Compressed Sensing Performance of Binary Matrices with Binary Column Correlations

Weizhi Lu, Tao Dai and Shu-Tao Xia

Graduate School at Shenzhen, Tsinghua University
Shenzhen, Guangdong, 518055, China
wzlusd@sz.tsinghua.edu.cn

Abstract

This paper studies a class of binary matrices with correlations between distinct columns only equal to zero or one, which has reported comparable performance with random matrices in recent studies of compressed sensing. For such matrix, we analyze its structure property and provide an improved performance estimation.

1 Introduction

Compressed sensing aims to recover a signal $\mathbf{x} \in \mathbb{R}^n$ with at most k nonzeros from a linear measurement $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^n$ with $m < n$, and this goal can be realized by solving a convex optimization problem

$$\min \|\hat{\mathbf{x}}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\hat{\mathbf{x}} \tag{1}$$

if the sparsity k of $\mathbf{x}$ is sufficiently small [1]. In practice, we usually wish to maximize the guaranteed sparsity k provided the matrix size $m \times n$, or minimize the compression ratio of m/n given the sparsity k. This needs us to optimize the measurement matrix $\mathbf{A}$. Now it is known that some random matrices, such as Gaussian matrices and Bernoulli matrices, can achieve the optimal performance $k \approx m/\log(n/m)$ with high probability [2]; however, the random structure restricts their practical applications. In terms of implementation complexity, we are more interested in deterministic matrices, especially those only with 0-1 binary elements.

Currently, binary matrices are mainly constructed with coding theory, based on the fact that codewords usually have large mutual distances and thus strong orthogonality [3–6]. Moreover, such matrix can also be obtained by sampling some known orthogonal systems [7]. However, the two construction methods mentioned above can only provide explicit performance guarantees for some particular matrix sizes, which cannot meet the needs of practical applications. To address this problem, we turn our attention to the parity-check matrices of LDPC codes, which have fixed structures and comparable performance to random matrices [8]. Furthermore, as detailed later, such kind of matrices allows to be constructed with arbitrary size, as the compression ratio m/n is sufficiently large [8]. The construction can be implemented with a greedy algorithm, known as the progressive edge-growth (PEG) algorithm [9]. Such matrix has a special structure property, the correlation between distinct columns is 0-1 binary; and so it is simply called *the binary matrix with binary correlation*

(BMBC) hereafter. For simplicity, in the paper we mainly study the BMBC matrix with a uniform column degree d and row degree dn/m, namely with d nonzeros per column and dn/m nonzeros per row.

For BMBC matrices, in the paper we show that there exists an upper bound for column degree d and a lower bound for compression ratio m/n. Beyond the two bounds, the BMBC matrices will not exist; in other words, the property of binary column correlations cannot be obtained. The theoretical performance of BMBC matrix was firstly analyzed by Dimakis, et al [10]. Then, Liu and Xia provided an explicit bound $k < d$ for the guaranteed sparsity [11]. However, this bound is still far from the true performance; empirically, the true guaranteed sparsity tends to increase with the matrix size $m \times n$, rather than upper bounded by the column degree d. The estimation error comes from the fact that the two known performance estimators, NSP and RIP, cannot be exactly derived [12], and often they are roughly estimated with the maximum column correlation (namely the coherence). To cope with this problem, in the paper we derive the distribution of column correlation for BMBC matrix. This parameter helps us derive a better estimation for the RIP, then a more reasonable upper bound for the guaranteed sparsity, which is not only related to the column degree but also to the matrix size. Consider little experimental work has been reported in the previous literature, we test the performance of BMBC matrix over varying column degrees and compression ratios. The experiments demonstrate that the performance of BMBC matrix with given size tends to increase with the increase of column degree, and the maximum column degree tends to yield better recovery performance than Gaussian random matrices at high compression ratios.

The rest of the paper is organized as follows. In Section 2, we show that the BMBC matrix has an upper bound for column degree and a lower bound for compression ratio, by studying the associated bipartite graph. Moreover, we derive the distribution of column correlation for BMBC matrix. With such parameter, we provide an improved estimation for the guaranteed sparsity of BMBC matrix in Section 3. The empirical performance of BMBC matrix is tested in Section 4. The paper is concluded in Section 5.

2 Structure Property of BMBC Matrix

In this section, we show that the BMBC matrix can be associated with a bipartite graph without cycles of length 4. With the help of such graph, we easily analyze the structure property of BMBC matrix, including the upper bound of column degree, the lower bound of compression ratio and the distribution of column correlation.

2.1 Equivalence between binary matrix and bipartite graph

Fig.1(a) illustrates an example of bipartite graph, which contains m check nodes, n variable nodes, and a number of edges placed between two kinds of nodes. Such graph can be uniquely associated with an $m \times n$-sized binary matrix, as we make the variable nodes and check nodes correspond respectively to the matrix's columns and rows, and the edges correspond to the matrix's nonzero positions. Recall this

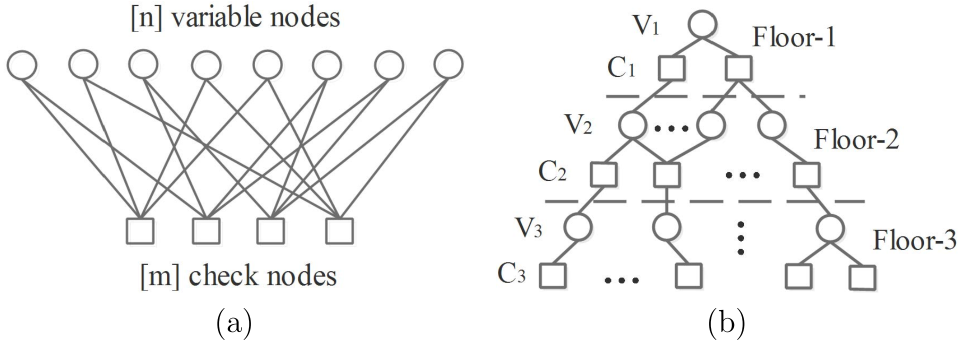

Figure 1: A bipartite graph in (a) and a tree expanded from a variable node in (b). The variable nodes and check nodes are represented with circles and squares, respectively.

paper mainly studies the BMBC matrix with a uniform column degree d and row degree nd/m, and the associated bipartite graph should have d edges incident to each variable node and nd/m edges linked to each check node.

From each variable node, as shown in Fig.1(b), we can expand a tree by traversing all reachable nodes through edges. The tree includes some cycles, each with length not smaller than 4. Here the length means the number of edges included in the cycle. For a bipartite graph with given size, it is known that the lengths of cycles tend to decrease with the increase of the number of edges. As detailed later, this implies that there should exist an upper bound for the number of edges, if we want to ensure the graph have no cycles of length 4. For the convenience of analysis, as shown in Fig.1(b), we divide the tree into multiple floors, each floor containing two kinds of nodes. For simplicity, we often use V_i (and C_i) to denote the set of variable nodes (and the set of check nodes) within the i-th floor of the tree. Observe that the set V_1 only contains the *root* variable node, which has the following property.

Property 1. Given a tree expanded from a variable node. Consider the equivalence between binary matrix and bipartite graph, we can say the *root* variable node of the tree correlates with each variable node in the 2nd floor (or say in V_2); and the correlation value is the number of the check nodes (in the 1st floor, or say in C_1) they commonly connect to.

The above property leads to the following equivalence relation.

Property 2. A BMBC matrix associates with a bipartite graph without cycles of length 4.

Proof. Consider a tree of the bipartite graph associated with the BMBC matrix (see Fig.1(b)). To make the column correlation of BMBC matrix not greater than one, with Property 1 we know each variable node within the 2nd floor of the tree (namely in V_2) can connect only one check node in the 1st floor (namely in C_1). In this case, the tree contains no cycles of length 4. Then the whole graph has no cycles of length 4. □

2.2 Structure property of BMBC matrix

In this part, we provide three crucial structure parameters for BMBC matrix, the upper bound of column degree, the lower bound of compression ratio and the distribution of column correlation, which are detailed as follows.

In Theorem 1, we derive an upper bound $d < \sqrt{m}$ for the column degree d, which implies a sparse matrix structure, friendly to hardware implementation. Meanwhile, we derive a lower bound $m/n \geq 4/(n+1)$ for the compression ratio m/n, in order to ensure the column degree $d \geq 2$. It is apparent that the lower bound on compression ratio will restrict the application of BMBC matrix at low compression ratios. However, it is easy to see that this bound tends to go down with the increase of signal length n, and this property is beneficial to high dimensional signals. Moreover, it is noteworthy that the two bounds we derive for column degree and compression ratio are not tight, and hard to be approached by the real BMBC matrices constructed with PEG algorithm.

Theorem 1. Consider a binary matrix $\mathbf{A} \in \{0,1\}^{m\times n}$ with column degree $d \geq 2$ and row degree nd/m. Suppose $\mathbf{A}_i^\top \mathbf{A}_{j\neq i} \in \{0,1\}$, where $\mathbf{A}_i$ denotes the i-th column of $\mathbf{A}$, $i \in [n]$. Then we can derive

$$d \leq \frac{1}{2n}(m + \sqrt{m^2 + 4mn(n-1)}) \tag{2}$$

and

$$\frac{m}{n} \geq \frac{4}{n+1} \tag{3}$$

Proof. Let us see the bipartite graph associated with $\mathbf{A}$, which should have n variable nodes and m check nodes, each variable node connecting $d \geq 2$ edges and each check node connecting nd/m edges. Expander a tree expanded from a variable node (see Fig.1(b)); and the tree should have no cycles of length 4 according to Property 2. As stated before, let us use V_i and C_i to denote the variable nodes in the $i-th$ floor of the tree. Apparently, $|V_1| = 1$ and $|C_1| = d|V_1| = d$. Since the three has no cycles of 4, $|V_2| = (\frac{dn}{m} - |V_1|)|C_1| = d(\frac{dn}{m} - 1)$. Note that the number of variable nodes in the first 2 floors of the tree cannot exceed the total number of variable nodes, namely $|V_1| + |V_2| \leq n$, which leads to Eq. (2). Moreover, to make sure $d \geq 2$, we need the right-hand side of Eq. (2) to be greater than 2, then derive Eq. (3). □

Theorem 2 provides the probability η that the column correlation of BMBC matrix takes the value 1, rather than the value 0. It is easy to see that the value of η is determined by the matrix size and column degree, and tends to decrease with the increase of matrix size, as the compression ratio and column degree are fixed. This parameter will be used in the next subsection to estimate the guaranteed sparsity of BMBC matrix.

Theorem 2. Consider a binary matrix $\mathbf{A} \in \{0,1\}^{m\times n}$ with column degree d and row degree nd/m. Suppose $\mathbf{A}_i^\top \mathbf{A}_{j\neq i} \in \{0,1\}$, where $\mathbf{A}_i$ denotes the i-th column of $\mathbf{A}$, $i \in [n]$. Then,

$$\mathbf{A}_i^\top \mathbf{A}_{j\neq i} = \begin{cases} 1 & \text{with probability } \eta = \frac{(nd-m)d}{(n-1)m} \\ 0 & \text{with probability } 1-\eta \end{cases} \tag{4}$$

Proof. This result is derived based on the fact that among $n-1$ columns $\mathbf{A}_{j\in[n]\setminus i}$, only $\frac{(nd-m)d}{m}$ columns correlate with the column $\mathbf{A}_i$. The proof is similar to that of Theorem 1. Expand a tree from the variable node corresponding to $\mathbf{A}_i$, see the example in Fig. 1(b). With Property 1, we know the *root* variable node correlates only with the variable nodes in the 2nd floor, namely the nodes in V_2; and all the correlation values are equal to 1 since the graph has no cycles of length 4 according to Property 1. The number of variable nodes in V_2 has been derived as $|V_2| = \frac{(nd-m)d}{m}$ in the proof of Theorem 1. Then the proof is completed. □

3 Theoretical Performance of BMBC Matrix

In Theorem 3, we derive an upper bound of the guaranteed sparsity k for BMBC matrix, which is determined by two parameters, the column degree d and the probability η that the column correlation takes the value 1 rather than the value 0. Since $\eta \in (0,1)$, this bound should be much greater than the column degree d, namely the previous upper bound derived in [11]. This statement is verified by the examples shown in Table 1. As can be seen, the guaranteed sparsity derived with Eq.(5) presents the tendency of increasing with the matrix size, as the compression ratio and column degree are fixed.

Moreover, it is noteworthy that the guaranteed sparsity shown in Eq.(5) is derived under the condition of $d \to \infty$. Empirically, the condition on d is not strict due to the following fact. In the proof of Theorem 3, we derive Eq.(8) and (10) under the condition of $k \to \infty$. This condition naturally leads to the condition of $d \to \infty$, in terms of the relation of k and d shown in Eq.(5). In practice, the approximation errors for Eq.(8) and (10) usually can be ignored as k is not too large. This implies that Eq.(5) can be approximately derived as d is not too large.

Theorem 3. Consider a binary matrix $\mathbf{A} \in \{0, 1/\sqrt{d}\}^{m\times n}$ with column degree d and row degree nd/m. Suppose $\mathbf{A}_i^\top \mathbf{A}_{j\neq i} \in \{0, 1/d\}$, where $\mathbf{A}_i$ denotes the i-th column of $\mathbf{A}$, $i \in [n]$; and the probability of $\mathbf{A}_i^\top \mathbf{A}_{j\neq i} = 1/d$ is equal to η. By Eq. (1), a signal

Table 1: The guaranteed sparsity k derived with Eq.(5) for $m \times n$ BMBC matrix with column degree d. Recall the parameter η in Eq.(5) is derived with Eq.(4).

$m = 500$ and $n = 1000$									
d	2	3	4	5	6	7	8	9	10
k	9	13	15	15	15	15	15	15	15

$m = 5000$ and $n = 10000$									
d	2	3	4	5	6	7	8	9	10
k	92	134	149	152	152	149	145	141	137

$\mathbf{x} \in \mathbb{R}^n$ with at most k nonzeros can be recovered from $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m$, if

$$k < \eta^{-1} \left(\sqrt{d + 3 - 4\eta} - 2\sqrt{1 - \eta} \right)^2 \tag{5}$$

where $d \to \infty$.

Proof. The proof is established based on the result from [13]: If the measurement matrix $\mathbf{A}$ satisfies the RIP condition $\delta_k < 1/3$, then all signals $\mathbf{x}$ with at most $k(> 1)$ nonzeros can be recovered from $\mathbf{y} = \mathbf{A}\mathbf{x}$ via solving Eq. (1), where δ_k denotes the restricted isometry constant of $\mathbf{A}$; more precisely, δ_k is defined to be the minimal value belonging to the interval of $(0, 1)$ such that

$$1 - \delta_k \leq \|\mathbf{A}_\psi \mathbf{x}\|^2 / \|\mathbf{x}\|^2 \leq 1 + \delta_k \tag{6}$$

where $\mathbf{A}_\psi$ denotes a submatrix of $\mathbf{A}$ formed by columns indexed by $\psi \subset [n]$, $|\psi| = k$.

To derive the guaranteed sparsity k, the proof only needs to provide the solution of δ_k. Since $\lambda_{min}(\mathbf{A}_\psi^\top \mathbf{A}_\psi) \leq \|\mathbf{A}_\psi \mathbf{x}\|^2 / \|\mathbf{x}\|^2 \leq \lambda_{max}(\mathbf{A}_\psi^\top \mathbf{A}_\psi)$, where $\lambda_{min}(\mathbf{A}_\psi^\top \mathbf{A}_\psi)$ and $\lambda_{max}(\mathbf{A}_\psi^\top \mathbf{A}_\psi)$ denote the two extreme eigenvalues of $\mathbf{A}_\psi^\top \mathbf{A}_\psi$, we can write

$$\delta_k = \frac{\lambda_{max}(\mathbf{A}_\psi^\top \mathbf{A}_\psi) - \lambda_{min}(\mathbf{A}_\psi^\top \mathbf{A}_\psi)}{\lambda_{max}(\mathbf{A}_\psi^\top \mathbf{A}_\psi) + \lambda_{min}(\mathbf{A}_\psi^\top \mathbf{A}_\psi)} \tag{7}$$

wherein the two extreme eigenvalues are derived as follows.

For simplicity, let us write $\mathbf{B} = (\mathbf{A}_\psi^\top \mathbf{A}_\psi - \mathbf{I})$, where $\mathbf{I}$ denotes an identity matrix. Then $\mathbf{B}$ is a symmetric matrix with the diagonal elements equal to zero and the off-diagonal elements taking the nonzero value $1/d$ with probability η. Further, write

$$\mathbf{Q} = \frac{1}{\sqrt{\eta(1 - \eta)}} (d\mathbf{B} - \eta \mathbf{J})$$

where $\mathbf{J}$ is an all-ones matrix. As in [14], with Wigner semicircle law one can derive

$$\lambda_{min} \left(\frac{1}{\sqrt{k}} \mathbf{Q} \right) \geq -2$$

namely,

$$\lambda_{min}(d\mathbf{B} - \eta \mathbf{J}) \geq -2\sqrt{k\eta(1 - \eta)}, \tag{8}$$

if $|\psi| = k \to \infty$.

Note that both $d\mathbf{B} - \eta \mathbf{J}$ and $\eta \mathbf{J}$ are Hermitian matrices, and $\eta \mathbf{J}$ is positive semi-definite and has the rank equal to 1. With the Cauchy interlacing inequality [15], we have

$$\lambda_{min}(d\mathbf{B}) \geq \lambda_{min}(d\mathbf{B} - \eta \mathbf{J}),$$

namely

$$\lambda_{min}(\mathbf{B}) \geq \frac{1}{d} \cdot \lambda_{min}(d\mathbf{B} - \eta \mathbf{J}) \geq -\frac{2}{d}\sqrt{k\eta(1 - \eta)}.$$

It follows that

$$\lambda_{min}(\mathbf{A}_\psi^\top \mathbf{A}_\psi) = \lambda_{min}(\mathbf{B}) + 1 \geq -\frac{2}{d}\sqrt{k\eta(1-\eta)} + 1. \tag{9}$$

Since $\lambda_{max}(d\mathbf{B}) \approx k\eta + 1$ when $k \to \infty$ [16], we have

$$\lambda_{max}(\mathbf{A}_\psi^\top \mathbf{A}_\psi) = \lambda_{max}(\mathbf{B}) + 1 \approx \frac{1}{d}(k-1)\eta + 1. \tag{10}$$

Combine Eq.(7), (9) and (10),

$$\delta_k \lesssim \frac{k\eta + 2\sqrt{k\eta(1-\eta)} + 1}{k\eta - 2\sqrt{k\eta(1-\eta)} + 2d + 1}.$$

By letting $\delta_k < 1/3$, we can simply derive Eq. (5). Recall Eq.(8) is derived under the condition of $k \to \infty$, which together with the result of Eq.(5) implies the condition of $d \to \infty$. Then the proof is completed.

□

4 Simulation

In this section, we test the recovery performance of BMBC matrices using the known recovery algorithm, the subspace pursuit algorithm [17]. The input sparse signals have nonzero elements drawn from $\{\pm 1\}$ with equal probability. The recovery error is measured with $\|x - \hat{x}\|/\|x\|$, for which we take the average of 10^4 simulation runs. The recovery is regarded successful, as the average error is smaller than 10^{-4}. By this means, we can find a guaranteed sparsity k for given matrix. As in [8], the BMBC matrices are constructed with PEG algorithm [9]. As shown later, the simulation investigates two interesting problems for BMBC matrices, the performance change over the varying column degree and the performance comparison with random matrices.

4.1 Performance change over the varying column degree

Fig.2 shows the guaranteed sparsity k of 500×1000 BMBC matrix with varying column degree d. It is apparent that the guaranteed sparsity tends to grow with the increase of column degree. Recall the BMBC matrix has an upper bound for the column degree, as the matrix size is given. Then we are motivated to maximize the column degree of BMBC matrix, so as to achieve its best performance. As in [8], we can construct such matrix with PEG algorithm.

4.2 Performance comparison with random matrices

Here we construct a group of BMBC matrices with different compression ratios m/n, $n = 1000$. To obtain the best performance for each BMBC matrix, we maximize their column degrees, whose values are detailed in Table 2. Then we compare these matrices with random matrices, including the Gaussian random matrices and binary

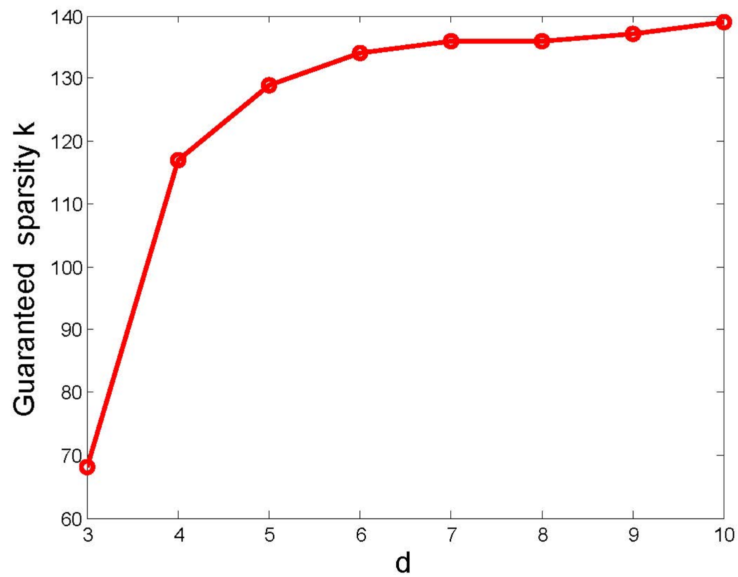

Figure 2: The guaranteed sparsity k of 500×1000 BMBC matrix with column degree d.

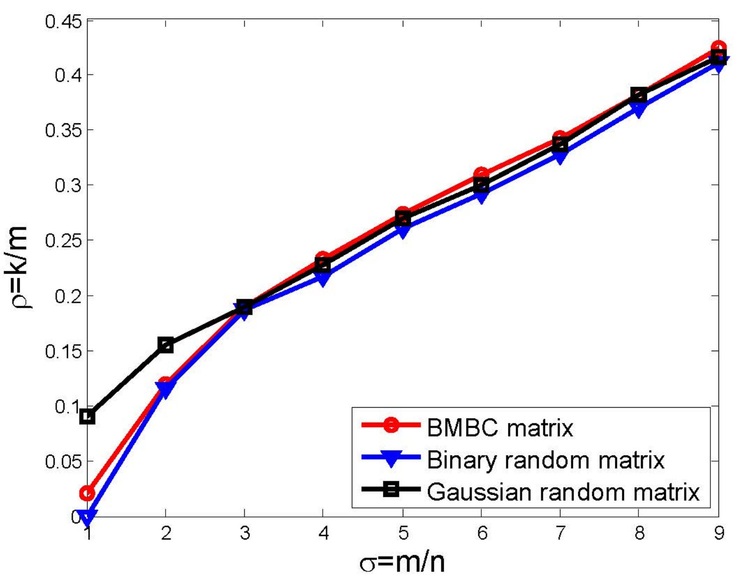

Figure 3: The guaranteed sparsity k (depicted with $\rho = k/m$) of three types of matrices with varying compression ratios $\sigma = m/n$, $n = 1000$. Note the BMBC matrix and binary random matrix share the same column degree as detailed in Table 2.

random matrices. For fair comparison, we make the binary random matrix share the same column degree with the BMBC matrix with the same size. From Fig.3, it can be seen that BMBC matrices always perform better than binary random matrices, and also outperform Gaussian random matrices as the the compression ratio is sufficiently large, i.e. $m/n > 0.3$.

Table 2: The column degree of $m \times 1000$ BMBC matrix constructed with PEG algorithm. Each column degree has achieved its upper bound that can be obtained with PEG algorithm.

m	100	200	300	400	500	600	700	800	900
d	3	5	7	8	10	11	12	14	15

Moreover, it is noteworthy that the maximal column degree we derive with PEG algorithm for BMBC matrix, as detailed in Table 2, is usually smaller than the theoretical bound we can derive with Eq.(2). As stated before, this is because the bound in Eq.(2) is not tight. This problem also occurs to the lower bound of compression ratio derived in Eq.(3). With PEG algorithm, we can hardly construct a BMBC matrix with compression ratio achieving the lower bound shown in Eq.(3).

5 Conclusion

This paper has studied the compressed sensing performance of binary matrix with binary column correlations, called BMBC matrix for shorthand. It is proved that such kind of matrices has an upper bound for column degree, as the matrix size is fixed. Empirically, the BMBC matrix with its maximum column degree achieved tends to perform better than the binary random matrix with the same column degree, and also outperforms Gaussian random matrices as the compression ratio m/n is sufficiently large. Besides performance advantage, the BMBC matrix also has the advantage of storage and computation, because its maximum column degree is much smaller than $\sqrt{m}$. The major drawback of such matrix is that it will not exist if the compression ratio is below some threshold, such as $n/m < 4/(n+1)$; fortunately, this lower bound tends to descend with the increase of signal length. In summary, we can say the BMBC matrix with high compression ratio is suitable for compressed sensing both in terms of performance and complexity.

References

[1] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, dec. 2005.

[3] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4-6, pp. 918 – 925, 2007.

[4] S. Howard, A. Calderbank, and S. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes," in *42nd Annual Conference on Information Sciences and Systems (CISS 2008)*, march 2008, pp. 11–15.

[5] A. Amini and F. Marvasti, "Deterministic construction of binary, bipolar, and ternary compressed sensing matrices," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2360–2370, April 2011.

[6] X.-J. Liu and S.-T. Xia, "Constructions of quasi-cyclic measurement matrices based on array codes," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2013, pp. 479–483.

[7] S. Li and G. Ge, "Deterministic sensing matrices arising from near orthogonal systems," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2291–2302, April 2014.

[8] W. Lu, K. Kpalma, and J. Ronsin, "Sparse binary matrices of LDPC codes for compressed sensing," in *Data Compression Conference (DCC), 2012*, April 2012, p. 405.

[9] X.-Y. Hu, E. Eleftheriou, and D. Arnold, "Regular and irregular progressive edge-growth Tanner graphs," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 386–398, jan. 2005.

[10] A. Dimakis, R. Smarandache, and P. Vontobel, "LDPC codes for compressed sensing," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3093–3114, May 2012.

[11] X.-J. Liu and S.-T. Xia, "Reconstruction guarantee analysis of binary measurement matrices based on girth," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2013, pp. 474–478.

[12] A. M. Tillmann and M. E. Pfetsch, "The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing," *IEEE Transactions on Information Theory*, Feb. 2014.

[13] T. T. Cai and A. Zhang, "Sharp rip bound for sparse signal and low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 74–93, 2013.

[14] L. V. Tran, V. H. Vu, and K. Wang, "Sparse random graphs: Eigenvalues and eigenvectors," *Random Structures & Algorithms*, vol. 42, no. 1, pp. 110–134, 2013.

[15] T. Tao and V. Vu, "Random matrices: Universality of local eigenvalue statistics," *Acta Mathematica*, vol. 206, pp. 127–204, 2011.

[16] T. Ando, Y. Kabashima, H. Takahashi, O. Watanabe, and M. Yamamoto, "Spectral analysis of random sparse matrices," *IEICE Transactions*, pp. 1247–1256, 2011.

[17] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.