Capacity Achieving Linear Codes with Random Binary Sparse Generating Matrices over the Binary Symmetric Channel

A. Makhdoumi Kakhaki[1,3], H. Karkeh Abadi[1,4], P. Pad[1,5], H. Saeedi[2], F. Marvasti[1], K. Alishahi[1]
[1]Advanced Communications Research Institute, Sharif University of Technology, Tehran, Iran
[2]Department of ECE, Tarbiat Modares University, Tehran, Iran
[3]Department of EECS, Massachusetts Institute of Technology, Cambridge, MA, USA
[4]Department of EE, Stanford University, Stanford, CA, USA
[5]School of Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
Email: makhdoum@mit.edu, hosseink@stanford.edu, pedram.pad@epfl.ch, hsaeedi@modares.ac.ir
{marvasti, alishahi}@sharif.edu

***Abstract*—In this paper, we prove the existence of capacity achieving linear codes with random binary sparse generating matrices over the Binary Symmetric Channel (BSC). The results on the existence of capacity achieving linear codes in the literature are limited to the random binary codes with equal probability generating matrix elements and sparse parity-check matrices. Moreover, the codes with sparse generating matrices reported in the literature are not proved to be capacity achieving for channels other than Binary Erasure Channel. As opposed to the existing results in the literature, which are based on optimal maximum a posteriori decoders, the proposed approach is based on a different decoder and consequently is suboptimal. We also demonstrate an interesting trade-off between the sparsity of the generating matrix and the error exponent (a constant which determines how exponentially fast the probability of error decays as block length tends to infinity). Based on our results, we also propose a channel coding rate achievable by linear codes at a given block length and error probability. Moreover, we prove the existence of capacity achieving linear codes with a given (arbitrarily low) density of ones on rows of the generating matrix. In addition to proving the existence of capacity achieving sparse codes, an important conclusion of our paper is to prove that any arbitrarily selected sequence of sparse generating matrices is capacity achieving with high probability.**

I. Introduction

The Shannon coding theorem [1] states that for a variety of channels with a given capacity C, if the information transmission rate R over the channel is below C, there exists a coding scheme for which the information can be transmitted with an arbitrarily low probability of error. For Discrete Memoryless Channels (DMC), it has been shown [2] that the probability of error can be bounded between two exponentially decaying functions of the codeword blocklength, n. In this theorem, there is no constraint on the codes in terms of linearity. In [3], a simpler proof of the Shannon theorem has been provided. The existence of capacity achieving linear codes over the Binary Symmetric Channel (BSC) was shown by Elias [4] where it was also proved that random linear codes have the same error exponent as random codes. A similar result has been obtained in [5]. It was recently shown in [6] that the error exponent of a typical random linear code can, in fact, be larger than a typical random code, implying a faster decaying of error as n increases. Some bounds on the decoding error probability of linear codes have been derived in [7]. The result reported in [4]-[7] are all based on the fact that the elements of generating matrices of the capacity achieving linear codes should be one or zero with equal probability; therefore the generating matrix of such approaches are not sparse.[1] Moreover, most papers on capacity achieving sparse linear codes are concentrated on codes with sparse parity-check matrices. In particular, an important class of codes called Low-Density Parity-Check (LDPC) codes [8], [9] have been of major interest in the past decade. While these codes have sparse parity-check matrices, they do not necessarily exhibit sparse generating matrices which are the focus of this paper. In [10]-[11], some Low-Density Generating-Matrix (LDGM) schemes have been proposed which have performance approaching the capacity.[2] Some other related literature on the codes with sparse generating matrices having performance close to capacity includes [12]-[14]; in [12], a capacity-achieving scheme has been proposed based on serially concatenated codes with an outer LDPC code and an inner LDGM code. However, the generating matrix corresponding to the concatenation is not necessarily sparse. On the other hand, rateless codes have been proposed in [13] and [14] which have sparse generating matrices but are only proved to be capacity achieving over the Binary Erasure Channel (BEC).

In this paper, using a novel approach, we prove the existence of capacity achieving linear codes with *sparse generating matrices* that can provide reliable communications over the BSC

[1]A sparse generating matrix is a matrix with a statistically low density of ones, see Section II for the exact definition.

[2]We distinguish between "capacity approaching" and "capacity achieving" codes. The former term is used when the performance of the code can be shown numerically to approach capacity without any guarantee to achieve it. The latter term is used if the performance can be rigorously proved to achieve the capacity. The subject of this paper is on the latter case.

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

at rates below the channel capacity. The proof is accomplished by first deriving a lower bound on the probability of correct detection for a given generating matrix and then by taking the expectation of that lower bound over all possible generating matrices with elements 1 and 0 with probability ρ and $1-\rho$, respectively. By showing that this expectation goes to one as n approaches infinity, we prove the existence of linear capacity achieving codes. To show the sparsity, we extend this result by taking the expectation over a subset of matrices for which the density of ones could be made arbitrarily close to any target ρ. We then prove a stronger result that indicates the existence of capacity achieving linear codes with the same low density of ones in each row of the generating matrix. In addition to proving the existence of capacity achieving sparse codes, we also show that for a sufficiently large code length, no search is necessary in practice to find the desired deterministic matrix. This means that a randomly chosen code can have the desired error correcting property with high probability. This is done by proving that the error probability of a sequence of codes, corresponding to a randomly selected sequence of sparse generating matrices tends to zero as n approaches infinity, in probability. This important result is then extended to generating matrices with low density rows. As opposed to the existing results in the literature, which are based on Maximum A Posteriori (MAP) decoders, the proposed proofs are based on a suboptimal decoder,[3] which makes our approach also novel from decoder point of view.

Although in reality the block length of codes is finite, in order to prove that a class of codes is capacity achieving, we assume that the block length goes to infinity. An interesting question is that for a given error probability and block length, how close to capacity the rate of the code can be. An upper bound for the coding rate achievable at a given block length and error probability is the sphere packing bound (see Equation (5.8.19) in [4]). In [15], for a given block length and error probability performance, the authors have obtained a lower bound on the achievable rate called Random Coding Union (RCU) bound. In this paper, we compare the rate of our sparse linear codes with these bounds. We also demonstrate an interesting trade-off between the sparsity of the generating matrix and the error exponent such that the sparser the matrix, the smaller the error exponent becomes.

It is important to note that we rigorously prove the existence of capacity achieving linear codes for a constant ρ resulting in a non-vanishing density of ones on the generating matrix as n tends to infinity. However, we have made a conjecture that if we choose $\rho(n)$ of $O(\frac{\log n}{\sqrt{n}})$, the resulting codes can still be capacity achieving, which implies a vanishing density of ones. It is worth mentioning that in the full version of this paper [16], we have proved similar results on the existence of capacity achieving linear sparse codes over the BEC. In particular, we have been able to prove that to have capacity achieving generating matrices, $\rho(n)$ can be of $O(\frac{\log n}{n})$. This implies that the number of ones in the generating matrix is about $n \log n$ which is asymptotically less than $n^{3/2} \log n$, the number of ones in the case of BSC.

[3]See the details in the next section.

The organization of the paper is as follows: In the next section, some preliminary definitions and notations are presented. In Sections III we present our theorems for BSC and Section IV concludes the paper. Due to lack of space, we have omitted the proof of theorems and propositions. They can be found in the full version of the paper [16].

II. Preliminaries

Consider a Discrete Memoryless channel (DMC) which is characterized by $\mathcal{X}$ and $\mathcal{Y}$ as its input and output alphabet sets, respectively, and the transition probability function $\mathbb{P}(y|x)$, where $x \in \mathcal{X}$ is the input, and $y \in \mathcal{Y}$ is the output of the channel. In this paper, we consider the binary case where $\mathcal{X} = \{0,1\}$. A binary code $\mathcal{C}(n,k)$ of rate R is a mapping from the set of 2^k k-tuples X_i to n-tuples Z_i, $0 \leq i \leq 2^k - 1$, where $X_i \in \{0,1\}^k$, $Z_i \in \{0,1\}^n$, and the code rate R is defined as the ratio of k by n. Since we are only interested in *Linear Codes*, the mapping is fully specified by an $n \times k$ binary matrix $\mathbf{A} = \{A_{ij}\}$ (the generating matrix), and encoding is accomplished by a left multiplication by $\mathbf{A}$:

$$Z_i = \mathbf{A} X_i,$$

where the calculations are in $\mathbb{GF}(2)$. The vector Z_i is then transmitted through the DMC. Decoding is defined as recovering the vector X_i from the possibly corrupted received version of Z_i.

In this paper the employed decoding scheme relies on the a posteriori probability distribution. Let $\mathbf{A}$ be the generating matrix. For a received vector $Y = y$, the decoder allocates a random vector such as $X = x$ as the original transmitted message with the conditional probability $\mathbb{P}(X = x|Y = y)$. Clearly, the probability of correct detection using $\mathbf{A}$ as the generating matrix is

$$p_c(\mathbf{A}) = \sum_{i,j} \mathbb{P}(x_i)\mathbb{P}(y_j|x_i)\mathbb{P}(x_i|y_j) =$$

$$\sum_{i,j} \mathbb{P}(x_i, y_j)\mathbb{P}(x_i|y_j) = \mathbb{E}_{X,Y}(\mathbb{P}(X|Y)), \qquad (1)$$

where $\mathbb{P}(X,Y)$ depends on $\mathbf{A}$.

Note that the optimal decoder is a MAP decoder which allocates $argmax_x \mathbb{P}(X = x|Y = y)$ and that the probability of correct detection using MAP is more than or equal to the probability of correct detection in (1). Throughout the paper, the index i in X_i and Z_i may be dropped for more clarity. For the sake of convenience, the following notations are used for the remainder of the paper.

***Definition** 1:* Let $\mathcal{A}_{n\times k}$ be the set of all binary $n \times k$ matrices. The density of an $\mathbf{A} \in \mathcal{A}_{n\times k}$ is defined as the total number of ones within the matrix divided by the number of its elements (nk). A matrix with a density less than 0.5 is called sparse; the smaller the density, the sparser the matrix becomes.

***Definition** 2:* Let each entry of each element of $\mathcal{A}_{n\times k}$ has a Bernoulli(ρ) distribution, $0 < \rho < 1$.[4] This scheme induces a *probability distribution* on the set $\mathcal{A}_{n\times k}$, denoted by Bernoulli(n, k, ρ). For the rest of paper, we consider this distribution on the set $\mathcal{A}_{n\times k}$.

Note that as n approaches infinity, the typical matrices of $\mathcal{A}_{n\times k}$ have a density close to ρ.

III. Capacity Achieving Sparse Linear Codes for Binary Symmetric Channel

Consider a BSC with cross-over probability ϵ. The capacity of this channel is given by $C = 1 - h(\epsilon)$, where $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon)$. We suppose that R, the rate of the code, is less than C. In this section, we prove the existence of capacity achieving linear codes with arbitrarily sparse generating matrices over the BSC. We prove the existence by showing that the average error probability over such generating matrices tends to zero as n approaches infinity.

A. Channel Modeling

Assume that we encode a message vector X to generate the codeword $\mathbf{A}X$. Note that X is chosen uniformly from the set $\{0,1\}^k$. Due to the effect of error in the BSC, each entry of the transmitted codeword $\mathbf{A}X$ can be changed from 0 to 1 and vice versa. These changes can be modeled by adding 1 to erroneous entries of $\mathbf{A}X$ (in $\mathbb{GF}(2)$). Therefore, the error of a BSC with cross-over probability ϵ can be modeled by a binary n-dimensional error vector N with i.i.d. entries with Bernoulli(ϵ) distribution. Thus, if the output of the channel is shown by Y, the following equation models the channel:

$$Y_{n\times 1} = \mathbf{A}_{n\times k} X_{k\times 1} + N_{n\times 1}. \tag{2}$$

Note that X and N are independent.

B. Capacity achieving sparse linear codes for the BSC

In the following theorem, a lower bound for the average probability of correct detection over the set $\mathcal{A}_{n\times k}$, is obtained.

***Theorem** 1:* Consider a BSC with cross-over probability ϵ. A lower bound for the average probability of correct detection over all $n \times k$ generating matrices with Bernoulli(n, k, ρ) distribution is given by

$$\mathbb{E}_{\mathbf{A}\in\mathcal{A}_{n\times k}}\left(p_c\left(\mathbf{A}\right)\right) \geq \sum_{i=0}^{n} \binom{n}{i} \times \frac{2^n \epsilon^{2i} (1-\epsilon)^{2(n-i)}}{\sum_{j=0}^{k} \binom{k}{j} (1-(1-2\epsilon)(1-2\rho)^j)^i (1+(1-2\epsilon)(1-2\rho)^j)^{n-i}}. \tag{3}$$

An important result of this theorem is that if we fix the average error probability, we can find the maximal achievable rate for a given block length over a given channel. Fig. 1 is a plot of the coding rate versus n for different values of ρ for a BSC with $\epsilon = .11$. Note that for this value of ϵ, the capacity of BSC is equal to 0.5. This plot is numerically evaluated from Theorem 1 where the average probability of error is set to 10^{-3}. As can be seen, at block length of 1000, we can achieve the rate of 0.4 which is about 80% of the capacity. Another interesting observation of this figure is that when the block length n increases, the achievable coding rate becomes independent of the density ρ. The significance of this observation is that sparse generating matrices can replace non-sparse ones for large block sizes which implies a simpler encoder structure.

Fig. 2 shows the comparison of our result to the sphere packing bound and the RCU bound of [15] for $\epsilon = .11$ and average probability error of 10^{-3}. As can be seen, the rate of our codes follows both bounds pretty closely. It is important to note that the RCU bound guarantees a lower bound on the achievable rate for codes which are not necessarily linear. Consequently, it is not surprising that the rate of our linear codes have not achieved the RCU bound.

Now using the results of Theorem 1, we want to prove the existence of capacity achieving linear codes. In the following theorem, we will show that the expected value of the correct detection probability over all generating matrices from $\mathcal{A}_{n\times k}$ approaches 1 indicting the existence of at least one linear capacity achieving code.

***Theorem** 2:* For any $0 < \rho < 1$, for a BSC we have

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{A}\in\mathcal{A}_{n\times k}}(p_c(\mathbf{A})) = 1. \tag{4}$$

The performance of linear codes is determined by the error exponent which is defined as follows:

***Definition** 3:* The error exponent of a family of codes $\mathcal{C}$ of rate R is defined as

$$E_{\mathcal{C}}(R) = \lim_{n\to\infty} -\frac{1}{n}\log p_e, \tag{5}$$

where p_e is the average probability of decoding error.

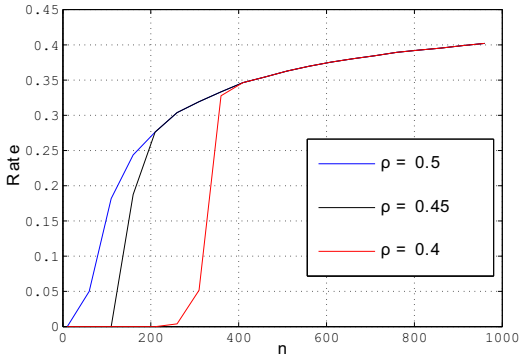

Fig. 1. Coding rate versus block length for different values of ρ with $\epsilon = .11$ and average error probability of 10^{-3}.

If the limit is greater than zero, the average error probability of the proposed codes decreases exponentially to zero as n increases. The error exponent is an index such that the larger

[4] A binary random variable has Bernoulli(ρ) distribution if it is equal to 1 with probability of ρ and equal to 0 with probability of $1-\rho$.

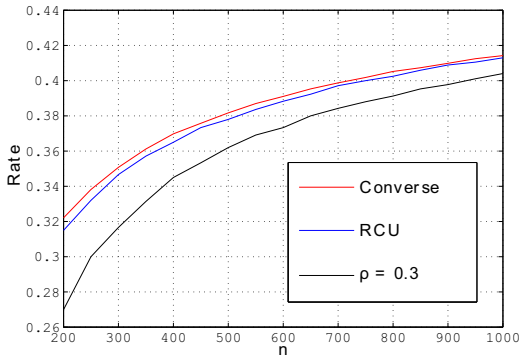

Fig. 2. Coding rate versus block length with $\epsilon = .11$ and average error probability of 10^{-3} for sphere-packing bound, Random Coding Union (RCU) bound [15] and our random linear code with $\rho = .3$.

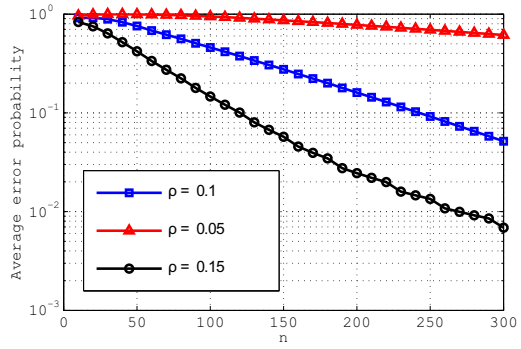

Fig. 3. The average error probability versus n for different values of ρ, $\epsilon = 0.05$, $R = 0.8C$.

the error exponent, the faster the probability of error decays as n increases. Based on our observation, there is an interesting relation between the error exponent of the codes constructed by generating matrices with Bernoulli(n, k, ρ) distribution and the values of ρ. In Fig. 3, we have plotted the average probability of error versus n for various values of ρ and a fixed code rate. As can be seen, the error exponent which is equal to the slope of the curves, increases as ρ increases (the generating matrix become less sparse). In other words, although the probability of error for sparse codes goes to to zero exponentially as n increases; this decrease is not as fast as high density codes.

***Definition** 4:* Let $W(A)$ be the number of ones in a given binary matrix A and η be an arbitrary positive constant. $\mathcal{T}^{\eta}_{n\times k}$ is defined as a subset of $\mathcal{A}_{n\times k}$ for which $|\frac{W(A)}{nk} - \rho| < \eta$, $\eta > 0$. By choosing a sufficiently small η, the set $\mathcal{T}^{\eta}_{n\times k}$ is in fact a subset of $\mathcal{A}_{n\times k}$ which contains matrices having density of ones arbitrarily close to any given ρ. Note that the probability distribution on $\mathcal{T}^{\eta}_{n\times k}$ is induced from the probability distribution on $\mathcal{A}_{n\times k}$.

In Theorems 1 and 2, we proved the existence of capacity achieving codes for any value of ρ. We did not explicitly prove the existence of sparse capacity achieving codes. However, using concentration theory [17], we can see that for a sufficiently large n, a randomly chosen matrix from $\mathcal{A}_{n\times k}$ is in the subset $\mathcal{T}^{\eta}_{n\times k}$ with high probability. In other words, we can state the following proposition which implies the existence of capacity achieving codes which are sparse.

***Proposition** 1:* Let $\mathcal{T}^{\eta}_{n\times k}$ be the set of typical matrices defined in Definition (4). We then have

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{A}\in\mathcal{T}^{\eta}_{n\times k}}(p_e) = 0. \tag{6}$$

***Definition** 5:* We define $\mathcal{R}_{n\times k}$ as the set of all binary $n\times k$ matrices with rows that have $k\rho$ ones. We also consider a uniform distribution on the set $\mathcal{R}_{n\times k}$ for the rest of the paper.

In the next theorem, we will prove a stronger result on capacity achieving sparse codes. We show the existence of capacity achieving matrices with rows containing exactly $k\rho$ ones. In other words, the density of ones in each row is exactly equal to ρ. This also implies that the generating matrix has a density of ones exactly equal to ρ. In Theorem 3, we shall derive a lower bound on the average probability of correct detection and in Theorem 4 we will prove that this lower bound tends to one. This shows that the average probability of error over the set $\mathcal{R}_{n\times k}$ approaches zero, implying the existence of capacity achieving codes with generating matrices taken from $\mathcal{R}_{n\times k}$.

***Theorem** 3:* For a binary symmetric channel with cross-over probability ϵ, a lower bound for the expected value of the probability of correct detection over all generating matrices in $\mathcal{R}_{n\times k}$ is given by

$$\mathbb{E}_{\mathbf{A}\in\mathcal{R}_{n\times k}}\left(p_c\left(\mathbf{A}\right)\right) \geq \sum_{i=0}^{n}\binom{n}{i}\epsilon^i(1-\epsilon)^{(n-i)}\times \frac{\epsilon^i(1-\epsilon)^{(n-i)}}{\sum_{j=0}^{k}\binom{k}{j}(\epsilon A_j+(1-\epsilon)B_j)^i((1-\epsilon)A_j+\epsilon B_j)^{n-i}}. \tag{7}$$

where

$$A_j = \sum_{\text{q is odd}} \frac{\binom{j}{q}\binom{k-j}{k\rho-q}}{\binom{k}{k\rho}}, \quad B_j = \sum_{\text{q is even}} \frac{\binom{j}{q}\binom{k-j}{k\rho-q}}{\binom{k}{k\rho}}.$$

***Theorem** 4:* For each $0 < \rho < 1$, we have

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{A}\in\mathcal{R}_{n\times k}}(p_c(\mathbf{A})) = 1. \tag{8}$$

In Theorems 1 and 2, we proved the existence of capacity achieving linear codes with generating matrices having Bernoulli(n, k, ρ) distribution by showing that the average probability of error over all generating matrices tends to zero as n approaches infinity. This implies that we may have to perform a search over $\mathcal{A}_{n\times k}$ to find such a matrix. Assume that we simply pick matrices randomly for each n from the set $\mathcal{A}_{n\times k}$. This constitutes a sequence of $n \times nR$ matrices. Now consider the resulting sequence of error probabilities corresponding to the sequence of generating matrices. In the following proposition, we shall prove that the limit of this sequence is zero in probability, i.e., a sequence of randomly chosen matrices is capacity achieving with high probability.

This suggests that for sufficiently large n, no search is necessary to find a desired deterministic generating matrix.

***Proposition** 2:* Let $\{\mathbf{A}_{n\times nR}\}_{n=0}^{\infty}$ be the sequence of matrices, where $\mathbf{A}_{n\times nR}$ is selected randomly from $\mathcal{A}_{n\times nR}$. If we denote the error probability of the generating matrix $\mathbf{A}_{n\times nR}$ over BSC by $p_e(\mathbf{A}_n)$, then $p_e(\mathbf{A}_n)$ converges in probability to zero as n tends to infinity.

***Note** 1:* If we use the result of Theorem 4, we can extend Proposition 2 to the case where we construct the matrix sequence by choosing the matrices from the set $\mathcal{R}_{n\times k}$. In other words, in order to have capacity achieving sequences of generating matrices for BSC with arbitrarily low density rows, we can simply pick generating matrices randomly from $\mathcal{R}_{n\times k}$.

At this stage, we have been able to rigorously prove the existence of capacity achieving sparse linear codes over the BSC. However for a given ρ, although the density of ones can be made arbitrarily small, it does not go to zero even when n approaches infinity. Let us assume the case where ρ is a decreasing function of n such that $\lim_{n\to\infty}\rho(n) = 0$, resulting in zero density of ones as n goes to infinity. In the following conjecture, we will propose a result indicating that this assumption can in fact be true. Although, we have not been able to rigorously prove the conjecture, a sketch of the proof has been presented in [16].

***Conjecture** 1:* For any $\rho(n)$ of $O(\frac{\log n}{\sqrt{n}})$, by assuming the Bernoulli$(n, k, \rho(n))$ distribution on the set $\mathcal{A}_{n\times k}$, we have

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{A}\in\mathcal{A}_{n\times k}}(p_c(\mathbf{A})) = 1 \tag{9}$$

IV. Conclusions

In this paper, a novel approach to prove the existence of capacity achieving sparse linear codes over the BSC was proposed. In Theorem 1, we derived a lower bound on the average probability of correct detection over the set $\mathcal{A}_{n\times k}$. In Theorem 2, we proved that the average probability of error over $\mathcal{A}_{n\times k}$ tends to zero. Then we proved the existence of sparse capacity achieving codes in Proposition 2. In Theorem 3, we derived a lower bound on the average probability of correct detection over the set $\mathcal{R}_{n\times k}$. Using this lower bound in Theorem 4, we proved the existence of capacity achieving codes with generating matrices with the same density in each row. In Proposition 2 and its preceding note, we showed that the error probability of codes corresponding to any randomly chosen sequence of generating matrices tends to zero in probability. This implies that for a sufficiently large n, a randomly chosen matrix from $\mathcal{A}_{n\times k}$ and $\mathcal{R}_{n\times k}$ will have the average error correcting capability. In addition, we conjectured that Theorem 2 can hold for the case where ρ is of $O(\frac{\log n}{\sqrt{n}})$. This implies that for a capacity achieving code over a BSC, the density of the generating matrix can approach zero. We also demonstrated an interesting trade-off between the sparsity of the generating matrix and the error exponent indicating that a sparser generating matrix results in a smaller error exponent. We also observed that fixing the average bit error rate and ϵ, the rates for the codes with generating matrices of higher densities are closer to capacity for small block sizes. For larger block sizes, however, the rate becomes independent of the generating matrix density. In our proofs, we have used a suboptimal decoder while previous works in the literature were based on a MAP decoder. This implies that we can get stronger results if we use the optimal MAP decoder.

For future work, one can try to rigorously prove Conjecture 1 and possibly extend it to the case of matrices in the set $\mathcal{R}_{n\times k}$. The improvement in the bounds using a MAP decoder can be an interesting topic to investigate. The extension of the results to other memoryless channels is another challenging topic to be explored. A very interesting work is to analytically derive the error exponent to prove the trade-off between error exponent and sparsity of the generating matrix.

Acknowledgment

The authors would like to thank Professor G. D. Forney for his valuable comments and suggestions and Mr. R. Farhoudi for his comments about proof of theorems.

References

[1] C. E. Shannon, A mathmatical theory of communications, *Bell Systems Technical Journal*, vol. 27, pp. 379-429, 1948.

[2] R. M. Fano, *Transmisson of Information*, The M.I.T. Press, Cambridge, 1961.

[3] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Transactions Information Theory*, vol. 11, no. 1, pp. 3-18, Jan. 1965.

[4] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons Inc. New York, NY, USA 1968, p. 204.

[5] M. Mezard and A. Montanari, *Information, physics, and computation,* Oxford University Press, USA, 2009, pp. 105-128.

[6] A. Barg and G. D. Forney, "Random codes: Minimum distances and error exponents," *IEEE Transactions Information Theory*, vol. 48, no. 9, pp. 2568-2573, Sept. 2002.

[7] G. Poltyrev, "Bounds on the decoding error probability of linear binary codes via their spectra," *IEEE Transactions Information Theory*, vol. 40, no. 4, pp. 1284-1292, Jul. 1994.

[8] R. G. Gallager, "Low density parity check codes," *IRE Transactions Information Theory*, vol. IT-8, pp. 21, Jan. 1964.

[9] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *IEE Electronics Letters*, vol. 33, no. 6, pp. 457-458, Jul. 1997.

[10] F. J. Vazquez-Araujo, M. Gonzalez-Lopez, L. Castedo, and J. Garcia-Frias, "Capacity approaching low-rate LDGM codes," *IEEE Transactions Communications*, vol. 59, no. 2, pp. 352-356, Feb 2011.

[11] J. Garcia-Frias and Z. Wei, "Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix," *IEEE Communications Letters*, vol. 7, no. 6, pp. 266-268, June 2003.

[12] H. Chun-Hao and A. Anastasopoulos, "Capacity-achieving codes with bounded graphical complexity and maximum likelihood decoding," *IEEE Transactions Information Theory*, vol. 56, no. 3, pp. 992-1006, March 2010.

[13] M. Luby, "LT codes," in Proc. *IEEE Symposium on Foundations of Computer Science*, pp. 271-280, 2002.

[14] A. Shokorollahi, "Raptor codes," *IEEE Transactions Information Theory*, vol. 52, no. 6, pp. 2551-2567, June 2006.

[15] Y. Polyanskiy, H. Vincent Poor, and S. Verdu, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions Information Theory*, vol. 56, issue 5, pp. 2307-2359, May 2010.

[16] A. Makhdoumi Kakhaki, H. Karkeh Abadi, P. Pad, H. Saeedi, F. Marvasti, and K. AlishahiA, "Capacity Achieving Linear Codes with Random Binary Sparse Generating Matrices," [Online], Available: http://arxiv.org/abs/1102.4099

[17] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Application*, Springer, 2009.