Optimality of LDGM-LDPC Compound Codes for Lossy Compression of Binary Erasure Source

Grégory Demay*, Vishwambhar Rathi*†, Lars K. Rasmussen*†
{gdem, vish, lkra}@kth.se
*School of Electrical Engineering, †KTH Linnaeus Centre ACCESS
KTH - Royal Institute of Technology, Stockholm, Sweden

***Abstract*—We consider the Binary Erasure Source (BES) introduced by Martinian and Yedidia. Based on the technique introduced by Martinian and Wainwright, we upper bound the rate-distortion performance of the check regular Poisson LDGM ensemble and the compound LDGM-LDPC ensemble for the BES. We also show that there exist compound LDGM-LDPC codes, with degrees independent of the block-length, which can achieve any given point on the Shannon rate-distortion curve of the BES.**

I. Introduction

Following the remarkable success of sparse graph codes for the channel coding problem, many researchers have explored their capabilities for various source coding problems. One of the first contributions in this direction was made in [1], where Martinian and Yedidia introduced the Binary Erasure Source (BES). They showed that Low-Density Generator Matrix (LDGM) codes, which are duals of capacity achieving Low-Density Parity-Check (LDPC) codes for the Binary Erasure Channel (BEC), are optimal zero-distortion compression codes for the BES.

The first lower bound for lossy compression of a Binary Symmetric Source (BSS) using LDGM ensembles was derived by Dimakis, Wainwright, and Ramchandran in [2]. In contrast, Kudekar and Urbanke derived lower bounds on the rate-distortion performance of individual LDGM codes for the BSS [3]. Based on the second moment method, Martinian and Wainwright further derived upper bounds on the lossy compression performance of Check Regular Poisson (CRP) LDGM ensembles for the BSS [4–6]. They also proposed a compound LDGM-LDPC ensemble based on CRP LDGM ensembles and regular LDPC ensembles [4]. In particular, they showed that a randomly chosen code from the compound ensemble under optimal decoding achieves Shannon's rate-distortion bound with high probability, and with degrees remaining independent of the block-length [7].

In an earlier paper, lower bounds on the rate distortion performance of LDGM codes were derived for the BES [8]. In the present paper we focus on obtaining upper bounds on the rate distortion performance of LDGM codes for the BES, with the further objective of gaining a deeper insight into the behaviour of sparse-graph codes used as lossy compressors. Based on the technique in [4–6], we derive upper bounds on the rate-distortion performance for the BES using codes from the CRP LDGM ensemble and from the compound LDGM-LDPC ensemble. We also prove the optimality of the compound construction for the BES.

The remainder of the paper is organized as follows. In Section II, we formally state the problem and provide the necessary background results and definitions. In Section III, we derive upper bounds on the rate-distortion function for the CRP LDGM ensemble and the compound LDGM-LDPC ensemble. The optimality of the compound LDGM-LDPC ensemble is proven in Section IV, and in Section V we conclude with some discussion.

II. Definitions and Notations

Consider the BES which was introduced in [1]. Its source alphabet is $\mathcal{A} = \{0, 1, \star\}$, where $\star$ is the erasure symbol. A BES whose source symbol can take on the value $\{\star\}$ with probability ϵ or the values $\{0, 1\}$ with equal probabilities is denoted by BES(ϵ). The source encoding of a BES corresponds to encoding it with a binary alphabet $\{0, 1\}$. The erasure symbol $\star$ can be encoded to either 0 or 1 without incurring any distortion penalty. Thus the distortion function $d(x, y)$ between x and y, where $x \in \mathcal{A}$ and $y \in \{0, 1\}$, is zero if $x = \star$ or $x = y$. Otherwise, it is equal to one. The BES models the situation where some of the source symbols are not relevant or corrupted by noise. These symbols correspond

to erasures and can be encoded by a zero or a one without incurring any penalty.

For the BES(ϵ), the Shannon rate-distortion function is described as $R_\epsilon^{\text{sh}}(D) = (1-\epsilon)[1-h(D/(1-\epsilon))]$ if $D < (1-\epsilon)/2$ and zero otherwise, where $h(x) \triangleq -x\log_2(x) - (1-x)\log_2(1-x), x \in [0,1]$ is the binary entropy function. We denote the natural logarithm by $\log$ and logarithm to base two by $\log_2$.

We consider binary LDGM codes, which are binary linear block codes defined by a sparse generator matrix. A code of rate $R \leq \frac{m}{n}$ maps binary sequences $w \in \{0,1\}^m$ into a codeword $c \in \{0,1\}^n$, $m \leq n$. More precisely, let $\mathbb{L}(\mathbf{G})$ be a binary LDGM code of rate $0 \leq R \leq 1$ and block-length n, generated by a sparse binary generator matrix $\mathbf{G} \in \{0,1\}^{m\times n}$

$$\mathbb{L}(\mathbf{G}) = \left\{c \in \{0,1\}^n : \exists w \in \{0,1\}^{nR} \text{ s.t. } c = w\mathbf{G}\right\}.$$

We consider the Check Regular Poisson (CRP) LDGM ensemble denoted by $\mathfrak{L}_P(d_c, m, n)$. A randomly chosen code $\mathbb{L}(\mathbf{G})$ belonging to $\mathfrak{L}_P(d_c, m, n)$ is generated by the following procedure. Each check node is connected to d_c information bits chosen uniformly at random and with replacement. The degree distribution of the information bits tends to a Poisson distribution as the block-length n increases.

We now define the LDGM-LDPC compound construction. A LDGM-LDPC code is denoted by $\mathbb{C}(\mathbf{G}, \mathbf{H})$, with the LDGM matrix $\mathbf{G} \in \{0,1\}^{m\times n}$ and the LDPC matrix $\mathbf{H} \in \{0,1\}^{k\times m}$, where n is the block-length of the overall code and m the length of the information sequence. Then,

$$\mathbb{C}(\mathbf{G},\mathbf{H}) = \{c \in \{0,1\}^n : \exists w \in \{0,1\}^m \text{ s.t. } c = w\mathbf{G} \text{ and } w\mathbf{H}^T = \mathbf{0}\}.$$

We denote the rate of the generator matrix $\mathbf{G}$ by R_G and the rate of the parity-check matrix $\mathbf{H}$ by R_H, then the rate of the compound LDGM-LDPC code $\mathbb{C}(\mathbf{G},\mathbf{H})$ is given by $R = R_G R_H$. An example of a LDGM-LDPC compound code is given in Figure 1.

We denote the compound LDGM-LDPC ensemble by $\mathfrak{C}(d_c, d_v, d_c', m, n)$. A random code $\mathbb{C}(\mathbf{G},\mathbf{H})$ from $\mathfrak{C}(d_c, d_v, d_c', m, n)$ is generated by choosing both uniformly at random, the generator matrix $\mathbf{G}$ from $\mathfrak{L}_P(d_c, m, n)$, and the parity-check matrix $\mathbf{H}$ from the standard (d_v, d_c')-regular LDPC ensemble [9].

A. Source Coding

We consider rate distortion encoding of a BES using the CRP LDGM ensemble and the compound LDGM-LDPC ensemble. For a given code $\mathbb{C}$, let N be the total number of codewords. We index the codewords as $\{C_1, \cdots, C_N\}$, where C_1 is the codeword generated by the all-zero information word.

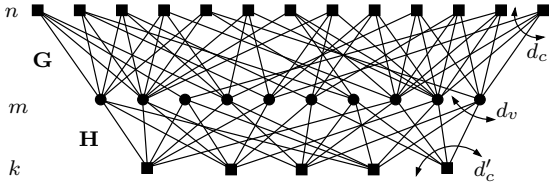

Fig. 1: Compound LDGM-LDPC code. The top layer is a (n, m) LDGM code whereas the bottom part is a (m, k) LDPC code. The top layer of nodes are the check bits of the LDGM code, the middle layer are the information bits of the LDGM code, and the bottom layer are the parity-check nodes.

For a source word $S \in \{0,1,\star\}^n$, let $X_i(\mathbb{C}, S, D)$ be the indicator function which evaluates to one if C_i is within distortion Dn from S. We define,

$$Z(\mathbb{C}, S, D) = \sum_{i=1}^{N} X_i(\mathbb{C}, S, D). \tag{1}$$

For the sake of notational simplicity, we drop the arguments and write X_i and Z. We want to derive a lower bound on the probability $\mathbb{P}\{Z > 0\}$. Let $\mathcal{S}_b$ be the set of source sequences with b erasures and let $\mathcal{S}_\mathcal{E}$ be the set of source sequences whose erasure positions are elements of the set $\mathcal{E}$, $\mathcal{E} \subset \{1,\ldots,n\}$. We write

$$\begin{aligned}\mathbb{P}\{Z>0\} &= \sum_{b=0}^{n}\sum_{\mathcal{E}:|\mathcal{E}|=b} \mathbb{P}\{S\in\mathcal{S}_\mathcal{E}\}\mathbb{P}\{Z>0|S\in\mathcal{S}_\mathcal{E}\},\\ &\overset{(a)}{=} \sum_{b=0}^{n}\binom{n}{b}\epsilon^b(1-\epsilon)^{n-b}\mathbb{P}\{Z>0|S\in\mathcal{S}_\mathcal{B}\},\end{aligned} \tag{2}$$

where $\mathcal{B} = \{1,\ldots,b\}$. (a) follows because of the i.i.d. behavior of BES. We are interested in finding a lower bound on the probability $\mathbb{P}\{Z > 0|S \in \mathcal{S}_\mathcal{B}\}$. Note that we are only interested in the growth rate of $\mathbb{P}\{Z > 0\}$ due to the following concentration result in [7].

Lemma II.1. *Assume that for a given distortion D and a given ensemble of codes, we have*

$$\lim_{n\to\infty}\frac{1}{n}\log_2\mathbb{P}\{Z(\mathbb{C},S,D)>0\} \geq 0. \tag{3}$$

Then, $\forall\theta > 0$, there exists a code in the ensemble such that for sufficiently large block-length n the average distortion is less than than $D+\theta$.

B. Second Moment Method

We derive a lower bound on $\mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\}$ by the second moment method,

$$\mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\} \geq \frac{E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right)^2}{E\left(Z^2|S \in \mathcal{S}_{\mathcal{B}}\right)}. \quad (4)$$

In the next lemma, we compute the first moment $E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right)$ for *any* linear code. We also show that if $nD \geq \frac{n-b}{2}$, then the growth rate of $\mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\}$ is zero.

Lemma II.2. *Consider any linear block code with rate R and block length n. The first moment $E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right)$ is given by*

$$E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right) = \begin{cases} o\left(2^{n-b}\right) 2^{nR - n(1-\beta)\left(1-h\left(\frac{D}{1-\beta}\right)\right)} & \text{if } nD \leq \frac{n-b}{2}, \\ o\left(2^{n-b}\right) 2^{nR} & \text{otherwise.} \end{cases}$$

where $\beta = \frac{b}{n}$. If $D \geq \frac{1-\beta}{2}$, then

$$\lim_{n\to\infty} \frac{\log_2\left(\mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\}\right)}{n} = 0.$$

Proof: By using the definition of Z in terms of random variables X_i given in (1), we obtain

$$\begin{aligned} E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right) &= \sum_{i=1}^{N} E\left(X_i|S \in \mathcal{S}_{\mathcal{B}}\right), \\ &= 2^{nR}\mathbb{P}\{X_1 = 1|S \in \mathcal{S}_{\mathcal{B}}\}. \end{aligned}$$

We obtain the expression for $E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right)$ by noting that

$$\mathbb{P}\{X_1 = 1|S \in \mathcal{S}_{\mathcal{B}}\} = \sum_{j=0}^{nD} \binom{n-b}{j} \frac{1}{2^{n-b}}$$

and the maximum of the summation term is attained at $j = nD$ if $nD \leq \frac{n-b}{2}$, otherwise it is attained at $\frac{n-b}{2}$.

Thus, when $D \geq \frac{1-\beta}{2}$, we obtain

$$\lim_{n\to\infty} \frac{\log_2\left(\mathbb{P}\{X_1 > 0|S \in \mathcal{S}_{\mathcal{B}}\}\right)}{n} = (1-\beta)\left(h\left(\frac{1}{2}\right) - 1\right) = 0.$$

The claim of the lemma follows by noting that $\mathbb{P}\{X_1 > 0|S \in \mathcal{S}_{\mathcal{B}}\} \leq \mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\}$ and probability of any event is upper bounded by one. ■

Remark: From the lemma above, we need to lower bound $\mathbb{P}\{Z > 0|S \in \mathcal{S}_{\mathcal{B}}\}$ only for $\beta < 1 - 2D$. In addition, the rate distortion performance analysis is non-trivial only for $D < (1-\epsilon)/2$.

In the next lemma, we compute the second moment of Z in terms of its expectation. The proof of this lemma is identical to that of Lemma 3 in [5].

Lemma II.3. *For any linear code, the second moment satisfies the relation*

$$E\left(Z^2|S \in \mathcal{S}_{\mathcal{B}}\right) = E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right) + E\left(Z|S \in \mathcal{S}_{\mathcal{B}}\right) \times \left(\sum_{j\neq 1} \mathbb{P}\{X_j = 1|X_1 = 1, S \in \mathcal{S}_{\mathcal{B}}\}\right). \quad (5)$$

In the next section, we derive upper bounds on the rate distortion performance of CRP LDGM ensembles and LDGM-LDPC ensembles for the BES.

III. Upper Bounds on the Rate Distortion Performance

Consider a randomly chosen code $\mathbb{L}(\mathbf{G}) \in \mathfrak{L}_P(d_c, m, n)$. Let $C(\nu)$ be a codeword which is generated by an information word of weight νm, $\nu \in [0,1]$. Then by the definition of the CRP LDGM ensemble each component of $C(\nu)$ is Bernoulli distributed with parameter $\delta(\nu, d_c) \triangleq \frac{1}{2}\left[1 - (1-2\nu)^{d_c}\right]$. Note that to simplify notation, we sometimes drop the arguments of $\delta(\nu, d_c)$ and denote it by δ. Define

$$\mathcal{Q}(\nu,\beta) \triangleq \mathbb{P}\{d\left(C(\nu), S\right) \leq nD \mid d\left(C_1, S\right) \leq nD, S \in \mathcal{S}_{\mathcal{B}}\}. \quad (6)$$

The following lemma bounds the exponential behavior of $\mathcal{Q}(\nu,\beta)$.

Lemma III.1. *Let $\beta < 1 - 2D$. For a randomly chosen code from the CRP LDGM ensemble $\mathfrak{L}_P(d_c, m, n)$ or the compound LDGM-LDPC ensemble $\mathfrak{C}(d_c, d_v, d_c', m, n)$, the exponential growth rate of the conditional probability defined in (6) is upper bounded as*

$$\frac{1}{n}\log_2 \mathcal{Q}(\nu,\beta) \leq F\left(\delta(\nu, d_c), \beta, D\right) + o(1), \quad (7)$$

where

$$\begin{aligned} F\left(\gamma,\beta,D\right) &= \inf_{\lambda<0} \max_{\tau\in[0,D]} G(\tau,\lambda,\gamma,\beta,D), \\ &= \max_{\tau\in[0,D]} G(\tau,\lambda^\star,\gamma,\beta,D). \end{aligned}$$

In order to define $G(\tau,\lambda,\gamma,\beta,D)$, consider $f_1(\gamma,\lambda) \triangleq (1-\gamma)e^{\lambda} + \gamma$. Then,

$$G(\tau,\lambda,\gamma,\beta,D) = (1-\beta)\left[h\left(\frac{\tau}{1-\beta}\right) - h\left(\frac{D}{1-\beta}\right)\right] + \tau\log_2\left(f_1(\gamma,\lambda)\right) + (1-\beta-\tau)\log_2\left(f_1(1-\gamma,\lambda)\right) - \frac{\lambda D}{\log 2}.$$

The definition of λ^* is based on the following quadratic equation

$$x^2(1-\beta-D)\delta(1-\delta)+x\left(\tau(1-2\delta)+\delta^2(1-\beta)\right) - x\left(D\left(\delta^2+(1-\delta)^2\right)\right)-D\delta(1-\delta)=0. \quad (8)$$

Let ρ^* be the only positive solution of (8). Then $\lambda^* = \min\left(0, \log\left(\rho^*\right)\right)$.

Proof: This proof is an extension of the proof for Lemma 5 in [7]. Since C_1 is the all-zero codeword, the condition $d\left(C_1, S\right) \leq nD$ is equivalent to $w_H(S) \leq nD$, where $w_H(S)$ denotes the Hamming weight of S (Hamming weight of an erasure is zero). Let T be the random variable corresponding to the Hamming weight of S, i.e., $T = w_H(S)$, knowing that S lies in $\mathcal{S}_{\mathcal{B}}$. Then

$$\mathbb{P}\{T=t\} = \frac{\binom{n-b}{t}}{\sum_{i=0}^{nD}\binom{n-b}{i}}.$$

Let Y be the random variable corresponding to the Hamming distance between $C(\nu)$ and S, when $S \in \mathcal{S}_{\mathcal{B}}$. Then

$$Y = \begin{cases} \sum_{j=1}^{T} U_j + \sum_{j=1}^{n-b-T} V_j, & \text{if } 1 \leq T \leq nD \\ \sum_{j=1}^{n-b-T} V_j, & \text{if } T = 0, \end{cases}$$

where U_j and V_j are independent Bernoulli random variables with parameters $1-\delta(\nu, d_c)$ and $\delta(\nu, d_c)$ respectively. Then,

$$\mathcal{Q}(\nu, \beta) = \mathbb{P}\{Y \leq nD\}. \quad (9)$$

To bound this probability we will use the Chernoff bound in the following manner

$$\frac{1}{n}\log_2 \mathbb{P}\{Y \leq nD\} \leq \inf_{\lambda<0}\left(\frac{1}{n}\log_2 \mathbb{M}_Y(\lambda) - \frac{\lambda D}{\log 2}\right), \quad (10)$$

where $\mathbb{M}_Y(\lambda)$ denotes the moment generating function of the random variable Y. Then we have

$$\mathbb{M}_U\left(\lambda\right) = (1-\delta)e^{\lambda}+\delta, \quad \mathbb{M}_V\left(\lambda\right) = \delta e^{\lambda}+1-\delta,$$

$$\mathbb{M}_Y\left(\lambda\right) = \sum_{t=0}^{nD}\mathbb{P}\{T=t\}\left[\mathbb{M}_U\left(\lambda\right)\right]^t\left[\mathbb{M}_V\left(\lambda\right)\right]^{n-b-t}.$$

Let $\tau = t/n$. Using Stirling's formula, we obtain

$$\frac{1}{n}\log_2 \mathbb{M}_Y\left(\lambda\right) = \frac{1}{n}\log_2\left\{\sum_{t=0}^{Dn} 2^{n(1-\beta)\left(h\left(\frac{\tau}{1-\beta}\right)-h\left(\frac{D}{1-\beta}\right)\right)} \times 2^{n(\tau\log_2(\mathbb{M}_U(\lambda))+(1-\beta-\tau)\log_2(\mathbb{M}_V(\lambda)))}\right\}+o(1).$$

Using this, (9), and (10) we have

$$\frac{1}{n}\log_2 \mathcal{Q}(\nu, D) \leq \inf_{\lambda<0}\max_{\tau\in[0,D]} G(\tau,\beta,\lambda,\delta,D)+o(1).$$

Using similar arguments as in the proof of Lemma 5 of [7], it can be shown that the order of infimum with respect to λ and maximum with respect to τ can be interchanged. Equating the partial derivative of $G(\tau,\beta,\lambda,\delta,D)$ with respect to λ to zero results in the quadratic equation (8) in terms of e^{λ}. Solving the quadratic equation gives the desired expression for $F\left(\gamma,\beta,D\right)$. This proves the lemma. ∎

Note that in the previous lemma we derived an upper bound on the growth rate of $\mathcal{Q}(\nu,\beta)$ for any fraction β of erasures such that $\beta < 1-2D$. However, from now on we will only consider $\beta = \epsilon$. This is because asymptotically the probability of having a source sequence with fraction of erasures equal to ϵ is almost equal to one. In the following lemma we derive an upper bound on the rate distortion performance of the CRP LDGM ensemble.

Lemma III.2. *Consider the CRP LDGM ensemble $\mathfrak{L}_P\left(d_c, m, n\right)$ with rate $R \leq \frac{m}{n}$. The rate distortion performance of $\mathfrak{L}_P\left(d_c, m, n\right)$ for the BES(ϵ) is upper bounded by*

$$R \geq \max_{\nu\in[0,1]} \frac{R_{\epsilon}^{sh}(D) + F(\delta\left(\nu, d_c\right), \epsilon, D)}{1-h(\nu)},$$

where $F(\delta\left(\nu, d_c\right), \epsilon, D)$ is defined in Lemma III.1.

Proof: Combining (4) and (5) we obtain

$$\frac{1}{n}\log_2\left(\mathbb{P}\{Z>0|S\in\mathcal{S}_{\mathcal{B}}\}\right) \geq \frac{1}{n}\log_2 E\left(Z|S\in\mathcal{S}_{\mathcal{B}}\right) - \frac{1}{n}\log_2\left(1+\sum_{j\neq 1}\mathbb{P}\{X_j=1|X_1=1, S\in\mathcal{S}_{\mathcal{B}}\}\right) \quad (11)$$

Assuming $\beta \leq 1-2D$, we can upper bound the last term in (11) by using (7).

$$\frac{1}{n}\log_2\left(1+\sum_{j\neq 1}\mathbb{P}\{X_j=1|X_1=1, S\in\mathcal{S}_{\mathcal{B}}\}\right) = \frac{1}{n}\log_2\left(\sum_{\substack{\nu\in[0,1]:\\ \nu m\in\mathbb{N}}}\binom{m}{\nu m}\mathcal{Q}\left(\nu,\beta\right)\right) \leq \max_{\nu\in[0,1]}\left\{Rh(\nu)+F\left(\delta(\nu,d_c),\beta,D\right)\right\} \quad (12)$$

Combining the last two equations and (2), and consid-

ering only the typical value $\beta = \epsilon$ we have

$$\frac{1}{n}\log_2 \mathbb{P}\{Z>0\} \geq R-(1-\epsilon)\left(1-h\left(\frac{D}{1-\epsilon}\right)\right)- \max_{\nu\in[0,1]}\{Rh(\nu)+F(\delta(\nu,d_c),\epsilon,D)\}+o(1). \quad (13)$$

As long as the RHS in (13) stays non-negative, we get the claim from Lemma II.1. ∎

In the following lemma we derive an upper bound on the rate distortion performance of the compound LDPC-LDGM ensemble.

Lemma III.3. *Consider the compound LDGM-LDPC ensemble $\mathfrak{C}(d_c,d_v,d_c',m,n)$ with overall rate R. Let $B(\nu)$ be an upper bound on the growth rate of the weight distribution of the (d_v,d_c')-regular LDPC ensemble. Then the rate distortion performance of $\mathfrak{C}(d_c,d_v,d_c',m,n)$ for the BES(ϵ) is upper bounded by*

$$R \geq \max_{\nu\in[0,1]} \frac{R_\epsilon^{sh}(D)+F(\delta(\nu,d_c),\epsilon,D)}{1-\frac{B(\nu)}{R_H}},$$

where $F(\delta(\nu,d_c),\epsilon,D)$ is defined in Lemma III.1.

Proof: The proof for this lemma is very similar to the proof of the Lemma III.2. The equivalent of (12) for the compound construction is

$$\sum_{j=1}^{N}\mathbb{P}\{X_j=1|X_1=1,S\in\mathcal{S}_{\mathcal{B}}\} = \sum_{\substack{\nu\in[0,1]:\\ \nu m\in\mathbb{N}}}\mathcal{A}_m(\nu)\mathcal{Q}(\nu,\beta),$$

where $\mathcal{A}_m(\nu)$ denotes the number of codewords of weight νm of the LDPC code. We then upper bound the growth rate of $\mathcal{A}_m(\nu)$ by $B(\nu)$ to obtain the desired result. ∎

In the next section, we prove the optimality of the compound construction for the BES(ϵ).

IV. Source coding optimality of the compound LDPC-LDGM ensemble

Before proving the optimality of the compound construction, we recall that $B(\nu)$ is an upper bound on the growth rate of the weight enumerator function for the LDPC code $A_m(\nu)$. Since the dependence of the function B on the degree pair (d_v,d_c') is obvious we will use both notations $B(\nu)$ or $B(\nu,d_v,d_c')$. The following lemma states some properties on the bounding function $B(\nu)$ which are proved in [7].

Lemma IV.1. *For a LDPC code with degrees (d_v,d_c'), where d_c' is even, an upper bound $B(\nu,d_v,d_c')$ on the growth rate of the weight enumerator function of the code can be defined for any $\nu\in\left[0,\frac{1}{2}\right]$ as*

$$B(\nu,d_v,d_c') = (1-d_v)h(\nu)-(1-R_H) + d_v \inf_{\lambda\leq 0}\left\{\frac{1}{d_c'}\log_2\left(\left(1+2^{\lambda}\right)^{d_c'}+\left(1-2^{\lambda}\right)^{d_c'}\right)-\nu\lambda\right\},$$

and $B(\nu)=B(1-\nu)$ for $\nu\in\left[\frac{1}{2},1\right]$. The function $B(\nu)$ we just defined satisfies the following conditions.

1) *$B(\nu)$ is symmetric around $\frac{1}{2}$.*
2) *$B(\nu)$ is twice differentiable on $(0,1)$ with $B'\left(\frac{1}{2}\right)=0$ and $B''\left(\frac{1}{2}\right)<0$.*
3) *$B(\nu)$ achieves its unique optimum at $\nu=\frac{1}{2}$, where $B\left(\frac{1}{2}\right)=R_H$.*
4) *$\exists\mu_1>0$ such that $\forall\nu\in(0,\mu_1)$, $B(\nu)<0$.*

The next lemma derives some properties of the function $F(\delta(\nu,d_c),\epsilon,D)$ (defined in Lemma III.1) which will be useful in proving the optimality of the compound construction.

Lemma IV.2. *For any even degree $d_c\geq 4$, the function $F(\delta(\nu,d_c),\epsilon,D)$ is differentiable in the neighborhood of $\nu=\frac{1}{2}$ with*

$$F\left(\delta\left(\frac{1}{2},d_c\right),\epsilon,D\right) = -R_\epsilon^{sh}(D) \quad (14)$$

$$\frac{\partial}{\partial\nu}F(\delta(\nu,d_c),\epsilon,D)\Big|_{\nu=\frac{1}{2}} = 0 \quad (15)$$

$$\frac{\partial^2}{\partial\nu^2}F(\delta(\nu,d_c),\epsilon,D)\Big|_{\nu=\frac{1}{2}} = 0 \quad (16)$$

Proof: Note that $\delta\left(\frac{1}{2}\right)=\frac{1}{2}$, $\forall d_c\geq 1$. Thus,

$$G\left(\tau,\lambda,\frac{1}{2},\beta,D\right) = (1-\beta)\log_2\left(f_1\left(\frac{1}{2},\lambda\right)\right)- \frac{\lambda D}{\log 2}+(1-\beta)\left[h\left(\frac{\tau}{1-\beta}\right)-h\left(\frac{D}{1-\beta}\right)\right]. \quad (17)$$

Moreover, $\max_{\tau\in[0,D]}\left\{h\left(\frac{\tau}{1-\beta}\right)-h\left(\frac{D}{1-\beta}\right)\right\}=0$, since $\frac{\tau}{1-\beta}\leq\frac{D}{1-\beta}\leq\frac{1}{2}$ and $h(\cdot)$ is an increasing function on $\left[0,\frac{1}{2}\right]$. Finally, the infimum of the first two terms in (17) is attained at $\lambda=-\log\left(\frac{1-\beta}{D}-1\right)$. Consequently we have

$$F\left(\delta\left(\frac{1}{2},d_c\right),\epsilon,D\right) = -R_\epsilon^{sh}(D).$$

Since $\delta(\nu,d_c)$ is twice differentiable in ν, $F(\delta(\nu,d_c),\epsilon,D)$ is also twice differentiable in ν. Since $\frac{\partial}{\partial\nu}\delta(\nu,d_c)\big|_{\nu=\frac{1}{2}} = \frac{\partial^2}{\partial\nu^2}\delta(\nu,d_c)\Big|_{\nu=\frac{1}{2}} = 0$ if $d_c\geq 4$, then by using chain rule we obtain (15) and (16). ∎

We are now ready to prove our main result.

Theorem IV.1. *Consider lossy compression of BES(ϵ) using the compound construction. Given any distortion D, $D \leq (1-\epsilon)/2$, $\forall \eta > 0$, let R be the desired rate of compression with $R < R_\epsilon^{sh}(D) + \eta$. Then there exist degrees (d_c, d_v, d_c') (independent of the block-length) and a compound code $\mathbb{C}(\mathbf{G},\mathbf{H}) \in \mathfrak{C}(d_c, d_v, d_c', m, n)$ with rate R which achieves average distortion D.*

Proof: To complete the proof, we need to show that compound codes with a top degree d_c independent of the block-length are sufficient. We will restrict ourselves to even d_c'. Similar to (13) but for the compound construction case, equation (3) is equivalent to

$$\Delta \geq \max_{\nu \in [0,1]} \{K(\nu, d_c)\},$$

where $\Delta = R - R_\epsilon^{\text{sh}}(D)$ and $K(\nu, d_c) = \frac{R}{R_H} B(\nu, d_v, d_c') + F(\delta(\nu, d_c), \epsilon, D)$. We now divide the proof into three steps.

1) $\exists \mu_1 > 0$, independent of d_c, such that $\forall \nu \in [0, \mu_1]$, $K(\nu, d_c) \leq \Delta$.
2) $\exists \mu_2 > 0$, independent of d_c, such that $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$, $K(\nu, d_c) \leq \Delta$.
3) $\exists d_c^* < \infty$ such that $\forall \nu \in [\mu_1, 1/2 - \mu_2]$, $K(\nu, d_c^*) \leq \Delta$.

Proof of 1): By the last property of $B(\nu)$ in Lemma IV.1, $\exists \mu_1 > 0$ such that $B(\nu) \leq 0$ for all $\nu \in [0, \mu_1]$. Since $F(\delta(\nu, d_c), \epsilon, D) \leq 0$ for all ν we have $K(\nu, d_c) \leq 0 < \Delta$. Note that μ_1 is independent of the LDGM part.

Proof of 2): We will use Taylor expansion of $K(\nu, d_c)$ around $\nu = \frac{1}{2}$ up to second order. Note that $K(\nu, d_c) \leq K(\nu, 4)$, $\forall d_c \geq 4$ since $\delta(\nu, d_c)$ is increasing in d_c and $F(\gamma, \epsilon, D)$ is decreasing in γ. Thus it suffices to show that $K(\nu, 4) \leq \Delta$, $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$. Using Lemma IV.1 and IV.2 we can calculate the derivative of $K(\nu, d_c)$ with respect to ν.

$$K\left(\frac{1}{2}, d_c\right) = R - R_\epsilon^{\text{sh}}(D) = \Delta,$$
$$\frac{\partial}{\partial \nu} K(\nu, d_c)\Big|_{\nu=\frac{1}{2}} = 0, \quad \frac{\partial^2}{\partial \nu^2} K(\nu, d_c)\Big|_{\nu=\frac{1}{2}} < 0.$$

By continuity of second derivative of $K(\nu, d_c)$, for some $\mu_2 > 0$ we have for any $\nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$ the second derivative is negative. For any $\nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$, there exists $\tilde{\nu} \in [\nu, 1/2]$ such that

$$K(\nu, 4) = \Delta + \frac{1}{2}\left(\tilde{\nu} - \frac{1}{2}\right)^2 \frac{\partial^2}{\partial \nu^2} K(\nu, d_c)\Big|_{\nu=\frac{1}{2}} \leq \Delta.$$

proof of 3): Using Lemma IV.1 there exists a function $\sigma(\mu_2)$ such that

$$B(\nu) \leq R_H[1 - \sigma(\mu_2)] \text{ for all } \nu \leq \frac{1}{2} - \mu_2. \quad (18)$$

Since $F(\gamma, \epsilon, D)$ is continuous in γ and $\lim_{d_c \to \infty} \delta(\mu_1, d_c) = \frac{1}{2}$, we have

$$\lim_{d_c \to \infty} F(\delta(\mu_1, d_c), \epsilon, D) = -R_\epsilon^{\text{sh}}(D).$$

As $F(\gamma, \epsilon, D)$ is a decreasing function in γ, for any $\mu_3 > 0$, $\exists d_c^* < \infty$ such that

$$F(\delta(\mu_1, d_c^*), \epsilon, D) \leq -R_\epsilon^{\text{sh}}(D) + \mu_3. \quad (19)$$

Combining the results of both equations (18) and (19) we have

$$\begin{aligned} K(\nu, d_c^*) &\leq R[1 - \sigma(\mu_2)] - R_\epsilon^{\text{sh}}(D) + \mu_3, \\ &= \Delta + (\mu_3 - R\sigma(\mu_2)). \end{aligned}$$

We complete the proof by choosing any μ_3 less than $R\sigma(\mu_2)$. ∎

V. Conclusion

We derived upper bounds on the rate distortion performance of the CRP LDGM ensemble and the compound LDGM-LDPC ensemble for the BES. We showed that the compound construction can achieve the Shannon rate-distortion function for lossy compression of the BES.

It is an interesting future research direction to study performance of sparse graph codes for rate-distortion encoding of sources other than the BSS and the BES.

Acknowledgment

The work has been supported in parts by the ARC Grant DP0986089, the Swedish Research Council under VR grant 621-2009-4666, and the European Research Council under the Seventh Framework Programme ERC Grant 228044.

References

[1] E. Martinian and J. Yedidia, "Iterative quantization using codes on graphs," in *Proc. 35th Annu. Allerton Conf. Commun., Control and Computing*, Monticello, IL, 2003.

[2] A. Dimakis, M. J. Wainwright, and K. Ramchandran, "Lower bounds on the rate-distortion function of LDGM codes," in *Proc. Inf. Theory Workshop*, 2007.

[3] S. Kudekar and R. Urbanke, "Lower bounds on the rate-distortion function of individual LDGM codes," in *5th Int. Symp. Turbo Codes and Related Topics*, Lausanne, Switzerland, 2008.

[4] E. Martinian and M. J. Wainwright, "Analysis of LDGM and compound codes for lossy compression and binning," in *Proc. Inf. Theory Workshop*, San Diego, CA, USA, Feb. 2006.

[5] ——, "Low-density codes achieve the rate-distortion bound," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2006.

[6] ——, "Low-density construction can achieve the Wyner-Ziev and Gelfand-Pinsker bounds," in *Proc. Int. Symp. Inf. Theory*, Jul. 2006, pp. 484–488.

[7] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

[8] G. Demay, V. Rathi, and L. K. Rasmussen, "Rate distortion bounds for binary erasure source using sparse graph codes," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2010.

[9] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.