The typical performance of irregular low-density generator-matrix codes for lossy compression

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **42** (2009) 135002 (11pp) doi:10.1088/1751-8113/42/13/135002

The typical performance of irregular low-density generator-matrix codes for lossy compression

Kazushi Mimura

Faculty of Information Sciences, Hiroshima City University, 3-4-1 Ohtsuka-Higashi,
Asaminami-Ku, Hiroshima 731-3194, Japan

E-mail: mimura@hiroshima-cu.ac.jp

Received 29 October 2008, in final form 16 February 2009
Published 12 March 2009
Online at stacks.iop.org/JPhysA/42/135002

Abstract

We evaluate the typical performance of irregular low-density generator-matrix (LDGM) codes, which is defined by sparse matrices with arbitrary irregular bit degree distribution and arbitrary check degree distribution, for lossy compression. We apply the replica method under a one-step replica symmetry breaking (1RSB) ansatz to this problem.

PACS numbers: 89.90.+n, 02.50.−r, 05.50.+q, 75.10.Hk

1. Introduction

The channel coding can be considered as the dual problem of lossy source coding in rate-distortion theory [1, 2]. Matsunaga and Yamamoto showed that it is possible to approach the binary rate-distortion bound using LDPC codes [3]. In recent years, the lossy source coding problem based on low-density generator-matrix (LDGM) codes has been widely investigated.

This scheme can attain high performance very close to the Shannon bound, however a combinatorial optimization problem needs to be solved to obtain optimal source coding. Some practical encoding algorithms are proposed for this scheme, e.g., a belief-propagation-based encoder proposed by Murayama [4] and a survey-propagation-based encoder proposed by Wainwright and Maneva [5].

The performance of this scheme is also explored using various approaches. Murayama and Okada applied replica methods to evaluate performance of LDGM codes defined by regular sparse matrices for lossy compression [6]. Ciliberti *et al* have used the cavity method to evaluate check-regular LDGM performance [7, 8]. On the other hand, Martinian and Wainwright derived rigorous upper bounds on the effective rate-distortion function of LDGM codes for the binary symmetric source [9]. Dimakis *et al* derived lower bounds for check-regular LDGM codes [10, 11].

With respect to irregular LDGM codes analyzed so far, elements of a reproduced message are given by exactly K elements chosen at random from a codeword. This implies that previous

1751-8113/09/135002+11$30.00 © 2009 IOP Publishing Ltd Printed in the UK

analyses treat only the case where a bit degree distribution is Poissonian. An irregular bit and check degree distributions of a generator matrix are not optimized for lossy source coding. The goal of this paper is to evaluate how typical performance of irregular LDGM codes for lossy compression depends on a bit degree distribution and a check degree distribution.

2. Background

Let us first provide the concepts of the rate-distortion theory [1]. Let x be a binary i.i.d. discrete source which takes in a source alphabet $\mathcal{X} = \{0, 1\}$ with $\mathbb{P}[x = 0] = \mathbb{P}[x = 1] = 1/2$, where $\mathbb{P}$ represents the probability of its argument. A source message of M random variables, $\boldsymbol{x} = {}^t(x_1, \ldots, x_M) \in \mathcal{X}^M$, is compressed into a shorter expression, where the operator ^{t} denotes the transpose. The encoder describes the source sequence $\boldsymbol{x} \in \mathcal{X}^M$ by a codeword $\boldsymbol{z} = \mathcal{F}(\boldsymbol{x}) \in \mathcal{X}^N$. The decoder represents $\boldsymbol{x}$ by a reproduced message $\hat{\boldsymbol{x}} = \mathcal{G}(\boldsymbol{z}) \in \mathcal{X}^M$. Note that M represents the length of a source sequence, while $N(<M)$ represents the length of a codeword. The code rate is $R = N/M$. The distortion between single letters is measured by the Hamming distortion defined by

$$d(x, \hat{x}) = \begin{cases} 0, & \text{if} \quad x = \hat{x}, \\ 1, & \text{if} \quad x \neq \hat{x}, \end{cases} \tag{1}$$

and the distortion between M-bit sequences $\boldsymbol{x} \in \mathcal{X}^M$ and $\hat{\boldsymbol{x}} \in \mathcal{X}^M$ is measured by the averaged single-letter distortion as $d(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \frac{1}{M} \sum_{\mu=1}^{M} d(x_\mu, \hat{x}_\mu)$. This results in the probability of error distortion, since $\mathbb{E}[d(x, \hat{x})] = \mathbb{P}[x \neq \hat{x}]$, where $\mathbb{E}$ represents the expectation. The distortion associated with the code is defined as $D = \mathbb{E}[d(\boldsymbol{x}, \hat{\boldsymbol{x}})]$, where the expectation is over the probability distribution on $\mathcal{X}^M \times \mathcal{X}^M$. A rate distortion pair (R, D) is said to be *achievable* if there exists a sequence of rate distortion codes $(\mathcal{F}, \mathcal{G})$ with $\mathbb{E}[d(\boldsymbol{x}, \hat{\boldsymbol{x}})] \leqslant D$ in the limit $M \to \infty$. The rate distortion function $R(D)$ is the infimum of rates R such that (R, D) is in the rate distortion region of the source for a given distortion D. The rate-distortion function of a Bernoulli$(1/2)$ i.i.d. source is given by

$$R(D) = 1 - h_2(D), \tag{2}$$

where $h_2(x) = -x \log_2(x) - (1 - x) \log_2(1 - x)$ is the binary entropy function.

3. Lossy compression scheme

A source message of M random variables, $\boldsymbol{x} \in \mathcal{X}^M$, is compressed into a shorter expression, where the operator ^{t} denotes the transpose. The encoder describes the source sequence $\boldsymbol{x} \in \mathcal{X}^M$ by a codeword $\boldsymbol{z} = \mathcal{F}(\boldsymbol{x}) \in \mathcal{X}^N$. The decoder represents $\boldsymbol{x}$ by a reproduced message $\hat{\boldsymbol{x}} = \mathcal{G}(\boldsymbol{z}) \in \mathcal{X}^M$. The code rate is $R = N/M \leqslant 1$.

Using a given $M \times N$ sparse matrix $\mathcal{A} = (a_{\mu i}) \in \{0, 1\}^{M \times N}$, the decoder is defined as

$$\mathcal{G}(\boldsymbol{z}) = \mathcal{A}\boldsymbol{z} \pmod 2. \tag{3}$$

The encoding is represented by

$$\mathcal{F}(\boldsymbol{x}) = \operatorname*{argmin}_{\hat{\boldsymbol{z}} \in \mathcal{X}^N} d(\boldsymbol{x}, \mathcal{G}(\hat{\boldsymbol{z}})), \tag{4}$$

where d is the distortion measure. In this paper, we use the Hamming distortion. Although the definition means that a computational cost of the encoding is of $O(e^N)$, we can utilize some suboptimal algorithms based on message passing to encode [4, 5].

4. Analysis

To simplify the calculations, we first introduce a simple isomorphism between the additive Boolean group $(\{0,1\},\oplus)$ and the multiplicative Ising group $(\{+1,-1\},\times)$ defined by $J \times \hat{J} = (-1)^{x\oplus\hat{x}}$, where $J,\hat{J} \in \{+1,-1\} = \mathcal{J}$ and $x,\hat{x} \in \{0,1\} = \mathcal{X}$. Hereafter, we use the following Ising (bipolar) representations: the Ising source message $\boldsymbol{J} \in \mathcal{J}^M$, the Ising reproduced message $\hat{\boldsymbol{J}} \in \mathcal{J}^M$ and the Ising codeword $\boldsymbol{\xi} \in \mathcal{J}^N$. The source bit can be described as a random variable with the probability:

$$P_J(J) = \tfrac{1}{2}\delta(J-1) + \tfrac{1}{2}\delta(J+1), \tag{5}$$

where $\delta(x)$ denotes Dirac's delta function. The μth element of the Ising reproduced message $\hat{J}_\mu$ is given by products of the elements of the tentative Ising codeword $\boldsymbol{s} \in \mathcal{J}^N$:

$$\hat{J}_\mu = \prod_{i\in\mathcal{L}(\mu)} s_i, \tag{6}$$

where $\mathcal{L}(\mu) = \{i|a_{\mu i} = 1, \mathcal{A} = (a_{\mu i})\}$.

The matrix $\mathcal{A}$ has K_μ nonzero elements in the μth row and C_i nonzero elements in the ith column. We consider the source length and the codeword length to be infinite, while code rate R is kept finite. The parameters $K_1,\ldots,K_M$ and $C_1,\ldots,C_N$ are usually of $O(N^0)$, therefore the matrix $\mathcal{A}$ becomes very sparse. In densely constructed cases, we also assume that these parameters are not of $O(N^0)$ but $K, C_1,\ldots,C_N \ll N$ holds. Counting the number of nonzero elements in the matrices leads to $K_1 + \cdots + K_M = C_1 + \cdots + C_N$. The code rate is therefore $R = \tilde{K}/\tilde{C}$, where $\tilde{K} = \frac{1}{N}\sum_{\mu=1}^M K_\mu$ and $\tilde{C} = \frac{1}{N}\sum_{i=1}^N C_i$. Code constructions are described by the connectivity parameter $\mathcal{D}^\mu_{i_1,\ldots,i_{K_\mu}} \in \{0,1\}$ which specifies a set of indices $i_1,\ldots,i_{K_\mu}$ corresponding to nonzero elements in the μth row of the sparse matrix $\mathcal{A}$. The connectivity parameter is defined by

$$\mathcal{D}^\mu_{i_1,\ldots,i_{K_\mu}} = \begin{cases} 1, & \text{if} \quad \{i_1,\ldots,i_{K_\mu}\} = \mathcal{L}(\mu) \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

An ensemble of codes is generated as follows. (i) Sets of $\{K_1,\ldots,K_M\}$ and $\{C_1,\ldots,C_N\}$ are sampled independently from an identical distributions $P_K(K)$ and $P_C(C)$, respectively. (ii) The connectivity parameters $\mathcal{D}^\mu_{i_1,\ldots,i_{K_\mu}}$ are generated such that

$$\sum_{\mu=1}^M \sum_{\langle i_1=i,i_2,\ldots,i_{K_\mu}\rangle} \mathcal{D}^\mu_{i,i_2,\ldots,i_{K_\mu}} = C_i, \tag{8}$$

where $\sum_{\langle i_1=i,i_2,\ldots,i_{K_\mu}\rangle}$ denote the summation over $\{(i_2,\ldots,i_{K_\mu}) \in \{1,\ldots,N\}^{K_\mu-1}|i_2 < \cdots < i_{K_\mu}, i_2 \neq i,\ldots,i_{K_\mu} \neq i\}$.

To analyze typical performance of rate-compatible LDGM codes for lossy compression, we apply a analytical method similar to [6, 12–15]. The Hamming distortion $d(\boldsymbol{J},\hat{\boldsymbol{J}})$ becomes $d(\boldsymbol{J},\hat{\boldsymbol{J}}) = \frac{1}{2} - \frac{1}{2M}\sum_{\mu=1}^M J_\mu\big\{\prod_{i\in\mathcal{L}(\mu)} s_i\big\}$, since $\boldsymbol{J},\hat{\boldsymbol{J}} \in \mathcal{J}^M$. Using the connectivity parameter $\mathcal{D}^\mu_{i_1,\ldots,i_{K_\mu}}$, we can rewrite this Hamming distortion in the form:

$$d(\boldsymbol{J},\hat{\boldsymbol{J}}) = \frac{1}{2} - \frac{1}{2M}\sum_{\mu=1}^M \sum_{\langle i_1,\ldots,i_{K_\mu}\rangle} \mathcal{D}^\mu_{i_1,\ldots,i_{K_\mu}} J_\mu s_{i_1}\cdots s_{i_{K_\mu}}, \tag{9}$$

where $\sum_{\langle i_1,\ldots,i_{K_\mu}\rangle}$ denote the summation over $\{(i_1,\ldots,i_{K_\mu}) \in \{1,\ldots,N\}^{K_\mu}|i_1 < \cdots < i_{K_\mu}\}$. We here define the Hamiltonian

$$\mathcal{H}(\boldsymbol{s},\boldsymbol{J}) = Md(\boldsymbol{J},\hat{\boldsymbol{J}}(\boldsymbol{s})), \tag{10}$$

to explore typical performance. The free energy is calculated from the partition function $Z(\beta) = \sum_{\boldsymbol{s}\in\mathcal{J}^N} \exp[-\beta\mathcal{H}(\boldsymbol{s},\boldsymbol{J})]$. From the free energy, we can obtain a distortion between an original message and a reproduction message D for a fixed code rate R. We follow the calculation of [6, 12, 16–19].

4.1. Replica symmetric solution

We first assume the replica symmetry (RS). Using the replica symmetric partition function $Z_{\mathrm{RS}}(\beta)$, we find the replica symmetric free energy as

$$f_{\mathrm{RS}}(\beta) = -\frac{1}{\beta n M}\ln\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z_{\mathrm{RS}}(\beta)^n] \tag{11}$$

$$\begin{aligned} &= \frac{1}{2} - \frac{1}{\beta}\,\underset{\pi,\hat{\pi}}{\mathrm{extr}}\left[\ln\cosh\frac{\beta}{2} - \bar{K}\int_{-1}^{1}\mathrm{d}x\,\pi(x)\int_{-1}^{1}\mathrm{d}\hat{x}\,\hat{\pi}(\hat{x})\ln(1+x\hat{x})\right. \\ &\quad + \sum_K P_K(K)\left(\prod_{k=1}^{K}\int_{-1}^{1}\mathrm{d}x_k\,\pi(x_k)\right)\mathbb{E}_J\left[\ln\left(1+\left(\tanh\frac{\beta J}{2}\right)\prod_{k=1}^{K}x_k\right)\right] \\ &\quad \left. + \frac{\bar{K}}{\bar{C}}\sum_C P_C(C)\left(\prod_{c=1}^{C}\int_{-1}^{1}\mathrm{d}\hat{x}_c\,\hat{\pi}(\hat{x}_c)\right)\ln\left(\sum_{\sigma=\pm1}\prod_{c=1}^{C}[1+\sigma\hat{x}_c]\right)\right], \end{aligned} \tag{12}$$

where the parameters are determined by the saddle-point equations obtained by calculating functional variations:

$$\pi(x) = \sum_C \frac{C}{\bar{C}}P_C(C)\left(\prod_{c=1}^{C-1}\int_{-1}^{1}\mathrm{d}\hat{x}_c\,\hat{\pi}(\hat{x}_c)\right)\delta\left(x - \tanh\left[\sum_{c=1}^{C-1}\tanh^{-1}\hat{x}_c\right]\right), \tag{13}$$

$$\hat{\pi}(\hat{x}) = \sum_K \frac{K}{\bar{K}}P_K(K)\left(\prod_{k=1}^{K-1}\int_{-1}^{1}\mathrm{d}x_k\,\pi(x_k)\right)\mathbb{E}_J\left[\delta\left(\hat{x} - \left(\tanh\frac{\beta J}{2}\right)\prod_{k=1}^{K-1}x_k\right)\right], \tag{14}$$

with $\bar{K} = \sum_K KP_K(K)$ and $\bar{C} = \sum_C CP_C(C)$ (see the outline of the derivation in appendix A). We can obtain the distortion, which is reproduction errors, $u_{\mathrm{RS}}(\beta) = \partial[\beta f_{\mathrm{RS}}(\beta)]/\partial\beta$ and the replica symmetric entropy $s_{\mathrm{RS}}(\beta) = \beta[u_{\mathrm{RS}}(\beta) - f_{\mathrm{RS}}(\beta)]$.

For arbitrary $P_K(K)$, $P_C(C)$ and β, $\pi(x) = \delta(x)$ and $\hat{\pi}(\hat{x}) = \delta(\hat{x})$ are always solutions of the saddle-point equations (13) and (14). These correspond to the paramagnetic solution. The paramagnetic free-energy, internal energy and entropy are given by $f_{\mathrm{PARA}}(\beta) = \frac{1}{2} - \frac{1}{\beta}\ln\cosh\frac{\beta}{2} - \frac{R}{\beta}\ln 2$, $u_{\mathrm{PARA}}(\beta) = \frac{1}{2} - \frac{1}{2}\tanh\frac{\beta}{2}$ and $s_{\mathrm{PARA}}(\beta) = \ln\cosh\frac{\beta}{2} - \frac{\beta}{2}\tanh\frac{\beta}{2} + R\ln 2$, respectively. However, this RS solution takes negative entropy while $R\ln 2 < \frac{\beta}{2}\tanh\frac{\beta}{2} - \ln\cosh\frac{\beta}{2}$. Especially, when the inverse temperature $\beta \to \infty$, the RS entropy becomes $s_{\mathrm{RS}}(\beta) = (R-1)\ln 2$. This means we have to look for the true solution beyond the RS ansatz for $R \leqslant 1$.

4.2. One-step replica symmetry breaking solution

The replica symmetric breaking (RSB) theory for sparse systems is still under development [20–25]. Therefore, as a first approach we introduce the frozen RSB to produce a solution with non-negative entropy [6, 12, 13]. The frozen RSB method is a limited version of full one-step RSB (1RSB) and includes the RS method as a special case. In this 1RSB scheme, n replicas are divided into n/m groups which contain m replicas each. The symmetry breaking

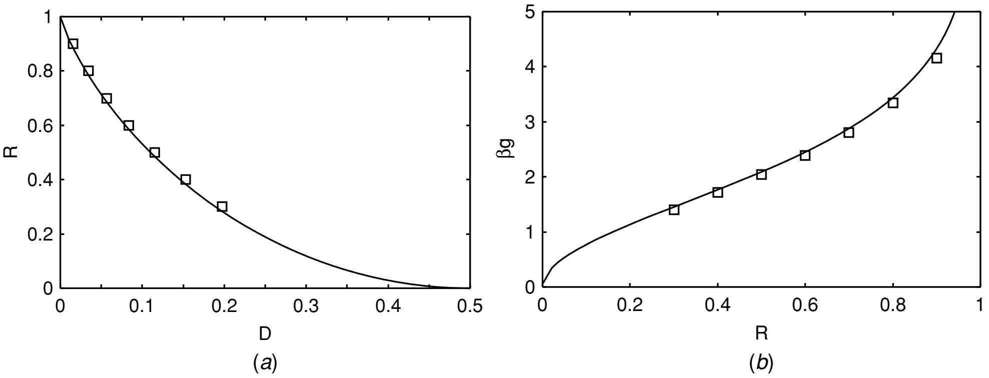


Figure 1. Example of numerical solutions for finite connectivity systems with $P_K(K) = \delta_{K,2}$ and $P_C(C) = \tilde{\mathcal{P}}_C(C)$. (*a*) Rate distortion performance for $r = 0.3, 0.4, \ldots, 0.9$ (squares). The solid line denotes the rate distortion performance in large $\bar{K}$ and $\bar{C}$ limits, which coincides with the Shannon bound. (*b*) Inverse temperature β_g for $r = 0.3, 0.4, \ldots, 0.9$ (squares). The solid line denotes the inverse temperature β_c, which is defined by $s_{\mathrm{PARA}}(\beta_c) = 0$.

parameter m was found to be $m = \beta_g/\beta$, where β_g is an inverse temperature at which the replica symmetric entropy vanishes, i.e., $s_{\mathrm{RS}}(\beta_g) = 0$ (see appendix B). This 1RSB scheme gives the exact solution for the random energy model (REM) [6, 26]. For $\beta > \beta_g$, the 1RSB free energy becomes $f_{1\mathrm{RSB}}(\beta) = f_{\mathrm{RS}}(\beta_g)$. It can be regarded as a constant with respect to the inverse temperature β. We assume that the 1RSB scheme is enough good to approximate the solution even if $\bar{K}$ and $\bar{C}$ are finite. Under this assumption, the distortion D is simply given by $D = \lim_{\beta\to\infty} u_{1\mathrm{RSB}}(\beta) = u_{\mathrm{RS}}(\beta_g)$.

5. Results and discussion

5.1. Basic results

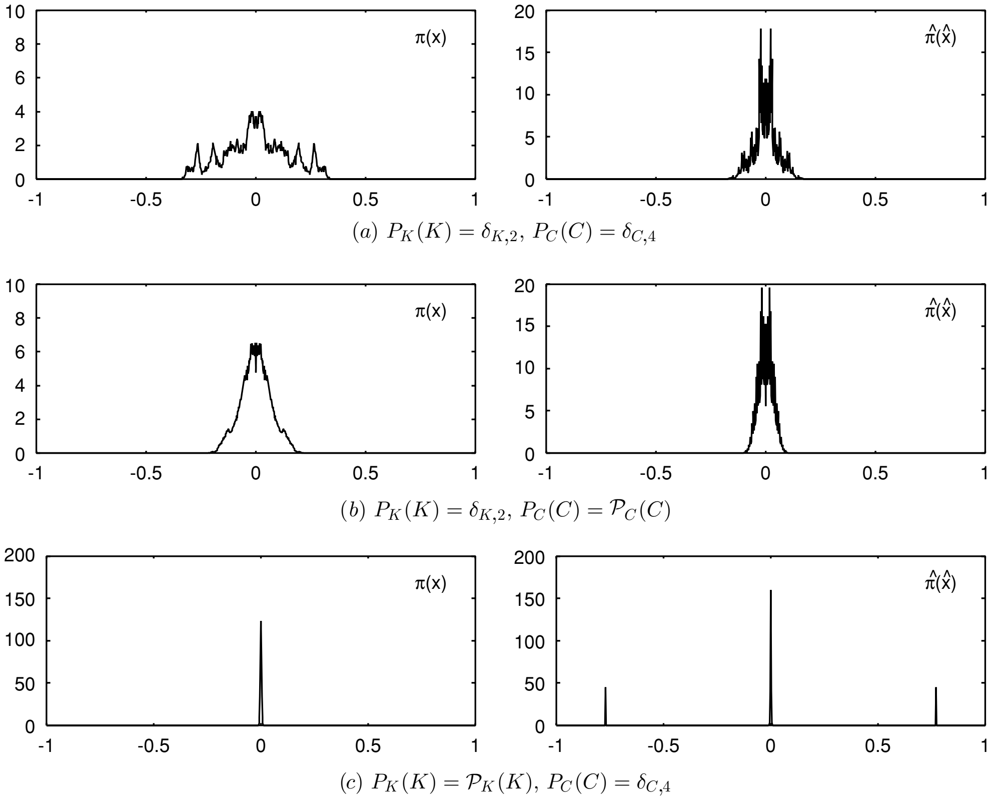
In large $\bar{K}$ and $\bar{C}$ limit, there are no other solutions except $\pi(x) = \delta(x)$ and $\hat{\pi}(\hat{x}) = \delta(\hat{x})$ for the saddle-point equations. We then found the relationship

$$R = 1 - h_2(D), \tag{15}$$

from $s_{\mathrm{RS}}(\beta_g) = \ln\cosh\frac{\beta_g}{2} - \frac{\beta_g}{2}\tanh\frac{\beta_g}{2} + R\ln 2 = 0$ and $D = u_{\mathrm{RS}}(\beta_g) = \frac{1}{2} - \frac{1}{2}\tanh\frac{\beta_g}{2}$.

In finite $\bar{K}$ and $\bar{C}$ cases, the solutions $\pi(x) = \delta(x)$ and $\hat{\pi}(\hat{x}) = \delta(\hat{x})$ also exist, but these are no longer stable [6]. We have to solve equations (13) and (14) numerically. We choose the proper value of the inverse temperature β_g which gives $s_{\mathrm{RS}}(\beta_g) = 0$ by using the numerical results of the saddle-point equations. Since the distortion D can be evaluated from $D = u_{\mathrm{RS}}(\beta_g)$, we can also obtain the relation between the code rate $R = \bar{K}/\bar{C}$ and the distortion D in the finite connectivity systems.

As one of the simplest examples to treat the arbitrary code rate, we here introduce degree distributions $P_K(K) = \delta_{K,2}$ and $P_C(C) = \frac{7r-2}{5r}\delta_{C,2} + \frac{2(1-r)}{5r}\delta_{C,7}(\equiv \tilde{\mathcal{P}}_C(C))$, which are valid for $\frac{2}{7} \leqslant r \leqslant 1$. Here, $\delta_{m,n}$ denotes Kronecker's delta taking 1 if $m = n$ and 0 otherwise. In this case, we can adjust the code rate $R(= r)$ via the parameter r. We apply the Monte Carlo integration to solve the saddle-point equations. Figure 1(*a*) shows the rate-distortion performance of this system. Figure 1(*b*) shows the inverse temperature β_g, which is a root of the replica symmetric entropy $s_{\mathrm{RS}}(\beta_g) = 0$.



(a) $P_K(K)=\delta_{K,2}$, $P_C(C)=\delta_{C,4}$

(b) $P_K(K)=\delta_{K,2}$, $P_C(C)=\mathcal{P}_C(C)$

(c) $P_K(K)=\mathcal{P}_K(K)$, $P_C(C)=\delta_{C,4}$

Figure 2. Snapshots of the order functions $\pi(x)$ and $\hat{\pi}(\hat{x})$. (*a*) A check-regular and bit-irregular case, (*b*) a regular case, and (*c*) a check-irregular and bit-regular case.

5.2. *Some typical irregular constructions*

We next apply some degree distributions as typical examples. It should be noted that these distributions discussed here are not optimized but heuristically chosen. All three examples have the code rate $R = 1/2$.

Firstly, we consider a regular case characterized as $P_K(K) = \delta_{K,2}$ and $P_C(C) = \delta_{C,4}$. Figure 2(*a*) shows stable solutions $\pi(x)$ and $\hat{\pi}(\hat{x})$ of the saddle-point equations for this case. It can be confirmed that the functions $\pi(x)$ and $\hat{\pi}(\hat{x})$ are broad in shape. In this case, the distortion becomes $D = 0.116$. The Shannon bound is $D_{SB} = 0.1100$.

Secondly, we treat a check-regular and bit-irregular case whose degree distributions are defined as $P_K(K) = \delta_{K,2}$ and $P_C(C) = \mathcal{P}_C(C)$, where

$$\mathcal{P}_C(C) = 0.04\delta_{C,1} + 0.15\delta_{C,2} + 0.22\delta_{C,3} + 0.22\delta_{C,4} + 0.18\delta_{C,5} + 0.11\delta_{C,6} + 0.08\delta_{C,7}. \qquad (16)$$

This $\mathcal{P}_C(C)$ is a rough approximation of the Poissonian distribution $\mathrm{e}^{-\lambda}\lambda^{C-1}/(C-1)!$ with $\lambda = 3$. The distortion is $D = 0.115$ for this case. It represents an ensemble which have at least one non-zero element in each row. In the check-regular case, when we choose the non-zero elements randomly, there exist some columns whose elements are all zero. In such a situation, the code rate essentially becomes small.

Lastly, for a check-irregular and bit-regular case, we apply $P_K(K) = \mathcal{P}_K(K)$ and $P_C(C) = \delta_{C,4}$, where

$$\mathcal{P}_K(K) = 0.36\delta_{K,1} + 0.36\delta_{K,2} + 0.20\delta_{K,3} + 0.08\delta_{K,4}. \tag{17}$$

This $\mathcal{P}_K(K)$ is a rough approximation of the Poissonian distribution $\mathrm{e}^{-\lambda}\lambda^{K-1}/(K-1)!$ with $\lambda = 1$. The reason why we consider this distribution is same to $\mathcal{P}_C(C)$. In this case, the distortion becomes $D = 0.115$. These three kinds of distributions give almost same distortion.

Figures 2(*b*) and (*c*) show stable solutions for these irregular cases. It can be confirmed that the distribution $\pi(x)$ and $\hat{\pi}(\hat{x})$ become a little bit narrow than the regular case. It is considered that the distortion can become small due to this.

6. Conclusions

We evaluate typical performance of LDGM codes with irregular bit and check degree distributions by applying the replica method under 1RSB ansatz. Our result shows that we can use an arbitrary code rate. It might be possible to investigate suboptimal irregular degree distributions by using the hill-climbing approach similar to the case of the density evolution [27, 28].

In the practical point of view, it must be important to evaluate some polynomial time encoding algorithms with arbitrary degree distributions. It should be noted that the analysis addressed here is based on an exact calculation of the encoder's definition. Therefore, it can be considered that the distortion obtained by this analysis provides the theoretical limit for given check and bit degree distributions.

Recently, the cavity method was introduced to evaluate the typical performance [7]. Since the cavity method does not need the replica trick, it might be able to avoid some assumptions. Applying the cavity method to this problem is also important and is a part of our future work.

Acknowledgments

The author would like to thank Kazutaka Nakamura, Tatsuto Murayama and Yoshiyuki Kabashima for their helpful comments. We also thank Toru Yano for giving us valuable preprints. This work was partially supported by a Grant-in-Aid for Encouragement of Young Scientists (B) No. 18700230 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Appendix A. Derivation of replica symmetric free energy

We assume that the event $\mathcal{D}^{\mu}_{i_1,\ldots,i_{K_\mu}} = 1$ occurs independently for every row μ. We then have

$$\mathbb{P}\big(\mathcal{D}^{\mu}_{i_1,\ldots,i_{K_\mu}} = 1\big) = p_\mu, \tag{A.1}$$

$$\mathbb{P}\big(\mathcal{D}^{\mu}_{i_1,\ldots,i_{K_\mu}} = 0\big) = 1 - p_\mu, \tag{A.2}$$

where $\mathbb{P}(\cdots)$ denotes the probability of the event $(\cdots)$ and $p_\mu = \binom{N}{K_\mu}^{-1} \simeq K_\mu!/N^{K_\mu}$. Introducing the constraint concerning the column (8) by using Dirac's delta function, the ensemble average over the codes is represented as

$$\begin{aligned}\mathbb{E}_{\mathcal{A}}[(\cdots)] &= \left(\sum_{\{K_\mu\}}\prod_{\mu=1}^{M} P_K(K_\mu)\right)\left(\sum_{\{C_i\}}\prod_{i=1}^{N} P_C(C_i)\right) \\ &\quad\times \frac{1}{\mathcal{N}_{\mathcal{D}}}\mathbb{E}_{\mathcal{D}}\left[\left\{\prod_{i=1}^{N}\delta\left(\sum_{\mu=1}^{M}\sum_{\langle i_1=i,i_2,\ldots,i_{K_\mu}\rangle}\mathcal{D}^{\mu}_{i_1=i,i_2,\ldots,i_{K_\mu}};C_i\right)\right\}(\cdots)\right], \\ &= \left(\sum_{\{K_\mu\}}\prod_{\mu=1}^{M} P_K(K_\mu)\right)\left(\sum_{\{C_i\}}\prod_{i=1}^{N} P_C(C_i)\right) \\ &\quad\times \frac{1}{\mathcal{N}_{\mathcal{D}}}\mathbb{E}_{\mathcal{D}}\left[\left\{\prod_{i=1}^{N}\oint\frac{\mathrm{d}Z_i}{2\pi\mathrm{i}}\frac{1}{Z_i^{C_i+1}}\prod_{\mu=1}^{N}\prod_{\langle i_1=i,i_2,\ldots,i_{K_\mu}\rangle} Z_i^{\mathcal{D}^{\mu}_{i_1=i,i_2,\ldots,i_{K_\mu}}}\right\}(\cdots)\right],\end{aligned} \qquad \text{(A.3)}$$

where $\mathbb{E}_{\mathcal{D}}$ denotes the average over the connectivity parameter. Observing that $\sum_{\langle i_1,\ldots,i_{K_\mu}\rangle}(\cdots) = \frac{1}{K_\mu!}\left(\sum_i(\cdots)\right)^{K_\mu}$ for large N, the normalization constant $\mathcal{N}_{\mathcal{D}}$ is given by

$$\mathcal{N}_{\mathcal{D}} = \mathbb{E}_{\mathcal{D}}\left[\prod_{i=1}^{N}\delta\left(\sum_{\mu=1}^{M}\sum_{\langle i_1=i,i_2,\ldots,i_{K_\mu}\rangle}\mathcal{D}^{\mu}_{i_1=i,i_2,\ldots,i_{K_\mu}};C_i\right)\right] = \frac{(N\bar{C})!}{N^{N\bar{C}}\prod_{i=1}^{N}C_i!}. \qquad \text{(A.4)}$$

To evaluate the free energy, we calculate the replicated partition function

$$\begin{aligned}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z(\beta)^n] &= \mathrm{e}^{-\frac{nM\beta}{2}}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}\left[\sum_{\boldsymbol{s}^1,\ldots,\boldsymbol{s}^n}\exp\left[\frac{\beta}{2}\sum_{\mu=1}^{M}\sum_{\langle i_1,\ldots,i_{K_\mu}\rangle}\mathcal{D}^{\mu}_{i_1,\ldots,i_{K_\mu}}J_\mu\sum_{\alpha=1}^{n}s_{i_1}^{\alpha}\cdots s_{i_{K_\mu}}^{\alpha}\right\}\right]\right] \\ &= \mathrm{e}^{-\frac{nM\beta}{2}}\left(\sum_{\{K_\mu\}}\prod_{\mu=1}^{M} P_K(K_\mu)\right)\left(\sum_{\{C_i\}}\prod_{i=1}^{N} P_C(C_i)\right) \\ &\quad\times\frac{1}{\mathcal{N}_{\mathcal{D}}}\left(\prod_{i=1}^{N}\oint\frac{\mathrm{d}Z_i}{2\pi\mathrm{i}}\frac{1}{Z_i^{C_i+1}}\right)\prod_{\mu=1}^{M}\left(p_\mu\sum_{\langle i_1,\ldots,i_{K_\mu}\rangle}\left(\cosh\frac{\beta}{2}\right)^n Z_{i_1}\cdots Z_{i_{K_\mu}}\right. \\ &\quad + p_\mu\sum_{m=1}^{n}\sum_{\langle\alpha_1,\ldots,\alpha_m\rangle}\left(\cosh\frac{\beta}{2}\right)^n\mathbb{E}_J\left[\left(\tanh\frac{\beta J}{2}\right)^m\right] \\ &\quad\left.\times\sum_{\langle i_1,\ldots,i_{K_\mu}\rangle}\left(Z_{i_1}s_{i_1}^{\alpha_1}\cdots s_{i_1}^{\alpha_m}\right)\cdots\left(Z_{i_{K_\mu}}s_{i_{K_\mu}}^{\alpha_1}\cdots s_{i_{K_\mu}}^{\alpha_m}\right)\right).\end{aligned} \qquad \text{(A.5)}$$

We next introduce order parameters $q_{\alpha_1,\ldots,\alpha_m}$ and q_0, defined by

$$q_{\alpha_1,\ldots,\alpha_m} = \frac{1}{N}\sum_{i=1}^{N}Z_i s_i^{\alpha_1}\cdots s_i^{\alpha_m}, \qquad \text{(A.6)}$$

$$q_0 = \frac{1}{N}\sum_{i=1}^{N}Z_i. \qquad \text{(A.7)}$$

Using the Fourier expression of the Dirac delta function, we find

$$\begin{aligned}
\mathbb{E}_{\mathcal{A},J}[Z(\beta)^n] = \mathrm{e}^{-\frac{nM\beta}{2}}&\left(\int\frac{\mathrm{d}q_0\,\mathrm{d}\hat{q}_0}{2\pi}\right)\left(\prod_{\langle\alpha_1\rangle}\int\frac{\mathrm{d}q_{\alpha_1}\,\mathrm{d}\hat{q}_{\alpha_1}}{2\pi}\right)\cdots\left(\prod_{\langle\alpha_1,\ldots,\alpha_n\rangle}\int\frac{\mathrm{d}q_{\alpha_1,\ldots,\alpha_n}\mathrm{d}\hat{q}_{\alpha_1,\ldots,\alpha_n}}{2\pi}\right)\\
&\times\left(\sum_{\{K_\mu\}}\prod_{\mu=1}^{M}P_K(K_\mu)\right)\left(\sum_{\{C_i\}}\prod_{i=1}^{N}P_C(C_i)\right)\frac{1}{\mathcal{N}_{\mathcal{D}}}\left(\prod_{i=1}^{N}\oint\frac{\mathrm{d}Z_i}{2\pi\mathrm{i}}\frac{1}{Z_i^{C_i+1}}\right)\\
&\times\exp\left[-N\left\{q_0\hat{q}_0+\cdots+\sum_{\langle\alpha_1,\ldots,\alpha_n\rangle}q_{\alpha_1,\ldots,\alpha_n}\hat{q}_{\alpha_1,\ldots,\alpha_n}\right\}\right.\\
&\left.+\hat{q}_0\sum_{i=1}^{N}Z_i+\cdots+\sum_{\langle\alpha_1,\ldots,\alpha_n\rangle}\hat{q}_{\alpha_1,\ldots,\alpha_n}\sum_{i=1}^{N}Z_i s_i^{\alpha_1}\cdots s_i^{\alpha_n}\right]\\
&\times\prod_{\mu=1}^{M}\left(T_0 q_0^{K_\mu}+\sum_{m=1}^{n}\sum_{\langle\alpha_1,\ldots,\alpha_m\rangle}T_m(q_{\alpha_1,\ldots,\alpha_m})^{K_\mu}\right),
\end{aligned}\tag{A.8}$$

with $T_m = \left(\cosh\frac{\beta}{2}\right)^n\mathbb{E}_J\left[\left(\tanh\frac{\beta J}{2}\right)^m\right]$. To proceed further, we introduce the replica-symmetric assumption:

$$q_{\alpha_1,\ldots,\alpha_m} = q\int_{-1}^{1}\mathrm{d}x\,\pi(x)x^m,\tag{A.9}$$

$$\hat{q}_{\alpha_1,\ldots,\alpha_m} = \hat{q}\int_{-1}^{1}\mathrm{d}\hat{x}\,\hat{\pi}(\hat{x})\hat{x}^m,\tag{A.10}$$

where $\pi(x)\geqslant 0$, $\hat{\pi}(\hat{x})\geqslant 0$ and $\int_{-1}^{1}\mathrm{d}x\,\pi(x) = \int_{-1}^{1}\mathrm{d}\hat{x}\,\hat{\pi}(\hat{x}) = 1$. This assumption means that the order parameters depend only on the number of indices. We write the replica symmetric partition function as $Z_{\mathrm{RS}}(\beta)$. Using the integral form of Dirac's delta function, we obtain

$$\begin{aligned}
\mathbb{E}_{\mathcal{A},J}[Z_{\mathrm{RS}}(\beta)^n] = \operatorname*{extr}_{\pi,\hat{\pi},q,\hat{q}}&\frac{\mathrm{e}^{-\frac{nM\beta}{2}}}{\mathcal{N}_{\mathcal{D}}}\left(\sum_{\{K_\mu\}}\prod_{\mu=1}^{M}P_K(K_\mu)\right)\left(\sum_{\{C_i\}}\prod_{i=1}^{N}P_C(C_i)\right)\\
\times&\left(\prod_{i=1}^{N}\left\{\frac{\hat{q}^{C_i}}{C_i!}\left(\prod_{c=1}^{C_i}\int_{-1}^{1}\mathrm{d}\hat{x}_c\,\hat{\pi}(\hat{x}_c)\right)\left(\sum_{\sigma=\pm1}\prod_{c=1}^{C_i}(1+\sigma\hat{x}_c)\right)^n\right\}\right)\\
\times&\exp\left[-Nq\hat{q}\int_{-1}^{1}\mathrm{d}x\,\pi(x)\int_{-1}^{1}\mathrm{d}\hat{x}\,\hat{\pi}(\hat{x})(1+x\hat{x})^n\right]\\
\times&\prod_{\mu=1}^{M}\left(q^{K_\mu}\left(\cosh\frac{\beta}{2}\right)^n\left(\prod_{k=1}^{K_\mu}\int_{-1}^{1}\mathrm{d}x_k\,\pi(x_k)\right)\mathbb{E}_J\left[\left(1+\left(\tanh\frac{\beta J}{2}\right)\prod_{k=1}^{K_\mu}x_k\right)^n\right]\right].
\end{aligned}\tag{A.11}$$

Finally, substituting this into (11) and taking the limit $n\to 0$, we arrive at (12). The saddle-point equations (13) and (14) are simply obtained as the extremization condition of (12).

Appendix B. One-step replica symmetry breaking solution

We follow the calculation of the [13]. We assume that the space of configuration is divided in n/m groups with m identical configurations in each

$$\frac{1}{N}\boldsymbol{s}^\alpha\cdot\boldsymbol{s}^\beta = \begin{cases}1, & \text{if}\quad \alpha \text{ and } \beta \text{ are in the same group}\\ q, & \text{otherwise.}\end{cases}\tag{B.1}$$

Using this ergodicity breaking assumption, the 1RSB replicated partition function becomes

$$\begin{aligned}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z_{1\mathrm{RSB}}(\beta)^n]|_{(B.1)} &= \mathbb{E}_{\mathcal{A},\boldsymbol{J}}\left[\left(\sum_{\boldsymbol{s}} \mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{s},\boldsymbol{J})}\right)^n\right]\Bigg|_{(B.1)} \\ &= \mathbb{E}_{\mathcal{A},\boldsymbol{J}}\left[\left(\sum_{\boldsymbol{s}} \mathrm{e}^{-\beta m\mathcal{H}(\boldsymbol{s},\boldsymbol{J})}\right)^{n/m}\right] \\ &= \mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z_{\mathrm{RS}}(\beta m)^{n/m}]. \end{aligned} \tag{B.2}$$

Then we obtain the 1RSB free energy as

$$\begin{aligned} f_{1\mathrm{RSB}}(\beta) &= -\frac{1}{\beta}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[\ln Z_{1\mathrm{RSB}}(\beta)] \\ &= -\frac{1}{\beta}\left(\frac{\partial}{\partial n}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z_{1\mathrm{RSB}}(\beta)^n]\right)\Bigg|_{n=0} \\ &= -\frac{1}{\beta m}\left(\frac{\partial}{\partial (n/m)}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[Z_{\mathrm{RS}}(\beta m)^{n/m}]\right)\Bigg|_{n/m=0} \\ &= -\frac{1}{\beta m}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[\ln Z_{\mathrm{RS}}(\beta m)] \\ &= f_{\mathrm{RS}}(\beta m). \end{aligned} \tag{B.3}$$

The symmetry breaking parameter m should be determined to extremize the 1RSB free energy as

$$\frac{\partial}{\partial m} f_{1\mathrm{RSB}}(\beta) = 0. \tag{B.4}$$

The left-hand side of this condition becomes

$$\begin{aligned} \frac{\partial}{\partial m} f_{1\mathrm{RSB}}(\beta) &= -\frac{\partial}{\partial m}\frac{1}{\beta m}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[\ln Z_{\mathrm{RS}}(\beta m)] \\ &= -\frac{1}{m}\left(\frac{\partial[\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[\ln Z_{\mathrm{RS}}(\beta m)]]}{\partial(\beta m)} - \frac{1}{\beta m}\mathbb{E}_{\mathcal{A},\boldsymbol{J}}[\ln Z_{\mathrm{RS}}(\beta m)]\right) \\ &= \frac{1}{m}\left(\frac{\partial[(\beta m) f_{\mathrm{RS}}(\beta m)]}{\partial(\beta m)} - f_{\mathrm{RS}}(\beta m)\right) \\ &= \frac{1}{\beta m^2} s_{\mathrm{RS}}(\beta m). \end{aligned} \tag{B.5}$$

Namely, the condition (B.4) is equivalent to $s_{\mathrm{RS}}(\beta m) = 0$. Therefore, the symmetry breaking parameter is given by $m = \beta_g/\beta$ with $s_{\mathrm{RS}}(\beta_g) = 0$.

References

[1] Cover T M and Thomas J A 2006 *Elements of Information Theory* 2nd edn (New York: Wiley)

[2] Csiszár I and Körner J 1981 *Information Theory: Coding Theorems for Discrete Memoryless Systems* (New York: Academic)

[3] Matsunaga Y and Yamamoto H 2003 *IEEE Trans. Inf. Theory* **49** 2225

[4] Murayama T 2004 *Phys. Rev.* E **69** 035105

[5] Wainwright M J and Maneva E 2005 *Proc. Int. Symp. Inf. Theory* 1493

[6] Murayama T and Okada M 2003 *J. Phys. A: Math. Gen.* **36** 11123

[7] Ciliberti S and Mézard M 2006 *J. Stat. Mech.* **3** 58
[8] Ciliberti S, Mézard M and Zecchina R 2006 *Complex Syst. Methods* **3** 58
[9] Martinian E and Wainwright M J 2006 *Workshop on Information, Theory and its Applications*
[10] Dimakis A G, Wainwright M J and Ramchandran K 2005 *Information Theory Workshop* 650
[11] Wainwright M J 2007 *IEEE Signal Process. Mag.* **24** 47
[12] Kabashima Y and Saad D 1999 *Europhys. Lett.* **45** 97
[13] Vicente R, Saad D and Kabashima Y 2000 *J. Phys. A: Math. Gen.* **33** 6527
[14] Nakamura K 2003 *Doctoral Thesis* Tokyo Institute of Technology
[15] Mimura K and Okada M 2006 *Phys. Rev.* E **74** 026108
[16] Krauth W and Mézard M 1989 *J. Phys. France* **50** 3057
[17] Tanaka T and Saad D 2003 *Technical Report* (unpublished)
[18] Yano T, Tanaka T and Saad D 2007 *Proc. SITA2007* 41.4
[19] Yano T 2008 *Doctoral Thesis* Keio University
[20] Wong K Y and Sherrington D 1988 *J. Phys. A: Math. Gen.* **21** L459
[21] de Dominicis C and Goldschmidt Y Y 1989 *J. Phys. A: Math. Gen.* **22** L775
[22] Lai P and Goldschmidt Y Y 1990 *J. Phys. A: Math. Gen.* **23** 3329
[23] Goldschmidt Y Y and de Deminicis C 1990 *Phys. Rev.* B **41** 2184
[24] Monasson R 1998 *J. Phys. A: Math. Gen.* **31** 513
[25] Parisi G and Tria F 2002 *Eur. Phys. J.* B **30** 4, 1434
[26] Derrida B 1981 *Phys. Rev.* B **24** 2613
[27] Richardson T J and Urbanke R L 2001 *IEEE Trans. Inf. Theory* **47** 2, 599
[28] Richardson T J, Shokrollahi M A and Urbanke R L 2001 *IEEE Trans. Inf. Theory* **47** 2, 619