elements are nonzero, then $|T| = \kappa \leq s \ll N$. The column sub matrix $\Phi_T$ defined in previous part can be set up and satisfies (9). Let us denote $E_\kappa$ one complementary event of condition in (1), i.e.

$$E_\kappa = \left\{|\|\Phi_T u\|^2 - 1\| \geq \delta\right\}$$

and denote $E$ the union of all possible complementary events, i.e. $E = \bigcup_{\kappa = 1}^{N} E_\kappa$. Then one obtains

$$\Pr[E] = \bigcup_{\kappa = 1}^{N} \Pr[E_\kappa] \leq 2\exp(-c(\delta)M)$$

$$\leq 2s\left(\frac{n}{s}\right)\exp(-c(\delta)M)$$

where, for a fixed constant $c_3 > 0$, whenever $s \leq c_3M/\log(N/s)$, the bound will only have the exponent with the exponential $-c_4M$ provided that $c_4 \leq c(\delta) - c_3[1 + (1 + (\log s)/s)/\log N/s]$. Hence we can choose $c_3$ sufficiently small to ensure that $c_4 > 0$.

Consequently, the probability for satisfying RIP is at least

$$1 - \Pr[E] \approx 1 - 2e^{-c_4M}.$$
Fig. 2. Maximum sparsity for fixed signal size $N = 800$ and variable number of measurements $M \in [100, 500]$ (left) and for variable signal size $N \in [300, 1000]$ and fixed number of measurements $M = 200$ (right).

Fig. 3. Probability of correct recovery for fixed signal size $N = 1000$.

V. CONCLUSION

In this paper, we firstly recall the statistical property of one special chaotic system - Logistic map and prove that the generated sequence is approximately independent with sampling distance large enough (for instance $d = 15$). Then we prove that matrix constructed with this sampled chaotic sequence also satisfies RIP with overwhelming probability. From the experiments, it shows that chaotic matrix has the similar performance to Sparse matrix, Gaussian random matrix and Bernoulli random matrix.

REFERENCES


Fig. 4. Recovery rate for different sensing matrix.

Fig. 5. Recovery rate for different initial conditions for chaotic sequence.