Average and reliability error exponents in low-density parity-check codes

N.S. Skantzos[†*]**, J. van Mourik**[†]**, D. Saad**[†] **and Y. Kabashima**[‡]

[†]Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK
[*] Institut for Theoretical Physics, Celestijnenlaan 200D, KULeuven, Leuven, B-3001 Belgium
[‡]Dept. of Computational Intelligence & Systems Science, Tokyo Institute of Technology, Yokohama 2268502, Japan

Abstract. We present a theoretical method for a *direct* evaluation of the average and reliability error exponents in low-density parity-check error-correcting codes using methods of statistical physics. Results for the binary symmetric channel (BSC) are presented for codes of both finite and infinite connectivity.

1. Introduction

Low-density parity-check codes (LDPC) have attracted significant interest in recent years due to their simplicity and exceptionally high performance [1]. Their simplicity and inherent randomness make them amenable to analysis using established methods in the area of statistical physics. These have been employed in a number of papers [2]-[9] to gain insight into the properties of LDPC codes and to evaluate their performance.

These studies include the evaluation of critical noise levels for given codes [2], an exact calculation of weight and magnetisation enumerators [4], the performance of irregular codes [3], properties of codes in real-valued channels [5], and the derivation of bounds for the reliability exponent [6], to name but a few. These studies also represent the interdisciplinary nature of this research area and illustrate the successful interaction between researchers in the two disciplines.

The evaluation of error exponents has been a long-standing problem in information theory [10, 11]. Efforts to obtain exact expressions and/or bounds to the error exponent resulted in partial success; although tight bounds have been derived in the case of random codes and LDPC with infinite connectivity [10], only limited results have been obtained for sparely connected codes. Main stream techniques to tackle the problem include sphere-packing and union-bound arguments [11, 10]. Below a certain code-rate value, the estimated bounds also become loose and require using the 'expurgated exponent' techniques [10] for obtaining a tighter bound.

In this paper, we employ methods of statistical physics to evaluate directly the average error exponent and typical reliability exponent in Gallager and MN [12] LDPC codes. The *average error exponent* is obtained by carrying out averages over the ensemble of randomly generated LDPC codes of given rate and connectivity; while the *reliability exponent* is obtained by selecting the best codes in that ensemble. Averages result in the emergence of macroscopic variables, representative of the ensemble properties, that can be obtained numerically and used to calculate the average error exponent (in the current calculation we assume that short loops, which contribute polynomially to the block error probability in LDPC codes [13], have been removed). Average error exponent solutions have been obtained for both finite and infinite connectivity vector ensembles, while reliability exponent solutions have been obtained only in the case of infinite connectivity.

As a reference point to test our theory, we use known results obtained in the information theory literature for solvable limits (e.g. codes of infinite connectivity), and find that our method reproduces them exactly. Perhaps not surprisingly, we also find that at fixed noise level and code rate, the reliability exponent for codes of finite connectivity is always upper-bounded by that of the infinite-connectivity case.

Before we proceed, the distinction between the typical bounds found previously using methods of statistical physics [6], and the current calculation should be clarified. In the former, one employs methods of statistical physics to calculate the typical value of a *bound* based on inequalities introduced by Gallager; while in the current calculation, a direct estimation of the average error exponent, rather than a bound, is sought. An additional advantage of the current approach is that it can be extended to provide reliability exponent values for LDPC codes by restricted averages over codes of high performance.

The paper is organised as follows: In section 2, we introduce the general coding framework and the technique used. In sections 3 and 4 we present an outline of the derivation and the solutions obtained in both finite and infinite connectivity cases respectively. In Section 5 we compare the error exponent results obtained for MN codes to those of Gallager codes in both finite and infinite connectivity cases. Discussion and conclusions are presented in section 6.

2. Definitions

A regular (k, j) Gallager error-correcting code is defined by the binary $(N - K) \times N$ (parity check) matrix $A = [C_1 | C_2]$, which is known to both sender and receiver. The $(N - K) \times (N - K)$ matrix C_2 is taken to be invertible. The number of non-zero elements in each row of A is given by k, while the number of non-zero elements per column is given by $j \equiv k(N - K)/N$.

Gallager's encoding scheme consists of generating a codeword $\boldsymbol{t} \in \{0, 1\}^N$ from an information (message) vector $\boldsymbol{s} \in \{0, 1\}^K$ (with $N > K$) via the linear operation $\boldsymbol{t} = G^T \boldsymbol{s}$ (mod 2) where G is the generator matrix defined by $G = [I | C_2^{-1} C_1]$ (mod 2).

The code rate is then given by $R \equiv K/N = 1 - j/k$, and measures the information redundancy of the transmitted vector.

Upon transmission of the codeword $\boldsymbol{t}$ via a noisy channel, taken here to be a BSC, the vector $\boldsymbol{r} = \boldsymbol{t} + \boldsymbol{n}^0 \pmod 2$ is received, where $\boldsymbol{n}^0 \in \{0,1\}^N$ is the true channel noise. The statistics of the BSC is fully determined by the flip rate $p \in [0,1]$:

$$P(n_i^0) = (1-p)\,\delta_{n_i^0,0} + p\,\delta_{n_i^0,1} \tag{1}$$

Decoding is carried out by multiplying $\boldsymbol{r}$ by A to produce the syndrome vector $\boldsymbol{z} = A\boldsymbol{r} = A\boldsymbol{n}^0$, since $AG^T = 0$ by construction. In order to reconstruct the original message $\boldsymbol{s}$, one has to obtain an estimate $\boldsymbol{n}$ for the true noise $\boldsymbol{n}^0$. First we select the parity check set of A and $\boldsymbol{n}^0$, i.e. all $\boldsymbol{n}$ that satisfy the parity check equations: $\mathcal{I}_{pc}(A, \boldsymbol{n}^0) \equiv \{\boldsymbol{n} \mid A\boldsymbol{n} = A\boldsymbol{n}^0\}$. Since all operations are performed in modulo 2 arithmetic, $\mathcal{I}_{pc}(A, \boldsymbol{n}^0)$ typically contains $\exp[NR\ln(2)]$ candidates for the true noise vector $\boldsymbol{n}^0$.

It was shown (see e.g. [2, 6, 8] for technical details) that this problem can be cast into a statistical mechanics formulation, by replacing the field $(\{0,1\}, +\mathrm{mod}(2))$ by $(\{1,-1\}, \times)$, and by adapting the parity checks correspondingly. From the parity check matrix A we construct the binary tensor $\mathcal{A} = \{\mathcal{A}_{\langle i_1 \cdots i_k \rangle}, 1 \le i_1 < i_2 \cdots < i_k \le N\}$, where $\mathcal{A}_{\langle i_1 \cdots i_k \rangle} = 1$ if A has a row in which the elements $\{i_c, c = 1, \ldots, k\}$ are all 1 (i.e. when the bits $\langle i_1 \cdots i_k \rangle$ are involved in the same parity check), and 0 otherwise. The fact that each bit $i_1 = 1, \ldots, N$ is involved in exactly j parity checks is then expressed by $\sum_{i_2 < \cdots < i_k} \mathcal{A}_{\langle i_1 \cdots i_k \rangle} = j, \quad \forall \; i_1 = 1, \ldots, N$ and the parity check equations become $\prod_{c=1}^k n_{i_c} = \prod_{c=1}^k n_{i_c}^0, \quad \forall \mathcal{A}_{\langle i_1 \cdots i_k \rangle} = 1$.

Decoding now consists in selecting an $\boldsymbol{n}$ from $\mathcal{I}_{pc}(\mathcal{A}, \boldsymbol{n}^0)$, on the basis of its noise statistics, which are fully described by its *magnetisation* $m(\boldsymbol{n}) = 1/N \sum_i n_i$ (corresponding to the weight in the information theory literature). Note that the number of flipped bits in a candidate noise vector $\boldsymbol{n}$ is given by $N(1 - m(\boldsymbol{n}))/2$. Therefore, we introduce a Hamiltonian or cost function for each noise candidate that is negatively proportional to its magnetisation:

$$H(\boldsymbol{n}) = -F \sum_{i=1}^N n_i = -FNm(\boldsymbol{n}) \tag{2}$$

where we take $F = \frac{1}{2}\log\frac{1-p}{p}$, such that up to normalisation $\exp(-H(\boldsymbol{n}))$ yields the correct prior for candidate noise vectors generated by the BSC [14]. Then, a vector $\boldsymbol{n}$ from $\mathcal{I}_{pc}(\mathcal{A}, \boldsymbol{n}^0)$ with the highest magnetisation (lowest weight) is selected as a solution; this corresponds to Maximum A Posteriori (MAP) decoding.

We are now interested in the probability that other candidate noise vectors are selected from the parity check set $\mathcal{I}_{pc}(\mathcal{A}, \boldsymbol{n}^0)$, other than the correct (i.e. true) noise vector $\boldsymbol{n}^0$, for any given combination $\{\boldsymbol{n}^0, \mathcal{A}\}$; this is termed the *block error probability*. In order to calculate this probability, we introduce an indicator function:

$$\Delta(\boldsymbol{n}^0, \mathcal{A}) = \lim_{\beta_{1,2} \to \infty} \lim_{\lambda_{1,2} \to \pm\lambda} \left[Z_1^{\lambda_1}(\boldsymbol{n}^0, \mathcal{A}; \beta_1) \, Z_2^{\lambda_2}(\boldsymbol{n}^0, \mathcal{A}; \beta_2) \right]\Big|_{\beta_1 = \beta_2 = \beta} \tag{3}$$

where

$$Z_1(\boldsymbol{n}^0, \mathcal{A}; \beta_1) = \sum_{\boldsymbol{n} \in \mathcal{I}_{pc}(\boldsymbol{n}^0, \mathcal{A}) \setminus \boldsymbol{n}^0} e^{-\beta_1 H(\boldsymbol{n})}, \qquad Z_2(\boldsymbol{n}^0, \mathcal{A}; \beta_2) = \sum_{\boldsymbol{n} \in \mathcal{I}_{pc}(\boldsymbol{n}^0, \mathcal{A})} e^{-\beta_2 H(\boldsymbol{n})}. \quad (4)$$

The partition functions $Z_1(\boldsymbol{n}^0, \boldsymbol{A}; \beta_1)$ and $Z_2(\boldsymbol{n}^0, \boldsymbol{A}; \beta_2)$ differ only in the exclusion of $\boldsymbol{n}^0$ from Z_1. If the true noise $\boldsymbol{n}^0$ has the highest magnetisation of all candidates in the parity check set (decoding success), the Boltzmann factor $\exp[-\beta H(\boldsymbol{n}^0)]$ will dominate the sum over states in Z_2 in the limit $\beta \to \infty$, and $\Delta(\boldsymbol{n}^0, \mathcal{A}) = 0$. Alternatively, if some other vector $\boldsymbol{n} \neq \boldsymbol{n}^0$ has the highest magnetisation of all candidates in the parity check set (decoding failure), its Boltzmann factor will dominate both Z_1 and Z_2 and $\Delta(\boldsymbol{n}^0, \mathcal{A}) = 1$. Note that the separate temperatures β_1 and β_2, which are put to be equal to β in the end, and the powers $\lambda_{1,2}$ which are taken to be $\pm\lambda$ in the end, have been introduced in order to allow us to determine whether obtained solutions are physical or not. The power $\lambda \geq 0$ have been introduced to restrict the indicator function results to $0/1$. In principle, this can be done by taking the limit $\lambda \to 0$; however, in section 3, we show that finite $0 < \lambda < 1$ values will be used due to various constraints.

To derive the *average error exponent*, we take the logarithm of the indicator function average with respect to all possible realisations of true noise vectors $\boldsymbol{n}^0$, and the ensemble of regular (k, j) codes $\mathcal{A}$:

$$Q = \lim_{N \to \infty} \frac{1}{N} \log \left\langle \left\langle \Delta(\boldsymbol{n}^0, \mathcal{A}) \right\rangle_{\boldsymbol{n}^0} \right\rangle_{\mathcal{A}} \quad (5)$$

where

$$\langle f(\boldsymbol{n}^0) \rangle_{\boldsymbol{n}^0} = \frac{1}{(2 \cosh F)^N} \sum_{\boldsymbol{n}^0} \exp(F \sum_i n_i^0) \, f(\boldsymbol{n}^0) \quad (6)$$

and

$$\langle f(\mathcal{A}) \rangle_{\mathcal{A}} = \frac{\sum_{\mathcal{A}} \prod_{i_1=1}^N \delta[\sum_{i_2 < \cdots < i_k} \mathcal{A}_{\langle i_1 \cdots i_k \rangle} - j] \, f(\mathcal{A})}{\sum_{\mathcal{A}} \prod_{i_1=1}^N \delta[\sum_{i_2 < \cdots < i_k} \mathcal{A}_{\langle i_1 \cdots i_k \rangle} - j]}. \quad (7)$$

To obtain an expression for the *reliability exponent* one carries out a similar calculation with one main difference: prior to averaging the indicator function over the ensemble of regular (k, j) codes $\mathcal{A}$, one takes the averaged expression with respect to realisations of true noise vectors $\boldsymbol{n}^0$ to a power r which favours code constructions with a low average error probability (i.e., $r < 1$). The logarithm of the expression averaged over the ensemble of codes $\mathcal{A}$ is then divided by r to remove the exponent. The expression calculated is:

$$Q_r = \lim_{N \to \infty} \frac{1}{Nr} \log \left\langle \left[\langle \Delta(\boldsymbol{n}^0, \mathcal{A}) \rangle_{\boldsymbol{n}^0} \right]^r \right\rangle_{\mathcal{A}} \quad (8)$$

Since there are only discrete degrees of freedom, physically meaningful solutions must have a non-negative entropy, requiring the disorder-averaged entropies of the two partition functions (4) to be non-negative. Note that due to the order of taking the logarithm vs the various averages, expressions (5) and (8) are not equivalent to a (quenched) disorder-averaged free energy. Using general principles one can show that for

general values of $\beta_{1,2}$ and $\lambda_{1,2}$, the disordered-averaged entropies (with averages taken over the joint distribution of code-constructions $\{\mathcal{A}\}$, true- and candidate-noise $\{\boldsymbol{n}^0, \boldsymbol{n}\}$ as suggested by (5) and (8)) are given, for both calculations (5) and (8), by

$$\langle S_x \rangle = \frac{\partial Q_r}{\partial \lambda_x} - \frac{\beta_x}{\lambda_x} \frac{\partial Q_r}{\partial \beta_x} \geq 0, \qquad x = 1, 2 \tag{9}$$

which have to be positive.

3. Average error exponent - general solution

Using standard statistical physics methods such as in [14], we perform the gauge transformation $n_i \to n_i n_i^0$, and the averages over true noise (6) and code constructions (7). In the case of $r \neq 1$, each quantity carries two indices (a replica index and another index coming from the power r); however, the two indices factorise *unless an explicit, more complex, symmetry breaking structure is introduced.* Here, we do not assume a more complex structure that entangles the two types of indices; we also assume the simplest replica symmetric scheme [15] to arrive at the following expression for the average error exponent ($r = 1$), and for the reliability exponent (optimised r):

$$Q_r(\beta_1, \beta_2, \lambda_1, \lambda_2) = \frac{1}{r} \mathrm{Extr}_{\pi, \hat{\pi}} \left[\frac{j}{k} \log\ I_1[\pi] - j \log\ I_2[\pi, \hat{\pi}] + \log\ I_3[\hat{\pi}] \right] \tag{10}$$

where

$$I_1 = \int \prod_{c=1}^{k} \{d\pi(x_c, y_c)\} \left(\frac{1 + \prod_{c=1}^{k} x_c}{2} \right)^{r\lambda_1} \left(\frac{1 + \prod_{c=1}^{k} y_c}{2} \right)^{r\lambda_2} \tag{11}$$

$$I_2 = \int \{d\pi(x, y)\ d\hat{\pi}(\hat{x}, \hat{y})\} \left(\frac{1 + x\hat{x}}{2} \right)^{r\lambda_1} \left(\frac{1 + y\hat{y}}{2} \right)^{r\lambda_2} \tag{12}$$

$$I_3 = \int \prod_{c=1}^{j} \{d\hat{\pi}(\hat{x}_c, \hat{y}_c)\} \left\langle \left[\sum_{u=\pm 1} e^{\beta_1 F n^0 u} \prod_{c=1}^{j} \left(\frac{1 + u\hat{x}_c}{2} \right) \right]^{\lambda_1} \right.$$
$$\left. \times \left[\sum_{v=\pm 1} e^{\beta_2 F n^0 v} \prod_{c=1}^{j} \left(\frac{1 + v\hat{y}_c}{2} \right) \right]^{\lambda_2} \right\rangle_{n^0}^{r} \tag{13}$$

where we have used the short-hand notation $df(x, y) = dxdy\ f(x, y)$. For $r = 1$, functional extremisation of (10) with respect to the densities $\pi(x, y)$ and $\hat{\pi}(\hat{x}, \hat{y})$ results in a closed set of equations (reminiscent of 'density evolution' equations [1]):

$$\hat{\pi}(\hat{x}, \hat{y}) = \int \prod_{c=1}^{k-1} \{d\pi(x_c, y_c)\} \delta \left[\hat{x} - \prod_{c=1}^{k-1} x_c \right] \delta \left[\hat{y} - \prod_{c=1}^{k-1} y_c \right] \tag{14}$$

$$\pi(x, y) = \frac{\left\langle \left\langle \delta \left[x - \frac{D_-(\hat{\boldsymbol{x}}; \beta_1)}{D_+(\hat{\boldsymbol{x}}; \beta_1)} \right] \delta \left[y - \frac{D_-(\hat{\boldsymbol{y}}; \beta_2)}{D_+(\hat{\boldsymbol{y}}; \beta_2)} \right] \right\rangle \right\rangle'}{\left\langle \left\langle 1 \right\rangle \right\rangle'}. \tag{15}$$

where

$$\left\langle\!\left\langle \cdot \right\rangle\!\right\rangle' \equiv \int \prod_{c=1}^{j-1} \{d\hat{\pi}(\hat{x}_c, \hat{y}_c)\} \left\langle D_+^{\lambda_1}(\hat{\boldsymbol{x}}; \beta_1) D_+^{\lambda_2}(\hat{\boldsymbol{y}}; \beta_2) \cdot \right\rangle_{n^0}, \tag{16}$$

$$D_{\pm}(\boldsymbol{z}; \beta) \equiv [e^{\beta F n^0} \prod_{c=1}^{j-1}(1 + z_c)] \pm [e^{-\beta F n^0} \prod_{c=1}^{j-1}(1 - z_c)]. \tag{17}$$

For given $(\beta_1, \beta_2, \lambda_1, \lambda_2)$ in general, solutions to (14) and (15) can only be obtained numerically. Inserting these solutions into (10) we then obtain $Q(\beta_1, \beta_2, \lambda_1, \lambda_2)$, which becomes the *average error exponent* for $\lambda_1 = -\lambda_2 = \lambda > 0$, and for $\beta_1 = \beta_2 = \beta \to \infty$.

We must recall, however, that physically meaningful solutions must satisfy the conditions (9) stating that the entropies related to the full and the restricted partition sums are non-negative.

We restrict ourselves to regions below the thermodynamic transition where the average case is dominated by the ferromagnetic solution, such that we can fix the system described by Z_2 in (4) to the ferromagnetic solution. This dominance is guaranteed if the following constraint is satisfied

$$\left.\frac{\partial Q}{\partial \beta}\right|_{\lambda_1 = -\lambda_2 = \lambda} \leq 0 . \tag{18}$$

It turns out that for given $\lambda > 0$, the largest value of β for which (18) is satisfied is given by the simple expression $\beta = 1/(1 + \lambda)$. Hence, in order to maximise β, we must look for the smallest value λ_* that satisfies the conditions on the non-negativity of the entropies (9). Unfortunately, in general this value λ_* can only be obtained numerically. The value obtained for the average error exponent by this analysis is then given by $Q(1/(1 + \lambda_*), 1/(1 + \lambda_*), \lambda_*, -\lambda_*)$ from (10).

In figure 1 we present the obtained average error exponent as a function of the flip rate for $(k, j) = (4, 3)$, $(R = 1/3)$ and $(k, j) = (6, 3)$ $(R = 1/2)$ codes. We observe that the error exponent indeed converges to zero, as it should, when the flip rate approaches its critical value.

Notice the similarity between the equations obtained here and in [6] in spite of the different starting points. It has been shown in [6] that the analysis should be refined in low rate regions by considering a more complex bound. The refined analysis resulted in tight bounds of the error exponent even in the region of low code-rates, similar to those obtained using expurgated exponent methods. In the next section we will show that the selection of 'best codes' through the optimisation of the power r, in calculating the reliability exponent, provides similar results to those obtained in [6].

4. An exactly solvable limit: $k, j \to \infty$

Whereas for finite density codes solutions for the average error exponent are obtained numerically, in the limit of $k, j \to \infty$ (while keeping the rate $R = 1 - j/k$ finite) one obtains two types of analytic solutions to equations (14) and (15), which can be verified
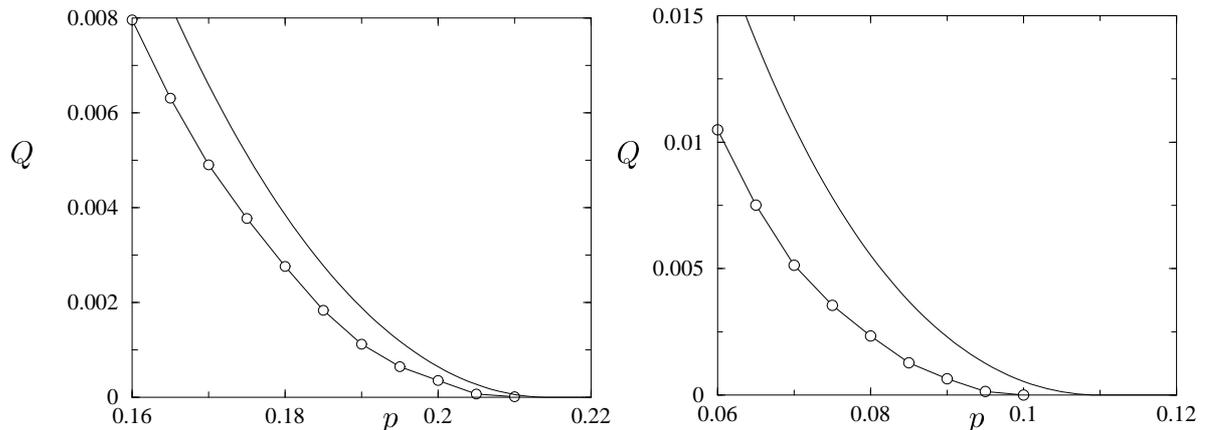
Figure 1. Average error exponent Q as function of the flip rate p for codes of $(k, j) = (4, 3)$ (left picture) and $(k, j) = (6, 3)$ (right picture). Lines with markers correspond to the finite (k, j) cases. For comparison we also present (thick solid lines) the value of the average error exponent in the case of $k, j \to \infty$ with $R = 1/4$ (left) and $R = 1/2$ (right) as described in the analysis of section 4. Note that the transition from type I to type II solution occurs at small p values outside the range of this figure.

by substitution. Moreover, in this limit one also obtains solutions in the reliability exponent calculation (8), which are generally difficult to obtain for finite k and j values.

4.1. Average error exponent

Solutions obtained in the average error exponent calculation take the following form:
Type I:

$$\pi(x, y) = \frac{1}{2} \left[\delta(x - 1) + \delta(x + 1) \right] \delta(y - 1)$$

$$\hat{\pi}(\hat{x}, \hat{y}) = \frac{1}{2} \left[\delta(\hat{x} - 1) + \delta(\hat{x} + 1) \right] \delta(\hat{y} - 1) \tag{19}$$

Type II:

$$\pi(x, y) = \delta(y - 1) \left[G_+(F(1 + \beta_2 \lambda_1)) \, \delta(x - \tanh(\beta_1 F)) \right.$$

$$\left. + \, G_-(F(1 + \beta_2 \lambda_2)) \, \delta(x + \tanh(\beta_1 F)) \right]$$

$$\hat{\pi}(\hat{x}, \hat{y}) = \delta(\hat{y} - 1) \, \delta(\hat{x}) \tag{20}$$

with $G_{\pm}(x) = \frac{1}{2}[1 \pm \tanh(x)]$.

Taking $\beta_1 = \beta_2 = \beta$ and $\lambda_1 = -\lambda_2 = \lambda$, the average error exponent as obtained from the type I solution is given by

$$Q_I = -\frac{j}{k} \log 2 - \log \cosh F + \log \cosh(\beta F \lambda) + \log 2 \cosh(F - \beta F \lambda) \,. \tag{21}$$

We find that the entropies (9) are always identically zero, and that the constraint (18)

requires that $\beta = 1/2$, such that $\lambda = 1$ and

$$Q_I = -\frac{j}{k} \log 2 - \log[e^F + e^{-F}] + \log[e^F + e^{-F} + 2] \tag{22}$$

which is exactly the Bhattacharyya limit [11].

The average error exponent as obtained from the type II solution is given by

$$Q_{II} = \lambda \left[-\frac{j}{k} \log 2 + \log 2 \cosh[\beta F] \right] + \log[2 \cosh(F - \beta F \lambda)] - \log 2 \cosh F \tag{23}$$

The condition on the entropy $\langle S_2 \rangle \geq 0$ is satisfied for all $\beta > 0$, whereas the condition $\langle S_1 \rangle \geq 0$ is violated below the critical (freezing) temperature $1/\beta^*$ obtained from

$$-\frac{j}{k} \log 2 - \beta^* F \tanh[\beta^* F] + \log 2 \cosh[\beta^* F] = 0 \tag{24}$$

This negative entropy is an artifact of the assumption about the symmetry between replicas, and is easily remedied by considering a 'frozen RSB' ansatz [2]. Using this ansatz and taking into account condition (18), the (frozen) average error exponent obtained from the type II solution, is finally given by

$$Q_{II}^{fr} = F \tanh[\beta^* F] + \frac{j}{k} \log 2 - \log 2 \cosh F \tag{25}$$

What remains is to determine whether the type I or type II solution is physically dominant, by using Q as a generating function for calculating the related free energies (through its derivative with respect to λ). Results for the case of $k, j \to \infty$ are presented in figure 2 for $p = 0.01$ and $p = 0.05$.

4.2. Reliability exponent

To obtain the reliability exponent we take equations (10)-(13) and optimise with respect to r. Deriving a general set of equations similar to (14,15), that can be solved iteratively, is difficult in this case. However, in the limit $k, j \to \infty$, we observe that we can restrict the possible solutions of $\hat{\pi}(\hat{x}, \hat{y})$ to two different types:
Type I:

$$\hat{\pi}(\hat{x}, \hat{y}) = \frac{1}{2} [\delta(\hat{x} - 1) + \delta(\hat{x} + 1)] \delta(\hat{y} - 1) \tag{26}$$

Type II:

$$\hat{\pi}(\hat{x}, \hat{y}) = \delta(\hat{x}) \delta(\hat{y} - 1) \tag{27}$$

In this case, knowledge of the solution for $\hat{\pi}(\hat{x}, \hat{y})$ is sufficient for calculation the reliability exponent (10). Furthermore, the expression obtained from the type II solution turns out to be identical to that of the average error exponent (25).

On the other hand, the reliability exponent obtained from the type I solution is somewhat different, and takes the form:

$$Q_{r\ I} = -\frac{1}{r} \frac{j}{k} \log 2 - \log[\cosh(F)] + \frac{1}{r} \log[\cosh^r(F) + \cosh^r((2\beta\lambda - 1)F)] . \tag{28}$$

Given the relation (18) and $\beta = 1/(1 + \lambda)$, one obtains $\lambda = 1$, $\beta = 1/2$, and the expression reduces to

$$Q_{r\ I} = -\frac{1}{r}\frac{j}{k}\log 2 - \log[\cosh(F)] + \frac{1}{r}\log[\cosh^r(F) + 1] \ . \qquad (29)$$

Optimising the expression with respect to r, one obtains a similar expression to the expurgated exponent result [10]

$$E_{ex}(r, R) = \max_r \left\{ \ln 2 \cosh F - \frac{1}{r}\ln\left[(2\cosh F)^r + 1\right] + \frac{1}{r}(1 - R)\ln 2 \right\} \ , (30)$$

which is also identical to the result obtained for the average bound of the reliability exponent in [6].

The reliability exponent is therefore identical to the average error exponent except for very low R values as shown in figure 2 for $p = 0.01$ and $p = 0.05$ (marked by a dotted line in the two cases considered).
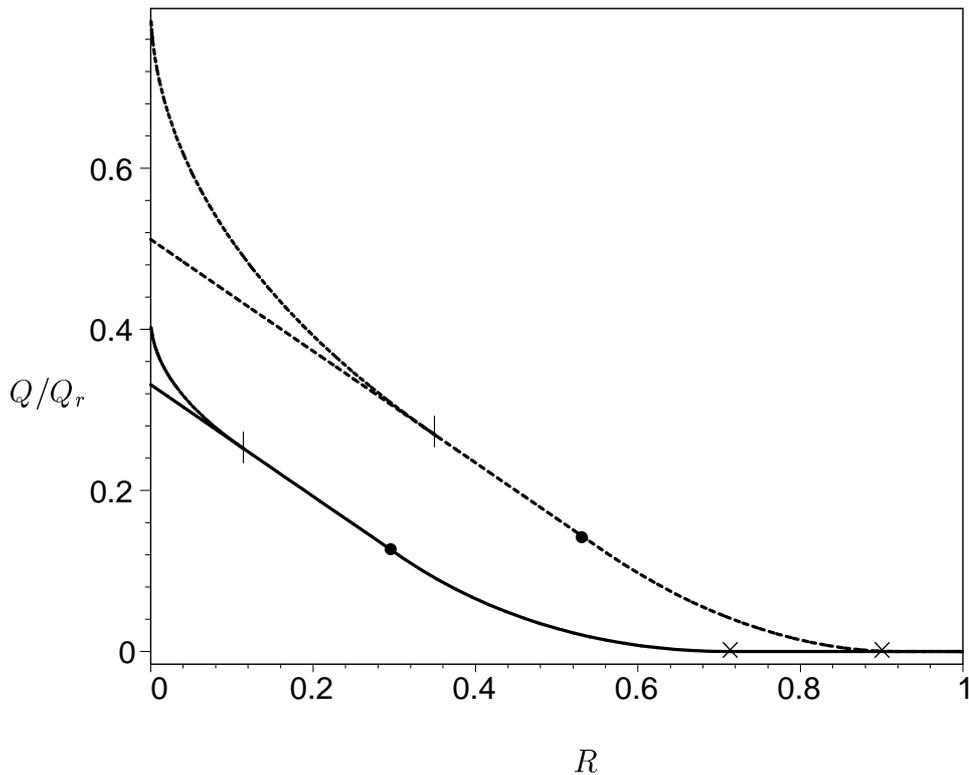

Figure 2. Reliability and average exponents as function of the code rate R for regular $k, j \to \infty$ Gallager codes for which analytical expressions can be derived; see (22) and (25) (dashed: $p = 0.01$ and solid: $p = 0.05$). The reliability exponent is identical to the average error exponent except for very low R values where it is represented by the curved solutions above the linear average exponent results marked by a dashed and solid lines, respectively. The transition point is marked by a vertical line. The transition between solutions of type I and II is marked by ● and the critical transition point by a ×.

5. MN codes

In this section, we extend our treatment of the average error and reliability exponent to regular MN codes [12], a variant of LDPC codes.

A regular MN code is defined by the binary $N \times (N + K)$ matrix $A = [C_s | C_n]$, concatenating two sparse matrices with the $N \times N$ matrix C_n assumed invertible. The $N \times K$ matrix C_s has k non-zero elements per row and j per column while C_n has t non-zero elements per row and per column. The code rate is given by $R \equiv K/N = k/j$. The encoding scheme consists of generating a codeword $\boldsymbol{t}^0 \in \{0, 1\}^N$ from an (unbiased) message vector $\boldsymbol{s}^0 \in \{0, 1\}^K$ via $\boldsymbol{t}^0 = (C_n^{-1} C_s)\boldsymbol{s}^0$. Upon sending $\boldsymbol{t}^0$ through the noisy channel the vector $\boldsymbol{r} = \boldsymbol{t}^0 + \boldsymbol{n}^0$ is received, where $\boldsymbol{n}^0$ is the true channel noise (1).

Decoding is carried out by multiplying the received vector $\boldsymbol{r}$ by C_n to produce the syndrome vector $\boldsymbol{z} = C_s \boldsymbol{s}^0 + C_n \boldsymbol{n}^0$. In order to reconstruct the original message, one selects the best estimate $(\boldsymbol{s}, \boldsymbol{n})$, for the true $(\boldsymbol{s}^0, \boldsymbol{n}^0)$ from the parity check set $\mathcal{I}_{pc} = \{(\boldsymbol{s}, \boldsymbol{n}) \, | \, C_s \boldsymbol{s} + C_n \boldsymbol{n} = \boldsymbol{z}\}$, on the basis of the message/noise statistics. Note that since we take the message vector $\boldsymbol{s}^0$ to be unbiased, the selection will only be based on the noise statistics.

Since most calculation steps are completely analogous (although lengthier) to those of Gallager codes, we only state the final general expression for MN codes:

$$Q_r(\beta_1, \beta_2, \lambda_1, \lambda_2) = \frac{1}{r} \text{Extr}_{\pi, \hat{\pi}, \rho, \hat{\rho}} \Bigg\{ \tag{31}$$

$$- k \log \int \{d\pi(x, y) \, d\hat{\pi}(\hat{x}, \hat{y})\} \left(\frac{1 + x\hat{x}}{2}\right)^{r\lambda_1} \left(\frac{1 + y\hat{y}}{2}\right)^{r\lambda_2}$$

$$- t \log \int \{d\rho(x, y) \, d\hat{\rho}(\hat{x}, \hat{y})\} \left(\frac{1 + x\hat{x}}{2}\right)^{r\lambda_1} \left(\frac{1 + y\hat{y}}{2}\right)^{r\lambda_2}$$

$$+ \log \int \prod_{i=1}^{k} \{d\pi(x_i, y_i)\} \prod_{l=1}^{t} \{d\rho(u_l, v_l)\} \left(\frac{1 + \prod_i x_i \prod_l u_l}{2}\right)^{r\lambda_1} \left(\frac{1 + \prod_i y_i \prod_l v_l}{2}\right)^{r\lambda_2}$$

$$+ \frac{k}{j} \log \int \prod_{c=1}^{j} \{d\hat{\pi}(\hat{x}_c, \hat{y}_c)\} \left[\sum_{\sigma=\pm} \prod_{c=1}^{j} \left(\frac{1 + \sigma \hat{x}_c}{2}\right)\right]^{r\lambda_1} \left[\sum_{\sigma=\pm} \prod_{c=1}^{j} \left(\frac{1 + \sigma \hat{y}_c}{2}\right)\right]^{r\lambda_2}$$

$$+ \log \int \prod_{l=1}^{t} \{d\hat{\rho}(\hat{x}_l, \hat{y}_l)\} \left\langle \left[\sum_{\tau=\pm} e^{\beta_1 \tau F n^0} \prod_{l=1}^{t} \left(\frac{1 + \tau \hat{x}_l}{2}\right)\right]^{\lambda_1} \right.$$

$$\left. \times \left[\sum_{\tau=\pm} e^{\beta_2 \tau F n^0} \prod_{l=1}^{t} \left(\frac{1 + \tau \hat{y}_l}{2}\right)\right]^{\lambda_2} \right\rangle_{n^0}^{r} \Bigg\}$$

with the short-hand notation $df(x, y) = dxdy \, f(x, y)$. As for Gallager codes, for $\beta_1 = \beta_2 = \beta$, and $\lambda_1 = -\lambda_2 = \lambda$, Q_r becomes the *average error exponent* for $r = 1$, while for optimised r it becomes the *reliability exponent*. Furthermore, the conditions (9) and (18) must always be satisfied.

Similarly to the case of Gallager codes one can derive a set of functional equations (reminiscent of 'density evolution' equations [1]) for $\pi, \hat{\pi}, \rho$ and $\hat{\rho}$.
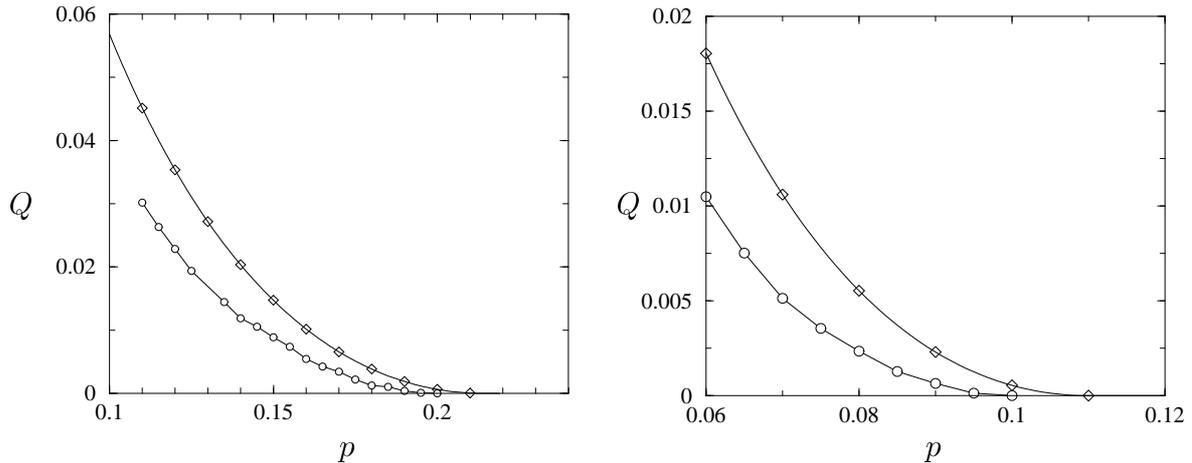
Figure 3. Average error exponent Q as function of the flip rate p obtained for Gallager (10) and MN codes (31). Left: results for MN codes of $R = 1/4$. Lower line and circles: $(k, j, t) = (1, 4, 2)$. Upper line and diamonds: $(2, 8, 3)$. Thick upper line: $k, j \rightarrow \infty$. Right: results for Gallager and MN codes of $R = 1/2$. Lower line and circles: Gallager, $(k, j) = (6, 3)$. Upper line and diamonds: MN, $(k, j, t) = (3, 6, 3)$. Upper line: the analytical solution of $k, j \rightarrow \infty$ obtained via (22,25).

5.1. Average error exponent - finite k, j and t

The average error exponent can be calculated numerically for finite k, j and t values; the average error exponent Q as function of the flip rate p is shown in figure 3.

On the left, we show results for MN codes of fixed rate $R = 1/4$ with three different sets of parameters $(k, j, t) = (1, 4, 2)$ (circles), $(2, 8, 3)$ (diamonds) and $k, j \rightarrow \infty$ (upper line). It is interesting to notice that average exponents for either $k > 2$ or $t > 2$ values coincide with that of the infinite connectivity case (which can be obtained analytically). This complements other interesting properties of MN codes, to do with their critical flip rate values, that have been obtained previously, distinguishing them from Gallager LDPC codes [2, 4, 5].

On the right, we see a comparison between average error exponents of Gallager and MN codes ($R = 1/2$). The Gallager code $(k, j) = (6, 3)$ average error exponent (circles) is significantly below the random code $k, j \rightarrow \infty$ value (thin upper line) and the equivalent MN code $(k, j, t) = (3, 6, 3)$ result (diamonds).

5.2. Average and reliability error exponents - $k, j, t \rightarrow \infty$

The case of $k, j, t \rightarrow \infty$ is solvable exactly for all transmission rates, and both average and reliability error exponents can be obtained analytically. The solutions obtained as well as the average and reliability error exponents calculated are identical to those of Gallager LDPC codes. Retrospectively, this is not surprising as both codes become random codes in this limit.

6. Discussion

In this paper we suggest a method for direct evaluation of the average and reliability error exponent over the ensemble of LDPC error-correcting codes of given rate and connectivity. An analytical solution has been obtained, for both Gallager and MN codes, using methods of statistical physics, which is in perfect agreement with known results in the limit $k, j(, t) \to \infty$ (with R finite). The results for MN and Gallager codes become identical in this limit as both become random codes.

Average error exponent results obtained by our method for codes of finite (k, j) values cannot be obtained using traditional approaches used in the information theory community. As expected, they seem to be upper bounded by the $k, j \to \infty$ curves, but suggest a profoundly different behaviour for Gallager and MN LDPC codes. Average error exponent results for Gallager codes show a gradually improved performance as the parameters (k, j) increase, until they finally coincide with the $k, j \to \infty$ error exponent result. The results for MN codes becomes identical to the $k, j \to \infty$ error exponent result for all $k > 2$ or $t > 2$. To some extent, this is in agreement with previous results obtained for the critical flip rate of MN codes [2, 4, 5] and is a result of the close-to-random codebook they generate.

An interesting feature of the present study is the similarity of our equations to those obtained in [6] in spite of the different approaches used. An important advantage offered by the current approach is a potential extension to select high performance codes to obtain *reliability exponent* values for LDPC codes of finite connectivity; obtaining such solutions remains a difficult task and is currently under study.

Acknowledgments

Reference

[1] T Richardson, A Shokrollahi and R Urbanke (2001), *IEEE Trans. on Info. Theory* **47** 619-637
[2] R Vicente, D Saad and Y Kabashima (1999) *Phys Rev E* **60** 5352-5366
[3] R Vicente, D Saad and Y Kabashima (2000) *J Phys A* **33** 6527-6542
[4] J van Mourik, D Saad and Y Kabashima (2002) *Phys Rev E* **66** 026705
[5] T Tanaka and D Saad, (2003) submitted in *Phys Rev E*
[6] Y Kabashima, N Sazuka, K Nakamura and D Saad (2001) *Phys Rev E* (2001) **64** 046113
[7] A Montanari and N Sourlas *Eur Phys J B* **18** 107-119
[8] A Montanari *Eur Phys J B* **23** 121-136
[9] S. Franz, M. Leone, A. Montanari and F. Ricci-Tersenghi, *Phys Rev E* **66** 046120
[10] R G Gallager (2001) 'Information theory and reliable communication' Wiley & Sons, NY
[11] A J Viterbi and J K Omura (1979) 'Principles of Digital Communication and Coding', McGraw-Hill Int Ed (Singapore)
[12] D.J.C. MacKay (1999) *IEEE Trans. Inf. Th.* **45** 399

[13] G. Miller and D. Burshtein (2001) *IEEE Trans. Infor. Theory* **47** 2696

[14] H Nishimori (2001) 'Statistical Physics of Spin Glasses and Information Processing', Oxford University Press, UK

[15] K Y M Wong and D Sherrington (1987) *J Phys A* **20** L793-L799