Applications of correlation inequalities to low density graphical codes

N. Macris[a]

Laboratoire de Théorie des Communications, École Polytechnique Fédérale de Lausanne, Station 14 - LTHC -EPFL,
1015 Lausanne, Switzerland

Abstract. This contribution is based on the contents of a talk delivered at the Next-SigmaPhi conference held in Crete in August 2005. It is adressed to an audience of physicists with diverse horizons and does not assume any background in communications theory. Capacity approaching error correcting codes for channel communication known as Low Density Parity Check (LDPC) codes have attracted considerable attention from coding theorists in the last decade. Surprisingly strong connections with the theory of diluted spin glasses have been discovered. In this work we elucidate one new connection, namely that a class of correlation inequalities valid for Gaussian spin glasses can be applied to the theoretical analysis of LDPC codes. This allows for a rigorous comparison between the so called (optimal) maximum a posteriori and the computationaly efficient belief propagation decoders. The main ideas of the proofs are explained and we refer to recent works for the more lengthy technical details.

1 Codes for communication through noisy channels

We consider a (simplified) communication system with three basic building blocks: the *encoder*, the *channel* and the *decoder*.

Encoder. Suppose that messages to be sent are labelled $\{1, ..., M\}$ and that $M = 2^K$. The messages can be represented by binary strings of length K, so that if a message is sent K *information bits* are transmitted. Because of channel imperfections these binary strings are encoded before they are fed into the channel. In general the encoder is a map $\mathbb{F}_2^K \to \mathbb{F}_2^N$, with $\mathbb{F}_2 = \{0, 1\}$ and $N > K$. So the codebook consists of 2^K code words that are binary strings of length N, $(x_1, ..., x_N) = \mathbf{x}$. In order to send K information bits we make N uses of the channel: one says that the rate of transmission is $R = \frac{K}{N}$.

Channel. We take a discrete (binary input) memoryless channel with general output alphabet (for example $\mathbb{F}_2$ or $\mathbb{R}$). Given a sent codeword $(x_1, ..., x_N)$ the received word is $(y_1, ..., y_N) = \mathbf{y}$ with probability $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{N} p_{Y|X}(y_i|x_i)$. In this context the choice of the transition probability $p_{Y|X}$ specifies the model for the channel and is supposed to be known to the sender and the receiver.

[a] e-mail: nicolas.macris@epfl.ch

Decoder. Given that $\mathbf{x}^{in}$ is sent, the receiver possesses a deformed version $\mathbf{y}$ (the channel observations or the channel output) and his task is to find estimates $D(\mathbf{y})$ so that the bit probability of error $P_{error}((D(\mathbf{y}))_i \neq x_i^{in})$ is as small as possible. One can show that the best decoder (the one which gives the smallest probability of error) is given by the Maximum a Posteriori (MAP) estimator $(\hat{x}_i)_{MAP} = \text{argmax}_{x_i} p_{X|\mathbf{Y}}(x_i|\mathbf{y})$ Unfortunately this cannot be computed efficiently and other suboptimal estimators must be considered. Of course it is important to compare their relationship and performance to the MAP estimator. This problem is adressed here for LDPC codes and the suboptimal estimator given by Belief Propagation (BP).

Shannon's *noisy channel coding theorem* asserts that one can communicate reliably as long as the rate R is smaller than the *channel capacity* $C = \max_{p_{\mathbf{x}}} I(X; Y)$. In this formula $I(X; Y)$ is the *mutual information* between random variables X and Y which can be interpreted as the information gained about X given that Y is observed. The maximization over the prior distribution of the codewords p_X corresponds to finding the best possible codebook. In formulas, $I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$, where the Shannon entropy of X is $H(X) = -\sum_x p_X(x) \ln p_X(x)$ and the conditional entropy $H(X|Y) = -\sum_{x,y} p_Y(y) p_{X|Y}(x|y) \ln p_{X|Y}(x|y)$ and similarly for X and Y exchanged. All marginals are computed from $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = p_{\mathbf{X}}(\mathbf{x}) p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$. Thus C is a
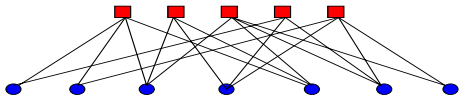
Fig. 1. A Tanner graph. Check nodes on the top row constrain the bits attached to variable nodes on the bottom row.

functional of the channel transition probability. Moreover there is no way to communicate reliably when $R > C$.

More precisely, let $R \leq C - \epsilon$ where $\epsilon > 0$ is as small as we wish. There exists an $N_0(\epsilon)$ such that for each $N > N_0(\epsilon)$ we can find encoding and decoding maps (the optimal decoder does the job) such that $P_{error} < \epsilon$. Conversely, if $R \geq R + \epsilon$ for any N and any encoding map $P_{error} \geq p_0 > 0$ for some p_0 independent of N.

For our purpose it is more convenient to fix a desired rate R once for all and translate the inequality $R < C$ as a condition on the channel noise $n < n_{sh}$ where n_{sh} is a (channel dependent) function of R. This means we can reliably transmit at rate R as long as the channel noise is lower than the Shannon threshold n_{sh}.

Shannon's theorem is not constructive in the sense that it guarantees the existence of an encoder in an ensemble of random codes, but does not allow to construct "good" (capacity approaching and computationaly efficient) encoders and decoders. One of the main themes of information and coding theory for the last fifty years has been to precisely define and address such questions. A fruitful idea is to restrict the encoder maps to the class of linear error correcting codes. Remarkably Shannon's theorem is still true if one restricts to the class of linear encoders and there is no loss in capacity. For more details we refer the reader to [1].

For us a linear code is a vector subspace of $\mathbb{F}_2^N$ of dimension $K < N$. The subspace can be defined as the kernel of a parity check $M \times N$ matrix H with $N - M = K$. In other words the set of code words satisfy M constraints (so called parity checks)

$$\sum_{k=1}^{N} H_{lk} x_k = 0 \bmod 2, \ l = 1, ..., M, \ H_{lk} = 0, 1 \quad (1)$$

Note that the rate of the code is $R = \frac{K}{N} = 1 - \frac{M}{N}$. A very useful graphical representation of a linear code is in terms of the Tanner graph (or factor graph). This is a bipartite graph with variable nodes $i \in \{1, ..., N\}$, check nodes $A \in \{1, ...M\}$, and edges connecting variable and check nodes. We say that a variable node i "belongs" to a check node A, $i \in A$, if and only if it appears in the parity check equation labeled by A. In this case an edge connects i and A (see Fig. 1). Low Density Parity Check (LDPC) codes are a special class of linear codes with sparse Tanner graphs: the degrees (or coordination number) of check and variable nodes are of $O(1)$ with respect to N. For such codes there is still a threshold phenomenon as in Shannon's theorem however in general the maximal rate at which error free communication is possible is below Shannon's capacity. On the other hand suboptimal but computationaly efficient decoding algorithms exist.

2 Low Density Parity check codes as diluted spin glasses

The close connection of the above formalism to random spin systems was first noticed by Sourlas [2]. While this connection is quite general and not limited to binary alphabets, memoryless channels and linear codes, here we rephrase it in the case of low density parity check codes. If code word bits are represented by spins through the mapping $s_i = (-1)^{x_i}$, the parity check equations (1) become

$$\frac{1}{2}(1 + s_A) = 1, \qquad s_A = \prod_{i \in A} s_i, \qquad A = 1, ..., M. \quad (2)$$

The a posteriori probability distribution used in MAP decoding is nothing else than the Gibbs measure of a spin system where the spins are attached to variable nodes while check nodes are a convenient way to represent their many-body interactions. By Bayes rule

$$p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{1_{\mathcal{C}}(\mathbf{x}) \prod_{i=1}^{N} p_{Y|X}(y_i|x_i)}{\sum_{\mathbf{x}} 1_{\mathcal{C}}(\mathbf{x}) \prod_{i=1}^{N} p(y_i|x_i)}. \quad (3)$$

This is a Gibbs Measure $\langle - \rangle_{\mathcal{C}} = \frac{e^{-H_{\mathcal{C}}}}{Z_{\mathcal{C}}}$ with hamiltonian

$$H_{\mathcal{C}} = -\sum_{A \in \mathcal{C}} J_A(s_A - 1) - \sum_{i=1}^{n} h_i s_i, \qquad s_A = \prod_{i \in A} s_i \quad (4)$$

where $J_A = +\infty$ and $h_i = \frac{1}{2} \ln \frac{p(y_i|0)}{p(y_i|1)}$. The channel observations enter through a quenched random magnetic field h_i whose distribution is induced by the distribution of channel observations. It can be shown that for symmetric channels (these satisfy $p(y|x) = p(-y| - x)$) there is no loss in generality to assume that the input word is $(x_1^{in} = 0, ..., x_N^{in} = 0)$, so that the distribution of channel observations is $\prod_{i=1}^{N} p(y_i|0)$. Another source of quenched randomness is given by the Tanner graph (defining the coupling constants J_A) which is taken from an ensemble of random graphs. Since our results are independent of the choice of this ensemble we do not discuss their construction in detail. Let us point out that the performance of a particular coding scheme depends on the choice of the ensemble. The expectation value with respect to the channel observations and the graphs are denoted $\mathbb{E}_{\mathcal{C},h}$.

The MAP decoding rule becomes

$$(\hat{s}_i)_{MAP} = \text{sign} \langle s_i \rangle_{\mathcal{C}} \quad (5)$$

and the average bit probability of error for the optimal decoder is basically the overlap of $(\hat{s}_1, ..., \hat{s}_N)$ with the fully ferromagnetic configuration $(1, ..., 1)$ (or the sent codeword)

$$P_{error} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{\mathcal{C},h} \left[1 - \text{sign} \langle s_i \rangle_{\mathcal{C}} \right]. \quad (6)$$

The replica or cavity methods can be applied to the calculation of such quantities and show that a phase transition occurs [3]. Namely there is a threshold n_{MAP} such
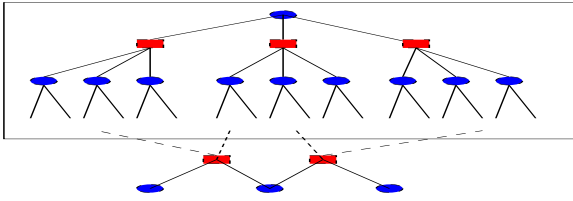
Fig. 2. Tree like neighborhood T_o of an arbitrary root node o. The loops are of size $O(N)$ with high probability.

that for $n < n_{MAP}$ the probability of error goes to zero in the thermodynamic limit (ferromagnetic phase), while for $n > n_{MAP}$ the probability of error is bounded away from zero. Sparse graphs are locally tree like in the sense that the typical size of loops is $O(N)$ and have no boundary. Hence it is reasonable to expect that mean field approaches such as the replica or cavity methods yield exact results. This is for the moment unproven although some progress in this direction has been made by the use of interpolation methods [6,7]. As explained below our use of correlation inequalities yields closely related results.

3 Efficient decoding

Although one can optimize the degrees of the Tanner graphs in order that n_{MAP} approaches n_{sh}, MAP decoding is computationally too expensive. However one can take advantage of the fact that low density graphs are locally tree like (see Fig. 2). Consider a specified root node o and its neighborhood of depth d. As long as $d = O(1)$ with respect to N this neighborhood is a tree with high probability. Thus one can expect that a good approximation is obtained by neglecting the loops and solving for the magnetization of the spin system on a tree. The sign of the magnetization on the tree defines the *Belief Propagation* (BP) estimate

$$(\hat{s}_o)_{BP} = \operatorname{sign} \tanh\left(h_o + \sum_{C \in o} u_{C \to o}^{(d-1)} \right). \qquad (7)$$

In this formula the fields $u_{C \to i}$ are computed from the iterative procedure

$$u_{C \to i}^{(t+1)} = \tanh^{-1} \prod_{j \in C \setminus i} \tanh h_{j \to C}^{(t)},$$

$$h_{i \to A}^{(t+2)} = h_i + \sum_{C \in V(i) \setminus A} u_{C \to i}^{(t+1)} \qquad (8)$$

with the initial conditions $h_{i \to C}^{(0)} = h_i$.

The belief propagation decoding algorithm is an iteration based on these exchanges of *messages* $u_{C \to i}$ from checks to variables and *messages* $h_{j \to C}$ from variables to checks. It is applied to the full Tanner graph and despite the presence of loops it converges and successfully decodes for $n < n_{BP}$. The relationship between the various thresholds is $n_{BP} < n_{MAP} < n_{sh}$. It should be clear that this

algorithm is closely related to the cavity equations of spin glass theory.

One of the main problems in the theory of LDPC codes is to optimize the codes so that the various thresholds come as close as possible to n_{sh}. A more basic problem is to compare the error probabilities given by the BP and MAP decoders. While this is difficult in general we show below how these decoders can be compared for closely related quantities - *the generalized EXIT curves* - through the use of correlation inequalities.

4 Correlation inequalities

Here we restrict ourselves to the case of the binary input additive white Gaussian noise channel (BIAWGNC) where the results are more transparent. Mathematically the channel is defined as $y_i = x_i + W_i$, W_i i.i.d $\mathcal{N}(0, n)$. Then the log-likelihood ratio (or magnetic field) $h_i = \frac{1}{2} \ln \frac{p(y_i|0)}{p(y_i|1)}$ has a *Gaussian* distribution with equal mean and variance $\mathbb{E}_h[h_i] = \mathbb{V}_h[h_i] = n^{-1/2}$. We soften the parity check constraints from $J_A = +\infty$ to independent Gaussian random variables with equal mean and variance $\mathbb{E}_J[J_A] = \mathbb{V}_J[J_A] = t_A$. The case of hard constraints (the parity checks) is recovered by making $t_A \to +\infty$. With soft random constraints the hamiltonian is a *Gaussian spin glass* with Nishimori gauge symmetry. Contucci, Graffi and Nishimori proved for such systems the following set of inequalities hold [8]

$$\mathbb{E}_J[\langle s_X \rangle] \geq 0, \qquad \frac{\partial}{\partial t_Y} \mathbb{E}_J[\langle s_X \rangle] \geq 0,$$
$$\text{any} \qquad X, Y \subset \{1, ..., N\}. \quad (9)$$

The reader will recognize the close similarity to the famous Griffith-Kelly-Sherman correlation inequalities valid for fully ferromagnetic systems.

This inequality can be applied to compare the magnetization on the initial Tanner graph and on a tree graph. In the coding context this allows a comparison between MAP and BP decoders. Consider the Gibbs measure defined by the Gaussian spin glass hamiltonian with some set of variances t_A, $A = 1, ..., M$. The neighborhood T_o of o (see Fig. 3) is a tree with probability $(1 - O(\frac{k^d}{N}))$ where k is a constant related to the maximal degree of the nodes. The second correlation inequality implies that, if for the checks outside of T_o we decrease t_A to zero, the average magnetization of site o decreases. This inequality is preserved if we increase t_A to infinity for the checks inside T_o. In other words

$$\mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}] \geq \mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{T_o} | T_o \text{ is a tree}] \Pr(T_o \text{ is a tree}). \tag{10}$$

The right hand side should also include a contribution coming from the probability that T_o is not a tree but by the first correlation inequality it is positive so that we can omit it. We refer to this procedure as the "check erasing" (see Fig. 3 for a pictorial illustration of check erasing). On the tree graph the statistical mechanical sums can be
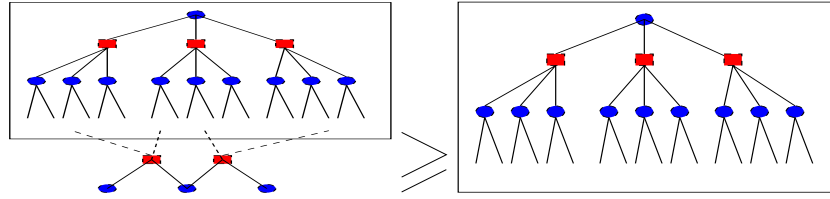
Fig. 3. A pictorial representation of the check erasing inequality.

performed exactly and yield in a natural way the Belief Propagation algorithm of the previous section. So

$$\mathbb{E}_{\mathcal{C},h}\left[\langle s_o \rangle_{\mathcal{C}}\right] \geq (\hat{s}_o)_{BP}\left(1 - O\left(\frac{k^d}{N}\right)\right) \qquad (11)$$

Finally one can take the thermodynamic limit $N \to +\infty$ and then the limit $d \to +\infty$. While on the right hand side these limits can be shown to exist, the existence of the thermodynamic limit for the left hand side is an open problem. Thus we really take the $\liminf_{N \to \infty}$.

5 Generalized EXIT curves

The probability of error (6) is technically cumbersome to handle. Another quantity called in coding theory the "extrinsic information transfer" is more convenient to study. It yields the same thresholds as the error probability and from the statistical mechanical perspective it is much more natural as will become clear below. Here we define the generalized EXIT curve associated to MAP decoding as [5]

$$g_{MAP}(n) = \liminf_{N \to \infty} \frac{1}{N} \frac{d}{dn} \mathbb{E}_{\mathcal{C}}\left[H\left(X_1, ..., X_N | Y_1, ..., Y_N\right)\right]. \tag{12}$$

The conditional entropy of the a posteriori distribution is nothing else than the average entropy of the Gibbs distribution for the spin glass. It should not come as a surprise that this can be related to the free energy

$$\mathbb{E}_{\mathcal{C}}\left[H(\mathbf{X}|\mathbf{Y})\right] = \mathbb{E}_{\mathcal{C},h}\left[\ln Z_{\mathcal{C}}\right] - \sum_{i=1}^{N} \mathbb{E}_{\mathcal{C},h}\left[h_i \langle s_i \rangle_C\right]. \tag{13}$$

In the case of a BIAWGNC the derivative with respect to the noise has a simple relation to the magnetization. This is not obvious a priori because the channel noise does not enter like an external field and for more general channels the corresponding relation is more complicated. The derivation of (13) is too lengthy to show here but let us note that the main point is to use Nishimori identities [4]

$$g_{MAP}(n) = \liminf_{N \to +\infty} \frac{1}{2n^3 N} \sum_{i=1}^{N} \mathbb{E}_{\mathcal{C},h}\left[1 - \langle s_i \rangle_{\mathcal{C}}\right]$$

$$= \frac{1}{2n^3} \mathbb{E}_{\mathcal{C},h}\left[1 - \langle s_o \rangle_{\mathcal{C}}\right], \qquad \text{any } o. \tag{14}$$

The following lemma shows that $g_{MAP}(n)$ and P_{error} have the same threshold.

Lemma. *Assume communication through a BIAWGNC with noise n and an ensemble of linear codes. We have that $g_{MAP}(n) = 0$ if and only if $P_e = \lim_{N \to +\infty} P_{error} = 0$.*

To show that $g_{MAP}(n) = 0$ implies $P_e = 0$ we note that if $1 = \mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}]$ then $\mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}^2] - \mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}]^2 = 0$ because of the Nishimori identity $\mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}^2] = \mathbb{E}_{\mathcal{C},h}[\langle s_o \rangle_{\mathcal{C}}]$. Thus the random variable $\langle s_o \rangle_{\mathcal{C}}$ does not fluctuate and equals 1 almost surely. Thus $\text{sign}\langle s_o \rangle_{\mathcal{C}} = +1$ and $P_e = 0$. For the converse we combine Fano's inequality [1] together with Jensen to get $0 \leq \frac{1}{N} H(\mathbf{X}|\mathbf{Y}) \leq h(P_{error})$ where h is the binary entropy function. Thus $\lim_{N \to +\infty} \frac{1}{N} H(\mathbf{X}|\mathbf{Y}) = 0$. If this is true for a whole range of n we can conclude $g_{MAP}(n) = 0$.

Combining (10) and (14) we obtain [11,10]

Theorem. *Assume communication through a BIAWGNC with noise n and an LDPC ensemble of codes. Then*

$$g_{MAP}(n) \geq \lim_{d \to +\infty} \frac{1}{2n^3} \mathbb{E}_{h,l,u_1,...,u_l}$$

$$\times \left[1 - \tanh\left(h + \sum_{c=1}^{l} u_c^{(d)}\right)\right] \tag{15}$$

where the right hand side is computed from the BP algorithm and defines the generalized EXIT curve associated to the BP decoder, $g_{BP}(n)$. The p.d.f of h is Gaussian with mean and variance $n^{-1/2}$, l is the random degree of variable nodes, the distribution of u_i is induced by the message passing algorithm.

Such bounds and the method used here extends to the class of (smooth) binary input symmetric channels [12]. These bounds have also been derived recently by the method of physical degradation [5,13]. To conclude we briefly discuss a number of consequences of the theorem.

General picture. In general the BP and MAP curves may have several discontinuities corresponding to several phase transitions in the spin glass. In the simplest (non trivial) case where there is only one discontinuity their behavior is as follows. For $0 < n < n_{BP}$ $g_{BP}(n) = 0$, there is a jump discontinuity at n_{BP} and for $n > n_{BP}$ $g_{BP}(n)$ is strictly positive. The same occurs for g_{MAP} but with the jump discontinuity at n_{MAP} and $n_{MAP} > n_{BP}$. Moreover the BP curve is always under the MAP curve.

Bound on MAP threshold. From the definition of the MAP generalized EXIT curve we see that

$$\int_{n_{MAP}}^{+\infty} g_{MAP}(n)dn = \liminf_{N\to\infty} \frac{1}{N} \left(H(\mathbf{X}|\mathbf{Y})|_{n=+\infty}\right.$$

$$\left. -H(\mathbf{X}|\mathbf{Y})|_{n=n_{MAP}}\right) = R. \quad (16)$$

Indeed for infinite noise we have no knowledge of the sent signal (so the conditional entropy is R) and just below the MAP threshold we have perfect knowledge (the conditional entropy is zero). The theorem then implies

$$R > \int_{n_{MAP}}^{+\infty} g_{BP}(n)dn \quad (17)$$

where the right hand side can be computed numerically. This then yields a lower bound on the MAP threshold. Numerical evaluations tend to show that this bound is tight which suggests that above the MAP threshold the BP and MAP curves should coincide [9].

Bounds on the conditional entropy. It is possible to obtain bounds on the conditional entropy itself by integration of the inequality (15). Let us set $h(\mathbf{X}|\mathbf{Y}) = \liminf_{N\to+\infty} \frac{1}{N} H(\mathbf{X}|\mathbf{Y})$. Integrating from 0 to n we get

$$h(\mathbf{X}|\mathbf{Y}) \geq \int_0^n g_{BP}(n)dn \quad (18)$$

and integrating from n to $+\infty$,

$$h(\mathbf{X}|\mathbf{Y}) \leq R - \int_n^{+\infty} g_{BP}(n)dn =$$

$$\int_0^n g_{BP}(n)dn + \left(R - \int_0^{+\infty} g_{BP}(n)dn\right). \quad (19)$$

In the case where there is no phase transition one can show that $R = \int_0^{+\infty} dn g_{BP}(n)$ so that we get an exact expression for the conditional entropy and its derivative satisfies $g_{MAP}(n) = g_{BP}(n)$. We have a situation where the model is exactly solved and the result of the cavity method (or replica symmetric expression) is proved to be exact. However there is no fully polarized phase and no error free communication. When there is one (or many)

phase transition the parenthesis in the last the right hand side of (19) is strictly positive so that the two bounds for $h(\mathbf{X}|\mathbf{Y})$ do not match. However it is believed that the upper bound (18) is tight above the MAP threshold because it coincides with the result of the replica symmetric calculation. The same bound has been obtained [7] using the interpolation methods developed by Guerra [6] for the Sherrington-Kirkpatrick model. Clearly, it would be desirable to prove the converse inequality.

References

1. R.B. Ash, *Information Theory* (Dover Publications, 1990)
2. N. Sourlas, Nature **339**, 693 (1989); in *From Statistical Physics to Statistical inference and back*, edited by. P. Grassberger, J.P. Nadal (Kluwer Academic, 1994), p. 195; in Mathematical Results in Statistical Mechanics, edited by S. Miracle-Sole, J. Ruiz, V. Zagrebnov (World Scientific, 1999), p. 475
3. A. Montanari, Eur. Phys. J. **23** (2001)
4. H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford Science Publications, 2001)
5. R. Urbanke, T. Richardson, in *Modern Coding Theory* (Cambridge University Press, in preparation)
6. F. Guerra, F. Toninelli, J. Math. Phys. **43**, 3704 (2002)
7. A. Montanari, *Tight bounds for LDPC and LDGM codes under MAP decoding*, preprint (2004)
8. S. Morita, H. Nishimori, P. Contucci, J. Phys. A **37**, L203 (2004)
9. C. Méasson, R. Urbanke, An upper-bound for the ML threshold of iterative coding systems over the BEC, *Proc. of the 41st Allerton Conference on Communications, Control and Computing, Allerton House, Monticello, USA, October 2003*, p. 3
10. N. Macris, Correlation inequalities: a useful tool in the theory of LDPC codes, *Proc. IEEE int symp inf theory, Adelaide, September (2005)*
11. N. Macris, *Griffith-Kelly-Sherman correlation inequalities: a useful tool in the theory of error correcting codes*, preprint (2005)
12. N. Macris, *Sharp bounds on generalized EXIT functions*, preprint (2005)
13. C. Méasson, A. Montanari,and R.Urbanke, *Proc of IEEE Int. Symp Inf. Theory, Adelaide, Australia* (2005)