Compressive Sensing with Chaotic Sequence

Lei Yu, Jean-Pierre Barbot, Gang Zheng, Hong Sun

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

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

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

In CS framework, finding a proper sensing matrix Φ 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 $s \le O(M/\log N)$[10]. And the randomly selected Fourier basis also retains RIP with overwhelming probability with sparsity $s \le O(M/(\log N)^6)$[10]. On the other hand, many researchers have employed some other techniques to construct deterministic sensing matrix: one group satisfying Statistical Isometry Property (StRIP) [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 $s \le O(M/\log(N/s))$.

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

$$z_{n+1} = rz_n(1-z_n) \tag{2}$$

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]:

$$z_n = \frac{1}{2}[1-\cos(2\pi\theta 2^n)] \tag{3}$$

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

$$x_n = \cos(2\pi\theta 2^n) \tag{4}$$

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.

This work was supported by "Bourses Doctorales en Alternance", PEPS-A2SDC of INSIS-CNRS, NSFC (No. 60872131) and FEDER through CPER 2007-2013 with INRIA Lille-Nord Europe.

L.Yu and H. Sun are with Signal Processing Laboratory, Electronic and Information School, Wuhan University, 430079, Wuhan, China. L. Yu and J-P. Barbot are with ECS ENSEA, 6 Avenue du Ponceau, 95014 Cergy-Pontoise, France and EPI ALIEN, INRIA, France. G. Zheng is with INRIA Lille-Nord Europe, 40 Avenue Halley, 59650, Villeneuve d'Ascq, France.

Email address: yuleiwhu@gmail.com, barbot@ensea.fr, gang.zheng@inria.fr, hongsun@whu.edu.cn.

A. Distribution

Function (4) possesses the following features: zero mean, values bounded within the interval $-1 \le x_n \le 1$, invariant density given by $\rho(x) = \frac{1}{\pi}(1-x^2)^{-1/2}$ and $x_0 = \cos(2\pi\theta)$.

B. Correlations

It can be checked that the m-th moment of x_n satisfies $E(x_n^m) = 0$ if m is odd and

$$E(x_n^m) = 2^{-m} \begin{pmatrix} m \\ \frac{m}{2} \end{pmatrix} \tag{5}$$

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 $X = \{x_n, x_{n+1}, ..., x_{n+k}, ...\}$ 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 $m_0, m_1 < 2^d$, it has

$$E(x_n^{m_0} x_{n+d}^{m_1}) = E(x_n^{m_0}) E(x_{n+d}^{m_1}). \tag{6}$$

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

$$\begin{aligned} & E(x_n^{m_0} x_{n+d}^{m_1}) \\ = & \int_{-1}^{1} \rho(x_0) x_n^{m_0} x_{n+d}^{m_1} dx_0 \\ = & \int_{0}^{1} \cos^{m_0}(2\pi\theta 2^n) \cos^{m_1}(2\pi\theta 2^{n+d}) d\theta \\ = & \frac{1}{2^{(m_0+m_1)}} \sum_{\sigma} \delta \left(2^n \sum_{i=1}^{m_0} \sigma_{n_i} + 2^{n+d} \sum_{i=1}^{m_1} \sigma_{(n+d)_i} \right) \end{aligned}$$

where the last equation uses the fact that $\cos\theta = (e^{i\theta} + e^{-i\theta})/2$ and $\int_0^1 e^{i2\pi\theta k} d\theta = \delta(k)$, with $\delta(k) = 0$ if $k \neq 0$ otherwise equals to 1. $\sum_\sigma$ is the summation over all possible configurations, where $\sigma_{n_i} = \pm 1$ and $\sigma_{(n+d)_i} = \pm 1$.

All possible cases for m_0 and m_1 are analysis as below:

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

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

$$E(x_n^{m_0} x_{n+d}^{m_1}) = 2^{-(m_0+m_1)} \begin{pmatrix} m_0 \\ m_0/2 \end{pmatrix} \begin{pmatrix} m_1 \\ m_1/2 \end{pmatrix}$$

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

■

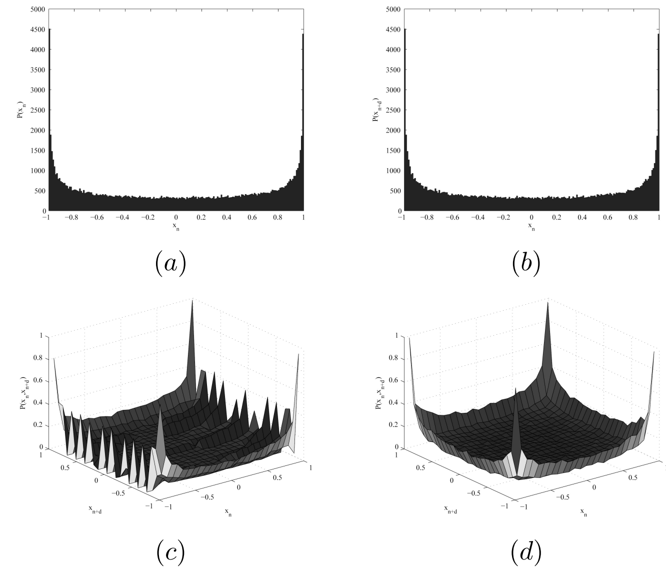

Fig. 1. Probability density $\rho(x_n)$ (a) and $\rho(x_{n+d})$ (b); (c) and (d) joint probability density $P(x_n, x_{n+d})$ for sampling distance $d = 5, 15$.

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 $d = 15$, $E(x_n^{m_0} x_{n+d}^{m_1}) = E(x_n^{m_0}) E(x_{n+d}^{m_1})$ 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 $Z(d, k, z_0) = \{z_n, z_{n+d}, ..., z_{n+kd}\}$ 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

$$x_k = 1 - 2z_{n+kd} \tag{7}$$

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

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

$$\Phi = \sqrt{\frac{2}{M}} \begin{pmatrix} x_0 & \dots & x_{M(N-1)} \\ x_1 & \dots & x_{M(N-1)+1} \\ \vdots & \vdots & \vdots \\ x_{M-1} & \dots & x_{MN-1} \end{pmatrix} \tag{8}$$

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 Φ 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 Φ 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 $h \in]0, 1/2[$,

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

where $Q = \langle \mathbf{x}, \mathbf{u} \rangle$ with $\mathbf{x}$ being any row vector of Φ and $\mathbf{u}$ being any unit vector.

Remark 3.4: In Lemma 3.3, $\lessapprox$ 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 Φ, then conclude the RIP using permutation theory.

1) **J-L property:**
Denote Φ_T the arbitrary column sub matrix of Φ, with index set $|T| = s$. For any unit vector $\mathbf{u} \in \mathbb{R}^s$, from Chernoff's inequality, given some positive value h, it has

$$\begin{aligned}
& \Pr\left[\|\Phi_T\mathbf{u}\|^2 \geq 1+\delta\right] \\
\leq\ & \exp\left(-hM(1+\delta)\right) E\left[\exp\left(hM\|\Phi_T\mathbf{u}\|^2\right)\right] \\
\approx\ & \exp\left(-hM(1+\delta)\right)\left(E\left[\exp\left(hQ^2\right)\right]\right)^M \\
\lessapprox\ & \exp\left(-hM(1+\delta)\right)\left(\frac{1}{\sqrt{1-2h}}\right)^M \\
\lessapprox\ & \exp\left(-\frac{M}{2}\left(\delta^2/2-\delta^3/3\right)\right) \\
=\ & \exp(-c_1(\delta)M)
\end{aligned}$$

where the last inequality is obtained by Taylor expansion and setting $h = \frac{1}{2}\frac{\delta}{1+\delta}$, 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

$$\begin{aligned}
& \Pr\left[\|\Phi_T\mathbf{u}\|^2 \leq 1-\delta\right] \\
\leq\ & \exp\left(hM(1+\delta)\right) E\left[\exp\left(-hM\|\Phi_T\mathbf{u}\|^2\right)\right] \\
\approx\ & \exp\left(hM(1+\delta)\right)\left(E\left[\exp\left(-hQ^2\right)\right]\right)^M \\
\lessapprox\ & \exp(hM(1-\delta))\left(1-h+\frac{3}{2}h^2\right)^M \\
=\ & \exp(-c_2(\delta)M)
\end{aligned}$$

where the last inequality is obtained by Taylor expansion and setting $h = h_{opt} = \frac{-2-\delta+\sqrt{4+8\delta-5\delta^2}}{3(1-\delta)}$, 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

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

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

$$\mathcal{E}_\kappa = \left\{\left|\|\Phi_T\mathbf{u}\|^2 - 1\right| \geq \delta\right\}$$

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

$$\begin{aligned}
\Pr[\mathcal{E}] &= \bigcup_i \Pr[\mathcal{E}_{T_i}] \lessapprox 2\exp(-c(\delta)M)\sum_{\kappa=1}^{s}\binom{N}{\kappa} \\
&\leq 2s\binom{n}{s}\exp(-c(\delta)M) \\
&\leq 2s(eN/s)^s\exp(-c(\delta)M) \\
&= \exp\left(\log 2 - c(\delta)M + s(\log(N/s)+1) + \log s\right)
\end{aligned}$$

where, for a fixed constant $c_3 > 0$, whenever $s \leq c_3 M/\log(N/s)$, the bound will only have the exponent with the exponential $\leq -c_4 M$ 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 $1 - \Pr[\mathcal{E}] \gtrapprox 1 - 2e^{-c_4 M}$. ∎

IV. Experiments

As presented in section III, we choose sampling distance $d = 15$, then generate the chaotic matrix following (8). The synthetic sparse signals $\mathbf{v}$ adopted throughout this section are with only ± 1 nonzero entries. The locations and signs of the peaks are chosen randomly. The measurement vector $\mathbf{y}$ is computed by $\mathbf{y} = \Phi\mathbf{v}$. Then the reconstruction $\mathbf{v}^*$ from $\mathbf{y}$ is solving by Linear Programming, which is accomplished using the SparseLab [13]. The decision for failure reconstruction is $\|\mathbf{v}^* - \mathbf{v}\| > 0.1$.

One interest is the maximum sparsity s which allows exact reconstruction of the signal. The results are given in Fig. 2 and show that the maximum sparsity s in the case of chaotic matrix is similar to that in the case of Sparse matrix [6], Gaussian random matrix and Bernoulli random matrix. The delicate experiment for the maximum sparsity s with respect to measurements M for chaotic matrix is given in Fig. 3.

Also the probability of successful recovery (recovery rate) for fixed signal size $N = 100$ and fixed measurement number $M = 50$ is compared among these matrices, shown in Fig. 4. The result shows that chaotic matrix performs similar to the other 3 matrices.

In addition, to evaluate the influence of the initial condition of the chaotic system (2), we set the initial state respectively to $z_0 \in \{0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.9\}$, and redo the experiment to test the recovery rate, shown in Fig. 5. The result shows that the initial state takes no influence to the recovery rate.

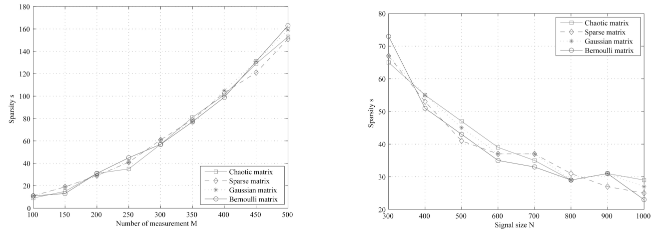

Fig. 2. Maximum sparsity for fixed signal size $N = 800$ and variable number of measurements $M \in [100, 500]$ (left) and for variable signal size $N \in [300, 1000]$ and fixed number of measurements $M = 200$ (right).

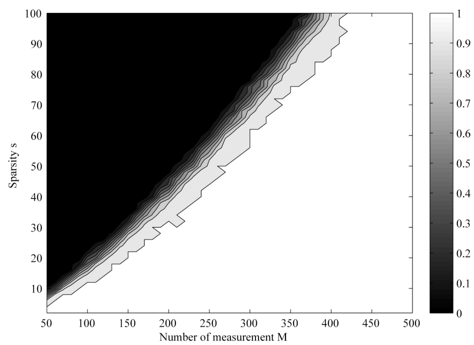

Fig. 3. Probability of correct recovery for fixed signal size $N = 1000$.

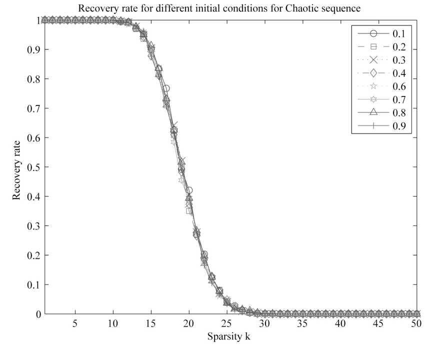

Fig. 5. Recovery rate for different initial conditions for chaotic sequence.

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

[1] Dimitris Achlioptas. Database-friendly random projections. In *Proc. 20th ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems (PODS)*, pages 274–281, New York, NY, USA, 2001. ACM.

[2] L. Applebaum, S. Howard, S. Searle, and R. Calderbank. Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery. *Applied and Computational Harmonic Analysis*, 26(2), March 2009.

[3] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, December 2008.

[4] D. Baron, S. Sarvotham, and R.G. Baraniuk. Bayesian compressive sensing via belief propagation. *Signal Processing, IEEE Transactions on*, 58(1):269 –280, jan. 2010.

[5] R. Berinde, A.C. Gilbert, P. Indyk, H. Karloff, and M.J. Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *46th Annual Allerton Conference on Communication, Control, and Computing,*, pages 798 –805, sept. 2008.

[6] R. Berinde and P. Indyk. Sparse recovery using sparse random matrices. Technical report, MIT, 2008.

[7] R. Calderbank, S. Howard, and S. Jafarpour. Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property. *Selected Topics in Signal Processing, IEEE Journal of*, 4(2):358 –374, april 2010.

[8] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Compte Rendus de l'Academie des Sciences, Series I*, 346(9-10), May 2008.

[9] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51(12):4203–4215, Dec. 2005.

[10] E.J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406 –5425, dec. 2006.

[11] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21 –30, march 2008.

[12] R. A. DeVore. Deterministic constructions of compressed sensing matrices. *Journal of Complexity*, 23(4-6), August 2007.

[13] David Donoho, Iddo Drori, Victoria Stodden, Yaakov Tsaig, and Morteza Shahram. Sparselab. http://sparselab.stanford.edu/.

[14] S. Howard, R. Calderbank, and S. Searle. A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes. *Conference on Information Sciences and Systems (CISS)*, March 2008.

[15] W. Johnson and J. Lindenstrauss. Extensions of Lipschitz maps into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

[16] Shahar Mendelson, Alain Pajor, and Nicole Tomczak-Jaegermann. Uniform uncertainty principle for bernoulli and subgaussian ensembles. *Constructive Approximation*, 28(3):277–289, December 2008.

[17] Wolfram Research. "logistic map". http://documents.wolfram.com/.

[18] Adriana Vlad, Adrian Luca, and Madalin Frunzete. Computational measurements of the transient time and of the sampling distance that enables statistical independence in the logistic map. In *Proc. International Conference on Computational Science and Its Applications (ICCSA)*, pages 703–718, Berlin, Heidelberg, 2009. Springer-Verlag.

[19] Kai Wang, Wenjiang Pei, Haishan Xia, Monica Garcia Nustes, and J.A. Gonzalez. Statistical independence in nonlinear maps coupled to non-invertible transformations. *Physics Letters A*, 372(44):6593 – 6601, 2008.

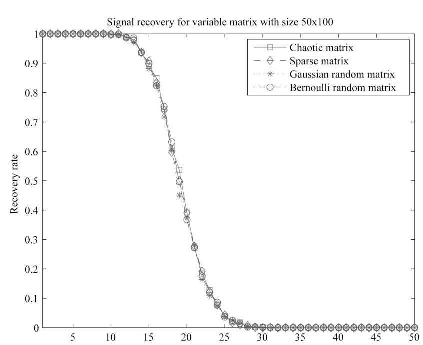

Fig. 4. Recovery rate for different sensing matrix.