Toeplitz-Structured Chaotic Sensing Matrix for Compressive Sensing

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

Toeplitz-Structured Chaotic Sensing Matrix for Compressive Sensing

Lei Yu[*†§], Jean Pierre Barbot[†§], Gang Zheng[‡§], Hong Sun[*]

[*]Signal Processing Laboratory, Electronic and Information School, Wuhan University, China
[†]ECS ENSEA, 6 Avenue du Ponceau, 95014 Cergy-Pontoise, France.
[‡]INRIA Lille-Nord Europe, 40 Avenue Halley, 59650, Villeneuve d'Ascq, France.
[§]EPI ALIEN, INRIA, France.
lei.yu@ensea.fr, barbot@ensea.fr, gang.zheng@inria.fr, hongsun@whu.edu.cn

Abstract— **Compressive Sensing (CS) is a new sampling theory which allows signals to be sampled at sub-Nyquist rate without loss of information. Fundamentally, its procedure can be modeled as a linear projection on one specific sensing matrix, which, in order to guarantee the information conservation, satisfies Restricted Isometry Property (RIP). Ordinarily, this matrix is constructed by the Gaussian random matrix or Bernoulli random matrix. In previous work, we have proved that the typical chaotic sequence - logistic map can be adopted to generate the sensing matrix for CS. In this paper, we show that Toeplitz-structured matrix constructed by chaotic sequence is sufficient to satisfy RIP with high probability. With the Toeplitz-structured Chaotic Sensing Matrix (TsCSM), we can easily build a filter with small number of taps. Meanwhile, we implement the TsCSM in compressive sensing of images.**

I. INTRODUCTION

In information field, traditional information capturing paradigm always follows the famous Shannon-Nyquist sampling theory, which says that signals must be sampled larger than Nyquist rate (2 times of the frequency bandwidth of the signal) to guarantee the exact recovery [16]. This leads that, even for a 2 minutes audio record, 5MB data are needed to retain the voice undistorted. Consequently, in order to reduce the storage consumption and the transmission bandwidth, conventionally, an immediate data compression takes place after the sensing procedure. The principle of this kind of *sensing-then-compression* scheme, such as JPEG or JPEG 2000, is to discard the small coefficients and retain the significant ones under some fixed transformation. In reality, the data stored and transmitted is only a very small part of the original ones, while we still have to sense the whole of the samples and then discard most of them. It is extremely wasteful of power consumption and memory storage. Moreover, this *sensing-then-compression* scheme is not practicable if the signal is presented at a high rate or if the measurement device has limited computational resources (such as in a sensor network).

Recently, a new theory called Compressive Sensing (CS) was proposed by Candès et. al. [7] and Donoho [9]. Unlike the traditional sampling methods, CS provides an information capturing paradigm with both sampling and compression. It permits signals being sampled below the conventional Nyquist rate, meanwhile still allowing optimal reconstruction of the signal. If signals are sparse or compressible, the required measurements are far less than those of traditional methods, and even more, the measure process is non-adaptive which makes the sampling process much more universal and faster.

A. Preliminary of Compressive Sensing

Mathematically, the procedure of CS can be expressed as a linear projection

$$\mathbf{y} = \Phi\mathbf{v} \tag{1}$$

where $\mathbf{v} \in \mathbb{R}^n$ is the original signal, $\Phi \in \mathbb{R}^{m \times n}$ is the sensing matrix and $\mathbf{y} \in \mathbb{R}^m$ is the measurement.

Remark 1.1: Particularly, the traditional sampling paradigm following Nyquist theory, is one special case of CS with sensing matrix Φ being identity matrix with $m = n$. From this point, CS improves the way of capturing signal information, typically by using random sensing matrix with $m \ll n$. Consequently, we say CS samples and compresses signals simultaneously.

Since $m \ll n$, CS highly reduces the length of the signal. However, it also leads equation (1) be highly underdetermined. With an additional assumption on original signal, supposing that sparsity is always fulfilled for natural signals (itself or in some basis), the solution of $\mathbf{v}$ for (1) can then be possibly uniquely determined.

Definition 1.2: (Sparsity) Given a vector $\mathbf{v} \in \mathbb{R}^n$, and denote $Card(\mathbf{v}) = Card\{v_i \neq 0, i \in [1,n]\}$ the cardinality of non zero entries of $\mathbf{v}$, then $\mathbf{v}$ is called to be an s-sparsity vector, if $Card(v) \leq s \ll n$.

Based on Definition 1.2, we can then define the set of sparse vector of s-sparsity Σ_s as follows:

$$\Sigma_s = \{\mathbf{v} \in \mathbb{R}^n \mid Card(\mathbf{v}) \leq s\} \tag{2}$$

Intuitively, equation (1) can be solved by searching the sparsest one of vectors lied on the null surface of Φ [7], i.e.

$$\mathbf{v}^* = \arg\min_{\Phi\mathbf{v}=y} \|\mathbf{v}\|_0 \tag{3}$$

where $\|\mathbf{v}\|_0$ is ℓ_0 norm and denotes the counter of nonzero entries of $\mathbf{v}$.

However, it was shown to be a NP-Hard problem in [7], and thus practically, it was relaxed in [7] to solve the following problem.

$$\mathbf{v}^* = \arg\min_{\Phi\mathbf{v}=y} \|\mathbf{v}\|_1 \tag{4}$$

where $\|\mathbf{v}\|_1$ is ℓ_1 norm and denotes the absolute sum of all entries of $\mathbf{v}$.

Definition 1.3 (Restricted Isometry Property [4]):
For matrix $\Phi \in \mathbb{R}^{m \times n}$, define $\delta_s > 0$ be the smallest constant satisfying following inequality,

$$(1 - \delta_s)\|\mathbf{v}\|_2^2 \leq \|\Phi\mathbf{v}\|_2^2 \leq (1 + \delta_s)\|\mathbf{v}\|_2^2 \qquad (5)$$

for all vectors $\mathbf{v} \in \Sigma_s$.
Then matrix Φ is said to satisfy the RIP of order s with constant δ_s.

Theorem 1.4 (Exact Recovery Theorem [4]):
- If RIP of order $2s$ for matrix Φ satisfies $\delta_{2s} < 1$, then for any vector $\mathbf{v} \in \Sigma_s$, there exists a unique solution for (3).
- If $\delta_{2s} < \sqrt{2} - 1$, then for any vector $\mathbf{v} \in \Sigma_s$, solutions for (3) and (4) are equivalent.

B. Motivation

Consequently, finding a proper sensing matrix Φ satisfying RIP with conditions in Theorem 1.4 is one of the most important problem in CS. Candès and Tao have proposed that matrix with elements drawn by Gaussian distribution or Bernoulli distribution satisfies RIP with overwhelming probability, providing that sparsity $s \leq O(m/\log n)$ [6]. And the randomly selected Fourier basis also retains RIP with overwhelming probability if sparsity $s \leq O(m/(\log n)^6)$ [6]. Even though there exist some deterministic sensing matrices, such as *Chirp Sensing Codes* by L. Applebaum, *et al.* [2], deterministic sensing matrix using finite fields by R. A. Devore [8] and using second order Reed-Muller codes by S. Howard *et al.* [11], their RIP properties are not either fulfilled certainly. Until now, the widely used sensing matrix in practice is generated randomly.

In previous work [19], we have proposed the chaotic sequence to construct such a sensing matrix, called chaotic matrix. Based on *Egodicity* and statistical property of chaotic sequence, we have proved that chaotic matrix also has RIP with overwhelming probability, providing that $s \leq O(m/\log(n/s))$. Meanwhile, the proof shows that the probability of satisfying RIP for the proposed chaotic matrix is larger than that for Gaussian random matrix and Bernoulli random matrix [6].

In this paper, we aim to construct a Toeplitz matrix with chaotic sequence and prove that the Toeplitz chaotic matrix also retains the RIP property with overwhelming probability. The use of Toeplitz matrix in CS application takes the following advantages:
1) Only $O(n)$ independent chaotic variables are needed;
2) The multiplication can be efficiently implemented using FFT;
3) Toeplitz-structured matrix arises naturally in certain applications such as filter design.

C. Paper Organization

In section II, the Toeplitz-structured Chaotic Sensing Matrix (TsCSM) is proposed for CS and its RIP property is proved. In section III, Iterative Threshold (IT) algorithm to reconstruct the original signals is introduced. In section IV, we compare the TsCSM with Chaotic Sensing Matrix (CSM) and Gaussian Sensing Matrix (GSM), meanwhile, some experiments on images are illustrated.

II. TOEPLITZ-STRUCTURED CHAOTIC SENSING MATRIX

A. Chaotic Sensing Matrix and its Johnson-Lindenstrauss (J-L) Embedding [13] Property

Considering one popular discrete chaotic system, logistic map, whose difference equation can be written as below.

$$z_+ = rz(1 - z), \quad \text{with} \quad r \in]0, 4] \qquad (6)$$

where $z \in]0, 1[\subset \mathbb{R}$ is the discrete state.

It is well known that for logistic map (6) with parameter $r = 4$, the state sequence $z(t)$ satisfies Beta distribution with parameter $\alpha = 0.5$ and $\beta = 0.5$ [18], with the following probability density function.

$$f(x; 0.5, 0.5) = \frac{1}{\pi}(x - x^2)^{-1/2} \qquad (7)$$

Set $z_i(t)$ as the output sequence generated by logistic map (6) with initial condition $z_i(0)$, and let $x_i(t)$ denote the regularization of $z_i(t)$ as below

$$x_i(t) = z_i(t) - 0.5 \qquad (8)$$

Approximately, $x_i(t)$ can be considered as a random variable and it satisfies the following distribution.

$$f(x) = \frac{1}{\pi}(0.25 - x^2)^{-1/2} \qquad (9)$$

which is called here Beta-like distribution.

Then, by selecting m different initial conditions $z(0) \in]0, 1[^m \subset \mathbb{R}^m$, one can obtain m vectors with dimension n, which enables us to construct the following matrix Φ scaled by $\sqrt{8/m}$

$$\Phi = \sqrt{\frac{8}{m}} \begin{pmatrix} x_0(0) & \dots & x_0(n-1) \\ \vdots & \ddots & \vdots \\ x_{m-1}(0) & \dots & x_{m-1}(n-1) \end{pmatrix}$$

which is called here the Beta-like matrix.

In the previous work [19], it has the following theorem:
Theorem 2.1 (J-L Embedding with Beta-like Matrix):
Given a Beta-like matrix $\Phi \in \mathbb{R}^{m \times s}$, there exists a constant $\delta_s > 0$, such that, for any vector $\mathbf{u} \in \mathbb{R}^s$, the following inequality

$$(1 - \delta_s)\|\mathbf{u}\|_2^2 \leq \|\Phi\mathbf{u}\|_2^2 \leq (1 + \delta_s)\|\mathbf{u}\|_2^2 \qquad (10)$$

is satisfied with overwhelming probability.

$$\Pr\left[\left|\|\Phi\mathbf{u}\|^2 - 1\right| \geq \delta_s\right] \leq 2\exp(-c(\delta_s)m) \qquad (11)$$

where $c(\delta_s)$ is a constant related to δ_s.

Remark 2.2: Note that we do not refer the final theorem in [19], and the formation of Φ is not exactly the same. Even though, the theorem is also verified because of, statistically, the *Egodicity* property of the chaotic system.

Remark 2.3: This theorem is also retained for matrix constructed with Gaussian or Bernoulli matrix, but with different constant $c(\delta_s)$ [1].

B. Toeplitz-structured Chaotic Sensing Matrix

For chaotic system (6), let us set one initial condition $z(0) \in \mathbb{R}$, then generate a sequence $x \in \mathbb{R}^n$ according to (8). Then we can construct a Toeplitz-structured matrix $\Phi \in \mathbb{R}^{m \times n}$ of the following form

$$\Phi = \sqrt{\frac{8}{m}} \begin{pmatrix} x(n-1) & x(n-2) & \dots & x(0) \\ x(0) & x(n-1) & \dots & x(1) \\ \vdots & \vdots & \vdots & \vdots \\ x(m-2) & x(m-3) & \dots & x(m-1) \end{pmatrix}$$

where $\sqrt{8/m}$ is for normalization.

In [3], the authors proposed to construct the Toeplitz-structured matrix with Gaussian or Bernoulli random sequences and according to the structure of Toeplitz matrix, it can be concluded that for Gaussian (or Bernoulli) Toeplitz-structured matrix, if $m \geq c_1 s^3 \ln(n/s)$, its RIP of order s with constant δ_s is verified with probability at least

$$1 - e^{-c_2 k/m^2}$$

where c_1 and c_2 are dependant only on δ_s.

According to Theorem 2.1, both the TsCSM and Toeplitz-structured Gaussian or Bernoulli matrix retains the J-L embedding property with overwhelming probability. Similarly, for TsCSM, if $m \geq c_1' s^3 \ln(n/s)$, the RIP of order s with constant δ_s is verified with probability at least

$$1 - e^{-c_2' k/m^2}$$

where c_1' and c_2' are dependant only on δ_s.

III. SIGNAL RECONSTRUCTION

From now on, there exist many reconstruction methods for signal reconstruction in CS framework, such as Basis Pursuit (BP) [5], Orthogonal Matching Pursuit (OMP) [17], CoSaMP [15], *et. al*. Since this paper does not focus on the reconstruction algorithms, we just briefly introduce one iterative algorithm - Iterative Thresholding (IT) [12], to reconstruct the original signal.

In reality, the nature signals are compressible, i. e. the signals can be sparsely represented by an specific basis, such as Fourier basis, Wavelet basis. Denote the nature signals as $f \in \mathbb{R}^n$ and a basis $B \in \mathbb{R}^{n \times n}$, then

$$f = B\mathbf{v} \qquad (12)$$

where $\mathbf{v} \in \Sigma_s$ is s-sparse. And the CS procedure is expressed as the following linear projection

$$y = \Phi f = \Phi B \mathbf{v} \qquad (13)$$

IT algorithm is to solve the following problem

$$f^* = \arg \min_{f \in \mathbb{R}^n} \|y - \Phi f\|_2^2 + \lambda \|B^* f\|_1$$

which is a relaxion of optimization (4). And IT algorithm proceeds as follows to find the optimal solution for (13), let $f^{(0)} = 0$ and use the iteration

$$f^{(i+1)} = B S_\lambda (B^* f^{(i)} + B^* \Phi^* (y - \Phi f^{(i)})) \qquad (14)$$

where i represents the ith iteration, B is a basis such as discrete wavelet basis, B^* is the inverse of B and $S_\lambda(*)$

is a non-linear shrinkage operator that keeps the larger coefficients, i.e.

$$S_\lambda(x) = \begin{cases} x - \frac{\lambda}{2} & \text{if } x \geq \frac{\lambda}{2} \\ 0 & \text{if } |x| < \frac{\lambda}{2} \\ x + \frac{\lambda}{2} & \text{if } x \leq -\frac{\lambda}{2} \end{cases} \qquad (15)$$

where λ is the median of all coefficients, i.e. $\lambda = \text{median}(x)$. The IT algorithm is illustrated as Algorithm 1.

Algorithm 1 Iterative Thresholding Algorithm (IT)

Input: measurements $y \in \mathbb{R}^m$ and measurement matrix $\Phi \in \mathbb{R}^{m \times n}$, $m \ll n$;
Output: The recovery image $f^* = f^{(t_{\max})}$
Initial: $f^{(0)} = 0$, maximum iteration $t_{\max}$
    for $i = 0$ to $t_{\max}$ **do**
      Update: $f^{(i)} = f^{(i)} + \Phi^*(y - \Phi f^{(i)})$;
      Transform: Compute the coefficients over basis B, $\mathbf{v} = B^* f^{(i)}$;
      Sparse promote: $\mathbf{v} = S_\lambda(\mathbf{v})$;
      Inverse transform: $f^{(i+1)} = B\mathbf{v}$.
    end for

IV. NUMERICAL SIMULATIONS

A. TsCSM, Chaos Sensing Matrix (CSM) and Gaussian Sensing Matrix (GSM)

In section II, it has been proved that the TsCSM also has the RIP property with overwhelming probability, i.e. it guarantees the exact recovery for any sparse signals. To verify the property, a comparison between TsCSM and CSM (or GSM) is carried out by considering their exact recovery rates. Let us describe the experiment setup: we build a TsCSM (or CSM, GSM) with dimension 50×100, which means that there are 50 measurements, then, for each sparsity level $s \in \{1, ..., 50\}$, we arbitrarily generate 1000 s-sparse signals with dimension 100, we compress them following the CS procedure (1) and we use the Algorithm 1 (here the basis B is identity matrix) to recover the sparse signals. For each trial, we denote it successful if its recovery error is less than a level (here set it as 0.001) and we can finally obtain a successful rate, i.e. recovery rate, see Fig. 1.

In Fig. 1, the recovery rate for TsCSM is similar with CSM but little larger than GSM, and the reason is that the probability of J-L embedding property for GSM is smaller than that for CSM which has been demonstrated in [19].

B. TsCSM for Images

Recently, the application of CS on imaging is attracted a lot researchers [10], [14]. In this simulation, we only analyze the performance of the TsCSM fundamentally on imaging rather than the physical implementation. Suppose TsCSM is used in the single imaging system introduced in [10], [14], we will give two experiments to illustrate the performance of TsCSM.

First, let's fix the sample rate at 0.5, i.e. for image with size 256×256, we only measure $32,768$ samples
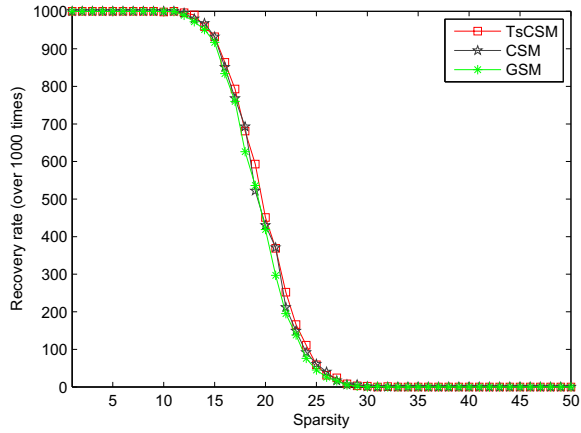
Fig. 1. Recovery rate for TsCSM, CSM and GSM (over 1 000 trials), with dimension 50×100.

by projection with TsCSM following CS procedure (1); then, we use Algorithm 1 (with B as wavelet basis) to reconstruct the original images from the compressed data. The result are listed in Fig. 2, where Peak Signal to Noise Ratio (PSNR) is defined as

$$\text{PSNR} = 20 \log \frac{\|f\|_2}{\|f - f^*\|_2} \tag{16}$$

Second, let's verify the efficiency of compression. For one image with dimension 256×256, we measure this image with sample rate at $0.3, 0.5, 0.7$ and 0.9 separately, then use Algorithm 1 (with B as wavelet basis) to reconstruct the original images, see Fig. 3. The error is the direct difference between the recovery and the original image.

V. Conclusion

This paper briefly introduced a new sampling theory - Compressive Sensing (CS) and proposed a method to construct a Toeplitz-structured Matrix for CS with chaotic sequence. We prove that Toeplitz-structured Chaotic Sensing Matrix (TsCSM) retains the Restricted Isometry Property (RIP) with overwhelming probability, which guarantees the exact recovery. Then Iterative Thresholding (IT) algorithm is introduced to reconstruct the original signal. After that, a simulation on 1-dimensional signal CS is implemented and the comparison between TsCSM and CSM (or GSM) shows that TsCSM has the similar (or better) performance. Meanwhile, we also simulate imaging system with TsCSM, and the results show a good perspective for TsCSM.

Acknowledgment

Fig. 2. Reconstructions of CS with TsCSM for different types of images.

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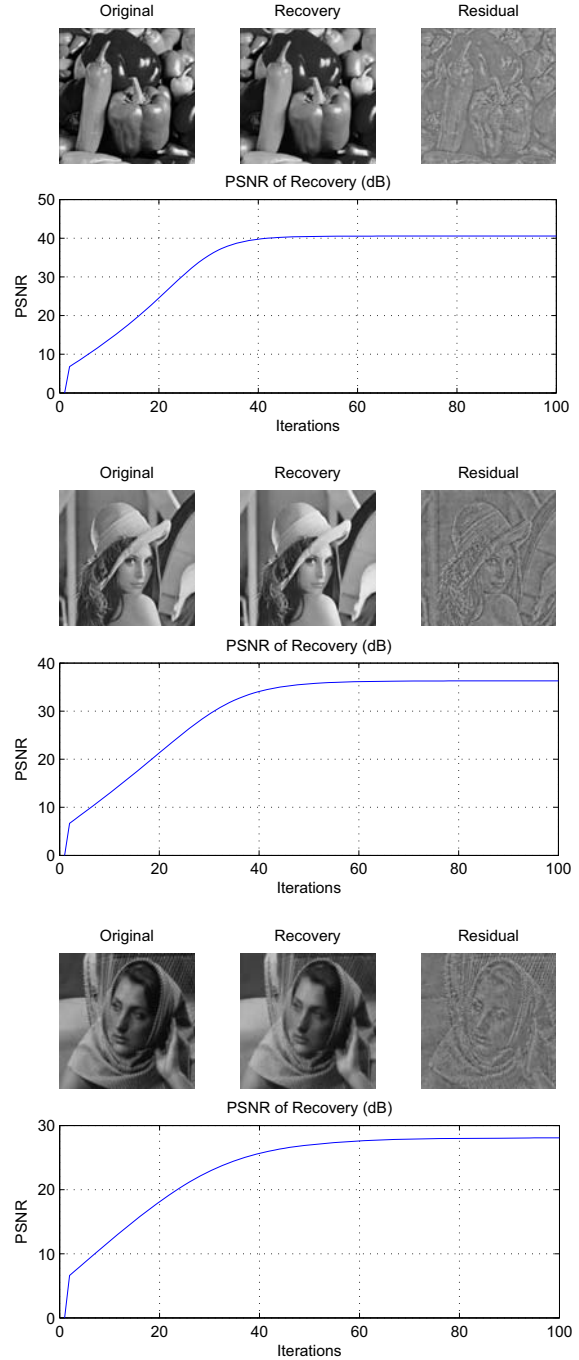
[2] L. Applebaum, S. Howard, S. Searle, and R. Calderbank. Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery. *Applied and Computational Harmonic Analysis*, 26(2), March 2009.

[3] Waheed U. Bajwa, Jarvis D. Haupt, Gil M. Raz, Stephen J. Wright, and Robert D. Nowak. Toeplitz-structured compressed sensing matrices. In *IEEE Workshop SSP*, 2007.

[4] 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.

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

[6] E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? 52(12):5406–5425, Dec. 2006.
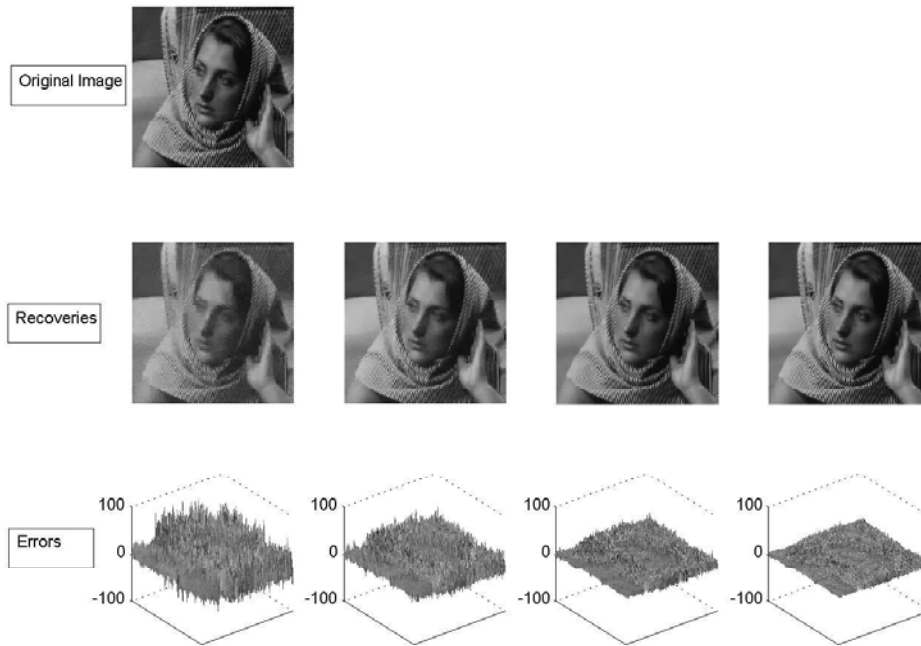
Fig. 3. Reconstructions of CS with TsCSM under different sample rate: from left to right, 0.3, 0.5, 0.7, 0.9.

[7] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. 25(2):21–30, March 2008.

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

[9] D. L. Donoho. Compressed sensing. 52(4):1289–1306, April 2006.

[10] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, Ting Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. 25(2):83–91, March 2008.

[11] 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.

[12] M. Defrise I. Daubechies and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, 57(11):1413–1457, Nov. 2004.

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

[14] Jianwei Ma. Single-pixel remote sensing. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, 6(2):199–203, April 2009.

[15] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.*, to appear.

[16] C. E. Shannon. Communication in the presence of noise. 72(9):1192–1201, September 1984.

[17] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. 53(12):4655–4666, Dec. 2007.

[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] Lei Yu, Jean Pierre Barbot, Gang Zheng, and Hong Sun. Chaotic sensing matrix for compressive sensing. Preprint, http://www-ecs.ensea.fr/.