EUROPHYSICS LETTERS 5 May 2000

Europhys. Lett., (), pp. (2000)

Error-correcting code on a cactus: a solvable model

R. VICENTE[1], D. SAAD[1] AND Y. KABASHIMA[2]

[1] *The Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK*
[2] *Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama 2268502, Japan*

(received ; accepted)

PACS. 89.90+n – Other areas of general interest to physicists..
PACS. 89.70+c – Information Theory.
PACS. 05.50+q – Lattice theory and statistics; Ising problems.

Abstract. – An exact solution to a family of parity check error-correcting codes is provided by mapping the problem onto a Husimi cactus. The solution obtained in the thermodynamic limit recovers the replica symmetric theory results and provides a very good approximation to finite systems of moderate size. The probability propagation decoding algorithm emerges naturally from the analysis. A phase transition between decoding success and failure phases is found to coincide with an information-theoretic upper bound. The method is employed to compare Gallager and MN codes.

The theory of error-correcting codes concentrates on the efficient introduction of redundancy to given messages for protecting the information content against corruption. The theoretical foundations of this area were laid by Shannon's seminal work [1] and have been developing ever since (see [2] and references therein). One of the main results obtained in this field is the celebrated *channel coding theorem* stating that there exists a code such that the average message error probability P_E, when maximum likelihood decoding is used, is upper bounded by $P_E < e^{-M\,E(R)}$, where M is the length of the encoded transmission and $R = (\text{ message information content })/M$ is the code rate. The exponent $E(R)$ is positive for code rates below the *channel capacity*, corresponding to the maximal mutual information between the received and the transmitted signals, and vanishes above it. For rates R below the channel capacity, commonly termed *Shannon's bound*, the error probability can be made arbitrarily small.

The channel coding theorem is based on unstructured random codes and impractical decoders as maximum likelihood [2] or typical sets [3]. In the last fifty years several practical methods have been proposed and implemented, but none has been able to saturate Shannon's bound. In 1963 Gallager [4] proposed a coding scheme involving sparse linear transformations of binary messages that was forgotten soon after, in part due to the success of convolutional codes [2] and the computational limitations of the time. Gallager codes have been recently rediscovered by MacKay and Neal (MN) that independently proposed a closely related code [3]. This almost coincided with the breakthrough discovery of the high-performance turbo codes [5].

Variations of Gallager codes have displayed performance comparable (and sometimes superior) to turbo codes [6], qualifying them as state-of-the-art codes.

Statistical physics has been applied to the analysis of error-correcting codes as an alternative to information theory methods yielding some new interesting directions and suggesting new high-performance codes [7]. Sourlas was the first to relate error-correcting codes to spin glass models [8], showing that the Random Energy Model (REM)[9, 10, 11] can be thought of as an ideal code, capable of saturating Shannon's bound at vanishing code rates. This work was extended recently to the case of finite code rates [12, 13] and has been further developed for analysing MN codes of various structures [14, 15, 16]. All of the analyses mentioned above, as well as the recent turbo code analysis [17], relied on the replica approach under the assumption of replica symmetry. It is also worthwhile mentioning a different approach, used in the analysis of convolutional codes [18], of employing the transfer-matrix formalism and power series expansions. However, to date, the only model that can be analysed exactly is the REM that corresponds to an impractical coding scheme of a vanishing code rate.

In this letter we present an *exact* analysis to the performance of Gallager error-correcting codes based on a generalisation of Bethe lattices known as the Husimi cactus [19]. We solve the model recovering results obtained by the replica symmetric theory and finding the noise level that corresponds to the phase transition between perfect decoding and a decoding failure phase, this appears to coincide with existing information-theoretic upper bounds. We experimentally show that the solution accurately approximates Gallager codes of moderate size. We also show that the probability propagation (PP) decoding algorithm emerges naturally from this framework allowing for the analysis of the practical decoding performance. Finally, we summarise the differences between Gallager and MN codes, which are somewhat obscure in the information theory literature but become explicit in this framework.

We will concentrate here on a simple communication model whereby messages are represented by binary vectors and are communicated through a Binary Symmetric Channel (BSC) where uncorrelated bit flips appear with probability f. A Gallager code is defined by a binary matrix $\boldsymbol{A} = [\boldsymbol{C_1} \mid \boldsymbol{C_2}]$, concatenating two very sparse matrices known to both sender and receiver, with $\boldsymbol{C_2}$ (of dimensionality $(M-N)\times(M-N)$) being invertible; the matrix $\boldsymbol{C_1}$ is of dimensionality $(M-N)\times N$.

Encoding refers to the production of an M dimensional binary code word $\boldsymbol{t} \in \{0,1\}^M$ $(M > N)$ from the original message $\boldsymbol{\xi} \in \{0,1\}^N$ by $\boldsymbol{t} = \boldsymbol{G^T\xi} \pmod 2$, where all operations are performed in the field $\{0,1\}$ and are indicated by (mod 2). The generator matrix is $\boldsymbol{G} = [\boldsymbol{I} \mid \boldsymbol{C_2^{-1}C_1}] \pmod 2$, where $\boldsymbol{I}$ is the $N\times N$ identity matrix, implying that $\boldsymbol{AG^T} \pmod 2 = 0$ and that the first N bits of $\boldsymbol{t}$ are set to the message $\boldsymbol{\xi}$. In *regular* Gallager codes the number of non-zero elements in each row of $\boldsymbol{A}$ is chosen to be exactly K. The number of elements per column is then $C = (1-R)K$, where the code rate is $R = N/M$ (for unbiased messages). The encoded vector $\boldsymbol{t}$ is then corrupted by noise represented by the vector $\boldsymbol{\zeta} \in \{0,1\}^M$ with components independently drawn from $P(\zeta) = (1-f)\delta(\zeta) + f\delta(\zeta-1)$. The received vector takes the form $\boldsymbol{r} = \boldsymbol{G^T\xi} + \boldsymbol{\zeta} \pmod 2$.

Decoding is carried out by multiplying the received message by the matrix $\boldsymbol{A}$ to produce the *syndrome* vector $\boldsymbol{z} = \boldsymbol{Ar} = \boldsymbol{A\zeta} \pmod 2$ from which an estimate $\widehat{\boldsymbol{\tau}}$ for the noise vector can be produced. An estimate for the original message is then obtained as the first N bits of $\boldsymbol{r} + \widehat{\boldsymbol{\tau}} \pmod 2$. The Bayes optimal estimator (also known as *marginal posterior maximiser*, MPM) for the noise is defined as $\widehat{\tau}_j = \mathrm{argmax}_{\tau_j} P(\tau_j \mid z)$. The performance of this estimator can be measured by the probability of bit error $p_b = 1 - 1/M \sum_{j=1}^{M} \delta[\widehat{\tau}_j; \zeta_j]$, where $\delta[;]$ is Kronecker's delta. Knowing the matrices $\boldsymbol{C_2}$ and $\boldsymbol{C_1}$, the syndrome vector $\boldsymbol{z}$ and the noise

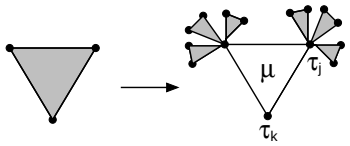

Fig. 1. – First step in the construction of Husimi cactus with $K = 3$ and connectivity $C = 4$.

level f it is possible to apply Bayes' theorem and compute the posterior probability

$$P(\boldsymbol{\tau} \mid \boldsymbol{z}) = \frac{1}{Z}\chi\left[\boldsymbol{z} = \boldsymbol{A}\boldsymbol{\tau}(\mathrm{mod}\ 2)\right] P(\boldsymbol{\tau}), \tag{1}$$

where $\chi[X]$ is an indicator function providing 1 if X is true and 0 otherwise. To obtain the MPM one has to compute the marginal posterior $P(\tau_j \mid \boldsymbol{z}) = \sum_{i \neq j} P(\boldsymbol{\tau} \mid \boldsymbol{z})$, which in general requires $\mathcal{O}(2^M)$ operations, and thus becomes impractical for long messages. To solve this problem one can use the sparseness of $\boldsymbol{A}$ to design algorithms that require $\mathcal{O}(M)$ operations to perform the same task. One of these methods is the probability propagation algorithm (PP), also known as belief propagation, sum-product algorithm (see [20]) or generalised distributive law [21].

The connection to statistical physics becomes clear when the field $\{0, 1\}$ is replaced by Ising spins $\{\pm 1\}$ and mod 2 sums by products [8]. The syndrome vector acquires the form of a multi-spin coupling $\mathcal{J}_\mu = \prod_{j \in \mathcal{L}(\mu)} \zeta_j$ where $j = 1, \cdots, M$ and $\mu = 1, \cdots, (M - N)$. The K indices of nonzero elements in the row μ of $\boldsymbol{A}$ are given by $\mathcal{L}(\mu) = \{j_1, \cdots, j_K\}$, and in a column l are given by $\mathcal{M}(l) = \{\mu_1, \cdots, \mu_C\}$.

The posterior (1) can be written as the Gibbs distribution [14, 15]:

$$\begin{aligned} P(\boldsymbol{\tau} \mid \mathcal{J}) &= \frac{1}{Z} \lim_{\beta \to \infty} \exp\left[-\beta \mathcal{H}_\beta(\boldsymbol{\tau}; \mathcal{J})\right] \\ \mathcal{H}_\beta(\boldsymbol{\tau}; \mathcal{J}) &= -\sum_{\mu=1}^{M-N} \left(\mathcal{J}_\mu \prod_{j \in \mathcal{L}(\mu)} \tau_j - 1 \right) - \frac{F}{\beta} \sum_{j=1}^{M} \tau_j \ . \end{aligned} \tag{2}$$

The external field corresponds to the prior probability over the noise and has the form $F = \mathrm{atanh}(1 - 2f)$. Note that the Hamiltonian itself depends on the inverse temperature β. The disorder is trivial and can be gauged as $\mathcal{J}_\mu \mapsto 1$ by using $\tau_j \mapsto \tau_j \zeta_j$. The resulting Hamiltonian is a multi-spin ferromagnet with finite connectivity in a random field $h_j = \beta^{-1} F \zeta_j$. The decoding process corresponds to finding *zero temperature* local magnetisations $m_j = \lim_{\beta \to \infty} \langle \tau_j \rangle_\beta$ and calculating estimates as $\widehat{\tau}_j = \mathrm{sgn}(m_j)$.

In the $\{\pm 1\}$ representation the probability of bit error, acquires the form

$$p_b = \frac{1}{2} - \frac{1}{2M} \sum_{j=1}^{M} \zeta_j \ \mathrm{sgn}(m_j), \tag{3}$$

connecting the code performance with the computation of local magnetisations.

A Husimi cactus with connectivity C is generated starting with a polygon of K vertices with one Ising spin in each vertex (generation 0). All spins in a polygon interact through a single coupling $\mathcal{J}_\mu$ and one of them is called the base spin. In figure 1 we show the first step in the construction of a Husimi cactus, in a generic step the base spins of the $n-1$ generation polygons,

numbering $(C-1)(K-1)$, are attached to $K-1$ vertices of a generation n polygon. This process is iterated until a maximum generation $n_{\max}$ is reached, the graph is then completed by attaching C uncorrelated branches of $n_{\max}$ generations at their base spins. In that way each spin inside the graph is connected to exactly C polygons. The local magnetisation at the centre m_j can be obtained by fixing boundary (initial) conditions in the 0-th generation and iterating recursion equations until generation $n_{\max}$ is reached. Carrying out the calculation in the thermodynamic limit corresponds to having $n_{\max} \sim \ln M$ generations and $M \to \infty$.

The Hamiltonian of the model has the form (2) where $\mathcal{L}(\mu)$ denotes the polygon μ of the lattice. Due to the tree-like structure, local quantities far from the boundary can be calculated recursively by specifying boundary conditions. The typical decoding performance can therefore be computed exactly without resorting to replica calculations [22].

We adopt the approach presented in [19] where recursion relations for the probability distribution $P_{\mu k}(\tau_k)$ for the base spin of the polygon μ is connected to $(C-1)(K-1)$ distributions $P_{\nu j}(\tau_j)$, with $\nu \in \mathcal{M}(j) \setminus \mu$ (all polygons linked to j but μ) of polygons in the previous generation:

$$P_{\mu k}(\tau_k) = \frac{1}{\mathcal{N}} \operatorname{Tr}_{\{\tau_j\}} \exp\left[\beta\left(\mathcal{J}_\mu \tau_k \prod_{j\in\mathcal{L}(\mu)\setminus k} \tau_j - 1\right) + F\tau_k\right] \prod_{\nu\in\mathcal{M}(j)\setminus\mu} \prod_{j\in\mathcal{L}(\mu)\setminus k} P_{\nu j}(\tau_j), \quad (4)$$

where the trace is over the spins τ_j such that $j \in \mathcal{L}(\mu) \setminus k$.

The effective field $\widehat{x}_{\nu j}$ on a base spin j due to neighbours in polygon ν can be written as :

$$\exp(-2\widehat{x}_{\nu j}) = \mathrm{e}^{2F} \frac{P_{\nu j}(-)}{P_{\nu j}(+)}, \quad (5)$$

Combining (4) and (5) one finds the recursion relation:

$$\exp(-2\widehat{x}_{\mu k}) = \frac{\operatorname{Tr}_{\{\tau_j\}} \exp\left[-\beta\mathcal{J}_\mu \prod_{j\in\mathcal{L}(\mu)\setminus k} \tau_j + \sum_{j\in\mathcal{L}(\mu)\setminus k}(F + \sum_{\nu\in\mathcal{M}(j)\setminus\mu} \widehat{x}_{\nu j})\tau_j\right]}{\operatorname{Tr}_{\{\tau_j\}} \exp\left[+\beta\mathcal{J}_\mu \prod_{j\in\mathcal{L}(\mu)\setminus k} \tau_j + \sum_{j\in\mathcal{L}(\mu)\setminus k}(F + \sum_{\nu\in\mathcal{M}(j)\setminus\mu} \widehat{x}_{\nu j})\tau_j\right]}. \quad (6)$$

By computing the traces and taking $\beta \to \infty$ one obtains:

$$\widehat{x}_{\mu k} = \operatorname{atanh}\left[\mathcal{J}_\mu \prod_{j\in\mathcal{L}(\mu)\setminus k} \tanh(F + \sum_{\nu\in\mathcal{M}(j)\setminus\mu} \widehat{x}_{\nu j})\right] \quad (7)$$

The effective local magnetisation due to interactions with the nearest neighbours in one branch is given by $\widehat{m}_{\mu j} = \tanh(\widehat{x}_{\mu j})$. The effective local field on a base spin j of a polygon μ due to $C-1$ branches in the previous generation and due to the external field is $x_{\mu j} = F + \sum_{\nu\in\mathcal{M}(j)\setminus\mu} \widehat{x}_{\nu j}$; the effective local magnetisation is, therefore, $m_{\mu j} = \tanh(x_{\mu j})$. Equation (7) can then be rewritten in terms of $\widehat{m}_{\mu j}$ and $m_{\mu j}$ and the PP equations [3, 12, 20] can be recovered:

$$m_{\mu k} = \tanh\left(F + \sum_{\nu\in\mathcal{M}(j)\setminus\mu} \operatorname{atanh}(\widehat{m}_{\nu k})\right) \qquad \widehat{m}_{\mu k} = \mathcal{J}_\mu \prod_{j\in\mathcal{L}(\mu)\setminus k} m_{\mu j} \quad (8)$$

Once the magnetisations on the boundary (0-th generation) are assigned, the local magnetisation m_j in the central site is determined by iterating (8) and computing :

$$m_j = \tanh\left(F + \sum_{\nu\in\mathcal{M}(j)} \operatorname{atanh}(\widehat{m}_{\nu j})\right) \quad (9)$$

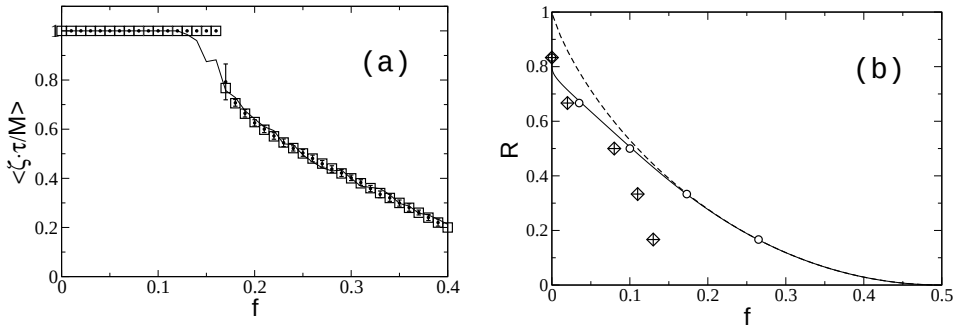

Fig. 2. – (a) Mean normalised overlap between the actual noise vector $\boldsymbol{\zeta}$ and decoded noise $\widehat{\boldsymbol{\tau}}$ for $K = 4$ and $C = 3$ (therefore $R = 1/4$). Theoretical values ($\Box$), experimental averages over 20 runs for code word lengths $M = 5000$ ($\bullet$) and $M = 100$ (full line). (b) Transitions for $K = 6$. Shannon's bound (dashed line), information theory upper bound (full line) and thermodynamic transition obtained numerically ($\circ$). Theoretical ($\diamond$) and experimental ($+$, $M = 5000$ averaged over 20 runs) PP decoding transitions are also shown. In both figures, symbols are chosen larger than the error bars.

The free energy can be obtained by integration as (8) represents extrema of the free energy [15, 16, 23].

By applying the gauge transformation $\mathcal{J}_\mu \mapsto 1$ and $\tau_j \mapsto \tau_j \zeta_j$, assigning the probability distributions $P_0(x)$ to boundary fields and averaging over random local fields $F\zeta$ one obtains from (7) the recursion relation in the space of probability distributions $P(x)$ [23]:

$$\begin{aligned} P_n(x) &= \int \prod_{l=1}^{C-1} d\widehat{x}_l \; \widehat{P}_{n-1}(\widehat{x}_l) \; \left\langle \delta \left[x - F\zeta - \sum_{l=1}^{C-1} \widehat{x}_l \right] \right\rangle_\zeta \\ \widehat{P}_{n-1}(\widehat{x}) &= \int \prod_{j=1}^{K-1} dx_j \; P_{n-1}(x_j) \; \delta \left[\widehat{x} - \operatorname{atanh} \left(\prod_{j=1}^{K-1} \tanh(x_j) \right) \right], \end{aligned} \tag{10}$$

where $P_n(x)$ is the distribution of effective fields at the n-th generation due to the previous generations and external fields, in the thermodynamic limit the distribution far from the boundary will be $P_\infty(x)$ (generation $n \to \infty$). The local field distribution at the central site is computed by replacing $C - 1$ by C in (10), taking into account C polygons in the generation just before the central site, and inserting the distribution $P_\infty(x)$. Equations (10) are identical to those obtained by the replica symmetric theory as in [14, 15, 16].

By setting initial (boundary) conditions $P_0(x)$ and numerically iterating (10), for $C \geq 3$ one can find, up to some noise level f_s, a single stable fixed point at infinite fields, corresponding to a totally aligned state (successful decoding). At f_s a bifurcation occurs and two other fixed points appear, stable and unstable, the former corresponding to a misaligned state (decoding failure). This situation is identical to that one observed in [14, 15, 16]. In terms of the local fields distribution $P_n(x)$, the aligned state corresponds to a runaway wave travelling to $x(n) \to \infty$ with n being the time variable. The misaligned state corresponds to a stable wave located at $x(n) \sim \mathcal{O}(1)$. Representing the distributions (10) by the first cummulants only, one can obtain a rough approximation in terms of one dimensional maps showing a bifurcation at some noise level $\tilde{f}_s$, this approach will be further exploited elsewhere.

The practical PP decoding is performed by setting initial conditions as $m_{\mu j} = 1 - 2f$ to

TABLE I. – *Gallager versus MN codes*

	Gallager	MN
dynamical variables	M	$N+M$
constraints	M-N	M
unbiased messages coding	*for all* K	K =1,2
Shannon's bound	$K \to \infty$	$K \geq$ 3 *and unbiased messages*

correspond to the prior probabilities and iterating (8) until stationarity or a maximum number of iterations is attained [3]. The estimate for the noise vector is then produced by computing $\widehat{\tau}_j = \text{sign}(m_j)$. At each decoding step the system can be described by histograms of the variables (8), this is equivalent to iterating (10) (a similar idea was presented in [3, 6]). Below f_s the process always converges to the successful decoding state, above f_s it converges to the successful decoding only if the initial conditions are fine tuned; in general the process converges to the failure state. In Fig.2a we show the theoretical mean overlap between actual noise $\boldsymbol{\zeta}$ and the estimate $\widehat{\boldsymbol{\tau}}$ as a function of the noise level f as well as results obtained with PP decoding.

Information theory provides an upper bound for the maximum attainable code rate by equalising the maximal information contents of the syndrome vector $\boldsymbol{z}$ and of the noise estimate $\widehat{\boldsymbol{\tau}}$ [3, 16]. The thermodynamic phase transition obtained by finding the stable fixed points of (10) and their free energies interestingly coincides with this upper bound within the precision of the numerical calculation. Note that this predicted performance is impractical as it requires $\mathcal{O}(2^M)$ operations for an exhaustive search for the global minimum of the free energy. In Fig.2b we show the thermodynamic transition for $K = 6$ compared with the upper bound, Shannon's bound and f_s values.

The geometrical structure of a Gallager code defined by the matrix $\boldsymbol{A}$ can be represented by a bipartite graph (*Tanner graph*) [20] with bit and check nodes. Each column j of $\boldsymbol{A}$ represents a bit node and each row μ represents a check node, $A_{\mu j} = 1$ means that there is an edge linking bit j to check μ. It is possible to show [24] that for a random ensemble of regular codes, the probability of completing a cycle after walking l edges starting from an arbitrary node is upper bounded by $\mathcal{P}[l; K, C, M] \leq l^2 K^l / M$. It implies that for very large M only cycles of at least order $\ln M$ survive. In the thermodynamic limit $M \to \infty$ the probability $\mathcal{P}[l; K, C, M] \to 0$ for any finite l and the bulk of the system is effectively tree-like. By mapping each check node to a polygon with K bit nodes as vertices, one can map a Tanner graph into a Husimi lattice that is effectively a tree for any number of generations of order less than $\ln M$. It is experimentally observed that the number of iterations of (8) required for convergence does not scale with the system size, therefore, it is expected that the interior of a tree-like lattice approximates a Gallager code with increasing accuracy as the system size increases. Fig.2a shows that the approximation is fairly good even for sizes as small as $M = 100$. Note that although the local magnetisations m_j for a loopy graph are not generally expected to converge to the values computed in a tree, $\text{sgn}(m_j)$ seems to do so. A thorough discussion on this respect for some specific graphical models can be found in [25].

In [3] MacKay and Neal introduced a variation on Gallager codes termed MN codes. The main difference between these codes is that for MN codes the syndrome vector contains also information on the original message in the form $\boldsymbol{z} = \boldsymbol{C_s}\boldsymbol{\xi} + \boldsymbol{C_n}\boldsymbol{\zeta}$. The message itself is directly estimated and there is no need for recovering the noise vector. MacKay has formulated and proved a number of theorems simultaneously for both codes using the fact that if both message and noise are sampled from the same distribution, these codes can be formulated as the same

estimation problem, i.e., finding the most probable vector $\boldsymbol{x}$ that satisfies $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x}$, given the matrix $\boldsymbol{A}$ and a prior distribution $P(\boldsymbol{x})$. Using statistical physics, we previously analysed MN codes [14, 15, 16]. It is interesting to note that in spite of the similarity between the two codes, there are some important differences in their dependence on the parameters K and C. In particular, Shannon's bound is only attainable by Gallager codes if $K \to \infty$, in contrast to results obtained for MN codes. Decoding of unbiased messages is generally possible with Gallager codes, but successful convergence is only guaranteed (in the thermodynamic limit) for $K = 1, 2$ in the MN codes. We outlined those differences in table I.

To summarise, we solved exactly, without resorting to the replica method, a system representing a Gallager code on a Husimi cactus. The results obtained are in agreement with the replica symmetric calculation and with numerical experiments carried out in systems of moderate size. The framework can be easily extended to MN and similar codes. We believe that methods of statistical physics are complimentary to those used in the statistical inference community and can enhance our understanding of general graphical models beyond error-correcting codes.

***

We acknowledge support from EPSRC (GR/N00562), The Royal Society (RV,DS) and from the JSPS RFTF program (YK).

REFERENCES

[1] Shannon C., *Bell Syst. Tech. J.*, **27** (1948) 379.
[2] Viterbi A.J. and Omura J. K., *Principles of Digital Communication and Coding*, (McGraw-Hill Book Co., Singapore) 1979.
[3] MacKay D.J.C. and Neal R.M., *Electr. Lett.*, **32** (1996) 1645; MacKay D.J.C., *IEEE Trans. Info. Theory*, **45**(1999) 399.
[4] Gallager R.G., *IRE Trans. Info. Theory*, **8** (1962) 21.
[5] Berrou G. et al., *Proc. of the IEEE Int. Conf. on Comm.*, (Geneva) 1993 p. 1064.
[6] Davey M.C., *Record-braking error-correction using low-density parity-check codes*, (Hamilton prize essay, University of Cambridge) 1998.
[7] Kanter I. and Saad D., *Phys. Rev. Lett.*, **83**(1999) 2660; *J. Phys A*, **33** (2000) 1675.
[8] Sourlas N., *Nature*, **339** (1989) 693; *Europhys. Lett.*, **25**(1994) 159.
[9] Derrida B., *Phys. Rev. B*, **24** (1981) 2613.
[10] Saakian D.B., *Pis'ma Zh. Éksp. Teor. Fiz.[JETP Lett.]*, **67** (1998) 440.
[11] Dorlas T.C. and Wedagedera J.R., *Phys. Rev. Lett.*, **83** (1999) 4441.
[12] Kabashima Y. and Saad D., *Europhys. Lett.*, **44** (1998) 668; *Europhys. Lett.*, **45** (1999) 97.
[13] Vicente R., Saad D. and Kabashima Y., *Phys. Rev. E*, **60** (1999) 5352.
[14] Kabashima Y., Murayama T. and Saad D., *Phys. Rev. Lett.*, **84**(2000) 1355.
[15] Murayama T. et al. , *Phys. Rev. E*, **62** (2000) in press.
[16] Vicente R., Saad D. and Kabashima Y., *J. Phys. A*, (2000) in press.
[17] Montanari A., Sourlas N., cond-mat/9909018 Preprint, 1999; Montanari A., cond-mat/0003218 Preprint, 2000.
[18] Dress C., Amic E. and Luck J.M., *J. Phys. A*, **28** (1995) 135.
[19] Rieger H. and Kirkpatrick T.R., *Phys. Rev. B*, **45**(1992) 9772.
[20] Kschischang F.R. and Frey B.J., *IEEE J. Select. Areas in Comm.*, **16**(1998) 153.
[21] Aji S.M. and McEliece R.J., *IEEE Trans. Info. Theory*, **46**(2000) 325.
[22] Gujrati P.D., *Phys. Rev. Lett.*, **74**(1995) 809.
[23] Bowman D.R. and Levin K., *Phys. Rev. B*, **25** (1982) 3438.
[24] Richardson T. and Urbanke R., Preprint, 1998.
[25] Weiss Y., *Neural Computation*, **12** (2000) 1.