Griffith–Kelly–Sherman Correlation Inequalities: A Useful Tool in the Theory of Error Correcting Codes

Nicolas Macris, *Member, IEEE*

***Abstract*—It is shown that a correlation inequality of statistical mechanics can be applied to linear low-density parity-check codes. Thanks to this tool we prove that, under a natural assumption, the exponential growth rate of regular low-density parity-check (LDPC) codes, can be computed exactly by iterative methods, at least on the interval where it is a concave function of the relative weight of code words. Then, considering communication over a binary input additive white Gaussian noise channel with a Poisson LDPC code we prove that, under a natural assumption, part of the GEXIT curve (associated to MAP decoding) can also be computed exactly by the belief propagation algorithm. The correlation inequality yields a sharp lower bound on the GEXIT curve. We also make an extension of the interpolation techniques that have recently led to rigorous results in spin glass theory and in the SAT problem.**

***Index Terms*—Correlation inequalities, density evolution, generalized EXIT (GEXIT) curve, growth rate, interpolation technique, iterative decoding, low-density parity-check (LDPC) codes, spin glasses.**

I. Introduction

THERE is a deep connection between the theory of linear error correcting codes and statistical mechanics of random spin systems (spin glasses). This connection was first uncovered by Sourlas [1] and was later applied to various coding schemes such as convolutional, turbo, low-density parity-check (LDPC) codes. In particular the replica method of spin glass theory, has been applied to LDPC ensembles and its intimate connection to density evolution equations and belief propagation algorithms has been recognized [2]–[4]. Giving a sound mathematical basis to the results of the replica method has been a long standing problem of spin glass theory, but recently progress in this direction has been accomplished by Guerra, which resulted in the so called interpolation techniques for the Sherrington–Kirkpatrick spin model [5]–[7]. These interpolation techniques have been succesfully applied to the satisfiability (SAT and XORSAT) problems [8], [9] and LDPC codes [10].

In statistical mechanics a very powerful tool is often provided by correlation inequalities: in this paper we demonstrate that this is also the case in coding. The correlation inequalities that we investigate in the context of coding are the Griffith–Kelly–Sherman (GKS) inequalities [15]. These have been known since a long time in the context of ferromagnetic nonrandom Ising type models and have been extended recently to the situation of spin glasses provided a certain "gauge symmetry" is present [16]. It will become clear later that this symmetry is implied by memoryless channel symmetry.

Instead of investigating the most general situation, here we limit ourselves to two special LDPC ensembles. We study two different problems: one concerns the exponential growth rate for regular codes; while the other one pertains to communication through a noisy channel. For each problem we demonstrate that: 1) the GKS inequalities are applicable, and 2) appropriate quantities can be calculated exactly by message passing algorithms. Let us stress that from the point of view of spin glass theory 2) amounts to show the exactness of the replica symmetric solution in a suitable range of parameters (along the so-called Nishimori line).

A summary of the present work has appeared in [17].

A. Growth Rate of Regular LDPC Codes

We study the growth rate of regular codes with arbitrary variable node degree l and even check node degree k. Our main result rests on an unproven assumption (called H1 in Section III) which is however very natural in statistical mechanics of spin systems, namely that a grand canonical free energy can be represented as the Legendre transform of a canonical one. The main result states that under this assumption, if on a certain interval the growth rate is a concave function of the relative weight of the codewords, then at least on part of that interval iterative message passing methods are exact (Theorem 3.3).

This is achieved through upper and lower bounds on the growth rate. For the upper bound we use the known fact that the combinatorial growth rate [27], [28] is an upper bound which happens to be sharp. For irregular ensembles it is known [29], [30] that such a bound is not sharp and this is essentially why we limit ourselves to regular codes. To obtain the lower bound we remark that the weight enumerator is the partition function of a ferromagnetic spin model so that the GKS inequality applies (see Lemma 3.1, at this point we need an even degree for the check nodes).

One would like to extend the main result to all values of the relative weight (i.e., to go beyond the concave region). Indeed numerical examples show that the curvature changes from concave to convex before the point where the growth rate vanishes

Manuscript received April 26, 2005; revised June 29, 2006. The material in this paper was presented in part at IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

The author is with the Ecole Polytechnique Fédérale de Lausanne, EPFL, CH-1015 Lausanne, Switzerland (e-mail: nicolas.macris@epfl.ch).

Communicated by T. J. Richardson, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2006.889002

(see, for example, Fig. 3), and therefore our result does not, unfortunately, allow for a rigorous computation of the typical minimum distance of the code ensemble (except for the trivial case where the left degree is equal to 2). We hope that our analysis is a first step toward this goal. Second, it would obviously be desirable to have a proof of assumption H1.

B. GEXIT Curves

We consider the problem of communication with a Poisson LDPC code through a binary input additive white Gaussian noise channel (BIAWGNC). We prove that a "Generalized EXIT curve"—defined as the derivative of the conditionnal entropy with respect to the inverse square noise—and denoted GEXIT, can be computed exactly by message passing algorithms at least for some range of noise values (Theorem 3.6). Here this result is conditional on an assumption (called H2 in Section III) which basicaly means that there are a finite set of noise values where the "number of density evolution fixed points can jump" and away from these "singularities" the replica symmetric functionnal is a differentiable function of the noise.

Such a result has been derived recently for the binary erasure channel (BEC) [20]–[22] by a combination of the notion of physical degradation and the area theorem [23].

Here we first prove a sharp lower bound to the GEXIT curve thanks to the GKS inequalities extended to random spin systems with channel symmetry (Lemma 3.4). This was in fact claimed in [21] (and proved by physical degradation for the binary symmetric channel), and then proved in full generality recently[1] in [11].

Next we prove a lower bound on the conditional entropy by adapting the interpolation method, as presented in [8], [9] for the SAT and XORSAT problems, to the case of LDPC codes. This case has been treated recently [10] for standard irregular ensembles ensembles having a generating function for the check node degree distribution satisfying a certain convexity requirement (for example if the checks have regular degree then it has to be even). One of our contributions is to prove the bound in the case of a Poisson distribution for the variable nodes and check nodes with any even or odd degree.

Generalized EXIT curves have been introduced recently in [21]. The first one denoted GEXIT is simply the derivative of the Shannon conditional entropy (of the input conditioned on the output) with respect to the channel entropy (or the noise parameter). In the special case of the BEC it is shown [20], with the help of the Area Theorem, that this is the same as the usual EXIT curve (defined as the average entropy of the ith input bit conditionned on the output bits except the ith one). For more general channels it turns out that the usual EXIT and GEXIT curves are numericaly very close. This GEXIT curve is said to be associated to MAP decoding because it involves the knowledge the probability distribution of the input conditionned on the output. In [11], [21] the authors define other "Generalized EXIT" curves associated to iterative decoding such as belief propagation. Moreover they provide a general bound stating that the MAP GEXIT curve always stays below the iterative one. Our result says that for the Poisson code ensemble GEXIT curves associated to MAP decoding and to belief propagation decoding are equal for some range of noise values. This range corresponds to noise values below the first discontinuity (if there is one). Note that here the curves do not trivialy vanish in this range because the Poisson ensemble does not have a MAP threshold. On the other hand, for "good" codes in that same range the EXIT curves typicaly vanish and a statement such as that of Theorem 3.6 would be of limited interest. However, a generalization of the intermediate lemmas and techniques would still be of interest.

C. Two Useful Identities

Another feature of this work is the use of two identities relating derivatives of the conditionnal entropy and bit correlations (Lemma 3.7). The first one is also closely related, but slightly different, to the relationship between mutual information and MMSE [24]. The second identity appears to be new in this context to the best of our knowledge. From the point of view of statistical mechanics it is a kind of fluctuation theorem: the left-hand side is a kind of "suceptibility" while the right-hand side is a kind of "correlation function." Because of these two identities which rely on a Gaussian channel, our result on GEXIT curves is limited to the Gaussian case. However we think that it is much more general and support for this conjecture comes from the fact that it is already known in the case of a BEC. We believe the above results should extend to general memoryless symmetric channels.

The paper is organized as follows: in Section II we formulate the problems to be studied in the language of spin systems, in Section III we state our main results and in Section IV we review and adapt the GKS inequalities. Sections V–VII are dedicated to the proofs of the main results. The proofs of some intermediate results are presented in the appendices.

II. LDPC Codes as Random Spin Systems

We consider two ensembles of codes, namely the regular LDPC(n,l,k) and Poisson LDPC(n,r,k) ensembles.

The regular ensemble is defined in a standard way through random bipartite graphs: the Tanner or factor graphs of the codes [25], [26]. We have n variable nodes of degree l labeled $i = 1,\dots,n$ connected to m check nodes of degree k labeled $c = 1,\dots,m$. The constraint $nl = mk$ must be satisfied and the design rate is fixed $r = 1 - \frac{l}{k}$. This ensemble of graphs is endowed with the uniform probability distribution.

In the Poisson ensemble we first fix a design rate $0 < r < 1$ and a number n of variable nodes. The number of check nodes m is a Poisson random variable of mean $n(1-r)$. There are k edges emanating from each check node and each edge is connected with uniform probability $\frac{1}{n}$ to a variable node. Given m, the probability that a variable node has degree l is $\binom{m}{l}\left(\frac{k}{n}\right)^l\left(1-\frac{k}{n}\right)^{m-l}$. In the limit where $n \to \infty$ this tends to a Poisson distribution of mean $k(1-r)$. Clearly, the Poisson ensemble as defined here does not provide a good code since there is a finite probability that a variable node is unconstrained. However from

[1] let us note that in the present paper the GEXIT is defined through a derivative with respect to the inverse noise in contrast to [11], [21] where it is defined through a derivative with respect to the noise itself. As a consequence inequalities are reversed. The method of correlation inequalities can also be extended to general symmetric channels [12], [13].

a technical point of view this ensemble is sufficiently simple that progress can be made toward rigorous results.

Each graph in the ensemble has an $m \times n$ adjacency matrix H whose matrix elements H_{ci} are equal to the number of edges $(\mathrm{mod} 2)$ connecting nodes i and c. Code words $x^n = (x_1, \ldots, x_n) \in \mathbf{F}_2^n$ satisfy m parity-check constraints

$$\sum_{i=1}^{n} H_{ci} x_i = 0, \qquad \mathrm{mod}\ 2 \qquad c = 1, \ldots, m.$$

In the "spin" language of statistical mechanics each bit is represented as a "spin" $s_i = (-1)^{x_i}$ taking values ± 1. The code words $s^n = (s_1, \ldots, s_n)$ then satisfy the constraints

$$\prod_{i \in \partial c} s_i = 1, \qquad c = 1, \ldots, m$$

where ∂c denotes the set of variable nodes that are adjacent to the check node c. It will be convenient to use the notation $s_X = \prod_{i \in X} s_i$ where X is any subset of $\{1, \ldots, n\}$. Then, given a factor graph and its associated code, s^n is a code word if and only if

$$\prod_{c=1}^{m} \frac{1}{2}(1 + s_{\partial c}) = 1. \tag{1}$$

A. Exponential Growth Rate

In terms of spin variables the relative weight of a code word is

$$w(x^n) = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{2}\left(1 - \frac{1}{n} \sum_{i=1}^{n} s_i\right).$$

It will be convenient to define

$$\omega(s^n) = \frac{1}{n} \sum_{i=1}^{n} s_i$$

so that $w(x^n) = \frac{1}{2}(1 - \omega(s^n))$.

We introduce the number of codewords with relative weight ω (for a given code in the ensemble LDPC(n, l, k))

$$A_n(\omega) = \sum_{s^n} \delta_{n\omega, n\omega(s^n)} \prod_{c=1}^{m} \frac{1}{2}(1 + s_{\partial c}) \tag{2}$$

and the generating function or weight enumerator

$$Z_n(h) = \sum_{\omega} A_n(\omega) e^{hn\omega} = \sum_{s^n} e^{h \sum_{i=1}^{n} s_i} \prod_{c=1}^{m} \frac{1}{2}(1 + s_{\partial c}). \tag{3}$$

In the language of statistical mechanics these two objects can be interpreted as partition functions of a spin system with a hard core interaction. The former (2) is the partition function in the "canonical ensemble" with fixed "magnetization per spin" ω, while the later (3) is the partition function in the "grand canonical ensemble" with fixed "magnetic field" h. By hard core interaction we mean the fact that the constraint (1) can be viewed as a Gibbs weight

$$\prod_{c=1}^{m} \frac{1}{2}(1 + s_{\partial c}) = \lim_{J_c \to \infty; c=1,\ldots,m} \exp(-H(s^n)) \tag{4}$$

where the "Hamiltonian"

$$H(s^n) = \sum_{c=1}^{m} J_c(1 - s_{\partial c}) \tag{5}$$

has infinitely large coupling constants $J_c \to +\infty, c = 1, \ldots, m$. The representation (4), (5), although not really necessary, will prove insightful in Section IV.

The growth rate of a code

$$g_n(\omega) = \frac{1}{n} \ln A_n(\omega)$$

is nothing else but the "canonical potential or free energy." The "grand canonical potential" (also called "pressure" or "free energy" depending on the interpretation; we adopt the later terminology) is the logarithm of the weight enumerator,

$$f_n(h) = \frac{1}{n} \ln Z_n(h). \tag{6}$$

We will be interested in upper and lower bounds for the expected value $\mathbf{E}_C[g_n(\omega)]$ over the code ensemble LDPC(n, l, k). For the upper bound we simply use Jensen's inequality

$$\mathbf{E}_C[g_n(\omega)] \leq \frac{1}{n} \ln \mathbf{E}_C[A_n(\omega)] \tag{7}$$

and the fact that the combinatorial growth rate on the right-hand side can be evaluated exactly [27], [28]. For the lower bound we will use a GKS inequality to estimate the expected value of the free energy $\mathbf{E}_C[f_n(h)]$. This then yields a bound for the growth rate through a Legendre transform. Indeed for large n from (3) and (6) we expect

$$f_n(h) \approx \sup_{\omega}(g_n(\omega) + h\omega) \tag{8}$$

and therefore

$$\inf_{h}(f_n(h) - h\omega) \approx g_n^*(\omega) \tag{9}$$

where $g_n^*(\omega)$ is the concave hull of $g_n(\omega)$. Thus an estimate for $f_n(h)$ can be translated into an estimate for $g_n^*(\omega)$. It turns out that $g_n(\omega)$ is not concave on the whole interval $-1 \leq \omega \leq 1$ so the estimate applies only on the restricted portion $[-\omega_c, \omega_c]$ where the function is equal to its concave envelope.

Note that in the heuristic (8),(9) we identify $f_n(h)$ and $g_n(\omega)$ with their expectation over the code ensemble because of the concentration phenomenon. A proof of concentration for these quantities is still an open problem, although a weak form of it for the growth rate has been obtained in [14].

We end this paragraph by stressing that (3) is the partition function of a random spin system (or spin glass). Here random refers to the fact that the underlying graph is sampled uniformly from an ensemble LDPC(n, l, k). The coupling constants are "ferromagnetic" meaning that $J_c > 0$ in (5); in the context of coding $J_c = +\infty$ as in (3). If furthermore $h > 0$ we say that the spin system is ferromagnetic. We will need later the notation $\langle - \rangle(h)$ for the Gibbs average associated to this spin system. More precisely the Gibbs average of any observable $s_X = \prod_{i\in X} s_i$ is defined as

$$\langle s_X\rangle(h) = \frac{1}{Z_n(h)}\sum_{s^n} s_X e^{h\sum_{i=1}^{n} s_i}\prod_{c=1}^{m}\frac{1}{2}(1+s_{\partial c}).$$

B. Conditional Entropy and EXIT Curves

Assume communication through a noisy binary input memoryless channel with output alphabet $\mathbf{R}$ and transition probability density $p_{Y|X}(y|x)$. The input is a code word $x^n = (x_1, \ldots, x_n)$ from $C \in$ LDPC(n, r, k), and the output $(y_1, \ldots, y_n) = y^n$ belongs to $\mathbf{R}^n$.

Suppose now that a code word $(x_1^{\text{input}}, \ldots, x_n^{\text{input}})$ is sent through the channel. Denoting expectations with respect to the probability density

$$\prod_{i=1}^{n} p_{Y|X}(y_i|x_i^{\text{input}})$$

of the observed output as $\mathbf{E}_Y$, the Shannon conditional entropy of the input given the output is

$$H(X^n|Y^n) = -\mathbf{E}_{Y^n}\Bigg[\sum_{x^n} p_{X^n|Y^n}(x^n|y^n) \times \ln p_{X^n|Y^n}(x^n|y^n)\Bigg]. \quad (10)$$

We will assume that the channel is symmetric: as verified below this implies that (10) does not depend on $(x_1^{\text{input}}, \ldots, x_n^{\text{input}})$. We will prove lower and upper bounds on the later quantity and also on its derivative with respect to the inverse noise parameter (genericaly called σ) namely the "Generalized EXIT curve" associated to MAP decoding

$$\frac{d}{d\sigma}\frac{1}{n}\mathbf{E}_C[H(X^n|Y^n)]. \quad (11)$$

We will now rewrite (10) in the language of statistical mechanics, and thereby recognize that it is nothing else than the Gibbs entropy of a random spin system. For the case where the code words are uniformly distributed, i.e.

$$p_X(x^n) = \frac{1}{|C|}1_C(x^n) = \frac{1}{|C|}\prod_{c=1}^{m}\frac{1}{2}\big(1+s_{\partial c}\big) \quad (12)$$

and a memoryless channel, Bayes formula yields

$$p_{X^n|Y^n}(x^n|y^n) = \frac{1}{Z'(y^n)}1_C(x^n)\prod_{i=1}^{n} p_{Y|X}(y_i|x_i) \quad (13)$$

where the normalization factor is

$$Z'(y^n) = \sum_{x^n} 1_C(x^n)\prod_{i=1}^{n} p_{Y|X}(y_i|x_i).$$

In terms of the log-likelihood ratios

$$h_i = \frac{1}{2}\ln\frac{p_{Y|X}(y_i|0)}{p_{Y|X}(y_i|1)} \quad (14)$$

and of the spin variable $s_i = (-1)^{x_i}$,

$$p_{Y|X}(y_i|x_i) = \frac{1}{2}p(y_i|0)\frac{1+e^{-2h_i}}{\cosh h_i}\exp(s_i h_i).$$

Thus (13) becomes

$$p_{X^n|Y^n}(x^n|y^n) = \frac{1}{Z_n(h^n)}\exp\Big(\sum_{i=1}^{n} h_i s_i\Big)\prod_{c=1}^{m}\frac{1}{2}\big(1+s_{\partial c}\big) \quad (15)$$

where

$$Z_n(h^n) = \sum_{s^n}\exp\Big(\sum_{i=1}^{n} h_i s_i\Big)\prod_{c=1}^{m}\frac{1}{2}\big(1+s_{\partial c}\big). \quad (16)$$

These are the Gibbs measure and the partition function of a finite random spin system. By random we mean that the code is taken from the ensemble LDPC(n, r, k) and the log-likelihood ratios h_i (or "magnetic fields") have a distribution induced by (14)

$$p(h_i|x_i^{\text{input}})|dh_i| = p_{Y|X}(y_i|x_i^{\text{input}})|dy_i|.$$

It will also be useful later to think of (15) and (16) in hamiltonian terms: in other words the hard parity-check constraint can be replaced by a Gibbs weight as in (4).

Let us now specialize the discussion to symmetric channels for which the transition probability satisfies $p_{Y|X}(y|x) = p_{Y|X}(-y| - x)$. In this case the spin system described above posseses an important symmetry group of so called "gauge transformations." One observes that for a given code C the Gibbs measure (15) is invariant under the transformations

$$s_i \to \tau_i s_i, \qquad h_i \to \tau_i h_i, \qquad i = 1, \ldots, n$$

where τ^n is any code word. These transformations form a group and are local in the sense that each spin is multiplied by a i dependent (phase) factor: one says that the spin system has a gauge symmetry. The symmetry of the channel and (14) implies

$$\begin{aligned} p(h_i|x_i^{\text{input}})|dh_i| &= p_{Y|X}(y_i|x_i^{\text{input}})|dy_i| \\ \to \tilde{p}(h_i|x_i^{\text{input}})|dh_i| &= p_{Y|X}(y_i|\tau_i x_i^{\text{input}})|dy_i|. \end{aligned}$$

It is therefore clear that the conditional entropy does not depend on the input word: indeed we can choose $\tau_i = (-1)^{x_i^{\text{input}}}$ so that (10) remains the same except that now $\mathbf{E}_{Y^n}$ is with respect to $\prod_{i=1}^{n} p_{Y|X}(y_i|0)$. In other words for symmetric memoryless channels we may assume that the input word is the all 0 codeword. In the sequel, it is appropriate to replace the notation $\mathbf{E}_{Y^n}$ by $\mathbf{E}_{h^n}$. In the particular case of the BIAWGNC with inverse square noise σ, assuming the all 0 input code word, $\mathbf{E}_h$ is with respect to

$$p(h) = \frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{(h-\sigma)^2}{2\sigma}).$$

The conditional entropy can be related to the free energy of the spin system. Indeed substitution of (15) into (10) easily leads to

$$H(X^n|Y^n) = \mathbf{E}_{h^n}[\ln Z_n(h^n)] - \sum_{i=1}^{n} \mathbf{E}_{h^n}[h_i\langle s_i\rangle] \qquad (17)$$

where $\langle - \rangle$ is the Gibbs average with respect to the measure (15). More precisely for any $X \subset \{1,\ldots,n\}$

$$\langle s_X\rangle = \frac{1}{Z_n(h^n)} \sum_{s^n} s_X \exp(\sum_{i=1}^{n} h_i s_i) \prod_{c=1}^{m} \frac{1}{2}(1 + s_{\partial c}).$$

In the case of the Gaussian channel the second term in (17) becomes very simple (see [3] for a proof)

$$\mathbf{E}_{h^n}[h_i\langle s_i\rangle] = \sigma. \qquad (18)$$

III. Main Results

In this section we formulate our main results. These state that under natural hypothesis the growth rate and the GEXIT curve of LDPC codes are (in some range of parameters) rigorously given by the fixed point solutions of the density evolution equations associated to iterative message passing algorithms. In the language of statistical mechanics this amounts to say that the replica symmetric solution is exact for the two LDPC ensembles considered in this paper. For the reader familiar with spin glass theory we remark that the Gibbs measures of interest here are defined on the Nishimori line where it is *a priori* known that the gauge symmetry precludes the breaking of replica symmetry (see, for example, [19], [3] for further discussion of this point in the context of coding).

A. Growth Rate for the Regular Ensemble

Here we restrict ourselves to the ensemble $\mathrm{LDPC}(n,l,k)$ with k even. Then the growth rate (resp the free energy) is an even function of ω (resp h) so that all the discussion will be limited to $\omega \geq 0$, $h > 0$, without loss of generality.

Consider the density evolution equations

$$y_{t+2} = x_{t+1}^{k-1}, \qquad x_{t+1} = \tanh(h + (l-1)\tanh^{-1} y_t) \quad (19)$$

with the initial condition $x_1 = \tanh h$, $y_0 = 0$, $(h \geq 0)$ and the associated fixed point equation

$$y = x^{k-1}, \qquad x = \tanh(h + (l-1)\tanh^{-1} y). \qquad (20)$$

These are written in the "difference domain": a check node of degree k receives $k-1$ messages x and transmits a y message to a variable node; a variable node of degree l receives $l-1$ messages y and transmits an x message to a check node.

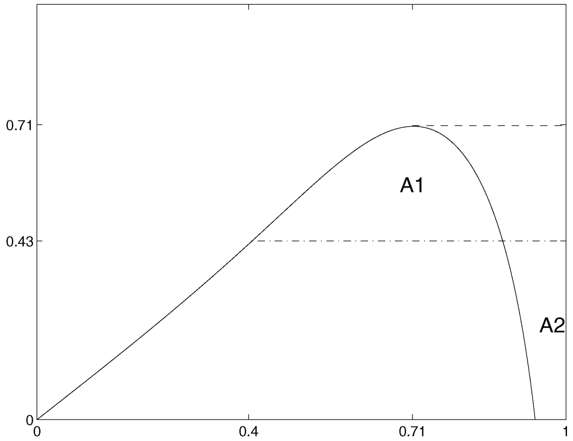

Fig. 1. $l = 5$, $k = 10$ code. Vertical axis is h, horizontal axis is ω. $h_{\mathrm{it}}(\omega)$ between 0 and $\omega_{\mathrm{it}} = 0.71$; $\omega_{\mathrm{it}}(h)$ increases from 0 to $h_{\mathrm{it}} = 0.71$ and jumps to $\omega_{\mathrm{it}}(h) = 1$ for $h > h_{\mathrm{it}}$. Full curve $h_{\mathrm{it,full}}(\omega)$. The Maxwell plateau at height $h_c = 0.43$ separates two equal areas $A_1 = A_2$.

In Section V we will prove that the sequence x_t (respectively, y_t) is increasing and tends toward the smallest solution of (20), greater than the initial condition. We call this particular fixed point $(x_*(h), y_*(h))$.

It is useful to notice that (20) gives the critical points of the function (sometimes called "replica symmetric free energy")

$$f_{RS}(x,y) = \ln[e^h(1+y)^l + e^{-h}(1-y)^l] - \frac{l}{k}\ln 2 + \frac{l}{k}\ln(1+x^k) - l\ln(1+xy). \quad (21)$$

The value of (21) at the particular fixed point $(x_*(h), y_*(h))$ is defined as the "iterative free energy,"

$$f_{\mathrm{it}}(h) = f_{RS}(x_*(h), y_*(h)). \qquad (22)$$

This definition is made here in order to stress that we look only at the particular fixed point reached by density evolution with initial condition $x_1 = \tanh h$, $y_0 = 0$. The "iterative magnetization" (the average relative weight given h) is

$$\omega_{\mathrm{it}}(h) = \tanh(h + l\tanh^{-1} y_*(h)). \qquad (23)$$

Our first application of a GKS inequality is

Lemma 3.1: Take a regular $\mathrm{LDPC}(n,l,k)$ ensemble with k even. Then for all h the weight enumerator satisfies

$$\liminf_{n\to\infty} \frac{d}{dh}\mathbf{E}_C[f_n(h)] \geq \omega_{\mathrm{it}}(h).$$

Remark 3.1: This lemma can be extended to irregular code ensembles provided the degrees of variable and check nodes are bounded, or their probabilities decay fast enough. It also extends to the situation where k is odd but $h > 0$.

It turns out that $\omega_{\mathrm{it}}(h)$ is increasing as a function of h so that the equation $\omega = \omega_{\mathrm{it}}(h)$ has at most one solution $h_{\mathrm{it}}(\omega)$. More precisely one should distinguish here the ensembles $\mathrm{LDPC}(n,2,k)$ and $\mathrm{LDPC}(n,l,k)$ with $l \geq 3$. For $l = 2$, $\omega_{\mathrm{it}}(h)$ is a continuous curve and $h_{\mathrm{it}}(\omega)$ is defined for every ω; moreover $f_{\mathrm{it}}(h)$ is continuous and differentiable. For $l \geq 3$ on

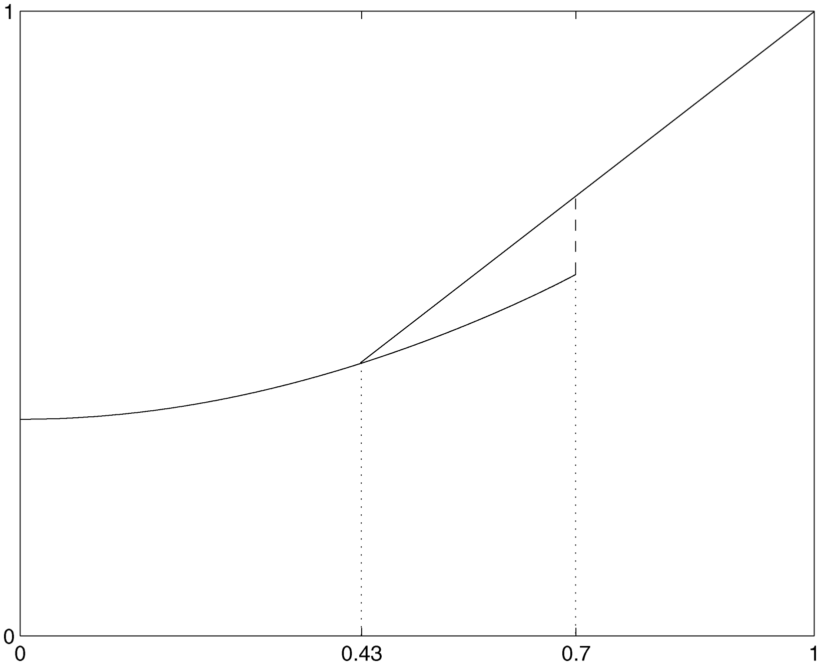

Fig. 2. Code $l = 5, k = 10$. Equilibrium free energy $f(h)$ goes from 0 to $h_c = 0.43$ where it is nondiffrentiable, and then is equal to h for $h > h_c$. The iterative free energy has a branch from 0 to $h_{\rm it} = 0.71$ where it has a jump discontinuity and is equal to h for $h > h_{\rm it}$.

the other hand $\omega_{\rm it}(h)$ is monotone increasing for $0 \leq h < h_{\rm it}$, has a unique jump discontinuity with a vertical slope at $h_{\rm it}$ and is equal to 1 for $h > h_{\rm it}$ (see Fig. 1 for the $(5, 10)$ ensemble). Thus when $l \geq 3$, $h_{\rm it}(\omega)$ is defined for $\omega \in [0, \omega_{\rm it}[$. The iterative free energy has a jump discontinuity at $h_{\rm it}$, a value above which it is simply linear (see Fig. 2).

From the above formulas it is possible to check that (23) and (22) are related. Namely, the total derivative of the iterative free energy is equal to the iterative magnetization at points where it is differentiable, but differs from it by a Dirac function at the jump discontinuity $h_{\rm it}$ (for $l \geq 3$). For $h \neq h_{\rm it}$,

$$\omega_{\rm it}(h) = \frac{d}{dh} f_{\rm it}(h).$$

Once $h_{\rm it}(\omega)$ is known, the iterative growth rate can be calculated as

$$g_{\rm it}(\omega) = f_{\rm it}(h_{\rm it}(\omega)) - h_{\rm it}(\omega)\omega. \tag{24}$$

For $l = 2$ (24) gives the full iterative growth rate for all ω. But for $l \geq 3$ this expression is defined only on the interval $[0, \omega_{\rm it}[$. In order to define iterative quantities for all ω one may plot the parametric curve

$$h(x) = \tanh^{-1} x - (l-1)\tanh^{-1} x^{k-1}$$
$$\omega(x) = \tanh[\tanh^{-1} x + \tanh^{-1} x^{k-1}] = \frac{x + x^{k-1}}{1 + x^k}.$$

This yields the full iteratite curve $h_{\rm it,full}(\omega)$ (see Fig. 1). The part of the curve corresponding to $\omega > \omega_{\rm it}$ comes from an unstable fixed point solution of (20). The iterative growth rate for $\omega \geq \omega_{\rm it}$ is then obtained from the same formulas as above applied to this unstable fixed point. We call the full iterative growth rate $g_{\rm it,full}(\omega)$ which is defined for all ω (see Fig. 2). We will also need $g^*_{\rm it,full}(\omega)$, the concave hull of $g_{\rm it,full}$, obtained by drawing a tangent to $g^*_{it,full}(\omega)$ passing through the point $\omega = 1$. The tangency point is ω_c (see Fig. 3). It is possible to see that necessarily $\omega_c \leq \omega_{\rm it}$ because $\omega_{\rm it}$ is the inflexion point of $g_{\rm it,full}$. We define h_c as the unique solution of $\omega_{\rm it}(h_c) = \omega_c$ (for $l = 2$, we have $\omega_c = 1$ and $h_c = +\infty$).

The first theorem can now be formulated.

Theorem 3.2: Let $f^*_{\rm it}(h)$ be the smallest convex function above $f_{\rm it}(h)$. For $l \geq 3$, $f^*_{\rm it}(h) = f_{\rm it}(h)$ for $|h| \leq h_c$ and $f^*_{\rm it}(h) = h$ for $|h| \geq h_c$. When $l = 2$, $f^*_{\rm it}(h) = f_{\rm it}(h)$. The $\limsup_{n\to\infty} \mathbf{E}_C[f_n(h)] = f(h)$ satisfies for all h

$$f(h) = f^*_{\rm it}(h).$$

Moreover $f'(h) = \omega_{\rm it}(h)$ for $|h| < h_c$ and $f'(h) = 1$ for $|h| > h_c$.

One expects that the limit of $\mathbf{E}_C[g_n(\omega)]$ also exists, but the foregoing estimates are not strong enough to prove this. We define the growth rate as

$$g(\omega) = \limsup_{n\to\infty} \mathbf{E}_C[g_n(\omega)]. \tag{25}$$

The next theorem relies on the natural assumption.

1) Hypothesis H1: $f(h) = \sup_\omega (g(\omega) + h\omega)$.

Theorem 3.3: Under hypothesis H1, for a regular ensemble LDPC(n, l, k) with k even, we have

$$g(\omega) = g_{\rm it}(\omega)$$

for all $|\omega| \leq \omega_c$.

We end this paragraph with a few informal comments in order to put these results into a broader perspective. The underpinning of this theorem is the Van der Waals picture of first order phase transitions [33]. A similar picture has been uncovered and

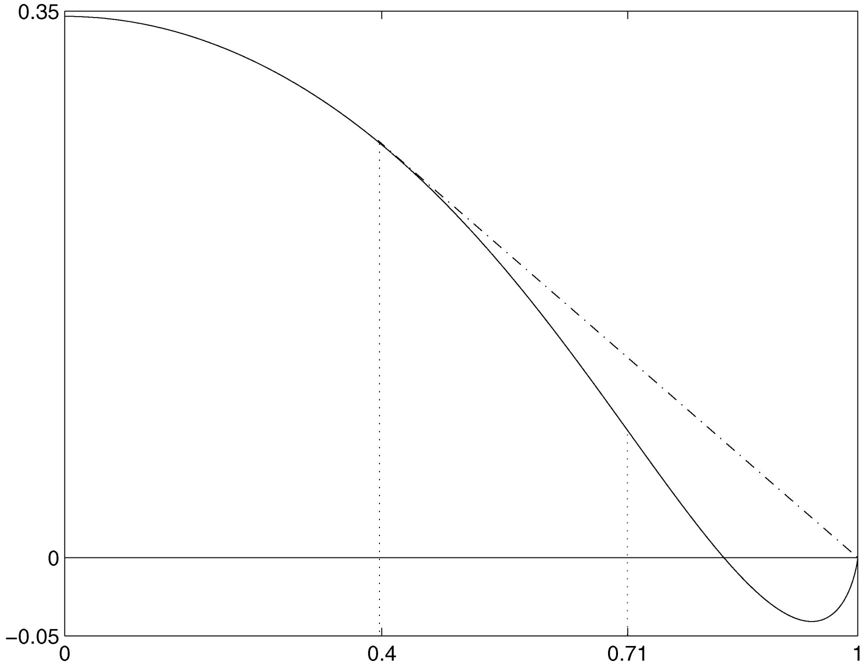

Fig. 3. Code ($l = 5, k = 10$). Full iterative growth rate $g_{\rm it,full}(\omega)$.The curve $g_{\rm it}(\omega)$ goes from 0 to $\omega_{\rm it} = 0.71$ where the concavity is lost. Theorem III-A holds up to $\omega_c = 0.4$ obtained by the Maxwell construction.

studied in detail in [21] in the context of communication through a BEC. The singularity of the free energy at h_c corresponds to the plateau $h_{\rm it}(\omega) = h_c$, or equivalently to the discontinuity of $\omega_{\rm it}(h)$ at h_c. The value h_c can be found from the "Maxwell construction": one draws the plateau such that the two areas A_1 and A_2 become equal. This statement is equivalent to the equality

$$\int_0^1 h_{\rm it,full}(\omega)d\omega = h_c - \int_0^{h_c} \omega_{\rm it}(h)dh.$$

In order to check this equality we note that the left-hand side is equal to

$$\int_0^1 \frac{d}{d\omega} g_{\rm it,full}(\omega)d\omega = g_{\rm it,full}(1) - g_{\rm it,full}(0) = -r\ln 2$$

and the right-hand side to

$$h_c - \int_0^{h_c} \frac{d}{dh} f_{\rm RS}(x_*(h), y_*(h)) = h_c - f_{\rm RS}(x_*(h_c), y_*(h_c)) \\ + f_{\rm RS}(x_*(0), y_*(0)) = f_{\rm RS}(x_*(0), y_*(0)) = r\ln 2.$$

The following fact is equivalent to the equality of areas: ω_c can be found on the graph of $g_{\rm it,full}$ by drawing the tangent passing through $\omega = 1$. From the point of view of the Van der Waals theory the part of the curve on $]\omega_c, \omega_{\rm it}[$ corresponds to a metastable state in the sense that the iterative free energy for $]h_c, h_{\rm it}[$ is the continuation of the left branch of the Gibbs free energy $f(h)$ (see Fig. 3) and is lower than the right linear branch (lower because here we have defined the free energy as minus the "physical" free energy for convenience). The part of the curve on $[\omega_{\rm it}, 1]$ can be obtained from the unstable fixed point solution of (20) and corresponds to a state which is unstable from the thermodynamic point of view [34], because on this interval $g_{\rm it,full}$ is *not* concave. The precise interpretation of stable, metastable and unstable states in the context of the code ensemble is for the moment unclear. For communication over a BEC [21] the authors discuss an interpretation of these states in terms of the complexity of decoding algorithms.

B. GEXIT Curves and Conditional Entropy for the Poisson Ensemble

Our notation will be as follows: $\mathbf{E}_Z$ is the expectation with respect to some random variable Z, which can be the log-likelihood variable (14), the degree l of variable nodes (l is Poisson with mean $(1-r)k$), likelihood variables x and u which are passed from variable to check nodes and vice versa. We will also need the probability densities $\zeta(x)$ and $\hat{\zeta}(u)$ of x and u.

Here it is convenient to write the density evolution equations associated to the iterative decoder in the "likelihood domain." A check node of degree k receives $k-1$ messages $x_1, \ldots, x_{k-1}$ and transmits the message u to a variable node; a variable node of degree l receives $l-1$ messages $u_1, \ldots, u_{l-1}$ and transmits the message x to a check node. The messsages are (half) log-likelihood ratios that can be interpreted as the "cavity magnetic fields" of spin glass theory

$$\hat{\zeta}^{(t+2)}(u) = \mathbf{E}_{x_1}^{(t+1)} \ldots \mathbf{E}_{x_{k-1}}^{(t+1)} \left[\delta(u - \tanh^{-1}(\prod_{i=1}^{k-1} \tanh x_i))\right] \tag{26}$$

$$\zeta^{(t+1)}(x) = \mathbf{E}_h \mathbf{E}_l \mathbf{E}_{u_1}^{(t)} \ldots \mathbf{E}_{u_{l-1}}^{(t)} \left[\delta(x - h - \sum_{c=1}^{l-1} u_c)\right]. \tag{27}$$

The sequence generated by the initial condition $\zeta^{(1)}(x) = p(x)$, $\hat{\zeta}^{(0)}(u) = \delta(u)$ converges (in a weak sense) to a limit probability measure. A proof of this fact, based on physical degradation, can be found in [22]. Here, for completeness, we give in Appendix IV a similar proof based on the GKS inequality. This limit probability measure, whose formal density we write $\zeta^*(x)$, $\hat{\zeta}^*(u)$, satisfies a fixed point equation given by (26)–(27) with the superscripts $t, t+1, t+2$ replaced by $*$. We will not make explicit use of this later fact, so its proof will be ommitted.

A simple calculation with functional derivatives shows that the solutions of the fixed point equations are the critical points of a functional (a "replica symmetric" free energy)

$$\begin{aligned} f_{\mathrm{RS}}[\zeta,\hat{\zeta}] =& (r-1)\ln 2 + (r-1)k\mathbf{E}_x\mathbf{E}_u\big[\ln(1 \\ &+ \tanh x \tanh u)\big] \\ &+ \mathbf{E}_h\mathbf{E}_l\mathbf{E}_{u_1}\ldots\mathbf{E}_{u_l}\Big[\ln\Big(e^h\prod_{c=1}^{l}(1+\tanh u_c) \\ &+ e^{-h}\prod_{c=1}^{l}(1-\tanh u_c)\Big)\Big] \\ &+(1-r)\mathbf{E}_{x_1}\ldots\mathbf{E}_{x_k}\Big[\ln(1+\prod_{i=1}^{k}\tanh x_i)\Big]. \end{aligned} \quad (28)$$

By definition the "iterative free energy" is

$$f_{\mathrm{it}}(\sigma) = \lim_{t\to\infty} f_{\mathrm{RS}}[\zeta^{(t+1)}, \hat{\zeta}^{(t+2)}] \quad (29)$$

and the "iterative GEXIT curve" (associated to a belief propagation decoder) is

$$\mathcal{E}_{\mathrm{it}}(\sigma) = \lim_{t\to\infty}\frac{1}{2}\mathbf{E}_h\mathbf{E}_l\mathbf{E}_{u_1}^{(t)}\ldots\mathbf{E}_{u_l}^{(t)}\Big[\tanh(h+\sum_{c=1}^{l}u_c)-1\Big]. \quad (30)$$

The argument for the existence of (30) is in the proof of Lemma (6.1). Concerning the existence of the limit in (29) the argument is a bit longer and is given in Appendix V. As we will show in Section VII there is a simple relation between the "iterative GEXIT curve" and the replica symmetric free energy

$$\lim_{t\to\infty}\frac{\partial f_{\mathrm{RS}}}{\partial\sigma}[\zeta^{t+1},\hat{\zeta}^{t+2}] - 1 = \mathcal{E}_{\mathrm{it}}(\sigma). \quad (31)$$

Our second application of a GKS inequality is

Lemma 3.4: Assume communication with a Poisson LDPC(n,r,k) code through a BIAWGNC with inverse square noise σ. For all σ,

$$\liminf_{n\to+\infty}\frac{1}{n}\frac{d}{d\sigma}\mathbf{E}_C[H(X^n|Y^n)] \ge \mathcal{E}_{\mathrm{it}}(\sigma). \quad (32)$$

Remark 3.2: This lemma can be extended to the situation of irregular codes as long as the degrees of variable and check nodes is bounded, or their probabilities decay fast enough. More recntly it has been proven for general channels again by correlation inequalities [12], [13]. As explained in the introduction such bounds have been derived earlier by physical degradation for the BSC [21], and proven in full generality more recently [11] by the data processing inequality.

Extending the interpolation method [8]–[10] to any value of k for a Gaussian channel we prove

Lemma 3.5: Assume communication with a Poisson LDPC(n,r,k) code through a BIAWGNC with inverse square noise σ. For all σ

$$\liminf_{n\to+\infty}\frac{1}{n}\mathbf{E}_C[H(X^n|Y^n)] \ge f_{\mathrm{it}}(\sigma) - \sigma. \quad (33)$$

Remark 3.3: In fact one first proves a sharper bound, see Corollary 6.3. This bound can be derived directly from (32) for $\sigma < \sigma_1$ where σ_1 is defined below.

For general LDPC ensembles one cannot exclude that $f_{\mathrm{it}}(\sigma)$ has jump discontinuities for some set D of values of σ (see the previous paragraph where this happens to $f_{\mathrm{it}}(h)$). For $\sigma \notin D$ we have formally

$$\begin{aligned} \frac{d}{d\sigma}f_{\mathrm{RS}}[\zeta^*,\hat{\zeta}^*] =& \frac{\partial f_{\mathrm{RS}}}{\partial\sigma}[\zeta^*,\hat{\zeta}^*] + \int\frac{\delta f_{\mathrm{RS}}}{\delta\zeta}[\zeta^*,\hat{\zeta}^*]\frac{\partial\zeta^*}{\partial\sigma} \\ &+ \int\frac{\delta f_{\mathrm{RS}}}{\delta\hat{\zeta}}[\zeta^*,\hat{\zeta}^*]\frac{\partial\hat{\zeta}^*}{\partial\sigma} = \frac{\partial f_{\mathrm{RS}}}{\partial\sigma}[\zeta^*,\hat{\zeta}^*]. \end{aligned} \quad (34)$$

In this formula the partial derivative is with respect to the σ dependence of $p(h)$. This motivates the natural assumption

1) Hypothesis H2: There exist at most a discrete set $\sigma_1,\ldots,\sigma_n$ of discontinuities of $f_{\mathrm{it}}(\sigma)$. For $\sigma \neq \sigma_i$, $i = 1,\ldots,n$ the left and right derivatives $\frac{d}{d\sigma}f_{\mathrm{it}}$, $\frac{\partial}{\partial\sigma}f_{\mathrm{RS}}$ exist and,

$$\frac{d}{d\sigma}f_{\mathrm{it}}(\sigma) = \lim_{t\to\infty}\frac{\partial f_{\mathrm{RS}}}{\partial\sigma}[\zeta^{t+1},\hat{\zeta}^{t+2}]. \quad (35)$$

In order to formulate our next result we make the definition

$$h_{\mathrm{av}}(\sigma) = \lim_{n\to\infty}\frac{1}{n}\mathbf{E}_C[H(X^n|Y^n)] \quad (36)$$

where the limit will be shown to exist for $\sigma > \sigma_n$.

Theorem 3.6: Assume H2 and communication through a BIAWGNC with the ensemble LDPC(n,r,k). For $\sigma > \sigma_n$ the limit (36) exists and we have

$$h_{\mathrm{av}}(\sigma) = f_{\mathrm{it}}(\sigma) - \sigma, \qquad \frac{d}{d\sigma}h_{\mathrm{av}}(\sigma) = \mathcal{E}_{\mathrm{it}}(\sigma). \quad (37)$$

If the set of discontinuities is empty the equalities hold for all σ.

Remark 3.4: The theorem follows by combining Lemmas 3.4 and 3.5. For $\sigma < \sigma_1$ both lemmas essentialy give the same information so that we are able to deduce equality only for $\sigma > \sigma_n$. For "good" generic LDPC ensembles we expect the iterative free energy to be discontinuous and the EXIT curves vanish in the low noise regime ($\sigma_n > 0$). However for the present case of a Poisson degree distribution of variable nodes the MAP threshold is at zero noise due to the finite fraction of unchecked nodes, and the EXIT curves are non trivial even at low noise.

We stress that here there is no restriction on the values of k. Although the result is restricted to the Gaussian channel it is probably more general. If we knew *a priori* that (35) holds for all σ the theorem would follow easily from the lower bound on the GEXIT curve (Lemma 3.4), and formula (31). Indeed if the equality in assumption H2 holds for every σ then [see (7.9)]

$$\begin{aligned} \int_0^\infty d\sigma\mathcal{E}_{\mathrm{it}}(\sigma) &= \int_0^\infty d\sigma\frac{d}{d\sigma}(f_{\mathrm{it}}(\sigma)-\sigma) \\ &= \lim_{\sigma\to\infty}(f_{\mathrm{it}}(\sigma)-\sigma) - f_{\mathrm{it}}(0) = -r\ln 2. \end{aligned} \quad (38)$$

So the area under the iterative EXIT curve is equal to $-r\ln 2$ which is also the area under the GEXIT curve. Since one curve is above the other, they must be equal. Note that if we do not assume continuity, the two members of the equality in H2 may differ by Dirac distributions for $\sigma \in D$ which destroy the above argument. This makes the proof of theorem 3.6 considerably more complicated.

As an application we give an estimate of the GEXIT curves in the low noise regime (i.e., $\sigma^{-1} \to 0$). Retaining only the term $l = 0$ in (30) we obtain

$$|\mathcal{E}_{\rm it}(\sigma)| \geq e^{-(1-r)k}\frac{1}{2}\mathbf{E}_h[1 - \tanh h].$$

As a particular application of a GKS inequality we will see (Section VII) that

$$|\mathcal{E}_{\rm it}(\sigma)| \leq \frac{1}{2}\mathbf{E}_h[1 - \tanh h].$$

For a Gaussian channel

$$\frac{1}{2}\mathbf{E}_h[1 - \tanh h] = \sum_{k=0}^{\infty}(-1)^k e^{-(k^2+3k+2)\sigma} \approx e^{-4\sigma}.$$

Thus, in the low noise regime

$$|\frac{d}{d\sigma}h_{\rm av}(\sigma)| = |\mathcal{E}_{\rm it}(\sigma)| = O(e^{-\frac{4}{\sigma^{-1}}}).$$

C. Two Useful Identities

Let us finaly state two identities that play an important role. The first is combined with GKS to prove Lemma 3.4, while the second is used in the Proof of Theorem 3.6.

Lemma 3.7: For a BIAWGNC and any linear code C

$$\frac{d}{d\sigma}H(X^n|Y^n) = \frac{1}{2}\sum_{i=1}^{n}(\mathbf{E}_h[\langle s_i\rangle] - 1) \tag{39}$$

$$\frac{d^2}{d\sigma^2}H(X^n|Y^n) = \frac{1}{2}\sum_{i,j=1}^{n}\mathbf{E}_h\left[\left(\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle\right)^2\right]. \tag{40}$$

This lemma is proved in Appendix I where we also discuss closely related formulas for bit and block error probabilities under ML decoding. These do not rely on the specific choice of the LDPC ensembles and are valid for any linear code in $\mathbf{F}_2^n$. The first identity is an instance of the relationship between mutual information and MMSE discussed in a different context by [24] (see also [21]). Note that the MMSE identity of [24] would involve $\langle s_i\rangle^2$ instead of $\langle s_i\rangle$; it turns out that for a symmetric channel the expected value of theses two quantities are equal (see the next section). The second identity appears to be new: in particular it suggests that the correlation function $\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle$ decays as $|i - j| \to \infty$.

IV. GKS Inequalities

In this section, we briefly review the GKS inequalities which are the main tool from which we will obtain the lower bounds. In general, the GKS inequalities express the positivity of certain correlations or equivalently the monotonicity of first and second derivatives of the free energy. The classical GKS inequalities [15] pertain to nonrandom ferromagnetic (positive coupling constants) spin systems. Although in the deterministic case they break down as soon as negative couplings are introduced, it turns out, quite surprisingly, that they are still true for expected values when couplings are random (with both signs) provided their distribution satisfies a certain symmetry condition (of which channel symmetry is a special case). The setting is given by a general spin Hamiltonian

$$H(s^n) = -\sum_{A\subset\{1,\ldots,n\}} J_A s_A \tag{41}$$

and the associated Gibbs averages

$$\langle s_B\rangle_J = \frac{\sum_{s^n} s_B e^{-H(s^n)}}{\sum_{s^n} e^{-H(s^n)}},\ s_B = \prod_{i\in B} s_i,\ B \subset \{1,\ldots,n\}.$$

A. Classical GKS Inequalities

We give the general statement and then specialize to the weight distribution problem. Suppose $J_A \geq 0$ for all A in the sum (41) (in other words the system is ferromagnetic). The first GKS inequality states that for any $X \subset \{1,\ldots,n\}$

$$\langle s_X\rangle_J \geq 0. \tag{42}$$

In particular it implies that for any J_A

$$\frac{\partial}{\partial J_A}\ln\sum_{s^n} e^{-H(s^n)} \geq 0.$$

The second GKS inequality states that for any pair of sets $X, Y \subset \{1,\ldots,n\}$

$$\langle s_X s_Y\rangle_J - \langle s_X\rangle_J\langle s_Y\rangle_J \geq 0. \tag{43}$$

In particular, this implies that

$$\frac{\partial}{\partial J_B}\langle s_A\rangle_J = \frac{\partial^2}{\partial J_B\partial J_A}\ln\sum_{s^n} e^{-H(s^n)} \geq 0. \tag{44}$$

In the weight enumerator problem, for a given code, the Hamiltonian (5) is of the form(41) with $J_A = J_{\partial c}$, for $A = \partial c \subset \{1,\ldots,n\}$ and $J_A = h$ for $A = \{i\}$. Therefore the GKS ineuqlities will be valid as long as $J_{\partial c} \geq 0$ and $h \geq 0$. For finite n we can trivially take the limit $J_{\partial c} \to +\infty$ so that they remain valid in the case of the hard core constraint which is really the case of interest. In fact the condition $h \geq 0$ is necessary only for k odd. Indeed when k is even the code word constraint (1) are invariant under $s_i \to -s_i$, $i = 1,\ldots n$ so that we are allowed to replace h by $|h|$.

B. GKS Inequality for Random Spin Systems

These have been derived recently in [16]. The formulation given here is slightly different but equivalent. The setting is now given by the Hamiltonian (41) where J_A are iid variables with a probability distribution satisfying

$$p(J_A) = p_0(|J_A|)e^{J_A} \tag{45}$$

for some function p_0. In the rest of the paper, it will be useful to keep in mind that (45) is equivalent to the class of ”symmetric distributions” [22], [32] defined by

$$p(-J_A) = p(J_A)e^{-2J_A}.$$

The expectation with respect to (45) will be denoted $\mathbf{E}_J$. As first shown by Nishimori [18], [19] when combined with gauge invariance of (41), condition (45) implies a host of exact identities (called Nishimori identities). Here we just state the identities that we will need latter on. For any subsets X, Y of $\{1,\dots,n\}$ we have

$$\mathbf{E}_J[\langle s_X\rangle_J^{2k}] = \mathbf{E}_J[\langle s_X\rangle_J^{2k-1}], \qquad k\geq 1 \tag{46}$$

$$\mathbf{E}_J[\langle s_X\rangle_J\langle s_X s_Y\rangle_J] = \mathbf{E}_J[\langle s_X\rangle_J\langle s_Y\rangle_J\langle s_X s_Y\rangle_J] \tag{47}$$

$$\mathbf{E}_J[\langle s_X\rangle_J\langle s_Y\rangle_J] = \mathbf{E}_J[\langle s_X\rangle_J\langle s_Y\rangle_J\langle s_X s_Y\rangle_J] \tag{48}$$

$$\mathbf{E}_J[\langle s_X\rangle_J^2\langle s_Y\rangle_J] = \mathbf{E}_J[\langle s_X\rangle_J^2\langle s_Y\rangle_J^2]. \tag{49}$$

In Appendix II, we give a proof of a general identity which contain all the above and many others as special cases. Although the method is standard we have not found such a general identity in the litterature.

In order to state the GKS inequalities a further assumption is needed, namely that *some subset* of $p(J_A)$ are Gaussian with mean and variance both equal to $t_A\geq 0$. Here the mean and variance are adjusted so that (45) is satisfied. Then using the Nishimori identities, it is shown in [16]

$$\frac{\partial}{\partial t_X}\mathbf{E}_J\left[\ln\sum_{s^n} e^{-H(s^n)}\right] = \mathbf{E}_J[\langle s_X\rangle_J]\geq 0 \tag{50}$$

and

$$\frac{\partial}{\partial t_Y}\mathbf{E}_J[\langle s_X\rangle_J] = 2t_Y\mathbf{E}_J\left[\left(\langle s_X s_Y\rangle_J - \langle s_X\rangle_J\langle s_Y\rangle_J\right)^2\right]\geq 0. \tag{51}$$

Let us discuss how this formalism can be applied to the random spin system defined by (15), (16). The first observation is that, remarkably, channel symmetry translates into (45) for the probability distribution of the log-likelihood variables (14). Thus for $A=i$, $i=1,\dots n$ we set $J_A=h_i$ (and $t_A=\sigma$ if the channel is Gaussian). Second, for $A=\partial c$, $c=1,\dots m$ we take independent and identically distributed (i.i.d.) Gaussian variables $J_A=J_c$, with $t_A=t_c$. We have

$$\begin{aligned}\langle s_X\rangle_J &= \frac{\sum_{s^n} s_X\exp\left(\sum_{c=1}^m(1-J_{\partial c}s_{\partial c})+\sum_{i=1}^n h_i s_i\right)}{\sum_{s^n}\exp\left(\sum_{c=1}^m(1-J_{\partial c}s_{\partial c})+\sum_{i=1}^n h_i s_i\right)}\\ &= \frac{\sum_{s^n} s_X\prod_{c=1}^m(1+s_{\partial c}\tanh J_c)e^{\sum_{i=1}^n h_i s_i}}{\sum_{s^n}\prod_{c=1}^m(1+s_{\partial c}\tanh J_c)e^{\sum_{i=1}^n h_i s_i}}.\end{aligned}$$

Taking the expectation with respect to all variables J_A

$$\begin{aligned}&\mathbf{E}_{h,J}[\langle s_X\rangle_J]\\ &=\mathbf{E}_h\left[\int\prod_{c=1}^m dJ_c\frac{e^{-\frac{J_c^2}{2}}}{\sqrt{2\pi}}\right.\\ &\qquad\left.\frac{\sum_{s^n}s_X\prod_{c=1}^m(1+s_{\partial c}\tanh(\sqrt{t_c}J_c+t_c))e^{\sum_{i=1}^n h_i s_i}}{\sum_{s^n}\prod_{c=1}^m(1+s_{\partial c}\tanh(\sqrt{t_c}J_c+t_c))e^{\sum_{i=1}^n h_i s_i}}\right].\end{aligned} \tag{52}$$

For any fixed J_c we have

$$\lim_{t_c\to\infty}\tanh(\sqrt{t_c}J_c+t_c)=1.$$

Moreover the ratio under the integral in (52) is bounded above by

$$2^{(n+m)}e^{\sum_{i=1}^n h_i s_i}.$$

Therefore, from dominated convergence, we obtain

$$\lim_{t_c\to\infty}\mathbf{E}_{h,J}[\langle s_X\rangle_J] = \mathbf{E}_h[\langle s_X\rangle].$$

The last formula implies that since the Nishimori identities (46), (47), (49) apply for t_c finite, they also apply to the spin system associated to a code in LDPC(n,r,k) (or for that matter any LDPC). Moreover, in the case of a Gaussian channel with noise parameter σ^{-1}, the probability distribution of the log-likelihoods is a Gaussian with mean and variance equal to σ: this means that we can apply the GKS inequalities (50) and (51) for finite t_c. Taking the limit $t_c\to\infty$ we conclude that the GKS inequalities also apply to any element of an LDPC ensemble. In particular if the channel is Gaussian

$$\frac{\partial}{\partial\sigma}\mathbf{E}_h[\langle s_X\rangle]\geq 0. \tag{53}$$

As will be shown later this is directly related to the monotonicity of the GEXIT curves.

V. Proof OF Theorem 3.3

We begin with the upper bounds which follow easily from results in the litterature. The combinatorial (or “annealed”) growth rate has been computed exactly using combinatorial methods [27], [28]. For regular codes

$$\lim_{n\to\infty}\frac{1}{n}\ln\mathbf{E}_C[A_n(\omega)] = g_{\mathrm{it,full}}(\omega). \tag{54}$$

Because of Jensen’s inequality this provides immediately a sharp upper bound (this is not true for irregular codes [29], [30])

$$g(\omega)\leq g_{\mathrm{it}}(\omega). \tag{55}$$

Note that in fact Gallager’s original upper bound [27]

$$\mathbf{E}_C[A_n(\omega)]\leq\exp(ng_{\mathrm{it,full}}(\omega)) \tag{56}$$

is sufficient to obtain (55). This bound (together with Jensen) gives also a sharp estimate for the free energy. Indeed

$$\begin{aligned}\exp(n\mathbf{E}_C[f_n(h)]) &\leq\sum_\omega\mathbf{E}_C[A_n(\omega)]\exp(nh\omega)\\ &\leq\sum_\omega\exp(n(g_{\mathrm{it,full}}(\omega)+h\omega))\\ &\leq n\exp(n\sup_\omega(g_{\mathrm{it,full}}(\omega)+h\omega))\end{aligned}$$

which implies

$$\begin{aligned}\limsup_{n\to\infty}\mathbf{E}_C[f_n(h)] &\leq\sup_\omega(g_{\mathrm{it,full}}(\omega)+h\omega)\\ &=\sup_\omega(g^*_{\mathrm{it,full}}(\omega)+h\omega) = f^*_{\mathrm{it}}(h).\end{aligned} \tag{57}$$

The last equality can be checked by explicit computation. This computation shows that for $l=2$, $f_{\mathrm{it}}^*(h)=f_{\mathrm{it}}(h)$ and for $l\geq 3$, $f_{\mathrm{it}}^*(h)=f_{\mathrm{it}}(h)$ if $|h|\leq h_c$, while $f_{\mathrm{it}}^*(h)=h$ if $|h|\geq h_c$.

We now proceed to prove the lower bounds. Consider $C\in \mathrm{LDPC}(n,l,k)$ and its associated Tanner graph. The distance between two nodes is defined as the minimal number of edges needed to connect the two nodes. Given some fixed variable node, say o, we define the neighborhood of depth d, denoted $\mathcal{N}_d(o)$, as the set of variable and check nodes which are at a distance less or equal to d from o. It is convenient to take d even. Tanner graphs of LDPC codes have the important property of being locally tree like. In [31], it is proven that there exists a numerical constant $c>0$ such that the probability that $\mathcal{N}_d(o)$ is a tree satisfies

$$\mathbf{P}[\mathcal{N}_d(o)\text{is a tree}]\geq 1-\frac{(cl)^{4d}}{n}. \tag{58}$$

Our first application of the GKS inequality is the proof of Lemma 5.1.

Lemma 5.1: Recall that $\langle - \rangle(h)$ denotes the Gibbs average associated to (3). We have

$$\liminf_{n\to\infty}\mathbf{E}_C[\langle s_o\rangle(h)]\geq \omega_{\mathrm{it}}(h). \tag{59}$$

Proof: Using (58) and (42)

$$\begin{aligned}&\mathbf{E}_C[\langle s_o\rangle(h)]\\ &\quad=\mathbf{E}_C[\langle s_o\rangle(h)|\mathcal{N}_d(o)\text{ a tree}]\mathbf{P}[\mathcal{N}_d(o)\text{ a tree}]\\ &\qquad+\mathbf{E}_C[\langle s_o\rangle(h)|\mathcal{N}_d(o)\text{ not a tree}]\mathbf{P}[\mathcal{N}_d(o)\text{ not a tree}]\\ &\quad\geq \mathbf{E}_C[\langle s_o\rangle(h)|\mathcal{N}_d(o)\text{ a tree}](1-\frac{(cl)^{4d}}{n}).\end{aligned}$$

Now consider the spin system defined by Hamiltonian (41) with $J_A=J_c$, for $A=\partial c\subset\{1,\ldots,n\}$ and $J_A=h$ for $A=\{i\}$. Set $J_c=0$ for $c\notin\mathcal{N}_d(o)$ and $J_c\to+\infty$ for $c\in\mathcal{N}_d(o)$. We call $\langle - \rangle_{\mathcal{N}_d(o)}$ the associated Gibbs measure. Since we choose d even all the spins attached to variable nodes not contained in $\mathcal{N}_d(o)$ are uncoupled so that $\langle s_o\rangle_{\mathcal{N}_d(o)}$ is the magnetization of the code restricted to the tree $\mathcal{N}_d(o)$. The GKS inequality in the form (44) implies that for each C for which $\mathcal{N}_d(o)$ is a tree

$$\langle s_o\rangle(h)\geq\langle s_o\rangle_{\mathcal{N}_d(o)}.$$

Therefore

$$\liminf_{n\to\infty}\mathbf{E}_C[\langle s_o\rangle(h)]\geq\mathbf{E}_C[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ a tree}]. \tag{60}$$

On a tree the iterative message passing procedure to compute $\langle s_o\rangle_{\mathcal{N}_d(o)}$ is exact, and yields

$$\mathbf{E}_C[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ a tree}]=\tanh(h+l\tanh^{-1}y_d). \tag{61}$$

The reader wishing to see a similar calculation in the context of statistical mechanics can consult the book of Baxter [35] where the Ising model on a tree is exactly solved. The result can also be inferred from the calculations reported in Appendix III.

To complete the proof of the lemma we have to derive the properties of the sequence y_d that were announced in Section III-A. In the case of a BEC, the problem is very similar to the present one and at this point Richardson and Urbanke use the concept of physical degradation [22]. Here there is no channel so there is no proper notion of physical degradation, but GKS turns out to be a convenient tool. By GKS, for a sequence of trees $\mathcal{N}_d(o)$, $\langle s_o\rangle_{\mathcal{N}_d(o)}$ is an increasing sequence (as a function of even d's). Since it is trivially bounded by 1 the sequence converges, thus (61) implies that y_d is an increasing and convergent sequence also. From (19) we conclude that x_{d-1} is also increasing and convergent. It is easy to see that the limit $(x_*(h),y_*(h))$ is necessarily one of the fixed point solutions of (20); and from (60), (61)

$$\liminf_{n\to\infty}\mathbf{E}_C[\langle s_o\rangle(h)]\geq\tanh(h+l\tanh^{-1}y_*(h)).$$

The following argument which is exactly the same than the one used on a BEC [22] shows how to select the right fixed point. Suppose that for some t, $x_t\leq x_*(h)$ (this is certainly the case for $t=0$). Then

$$\begin{aligned}x_{t+2}&=\tanh(h+(l-1)\tanh^{-1}x_t^k)\\ &\leq\tanh(h+(l-1)\tanh^{-1}x_*^k)=x_*(h).\end{aligned}$$

So the limit is equal to the smallest fixed point which is greater than the initial condition. This completes the proof of the lemma. □

Remark 5.1: Combining this lemma with (63) below, we obtain Lemma 3.1.

Lemma 5.2: For $0\leq h<h_{\mathrm{it}}$,

$$\liminf_{n\to\infty}\mathbf{E}_C[f_n(h)]\geq f_{\mathrm{it}}(h) \tag{62}$$

while for $h>h_{\mathrm{it}}$, $\lim_{n\to\infty}\mathbf{E}_C[f_n(h)]=f(h)$ exists and $f(h)=f_{\mathrm{it}}(h)=h$.

Proof: By symmetry

$$\begin{aligned}\mathbf{E}_C[\langle s_o\rangle(h)]&=\frac{1}{n}\sum_{i=1}^n\mathbf{E}_C[\langle s_i\rangle(h)]\\ &=\frac{d}{dh}\mathbf{E}_C[f_n(h)],\ 0\leq h<h_{\mathrm{it}}\end{aligned} \tag{63}$$

so that (59) becomes

$$\liminf_{n\to\infty}\frac{d}{dh}\mathbf{E}_C[f_n(h)]\geq\omega_{\mathrm{it}}(h)=\frac{d}{dh}f_{\mathrm{it}}(h). \tag{64}$$

For $h'<h_{\mathrm{it}}$ we integrate this inequality from 0 to h', use dominated convergence and $\mathbf{E}_C[f_n(0)]\geq r\ln 2$, $f_{\mathrm{it}}(0)=r\ln 2$ (for $l=2$ this argument holds for all h'). This yields (62). For $h>h_{\mathrm{it}}$, we remark that $\omega_{\mathrm{it}}(h)=1$, so that

$$1\geq\limsup_{n\to\infty}\frac{d}{dh}\mathbf{E}_C[f_n(h)]\geq\liminf_{n\to\infty}\frac{d}{dh}\mathbf{E}_C[f_n(h)]\geq 1.$$

This is equivalent to

$$\lim_{n\to\infty}\frac{d}{dh}(\mathbf{E}_C[f_n(h)]-h)=0.$$

Integrating over the interval $[h,\Lambda]$ ($h \geq h_{\text{it}}$) and using dominated convergence

$$\lim_{n\to\infty}(\mathbf{E}_C[f_n(\Lambda)] - \Lambda - \mathbf{E}_C[f_n(h)]) + h) = 0. \tag{65}$$

As shown by Gallager when $l \geq 3$ the bound (56) implies that the code has a linear minimum distance. More precisely there exists a $0 < \omega_{l,k} < 1$ such that the probability that a code word has relative weight $\omega_{l,k} < \omega < 1$ is less than $C_{l,k}n^{-(l-2)}$ for some numerical constant $C_{l,k}$. From this it is easily shown that for $\Lambda \geq \frac{\ln 2}{1-\omega_{l,k}}$

$$\mathbf{E}_C[f_n(\Lambda)] - \Lambda = O(n^{-(l-2)}).$$

We conclude the proof by taking the limit $n \to \infty$ in (65). □

Proof of Theorem 3.2: From (62) and (57)

$$f_{\text{it}}(h) \leq \liminf_{n\to\infty} \mathbf{E}_C[f_n(h)] \leq \limsup_{n\to\infty} \mathbf{E}_C[f_n(h)] \leq f^*_{\text{it}}(h).$$

Since $f_n(h)$ is convex, $\limsup_{n\to\infty} \mathbf{E}_C[f_n(h)]$ is convex. But $f^*_{\text{it}}(h)$ is by definition the smallest convex function above $f_{\text{it}}(h)$ so we must have

$$f^*_{\text{it}}(h) \leq \limsup_{n\to\infty} \mathbf{E}_C[f_n(h)].$$

Therefore

$$f^*_{\text{it}}(h) = \limsup_{n\to\infty} \mathbf{E}_C[f_n(h)]$$

which proves the theorem.

With the next lemma the Proof of Theorem 3.3 is complete.

Lemma 5.3: The equation $\omega = \omega_{\text{it}}(h)$ (see (23)) has at most one solution, and for every ω such that the solution exists we have

$$g(\omega) = g_{\text{it}}(\omega). \tag{66}$$

Proof: The equation

$$\omega = \omega_{\text{it}}(h)$$

has at most one solution because $\omega_{\text{it}}(h)$ is increasing: indeed by GKS $\langle s_0\rangle_{\mathcal{N}_d(o)}$ is an increasing function of h (at fixed d). For some values of ω the equation might not have a solution because $\omega_{\text{it}}(h)$ might be (and in practice is) discontinuous. From now on we look at the interval of ω for which the (unique) solution exists and call it $h_{\text{it}}(\omega)$. Because of Theorem 3.2

$$\inf_h(f(h) - h\omega) = \inf_h(f^*_{\text{it}}(h) - h\omega). \tag{67}$$

By assumption H1 the left-hand side is equal to $g^*(\omega)$ the convex hull of g. On the other hand the right-hand side of (67) is equal to $g^*_{\text{it,full}}(\omega) = g^*_{\text{it}}(\omega)$. Thus

$$g^*(\omega) = g^*_{\text{it}}(\omega).$$

By definition of ω_c, $g(\omega)_{\text{it}}$ is strictly concave on $[0,\omega_c]$. Therefore $g^*(\omega)$ is also strictly concave on $[0,\omega_c]$ and so must be $g(\omega)$. Thus $g(\omega) = g_{\text{it}}(\omega)$ on this interval. □

VI. Bounds on GEXIT and the Conditional Entropy

We begin with the lower bound on GEXIT for which the method is similar to that of the previous section.

A. Lower Bound on the GEXIT Curve

Given a code $C \in \text{LDPC}(n,r,k)$ and a specified variable node o we consider again a neighborhood $\mathcal{N}_d(o)$. The probability that all nodes of the factor graph have a degree less than $\ln n$ is (for n large)

$$\left(1 - \sum_{l>\ln n}\binom{n(1-r)}{l}\left(\frac{k}{n}\right)^l\left(1-\frac{k}{n}\right)^{n(1-r)-l}\right)^n \geq 1 - \frac{n}{2n^{\frac{1}{2}\ln(\ln n)}}.$$

Then following [31] one can show that the probability of $\mathcal{N}_d(o)$ being a tree again satisfies an inequality of the type (58) with l replaced by $\ln n$.

Lemma 6.1: For a binary input symmetric channel we have

$$\liminf_{n\to\infty} \mathbf{E}_{C,h^n}[\langle s_o\rangle] \geq \mathbf{E}_h\mathbf{E}_l\mathbf{E}^*_{u_1}\ldots\mathbf{E}^*_{u_l}\left[\tanh(h + \sum_{c=1}^{l} u_c)\right].$$

Remark 6.1: The proof below immediately extends to irregular ensembles as long as the degrees are bounded.

Proof: Proceeding as in the proof of 5.1, we get

$$\mathbf{E}_{C,h^n}[\langle s_o\rangle] \geq \mathbf{E}_{C,h}[\langle s_o\rangle|\mathcal{N}_d(o)\text{ tree}]\left(1 - \frac{(c\ln n)^{4d}}{n}\right).$$

Next we construct a new Gibbs measure for a random spin system on the tree $\mathcal{N}_d(o)$. Take the hamiltonian (41) and set the Gaussian coupling constants $J_A = J_c$ with $t_A = t_c$ for the check nodes $c = 1,\ldots,m$ and $J_A = h_i$ for the variable nodes $i = 1,\ldots,n$. Now for $c \notin \mathcal{N}_d(o)$ set $t_c = 0$ which means that $J_c = 0$ with probability one; and for $c \in \mathcal{N}_d(o)$ make $t_c \to \infty$. The associated Gibbs measure is denoted $\langle - \rangle_{\mathcal{N}_d(o)}$. The GKS inequality in the form (51) implies that

$$\mathbf{E}_{h^n}[\langle s_o\rangle|\mathcal{N}_d(o)\text{ tree}] \geq \mathbf{E}_{h^n}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ tree}].$$

Therefore

$$\liminf_{n\to\infty} \mathbf{E}_{C,h^n}[\langle s_o\rangle] \geq \mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ is a tree}]. \tag{68}$$

On a tree the right-hand side of this inequality can be computed exactly by iterative equations which yield (see Appendix C for a derivation)

$$\mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ tree}] = \mathbf{E}_h\mathbf{E}_l\mathbf{E}^{(d)}_{u_1}\ldots\mathbf{E}^{(d)}_{u_l}\left[\tanh(h + \sum_{c=1}^{l} u_c)\right]. \tag{69}$$

To complete the proof of the lemma we must show that the right-hand side has a well defined limit. One could proceed as in [31], [22] by using the concept of physical degradation. Here instead we use a GKS inequality

$$\mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{is a tree}] \leq \mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_{d+2}(o)}|\mathcal{N}_{d+2}(o)\text{is a tree}].$$

Thus the right-hand side of (69) is an increasing bounded sequence. Hence the result of the lemma follows by taking the limit $d \to \infty$ limit on both sides of inequality (68). □

The previous lemma holds for the rather general class of symmetric channels. For the next one however, we specify a Gaussian channel. The probability distribution of the log-likelihood variables is a Gaussian with mean and variance both equal to σ.

Proof: of Lemma 3.4: First of all, we note that from (69)

$$\mathcal{E}_{\text{it}}(\sigma) = \frac{1}{2}\big(\lim_{d\to+\infty}\mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ tree}] - 1\big).$$

Applying GKS in the form (53) to the Gibbs average on $\mathcal{N}_d(o)$ we conclude that $\mathbf{E}_{C,h}[\langle s_o\rangle_{\mathcal{N}_d(o)}|\mathcal{N}_d(o)\text{ tree}]$ is an increasing function of σ (d fixed). Thus $\mathcal{E}_{\text{IT}}$ increases as a function of σ.

By symmetry

$$\liminf_{n\to\infty}\frac{1}{2n}\sum_{i=1}^{n}\big(\mathbf{E}_{C,h^n}[\langle s_i\rangle] - 1\big) \geq \mathcal{E}_{\text{it}}(\sigma)$$

and because of identity (39)

$$\liminf_{n\to\infty}\frac{1}{n}\mathbf{E}_C[\frac{d}{d\sigma}H(X^n|Y^n)] \geq \mathcal{E}_{\text{it}}(\sigma).$$

This proves (32). □

B. Lower Bound on the Conditional Entropy

We use an extension of the interpolation techniques recently developed in [8] for the SAT and XORSAT problem. The case of LDPC codes has been considered in [10] for convex (on $[-e, +e]$) generating function of the check node degree distribution. In particular for regular check degree only even degree is allowed. Here we apply the interpolation technique following the setting of [8], [9] for Poisson LDPC and extend it to the case covering also *odd* degree. This is the only paragraph where the Poisson nature of the variable node degrees is crucialy used.

Let $0 \leq t \leq 1$ be an interpolation parameter, m_t a Poisson random variable (RV) with mean $\mathbf{E}[m_t] = t(1-r)m$, $l_{i,t}$, $i = 1, \dots, n$ i.i.d. Poisson RV with mean $\mathbf{E}[l_{i,t}] = (1-t)(1-r)k$. Let x be an RV distributed according to some arbitrary density $d(x)$. The later distribution is a "variational parameter" which will be adjusted later on. Consider an RV u distributed according to

$$\hat{d}(u) = \mathbf{E}_{x_1}\dots\mathbf{E}_{x_{k-1}}\left[\delta(u - \tanh^{-1}(\prod_{j=1}^{k-1}\tanh x_j)\right].$$

In this section $\mathbf{E}_x$ and $\mathbf{E}_u$ are expectations with respect to $d(x)$ and $\hat{d}(u)$, not to be confused with $\zeta(x)$ and $\hat{\zeta}(u)$ of previous sections. We introduce nm independent copies x_j^c and u_c^i and define the interpolating partition function

$$Z_n(t) = \left[\prod_{i=1}^{n}\prod_{c=1}^{l_{i,t}} 2\cosh u_c^i\right]^{-1}\sum_{s^n}\prod_{c=1}^{m_t}\frac{1}{2}(1 + s_{\partial c}) \times \exp\bigg(\sum_{i=1}^{n} s_i(h_i + \sum_{c=1}^{l_{i,t}} u_c^i)\bigg).$$

The corresponding Gibbs measure $\langle - \rangle_t$ interpolates between the $t = 0$ product measure (decoupled spins) and the $t = 1$ measure associated to an LDPC code. The average free energy can be computed as follows:

$$\frac{1}{n}\mathbf{E}_{C,h^n}[\ln Z_n(h^n)] = \frac{1}{n}\mathbf{E}_{l_{i,0},u_c^i}[\ln Z_n(t=0)] + \frac{1}{n}\int_0^1 dt\frac{d}{dt}\mathbf{E}_{C,h,m_t,l_{i,t},u_c^i}[\ln Z_n(t)].$$

For $t = 0$, the free energy of the code ensemble is easily computed

$$\frac{1}{n}\mathbf{E}[\ln Z_n(t=0)] = \mathbf{E}_l\mathbf{E}_h\mathbf{E}_{u_1}\dots\mathbf{E}_{u_l} \times\left[\ln 2\cosh(h + \sum_{c=1}^{l} u_c)\right] - (1-r)k\mathbf{E}_u[\ln 2\cosh u].$$

Following [8], [9] the calculation of the derivative with respect to t leads to

$$\frac{1}{n}\mathbf{E}_{C,h}[\ln Z_n(h^n)] = f_{RS}[d, \hat{d}] + \int_0^1 dt R_n(t) \tag{70}$$

with the remainder term

$$R_n(t) = \sum_{p=1}^{\infty}\frac{(-1)^{p+1}}{p}\bigg(\mathbf{E}[\langle Q_p^k\rangle_{p,t}] - k\mathbf{E}[\langle Q_p\rangle_{p,t}]T_p^{k-1} + (k-1)T_p^k\bigg).$$

In this expression we use the shorter notation $\mathbf{E}$ for the expectation over C, h_i, u_c^i, $l_{i,t}$, m_t. The numbers T_p are computed from $d(x)$ as

$$T_p = \mathbf{E}_x[(\tanh x)^p].$$

The ”overlap parameter” Q_p is defined as

$$Q_p = \frac{1}{n}\sum_{i=1}^{n} s_i^{(1)}\dots s_i^{(p)} \tag{71}$$

where s_i^α $\alpha = 1, \dots, p$ are p independent copies (or replicas) of the spin s_i. The average $\langle - \rangle_{p,t}$ has to be understood as the interpolating Gibbs measure $\langle - \rangle_t$ replicated p times. For example the average of each term in (71) is

$$\langle s_i^{(1)}\dots s_i^{(p)}\rangle_{p,t} = \langle s_i\rangle_t^p.$$

We will prove the following statement.

Proposition 6.2: For any symmetric probability distribution $d(x)$ and for a Gaussian channel we have

$$\liminf_{n\to+\infty}\int_0^1 dtR_n(t)\geq 0, \qquad \text{any } k. \tag{72}$$

The proof of the proposition will clearly show that in the case of k even it is sufficient to have $d(x)$ symmetric and any symmetric channel. For k odd we need to specify a Gaussian channel, but we believe that this should be true also more generaly.

Corollary 6.3: Let $\mathcal{S}$ denote the set of symmetric probability distributions. For a Gaussian channel we have

$$\liminf_{n\to\infty}\frac{1}{n}\mathbf{E}_C[H(X^n|Y^n)]\geq\sup_{d\in\mathcal{S}}f_{\mathrm{RS}}[d,\hat{d}]-\sigma\geq f_{\mathrm{it}}(\sigma)-\sigma.$$

Proof: An immediate consequence of the proposition is that for any symmetric probability distribution $d(x)$ the free energy is lower bounded by $f_{\mathrm{RS}}[d,\hat{d}]$. This shows the first inequality. Since $\zeta^{t+1}(x)$ is a symmetric probability distribution [32] the bound holds also for $d(x)=\zeta^{t+1}(x)$ and $\hat{d}(u)=\hat{\zeta}^{t+2}(u)$. Because of the identity (17) we also get a lower bound on the conditional entropy. Finally, we perform the limit $d\to\infty$. □

In order to prove Proposition 6.2 we first need to show that the overlap parameter does not fluctuate. This is expressed as follows.

Lemma 6.4: Recall $\mathbf{E}=\mathbf{E}_{C,h,m_t,l_{i,t},u_c^i}$ and $\langle -\rangle_{p,t}$ is the p times replicated interpolation measure. Let $\mathbf{P}$ denote the probability distribution $\mathbf{P}[X]=\mathbf{E}\langle 1_X\rangle_{p,t}$ and fix some $0<\delta<\frac{1}{4}$. For a Gaussian channel we have that for almost every σ

$$\lim_{n\to\infty}\int_0^1 dt\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]=0. \tag{73}$$

Proof: Using

$$b^k-a^k=(b-a)\sum_{l=0}^{k-1}b^{k-l-1}a^l \tag{74}$$

and $|Q_p|\leq 1$ we get

$$|Q_p^k-\langle Q_p\rangle_{p,t}^k|\leq k|Q_p-\langle Q_p\rangle_{p,t}|$$

therefore from Tchebycheff inequality applied to $\langle 1_X\rangle_{p,t}$

$$\begin{aligned}\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]&\leq\mathbf{P}\big[|Q_p-\langle Q_p\rangle_{p,t}|>\frac{p}{kn^\delta}\big]\\&\leq\frac{k^2n^{2\delta}}{p^2}\mathbf{E}\Big[\langle Q_p^2\rangle_{p,t}-\langle Q_p\rangle_{p,t}^2\Big].\end{aligned} \tag{75}$$

Let us estimate the right-hand side of (75). We notice

$$\begin{aligned}\mathbf{E}[\langle Q_p^2\rangle_{p,t}]&=\frac{1}{n^2}\sum_{i,j=1}^n\mathbf{E}[\langle s_i^{(1)}\dots s_i^{(p)}s_j^{(1)}\dots s_j^{(p)}\rangle_{p,t}]\\&=\frac{1}{n^2}\sum_{i,j=1}^n\mathbf{E}[\langle s_is_j\rangle_t^p]\\\mathbf{E}[\langle Q_p\rangle_{p,t}^2]&=\frac{1}{n^2}\sum_{i,j=1}^n\mathbf{E}[\langle s_i^{(1)}\dots s_i^{(p)}\rangle_{p,t}\langle s_j^{(1)}\dots s_j^{(p)}\rangle_{p,t}]\\&=\frac{1}{n^2}\sum_{i,j=1}^n\mathbf{E}[\langle s_i\rangle_t^p\langle s_j\rangle_t^p].\end{aligned}$$

Thus using (74) and then the Schwarz inequality we obtain

$$\begin{aligned}\mathbf{E}[\langle Q_p^2\rangle_{p,t}-\langle Q_p\rangle_{p,t}^2]&=\frac{1}{n^2}\sum_{i,j=1}^n\mathbf{E}[\langle s_is_j\rangle_t^p-\langle s_i\rangle_t^p\langle s_j\rangle_t^p]\\&\leq\frac{p}{n^2}\sum_{i,j=1}^n\mathbf{E}[|\langle s_is_j\rangle_t-\langle s_i\rangle_t\langle s_j\rangle_t|]\\&\leq\frac{p}{n}\Big\{\sum_{i,j=1}^n\mathbf{E}\Big[(\langle s_is_j\rangle_t-\langle s_i\rangle_t\langle s_j\rangle_t)^2\Big]\Big\}^{1/2}.\end{aligned} \tag{76}$$

For a Gaussian channel the identity (40) holds so that

$$\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]\leq\frac{k^2n^{2\delta}}{\sqrt{2}pn}\Big\{\frac{d^2}{d\sigma^2}\mathbf{E}[\ln Z(t)]\Big\}^{1/2}.$$

Integrating both side against a C_0^∞ positive test function $\varphi(\sigma)$ we get

$$\begin{aligned}&\int d\sigma\varphi(\sigma)\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]\\&\quad\leq\frac{k^2}{\sqrt{2}pn^{1-2\delta}}\int d\sigma\varphi(\sigma)\Big\{\frac{d^2}{d\sigma^2}\mathbf{E}[\ln Z(t)]\Big\}^{1/2}\\&\quad\leq\frac{k^2}{\sqrt{2}pn^{1-2\delta}}\Big\{\int d\sigma\varphi(\sigma)\Big\}^{1/2}\\&\qquad\times\Big\{\int d\sigma\varphi(\sigma)\frac{d^2}{d\sigma^2}\mathbf{E}[\ln Z(t)]\Big\}^{1/2}\\&\quad\leq\frac{k^2}{\sqrt{2}pn^{1-2\delta}}\Big\{\int d\sigma\varphi(\sigma)\Big\}^{1/2}\\&\qquad\times\Big\{\int d\sigma|\varphi'(\sigma)|\frac{d}{d\sigma}\mathbf{E}[\ln Z(t)]\Big\}^{1/2}\end{aligned}$$

where we have used Schwarz inequality and an integration by parts. The identity (39) can be extended to the interpolating system so that

$$\begin{aligned}&\int d\sigma\varphi(\sigma)\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]\\&\quad\leq\frac{k^2}{\sqrt{2}pn^{1-2\delta}}\Big\{\int d\sigma\varphi(\sigma)\Big\}^{1/2}\\&\qquad\times\Big\{\int d\sigma|\varphi'(\sigma)|\frac{1}{2}\sum_{i=1}^n(1+\mathbf{E}[\langle s_i\rangle_t])\Big\}^{1/2}\\&\quad\leq\frac{k^2}{\sqrt{2}pn^{\frac{1}{2}-2\delta}}\Big\{\int d\sigma\varphi(\sigma)\Big\}^{1/2}\Big\{\int d\sigma|\varphi'(\sigma)|\Big\}^{1/2}.\end{aligned}$$

Integrating over $t\in[0,1]$, using Fubini's theorem on the left to exchange the t and σ integrals, and then using dominated convergence we obtain

$$\int d\sigma\varphi(\sigma)\lim_{n\to\infty}\int_0^1 dt\mathbf{P}\big[|Q_p^k-\langle Q_p\rangle_{p,t}^k|>\frac{p}{n^\delta}\big]=0.$$

Since this is true for any positive test function we conclude that (73) holds for almost every σ. □

Proof of Proposition 6.3: The first step consists in combining the terms in the remainder with p odd and p even. First of all since $d(x)$ is a symmetric probability distribution

$$T_{2q} = T_{2q-1} \tag{77}$$

(see Appendix II). A similar identity holds for the overlaps

$$\langle Q_{2q}^k \rangle_{2q,t} = \langle Q_{2q-1}^k \rangle_{2q-1,t}. \tag{78}$$

Indeed from (71)

$$\begin{aligned}\langle Q_p^k \rangle_{p,t} &= \frac{1}{n^k} \sum_{i_1 \dots i_k = 1}^{n} \langle (s_{i_1}^{(1)} \dots s_{i_k}^{(1)}) \dots (s_{i_1}^{(p)} \dots s_{i_k}^{(p)}) \rangle_{p,t} \\ &= \frac{1}{n^k} \sum_{i_1 \dots i_k = 1}^{n} \langle s_{i_1} \dots s_{i_k} \rangle_t^p.\end{aligned}$$

Symmetry of $d(x)$ implies symmetry of $\hat{d}(u)$. Thus the interpolation measure satisfies the Nishimori identities (46)–(49) one of which tells us that for $p = 2q$

$$\mathbf{E}\big[\langle s_{i_1} \dots s_{i_k} \rangle_t^{2q}\big] = \mathbf{E}\big[\langle s_{i_1} \dots s_{i_k} \rangle_t^{2q-1}\big]$$

therefore (78) follows. With the help of (77) and (78) the terms in the remainder can be rearranged

$$\begin{aligned}R_n(t) = \sum_{q=1}^{\infty} \frac{1}{2q(2q-1)} \bigg(&\mathbf{E}[\langle Q_{2q}^k \rangle_{2q,t}] \\ &- k\mathbf{E}[\langle Q_{2q} \rangle_{2q,t}] T_{2q}^{k-1} + (k-1) T_{2q}^k \bigg).\end{aligned}$$

Clearly the sum over q converges because the term in the parenthesis is bounded by than $2k$; so it remains to prove that the parenthesis is non negative. For even k this easily follows from the convexity of the function $f(z) = z^k$. Indeed convexity implies

$$\begin{aligned}Q_{2q}^k - kQ_{2q}T_{2q}^{k-1} + (k-1)T_{2q}^k = (Q_{2q}^k - T_{2q}^k)& \\ - kT_{2q}^{k-1}(Q_{2q} - T_{2q}) \geq 0.& \quad (79)\end{aligned}$$

This is the argument used in [8]–[10].

Here we obtain that for any k the slightly weaker result (72) is true. We split the sum over q in two parts $2q \geq n^{\delta/2}$ and $2q < n^{\delta/2}$. The fisrt sum is clearly smaller than

$$\sum_{2q \geq n^{\delta/2}} \frac{2k}{2q(2q-1)} \leq C_1 \frac{k}{n^{\delta/2}} \tag{80}$$

for some positive numerical constant. We split the second sum over $2q < n^{\delta/2}$ in two more sums as follows:

$$\begin{aligned}\sum_{2q<n^{\delta/2}} \frac{1}{2q(2q-1)} \mathbf{E}\bigg[&\langle Q_{2q} \rangle_{2q,t}^k - k\langle Q_{2q} \rangle_{2q,t} T_{2q}^{k-1} \\ &+ (k-1)T_{2q}^k \bigg] \\ &+ \sum_{q<n^{\delta/2}} \frac{1}{2q(2q-1)} \mathbf{E}[\langle Q_{2q}^k \rangle_{2q,t} - \langle Q_{2q} \rangle_{2q,t}^k]. \quad (81)\end{aligned}$$

Since $\langle Q_{2q} \rangle_{2q,t}$ is positive we can use the convexity of the function $f(z) = |z|^k$ to show that

$$\langle Q_{2q} \rangle_{2q,t}^k - k\langle Q_{2q} \rangle_{2q,t} T_{2q}^{k-1} + (k-1)T_{2q}^k \geq 0 \tag{82}$$

(notice the difference between (79) and (82)) so that the first sum in (81) is non negative. The second sum can be estimated by

$$\begin{aligned}&\sum_{2q<n^{\delta/2}} \frac{1}{2q(2q-1)} \frac{2q}{n^\delta} \\ &\qquad + \sum_{2q<n^{\delta/2}} \frac{1}{2q(2q-1)} \mathbf{P}\big[|Q_{2q}^k - \langle Q_{2q} \rangle_{2q,t}^k| > \frac{2q}{n^\delta} \big].\end{aligned}$$

The first sum is smaller than $C_2 n^{-\delta} \ln n$ (C_2 a positive numerical constant); and because of Lemma 6.4, for almost everywhere (a.e.) σ the second sum has a t integral which tends to zero as $n \to \infty$. Combining these remarks with (80) and (82) we conclude that

$$\liminf_{n \to \infty} \int_0^1 dt R_n(t) \geq 0.$$

VII. Proof of Theorem 3.6

We first derive a consequence of Lemma 3.4

Lemma 7.1: For a Gaussian channel with inverse noise σ we have

$$\liminf_{n \to +\infty} \frac{1}{n} \mathbf{E}_C[H(X^n|Y^n)] \geq f_{\mathrm{it}}(\sigma) - \sigma, \quad \sigma < \sigma_1 \tag{83}$$

$$\limsup_{n \to +\infty} \frac{1}{n} \mathbf{E}_C[H(X^n|Y^n)] \leq f_{\mathrm{it}}(\sigma) - \sigma, \quad \sigma > \sigma_n. \tag{84}$$

Proof: Let us first compute $\frac{\partial f_{RS}}{\partial \sigma}[\zeta^{t+1}, \hat{\zeta}^{t+2}]$. Using the identity

$$\frac{\partial}{\partial \sigma} \frac{e^{-\frac{(h-\sigma)^2}{2\sigma}}}{\sqrt{2\pi\sigma}} = \Big(-\frac{\partial}{\partial h} + \frac{1}{2}\frac{\partial^2}{\partial h^2}\Big) \frac{e^{-\frac{(h-\sigma)^2}{2\sigma}}}{\sqrt{2\pi\sigma}} \tag{85}$$

and an integration by parts

$$\begin{aligned}&\frac{\partial f_{RS}}{\partial \sigma}[\zeta^{(t+1)}, \hat{\zeta}^{(t+2)}] \\ &\quad = \mathbf{E}_l \mathbf{E}_h \mathbf{E}_{u_1}^{(t+2)} \dots \mathbf{E}_{u_l}^{(t+2)} \bigg[\Big(\frac{\partial}{\partial h} \\ &\qquad + \frac{1}{2}\frac{\partial^2}{\partial h^2} \Big) \ln\Big(e^h \prod_{c=1}^{l} (1 + \tanh u_c) + e^{-h} \prod_{c=1}^{l} (1 - \tanh u_c) \Big) \bigg] \\ &\quad = \mathbf{E}_l \mathbf{E}_h \mathbf{E}_{u_1}^{(t+2)} \dots \mathbf{E}_{u_l}^{(t+2)} \bigg[\tanh(h + \sum_{c=1}^{l} u_c) \\ &\qquad + \frac{1}{2}(1 - \tanh^2(h + \sum_{c=1}^{l} u_c)) \bigg].\end{aligned}$$

The probability distributions $p(h)$ and $\zeta^{(t+1)}$, $\hat{\zeta}^{(t+2)}$ are symmetric [22], [32] which implies that (see Appendix B where this is seen as a special case of a Nishimori identity)

$$\mathbf{E}_h\mathbf{E}_l\mathbf{E}_{u_1}^{(t+2)}\dots\mathbf{E}_{u_l}^{(t+2)}\big[\tanh(h+\sum_{c=1}^{l}u_c)\big] = \mathbf{E}_h\mathbf{E}_{u_1}^{(t+2)}\dots\mathbf{E}_{u_l}^{(t+2)}\big[\tanh^2(h+\sum_{c=1}^{l}u_c)\big]. \quad (86)$$

Thus we obtain

$$\lim_{t\to\infty}\frac{\partial f_{RS}}{\partial\sigma}[\zeta^{(t+1)},\hat{\zeta}^{(t+2)}] = \mathcal{E}_{\rm it}(\sigma)+1 \quad (87)$$

and because of the assumption H2

$$\mathcal{E}_{\rm it}(\sigma) = \frac{d}{d\sigma}(f_{\rm it}(\sigma)-\sigma), \qquad \sigma\neq\sigma_1,\dots,\sigma_n. \quad (88)$$

To finish the proof we will now integrate the lower bound of lemma 3.4. We first integrate from 0 to $\sigma' < \sigma_1$ and apply dominated convergence, to get

$$\liminf_{n\to\infty}\frac{1}{n}\int_0^{\sigma'}\frac{d}{d\sigma}\mathbf{E}_C[H(X^n|Y^n)] \geq \int_0^{\sigma'}d\sigma\mathcal{E}_{\rm it}(\sigma) = \int_0^{\sigma'}d\sigma\frac{d}{d\sigma}(f_{\rm it}(\sigma)-\sigma).$$

For $\sigma = 0$ we have $\mathbf{E}_C[H(X^n|Y^n)] = f_{\rm it}(0) = r\ln 2$. Thus

$$\liminf_{n\to\infty}\frac{1}{n}\mathbf{E}_C[H(X^n|Y^n)] \geq f_{\rm it}(\sigma')-\sigma', \qquad \sigma'<\sigma_1. \quad (89)$$

For $\sigma' > \sigma_n$, we proceed similarly by integrating from σ' to $+\infty$. This time we must use that the conditional entropy vanishes as $\sigma\to\infty$ (limit of zero noise), and

$$\lim_{\sigma\to\infty}(f_{\rm it}(\sigma)-\sigma) = 0.$$

This last formula follows easily by integrating the following inequality:

$$\frac{1}{2}\mathbf{E}_h[\tanh h - 1] \leq \mathcal{E}_{\rm it}(\sigma) \leq e^{-(1-r)k}\frac{1}{2}\mathbf{E}_h[\tanh h - 1].$$

The left-hand side is an application of GKS: consider formula (C.15) and remove all check nodes below the root of the tree. The right-hand side is immediately obtained by retaining only the $l=0$ term (unchecked nodes). □

End of Proof of Theorem 3.6: From Lemma 83 and Corollary 6.3, we have

$$h_{\rm av}(\sigma) = f_{\rm it}(\sigma)-\sigma, \qquad \text{and} \qquad \sigma>\sigma_n. \quad (90)$$

It remains to compute the total derivative as in (85)–(87)

$$\frac{dh_{av}}{d\sigma} = \lim_{t\to\infty}\frac{1}{2}\mathbf{E}_h\mathbf{E}_l\mathbf{E}_{u_1}^{(t)}\dots\mathbf{E}_{u_l}^{(t)}\big[\tanh(h+\sum_{c=1}^{l}u_c)-1\big] = \mathcal{E}_{\rm it}(\sigma).$$

This proves the theorem.

VIII. Conclusion

Correlation inequalities often provide a powerful tool in statistical mechanics of spin systems. A major aim of this paper was to demonstrate that one of them, the GKS inequality, is useful to analyze LDPC codes. For the regular codes it provides a way to prove a sharp lower bound on the growth rate. For the Poisson ensemble and communication over a Gaussian channel it yields a sharp lower bound on the the GEXIT curve. As pointed out in the proofs of these results, GKS turns out to be an alternative tool to physical degradation in the later case; but can also be used when there is no channel (as in the growth rate problem). An important issue is to clarify what is the precise connection between GKS and physical degradation. One should also investigate if GKS and/or other correlation inequalities apply to other coding schemes and channels: this is in fact very likely in view of the intimate connection between linear codes and spin systems.

The extension of the interpolation technique to odd degree for check nodes works on a Gaussian channel because it relies on identity (40). Presumably the later identity can be generalized to other symmetric channels so that one can hope to extend the present results to general symmetric channels. We wish to point out that an extension to more general irregular LDPC(n,λ,ρ) ensembles might also be achieved by using a version of the interpolation techniques developped in [36] or [10].

Finaly let us point out that it would be desirable to improve on the present results in order to remove assumptions H1 and H2.

Appendix I
Proof of Lemma 3.7

We first prove the two identities in Lemma 3.7 and then comment on closely related formulas for the bit error probability.

Consider a fixed code C. Because of (85), using an integration by parts we have

$$\frac{d}{d\sigma}\mathbf{E}_h[\ln Z_n(h^n)] = \sum_{i=1}^{n}\int\mathbf{E}_h\left[\left(\frac{\partial}{\partial h_i}+\frac{1}{2}\frac{\partial^2}{\partial h_i^2}\right)\ln Z_n(h^n)\right].$$

Using the Nishimori identity (valid for any symmetric channel) $\mathbf{E}_h[\langle s_i\rangle] = \mathbf{E}_h[\langle s_i\rangle^2]$

$$\left(\frac{\partial}{\partial h_i}+\frac{1}{2}\frac{\partial^2}{\partial h_i^2}\right)\ln Z_n(h^n) = \mathbf{E}_h\left[\langle s_i\rangle+\frac{1}{2}(\langle s_i\rangle^2-\langle s_i\rangle)\right] = \frac{1}{2}(1+\mathbf{E}_h[\langle s_i\rangle]).$$

Thus for a fixed code and a symmetric memoryless channel

$$\frac{d}{d\sigma}\mathbf{E}_h[\ln Z_n(h^n)] = \sum_{i=1}^{n}\frac{1}{2}(1+\mathbf{E}_h[\langle s_i\rangle]).$$

If furthermore the channel is Gaussian we have (18). Thus (17) implies

$$\frac{d}{d\sigma}H(X^n|Y^n) = \sum_{i=1}^{n}\frac{1}{2}(\mathbf{E}_h[\langle s_i\rangle]-1).$$

Using again an integration by parts the second derivative is equal to

$$\frac{d^2}{d\sigma^2}\mathbf{E}_h[\ln Z_n(h^n)] = \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{E}_h\left[\left(\frac{\partial}{\partial h_i} + \frac{1}{2}\frac{\partial^2}{\partial h_i^2}\right)\langle s_j\rangle\right].$$

A straightforward computation yields

$$\left(\frac{\partial}{\partial h_i} + \frac{1}{2}\frac{\partial^2}{\partial h_i^2}\right)\langle s_j\rangle = \langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle \\ -\langle s_i s_j\rangle\langle s_j\rangle + \langle s_i\rangle\langle s_j\rangle^2.$$

Using the four Nishimori identities of Section IV-B we obtain

$$\mathbf{E}_h\left[\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle - \langle s_i s_j\rangle + \langle s_i\rangle\langle s_j\rangle^2\right] \\ = \mathbf{E}_h\left[\left(\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle\right)^2\right]$$

which leads to

$$\frac{d^2}{d\sigma^2}\mathbf{E}_h[\ln Z_n(h^n)] = \mathbf{E}_h\left[\left(\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle\right)^2\right].$$

The identity (40) of the lemma now follows immediately.

Here we wish to point out a similarity between (39) and the error probability for bit decoding. In the present setting the ML or MAP estimate for the ith bit is

$$\hat{x}_i = \arg\max \sum_{x_1\dots x_{i-1},x_{i+1}\dots x_n} p_{X|Y}(x^n|y^n).$$

In the spin language this becomes (with $s_i = (-1)^{x_i}$ and $\hat{s}_i = (-1)^{\hat{x}_i}$)

$$\hat{s}_i = s_i^{\text{input}}\text{sign}\langle s_i\rangle.$$

The average fraction of wrong bits is (for a fixed code)

$$P_e = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}(1 - s_i^{\text{input}}\mathbf{E}_Y[\text{sign}\langle s_i\rangle]).$$

Because of channel symmetry one can again show that this probability does not depend on the input word, so that we may assume $s_i^{\text{input}} = 1$ (with the appropriate $\mathbf{E}_h$)

$$P_e = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}(1 - \mathbf{E}_h[\text{sign}\langle s_i\rangle]).$$

APPENDIX II
THE USE OF GAUGE INVARIANCE

In this Appendix, we give a streamlined proof of a general Nishimori identity. Then we give the list of special cases that are explicitely used in the present work.

Lemma 2.1: Consider a spin system with Hamiltonian (41) with i.i.d. coupling constants whose distribution satisfies (45). Then for any collection of subsets $X_1,\dots,X_l$ and integers $m_1,\dots,m_l$

$$\mathbf{E}_J\left[\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l}\right] \\ = \mathbf{E}_J\left[\langle s_{X_1}^{m_1}\dots s_{X_l}^{m_l}\rangle\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l}\right]. \quad (91)$$

Proof: Because of (45) the left-hand side of (91) is equal to

$$\int\prod_A dJ_A p_0(|J_A|)e^{\sum_A J_A}\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l}.$$

We make a first gauge transformation $J_Y \to \tau_Y J_Y$, $s_Y \to \tau_Y s_Y$ which shows that the last expression is equal to

$$\int\prod_A dJ_A p_0(|J_A|)e^{\sum_A J_A\tau_A}\tau_{X_1}^{m_1}\dots\tau_{X_l}^{m_l}\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l}. \quad (92)$$

We sum over $\tau^n = (\tau_1,\dots,\tau_n)$, divide by 2^N, and then insert in the integral $1 = Z_J/Z_J$ where Z_J is the partition function

$$Z_J = \sum_{\tau^n} e^{\sum_A J_A\tau_A}.$$

Then (92) becomes

$$\frac{1}{2^N}\int\prod_A dJ_A p_0(|J_A|)Z_J\langle\tau_{X_1}^{m_1}\dots\tau_{X_l}^{m_l}\rangle\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l} \\ = \frac{1}{2^N}\sum_{\lambda^n}\int\prod_A dJ_A p_0(|J_A|)e^{\sum_A J_A\lambda_A}\langle\tau_{X_1}^{m_1}\dots\tau_{X_l}^{m_l}\rangle \\ \times\langle s_{X_1}\rangle^{m_1}\dots\langle s_{X_l}\rangle^{m_l}.$$

The last step is a second gauge transformation on each term of the sum over λ^n: $J_Y \to \lambda_Y J_Y$, $s_Y \to \lambda_Y s_Y$, $\tau_Y \to \lambda_Y\tau_Y$. This yields the right-hand side of (91) □

To obtain the first identity (46) we set $l = 2k-1$ and $X_i = X$, $m_i = 1$ for $i = 1,\dots,l$. To obtain the second (47) we take two sets $X_1 = X$, $X_2 = X\cup Y$, $m_1 = m_2 = 1$. For the third we set $X_1 = X$, $X_2 = Y$, $m_1 = m_2 = 1$ and for the fourth (49) $X_1 = X$, $X_2 = Y$ and $m_1 = 2$, $m_2 = 1$.

Finaly the identity (77) is a special case of (46) for the simplest spin system consisting of a single spin $H(s_1) = J_1 s_1$, $X = \{1\}$.

APPENDIX III
RECURSIVE EVALUATION OF GIBBS AVERAGES ON TREES

The goal is to compute quantities of the type $\langle s_0\rangle_{\mathcal{N}_d(o)}$ when the neighborhood $\mathcal{N}_d(o)$ is a tree. The computation presented here for completeness is in fact equivalent to the methods fond for example in [22].

We label the tree in the following way: the variable node root is o, the set of level 1 check nodes is $L_1 = \partial o$, the set of level 2 variable nodes is $L_2 = \partial L_1\backslash o$, the set of level 3 check nodes is $L_3 = \partial L_2\backslash L_1$, and so on until the set of level $d-1$ check nodes $L_{d-1} = \partial L_{d-2} \backslash L_{d-3}$ and the set of variable node leaves $L_d = \partial L_{d-1} \backslash L_{d-2}$. Introducing partial Gibbs weights

$$W_{k+1} = e^{\sum_{j\in L_{k+1}} h_j s_j} \prod_{c\in L_k} \frac{1}{2}(1+s_{\partial c}), \quad k = 1,3,\ldots,d-1$$

the statistical sum can be organized as follows:

$$\begin{aligned} \langle s_0\rangle_{\mathcal{N}_d(o)} &= \frac{1}{Z}\sum_{s_o=\pm1} s_o e^{h_o s_o} \sum_{s_i\in L_2} W_2 \\ &\quad \cdots \sum_{s_i\in L_{d-2}} W_{d-2} \sum_{s_i\in L_d} W_d \end{aligned} \tag{93}$$

$$Z = \sum_{s_o=\pm1} e^{h_o s_o} \sum_{s_i\in L_2} W_2 \ldots \sum_{s_i\in L_{d-2}} W_{d-2} \sum_{s_i\in L_d} W_d. \tag{94}$$

The above sums can be performed in a recursive way by first summing over the spins $s_i \in L_d$, then $s_i \in L_{d-1}$ and so on. Let us do explicitly the sum over level L_d.

$$\begin{aligned} \sum_{s_j\in L_d} W_d &= \prod_{i\in L_{d-2}} \prod_{c\in\partial i\cap L_{d-1}} \sum_{s_j\in L_d} e^{\sum_{j\in\partial c\backslash i} h_j s_j} \\ &\quad \times \frac{1}{2}(1+s_i \prod_{j\in\partial c\backslash i} s_j) \\ &= \prod_{j\in L_d} 2\cosh h_j \\ &\quad \times \prod_{i\in L_{d-2}} \prod_{c\in\partial i\cap L_{d-1}} \frac{1}{2}\Big(1+s_i \prod_{j\in\partial c\backslash i} \tanh h_j\Big) \\ &= \Bigg[\frac{\prod_{j\in L_d} 2\cosh h_j}{\prod_{c\in L_{d-1}} 2\cosh u^{(2)}_{c\to i}}\Bigg] \\ &\quad \times \exp\Big(\sum_{i\in L_{d-2}} \big(\sum_{c\in\partial i\cap L_{d-1}} u^{(2)}_{c\to i}\big) s_i\Big) \end{aligned}$$

with

$$u^{(2)}_{c\to i} = \tanh^{-1}\Big(\prod_{j\in\partial c\backslash i} \tanh x^{(1)}_{j\to c}\Big), \qquad x^{(1)}_{j\to c} = h_j$$

and A a normalization constant (independent of the spins, but depending on the log-likelihoods). Here $x^{(1)}_{j\to c}$ and $u^{(2)}_{c\to i}$ have the usual interpretation of messages transmitted between variable to check and check to variable nodes. Equations (93) and (94) become

$$\begin{aligned} \langle s_0\rangle_{\mathcal{N}_d(o)} &= \frac{1}{Z}\sum_{s_o=\pm1} s_o e^{h_o s_o} \sum_{s_i\in L_2} W_2 \\ &\quad \cdots \sum_{s_i\in L_{d-2}} W_{d-3} \sum_{s_i\in L_{d-2}} W'_{d-2} \\ Z &= \sum_{s_o=\pm1} e^{h_o s_o} \sum_{s_i\in L_2} W_2 \ldots \sum_{s_i\in L_{d-2}} W_{d-3} \sum_{s_i\in L_{d-2}} W'_{d-2}. \end{aligned} \tag{95}$$

Now the tree has one level less and the Gibbs weight of the last level is

$$\begin{aligned} W'_{d-2} &= e^{\sum_{i\in L_{d-2}} (h_i + \sum_{c\in\partial i\cap L_{d-1}} u^{(2)}_{c\to i}) s_i} \prod_{c\in L_{d-3}} \frac{1}{2}(1+s_{\partial c}) \\ &= \prod_{i\in L_{d-4}} \prod_{c\in\partial i\cap L_{d-3}} e^{\sum_{j\in\partial c\backslash i} x^{(3)}_{j\to c} s_j} \prod_{c\in L_{d-3}} \frac{1}{2}(1+s_{\partial c}) \end{aligned}$$

with

$$x^{(3)}_{j\to c} = h_j + \sum_{c\in\partial j\cap L_{d-1}} u^{(2)}_{c\to j}.$$

Iterating this computation we find

$$\langle s_0\rangle_{\mathcal{N}_d(o)} = \tanh\big(h_o + \sum_{c\in\partial o} u^{(d)}_{c\to o}\big)$$

where $u^{(d)}_{c\to o}$ is given by the message passing equations

$$\begin{aligned} u^{(t+2)}_{c\to i} &= \tanh^{-1}\Big(\prod_{j\in\partial c\backslash i} \tanh x^{(t+1)}_{j\to c}\Big) \\ x^{(t+1)}_{j\to c} &= h_j + \sum_{c'\in\partial j\backslash c} u^{(t)}_{c'\to j} \end{aligned}$$

with the initial condition $x^{(1)}_{j\to c} = h_j$, $u^{(0)}_{c\to j} = 0$. The probability distribution of the messages evolves according to

$$\begin{aligned} \hat{\zeta}^{(t+2)}(u) &= \mathbf{E}^{(t+1)}_{x_1} \ldots \mathbf{E}^{(t+1)}_{x_{k-1}} \Big[\delta\big(u - \tanh^{-1}(\prod_{j=1}^{k-1} \tanh x_j)\big)\Big] \\ \zeta^{(t+1)}(x) &= \mathbf{E}_h \mathbf{E}_l \mathbf{E}^{(t)}_{u_1} \ldots \mathbf{E}^{(t)}_{u_{l-1}} \Big[\delta\big(x - (h + \sum_{c=1}^{l-1} u_c)\big)\Big] \end{aligned}$$

with the initial condition $\zeta^{(1)}(x) = p(x)$ (in our case the Gaussian distribution of log-likelihoods) and $\zeta^{(0)} = \delta(u)$. The average value of the spin at the root is

$$\begin{aligned} &\mathbf{E}_{C,h}\big[\langle s_0\rangle_{\mathcal{N}_d(o)} | \mathcal{N}_d(0) \text{ is tree}\big] \\ &\qquad = \mathbf{E}_h \mathbf{E}_l \mathbf{E}^{(d)}_{u_1} \ldots \mathbf{E}^{(d)}_{u_l} \Big[\tanh\big(h + \sum_{c=1}^{l} u_c\big)\Big]. \end{aligned}$$

We end the appendix by remarking that (19) is obtained as a special case by specifying the message passing equations to the case of a regular tree with constant initial condition $x^{(1)}_{j\to c} = h$, $u^{(2)}_{c\to j} = 0$. Then (dropping the $j \to c$ and $c \to j$ subscript)

$$\begin{aligned} u^{(t+2)} &= \tanh^{-1}(\tanh x^{(t+1)})^k \\ x^{(t+1)} &= h + (l-1)u^{(t)}. \end{aligned}$$

Using "conjugate variables" $x_{t+1} = \tanh x^{(t+1)}$ and $y_{t+2} = \tanh u^{(t+2)}$ we get

$$\begin{aligned} y_{t+2} &= (x_{t+1})^{k-1} \\ x_{t+1} &= \tanh(h + (l-1)\tanh^{-1} y_t). \end{aligned}$$

The initial condition is now $x_1 = \tanh h$ and $y_0 = 0$.

Appendix IV
Existence of a Limiting Measure Under Density Evolution

We show that the sequence of probability measures $\zeta^d(x)dx$, $\hat{\zeta}^d(u)du$ have a limit. It is sufficient to show for any integer m the existence of the limit

$$\lim_{d\to+\infty} \mathbf{E}_u^{(d)}[(\tanh u)^m].$$

Indeed, then the moments of the random variable $v = \tanh^{-1} u$ all have a well defined limit and are bounded. Thus by a criterion of Carleman, they define a unique probability measure (whose formal density we called $\hat{\zeta}^*(u)$). The first density evolution for the moments

$$\mathbf{E}_u^{(d+2)}\left[(\tanh u)^m\right] = \left(\mathbf{E}_x^{(d+1)}\left[(\tanh x)^m\right]\right)^{k-1} \qquad (96)$$

implies that the moments $\mathbf{E}_x^{(d)}[(\tanh x)^m]$ converge to a limit which defines uniquely a probability measure (with formal density $\zeta(x)$).

To prove that the limit of the moments exists we use a GKS inequality. The tree $\mathcal{N}_d(o)$ is the union of l subtrees containing the edges $(o,c), c \in \partial o$. We take any one of these subtrees, call it $\mathcal{N}_d(o,c)$, and consider the sequence of such trees as d increases by two units. By GKS

$$\mathbf{E}_{C,h}[\langle s_o\rangle^m_{\mathcal{N}_d(o,c)} | \mathcal{N}_d(o) \text{is a tree}] \\ \leq \mathbf{E}_{C,h}[\langle s_o\rangle^m_{\mathcal{N}_{d+2}(o,c)} | \mathcal{N}_{d+2}(o) \text{is a tree}].$$

Note that here we really invoke a slight generalization of GKS because we consider all integer moments: this case is covered by the results in [16]. A calculation similar to the one performed in Appendix C shows that

$$\mathbf{E}_{C,h}[\langle s_o\rangle^m_{\mathcal{N}_d(o,c)} | \mathcal{N}_d(o) \text{is a tree}] = \mathbf{E}_u^{(d)}\left[(\tanh u)^m\right].$$

Therefore the moments form an increasing bounded sequence which converges.

Appendix V
Existence of the Limit $\lim_{t\to\infty} f_{rs}[\zeta^{(t+1)}, \zeta^{(t+2)}]$

Using the density evolution equation, the replica symmetric free energy can be expressed in the form

$$\begin{aligned} f_{\mathrm{RS}}&[\zeta^{(t+1)}, \zeta^{(t+2)}] \\ &= (1-r)(k-1)\ln 2 \\ &\quad - (1-r)(k-1)\mathbf{E}_{x_1}^{(t+1)} \dots \mathbf{E}_{x_k}^{(t+1)}\left[\ln(1 + \prod_{i=1}^{k} \tanh x_i\right] \\ &\quad + \mathbf{E}_l \mathbf{E}_h \mathbf{E}_{u_1}^{(t+2)} \dots \mathbf{E}_{u_l}^{(t+2)}\left[\ln 2\cosh(h + \sum_{c=1}^{l} u_c)\right. \\ &\quad \left. - \ln \prod_{c=1}^{l} 2\cosh u_c\right]. \end{aligned} \qquad (97)$$

We will prove that each of the two separate terms on the right-hand side has a limit.

We start with the first one (call it A_t) which is more straightforward. Expanding the logarithm and using a Nishimori identity (see for example (77)) we find

$$A_t = (1-r)(k-1)\sum_{q=1}^{\infty} \frac{1}{2q(2q-1)}(\mathbf{E}_x^{(t+1)}[(\tanh x)^{2q})])^k.$$

By the results of appendix IV each term of the series has a well defined limit as $t \to \infty$, so by dominated convergence this is also the case for A_t.

For the second term (call it B_t) the main idea is to represent it as the "difference of two free energies." This difference is then related to a "magnetization" on which we can apply a GKS argument. Consider a realization of the tree of depth t, $T_0 = \mathcal{N}_t(o)$ with root o and the spin system on the tree with degree l at the root. There are lk subtrees rooted at the nodes of level L_2. We call these lk subtrees $T_1, \dots T_{lk}$. Let us denote $Z(T_i)$, $i = 0, 1, \dots, lk$ the partition functions of the spin systems on each tree. The recursive method of Appendix III leads to the formula

$$\mathbf{E}_{C,h^n}[\ln Z(T_0) - \sum_{i=0}^{lk} \ln Z(T_i) | \mathcal{N}_t(0) \text{ is tree}] = \mathrm{B_t}. \qquad (98)$$

The next observation is that if we "delete" the l check nodes of level L_1 from the tree T_0, we get a disconnected graph constituted of lk trees rooted at level L_2 plus a single point o. The free energy of the disconnected graph is for a given realization $\ln 2\cosh h_o + \sum_{i=0}^{lk} \ln Z(T_i)$. We can interpolate between T_0 and the disconnected graphs by replacing the hard constraints of level L_1 by a Gibbs weight

$$e^{\sum_{c\in L_1} J_c(s_{\partial c} - 1)}.$$

The interpolation parmeters J_c are iid gaussian with mean and variance both equal to s. This adjustment makes it possible to use Nishimori identities as well as GKS inequalities. The partition function of the spin system on the tree T_0 with soft constraints at level L_1 is denoted by Z_{soft}. We have

$$\ln Z(T_0) - \sum_{i=0}^{lk} \ln Z(T_i) - \ln 2\cosh h_o \\ = \int_0^\infty ds \frac{d}{ds} \int_{-\infty}^{\infty} \prod_{c\in L_1} dJ_c \frac{e^{\frac{(J_c - s)^2}{2s}}}{\sqrt{2\pi s}} \ln Z_{\mathrm{soft}}.$$

The derivative with respect to s is performed using (85) and then integrating by parts. This leads to

$$\ln Z(T_0) - \sum_{i=0}^{lk} \ln Z(T_i) - \ln 2\cosh h_o \\ = \int_0^\infty ds \int_{-\infty}^{\infty} \prod_{c\in L_1} dJ_c \frac{e^{-\frac{(J_c - s)^2}{2s}}}{\sqrt{2\pi s}} \sum_{c\in L_1} \Big(\langle s_{\partial c} - 1\rangle_{\mathrm{soft}} \\ + \frac{1}{2}(\langle (s_{\partial c} - 1)^2\rangle_{\mathrm{soft}} - \langle s_{\partial c} - 1\rangle^2_{\mathrm{soft}})\Big).$$

Now we average over the h_i, use the Nishimori identities to simplify the right-hand side, and then average of the Tanner graphs given that $\mathcal{N}_d(o)$ is a tree. We obtain

$$\begin{aligned}\mathbf{E}_{C,h^n}[\ln Z(T_o) - \sum_{i=0}^{lk} \ln Z(T_i) | \mathcal{N}_d(0) \text{ is tree}] \\ - \mathbf{E}_h[\ln 2\cosh h] \\ = \int_0^\infty ds \int_{-\infty}^\infty \prod_{c\in L_1} dJ_c \frac{e^{-\frac{(J_c-s)^2}{2s}}}{\sqrt{2\pi s}} \\ \times \sum_{c\in L_1} \frac{1}{2}(\mathbf{E}_{C,h^n}[\langle s_{\partial c}\rangle_{\text{soft}} | \mathcal{N}_t(0) \text{ is tree}] - 1). \quad (99)\end{aligned}$$

Finaly combining this formula with (98)

$$\begin{aligned}B_t =& \mathbf{E}_h[\ln 2\cosh h] \\ &+ \int_0^\infty ds \int_{-\infty}^\infty \prod_{c\in L_1} dJ_c \frac{e^{\frac{(J_c-s)^2}{2s}}}{\sqrt{2\pi s}} \\ &\times \sum_{c\in L_1} \frac{1}{2}(\mathbf{E}_{C,h^n}[\langle s_{\partial c}\rangle_{\text{soft}} | \mathcal{N}_t(0) \text{ is tree}] - 1). \quad (100)\end{aligned}$$

Now GKS tells us that the Gibbs average in the integrand is monotone increasing as a function of (even) t, thus B_t is an increasing sequence. That the limit exists follows by a uniform bound most easily obtained from the formula

$$\begin{aligned}B_t =& -(1-r)k\ln 2 \\ &+ \mathbf{E}_l \mathbf{E}_h \mathbf{E}_{u_1}^{(t+2)} \ldots \mathbf{E}_{u_l}^{(t+2)} \left[\ln\left(e^h \prod_{c=1}^{l}(1+\tanh u_c)\right.\right. \\ &\left.\left. + e^{-h}\prod_{c=1}^{l}(1-\tanh u_c)\right)\right].\end{aligned}$$

ACKNOWLEDGMENT

The author would like to thank Ruediger Urbanke for very helpful discussions and encouragement. He also thanks Cyril Measson for many explanations on his work on the BEC and Henry Pfister for useful comments.

REFERENCES

[1] N. Sourlas, "Spin glass models as error correcting codes," *Nature*, vol. 339, pp. 693–695, 1989.

[2] T. Murayama, Y. Kabashima, D. Saad, and R. Vicente, "Statistical physics of regular low-density parity-check error correcting codes," *Phys. Rev. E*, vol. 62, pp. 1577–1591, 2000.

[3] A. Montanari, "The glassy phase of Gallager codes," *European Phys. J.*, vol. 23, 2001.

[4] Y. Kabashima, N. Sazuka, K. Nakamura, and D. Saad, "Evaluating zero error noise threshholds by the replica method for Gallager code ensembles," in *Proc. ISIT*, Lausanne, Switzerland, Jun. 2002, p. 255.

[5] F. Guerra and F. Toninelli, "Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model," *J. Math. Phys*, vol. 43, p. 3704, 2002.

[6] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model," *Commun. Math. Phys.*, vol. 233, pp. 1–12, 2003.

[7] M. Talagrand, "The Parisi formula," *Ann. Math.*, vol. 163, pp. 221–263, 2005.

[8] S. Franz and M. Leone, "Replica bounds for optimization problems and diluted spin systems," *J. Stat. Phys.*, vol. 111, pp. 535–564, 2003.

[9] M. Talagrand and D. Panchenko, "Bounds for diluted mean-fields spin glass models," *Prob. Theory Rel. Fields*, vol. 130, pp. 319–336, 2004.

[10] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3221–3246, 2005.

[11] C. Measson, A. Montanari, T. Richardson, and R. Urbanke, The Generalized Area Theorem and Some of its Consequences 2005 [Online]. Available: ArXiv, cs.IT/0511039, to be published

[12] N. Macris, "On the relation between MAP and BP GEXIT functions of low density parity check codes," in *Proc. IEEE Inf. Theory Workshop*, Punta del Este, Uruguay, Mar. 13–17, 2006.

[13] ——, *Sharp Bounds on Generalized EXIT Functions*, 2005, to be published.

[14] V. Rathi, "On the asymptotic weight and stopping set distribution of regular LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 52, pp. 4212–4218, 2006.

[15] R. B. Griffiths, *Phase Transitions and Critical Phenomena*, C. Domb and M. S. Green, Eds. New York: Academic, 1972, vol. 1.

[16] S. Morita, H. Nishimori, and P. Contucci, "Griffiths inequalities for the Gaussian spin glass," *J. Phys. A*, vol. 37, p. L203, 2004.

[17] N. Macris, "Correlation inequalities: A useful tool in the theory of LDPC codes," in *Proc. Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 4–9, 2005, pp. 2369–2373.

[18] H. Nishimori, *Progr. Theor. Phys.*, vol. 66, pp. 1169–1169, 1981.

[19] ——, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford, U.K.: Oxford Science, 2001.

[20] C. Measson and R. Urbanke, "An upper-bound for the ML threshold of iterative coding systems over the BEC," in *Proc. 41st Allerton Conf Commun, Contr. Comput.*, Monticello, NY, Oct. 2003, p. 3.

[21] C. Measson, A. Montanari, T. Richardson, and R. Urbanke, "Life above threshold: From list decoding to area theorem and MSE," in *IEEE Inf. Theory Workshop*, San Antonio, TX, Oct. 2004.

[22] R. Urbanke and T. Richardson, *Modern Coding Theory*, to be published.

[23] A. Ashikhmin, G. Kramer, and S. ten Brink, "Code rate and the area under extrinsic information transfer curves," in *Proc. ISIT*, Lausanne, Switzerland, Jun. 2002.

[24] D. Guo, S. Shamai, and S. Verdu, "Mutual information and MMSE in Gaussian channels," in *Proc. 2004 ISIT*, Chicago, IL, p. 347.

[25] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, pp. 533–547, 1981.

[26] F. Kschischang, B. Frey, and H. A. Loeliger, "Factor graphs and the sum product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, pp. 498–519, 2001.

[27] R. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[28] S. Listyn and V. Shevelev, "On ensembles of low-density parity-check codes: Asymptotic distance distributions," *IEEE Trans. Inf. Theory*, vol. 48, pp. 887–908, 2002.

[29] C. Di, "Asymptotic and Finite Length Analysis of Low-Density Parity-Check Codes," Thesis No 3072, Ecole Polytechnique Federale Lausanne, Lausanne, Switzerland, 2004.

[30] C. Di, A. Montanari, and R. Urbanke, "Weight distribution of LDPC codes: Combinatorics meets statistical physics," in *Proc. 2004 ISIT*, Chicago, IL.

[31] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, pp. 638–656, 2001.

[32] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, pp. 619–637, 2001.

[33] L. D. Landau and I. M. Lifshitz, *Statistical Physics*. Reading, MA: Addison Wesley, 1969.

[34] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*. Cambridge, U.K.: Cambridge University Press, 2000, ch. 8.

[35] R. Baxter, *Exactly Solved Models in Statistical Mechanics*. New York: Academic, ch. 3.

[36] S. Franz, M. Leone, and F. Toninelli, "Replica bounds for diluted non-Poissonian spin systems," *J. Phys. A*, vol. 36, pp. 10,967–10,985, 2003.