

# Tighter Decoding Reliability Bound for Gallager's Error-Correcting Code

Yoshiyuki Kabashima<sup>1</sup>, Naoya Sazuka<sup>1</sup>, Kazutaka Nakamura<sup>1</sup> and David Saad<sup>2</sup>

<sup>1</sup> *Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology,  
Yokohama 2268502, Japan.*

<sup>2</sup> *The Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK.*

## Abstract

Statistical physics is employed to evaluate the performance of error-correcting codes in the case of finite message length for an ensemble of Gallager's error correcting codes. We follow Gallager's approach of upper-bounding the average decoding error rate, but invoke the replica method to reproduce the tightest general bound to date, and to improve on the most accurate zero-error noise level threshold reported in the literature. The relation between the methods used and those presented in the information theory literature are explored.

Typeset using REVTeX

Many of the problems addressed in the Information Theory (IT) literature show great similarity to those treated in statistical physics. One of the main areas where these links are particularly strong is that of digital communication and coding theory; these links have been recently examined in the area of Low Density Parity Check (LDPC) [12,6] and turbo [8] error-correcting codes. It is only natural to expect that some relations between the analytical methods used in the two disciplines will emerge, and that advances in one could be employed to improve results in the other. In this Letter we focus on such an example. We utilize the replica method of statistical physics to assess the performance of Gallager's error correcting code in the case of finite message length, generalizing an established method in the IT community. The analysis reproduces the tightest general bound to date, but more importantly, it provides exact results to specific code constructions.

Error correcting codes play a vital role in facilitating reliable data transmission, ranging from cellular communication to data storage on magnetic media. In a general scenario, the  $N$  dimensional Boolean message  $\boldsymbol{\xi} \in \{0, 1\}^N$  is encoded to the  $M (> N)$  dimensional Boolean vector  $\boldsymbol{z}_0$ , and transmitted via a noisy channel, which is taken here to be a Binary Symmetric Channel (BSC) characterized by flip probability  $p$  per bit; other transmission channels may also be examined within a similar framework. At the other end of the channel, the corrupted codeword is decoded utilizing the structured codeword redundancy.

The block error rate  $P_E$ , defined as the probability for a decoding error, serves as a performance measure for the success of the coding method. In his seminal work [13], Shannon showed that the error rate can vanish for code rates  $R$  below the channel capacity in the limit  $N, M \rightarrow \infty$ ; in the case of the BSC and unbiased messages  $R = N/M < 1 - H_2(p)$ , where  $H_2(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ . The upper bound, for infinitely long messages, is often termed *Shannon's limit* to the error correcting ability. Evaluating  $P_E$  for practical codes of finite length became one of central topics in IT.

For maximum likelihood (ML) decoding where the most probable message given the possibly corrupted codeword defines the message estimate, it is believed that  $P_E$  of the best code scales as  $\exp[-ME(R)]$ . The non-negative exponent  $E(R)$  is termed *reliability function*

(RF); it becomes positive below the channel capacity defining the sensitivity of the optimal error rate to the message length, complementing Shannon's result.

Unfortunately, assessing the RF directly is generally difficult. Instead, Gallager's powerful method [3] bounds  $E(R)$  from the below utilizing the inequality

$$P_E \leq \underset{\{\mathbf{y}, \mathbf{x}\}}{\text{Tr}} P^{\frac{1}{1+\rho}}(\mathbf{y}, \mathbf{x}) \left( \underset{\{\mathbf{x}' \neq \mathbf{x}\}}{\text{Tr}} P^{\frac{1}{1+\rho}}(\mathbf{y}, \mathbf{x}') \right)^\rho, \quad (1)$$

which holds for any arbitrary ML estimation, inferring a binary vector  $\mathbf{x}$  after observing a vector  $\mathbf{y}$ , and a positive variable  $\rho > 0$ .

The average error rate  $\bar{P}_E$  for a certain ensemble of codes is greater than the ensemble minimum. Therefore, averaging the RHS of Eq.(1) over the ensemble, one obtains an upper-bound to the minimum error rate that scales exponentially with  $M$  for large but finite  $N$  and  $M$ ,  $\exp[-ME_{av}(\rho, R)]$ ; the exponent  $E_{av}(\rho, R)$  serves as a *lower-bound* of  $E(R)$ . One can tighten the lower bound by maximizing  $E_{av}(\rho, R)$  with respect to  $\rho > 0$ .

Evaluating  $E_{av}(\rho, R)$  is also difficult (except for  $\rho \in \mathbb{N}$ ). The strategy used by Gallager [3] is to further upper-bound the RHS of Eq.(1) utilizing Jensen's inequality  $\langle x^\rho \rangle \leq \langle x \rangle^\rho$ , which holds for any  $0 \leq \rho \leq 1$  with respect to the expectation over any arbitrary distribution of a positive variable. The added inequality presumably makes the bound looser. It is therefore surprising that maximizing the exponent with respect to  $\rho \in [0, 1]$  in the ensemble of all random codes having the same rate  $R$ , which results in the *random coding exponent*  $E_r(R)$ , provides an exact evaluation of the RF for high  $R$  values.

However, the bound by  $E_r(R)$  becomes loose once the optimal value of  $\rho$  reaches the upper limit of the interval, i.e.,  $\rho = 1$  (corresponding to Bhattacharyya's bound). It is not clear whether Jensen's inequality or Gallager's inequality (1) is responsible for this breakdown. Moreover, it is unclear how to devise a similar method for deriving bounds for other (non-random) codes, a question of high practical significance.

In this Letter we demonstrate how the methods of statistical physics may be employed to obtain tighter bounds for specific codes. This is carried out by a direct evaluation of  $E_{av}(\rho, R)$  for the ensemble of Gallager error-correcting codes [2]. This (linear) code was

rediscovered only recently [7], showing outstanding performance, competitive to other state-of-the-art techniques. It is characterized by a randomly generated  $(M - N) \times M$  Boolean sparse parity check matrix  $H$ , composed of  $K$  and  $C$  ( $\geq 3$ ) non-zero (unit) elements per row and column, respectively. Encoding the message vector  $\boldsymbol{\xi}$ , is carried out using the  $M \times N$  generating matrix  $G^T$ , satisfying the condition  $HG^T = 0$ , where  $\mathbf{z}_0 = G^T \boldsymbol{\xi} \pmod{2}$ . The  $M$  bit codeword  $\mathbf{z}_0$  is transmitted via a noisy channel, BSC in the current analysis; the corrupted vector  $\mathbf{z} = \mathbf{z}_0 + \boldsymbol{\zeta} \pmod{2}$  is received at the other end, where  $\boldsymbol{\zeta} \in \{0, 1\}^M$  represents a noise vector with an independent probability  $p$  per bit of having a value 1. Decoding is carried out by multiplying  $\mathbf{z}$  by the parity check matrix  $H$ , to obtain the syndrome vector  $\mathbf{J} = H\mathbf{z} = H(G^T \boldsymbol{\xi} + \boldsymbol{\zeta}) = H\boldsymbol{\zeta} \pmod{2}$ , and to find the most probable solution to the parity check equation  $H\mathbf{n} = \mathbf{J} \pmod{2}$ , for estimating the true noise vector  $\boldsymbol{\zeta}$ . One retrieves the original message using the equation  $G^T \mathbf{S} = \mathbf{z} - \mathbf{n} \pmod{2}$ ;  $\mathbf{S}$  to estimate of the original message.

To facilitate the analysis we map the Boolean  $(0, 1)$  variables onto the binary  $(\pm 1)$  representation. The binary vectors  $\mathbf{n}$  and  $\mathbf{J}$ , represent the noise estimate and syndrome vectors respectively; the latter is generated by taking products of the relevant noise bits  $J_\mu = \zeta_{i_{1\mu}} \dots \zeta_{i_{K\mu}}$ , where the indices  $i_{1\mu}, \dots, i_{K\mu}$  correspond to the nonzero elements in row  $\mu$  of the parity check matrix  $H$ .

The similarity between error-correcting codes and physical systems was first pointed out by Surlas [12], mapping a simple Boolean code onto Ising spin models with multi-spin interactions. We recently extended his work to more practical parity check codes [6]. We employ a similar formulation using the Hamiltonian

$$\mathcal{H}(\mathbf{n}; \mathbf{J}) = \gamma \sum_{\mathcal{G}} D_{\mathcal{G}} \delta \left( J_{\mathcal{G}}; - \prod_{i \in \mathcal{G}} n_i \right) - F \sum_{i=1}^M n_i, \quad (2)$$

to evaluate the joint probability for  $\mathbf{J}$  and  $\mathbf{n}$

$$P(\mathbf{J}, \mathbf{n}) = \lim_{\gamma \rightarrow \infty} \frac{\exp[-\beta \mathcal{H}(\mathbf{n}; \mathbf{J})]}{(2 \cosh F)^M}. \quad (3)$$

Here,  $\mathcal{G} \equiv \langle i_1, \dots, i_K \rangle$  runs over all combinations of  $K$  indices out of  $M$ ;  $J_{\mathcal{G}} \equiv \prod_{i \in \mathcal{G}} \zeta_i$  and the sparse tensor  $D_{\mathcal{G}}$  becomes non-zero (unit) only when all indices in  $\mathcal{G}$  correspond to non-zero

(unit) elements in a certain row of the parity check matrix  $H$ . Taking  $\gamma \rightarrow \infty$  enforces the parity check equation. The additive field  $F = (1/2) \ln [(1-p)/p]$  corresponds to the true prior probability in the Bayesian framework, reflecting the flip rate  $p$ . The inverse temperature  $\beta$  is introduced to emphasize the link with the statistical mechanics formulation and is generally fixed to  $\beta = 1$  unless specified otherwise.

One can then use (3) to evaluate  $\bar{P}_E$  from (1) by calculating the bound without invoking Jensen's inequality. The first part of the Hamiltonian (2) is invariant under gauge transformations of the form  $n_i \rightarrow n_i \zeta_i$ , and  $J_{\mathcal{G}} \rightarrow J_{\mathcal{G}} \prod_{i \in \mathcal{G}} \zeta_i = 1$ , which decouple the correlation between the dynamical vector  $\mathbf{n}$  and the true noise  $\boldsymbol{\zeta}$ . Rewriting the Hamiltonian one obtains a similar expression to Eq. (2) apart from the last term on the right which become  $F \sum_i \zeta_i n_i$ .

Quenched averages over the ensemble of codes is carried out with respect to the current random selection of the sparse tensor  $D$  and the noise vector, which eventually results in a similar procedure to the replica method in statistical mechanics. This gives rise to a set of order parameters  $q_{\alpha, \beta, \dots, \gamma} = \frac{1}{M} \sum_{i=1}^M Z_i n_i^\alpha n_i^\beta \dots n_i^\gamma$ , where  $\alpha, \beta \dots$  represent replica indices, and the variable  $Z_i$  comes from enforcing the restriction of  $C$  and  $L$  connections per index respectively as in [6]. This interesting similarity between Gallager's method and the replica method has been pointed out by Iba in [4].

To proceed further one has to make an assumption about the order parameter symmetry. As a first approximation we assume replica symmetry (RS) in the following order parameters and the related conjugate variables

$$q_{\alpha, \beta, \dots, \gamma} = q \int dx \pi(x) x^l, \quad \hat{q}_{\alpha, \beta, \dots, \gamma} = \hat{q} \int d\hat{x} \hat{\pi}(\hat{x}) \hat{x}^l, \quad (4)$$

where  $l$  is the number of replica indices,  $q$  and  $\hat{q}$  are normalization variables ( $\pi(x)$  and  $\hat{\pi}(\hat{x})$  are probability distributions). Unspecified integrals are over the range  $[-1, +1]$ .

Originally, the summation  $\text{Tr}_{\{\mathbf{n} \neq \boldsymbol{\zeta}\}}(\cdot)$  excludes the case of  $\mathbf{n} \neq \boldsymbol{\zeta}$ ; however, it can be shown that in the limit of large  $M$  this becomes identical to the full summation in the non-ferromagnetic phase, where  $\pi(x) \neq \delta(x - 1)$  and  $\hat{\pi}(\hat{x}) \neq \delta(\hat{x} - 1)$ . Then, one obtains

the expression

$$E_{av}(\rho, R) = -\frac{1}{M} \ln \left[ \left\langle \left\langle \text{Tr}_{\{\mathbf{J}, \boldsymbol{\zeta}\}} P^{\frac{1}{1+\rho}}(\mathbf{J}, \boldsymbol{\zeta}) \left( \text{Tr}_{\{\mathbf{n} \neq \boldsymbol{\zeta}\}} P^{\frac{1}{1+\rho}}(\mathbf{J}, \mathbf{n}) \right)^\rho \right\rangle_D \right\rangle \right] \\ = \ln(2 \cosh F) - \ln \left( 2 \cosh \left( \frac{F}{1+\rho} \right) \right) - \frac{1}{M} \ln \left\langle \mathcal{Z}_{\text{NF}}^\rho \left( \boldsymbol{\zeta}, D; \frac{F}{1+\rho} \right) \right\rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}, D}, \quad (5)$$

where  $\mathcal{Z}_{\text{NF}}(\boldsymbol{\zeta}, D; \frac{F}{1+\rho})$  denotes the partition function  $\text{Tr}_{\mathbf{n}} \lim_{\gamma \rightarrow \infty} \exp[-\beta \mathcal{H}]$  in the non-ferromagnetic phase for a system with an effective additive field  $F/(1+\rho)$ . Averages  $\langle \cdot \rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}, D}$  are over the distribution  $P(\boldsymbol{\zeta}; \frac{F}{1+\rho}) = \exp[\frac{F}{1+\rho} \sum_{i=1}^M \zeta_i] / (2 \cosh(\frac{F}{1+\rho}))^M$  and the uniform distribution of  $D$ . Extremizing  $\langle \mathcal{Z}_{\text{NF}}^\rho(\boldsymbol{\zeta}, D; \frac{F}{1+\rho}) \rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}, D}$  with respect to the order parameters  $q$ ,  $\hat{q}$ ,  $\pi(\cdot)$  and  $\hat{\pi}(\cdot)$ , under the replica symmetry ansatz (4), one obtains for the final term in (5)

$$\frac{1}{M} \ln \left\langle \mathcal{Z}_{\text{NF}}^\rho \left( \boldsymbol{\zeta}, D; \frac{F}{1+\rho} \right) \right\rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}, D} = \underset{\{q, \hat{q}, \pi(\cdot), \hat{\pi}(\cdot)\}}{\text{Ext}^*} \left\{ \frac{C q^K}{K} \int \prod_{i=1}^K dx_i \pi(x_i) \left( \frac{1 + \prod_{i=1}^K x_i}{2} \right)^\rho \right. \\ \left. + \ln \left[ \int \prod_{\mu=1}^C d\hat{x}_\mu \hat{\pi}(\hat{x}_\mu) \left\langle \left( e^{\frac{F}{1+\rho} \zeta} \prod_{\mu=1}^C \left( \frac{1 + \hat{x}_\mu}{2} \right) + e^{-\frac{F}{1+\rho} \zeta} \prod_{\mu=1}^C \left( \frac{1 - \hat{x}_\mu}{2} \right) \right)^\rho \right\rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}} \right] \right. \\ \left. + C \ln \hat{q} - C q \hat{q} \int dx d\hat{x} \pi(x) \hat{\pi}(\hat{x}) \left( \frac{1 + x \hat{x}}{2} \right)^\rho - \left( \frac{C}{K} - C \right) \right\}, \quad (6)$$

where  $\text{Ext}^*$  denotes extremization which excludes the ferro-magnetic solution and  $\langle \cdot \rangle_{\boldsymbol{\zeta} | \frac{F}{1+\rho}}$  is over  $P(\boldsymbol{\zeta}; \frac{F}{1+\rho})$ .

Before proceeding any further, we would like to mention some general properties of  $E_{av}(\rho, R)$ . From Eqs. (5) and (6), it can be shown that  $\lim_{\rho \rightarrow 0} E_{av}(\rho, R) = 0$  and  $\partial^2 E_{av}(\rho, R) / \partial \rho^2 < 0$ . This implies that  $\text{Max}_{\rho > 0} E_{av}(\rho, R)$ , becomes positive if and only if  $\partial E_{av}(\rho, R) / \partial \rho |_{\rho=0} > 0$ , for which  $\lim_{M \rightarrow \infty} \bar{P}_E = 0$  holds. Therefore, the zero error threshold, defined as the critical flip rate below which the average error rate vanishes as  $M \rightarrow \infty$ , is obtained by the condition  $\partial E_{av}(\rho, R) / \partial \rho = 0$ . From (5), this becomes

$$F \tanh F - \frac{1}{M} \langle \ln \mathcal{Z}_{\text{NF}}(\boldsymbol{\zeta}, D; F) \rangle_{\boldsymbol{\zeta} | F, D} = 0. \quad (7)$$

The second term is the averaged free energy for the Hamiltonian (2) with respect to the quenched randomness  $\boldsymbol{\zeta}$  and  $D$ , in the non-ferromagnetic phase. Employing the ferromagnetic gauge [10] one obtains the following expression for the ferromagnetic free energy (where

$\bar{P}_E=0$ ):  $(1/M) \langle \ln \mathcal{Z}_F(\zeta, D; F) \rangle_{\zeta|_{F,D}} = F \tanh F$ . Since the correct prior information about the flip rate  $p$  is used in the calculation, these two free energies are actually obtained in Nishimori's *finite* decoding temperature ( $\beta=1$ ) [12,11,10,5] for which the bit error probability is minimized. By satisfying (7), the zero error threshold for ML decoding, which corresponds to the zero temperature limit ( $\beta \rightarrow \infty$ ) [12,5], is determined by the phase boundary between the ferromagnetic and non-ferromagnetic phases at  $\beta=1$ .

Using the ferromagnetic gauge provides insight into the physical properties of the system. As the internal energy per bit in the non-ferromagnetic system is  $-F \tanh F$  under Nishimori's condition, Eq. (7) implies that the entropy of the non-ferromagnetic phase vanishes at the phase boundary for  $\beta=1$ , suggesting that this phase exhibits a replica symmetry breaking (RSB) at lower temperatures in general, and at  $\beta \rightarrow \infty$  in particular. In this sense, the zero-error threshold prediction obtained from Gallager's method and ML decoding, is surprising as it provides information about the ferro/non-ferro phase boundary at  $\beta \rightarrow \infty$  which is not easily obtained via the methods of statistical physics due to RSB effects. This argument can be extended to the case of general  $\beta \geq 1$ , as will be presented elsewhere.

An analytical expression to  $E_{av}(\rho, R)$  can be obtained in the limit  $K, C \rightarrow \infty$ , keeping the code rate  $R=1-C/K$  finite; for the non-ferromagnetic solution one then obtains  $q=2^{\rho/K}$ ,  $\hat{q}=2^{\rho(1-1/K)}$ ,  $\pi(x)=\delta(x)$  and  $\hat{\pi}(\hat{x})=(1/2)(1+\tanh F)\delta(\hat{x}-\tanh F)+(1/2)(1-\tanh F)\delta(\hat{x}+\tanh F)$ . Using Eqs. (5) and (6), one obtains the explicit expression  $E_{av}(\rho, R) = \ln 2 \cosh F - (1+\rho) \ln \left( 2 \cosh \frac{F}{1+\rho} \right) + \rho(1-R) \ln 2$ . In addition, there exists another solution for  $\rho \geq 1$ ,  $q=2^{1/K}$ ,  $\hat{q}=2^{1-1/K}$ ,  $\pi(x)=(1/2)\delta(x-1)+(1/2)\delta(x+1)$  and  $\hat{\pi}(\hat{x})=(1/2)\delta(\hat{x}-1)+(1/2)\delta(\hat{x}+1)$  providing  $E_{av}(\rho, R) = \ln 2 \cosh F - \ln \left( 2 \cosh F + 2 \cosh \left( \frac{1-\rho}{1+\rho} F \right) \right) + (1-R) \ln 2$ . Employing a method similar to that in [9,8], it can be shown that both RS solutions are locally stable against perturbations to the replica symmetric solution.

The relation between  $E_{av}(\rho, R)$  and the entropy of non-ferromagnetic solutions  $\mathcal{S}_{\text{NF}}$

$$\frac{\partial E_{av}(\rho, R)}{\partial \rho} = - \frac{\langle \mathcal{Z}_{\text{NF}}^\rho \left( \zeta, D; \frac{F}{1+\rho} \right) \mathcal{S}_{\text{NF}} \left( \zeta, D; \frac{F}{1+\rho} \right) \rangle_{\zeta|_{\frac{F}{1+\rho}, D}}}{\langle \mathcal{Z}_{\text{NF}}^\rho \left( \zeta, D; \frac{F}{1+\rho} \right) \rangle_{\zeta|_{\frac{F}{1+\rho}, D}}},$$

suggests another type of RSB, indicated by the negative entropy. This implies that the

entropy of the non-ferromagnetic RS solutions vanishes at  $\rho = \rho^*(R)$  which maximizes  $E_{av}(\rho, R)$ ; and the tightest lower bound of  $E(R)$  is therefore obtained at the RSB transition, which can be calculated from the locally stable RS solutions.

Solving the maximization problem one obtains

$$\text{Max}_{\rho>0} E_{av}(\rho, R) = \begin{cases} \ln 2 \cosh F - (1-R) \ln 2 & F \geq 2F^*(R) \\ -\ln(2 \cosh F + 2) , & \\ \ln 2 \cosh F - (1-R) \ln 2 & 2F^*(R) \geq F \geq F^*(R) \\ -F \tanh F^*(R) , & \\ 0 , & \text{otherwise} \end{cases} \quad (8)$$

where  $F^*(R)$  is the solution of the equation  $\ln 2 \cosh F^* - F^* \tanh F^* - (1-R) \ln 2 = 0$ . The position of the maximum is given as  $\rho^*(R) = 1$  for  $F \geq 2F^*(R)$ ,  $F/F^*(R) - 1$  for  $2F^*(R) \geq F \geq F^*(R)$  and 0, otherwise. Using the relation between  $F$  and  $p$ , this indicates that  $E(R)$  becomes positive if and only if  $R < 1 - H_2(p)$ , which corresponds to Shannon's limit.

Equation (8) is identical to the random coding exponent  $E_r(R)$  obtained in the IT literature [3], although one should emphasize the main differences between the two approaches: a) Starting from Gallager's inequality (1) we directly average over the ensemble while the  $E_r(R)$  result is obtained by invoking Jensen's inequality. b) Our result is obtained for an ensemble of a specific code.

With some hindsight, this is not very surprising as Gallager codes become similar to random codes in the limit  $K, C \rightarrow \infty$  [7,6]; this also implies that using Jensen's inequality does not produce a looser bound as initially thought.

To get a tighter bound for low  $R$  values we employ a refined inequality, upper-bounding the ensemble minimum of  $P_E$  by  $\left\langle \left( \text{Tr}_{\{\mathbf{J}, \boldsymbol{\zeta}\}} P^{\frac{1}{1+\rho}}(\mathbf{J}, \boldsymbol{\zeta}) \left( \text{Tr}_{\{\mathbf{J}, \mathbf{n} \neq \boldsymbol{\zeta}\}} P^{\frac{1}{1+\rho}}(\mathbf{J}, \mathbf{n}) \right)^\rho \right)^m \right\rangle_D^{\frac{1}{m}}$  ( $\rho > 0$ ,  $m > 0$ ), as in (1). A similar calculation along the lines described here (details will be shown elsewhere) provides the *expurgated exponent* bound [3] result for low  $R$  values (see Fig.1); this links our results to the best bounds reported in the IT literature to date.



Without trivializing the results obtained in the case of  $K, C \rightarrow \infty$ , the main achievement of our approach is the ability to investigate analytically the performance of Gallager (or similar) codes of finite  $K$  and  $C$ . To demonstrate the accuracy of the bounds obtained we examine the case of  $K=6$  and  $C=3$ . We numerically evaluated  $E_{av}(\rho, R)$  (5) for  $p=0.0915$ , a recent highly accurate estimate of the error threshold for this parameter [1], and for  $p=0.0990$ , which is the threshold predicted by our analysis. The numerical results were obtained by approximating  $\pi(\cdot)$  and  $\hat{\pi}(\cdot)$  using  $10^6$  dimensional vectors and iterating the saddle point equations until convergence. The results are shown in the inset; they indicate that  $\text{Max}_{\rho \geq 0} E_{av}(\rho, R) \simeq 1.0 \times 10^{-4} > 0$  for  $p=0.0915$  while  $E_{av}(\rho, R)$  is maximized (to zero) in the vicinity of  $\rho=0$  for  $p=0.0990$ , suggesting a tighter estimate for the error threshold than those reported in the IT literature.

In summary, we have developed a method to tightly upper-bound the dependence of the decoding error rate on the message length for Gallager codes. In the limit of infinite connectivity our result collapses onto the best general random coding exponents reported in the IT literatures, the *random coding exponent* and the *expurgated exponent* for high and low  $R$  values respectively. The method provides one of the only tools available for examining codes of finite connectivity; and predicts the tightest estimate of the zero error noise level threshold to date for Gallager codes. It can be easily extended to investigate other linear codes of a similar type and is clearly of high practical significance.

We demonstrated how the methods of statistical physics may complement and improve results obtained in the IT literature. These methods are applicable to a broad range of problems, especially within the sub-field of coding, and may be instrumental in improving existing results; some of these studies are already under way.

**Acknowledgement** We acknowledge support from the JSPS-RFTF program (YK), EPSRC (GR/N00562) and the Royal Society (DS). YK would like to thank Y. Iba for kindly showing him an unpublished manuscript [4] and D.J.C. MacKay for informing us of [1] prior to publication.

## REFERENCES

- [1] S. Aji, H. Jin, A. Khandekar, D.J.C. MacKay and R.J. McEliece, BSC Thresholds Code Ensembles Based on “Typical Pairs” Decoding, preprint (1999).
- [2] R.G. Gallager, IRE. Trans. on IT, **IT-8**, 21 (1962).
- [3] R.G. Gallager, IEEE Trans. IT, **11**, 3 (1965).
- [4] Y. Iba, Gallager Formalism in Information Theory and Replica Method (in Japanese), unpublished (1989).
- [5] Y. Iba, J. Phys. A: Math. and Gen., **32**, 3875 (1999).
- [6] Y. Kabashima, T. Murayama and D. Saad, Phys. Rev. Lett., **84**, 1355 (2000); T. Murayama, Y. Kabashima, D. Saad and R. Vincente, Phys. Rev. E, **62**, 1577 (2000).
- [7] D.J.C. MacKay, IEEE Trans. IT, **45**, 399 (1999); D.J.C. MacKay and R.M. Neal, Elect. Lett., **33**, 457 (1997).
- [8] A. Montanari and N. Surlas, Eur. Phys. J. B **18**, 107 (2000); A. Montanari, Eur. Phys. J. B **18**, 121 (2000)
- [9] P. Mottishaw and C. De Dominicis, J. Phys. A: Math. and Gen., **20**, L375 (1987).
- [10] H. Nishimori, J. Phys. Soc. Jpn., **62**, 2973 (1993); Prog. Theor. Phys., **66**, 1169 (1981).
- [11] P. Rujan, Phys. Rev. Lett., **70**, 2968 (1993).
- [12] N. Surlas, Nature, **339**, 693 (1989); Europhys. Lett., **25**, 693 (1994).
- [13] C.E. Shannon, Bell Sys. Tech. J., **27**, 379 (1948); **27**, 623 (1948).

FIGURES

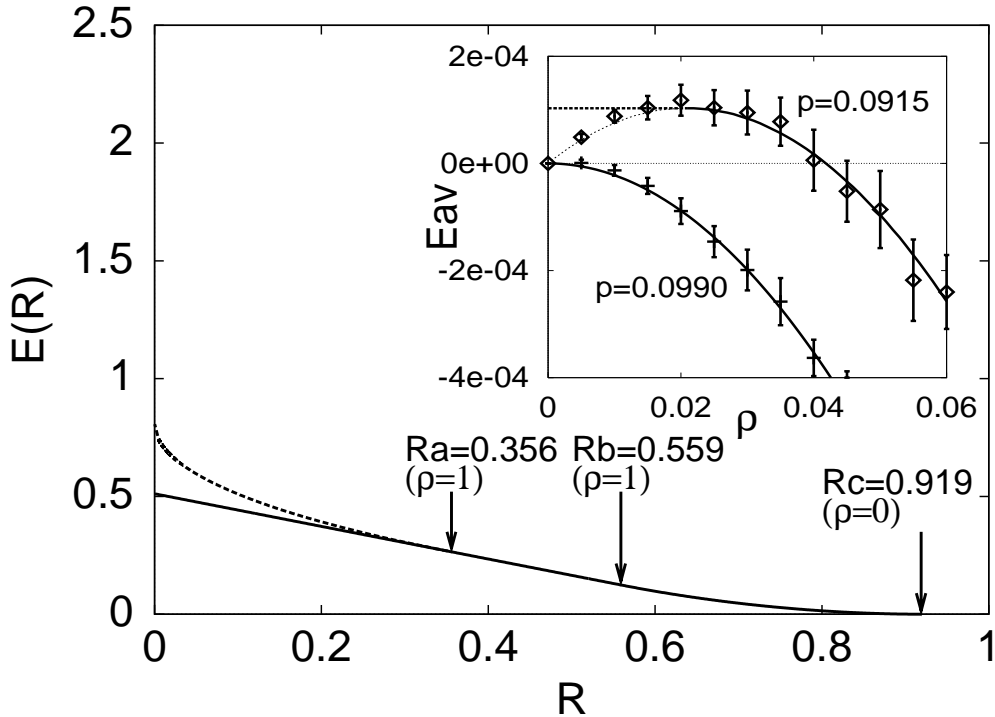


FIG. 1. Lower-bounds on the reliability exponent  $E(R)$  obtained for  $p = 0.01$  in the limit  $K, C \rightarrow \infty$ . Our method produces the same result as the random coding exponent  $E_r(R)$  (solid line) which provides an excellent bound for  $R > R_b$ . For low  $R < R_a$  values the bound becomes loose, and a better result (dashed line), identical to the expurgated exponent bound, is obtained (see text) by employing a refined inequality in (1). Inset - The exponent  $E_{av}(\rho, R)$  obtained numerically for a choice of finite parameters  $K=6$  and  $C=3$  ( $R=1/2$ ). Symbols and standard deviations are computed using 50 numerical solutions. Curves are obtained via a quadratic fit. For  $p=0.0915$ ,  $\rho^*(R) \simeq 0.02$ , suggesting that this flip rate is still below the threshold. Left of the peak, the RS solution (thin broken curve) is unstable. For  $p=0.0990$ , our predicted threshold, the maximum  $E_{av}(\rho, R) \simeq 0$  is obtained at  $\rho \simeq 0$ , implying that this is the correct threshold.