Sharp Bounds for MAP Decoding of General Irregular LDPC Codes

Shrinivas Kudekar, Nicolas Macris
LTHC, EPFL, I&C, CH-1015 Lausanne
email : shrinivas.kudekar@epfl.ch, nicolas.macris@epfl.ch

***Abstract*— Consider communication over a binary input memoryless output symmetric channel with LDPC codes and MAP decoding. Recently Montanari proved that the replica solution is a lower bound to the conditional entropy for a class of LDPC ensembles. Here we extend this lower bound to *any irregular* LDPC ensemble for the BEC, BIAWGNC, BSC. Our work combines an analysis of the second derivative of the conditional entropy with respect to the noise and the interpolation method.**

I. Introduction

Linear codes based on sparse random graphs are useful because of low-complexity decoding schemes and good performance [1],[2]. One quantity of interest is the maximum a posteriori (MAP) threshold, above which reliable communication is not possible. Fano's inequality tells us that the block error probability for a code having length n and rate r is lower bounded by $H(X^n|Y^n)/(nr)$ where $H(X^n|Y^n)$ is the entropy of the transmitted message X^n conditional to the received message Y^n. Thus lower bounds for the conditional entropy per bit $h_n = H(X^n|Y^n)/n$ also give upper bounds on the MAP threshold.

Recent techniques of statistical physics applied to communications have provided tight bounds on h_n for a class of LDPC(n, Λ, P) code ensembles. Here $\Lambda(x) = \sum_d \Lambda_d x^d$, $P(x) = \sum_k P_k x^k$ are the variable and check node degree distributions from the node perspective [1]. These results are based on the *interpolation method*, developed recently in the theory of mean field spin glasses [3], [4]. In [5] the interpolation method is extended to the standard LDPC ensembles with any polynomial $\Lambda(x)$ but $P(x)$ restricted to be a convex polynomial in a region $-e \leq x \leq e$. In particular if the right degree is constant this means it has to be *even*. *In this paper we drop the convexity requirement of $P(x)$ for the BEC, BIAWGNC and BSC. Our result holds for any standard regular or irregular code ensemble*. The general scheme of our analysis holds for any binary memoryless symmetric (BMS) channel - except for one inequality limited to BEC, BIAWGNC, BSC - and we expect that it should be possible to further extend the present results to more general channels.

We consider communication through a BMS channel with transition probability $p_{Y|X}(y|x)$ and noise parameter ϵ understood to vary in the appropriate range. We will work in terms of both the likelihood $l = \ln \frac{p_{Y|X}(y|1)}{p_{Y|X}(y|0)}$ and difference $t = \tanh \frac{l}{2}$ domains. Let V be some random variable with an arbitrary *symmetric* density $d_V(v)$. In the present context a r.v is said to be symmetric if $d_V(v) = e^v d_V(-v)$. Also let $U = \tanh^{-1}\left[\prod_{i=1}^{k-1} \tanh V_i\right]$ where V_i are i.i.d copies of V and k is the degree of check nodes distributed as P_k (the degree of variable nodes d is distributed as Λ_d). Notice that the r.v U appears as the check node message in the belief propagation (BP) decoding algorithm. The mathematically ill defined *replica calculations* of spin glass theory lead to a conjectured formula for the entropy

$$\mathsf{E}_{\mathcal{C}}[h_n] = \sup_{d_V} h_{RS}[d_V]$$

where the functional $h_{RS}[d_V]$ is known as the "replica symmetric" or "trial" entropy

$$\begin{aligned} h_{RS}[d_V] = &-\Lambda'(1)\mathsf{E}_{V,U}\big[\ln(1+\tanh V \tanh U)\big] \\ &+ \mathsf{E}_{l,d,U_c}\Big[\ln\Big(e^{\frac{l}{2}}\prod_{c=1}^{d}(1+\tanh U_c) \\ &+ e^{-\frac{l}{2}}\prod_{c=1}^{d}(1-\tanh U_c)\Big)\Big] \\ &+ \frac{\Lambda'(1)}{P'(1)}\mathsf{E}_{k,V_i}\big[\ln(1+\prod_{i=1}^{k}\tanh V_i)\big] - \frac{\Lambda'(1)}{P'(1)}\ln 2 \end{aligned}$$

Our main result is

Theorem 1.1: Assume communication using a standard irregular LDPC(n, Λ, P) code ensemble, through a BEC or BIAWGNC with any noise level and a BSC with crossover parameter $0.421 \leq \epsilon < 0.5$. For almost all ϵ in the above ranges we have,

$$\liminf_{n\to+\infty} \mathsf{E}_{\mathcal{C}}[h_n] \geq \sup_{d_V} h_{RS}[d_V]$$

It is strongly suspected that in fact the equality holds, so a numerical implementation of the replica expression provides a practical means to compute a precise value for the MAP threshold. Besides, the equations for the critical points of the functional $h_{RS}[d_V]$ are closely related to *density evolution*, so the above bound - and the conjectured equality - are of theoretical importance for the elucidation of the relationship between MAP and BP decoding.

Our proof of the theorem uses the *second* derivative of $\mathsf{E}_{\mathcal{C}}[h_n]$ with respect to the noise parameter. By analyzing this second derivative we are able to extend the domain of applicability of the interpolation method. The first derivative of the conditional entropy is known as the Generalized EXIT function [6] and plays an important role in the analysis of LDPC codes (namely the relationship between MAP and BP decoding). We find it interesting to see that the second derivative seems to also be of some use, and are not aware whether it has been investigated before.

II. Derivatives of The Conditional Entropy

A. Conditional Entropy and Free Energy

Consider a fixed code of an LDPC(n, Λ, P) ensemble. The Tanner graph has variable nodes, denoted by (i,j,...), that are connected to check nodes c. The *posterior distribution* $p_{X^n|Y^n}(x^n|y^n)$ used in MAP decoding can be viewed as the Gibbs measure of a particular random spin system. For this it is convenient to use the usual mapping of bits onto spins $\sigma_i = (-1)^{x_i}$. For a uniform prior over the code words and a BMS channel, Bayes rule implies $p_{X^n|Y^n}(x^n|y^n) = \mu_{\mathcal{C}}(\sigma^{(n)})$ with

$$\mu_{\mathcal{C}}(\sigma^{(n)}) = \frac{1}{Z_{\mathcal{C}}} \prod_{c\in\mathcal{C}} \frac{1}{2}(1+\sigma_{\partial c}) \prod_{i=1}^{n} e^{\frac{l_i}{2}\sigma_i}$$

where $\sigma_{\partial c} = \prod_{i\in c} \sigma_i$ and $Z_{\mathcal{C}}$ is the normalization factor or partition function. We can assume that the input is the all zero codeword which induces a distribution $c(l)$ for the likelihood variables (this is ϵ dependent). Expectations with respect to the Gibbs measure for a fixed graph and a fixed channel output are denoted by the bracket $\langle - \rangle$. More precisely for any $X \subset \{1, ..., n\}$,

$$\langle \sigma_X \rangle = \sum_{\sigma^n} \sigma_X \mu_{\mathcal{C}}(\sigma^n), \qquad \sigma_X = \prod_{i\in X} \sigma_i$$

Expectations with respect to the code ensemble and the channel outputs will be denoted by $\mathsf{E}_{\mathcal{C},l^n}[-]$.

The correspondence between communications and statistical mechanical quantities is as follows. The extrinsic soft bit estimate is

$$p_{X_1|Y^n\setminus Y_1}(0|y^n \setminus y_1) - p_{X_1|Y^n\setminus Y_1}(1|y^n \setminus y_1) = \langle \sigma_1 \rangle_0$$

where the subscript in the Gibbs bracket indicates that it is computed for $l_1 = 0$. We will also need extrinsic soft estimates for $X_1 \oplus X_j$, $j \neq 1$. These are computed with $p_{X_1\oplus X_j|Y^n\setminus Y_1,Y_j}$. In the statistical mechanics formalism they are simply expressed as $\langle \sigma_1\sigma_j \rangle_{00}$ where the subscript in the Gibbs bracket means that we set $l_1 = l_j = 0$. It is possible to show [7]

$$\mathbb{E}_{\mathcal{C}}[h_n] = \frac{1}{n}\mathbb{E}_{\mathcal{C},l^n}[\ln Z_{\mathcal{C}}] - \int_{-\infty}^{+\infty} dl c(l) \frac{l}{2}$$

Therefore the evaluation of the conditional entropy and soft bit estimate reduce to that of the average free energy $\frac{1}{n}\mathbb{E}_{\mathcal{C},l^n}[\ln Z_{\mathcal{C}}]$ and magnetization $\langle \sigma_1 \rangle_0$ of the corresponding spin system.

B. First Derivative

Differentiating the relationship between entropy and free energy we get [8],

$$\frac{d}{d\epsilon}\mathbb{E}_{\mathcal{C}}[h_n] = \int_{-\infty}^{+\infty} dl_1 \frac{dc}{d\epsilon}(l_1) g_1(t_1) \qquad (1)$$

where we recall the notation $t_1 = \tanh\frac{l_1}{2}$ and

$$g_1(t_1) = \mathsf{E}_{\mathcal{C},l^n\setminus 1}\left[\ln\left(\frac{1 + t_1\langle\sigma_1\rangle_0}{1+t_1}\right)\right]$$

Note that this is nothing else than the GEXIT function [6].

Channel symmetry implies a set of general *Nishimori identities* [9]. Expanding the logarithm and using these identities we obtain a useful power series expansion for (1),

$$\sum_{k=1}^{\infty} \frac{m_1^{(2k)}}{2k(2k-1)} \mathsf{E}_{\mathcal{C},l^n\setminus 1}[\langle\sigma_1\rangle_0^{2k} - 1] \qquad (2)$$

where

$$m_1^{(2k)} = \int_{-\infty}^{+\infty} dl_1 \frac{dc}{d\epsilon}(l_1) t_1^{2k}$$

The derivation of (2) and explanations on Nishimori identities can be found in [10].

C. Second Derivative of the Conditional Entropy

A calculation of the second derivative yields the formula

$$\frac{d^2}{d\epsilon^2}\mathbb{E}_{C}[h_n] = \int_{-\infty}^{+\infty} dl_i \frac{d^2c}{d\epsilon^2}(l_1) g_1(t_1)$$
$$+ \sum_{j\neq 1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} dl_1 dl_j \frac{dc}{d\epsilon}(l_1) \frac{dc}{d\epsilon}(l_j) g_2(t_1, t_j) \qquad (3)$$

with

$$g_2(t_1, t_j) = \mathsf{E}_{\mathcal{C},l^n\setminus 1,j}\Bigg[$$
$$\ln\left(\frac{1 + t_1\langle\sigma_1\rangle_{00} + t_j\langle\sigma_j\rangle_{00} + t_1t_j\langle\sigma_1\sigma_j\rangle_{00}}{1 + t_1\langle\sigma_1\rangle_{00} + t_j\langle\sigma_j\rangle_{00} + t_1t_j\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}}\right)\Bigg]$$

Remarkably the same formula (3) hold if in g_1, g_2 we replace the brackets $\langle - \rangle_0$, $\langle - \rangle_{00}$ by $\langle - \rangle$, and replace $g_1(t_1), g_2(t_1, t_j)$ by $-g_1(-t_1), -g_2(-t_1, -t_j)$.

Expanding the logarithms, using the Nishimori identities and reorganizing the series suitably we get a useful expansion for

$$\frac{d^2}{d\epsilon^2}\mathbb{E}_{C}[h_n] = S_1 + S_2$$

The term S_1 has the same form as (2) with $\frac{d^2c}{d\epsilon^2}(l_1)$ replacing $\frac{dc}{d\epsilon}(l_1)$. The second part S_2 is obtained after some tedious algebra,

$$\sum_{j\neq 1}\sum_{l=1}^{\infty}\sum_{k\geq l} m_1^{(2k-1)} m_j^{(2l-1)} \mathsf{E}_{\mathcal{C},l^n\setminus 1,j}\Big[\big(\langle\sigma_1\sigma_j\rangle_{00}$$

$$-\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\big)^2\langle\sigma_1\rangle_{00}^{2k-2l}\sum_{r=0}^{2l-2}A_{r,l}\langle\sigma_1\rangle_{00}^r\langle\sigma_j\rangle_{00}^r$$

$$\times\big(\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}-\langle\sigma_1\sigma_j\rangle_{00}\big)^{2l-r-2}\Big]$$

$$+\sum_{j\neq 1}\sum_{k=1}^{\infty}\sum_{l>k} m_1^{(2k-1)} m_j^{(2l-1)}\big((1\leftrightarrow j),(k\leftrightarrow l)\big)$$

where $A_{r,l}=\dfrac{1}{(2l)!}\binom{2l-2}{r}[2l]_r[2k-2]_{2l-2-r}$

and $[m]_r=(m)\cdots(m-r+1)$. This expansion, which is used later on, can be shown to converge.

III. THE DECAY OF SPIN-SPIN CORRELATIONS

One expects that in the limit $n\to\infty$ $\mathbb{E}_{\mathcal{C}}[h_n]$ remains continuous with a finite jump in the first derivative at the MAP (or phase transition) threshold(s). In other words the first derivative should remain finite uniformly in n. This is the content of the following lemma.

Lemma 3.1: For BEC, BSC and BIAWGN we have

$$0\leq\frac{d}{d\epsilon}\mathbb{E}_C[h_n]\leq a$$

where $a=1$ for BIAWGN, $a=2\ln 2$ for BEC and $a=(1-2\epsilon)/(2\epsilon(1-\epsilon))$ for BSC.

Proof: Using $|\langle\sigma_1\rangle_{00}|\leq 1$ and compute $m_1^{(2k)}$ in the expansion (2) we obtain the upper bound. The lower bound follows from $m_1^{(2k)}\leq 0$ [1] and $\langle\sigma_1\rangle_0^{2k}\leq 1$. ■

The second derivative remains finite except at the thresholds where it diverges as a function of n. This divergence is intimately related to the absence of decay of the spin-spin correlation $\langle\sigma_1\sigma_j\rangle-\langle\sigma_1\rangle\langle\sigma_j\rangle$ as a function of the distance between nodes 1 and j (on the Tanner graph the distance between two nodes is the length of the shortest path joining them). This is basically the content of lemma 3.2.

Depending on the situation it is more convenient to work with the brackets $\langle-\rangle_{00}$ or $\langle-\rangle$. This is why we will also need the following which we state here without proof.

Lemma 3.2: For any BMS channel there is a function $R(l_1,l_j)$ such that

$$\langle\sigma_1\sigma_j\rangle-\langle\sigma_1\rangle\langle\sigma_j\rangle=R(l_1,l_j)\big(\langle\sigma_1\sigma_j\rangle_{00}-\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\big)$$

In particular for the BEC with any noise level and the BSC with $0.421\leq\epsilon<0.5$ we have

$$R(l_1,l_j)\leq\rho$$

where ρ is a finite positive constant.

Remark: for the BIAWGNC we do not need this lemma. For the BEC and BSC the values of ρ are 1 and 16.83.

Lemma 3.3: For the BEC and the BIAWGNC with any noise level and the BSC with $0.421\leq\epsilon<0.5$, there exist finite positive constants b and c, possibly dependent on ϵ, such that

$$\frac{d^2}{d\epsilon^2}\mathbb{E}_C[h_n]\geq -c+b\sum_{j\neq 1}\mathsf{E}_{\mathcal{C},l^n}\left[(\langle\sigma_1\sigma_j\rangle-\langle\sigma_1\rangle\langle\sigma_j\rangle)^2\right]\tag{4}$$

Remark: A similar upper bound holds with other constants c',b'.

Proof: The proof can be based on the expansion of $\frac{d^2}{d\epsilon^2}\mathbb{E}_C[h_n]$. However for the BIAWGNC and the BEC we have a more elegant argument.

BIAWGNC. Using the identity $\frac{dc}{d\epsilon}(l)=2\epsilon^{-2}(c'(l)-c''(l))$ and integration by parts in the formula (3) expressed with the bracket $\langle-\rangle$, leads to the simple expression

$$\frac{d^2}{d\epsilon^2}\mathbb{E}_C[h_n]=\frac{1}{2}\sum_{j=1}^{n}\mathsf{E}(\langle\sigma_1\sigma_j\rangle-\langle\sigma_1\rangle\langle\sigma_j\rangle)^2$$

Thus we get (4) with $c=0$ and $b=\frac{1}{2}$.

BEC. We have $\frac{dc(l)}{d\epsilon}=\delta_0(l)-\delta_\infty(l)$, $\frac{d^2c(l)}{d\epsilon^2}=0$ so the second derivative is equal to

$$\sum_{j\neq 1}\mathsf{E}_{l^n\setminus 1,j}\left[\ln\frac{1+\langle\sigma_1\rangle_{00}+\langle\sigma_j\rangle_{00}+\langle\sigma_1\sigma_j\rangle_{00}}{1+\langle\sigma_1\rangle_{00}+\langle\sigma_j\rangle_{00}+\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}}\right]$$

On the BEC the spin system has positive coupling constants so that we can apply the Griffiths-Kelly-Sherman correlation inequalities [10] we know that $\langle\sigma_1\sigma_j\rangle_{00}-\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\geq 0$, $\langle\sigma_1\rangle_{00}\geq 0$, $\langle\sigma_j\rangle_{00}\geq 0$. Thus

$$\begin{aligned}1&\leq 1+\langle\sigma_1\rangle_{00}+\langle\sigma_j\rangle_{00}+\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\\&\leq 1+\langle\sigma_1\rangle_{00}+\langle\sigma_j\rangle_{00}+\langle\sigma_1\sigma_j\rangle_{00}\leq 4\end{aligned}$$

Inequality (4) then follows from $\ln u-\ln v\geq\frac{1}{4}(u-v)$ for $4\geq u\geq v\geq 1$ and lemma 3.2. We get $c=0$ and $b=1/4$.

BSC. Thanks to an expansion similar to (2) it is easy to show that the contribution S_1 to (3) is greater than $-c$ for $c=1/(2\epsilon(1-\epsilon))$. Let us now look at the contribution from S_2. In the expansion of the later we can isolate the term $k=l=1$

$$16(1-2\epsilon)^2\sum_{j\neq 1}\mathsf{E}_{\mathcal{C},l^n\setminus 1,j}\left[\big(\langle\sigma_1\sigma_j\rangle_{00}-\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\big)^2\right]$$

If we could prove that the rest of the series is strictly positive we would have the result for all values of ϵ. We have been unable to show this but we can easily bound the power series expansion term by term to show that it cannot be more negative than

$$-\frac{16(1-2\epsilon)^2}{2(1-\frac{25}{4}(1-2\epsilon)^2)^4}$$
$$\times\sum_{j\neq 1}\mathsf{E}_{\mathcal{C},l^{n\setminus 1,j}}\big[\big(\langle\sigma_1\sigma_j\rangle_{00}-\langle\sigma_1\rangle_{00}\langle\sigma_j\rangle_{00}\big)^2\big]$$

Combining these remarks with lemma 3.2 we obtain (4) with $b = 0.028(1-2\epsilon)^2\big[2-\frac{1}{(1-(2.5(1-2\epsilon))^2)^4}\big] > 0$ as long as $0.421 \le \epsilon < 0.5$. ■

IV. Proof of Theorem by Interpolation Method

A. A Brief Survey

We use the interpolation method in the form developed by Montanari. As explained in [5] it is difficult to deal directly with general irregular ensembles. Rather one introduces a *multi-poisson ensemble* which approximates the general ensemble. Once the bounds are derived for the multi-Poisson ensemble a limiting procedure permits to extend them to the general irregular ensemble. The multi-Poisson ensemble is a technical elaboration of the *Poisson ensemble* and due to lack of space we present the analysis here for the later. The extension of the estimates that follow to the multi-Poisson ensemble and thus to the standard irregular ensembles does not involve any extra difficulty except for technicalities.

The Poisson ensemble LDPC($n, 1-r, P$) has a fixed n number of variable nodes while the number of check nodes is Poisson with mean $n(1-r)$ where r is a fixed design rate. The variables nodes are connected to checks uniformly at random and their degree becomes Poisonnian as $n \to \infty$. The check node degree is distributed according to $P(x)$.

The main idea behind the interpolation technique is to recursively remove the check node constraints and compensate for the change of rate with extra observations U_c coming from an auxiliary channel. More precisely let $s \in [0,1]$ be an interpolating parameter. At "time" s the number of check nodes is a Poisson r.v with mean $n(1-r)s$ and variable nodes i receive d_i extra observations $\{U_a^i\}$ which are i.i.d copies of the r.v U. For each i, d_i is a Poisson r.v with mean $n(1-r)(1-s)$. The interpolating Gibbs measure is

$$\mu_{\mathcal{C}_s}(\sigma^n)=\frac{1}{Z_{\mathcal{C}_s}}\prod_{c\in\mathcal{C}_s}\frac{1}{2}(1+\sigma_{\partial c})\prod_{i=1}^{n}e^{(\frac{l_i}{2}+\sum_{c=1}^{d_i}U_a^i)\sigma_i}$$

where $\mathcal{C}_s$ is a Tanner graph at "time" s in LDPC($n,(1-r)s,P$). At $s = 1$ one recovers the original measure while at $s = 0$ we have a simple product measure which is tailored to yield the replica symmetric entropy $h_{RS}[d_V]$ up to a remainder term.

The central result that we use is

$$\mathsf{E}_{\mathcal{C}}[h_n]=h_{RS}[d_V]+\int_0^1 R_n(s)ds$$

The remainder term $R_n(s)$ is given by

$$R_n(s)=\sum_{p=1}^{\infty}\frac{1}{2p(2p-1)}$$
$$\times\,\mathsf{E}[\langle P(Q_{2p})-P'(q_{2p})(Q_{2p}-q_{2p})-P(q_{2p})\rangle_{2p,s}] \tag{5}$$

where $q_p = \mathsf{E}_V[(\tanh V)^p]$ and Q_p are *overlap parameters* defined as

$$Q_p=\frac{1}{n}\sum_{i=1}^{n}\sigma_i^{(1)}\sigma_i^{(2)}\cdots\sigma_i^{(p)} \tag{6}$$

Here $\sigma_i^{(\alpha)}, \alpha = 1, 2, \ldots, p$ are p independent copies (replicas) of the spin σ_i and $\langle - \rangle_{p,s}$ is the Gibbs bracket associated to the product measure (replica measure) $\prod_{\alpha=1}^{p}\mu_{\mathcal{C}_s}(\sigma_1^{(\alpha)}...\sigma_n^{(\alpha)})$. Finally we use the shorthand $\mathsf{E}[-]$ for the expectation with respect to $\mathcal{C}_s$, l^n, d_i, U_a^i.

B. Conjecture on Overlaps

We conjecture the following:

Conjecture 4.1: For any BMS channel, there exists a small enough number $\delta > 0$ such that for *Lebesgue almost every* ϵ

$$\lim_{n\to\infty}\int_0^1 ds\mathrm{P}\big[|Q_p^k-\langle Q_p\rangle_{p,s}^k|>\frac{p}{n^\delta}\big]=0 \tag{7}$$

where P is the probability distribution $\mathsf{E}\langle 1_X\rangle_{p,s}$.

Using lemmas 3.1 and 3.3 we will prove this conjecture for the BIAWGNC, BEC and the BSC (in the appropriate noise interval). At the threshold values of the noise one expects the overlap fluctuations to grow because the spin-spin correlation does not decay and this is why we have the *almost all* ϵ condition.

C. Proof of Main Theorem

Our aim is to show that

$$\liminf_{n\to+\infty}\int_0^1 R_n(s)ds\ge 0 \tag{8}$$

Since $|Q_{2p}| \le 1$ and $|q_{2p}| \le 1$ the contribution of the terms with $2p > n^\delta$ is smaller than $O(n^{-\delta})\sum_k kP_k$. Thus this contribution tends to 0 as $n \to +\infty$ and it is sufficient to look at the terms with $2p < n^\delta$.

Consider the term inside the Gibbs bracket of (5)

$$P(Q_{2p})-P'(q_{2p})(Q_{2p}-q_{2p})-P(q_{2p})$$
$$=\sum_k P_k\big(Q_{2p}^k-kQ_{2p}q_{2p}^{k-1}+(k-1)q_{2p}^k\big) \tag{9}$$

Even degrees k. For these terms we use the standard argument: the convexity of the function x^k on the whole real line implies $(Q_{2p}^k - kQ_{2p}q_{2p}^{k-1} + (k-1)q_{2p}^k) \ge 0$. Therefore the contribution of even terms to the remainder is non negative.

Odd degrees k. We decompose the Gibbs bracket as

$$\langle Q_{2p}^k-kQ_{2p}q_{2p}^{k-1}+(k-1)q_{2p}^k\rangle_{2p,s}=C_{2p}+F_{2p}$$

where

$$C_{2p} = \langle Q_{2p}\rangle_{2p,s}^k - k\langle Q_{2p}\rangle_{2p,s} q_{2p}^{k-1} + (k-1)q_{2p}^k$$

and

$$F_{2p} = \langle Q_{2p}^k\rangle_{2p,s} - \langle Q_{2p}\rangle_{2p,s}^k$$

Since $\langle Q_{2p}\rangle$ is positive, the convexity of x^k on the *positive real axis* (remember k is odd) implies that $C_{2p} \geq 0$ and the contribution to the remainder is non negative. The fluctuation term F_{2p} on the other hand can be negative. However we can control its effect thanks to (7). Its contribution to the remainder is

$$\Big| \sum_{2p<n^\delta} \frac{1}{2p(2p-1)} \mathsf{E}[\langle Q_{2p}^k\rangle_{2p,s} - \langle Q_{2p}\rangle_{2p,s}^k] \Big|$$
$$\leq \sum_{2p<n^\delta} \frac{1}{2p(2p-1)} \frac{2p}{n^\delta} + \sum_{2p<n^\delta} \frac{2}{2p(2p-1)}$$
$$\times \mathsf{P}[|Q_{2p}^k - \langle Q_{2p}^k\rangle_{2p,s}| > \frac{2p}{n^\delta}]$$

We can bound the first term above by $O(n^{-\delta}\ln n)$. From (7) and dominated convergence we get that the s integral of the second term goes to zero as $n \to \infty$.

Combining all the above results we obtain (8) and thus the theorem.

V. Overlap Fluctuation For BEC, BSC, BIAWGN

In this section we sketch the proof of the conjecture (7) for these three channels. The proof rests on lemmas 3.1 and 3.3. The identity

$$b^k - a^k = (b-a)\sum_{l=0}^{k-1} b^{k-l-1}a^l \tag{10}$$

and $|Q_p| \leq 1$ imply $|Q_p^k - \langle Q_p\rangle_p^k| \leq k|(Q_p - \langle Q_p\rangle_{p,s})|$. Then by Chebychev's inequality

$$\mathsf{P}[|Q_p^k - \langle Q_p\rangle_{p,s}^k| > \frac{p}{n^\delta}] \leq \mathsf{P}[|Q_p - \langle Q_p\rangle_{p,s}| > \frac{p}{kn^\delta}]$$
$$\leq O(\frac{4k^2n^{2\delta}}{p^2})\mathsf{E}[\langle Q_p^2\rangle_{p,s} - \langle Q_p\rangle_{p,s}^2]$$

Writing down explicitly the overlaps and using again (10) we have

$$\mathsf{E}[\langle Q_p^2\rangle_{p,s} - \langle Q_p\rangle_{p,s}^2]$$
$$\leq O(\frac{p}{n^2}) \sum_{i,j=1}^n \mathsf{E}[|\langle\sigma_i\sigma_j\rangle_s - \langle\sigma_i\rangle\langle\sigma_j\rangle_s|]$$

Swchartz's inequality shows that the right hand side is smaller than

$$O(\frac{p}{n})\Bigg(\sum_{i,j=1}^n \mathsf{E}\big[\big(\langle\sigma_i\sigma_j\rangle_s - \langle\sigma_i\rangle_s\langle\sigma_j\rangle_s\big)^2\big]\Bigg)^{1/2}$$

Finally we arrive at

$$\mathsf{P}[|Q_p^k - \langle Q_p\rangle_{p,s}^k| > \frac{p}{n^\delta}]$$
$$\leq O(\frac{k^2n^{2\delta-\frac{1}{2}}}{p})\Bigg(\sum_{j=1}^n \mathsf{E}[(\langle\sigma_1\sigma_j\rangle_s - \langle\sigma_1\rangle_s\langle\sigma_j\rangle_s)^2]\Bigg)^{\frac{1}{2}}$$

The results of the preceding sections are also valid for the bracket $\langle - \rangle_s$ because since V is a symmetric random variable U_a^i also is. Therefore from lemma 3.3 we have

$$\mathsf{P}\Big[|Q_p^k - \langle Q_p\rangle_p^k| > \frac{p}{n^\delta}\Big]$$
$$\leq O(\frac{k^2n^{2\delta-\frac{1}{2}}}{p})\Bigg(1 + cb^{-1} + b^{-1}\frac{d^2}{d\epsilon^2}\mathsf{E}_{\mathcal{C}}[h_n]\Bigg)^{\frac{1}{2}}$$

Let $\psi(\epsilon)$ be a C_0^∞ positive normalized test function. Using the Schwartz inequality, integration by parts over ϵ, and lemma 3.1 we can show

$$\int d\epsilon\psi(\epsilon)\Bigg(1 + cb^{-1} + b^{-1}\frac{d^2}{d\epsilon^2}\mathsf{E}_{\mathcal{C}}[h_n]\Bigg)^{1/2}$$
$$\leq 1 + cb^{-1} + b^{-1}\Big|\int d\epsilon\psi'(\epsilon)\frac{d}{d\epsilon}\mathsf{E}_{\mathcal{C}}[h_n]\Big|^{1/2}$$
$$\leq O(1)\big(1 + \int d\epsilon|\psi'(\epsilon)|\big)^{1/2}$$

From this we conclude (using Fubini and dominated convergence) that for $0 < \delta < \frac{1}{4}$

$$\int d\epsilon\psi(\epsilon)\lim_{n\to\infty}\int_s^\gamma ds\mathsf{P}[|Q_p^k - \langle Q_p\rangle_p^k| > \frac{p}{n^\delta}] = 0$$

Since this is true for any positive test function we conclude that (7) holds for almost every ϵ. Note that this proof would work for any channel satisfying (4).

Acknowledgment

The work of Shrinivas Kudekar has been supported by a grant of the Swiss National Foundation 200021-105604.

References

[1] T. Richardson, R. Urbanke "Modern Coding Theory," *Cambridge University Press*, in preparation.

[2] T. Richardson, R. Urbanke "The Capacity of LDPC codes under Message-Passing Decoding," *IEEE Trans. Inf. Theory.*, pp. 638–656, 2001.

[3] F. Guerra, F. Toninelli "Quadratic Replica Coupling in the Sherrington-Kirkpatrick Mean Field Spin Glass Model", *J. Math. Phys.* **43** p. 3704 (2002).

[4] S. Franz, M. Leone "Replica Bounds for Optimization Problems and Diluted Spin Systems," *J. Stat. Phys.*, **111** p. 535-564 (2003).

[5] A. Montanari, "Tight Bounds for LDPC and LDGM Codes Under MAP Decoding," *IEEE Trans. Inf. Theory.*, **51**, no. 9, pp. 3221–3246, (2005).

[6] R. Urbanke, "EXIT functions: Fundamental Properties and Practical Implications", *Tutorial, ISIT, Adelaide*, Sept 2005, http://lthcwww.epfl.ch/isit05/index.php.

[7] A. Montanari, "The glassy phase of Gallager codes", *European Phys. Journal*, **23** 2001.

[8] N. Macris, "On the Relation between MAP and BP GEXIT functions of LDPC codes," Proc of *IEEE Inf. Theory Workshop* (2006).

[9] H. Nishimori, "Statistical Physics of Spin Glasses and Information Processing: An Introduction", *Oxford Science Publications*, (2001).

[10] N. Macris, "Griffith-Kelly-Sherman Correlation Inequalities: A Useful Tool in the Theory of Error Correcting Codes," *IEEE Trans. Inf. Theory.*, submitted for publication.