Typical performance of regular low-density parity-check codes over general symmetric channels

2003 J. Phys. A: Math. Gen. 36 11143

(http://iopscience.iop.org/0305-4470/36/43/033)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 11143–11157 PII: S0305-4470(03)61622-7

Typical performance of regular low-density parity-check codes over general symmetric channels

Toshiyuki Tanaka[1,2] **and David Saad**[2]

[1] Department of Electronics and Information Engineering, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji-shi, Tokyo 192-0397, Japan
[2] Neural Computing Research Group, Aston University, Aston Triangle, Birmingham B4 7ET, UK

Received 27 March 2003, in final form 7 July 2003
Published 15 October 2003
Online at stacks.iop.org/JPhysA/36/11143

Abstract

Typical performance of low-density parity-check (LDPC) codes over a general binary-input output-symmetric memoryless channel is investigated using methods of statistical mechanics. Relationship between the free energy in statistical-mechanics approach and the mutual information used in the information-theory literature is established within a general framework; Gallager and MacKay–Neal codes are studied as specific examples of LDPC codes. It is shown that basic properties of these codes known for particular channels, including their potential to saturate Shannon's bound, hold for general symmetric channels. The binary-input additive-white-Gaussian-noise channel and the binary-input Laplace channel are considered as specific channel models.

PACS numbers: 02.50.−r, 75.10.Hk, 89.70.+c, 89.20.Kk

1. Introduction

Low-density parity-check (LDPC) codes have drawn considerable attention in the information-theory literature due to their excellent performance, close to the information-theoretic upper bound (Shannon's bound) defined by the channel coding theorem [1]; this can be achieved with a feasible decoding effort by using the so-called sum-product algorithm. The standard approach in the information-theory literature to the analysis of the typical performance of LDPC codes is the density evolution analysis [2]. Density evolution analysis provides the decoding threshold, a critical noise level of a given communication channel below which decoding by the sum-product algorithm is typically successful, which means that the average decoding error probability is exponentially small in the codelength; above this value, decoding by the sum-product algorithm typically fails.

The statistical-mechanics approach to the analysis of LDPC codes has also attracted much interest in the literature [3–8, 10–12]. The statistical-mechanics-based analysis also gives us similar results to density evolution, but it tells us more. Equipped with the free energy,

it provides information about the structure of solutions, and in particular, the information-theoretic threshold, a second critical noise level below which optimal decoding (which in general requires a computational effort that grows exponentially with the codelength) provides a perfect error-correction and above which even optimal decoding fails. For example, statistical-physics-based analyses have discovered that for a certain type of LDPC codes the information-theoretic threshold is equal to Shannon's bound even for low connectivity. Existing statistical-mechanical studies on LDPC codes, however, have been mostly confined to the case of binary symmetric channel (BSC), which fits into the statistical-mechanical framework in a natural way [4–7]. Notable exceptions are the work by Montanari [8] that discusses a binary-input output-symmetric (BIOS) channel, which is a generic class of channel models including the additive-white-Gaussian-noise channel (BIAWGNC) as well as the BSC[3], and the study of Sourlas codes [3], a simple LDPC code, in which non-BSC channels are addressed [10–12].

Motivated by these observations, we relate the free energy to the mutual information, a measure of information transmission commonly used in the information-theory literature in a general setting, based on which we investigate the typical performance of LDPC codes over a BIOS memoryless channel. From the statistical-mechanical point of view, LDPC codes are regarded as dilute random spin systems, the fact that has motivated the statistical-mechanics studies of LDPC codes; it is therefore natural to expect that they will exhibit some sort of universality, just as typical statistical-mechanical systems do, so that general properties of LDPC codes observed in the BSC case will be preserved when different, more realistic, communication channels are considered. In this paper, we show that this is generally the case. In particular, we show that the LDPC codes can potentially saturate Shannon's bound for general BIOS channel.

The paper is organized as follows: in section 2 we introduce the general framework, notation, codes and the channels that we will focus on. In section 3 we will establish relationship between mutual information and free energy in a general setting, and will briefly present the analysis. Results obtained for the Gallager and MacKay–Neal (MN) codes will be described in sections 4 and 5, respectively, followed by the conclusions.

2. The general framework

2.1. Symmetric channels

We consider the general class of BIOS memoryless channels. The channel input is binary (± 1), and the output may take any real value. The characteristics of a channel are described by the channel transition probabilities, $P(y|x = 1)$ and $P(y|x = -1)$. Let $p(y) \equiv P(y|x = 1)$. A symmetric channel is characterized as a channel whose transition probabilities satisfy $P(y|x = -1) = P(-y|x = 1) = p(-y)$, which yields $P(y|x) = p(xy)$. Various types of channel models of practical interest fall into the class of BIOS channels, including the binary symmetric channel (BSC)

$$p_{\mathrm{BSC}}(y) = (1 - p)\delta(y - 1) + p\delta(y + 1) \tag{1}$$

the binary-input additive-white-Gaussian-noise channel (BIAWGNC)

$$p_{\mathrm{BIAWGNC}}(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-(y-1)^2/2\sigma^2} \tag{2}$$

[3] One should also mention a recent study using the weight and magnetization enumerator methods [9].

and the binary-input Laplace channel (BILC)

$$p_{\mathrm{BILC}}(y) = \frac{1}{\lambda}\,\mathrm{e}^{-|y-1|/\lambda}. \tag{3}$$

Each of the parameters p, σ^2 and λ represents the degree of degradation induced by the channel noise. We call each of them the noise level and let d denote the generic one. BIAWGNC and BILC channels are arguably more realistic and relevant than BSC in the case of practical communication channels, and have been studied extensively in the information-theory literature.

2.2. Gallager code

LDPC codes have been originally introduced by Gallager in his seminal work from 1963 [13]. Gallager's original construction [13] is one of the most extensively studied LDPC codes in the information-theory literature. It is defined by its parity-check matrix $A = [C_1|C_2]$ of dimensionality $(M - N) \times M$, which is taken to be random and very sparse. The submatrix C_2, of dimensionality $(M - N) \times (M - N)$, is assumed invertible.

In the encoding step, the encoder computes a codeword from the information vector $\boldsymbol{\xi} \in \{0, 1\}^N$ by employing a generator matrix G

$$\boldsymbol{x} = G^T\boldsymbol{\xi} \mod 2 \tag{4}$$

where the generator matrix is defined by

$$G^T = \begin{bmatrix} I \\ C_2^{-1}C_1 \end{bmatrix} \mod 2. \tag{5}$$

This construction ensures $AG^T = O \mod 2$. The information code rate for unbiased messages is $R = N/M$.

In regular Gallager codes, the number of non-zero elements per row of A is fixed to be K. Average number of non-zero elements per column is then $C \equiv K(M - N)/M$, whereas we will consider the case in which the number of non-zero elements in each column is forced to be exactly C. *Irregular* Gallager codes can be defined by relaxing these constraints on the numbers of non-zero elements. It has been known that making code construction irregular may improve performance significantly [14], but we will not discuss irregular codes in the current paper. We call the resulting regular Gallager code a (C, K)-Gallager code[4].

2.3. MN code

We also discuss a variant of LDPC codes, called the MacKay–Neal (MN) code [15, 16]. The generator matrix G of the MN code is defined by $G^T = C_n^{-1}C_s \mod 2$, where C_s and C_n are sparse matrices of dimensionality $M \times N$ and $M \times M$, respectively; C_n is assumed invertible. The information rate for the code is $R = N/M$ for unbiased message.

In regular MN codes the same constraints on the numbers of non-zero elements are imposed on both matrices C_s and C_n. The numbers of non-zero elements per row of C_s and C_n should be exactly K and L, respectively. Also here, we do not discuss irregular MN codes [17] in this paper. The numbers of non-zero elements per column of C_s and C_n are set to C and L, respectively, where $C = KM/N$ holds. We call the resulting code a (K, C, L)-MN code.

[4] Difference exists in the convention used to specify the regular Gallager code between the statistical-mechanics and information-theory literatures: (C, K)-Gallager code as defined in this paper is called the (K, C)-Gallager code in the statistical-mechanics literature. In this paper we are following the convention of the information-theory literature.

3. Analysis

3.1. Mutual information

The mutual information is a fundamental quantity in information theory, which measures the information transmitted through a communication channel. Consider a generic situation in which random variables S, X and Y represent the data to be estimated, the sent signal and the received signal, respectively. We assume that mapping from S to X is one-to-one and deterministic. For the mutual information $I(S;Y)$ we have

$$I(S;Y) = I(X;Y) = H(Y) - H(Y|X) \tag{6}$$

where $H(Y)$ is the entropy of Y, and where $H(Y|X)$ is the conditional entropy of Y conditioned on X (i.e., of the degradation process).

Let the data prior be $P(S)$ and the channel characteristics be $P(Y|X)$. The posterior probability of S conditioned on Y becomes

$$P(S|Y) = \frac{P(Y|X(S))P(S)}{P(Y)} \tag{7}$$

where $P(Y) = \sum_S P(Y|X(S))P(S)$ is the marginal probability of Y. The mutual information $I(S;Y)$ becomes

$$I(S;Y) = -\langle \log P(Y) \rangle_Y + \langle\langle \log P(Y|X) \rangle_{Y|X} \rangle_X \tag{8}$$

where the first and second terms on the right-hand side are the entropy of Y and that of the channel noise, respectively. Since $-\log P(Y)$ is nothing but the free energy utilized in various statistical-mechanical studies of information processing, this argument establishes the relationship between the mutual information and the free energy (averaged over the received signal) in various cases, including those of Gallager and MN codes to be presented in the following.

It should also be noted that absolute values of the mutual information have a proper operational meaning (the amount of information transfer, the quantity commonly measured in *bits* or *nats*), in contrast to the free energy, where only relative values are of relevance. This arguably implies that the mutual information, rather than the conventional free energy, is the more fundamental quantity when it comes to problems of information transmission.

3.2. Gallager code

The decoding problem of the Gallager code is to find $\boldsymbol{\tau}$ which is best supported (i.e., most probable) by the received signal $\boldsymbol{y}$ among the set $T_{\boldsymbol{J}}$ of $\boldsymbol{\tau}$ satisfying the parity-check equation ($A\boldsymbol{\zeta} = A\boldsymbol{\tau} \bmod 2$ if we write it in the $\{0, 1\}$-notation). Let

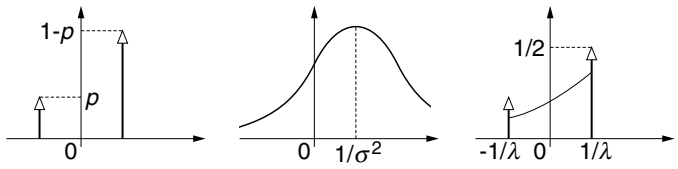
$$P_\gamma(\boldsymbol{\tau}; \boldsymbol{J}) = \mathcal{Z}_\gamma^{-1} \exp\left[-\gamma \sum_{\mu=1}^{M-N} \left(J_\mu \prod_{j \in \mathcal{L}(\mu)} \tau_j - 1 \right) \right] \tag{9}$$

where $\mathcal{L}(\mu) = \{j | A_{\mu j} = 1\}$ denotes the set of indices for which the parity-check matrix A has 1 in the μth row, and $J_\mu \equiv \prod_{j \in \mathcal{L}(\mu)} \zeta_j$ is the μth check. The set $T_{\boldsymbol{J}}$ is then expressed as the support of the 'prior' distribution $\lim_{\gamma \to \infty} P_\gamma(\boldsymbol{\tau}; \boldsymbol{J})$, as

$$T_{\boldsymbol{J}} = \left\{ \boldsymbol{\tau} \,\middle|\, \lim_{\gamma \to \infty} P_\gamma(\boldsymbol{\tau}; \boldsymbol{J}) > 0 \right\}. \tag{10}$$

Application of the Bayes formula yields the posterior probability of $\boldsymbol{\tau}$ conditioned on $\boldsymbol{y}$, as

$$P_\gamma(\boldsymbol{\tau}|\boldsymbol{y}; \boldsymbol{J}) = \frac{P(\boldsymbol{y}|\boldsymbol{\tau}) P_\gamma(\boldsymbol{\tau}; \boldsymbol{J})}{P_\gamma(\boldsymbol{y}; \boldsymbol{J})} \tag{11}$$



(*a*) BSC (*b*) BIAWGNC (*c*) BILC

Figure 1. Field distributions corresponding to various BIOS channels.

where

$$P_\gamma(\boldsymbol{y};\boldsymbol{J}) \equiv \sum_{\boldsymbol{\tau}} P(\boldsymbol{y}|\boldsymbol{\tau})P_\gamma(\boldsymbol{\tau};\boldsymbol{J}) \tag{12}$$

is the marginal probability of the received signal $\boldsymbol{y}$. The posterior acquires the Gibbs–Boltzmann form

$$P_\gamma(\boldsymbol{\tau}|\boldsymbol{y};\boldsymbol{J}) = \frac{1}{Z}\exp[-\beta\mathcal{H}_\gamma(\boldsymbol{\tau};\boldsymbol{y},\boldsymbol{J})] \tag{13}$$

with the Hamiltonian

$$\mathcal{H}_\gamma(\boldsymbol{\tau};\boldsymbol{y},\boldsymbol{J}) = -\gamma\sum_{\mu=1}^{M-N}\left(J_\mu\prod_{j\in\mathcal{L}(\mu)}\tau_j - 1\right) - \sum_{j=1}^{M}\log p(\tau_j y_j) + \log\mathcal{Z}_\gamma. \tag{14}$$

We have to take the limit $\gamma\to\infty$ and consider it at $\beta = 1$ (Nishimori's temperature [10, 19–21]) in order to obtain the true posterior. The marginal $P_\gamma(\boldsymbol{y};\boldsymbol{J})$ plays a role of the partition function.

As shown in [8, 18], the channel characteristics enter into the Hamiltonian (14) as the term $\log p(\tau_j y_j)$ which, in view of $\tau_j = \pm 1$, can be rewritten as

$$\log p(\tau_j y_j) = \tau_j\frac{1}{2}\log\frac{p(y_j)}{p(-y_j)} + \frac{1}{2}\log p(y_j)p(-y_j). \tag{15}$$

This means that it is the log-likelihood ratio $h(y_j) \equiv (1/2)\log(p(y_j)/p(-y_j))$ that serves as the external field acting on site j, and that the channel characteristics define the field distribution. Analysing the effect of having different communication channels on the properties of LDPC codes therefore reduces to investigating the effect of different field distributions on the physical properties of the system. Sketches of the field distribution for the BSC, BIAWGNC and BILC are shown in figure 1.

The assumption that the channel is memoryless is standard in the information-theory literature, and is also essential in the following analysis. In various application areas of digital communications, including those of wireless communications, one sometimes has to consider a channel with memory, since principal physical process of signal degradation may have larger time constant than the bit interval. In such cases, the memory effect induces correlations between the external field of different sites, which makes the analysis much more difficult.

We evaluate the mutual information $\mathcal{I}$ per transmitted symbol, of the variable to be estimated $\boldsymbol{\tau}$ (which is also the sent signal in the cases of Gallager and MN codes treated in this paper), and that of the received signal $\boldsymbol{y}$, in the infinite-codelength limit $M\to\infty$. From (8) we have

$$\mathcal{I} = -\lim_{M\to\infty} M^{-1}\langle\log P_\infty(\boldsymbol{y};\boldsymbol{J})\rangle_{\boldsymbol{y}} + \langle\log p(y)\rangle_y. \tag{16}$$

The difficulty in the evaluation of $\mathcal{I}$ lies in that of its 'free-energy' part

$$f \equiv -\lim_{M\to\infty} M^{-1}\langle \log P_\infty(\boldsymbol{y};\boldsymbol{J})\rangle_{\boldsymbol{y}} \tag{17}$$

for which we assume self-averaging over randomness of the parity-check matrix A and of $\boldsymbol{\zeta}$, and evaluate its average using the replica method, that is,

$$f = -\lim_{M\to\infty}\lim_{n\to 0} M^{-1}\frac{\partial}{\partial n}\log\langle [P_\infty(\boldsymbol{y};\boldsymbol{J})]^n\rangle_{A,\boldsymbol{y},\boldsymbol{\zeta}}. \tag{18}$$

In calculating the free energy, we perform the gauge transformation $\tau_j \to \zeta_j\tau_j$, $y_j \to \zeta_j y_j$, after which the average over $\boldsymbol{y}$ can be taken with respect to $\prod_{j=1}^M p(y_j)$.

The replica calculation basically follows the same line as in [5], we therefore omit details of the calculation. Exchanging the order of the two limits in (18) so that the limit $M\to\infty$ is taken first to apply the saddle-point method, one obtains

$$f = -\lim_{n\to 0}\frac{\partial}{\partial n}\operatorname*{Extr}_{\boldsymbol{q},\hat{\boldsymbol{q}}}\left[\frac{C}{K}\mathcal{G}_1(\boldsymbol{q}) - \mathcal{G}_2(\boldsymbol{q},\hat{\boldsymbol{q}}) + \mathcal{G}_3(\hat{\boldsymbol{q}})\right] + R\log 2 \tag{19}$$

where

$$\begin{aligned}
&\mathcal{G}_1(\boldsymbol{q}) \equiv \log\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle} q^K_{\alpha_1\cdots\alpha_m} - n\log 2 \qquad \mathcal{G}_2(\boldsymbol{q},\hat{\boldsymbol{q}}) \equiv \sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle} q_{\alpha_1\cdots\alpha_m}\hat{q}_{\alpha_1\cdots\alpha_m}\\
&\mathcal{G}_3(\hat{\boldsymbol{q}}) \equiv \log\left[\sum_{\tau^1,\ldots,\tau^n}\left\langle\prod_{\alpha=1}^{n} p(\tau^\alpha y)\right\rangle_y \frac{1}{C!}\left(\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}\hat{q}_{\alpha_1\cdots\alpha_m}\tau^{\alpha_1}\cdots\tau^{\alpha_m}\right)^C\right].
\end{aligned} \tag{20}$$

To proceed further we adopt the replica-symmetric (RS) ansatz, which assumes that the relevant quantities are invariant under any permutation of the replica indices $\{1,\ldots,n\}$. Following [5], we specifically let

$$q_{\alpha_1\cdots\alpha_m} = q_0\int u^m\pi(u)\,\mathrm{d}u \qquad \hat{q}_{\alpha_1\cdots\alpha_m} = \hat{q}_0\int \hat{u}^m\hat{\pi}(\hat{u})\,\mathrm{d}\hat{u}. \tag{21}$$

We will use the following simplifying notation:

$$\mathcal{D}^K_\pi(\boldsymbol{u}) \equiv \prod_{j=1}^{K}\pi(u_j)\,\mathrm{d}u_j. \tag{22}$$

The replica-symmetric free energy f^{RS} becomes

$$\begin{aligned}
f^{\mathrm{RS}} = \operatorname*{Extr}_{\pi,\hat{\pi}}\Bigg\{&\log 2 + C\iint\log(1+u\hat{u})\pi(u)\hat{\pi}(\hat{u})\,\mathrm{d}u\,\mathrm{d}\hat{u} - \frac{C}{K}\int\log\left(1+\prod_{j=1}^{K}u_j\right)\mathcal{D}^K_\pi(\boldsymbol{u})\\
&-\int\left\langle\log\left[p(y)\prod_{l=1}^{C}(1+\hat{u}_l) + p(-y)\prod_{l=1}^{C}(1-\hat{u}_l)\right]\right\rangle_y\mathcal{D}^C_{\hat{\pi}}(\hat{\boldsymbol{u}})\Bigg\}
\end{aligned} \tag{23}$$

in which q_0 and $\hat{q}_0$ have been eliminated using the extremization condition $q_0\hat{q}_0 = C$. The above free energy is similar to that of [8], and to that obtained for other channels [4, 6]; however, here we point to the direct relationship between the free energy and the mutual information (equation (16)). Heuristic construction of a sufficient condition to the extremization problem with respect to π and $\hat{\pi}$ is possible, and it gives the following saddle-point equations:

$$\begin{aligned}
\pi(u) &= \int\left\langle\delta\left[u - \tanh\left(h(y) + \sum_{l=1}^{C-1}\tanh^{-1}\hat{u}_l\right)\right]\right\rangle_y\mathcal{D}^{C-1}_{\hat{\pi}}(\hat{\boldsymbol{u}})\\
\hat{\pi}(\hat{u}) &= \int\delta\left(\hat{u} - \prod_{j=1}^{K-1}u_j\right)\mathcal{D}^{K-1}_{\pi}(\boldsymbol{u}).
\end{aligned} \tag{24}$$

There may exist more than one solution to the saddle-point equations. In such cases, in view of the saddle-point method used in deriving the replica-symmetric free energy, the solution which minimizes the mutual information $\mathcal{I}$ gives the globally stable state, defining the thermal equilibrium properties of the system. The significance of metastable states—stable solutions other than the globally stable state—is that they are regarded as defining the *practical* performance limits (i.e., the decoding threshold) rather than the information-theoretic threshold, as discussed later.

The performance of the code is quantified by the overlap $m = M^{-1}\sum_{k=1}^{M}\zeta_j\langle\tau_j\rangle$, which is given as $m = \int \mathrm{sign}(z)P(z)\,\mathrm{d}z$, where

$$P(z) = \int \left\langle \delta\left[z - \tanh\left(h(y) + \sum_{l=1}^{C}\tanh^{-1}\hat{u}_l\right)\right]\right\rangle_y \mathcal{D}_{\hat{\pi}}^{C}(\hat{\boldsymbol{u}}). \tag{25}$$

3.3. MN code

The decoding problem for the MN code is to find $\boldsymbol{S}$ and $\boldsymbol{\tau}$ which are the best suitable in view of the received signal $\boldsymbol{y}$ among the sets of $\boldsymbol{S}$ and $\boldsymbol{\tau}$ satisfying the parity-check equation ($C_s\boldsymbol{S} + C_n\boldsymbol{\tau} = C_s\boldsymbol{\xi} + C_n\boldsymbol{\zeta} \bmod 2$ if written in the {0, 1}-notation). Defining the μth component of the check $\boldsymbol{J}$ as $J_\mu = \prod_{j\in\mathcal{L}_s(\mu)}\xi_j\prod_{l\in\mathcal{L}_n(\mu)}\zeta_l$, where $\mathcal{L}_s(\mu) = \{j|(C_s)_{\mu j} = 1\}$ and $\mathcal{L}_n(\mu) = \{l|(C_n)_{\mu l} = 1\}$, the posterior probability of $\boldsymbol{S}$ and $\boldsymbol{\tau}$ conditioned on the received signal $\boldsymbol{y}$ and the check $\boldsymbol{J}$ is given by

$$P_\gamma(\boldsymbol{S},\boldsymbol{\tau}|\boldsymbol{y};\boldsymbol{J}) = \frac{1}{Z}\exp[-\beta\mathcal{H}_\gamma(\boldsymbol{S},\boldsymbol{\tau};\boldsymbol{y},\boldsymbol{J})] \tag{26}$$

in the limit $\gamma\to\infty$ and at $\beta = 1$, where the Hamiltonian $\mathcal{H}_\gamma(\boldsymbol{S},\boldsymbol{\tau};\boldsymbol{y},\boldsymbol{J})$ is defined as

$$\mathcal{H}_\gamma(\boldsymbol{S},\boldsymbol{\tau};\boldsymbol{y},\boldsymbol{J}) = -\gamma\sum_{\mu=1}^{M}\left(J_\mu\prod_{j\in\mathcal{L}_s(\mu)}S_j\prod_{l\in\mathcal{L}_n(\mu)}\tau_l - 1\right) - F_s\sum_{j=1}^{N}S_j - \sum_{l=1}^{M}\log p(\tau_l y_l) + \text{const} \tag{27}$$

where F_s is a parameter representing the bias of the information vector $\boldsymbol{\xi}$ in such a way that $P(\xi_j = \pm 1) = (1 \pm \tanh F_s)/2$ holds. Again, the channel characteristics define the random field acting on $\{\tau_l\}$ via the log-likelihood ratio $(1/2)\log(p(y)/p(-y))$.

We perform the gauge transformation $S_j \to \xi_j S_j$, $\tau_j \to \zeta_j\tau_j$ and $y_j \to \zeta_j\tau_j$, and introduce the mutual information per transmitted symbol $\mathcal{I} = -\lim_{M\to\infty}M^{-1}\langle\log P_\infty(\boldsymbol{y};\boldsymbol{J})\rangle_{\boldsymbol{y}} + \langle\log p(y)\rangle_y$. Assuming the self-averaging property to hold, the replica calculation can be done along the same way as in [5]. The 'free-energy' part of the mutual information $\mathcal{I}$ becomes

$$f = -\lim_{n\to 0}\frac{\partial}{\partial n}\operatorname*{Extr}_{\boldsymbol{q},\hat{\boldsymbol{q}},\boldsymbol{r},\hat{\boldsymbol{r}}}[\mathcal{G}_1(\boldsymbol{q},\boldsymbol{r}) - \mathcal{G}_2(\boldsymbol{q},\hat{\boldsymbol{q}},\boldsymbol{r},\hat{\boldsymbol{r}}) + \mathcal{G}_3(\hat{\boldsymbol{q}},\hat{\boldsymbol{r}})] + R\log 2 \tag{28}$$

where

$$\begin{aligned}\mathcal{G}_1(\boldsymbol{q},\boldsymbol{r}) &\equiv \log\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}q^K_{\alpha_1\cdots\alpha_m}r^L_{\alpha_1\cdots\alpha_m} - n\log 2\\ \mathcal{G}_2(\boldsymbol{q},\hat{\boldsymbol{q}},\boldsymbol{r},\hat{\boldsymbol{r}}) &\equiv \frac{N}{M}\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}q_{\alpha_1\cdots\alpha_m}\hat{q}_{\alpha_1\cdots\alpha_m} + \sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}r_{\alpha_1\cdots\alpha_m}\hat{r}_{\alpha_1\cdots\alpha_m}\end{aligned} \tag{29}$$

and

$$\mathcal{G}_3(\hat{\boldsymbol{q}},\hat{\boldsymbol{r}}) \equiv \frac{N}{M}\log\left[\sum_{S^1,\ldots,S^n}\left\langle \mathrm{e}^{F_s\sum_{\alpha=1}^n \xi S^\alpha}\right\rangle_\xi \frac{1}{C!}\left(\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}\hat{q}_{\alpha_1\cdots\alpha_m}S^{\alpha_1}\cdots S^{\alpha_m}\right)^C\right] + \log\left[\sum_{\tau^1,\ldots,\tau^n}\left\langle\prod_{\alpha=1}^n p(\tau^\alpha y)\right\rangle_y \frac{1}{L!}\left(\sum_{m=0}^{n}\sum_{\langle\alpha_1\cdots\alpha_m\rangle}\hat{r}_{\alpha_1\cdots\alpha_m}\tau^{\alpha_1}\cdots\tau^{\alpha_m}\right)^L\right]. \tag{30}$$

We adopt the RS ansatz as before, under which we have [5]

$$r_{\alpha_1\cdots\alpha_m} = r_0\int v^m\rho(v)\,\mathrm{d}v \qquad \hat{r}_{\alpha_1\cdots\alpha_m} = \hat{r}_0\int \hat{v}^m\hat{\rho}(\hat{v})\,\mathrm{d}\hat{v} \tag{31}$$

in addition to (21). The replica-symmetric free energy f^{RS} becomes

$$\begin{aligned} f^{\mathrm{RS}} = \operatorname*{Extr}_{\pi,\hat{\pi},\rho,\hat{\rho}}\Bigg\{ & (1+R)\log 2 + K\iint\log(1+u\hat{u})\pi(u)\hat{\pi}(\hat{u})\,\mathrm{d}u\,\mathrm{d}\hat{u} \\ & + L\iint\log(1+v\hat{v})\rho(v)\hat{\rho}(\hat{v})\,\mathrm{d}v\,\mathrm{d}\hat{v} \\ & - \iint\log\left(1+\prod_{k=1}^K u_k\prod_{l=1}^L v_l\right)\mathcal{D}_\pi^K(\boldsymbol{u})\mathcal{D}_\rho^L(\boldsymbol{v}) \\ & - \frac{K}{C}\int\left\langle\log\left[\sum_{S=\pm1}\mathrm{e}^{F_s\xi S}\prod_{k=1}^C(1+S\hat{u}_k)\right]\right\rangle_\xi \mathcal{D}_{\hat{\pi}}^C(\hat{\boldsymbol{u}}) \\ & - \int\left\langle\log\left[\sum_{\tau=\pm1}p(\tau y)\prod_{l=1}^L(1+\tau\hat{v}_l)\right]\right\rangle_y \mathcal{D}_{\hat{\rho}}^L(\hat{\boldsymbol{v}})\Bigg\} \end{aligned} \tag{32}$$

in which q_0, $\hat{q}_0$, r_0 and $\hat{r}_0$ have been eliminated using the extremization conditions, $q_0\hat{q}_0 = C$ and $r_0\hat{r}_0 = L$.

Construction of a heuristic solution to the extremization problem can be done in the same manner, which yields the following saddle-point equations:

$$\begin{aligned} \pi(u) &= \int\left\langle\delta\left[u-\tanh\left(F_s\xi+\sum_{l=1}^{C-1}\tanh^{-1}\hat{u}_l\right)\right]\right\rangle_\xi \mathcal{D}_{\hat{\pi}}^{C-1}(\hat{\boldsymbol{u}}) \\ \hat{\pi}(\hat{u}) &= \iint\delta\left(\hat{u}-\prod_{k=1}^{K-1}u_k\prod_{l=1}^L v_l\right)\mathcal{D}_\pi^{K-1}(\boldsymbol{u})\mathcal{D}_\rho^L(\boldsymbol{v}) \\ \rho(v) &= \int\left\langle\delta\left[v-\tanh\left(h(y)+\sum_{l=1}^{L-1}\tanh^{-1}\hat{v}_l\right)\right]\right\rangle_y \mathcal{D}_{\hat{\rho}}^{L-1}(\hat{\boldsymbol{v}}) \\ \hat{\rho}(\hat{v}) &= \iint\delta\left(\hat{v}-\prod_{k=1}^{K}u_k\prod_{l=1}^{L-1}v_l\right)\mathcal{D}_\pi^K(\boldsymbol{u})\mathcal{D}_\rho^{L-1}(\boldsymbol{v}). \end{aligned} \tag{33}$$

Also here, the saddle-point equations may have more than one solution. The globally stable state and metastable states are defined in the same way as in the Gallager code case.

The overlap is then evaluated by $m = \int \mathrm{sign}(z) P(z)\, \mathrm{d}z$, with

$$P(z) = \int \left\langle \delta \left[z - \tanh \left(F_s \xi + \sum_{l=1}^{C} \tanh^{-1} \hat{u}_l \right) \right] \right\rangle_{\xi} \mathcal{D}_{\hat{\pi}}^{C}(\hat{\boldsymbol{u}}). \tag{34}$$

When the message is unbiased ($F_s = 0$) and K is even, saddle-point solutions have the following symmetry: for each solution $\{\pi(u), \hat{\pi}(\hat{u}), \rho(v), \hat{\rho}(\hat{v})\}$ there is another solution $\{\pi(-u), \hat{\pi}(-\hat{u}), \rho(v), \hat{\rho}(\hat{v})\}$. The latter has the same overlap as that of the former with the opposite sign.

4. Results—Gallager code

4.1. Analytical solutions

It has been shown [8] that there are two analytical solutions to the saddle-point equation (24) for the general BIOS channel. One is the ferromagnetic solution, valid for any values of K and C (provided that $K, C \geqslant 2$), and another is the sub-optimal ferromagnetic solution (which was termed the paramagnetic phase in [8]), which is valid in the limit $K \to \infty$ (C may be finite, although [8] requires $C \to \infty$ as well).

The ferromagnetic solution is given by $\pi(u) = \delta(u - 1)$ and $\hat{\pi}(\hat{u}) = \delta(\hat{u} - 1)$. For the ferromagnetic solution we have $m_{\mathrm{ferro}} = 1$ and $\mathcal{I}_{\mathrm{ferro}} = R \log 2$, the former means that this solution corresponds to an error-free communication; and, by noting that $R \log 2$ is equal to the rate of the sent information, the latter means that asymptotically (as $M \to \infty$) there is no loss of information due to encoding and/or transmission, provided that the ferromagnetic solution is the globally stable state. It should be noted that the absence of information loss does not necessarily mean that practical decoding is possible.

The sub-optimal ferromagnetic solution is given by $\pi(u) = \langle \delta[u - \tanh h(y)] \rangle_y$, $\hat{\pi}(\hat{u}) = \delta(\hat{u})$, for which $m_{\mathrm{sf}} = \langle \mathrm{sign}[p(y) - p(-y)] \rangle_y$ and $\mathcal{I}_{\mathrm{sf}} = \mathsf{C}$ hold, where C is the channel capacity of the BIOS channel defined as

$$\mathsf{C} = \log 2 - \langle \log[p(y) + p(-y)] \rangle_y + \langle \log p(y) \rangle_y. \tag{35}$$

If $R \log 2 < \mathsf{C}$, we have $\mathcal{I}_{\mathrm{ferro}} < \mathcal{I}_{\mathrm{sf}}$, so that the ferromagnetic solution gives the globally stable state. On the other hand, if $R \log 2 > \mathsf{C}$, the sub-optimal ferromagnetic solution gives the globally stable state (no other solution has been identified in this case). This proves that the thermodynamic transition between the ferromagnetic and sub-optimal ferromagnetic solutions occurs at the theoretical limit $R \log 2 = \mathsf{C}$; and the maximum rate $R_{\max}$, up to which error-free communication is theoretically possible, asymptotically achieves the theoretical limit as $K \to \infty$. This result has been known for BSC [6, 7] and for BIAWGNC [8] in the physics literature and is in agreement with results reported in the information-theory literature [16]. The current result is an extension to the case of a *general BIOS channel*.

4.2. Numerical solutions of saddle-point equations

In finite-K cases no simple analytical solution exists other than the ferromagnetic one, so one has to solve the saddle-point equations numerically. We have done it for the BIAWGNC and BILC. The dependence of the overlap m on the noise level d (σ^2 for BIAWGNC, and λ for BILC) is qualitatively the same as that observed in BSC: for $K \geqslant 3$ the ferromagnetic solution with $m = 1$ is locally stable over the whole range of noise levels. At $d = d_s$, another solution with $m < 1$ appears, which defines the spinodal point; this is also termed the *dynamical transition* point in the physics literature. At a higher noise level $d = d_t > d_s$

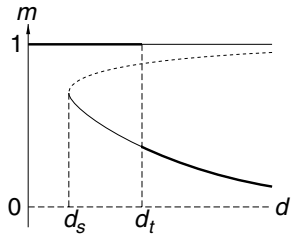


Figure 2. Noise-overlap diagram for Gallager code. Thick solid lines stand for the stable state, thin solid lines for metastable state and broken lines for unstable states. The ferromagnetic solution is the one with $m = 1$, while $m < 1$ defines the suboptimal ferromagnetic solution.

Table 1. The parameter values d_s and d_t at the spinodal point and thermodynamic transition, respectively, for (C, K)-Gallager codes over the BIAWGNC ($d \equiv \sigma^2$) and BILC ($d \equiv \lambda$) for various code parameters; d_0, denoting Shannon's bound for error-free communication, is also shown.

C	K	R	σ_s^2	σ_t^2	σ_0^2	λ_s	λ_t	λ_0
3	6	0.5	0.775	0.899	0.958	0.651	0.712	0.752
4	8	0.5	0.701	0.943	0.958	0.618	0.741	0.752
5	10	0.5	0.629	0.952	0.958	0.581	0.746	0.752
3	5	0.4	1.017	1.253	1.321	0.773	0.875	0.914
4	6	0.333	1.020	1.666	1.681	0.782	1.045	1.055
3	4	0.25	1.598	2.325	2.401	1.018	1.260	1.298

thermodynamic transition takes place, beyond which the ferromagnetic solution with $m = 1$ becomes metastable (see figure 2). The thermodynamic transition point d_t is upper bounded by the Shannon's bound of the noise level d_0 allowing error-free communication, which is defined by $R \log 2 = \mathsf{C}$. Thus, in general, $d_s < d_t \leqslant d_0$ holds. Table 1 summarizes the results for the BIAWGNC and BILC cases, showing the spinodal point d_s, the thermodynamic transition point d_t and Shannon's bound d_0 allowing error-free communication.

It should be noted that the results for the spinodal point agree well with the results obtained by the density evolution approach [2], as expected, since the saddle-point equations by the replica analysis happen to coincide with the time evolution equations in the density evolution. One can therefore expect that the spinodal point d_s defines the decoding threshold. Further discussion regarding this point can be found in [22]. On the other hand, the thermodynamic transition point d_t defines the information-theoretic threshold. Empirically, the decoding threshold d_s can be achieved by linear complexity decoders, such as the belief-propagation decoder [2, 16]. However, one should note that the information-theoretic threshold d_t can only be achieved by exhaustive computation of the posterior distribution, which is infeasible in practice.

Our results in this paper are based on the RS ansatz. However, it has been reported (e.g., [22]) that the suboptimal solution appearing at d_s has negative entropy so that one has to consider replica-symmetry-breaking (RSB) in order to characterize the suboptimal solution near d_s precisely. Although this may affect the values of d_s, accumulated evidences in the information-theory literature [2] as well as the analysis on Gallager codes over the binary erasure channel [22] suggests that the values of d_s are not significantly affected by taking RSB into account. We therefore assume that our results based on the RS ansatz represent the true decoding threshold of the system. General analysis based on RSB is quite complicated and

beyond the scope of this paper. The same argument also applies to the case of MN codes presented in the following.

5. Results—MN code

5.1. Analytical solutions

In the following we restrict our discussion of the MN code to the unbiased case $F_s = 0$. The ferromagnetic solution, corresponding to the error-free communication, can be constructed for the MN code with $L \geqslant 2$. (In fact, in the case $L = 1$ the matrix C_n reduces to a simple permutation matrix, so that we have to estimate each element of noise separately. This case is not at all interesting and therefore we will not discuss it any more.) It is given by

$$\pi(u) = \delta(u-1) \qquad \hat{\pi}(\hat{u}) = \delta(\hat{u}-1) \qquad \rho(v) = \delta(v-1) \qquad \hat{\rho}(\hat{v}) = \delta(\hat{v}-1) \quad (36)$$

for which $m_{\text{ferro}} = 1$ and $\mathcal{I}_{\text{ferro}} = R\log 2$, which again means that theoretically there is no loss of information, provided that the ferromagnetic solution is the globally stable state.

The MN code has the following paramagnetic solution for $K \geqslant 2$:

$$\pi(u) = \delta(u) \qquad \hat{\pi}(\hat{u}) = \delta(\hat{u}) \qquad \rho(v) = \langle \delta[v - \tanh h(y)] \rangle_y \qquad \hat{\rho}(\hat{v}) = \delta(\hat{v}) \quad (37)$$

which yields $m_{\text{para}} = 0$ and $\mathcal{I}_{\text{para}} = \mathsf{C}$. Again, since $\mathcal{I}_{\text{ferro}} < \mathcal{I}_{\text{para}}$ holds for $R\log 2 < \mathsf{C}$, we conclude that for the MN code the maximum rate $R_{\max}$, theoretically allowing error-free communication, achieves the theoretical limit as long as $K \geqslant 2, L \geqslant 2$, provided that there is no locally stable solution other than the ferromagnetic and paramagnetic solutions (for the case $K = 2$, however, there do exist stable solutions other than these two, as shown in the following). This result is an extension of the result reported in [4, 5] to the case of a *general BIOS channel*.

It should be noted that the paramagnetic solution (36) is also valid in the limit $L \to \infty$ for the case $K = 1$. This means that the above-mentioned result also holds for the case of $K = 1$ asymptotically in the limit $L \to \infty$. Note also that these statements do not imply that the error-free communication is possible at the theoretical limit in the practical sense. As in the case of the Gallager codes, the practical performance is affected by the existence of metastable states.

5.2. Numerical solutions of saddle-point equations

In order to explore solutions other than the ferromagnetic and paramagnetic solutions, we have to solve the saddle-point equations numerically. We have done it for the BIAWGNC and BILC cases. We observed qualitatively the same characteristics as those reported in [5].

The obtained numerical results suggest that the qualitative physical properties are categorized into three types according to the K value: cases with $K = 1, K = 2$ and $K \geqslant 3$, whereas it is only affected quantitatively by the values of C and L, as described in the following.

The structure of noise-overlap diagram for the MN code with $K = 1$ is qualitatively the same as that for Gallager code (see figure 2). At very low noise level only the ferromagnetic solution with $m = 1$ exists. At a certain noise level $d = d_s$ another metastable solution with $m < 1$ appears, and it becomes dominant beyond $d = d_t > d_s$. The spinodal point d_s again defines the decoding threshold. Numerical results show (see table 2) that in general the thermodynamic transition point d_t *is* smaller than Shannon's bound d_0. It is also observed that, for fixed C, increasing L makes d_s smaller and d_t larger, the latter of which approaches Shannon's bound d_0 as $L \to \infty$, as discussed at the end of the previous subsection. Even

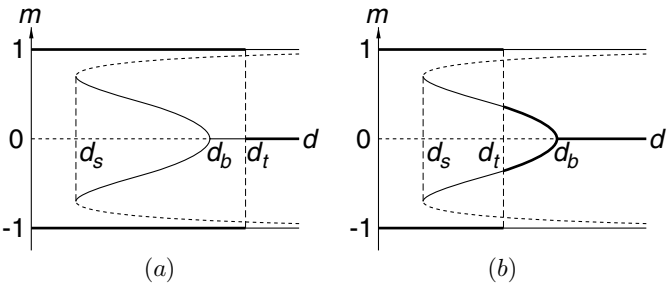


Figure 3. Noise-overlap diagrams for MN code with $K = 2$.

Table 2. The parameter values d_s, d_t and d_b at the spinodal point and thermodynamic transition, and at bifurcation of paramagnetic solution, respectively, for (K, C, L)-MN codes over the BIAWGNC ($d \equiv \sigma^2$) and BILC ($d \equiv \lambda$) with various code parameters; d_0, denoting Shannon's bound for error-free communication, is also shown.

K	C	L	R	σ_s^2	σ_t^2	σ_b^2	σ_0^2	λ_s	λ_t	λ_b	λ_0
1	2	3	0.5	0.775	0.901	–	0.958	0.652	0.714	–	0.752
1	2	4	0.5	0.703	0.944	–	0.958	0.619	0.740	–	0.752
1	2	5	0.5	0.630	0.955	–	0.958	0.582	0.748	–	0.752
1	3	2	0.333	1.338	1.423	–	1.681	0.903	0.934	–	1.055
1	3	3	0.333	1.129	1.659	–	1.681	0.831	1.040	–	1.055
1	3	4	0.333	0.913	1.672	–	1.681	0.735	1.051	–	1.055
2	3	2	0.667	0.536	0.587	0.612	0.588	0.525	0.551	0.597	0.553
2	3	3	0.667	0.430	0.588	0.459	0.588	0.464	0.553	0.493	0.553
2	3	4	0.667	0.368	0.588	0.385	0.588	0.419	0.553	0.437	0.553
2	4	2	0.5	0.809	0.958	0.919	0.958	0.689	0.751	0.771	0.752
2	5	2	0.4	1.039	1.321	1.175	1.321	0.807	0.914	0.894	0.914

for finite L the value of d_t may be numerically very close to d_0, especially when the rate R is small. These properties have already been reported for the BSC case [5], so that our finding implies that they also hold for the BIAWGNC and BILC cases, revealing some sort of universality.

The noise-overlap diagram for the cases with $K = 2$ has the general structure shown in figure 3. The diagram is characterized by three transition points: the spinodal point d_s (which defines the decoding threshold), the thermodynamic transition point d_t (which defines the information-theoretic threshold) and the bifurcation point d_b (at which the stability of the paramagnetic solution changes). The order of the thermodynamic transition point d_t and the bifurcation point d_b varies with the values of C and L, so that the bifurcation pattern for the cases with $K = 2$ is further divided into two sub-categories depending on the order of the two transitions: $d_s < d_b < d_t$ for the first group and $d_s < d_t < d_b$ for the second group. The noise-overlap diagrams for these groups are illustrated in figures 3(a) and (b), respectively. By the local stability analysis the bifurcation point d_b is determined by

$$\int v\rho(v)\,\mathrm{d}v = (C-1)^{-1/L} \tag{38}$$

with $\rho(v)$ as given in (37), which allows us to decide the type of bifurcation of a particular case. See the appendix for derivation of (38). As a result, we found that only a few cases with small values of C and L fall into the second category. The values of C and L for which the

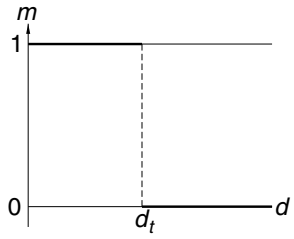


Figure 4. Noise-overlap diagram for MN code with $K \geqslant 3$.

$(2, C, L)$-MN code falls into the second category depend on the channel characteristics; as far as we have observed, only the cases with $L = 2$ fall into the second group. For the BIAWGNC case, the $(2, 3, 2)$-MN code is the only one instance, whereas for the BILC case, both $(2, 3, 2)$- and $(2, 4, 2)$-MN codes belong to this group. (For the BSC case, $(2, 3, 2)$-, $(2, 4, 2)$- and $(2, 5, 2)$-MN codes belong to this group.) All the $(2, C, L)$-MN codes but those mentioned above are in the first group. For the cases in the second group, the thermodynamic transition point d_t must be less than the information-theoretic limit d_0, but it is observed numerically that d_t is very close to d_0.

We observed that the noise-overlap diagram for the cases with $K \geqslant 3$ is relatively simple for the BIAWGNC and BILC cases, just as in the BSC case (see figure 4). The ferromagnetic solution with $m = 1$ (and its mirror image with $m = -1$ when K is even) and the paramagnetic solution are the only stable solutions found, both of which are locally stable over the whole range of the noise level. The system exhibits a first-order transition at Shannon's bound d_t. We did not find any solutions other than the ferromagnetic and paramagnetic solutions.

6. Conclusions

We have analysed the typical performance of LDPC codes over BIOS channel using statistical mechanics. Since it has been shown in [8, 18] that the log-likelihood ratio of the received signal acts as an external random field acting on each site, and that channel characteristics define the distribution of the random field, our analysis amounts to investigating the effect of the random field distribution on the performance. Relationship between the mutual information and the free energy has been established in a quite general setting, including the case of LDPC codes. Both Gallager and MN codes are analysed under the RS ansatz, to find that the basic properties of these codes remain unchanged regardless of channel characteristics. In particular, it has been shown for general BIOS channel that these codes potentially saturate Shannon's limit asymptotically, as $K \to \infty$, for the Gallager code; and when $K, L \geqslant 2$—with a few exceptions with small C and L values—and asymptotically as $L \to \infty$ for $K = 1$, for the MN code. Saddle-point solutions have also been numerically evaluated extensively for the cases of BIAWGNC and BILC channels, from which noise-overlap diagrams, as well as the spinodal, thermodynamic transition, and other bifurcation points, have been characterized.

Some of the results obtained, in particular the dynamical transition point values, have been linked to the limitation of practical decoding algorithms; more specifically, to the difficulty of finding optimal solutions in the presence of metastable states. It would be extremely interesting to study the practical decoding capabilities of novel decoding algorithms based on survey propagation [24], that have been shown to find optimal solutions even when metastable states are present.

Acknowledgments

We would like to thank Yoshiyuki Kabashima for his helpful suggestions, Jort van Mourik for providing his computer programs and Nikos Skantzos for helpful discussions. Supports from EPSRC research grant GR/N00562, from Bilateral Program between the UK and Japan, JSPS, Japan, and from grant-in-aid for scientific research on priority areas 14084209, MEXT, Japan, are acknowledged.

Appendix. Stability of paramagnetic solution of MN codes with $K \geqslant 2$

To probe the stability of paramagnetic solution, which exists for MN codes with $K \geqslant 2$, we analyse the stability with respect to $\boldsymbol{q}$ and $\boldsymbol{r}$ only, and do not consider stability with respect to $\hat{\boldsymbol{q}}$ and $\hat{\boldsymbol{r}}$. As discussed in [23], these conjugate variables are subsidiary to their counterparts, $\boldsymbol{q}$ and $\boldsymbol{r}$, respectively, so that the former should not be considered as independent variables.

Let $A, B, \ldots$ denote sets of replica indices such as $\langle \alpha_1 \cdots \alpha_m \rangle$, $m \geqslant 1$. We first evaluate the Hessian of the free energy (28) with respect to $4 \times (2^n - 1)$ variables $\{q_A, \hat{q}_A, r_A, \hat{r}_A\}$:

$$H = \begin{pmatrix} H_{qq} & H_{q\hat{q}} & & O \\ H_{q\hat{q}} & H_{\hat{q}\hat{q}} & & \\ & O & O & H_{r\hat{r}} \\ & & H_{r\hat{r}} & H_{\hat{r}\hat{r}} \end{pmatrix} \tag{39}$$

where

$$\left(H_{qq}\right)_{AB} = \begin{cases} 0 & (K \geqslant 3) \\ -\frac{K}{q_0^2}\left(\frac{r_A}{r_0}\right)^L \delta_{AB} & (K = 2) \end{cases} \tag{40}$$

and where $l(H_{q\hat{q}})_{AB} = (N/M)\delta_{AB}$, $(H_{\hat{q}\hat{q}})_{AB} = -\left[K(C-1)/\hat{q}_0^2\right]\delta_{AB}$, $(H_{r\hat{r}})_{AB} = \delta_{AB}$, $(H_{\hat{r}\hat{r}})_{AB} = -\left[L(L-1)/\hat{r}_0^2\right]\delta_{AB}$. The block-diagonal structure of the Hessian allows us to decompose the stability problem into two, one with respect to $\boldsymbol{q}$, and another with respect to $\boldsymbol{r}$.

Following the argument in the appendix of [23], one can say that the system is stable with respect to $\boldsymbol{q}$ (with $\hat{\boldsymbol{q}}$ depending on $\boldsymbol{q}$) if the matrix $H_c \equiv H_{qq} - H_{q\hat{q}}(H_{\hat{q}\hat{q}})^{-1}H_{q\hat{q}}$ is positive definite. A corresponding statement holds for the stability with respect to $\boldsymbol{r}$.

The stability with respect to $\boldsymbol{r}$ is straightforward, by noting that the matrix $H_{\hat{r}\hat{r}}$ is negative definite, which means that $H_c = -(H_{\hat{r}\hat{r}})^{-1}$ is positive definite.

We consider the stability with respect to $\boldsymbol{q}$. For $K \geqslant 3$, we have $H_c = \hat{q}_0^2[K/C^2(C-1)]I$, where I is the identity matrix, so that the stability immediately follows, irrespective of the noise level of the channel. For $K = 2$, the matrix H_c is diagonal, and its Ath element is

$$(H_c)_{AA} = -\frac{K}{q_0^2}\left(\frac{r_A}{r_0}\right)^L + \frac{K\hat{q}_0^2}{C^2(C-1)}. \tag{41}$$

Using the equality which holds under the RS ansatz,

$$\frac{r_A}{r_0} = \int_{-1}^{1} v^m \rho(v)\,\mathrm{d}v \tag{42}$$

where $A = \alpha_1 \cdots \alpha_m$, we have, as the stability condition,

$$E_m \equiv \int_{-1}^{1} v^m \rho(v)\,\mathrm{d}v < (C-1)^{-1/L} \tag{43}$$

for $m = 1, \ldots, n$. Since it can be shown that $E_{2m-1} = E_{2m}$ and $E_{2m} \geqslant E_{2m+2}$, the critical condition determining the stability is $E_1 < (C-1)^{-1/L}$, which yields (38).

References

[1] Shannon C E 1948 *Bell Syst. Tech. J.* **27** 379, 623
[2] Richardson T J and Urbanke R L 2001 *IEEE Trans. Inform. Theory* **47** 599
[3] Sourlas N 1989 *Nature (London)* **339** 693
[4] Kabashima Y, Murayama T and Saad D 2000 *Phys. Rev. Lett.* **84** 1355
[5] Murayama T, Kabashima Y, Saad D and Vicente R 2000 *Phys. Rev.* E **62** 1577
[6] Vicente R, Saad D and Kabashima Y 2000 *Europhys. Lett.* **51** 698
[7] van Mourik J, Saad D and Kabashima Y 2002 *Phys. Rev.* E **66** 026705
[8] Montanari A 2001 *Eur. Phys. J.* B **23** 121
[9] Skantzos N, van Mourik J and Saad D 2003 *Phys. Rev.* E **67** 037101
[10] Ruján P 1993 *Phys. Rev. Lett.* **70** 2968
[11] Nishimori H and Michael Wong K Y 1999 *Phys. Rev.* E **60** 132
[12] Vicente R, Saad D and Kabashima Y 1999 *Phys. Rev.* E **60** 5352
[13] Gallager R G 1962 *IRE Trans. Inform. Theory* **IT-8** 21
[14] Richardson T J, Shokrollahi M A and Urbanke R L 2001 *IEEE Trans. Inform. Theory* **47** 619
[15] MacKay D J C and Neal R M 1995 *Cryptography and Coding, 5th IMA Conf. (Lecture Notes in Computer Science)* vol 1025 ed C Boyd (Berlin: Springer) p 100
[16] MacKay D J C 1999 *IEEE Trans. Inform. Theory* **45** 399
[17] Kanter I and Saad D 1999 *Phys. Rev. Lett.* **83** 2660
[18] Sourlas N 1999 *Proc. Marseille Satellite Colloquium Mathematical Results in Stat. Mechanics Preprint* cond-mat/9811406
[19] Nishimori H 1993 *J. Phys. Soc. Jpn.* **62** 2973
[20] Iba Y 1999 *J. Phys. A: Math. Gen.* **32** 3875
[21] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing* (Oxford: Oxford University Press)
[22] Franz S, Leone M, Montanari A and Ricci-Tersenghi F 2002 *Phys. Rev.* E **66** 046120
[23] Tanaka T 2002 *IEEE Trans. Inform. Theory* **48** 2888
[24] Mézard M, Parisi G and Zecchina R 2002 *Science* **297** 812