Symmetric Toeplitz-Structured Compressed Sensing Matrices

Tao Huang¹,² • Yi-Zheng Fan¹,² • Ming Zhu¹,²

Received: 6 October 2014 / Revised: 2 March 2015 / Published online: 17 April 2015
© Springer Science+Business Media New York 2015

Abstract How to construct a suitable measurement matrix is an important topic in compressed sensing. A significant part of the recent work is that the measurement matrices are not completely random on the entries but exhibit some considerable structures. In this paper, we proved that a symmetric Toeplitz matrix and its variant can be used as measurement matrices and recovery signal with high probability. Compared with random matrices (e.g. Gaussian and Bernoulli matrices) and some structured matrices (e.g. Toeplitz and circulant matrices), we need to generate fewer independent entries to obtain the measurement matrix while the effectiveness of the recovery keeps good.

Keywords Compressed sensing • Restricted isometry property • Measurement matrix • Symmetric Toeplitz-structured matrix

1 Introduction

Compressed sensing has a number of potential applications in image processing, geophysics, medical imaging, computer science as well as other areas of science and technology. Due to [7] by Donoho and [4] by Candès, Romberg and Tao, this area makes a great progress since 2006. Compressed sensing focuses on the problem of
recovering $x$ from the knowledge of $y = Ax$, where $A$ is a suitable $k \times n$ measurement matrix (also called CS matrix) and $k \ll n$. As we know, when the signal $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ is $m$-sparse (i.e. the number of non-zero coefficients of $x$ is at most $m$) and the CS matrix $A$ holds certain conditions such as restricted isometry property (RIP) [3] for all $m$-sparse vectors $x$, the solution $x^*$ can recover $x$ exactly. Actually, recovering $x$ can be solved by $\ell_1$-minimization instead of $\ell_0$-minimization:

$$\min \|x\|_1, \; \text{subject to} \; y = Ax,$$

where the $\ell_p$-norm is defined by $\|x\|_p = \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}$, as usual.

As a weaker version of RIP, RIP of order $3m$ can be expressed as follows. Let $T \subset \{1, 2, \ldots, n\}$ and $A_T$ be the $k \times |T|$ submatrix obtained by retaining the columns of $A$ corresponding to the indices in $T$. If there exists a constant $\delta_{3m} \in (0, 1/3)$ such that

$$(1 - \delta_{3m})\|x\|_2^2 \leq \|A_T x\|_2^2 \leq (1 + \delta_{3m})\|x\|_2^2$$

holds for all subsets $T$ with $|T| \leq 3m$, then $A$ satisfies RIP of order $3m$. It was shown that random matrices with entries drawn independently from certain probability distributions satisfy RIP of order $3m$ with high probability for every $\delta_{3m} \in (0, 1/3)$ provided $k \geq \text{const} \cdot m \ln(n/m)$ [5, 6]. We refer to such matrices as independent and identically distributed CS matrices or simply IID CS matrices.

The problem that how to choose a suitable CS matrix is a main topic in compressed sensing. It had been proved that Gaussian or Bernoulli matrices can be used as IID CS matrices if we are allowed to choose the entries of sensing matrix freely; see [2, 4, 5, 13, 19]. However, most applications do not allow a free choice of the entries of the sensing matrix and enforce a particularly structure on the matrix. Recently, the random Toeplitz or circulant matrices introduced in [1, 14, 15] are estimated as CS matrix, where the matrix is generated by a random vector whose entries are chosen independently according to a suitable probability distribution. The authors proved that such matrix can provide recovery guarantees for $\ell_1$-minimization, and illustrated its application in system identification. Compared to Bernoulli or Gaussian matrices, random Toeplitz and circulant matrices have the advantages that they require an order of $n$ instead of $n^2$ random entries to be generated. In [8, 9] Fan et. al show that some random symmetric (not necessarily IID) matrices satisfy RIP.

Motivated by the work [1], we show that if a probability distribution $P(a)$ yields an IID CS matrix (having unit-norm columns in expectation) then a $k \times n$ partial symmetric Toeplitz matrix $A$ (also having unit-norm columns in expectation) of the form

$$\begin{bmatrix}
a_n & a_{n-1} & \cdots & a_2 & a_1 \\
a_{n-1} & a_n & \cdots & a_3 & a_2 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
a_{n-k+1} & a_{n-k+2} & \cdots & a_{k+1} & a_k
\end{bmatrix}$$

is also a measurement matrix in the sense that it satisfies RIP of order $3m$ with high probability for every $\delta_{3m} \in (0, 1/3)$ provided $k \geq \text{const} \cdot m^3 \cdot \ln(n/m)$, where the
entries $a_i$ ($i = 1, 2, \ldots, n$) are drawn independently from $P(a)$. Compared with Toeplitz matrix in [1], we only need to generate $n$ independent entries rather than $n + k - 1$ independent entries. So, $k - 1 = O(\ln n)$ less entries are reduced to form CS matrices, while the effectiveness of recovery keeps good.

## 2 Main Result

The proof of the main result adopts the technique used in [1] with a slight modification. We use the celebrated Hajnal-Szemerédi theorem on equitable coloring of graphs [10] to partition a $k \times |T|$ partial symmetric Toeplitz-structured matrix $A_T$ into roughly $O(m^2)$ IID submatrices with dimensions approximately equal to $O(k/m^2) \times |T|$. We also extend the main result to some variants of symmetric Toeplitz matrices.

**Lemma 2.1** For fixed $n, m$, let $P(a)$ be a probability distribution that generates a $k \times n$ IID matrix having unit-norm columns in expectation such that, for every $\delta_{3m} \in (0, 1/3)$ and every $T \subset \{1, 2, \ldots, n\}$ with $|T| = 3m$, the $k \times |T|$ IID submatrix obtained by retaining the columns corresponding to the indices in $T$ satisfies (1.2) with probability at least

$$1 - e^{-f(k,m,\delta_{3m})}, \quad (2.1)$$

where $f(k,m,\delta_{3m})$ is some real-valued function of $k, m$ and $\delta_{3m}$.

Let $\{a_i\}_{i=1}^n$ be a sequence of random variables drawn independently from the same distribution, and $A$ be a $k \times n$ symmetric Toeplitz matrix of the form given in (1.3). Then, for every $\delta_{3m} \in (0, 1/3)$ and every $T \subset \{1, 2, \ldots, n\}$ with $|T| = 3m$, the symmetric Toeplitz submatrix $A_T$ satisfies (1.2) with probability at least

$$1 - e^{-f([k/q],m,\delta_{3m})+ln(q)}, \quad (2.2)$$

where $q = 3m(6m - 1) + 1$.

**Remark 1** Let the probability distribution $P(a)$ be given by

$$N(0,1/k), \begin{cases} +\sqrt{1/k} & \text{with probability 1/2} \\ -\sqrt{1/k} & \text{with probability 1/2} \end{cases}, \text{ or } \begin{cases} +\sqrt{3/k} & \text{with probability 1/6} \\ 0 & \text{with probability 2/3} \\ -\sqrt{3/k} & \text{with probability 1/6} \end{cases} \quad (2.3)$$

Then

$$f(k,m,\delta_{3m}) = c_0 k - 3m \ln(12/\delta_{3m}) - \ln 2, \quad (2.4)$$

where $c_0 = c_0(\delta_{3m}) = \delta_{3m}^2/16 - \delta_{3m}^3/48$ (see [2]).

**Proof** Fix $\delta_{3m} \in (0, 1/3)$ and $T \subset \{1, 2, \ldots, n\}$ with $|T| = 3m$. Let $A_{T,i}$ denote the $i$-th row of $A_T$ for each $i = 1, 2, \ldots, k$. Construct an undirected graph $G = (V,E)$ with $V = \{1, 2, \ldots, k\}$ and $E = \{(i,i') \in V \times V : i \neq i', A_{T,i}$ and $A_{T,i'}$ are dependent $\}$. From the properties of symmetric Toeplitz matrices, $A_{T,i}$ can be
dependent with at most \(|T|/(2|T| - 1) =: q - 1\) other rows of \(A_T\), which implies that the maximum degree \(\Delta\) of \(G\) is given by \(\Delta \leq q - 1\). By Hajnal-Szemerédi theorem on equitable coloring of graphs [10], we can always partition \(G\) using \(q\) colors such that

\[
|k/q| \leq \min_{j \in \{1,2,\ldots,n\}} \max_{j \in \{1,2,\ldots,n\}} |C_j| \leq \max_{j \in \{1,2,\ldots,n\}} |C_j| \leq \lfloor k/q \rfloor,
\]

where \(\{C_j\}_{j=1}^q\) correspond to the different color classes.

Next, let \(A_T^j\) be the \(|C_j| \times |T|\) partition submatrix obtained by retaining the rows of \(A_T\) corresponding to the indices in \(C_j\), which is an IID submatrix of \(A_T\). Observe that for any \(x \in \mathbb{R}^{|T|}\),

\[
\|A_Tx\|_2^2 = \sum_{j=1}^q \|A_T^jx\|_2^2 = \sum_{j=1}^q \frac{|C_j|}{k} \|\tilde{A}_T^jx\|_2^2, \quad (2.5)
\]

where \(\tilde{A}_T^j := \sqrt{k/|C_j|}A_T^j\) (to ensure unit-norm columns in expectation). So, each \(\tilde{A}_T^j\) is an \(|C_j|\times |T|\) submatrix with IID entries from the distribution \(P(a)\) and hence, satisfies (1.2) with probability at least

\[
1 - e^{-f(|C_j|m,\delta_{3m})} \geq 1 - e^{-f([k/q]m,\delta_{3m})}. \quad (2.6)
\]

Noting that \(\sum_{j=1}^q \frac{|C_j|}{k} = 1\), from (2.5) the occurrence of the event \(\bigcap_{j=1}^q \tilde{A}_T^j\) satisfies (1.2) implies that for any \(x \in \mathbb{R}^{|T|}\),

\[
\sum_{j=1}^q \frac{|C_j|}{k} (1 - \delta_{3m})\|x\|_2^2 \leq \sum_{j=1}^q \frac{|C_j|}{k} \|\tilde{A}_T^jx\|_2^2 \leq \sum_{j=1}^q \frac{|C_j|}{k} (1 + \delta_{3m})\|x\|_2^2,
\]

and hence

\[
(1 - \delta_{3m})\|x\|_2^2 \leq \|A_Tx\|_2^2 \leq (1 + \delta_{3m})\|x\|_2^2.
\]

So

\[
\bigcap_{j=1}^q \tilde{A}_T^j\) satisfies (1.2) \(\subset \{A_T\) satisfies (1.2) \}. \quad (2.7)
\]

Consequently, from (2.6) and (2.7) we have

\[
\Pr(\{A_T\) satisfies (1.2) \}) = 1 - \Pr(\{A_T\) does not satisfies (1.2) \})
\geq 1 - \Pr(\{\bigcap_{j=1}^q \tilde{A}_T^j\) does not satisfies (1.2) \})
\geq 1 - \sum_{j=1}^q \Pr(\{\tilde{A}_T^j\ does not satisfies (1.2) \})
\geq 1 - \sum_{j=1}^q e^{-f([k/q]m,\delta_{3m})}
\geq 1 - e^{-f([k/q]m,\delta_{3m}) + \ln q}.
\]

This completes the proof. \(\square\)
**Theorem 2.2** Suppose that $n,m$ are given, and let $A$ be a $k \times n$ partially symmetric Toeplitz matrix of the form given in (1.3), where the entries $\{a_i\}_{i=1}^n$ are drawn independently from one of the probability distributions given in (2.3). Then, there exist constants $c_1,c_2 > 0$ depending only on $\delta_{3m}$ such that for any $k > c_1 m^3 \ln(n/m)$, $A$ satisfies RIP of order $3m$ for every $\delta_{3m} \in (0,1/3)$ with probability at least

$$1 - e^{-c_3 k/m^2}.$$  \hspace{1cm} (2.8)

**Proof** For given $\delta_{3m} \in (0,1/3)$, from Lemma 2.1, $A$ satisfies (1.2) for any $T \subset \{1,2,\ldots,n\}$ of cardinality $3m$ with probability at least

$$1 - e^{-c_0 [k/q] + 3m \ln(12/\delta_{3m}) + \ln 2 + \ln q} \geq 1 - e^{-c_0 k/18m^2 + 3m \ln(12/\delta_{3m}) + \ln 2 + \ln(18m^2) + c_0},$$

and there are $\left(\frac{n}{3m}\right) \leq (en/3m)^{3m}$ such subsets. Consequently, union bounding over these subsets yields that $A$ satisfies RIP of order $3m$ with probability at least

$$1 - e^{-c_0 k/18m^2 + 3m \ln(12/\delta_{3m}) + \ln(n/3m) + 1] + \ln 2 + \ln(18m^2) + c_0}.$$ \hspace{1cm} (2.9)

Next, fix $0 < c_2 < c_0/18$ and pick $c_1 > 54c_3/(c_0 - 18c_2)$, where $c_3 = \ln(12/\delta_{3m}) + 2 \ln 2 + c_0 + 4$. Then, for any $k \geq c_1 m^3 \ln(n/m)$, the exponent in the exponential in (2.9) is upper bounded by $-c_2 k/m^2$ and this completes the proof of the theorem. \hfill \square

By a similar discussion as in Theorem 2.2, we have the following two corollaries immediately.

**Corollary 2.3** The results of Theorem 2.2 apply equally well to the left-shifted Toeplitz matrix of the form

$$\begin{pmatrix}
a_1 & a_2 & \cdots & a_{n-1} & a_n \\
a_2 & a_3 & \cdots & a_n & a_{n-1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_k & \cdots & \cdots & a_{n-k+2} & a_{n-k+1}
\end{pmatrix}.$$}

**Corollary 2.4** Let $\Theta \subset \{1,2,\ldots,n\}$ with $|\Theta| = k$. Then the submatrix $A_{\Theta}$ obtained from a symmetric Toeplitz or left-shifted symmetric Toeplitz matrix by retaining the rows corresponding to the indices in $\Theta$, satisfies Theorem 2.2.

### 3 Applications

We now describe how the results for symmetric Toeplitz CS matrices lend themselves to (1) identification of LTI systems having sparse impulse responses; and (2) recovery of signals that either piecewise constant (PWC) or sparse in the Harr wavelet domain.
3.1 System Identification

The application of Toeplitz CS matrices in identifying a LTI system with finite sparse impulse responses has been discussed in [1]. Here we discuss the symmetric Toeplitz CS matrices in the identification of such system. Let \(x[0], x[1], \ldots, x[n-1]\) be \(m\)-sparse impulse responses of a LTI system (of duration \(n\)). Let \(a[0], a[1], \ldots, a[n-2], a[n-1], a[n-2], \ldots, a[n-k]\) be a sequence of duration \((n+k-1)\) which are symmetric to the time \((n-1)\) (i.e. \(a[n-1+t] = a[n-1-t]\) for \(t = 1, 2, \ldots, k-1\)), where \(a[0], a[1], \ldots, a[n-1]\) have been drawn from one of the probability distributions given in (2.3). So, \(k-1 = O(\ln n)\) less entries are reduced to be generated. Then, probing the given system with \(a[l]\) yields \(y[l] = a[l] * x[l]\) and the theory of CS along with Theorem 2.2 guarantees that, with high probability, \(x[l]\) can be exactly recovered by solving the convex program

\[
x[l] = \arg \left( \min_{z \in \mathbb{R}^n} ||z||_1, \text{ subject to } y = Az \right),
\]

where \(y[n-1]
\begin{bmatrix}
    y[n-1] \\
    y[n] \\
    \vdots \\
    y[n+k-2]
\end{bmatrix}
\]

and \(A = \begin{bmatrix}
    a[n-1] & a[n-2] & \cdots & a[1] & a[0] \\
    a[n-2] & a[n-1] & \cdots & a[2] & a[1] \\
    \vdots & \vdots & \ddots & \ddots & \vdots \\
    a[n-k] & a[n-k-1] & \cdots & \cdots & a[k-1]
\end{bmatrix}\).

3.2 Beyond Sparse Signals

When signals are sparse in some transform domain \(Ψ ≠ I\), i.e., \(x = Ψθ\) and \(θ \in \mathbb{R}^n\) is \(m\)-sparse, we need guarantee the product matrix \(AΨ\) satisfies RIP of order \(3m\) for successful recovery of \(θ\). It is unquestionable when \(A\) is a random matrix and \(Ψ\) is any orthonormal basis [2]. As described in [1], the Toeplitz matrices can still be used as CS matrices for some fixed transformations, even though they lack the universality property because of their highly structured nature. Here we still use the example in [1] for recovery of PWC signals by symmetric Toeplitz matrices. We benefit from this kind of matrices such as generation of only \(n\) independent random variables \((k-1 = O(\ln n)\) less random variables than Toeplitz matrices), and faster acquisition and reconstruction algorithms.

As an illustration, let \(x\) be a \(m\)-piece PWC signal with form \(x = Lθ\), where \(θ \in \mathbb{R}^n\) is \(m\)-sparse, and \(L \in \mathbb{R}^{n×n}\) is the matrix whose entries are 0’s above the diagonal and 1’s everywhere else. Further, let \(\{a_i\}_{i=1}^n\) be a sequence of random variables drawn independently from a distribution that yields an IID CS matrix, and \(A_L \in \mathbb{R}^{k×n}\) be a cascade of a \(k \times n\) symmetric Toeplitz matrix \(A\) and the \(n \times n\) differencing operator \(D\) (see [1]). Then, (i) \(A_L = AD\) has only \(n\) degrees of freedom, (ii) multiplication with \(A_L = AD\) (a Toeplitz matrix) requires only \(O(n \ln n)\) operations instead of \(O(n^2)\) work and (iii) the product matrix \(A_LL = ADL = A\) is a symmetric Toeplitz CS matrix satisfied RIP with high probability.
Now we compare the performances between the Gaussian matrices (Gauss), Toeplitz matrices (Toeplitz) and symmetric Toeplitz matrices (S-Toeplitz), where the entries of Toeplitz or S-Toeplitz matrices all obey Gaussian distribution. Let $x$ be a discrete signal with length 512 and the sparsity $m = 20$ whose nonzero entries are 1 or $-1$. The $\ell_1$-minimization is solved by the classical linear programming techniques; see [6]. The results of 100 experiments are summarized in Fig. 1. Next, we compare the performances through the real image reconstruction experiment. The original image is shown in Fig. 2, the size of this image is $64 \times 64$ and the sparsity is $m = 739$ in frequency domain. By Theorem 2.2, the measurement number $k = O(\ln n)$ for sufficiently large $n$. According to the experimental result in Fig. 1 and also the

![Fig. 1 Successful rate as a function of measurement number](image)

### 4 Experiment

The original image is shown in Fig. 2, the size of this image is $64 \times 64$ and the sparsity is $m = 739$ in frequency domain. By Theorem 2.2, the measurement number $k = O(\ln n)$ for sufficiently large $n$. According to the experimental result in Fig. 1 and also the

![Fig. 2 Real world data reconstruction](image)

- (a) original image, (b) Toeplitz (MSE=0.0668), (c) S-Toeplitz (MSE=0.0688)
empirical value, we set $k = 2400$. It is possible that $k$ could be taken a smaller value. However we aim to compare the reconstruction effects between Toeplitz and S-Toeplitz under the same number of measurements. The mean square error (MSE) is defined as $\text{MSE} = \frac{||X - M||_F}{||M||_F}$, where $||\cdot||_F$ is the Frobenius norm, $M$ is the original image and $X$ is the reconstruction. The $\ell_1$-minimization here is solved by the subspace pursuit algorithm proposed in [22]. The experimental results show that symmetric Toeplitz matrices are suitable CS matrices.

5 Conclusion

We show that symmetric Toeplitz-structured matrices and its variants are also sufficient to recover undersampled sparse signals. It provides a desirable alternative of measurement matrices for a number of applications and greatly reduces the computational and storage complexity in some large dimensional problems.

Acknowledgments Supported by National Natural Science Foundation of China (11071002, 11371028), Program for New Century Excellent Talents in University (NCET-10-0001), Key Project of Chinese Ministry of Education (210091), Specialized Research Fund for the Doctoral Program of Higher Education (20103401110002), Scientific Research Fund for Fostering Distinguished Young Scholars of Anhui University(KJIIQ1001), Academic Innovation Team of Anhui University Project (KJTD001B), Open Project of Key Laboratory of Opto-Electronic Information Acquisition and Manipulation Ministry of Education (OEIAM201406).

References

1. Bajwa, W., Haupt, J., Raz, G., Wright, S., & Nowak, R. (2007). Toeplitz structured compressed sensing matrices, In IEEE/SP Workshop on Statistical Signal Processing (SSP).
2. Baraniuk, R., Davenport, M., DeVore, R., & Wakin, M. (2008). A simple proof of the restricted isometry property for random matrices. Constructive Approximation, 28(3), 253–263.
3. Candès, E. J. (2008). The restricted isometry property and its implication for compressed sensing. Comptes Rendus Mathematique, 346(9), 589–592.
4. Candès, E., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete Fourier information. IEEE Transactions on Information Theory, 52(2), 489–509.
5. Candès, E. J., & Tao, T. (2006). Near optimal signal recovery from random projections: Universal encoding strategies. IEEE Transactions on Information Theory, 52(12), 5406–5425.
6. Candès, E. J., & Tao, T. (2005). Decoding by linear programming. IEEE Transactions on Information Theory, 51, 4203–4215.
7. Donoho, D. L. (2006). Compressed sensing. IEEE Transactions on Information Theory, 52(4), 1289–1306.
8. Fan, Y.-Z., Huang, T., & Zhu, M. (2012). Compressed sensing based on random symmetric Bernoulli matrix, arXiv:1212.3799.
9. Fan, Y.-Z., Huang, T., & Zhu, M. (2013). Mixed compressed sensing based on random graphs, arXiv:1307.2117.
10. Hajnal, A., & Szemerédi, E. (1970). Proof of a conjecture of P Erdős. In P. Erdős, A. Rényi, & V. T. Sós (Eds.), Combinatorial theory and its application (pp. 601–623). Amsterdam: north-holland.
11. Haykin, S. (2001). Adaptive filter theory. Upper Saddle River, NJ: Prentice Hall.
12. Rauhut, H. (2007). Random sampling of sparse trigonometric polynomials. Applied and Computational Harmonic Analysis, 22(1), 16–42.
13. Rauhut, H. (2008). Stability results for random sampling of sparse trigonometric polynomials. *IEEE Transactions on Information Theory, 54*(12), 5661–5670.
14. H. Rauhut, & E. Allee. (2009). Circulant and Toeplitz matrices in compressed sensing, *Computing Research Repository*, vol. abs/0902.4.
15. Rauhut, H., Romberg, J., & Tropp, J. (2012). Restricted isometries for partial random circulant matrices. *Applied and Computational Harmonic Analysis, 32*(2), 242–254.
16. Pemmaraju, S. (2001). Equitable coloring extends Chernoff-Hoeffding bounds, In *Proceedings of RANDOM-APPROX 2001*, Berkeley, CA, pp. 285–296.
17. Pfander, G., Rauhut, H., & Tropp, J. (2013). The restricted isometry for time-frequency structured random matrices, *Probability Theory Related Fields, 156*(3–4), 707–737.
18. Romberg, J., Raz, G., Wright, S., & Nowak, R. (2009). Compressive sensing by random convolution. *SIAM Journal on Imaging Sciences, 2*(4), 1098–1128.
19. Rudelson, M., & Vershynin, R. (2006). Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements, In *Conference on Information Sciences and Systems (CISS)*.
20. Rabiner, L. R., Crochiere, R. E., & Allen, J. B. (1978). FIR system modeling and identification in the presence of noise and with band-limited inputs. *IEEE Transactions on Acoustics, Speech and Signal Processing, 26*(4), 319–333.
21. Ljung, L. (1987). *System identification: Theory for the user*. Englewood Cliffs NJ: Prentice Hall.
22. Dai, W., & Milenkovic, O. (2008). Subspace pursuit for compressive sensing: Closing the gap between performance and complexity, Computing Research Repository, vol. abs/0803.0.
23. Tropp, J., Wakin, M., Duarte, M., Baron, D. & Baraniuk, R. (2006). Random filters for compressive sampling and reconstruction, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France.