Proof of replica formulas in the high noise regime for communication using LDGM codes

Shrinivas Kudekar, Nicolas Macris
Ecole Polytechnique Fédérale de Lausanne
School of Computer and Communication Science
EPFL, I&C, LTHC, Station 14
Lausanne CH-1015, Switzerland

***Abstract*— We consider communication over a binary input memoryless output symmetric channel with low density generator matrix codes and optimal maximum a posteriori decoding. It is known that the problem of computing the average conditional entropy, over such code ensembles in the asymptotic limit of large block length, is closely related to computing the free energy of a mean field spin glass in the thermodynamic limit. Tentative explicit formulas for these quantities have been derived thanks to the replica method (of spin glass theory) and are generally conjectured to be exact. In this contribution we show that the replica solution is indeed exact in the high noise regime, where it coincides with density evolution equations. Our method uses ideas coming from high temperature expansions in spin glass theory.**

I. MOTIVATION

We consider communication over a binary input memoryless output symmetric channel (BMS) with low density generator matrix (LDGM) codes and optimal maximum a posteriori (MAP) decoding. Let $U_1, U_2, \ldots, U_n$ denote the information bits from which we create m generator bits $X_1, X_2, \ldots, X_m$ using an LDGM code. The generator bits are transmitted through the channel with transition probability $p_{Y|X}$ and $Y_1, Y_2, \ldots, Y_m$ is received. We are interested in computing the average (over the code ensemble) entropy of the information word U^n given the received word Y^m.

The average over the code ensemble conditional entropy (per information bit) $\mathbb{E}_C[h_n] = n^{-1}\mathbb{E}_C[H(U^n|Y^m)]$ of the transmitted information word U^n conditional to the received message Y^m can be formally computed by the ill-defined replica method of statistical mechanics. From the rather explicit formula obtained in this way one may also compute its derivative with respect to the noise level, a quantity that is also called MAP GEXIT curve [1]. It is believed that replica symmetry breaking is absent for symmetric channels (for bit MAP decoding) and is conjectured that the replica symmetric equations are rigorously exact. It is well known that away from the intervals between BP and MAP thresholds, the replica formulas coincide with the ones given by density evolution. Thus the conjecture also tells that density evolution gives the exact conditional entropy and MAP GEXIT curve away from intervals separating BP and MAP thresholds

While the general proof of this conjecture is still an open problem, some progress has been made in the last years. Tight bounds have been derived using a variety of tools (physical degradation [1], interpolation method [2], [3], correlation inequalities [4], [5]); a full proof has been achieved for the BEC for a class of LDPC codes (combinatorial methods [6], interpolation method [7]).

In this paper we provide a full proof of the conjecture in high noise regimes for the case of general LDGM code ensembles with bounded degrees and BIAWGN, BEC, BSC channels (and convex combinations of them). We believe that our proof can be extended to a more general class of BMS channels although some of the estimates become more technical. Apart from this result the interest of this work also lies in the method which departs from all the ones previously used. Our main tool is an expansion, that has its roots in high temperature expansions of statistical mechanics, and allows to estimate correlations between bits assigned to nodes of the factor graph. This expansion basically converges in the high noise regime and allows to prove that distant bits have exponentially small correlation as a function of the graph distance. From such a basic result one can then prove that the replica (or density evolution) formulas for $\mathbb{E}_C[h_n]$ hold.

In the sequel we use the standard notations Λ, P for the usual degree distributions of the variable and check nodes of the LDGM code ensemble, and λ, ρ for these distributions from the edge perspective [1].

II. NEW RESULTS

A. Random spin system formulation of LDGM codes

The Tanner graph has variable nodes denoted $i, j, k = 1, ..., n$ that are connected to check nodes denoted $a, b, c = 1, ..., m$. We write $i \in a$ for the variable nodes that are connected to a check a. We will work in terms of the *half*-loglikelihood ratios $l_a = \frac{1}{2}\ln\frac{p_{Y|X}(y_a|1)}{p_{Y|X}(y_a|0)}$ attached to each check, and call their distribution $c(l)$. The distribution depends on the noise level ϵ. High noise means that $c(l)$ has most of its weight on small likelihood ratios. For the moment one may keep in mind that ϵ is such that one may find $H(\epsilon) << 1$ such that $\int_{|l|\geq H(\epsilon)} dl\, c(l) = \mathbb{P}[|l| \geq H(\epsilon)] << 1$. The posterior distribution used in MAP decoding is (for a uniform prior over the code words, a memoryless binary-input output-symmetric channel, and assuming the input is the all zero codeword)

$$p(u^n|y^m) = \frac{\prod_{a=1}^m p(y_a|\oplus_{i\in a} u_i)}{\sum_{u^n}\prod_{a=1}^m p(y_a|\oplus_{i\in a} u_i)}$$

This can be viewed as the random Gibbs measure of a random spin system (the measure is over u^n and the randomness is y^n and the graph). It is convenient to use the mapping of bits to spins $\sigma_i = (-1)^{u_i}$, and one can interpret the half-loglikelihood variables as random interaction coupling constants between the spins. We have

$$p(u^n|y^m) = \mu_{\mathcal{C}}(\sigma^n) \triangleq \frac{1}{Z}\prod_{a=1}^{m} e^{l_a\sigma_a}$$

where

$$\sigma_a = \prod_{i\in a}\sigma_i \qquad \text{and} \qquad Z = \sum_{\sigma^n}\prod_{a=1}^{m} e^{l_a\sigma_a}$$

is the normalization factor or partition function. 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 $A \subset \{1,...,n\}$, $\langle\sigma_A\rangle = \sum_{\sigma^n}\sigma_A\mu_{\mathcal{C}}(\sigma^n)$ where $\sigma_A = \prod_{i\in A}\sigma_i$. Expectations with respect to the code ensemble and the channel outputs will be denoted by $\mathbb{E}_{\mathcal{C},l^m}[-]$.

It follows from the definition of the Shannon conditional entropy that (see [2])

$$h_n = \frac{1}{n}H(U^n|Y^m) = \frac{1}{n}\mathbb{E}_{l^m}[\ln Z] - \frac{1}{n}\sum_{a=1}^{m}\mathbb{E}_{l^m}[l_a]$$

The quantity $n^{-1}\ln Z$ is known as the average *free energy* in statistical mechanics and we are interested in computing its average over the channel outputs and the code ensemble. Differentiating the relation between entropy and free energy we get

$$\frac{d\mathbb{E}_{\mathcal{C}}[h_n]}{d\epsilon} = \frac{\Lambda'(1)}{P'(1)}\int dl_a \frac{dc(l_a)}{d\epsilon}\mathbb{E}_{\mathcal{C},l^{m\setminus a}}\ln\left\{\frac{1+\langle\sigma_a\rangle_0\tanh l_a}{1+\tanh l_a}\right\} \tag{1}$$

where a is any (single) check node, and the subscript in the Gibbs average $\langle-\rangle_0$ means that we set $l_a = 0$. For the BI-AWGN channel the above expression simplifies considerably: using integration by parts and the Nishimori identities we get[1]

$$\frac{d\mathbb{E}_C[h_n]}{d\epsilon} = \frac{1}{\epsilon^3}\frac{\Lambda'(1)}{P'(1)}\big(1-\mathbb{E}_{C,l^m}[\langle\sigma_a\rangle]\big) \tag{2}$$

For the BEC channel, the fact that a generator bit is either received perfectly or erased simplifies the above expression to

$$\frac{d\mathbb{E}_C[h_n]}{d\epsilon} = \ln 2\frac{\Lambda'(1)}{P'(1)}(1-\mathbb{E}_{C,l^{m\setminus a}}[\langle\sigma_a\rangle_0]) \tag{3}$$

B. Correlation Decay

One aim of statistical mechanics is to determine the correlations between distant spins given that the interactions between spins are local. For a fairly general class of spin systems correlation decay is equivalent to the unicity of the measure in the large system size limit. When the later is non unique the system is described by a convex combination of extremal measures (mixed state), and the correlations for the mixed state do not decay. The convex combination corresponds to the coexistence of pure thermodynamic phases. From this point of view it is clear that in the high noise regime (where one is in the "undecodable phase") the correlations between distant bits should decay. However the standard Dobrushin theory [8], or the usual cluster expansions [9], are applicable only when the coupling constants (here the l_a) are *uniformly small*. Over general BMS channels like the BIAWGN or the BEC, where the loglikelihoods potentially take unbounded values this is not the case. However if the probability that the loglikelihood takes unbounded values is very small, so that the domains with large loglikelihood do not percolate, one expects that the correlation still decays on average (over the noise realizations and code ensemble). There are a number of methods to address the issue of uniqueness of Gibbs measure and correlation decay in random spin systems when there are unbounded interactions [10],[11],[12]. Here we use an expansion technique that goes back to Dreifus, Klein and Perez [12] to show the correlation decay for general BMS channels in the high noise regime. In general these expansions are non trivial because, although the domains with large loglikelihoods do not percolate their size is unbounded, so there are arbitrarily large regions of the graph where one does not expand around the "correct point".

Definition 1: A *walk* w between two variable nodes v_α, v_β, is a sequence $v_1, c_1, v_2, c_2, \ldots, c_l, v_{l+1}$ of variable nodes (denoted by $v_1, v_2, \ldots, v_{l+1}$) and checks (denoted by $c_1, c_2, \ldots, c_l$) such that $v_1 = v_\alpha, v_{l+1} = v_\beta$ and $\{v_i, v_{i+1}\} \in c_i$. We say that the walk is *self-avoiding* if $v_i \neq v_j, c_i \neq c_j$ for $i \neq j$. We also say that two variable nodes v_α, v_β are connected if and only if there exists a self-avoiding walk from v_α to v_β.

The length of the walk is the number of clauses in it. If $v_\alpha = v_\beta$ then a self-avoiding walk from v_α to v_β is the trivial walk v_α: we define its length as zero.

Let $W_{\alpha\beta}$ denote the set of all self-avoiding walks between variable nodes v_α, v_β, and $W_{AB} = \cup_{v_\alpha\in A, v_\beta\in B}W_{ab}$ where $A, B \subset \{1,...,n\}$.

Now fix some number $H > 0$. Denote by set $\mathcal{B}$ the set of all checks (generator bits) a, such that $|l_a| > H$. Thus $\mathcal{B} = \{a \mid |l_a| > H\}$.

Lemma 1: (Correlation bound) Consider any LDGM code and two fixed non intersecting sets A, $B \subset \{1,...,n\}$. We have

$$\big|\langle\sigma_A\sigma_B\rangle - \langle\sigma_A\rangle\langle\sigma_B\rangle\big| \leq 2\sum_{w\in W_{AB}}\prod_{a\in w}\rho_a$$

where $\rho_a = 1$, if $a \in \mathcal{B}$ and $\rho_a = e^{4|l_a|} - 1$, if $a \notin \mathcal{B}$. Notice that ρ_a are independent random variables.

The right hand side of this bound involves a sum over all self-avoiding walks connecting the two sets A and B where each walk carries a weight depending on the loglikelihood values it meets (see figure 1). The proof proceeds by an expansion of the Gibbs weight around the point $l_a = 0$ for all a. This high noise expansion can be organized as a sum over walks connecting nodes in A with nodes in B. It turns out that only selfavoiding walks survive, the other ones giving a zero

[1]Detailed derivations of these formulas can be found in [5] for LDPC codes. Note that for the Gaussian case it is $\langle\sigma_a\rangle$ that enters (2) and not $\langle\sigma_a\rangle$.

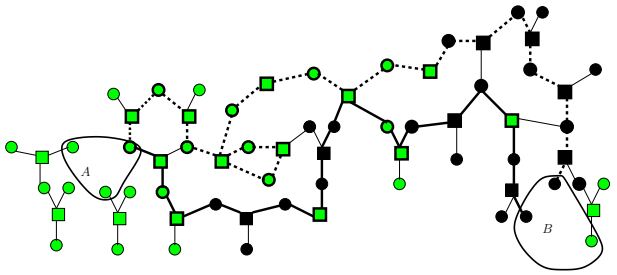

Fig. 1. Each set A and B contains three variable nodes. The light squares denote the generator bits in the complement of $\mathcal{B}$ and the dark squares denote the generator bits in $\mathcal{B}$. The thick path is an example of a self-avoiding path between A and B which contributes to the upper bound. The dashed path is a non-self-avoiding path and does not contribute to the bound.

contribution. A walk can traverse the bad set $\mathcal{B}$ in which case the expansion terms are not small, hence the weight $\rho_a = 1$; and it can traverse the complement of $\mathcal{B}$ in which case the expansion terms are small, hence the weight $\rho_a = e^{4l_a} - 1$. Details will be given elsewhere.

One can use this general bound to prove a correlation decay statement in the high noise regime, valid for any symmetric channel. For A and B two subsets of $\{1, ..., n\}$ let the graph distance between them be $d_{AB} = \min_{i\in A, j\in B} d(i, j)$ where $d(i, j)$ is the minimum length among all walks connecting i and j.

Corollary 1: (Correlation decay) Consider any LDGM code with l_{max}, r_{max} denoting the largest left and right degrees respectively and let $K \triangleq l_{max} r_{max}$. Let the noise level ϵ be such that there exists a function $H(\epsilon)$ satisfying

$$\delta_H(\epsilon) \triangleq e^{4H(\epsilon)} - 1 + \mathbb{P}[|l| > H(\epsilon)] < 1/K$$

Then

$$\mathbb{E}_{l^m}\left[\left|\langle\sigma_A\sigma_B\rangle - \langle\sigma_A\rangle\langle\sigma_B\rangle\right|\right] \leq \frac{2|A||B|}{1 - K\delta_H(\epsilon)}(K\delta_H(\epsilon))^{d_{AB}}$$

where $|A|, |B|$ denotes the cardinality of the sets A, B.

Remark 1: The hypothesis of this corollary is satisfied

- for BIAWGNC with $\epsilon^{-2} + \sqrt{2\epsilon^{-2}\ln 2K} < \frac{1}{4}\ln\left(1 + \frac{1}{2K}\right)$,
- for BEC with erasure probability $\epsilon > 1 - \frac{1}{2K}$,
- for BSC with $|\epsilon - \frac{1}{2}| < \frac{1}{2(2K+1)}$.

C. Exactness of replica solution

Our main theorem is valid for the BEC, BSC, BIAWGN channels in the noise ranges

- BIAWGNC with $\epsilon^{-2} + \sqrt{2\epsilon^{-2}\ln 2K^2} < \frac{1}{4}\ln\left(1 + \frac{1}{2K^2}\right)$
- BEC with $\epsilon > 1 - \frac{1}{2K^2}$
- BSC with $K^2|2\epsilon - 1|/\epsilon^2 < 1$

Note that here we do not attempt to obtain optimal values. The theorem is also valid for general channels with bounded loglikelihood ratios (e.g convex combinations of BSC).

Theorem 1 (Main theorem): Consider transmission using a LDGM(Λ, P) ensemble over BIAWGN(ϵ), BEC(ϵ) and BSC(ϵ) in the above range of noise. The MAP GEXIT function is given by

$$\lim_{n\to\infty}\frac{d\mathbb{E}_{\mathcal{C}}[h_n]}{d\epsilon} = \lim_{d\to\infty}\frac{\Lambda'(1)}{P'(1)}\int dh\frac{dc(h)}{d\epsilon}\mathbb{E}_d\ln\left\{\frac{1+\tanh\Delta}{1+\tanh h}\right\}$$

where both limits exist and where $\mathbb{E}_d$ is the average w.r.t the distribution for Δ given by

$$\Delta = \tanh^{(-1)}\left(\tanh l\prod_{i=1}^{k}\tanh v_i\right)$$

where the v_i are i.i.d random variables with distribution obtained from the iterative system of equations

$$\eta^{(d)}(v) = \sum_l \lambda_l \int\prod_{c=1}^{l-1} du_c\widehat{\eta}^{(d)}(u_c)\delta(v - \sum_{c=1}^{l-1}u_c)$$

$$\widehat{\eta}^{(d)}(u) = \sum_k \rho_k\int dhc(h)\prod_{i=1}^{k-1}dv_i\eta^{(d-1)}(v_i)$$
$$\times\,\delta(u - \tanh^{-1}(\tanh h\prod_{i=1}^{k-1}\tanh v_i)$$

with the initial condition $\eta^{(0)}(v) = \delta(0)$.

Remark 2: These equations are the iterative version of the replica fixed point equations, or equivalently the density evolution equations.

Remark 3: In the case of the BEC the formulas simplify drastically,

$$\lim_{n\to\infty}\frac{d\mathbb{E}_{\mathcal{C}}[h_n]}{d\epsilon} = \ln 2\frac{\Lambda'(1)}{P'(1)}\lim_{d\to\infty}(1 - P(1 - \lambda(y^{(d)})))$$

with

$$x^{(d)} = \lambda(y^{(d)});\ y^{(d)} = 1 - (1-\epsilon)\rho(1 - x^{(d-1)});\ x^{(0)} = 1$$

Also, in the case of BIAWGNC

$$\lim_{n\to\infty}\frac{d\mathbb{E}_{\mathcal{C}}[h_n]}{d\epsilon} = \lim_{d\to\infty}\frac{1}{\epsilon^3}\frac{\Lambda'(1)}{P'(1)}\int dlc(l)$$
$$\times\left\{1 - \mathbb{E}_d\left[\frac{\tanh\Delta + \tanh^2 l}{\tanh l + \tanh\Delta\tanh l}\right]\right\}$$

III. Sketch of proof for the main theorem

A. General strategy

Expanding the logarithm and using Nishimori identities we obtain the convergent expansion

$$\frac{\Lambda'(1)}{P'(1)}\sum_{p=1}^{+\infty}\frac{1}{2p(2p-1)}\left(\mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}] - 1\right) \qquad (4)$$
$$\times\int dl_a\frac{dc(l_a)}{d\epsilon}\tanh^{2p} l_a$$

The proof will be complete if we show that

$$\lim_{n\to+\infty}\mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}] = \lim_{d\to+\infty}\mathbb{E}_d[(\tanh\Delta)^{2p}] \qquad (5)$$

Indeed one can then resum the resulting series in (4) to obtain the replica formula.

Consider a neighborhood $N_d(a)$ of radius d around check a. Since the Tanner graph has bounded degrees, this is a tree with probability $1 - O(\frac{\gamma^d}{n})$ for some constant γ. Calling this event T_d, it is easy to see that,

$$\mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}] = \mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}|T_d] + O(\frac{\gamma^d}{n})$$

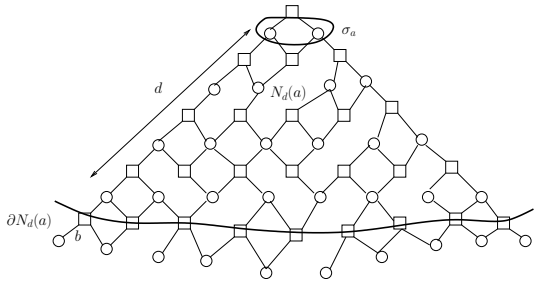

Fig. 2. Corollary 1 ensures that in the high noise regime the generator bit σ_a has weak correlations to the bits at the boundary of $N_d(a)$ since the shortest path connecting them has length d.

Therefore,

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}] = \lim_{d\to\infty}\lim_{n\to\infty} \mathbb{E}_{\mathcal{C},l^{m\setminus a}}[\langle\sigma_a\rangle_0^{2p}|T_d] \quad (6)$$

Notice that all paths connecting the bits in a with those outside $N_d(a)$ have a length at least equal to d, so because of corollary 1 in the high noise regime σ_a is very weakly correlated to the complement of $N_d(a)$. Informally speaking one should have that

$$\langle\sigma_a\rangle_0 = \langle\sigma_a\rangle_{0,N_d(a)} + \text{correlations of order } O(e^{-\frac{d}{\xi}}) \quad (7)$$

where ξ is a correlation length of order $O(1)$ and $\langle-\rangle_{0,N_d(a)}$ is the Gibbs measure restricted to $N_d(a)$. This is illustrated on figure 2. When $N_d(a)$ is a tree, the first term on the r.h.s of (7) is explicitly computable,

$$\langle\sigma_a\rangle_{0,N_d(a)} = \prod_{i\in a}\langle\sigma_i\rangle_{0,N_d(a)} = \prod_{i\in a}\tanh v_i^{(d)} \quad (8)$$

where $v_i^{(d)}$ are i.i.d with distribution given by the density evolution. Thus (5) follows from (6) and (8). In the next paragraphs we make (7) precise for the three kind of channels considered here.

B. BSC and channels with bounded likelihood ratios

Let $\partial N_d(a)$ be the set of checks that are at distance precisely d from a. We order the checks $b \in \partial N_d(a)$ in a given (arbitrary) way. For the first one we use $e^{l_b\sigma_b} = \cosh l_b + \sigma_b \sinh l_b$ to find

$$\langle\sigma_a\rangle_0 = \langle\sigma_a\rangle_{l_b=0} + \frac{\tanh l_b\big(\langle\sigma_a\sigma_b\rangle_{l_b=0} - \langle\sigma_a\rangle_{l_b=0}\langle\sigma_b\rangle_{l_b=0}\big)}{1+\langle\sigma_b\rangle_{l_b=0}\tanh l_b}$$

To lighten the notation $\langle-\rangle_{l_b=0}$ abusively denotes the Gibbs average where $l_a = l_b = 0$. Raising this equation to the power $2p$ we get

$$\langle\sigma_a\rangle_0^{2p} = \langle\sigma_a\rangle_{l_b=0}^{2p} + \Bigg[\sum_{q=1}^{2p}\binom{2p}{q}\langle\sigma_a\rangle_{l_b=0}^{2p-q} \left(\frac{\tanh l_b}{1+\langle\sigma_b\rangle_{l_b=0}\tanh l_b}\right)^q \big(\langle\sigma_a\sigma_b\rangle_{l_b=0} - \langle\sigma_a\rangle_{l_b=0}\langle\sigma_b\rangle_{l_b=0}\big)^q\Bigg]$$

We apply this formula iteratively to the first term of the right hand side above, for all further checks $b \in \partial N_d(a)$ in the specified order. Once this has been done for all checks of the boundary, the complement of $N_d(a)$ does not contribute anymore to Gibbs average as can be seen from its definition. This finally yields (after averaging over the code ensemble)

$$\mathbb{E}_{\mathcal{C},l^m}[\langle\sigma_a\rangle_0^{2p}|T_d] = \mathbb{E}_{\mathcal{C},l^m}[\langle\sigma_a\rangle_{0,N_d(a)}^{2p}|T_d] + S_{\text{CORR}} \quad (9)$$

where

$$S_{\text{CORR}} = \mathbb{E}_{\mathcal{C}}\Bigg[\sum_{b\in\partial N_d(a)}\mathbb{E}_{l^m}\sum_{q=1}^{2p}\binom{2p}{q}(\langle\sigma_a\rangle_b^*)^{2p-q} \left(\frac{\tanh l_b}{1+\langle\sigma_b\rangle_b^*\tanh l_b}\right)^q\big(\langle\sigma_a\sigma_b\rangle_b^* - \langle\sigma_a\rangle_b^*\langle\sigma_{b_i}\rangle_b^*\big)^q\Bigg|T_d\Bigg]$$

where $\langle-\rangle_b^*$ denotes a Gibbs average such that: for check b and all those occurring before it we have set $l = 0$ (and also $l_a = 0$).

For the BSC we have $|\tanh l_b| = 1-2\epsilon$ so that using corollary 1 we easily obtain that S_{CORR} is upper bounded by (uniformly in n)

$$\left(\frac{2r_{max}^2K^d(K\delta_H(\epsilon))^d}{1-K\delta_{H(\epsilon)}}\right)\sum_{q=1}^{2p}\binom{2p}{q}\frac{(1-2\epsilon)^q}{(2\epsilon)^q}2^{q-1} \quad (10)$$

Choosing $H(\epsilon) = \frac{1}{2}\ln\frac{1-\epsilon}{\epsilon}$ one deduces that $\lim_{d\to+\infty}\lim_{n\to+\infty} S_{\text{CORR}} = 0$ and hence the theorem as long as $K^2\frac{|1-2\epsilon|}{\epsilon^2} < 1$.

It is clear that the same kind of arguments go through for channels that have bounded loglikelihoods, hence the result of the theorem for such channels. However when the loglikelihood is unbounded it is not clear how to control the terms

$$\frac{\tanh l_b}{1+\langle\sigma_b\rangle_b^*\tanh l_b}$$

in S_{CORR}, whose denominator can (a priori) vanish with non zero-probability (this corresponds to the event $\langle\sigma_b\rangle_b^* = 1$). For the BEC (the loglikelihood takes values 0 and ∞) we are assured that $\langle\sigma_b\rangle_b^* \geq 0$ (in fact 0 or 1) so that this is not a real problem. However the whole proof can be made altogether more simply. For the gaussian channel however it is not clear how to go about with the present expansion. In the next paragraph we show how to treat the gaussian case thanks to a different starting point.

C. BIAWGN channel

Starting with expression (2), $\lim_{n\to+\infty}\frac{d\mathbb{E}_{\mathcal{C}}h_n}{d\epsilon}$ is given by,

$$\lim_{n\to\infty}\frac{1}{\epsilon^3}\frac{\Lambda'(1)}{P'(1)}\big(1-\mathbb{E}_{C,l^m}[\langle\sigma_a\rangle|\,T_d]\big)$$

We again order the generator bits in $\partial N_d(a)$. We set the noise level of the first generator bit to $\epsilon \leq \nu \leq \infty$. From the fundamental theorem of calculus one can write

$$\mathbb{E}_{l^m}[\langle\sigma_a\rangle] = \mathbb{E}_{l^m}[\langle\sigma_a\rangle_{l_b=0}] + \int_\infty^\epsilon d\nu\frac{d}{d\nu}\mathbb{E}_{l^m}[\langle\sigma_a\rangle_{l_b\sim\nu}]$$

where in the integral the expectation over l_b is w.r.t noise level ν. Using gaussian integration by parts and channel symmetry

we get

$$\mathbb{E}_{l^m}[\langle\sigma_a\rangle] = \mathbb{E}_{l^m}[\langle\sigma_a\rangle_{l_b=0}] + 2\int_\epsilon^\infty \frac{d\nu}{\nu^3} \times \mathbb{E}_{l^m}[\big(\langle\sigma_a\sigma_b\rangle_{l_b\sim\nu} - \langle\sigma_a\rangle_{l_b\sim\nu}\langle\sigma_b\rangle_{l_b\sim\nu}\big)^2]$$

Iterating this procedure for all the generator bits in $\partial N_d(a)$ in the specified order we get

$$\mathbb{E}_{\mathcal{C},l^m}[\langle\sigma_a\rangle|\ T_d] = \mathbb{E}_{\mathcal{C},l^m}[\langle\sigma_a\rangle_{N_d(a)}|T_d] + S_{\text{CORR}} \tag{11}$$

where

$$S_{\text{CORR}} = 2\mathbb{E}_{\mathcal{C}}\Bigg[\sum_{b\in\partial N_d(a)}\int_\epsilon^\infty d\nu\nu^{-3} \times \mathbb{E}_{l^m}\Bigg[\big(\langle\sigma_a\sigma_b\rangle^*_{l_b\sim\nu} - \langle\sigma_a\rangle^*_{l_b\sim\nu}\langle\sigma_b\rangle^*_{l_b\sim\nu}\big)^2\Bigg|T_d\Bigg] \tag{12}$$

From the corollary 1 we get

$$|S_{\text{CORR}}| \le \frac{2r_{max}^2 K^d (K\delta_H(\epsilon))^d}{\epsilon^2(1-K\delta_H(\epsilon))}$$

which is uniform in n. Choosing $H(\epsilon)$ such that $\epsilon^{-2} + \sqrt{2\epsilon^{-2}\ln 2K^2} < H(\epsilon) < \frac{1}{4}\ln(1+\frac{1}{2K^2})$ one deduces that $\lim_{d\to\infty}\lim_{n\to\infty} S_{\text{CORR}} = 0$.

It is easy to check that

$$\langle\sigma_a\rangle_{N_d(a)} = \frac{\langle\sigma_a\rangle_{0,N_d(a)} + \tanh l_a}{1+\langle\sigma_a\rangle_{0,N_d(a)}\tanh l_a}$$

where $\langle-\rangle_{0,N_d(a)}$ denotes the gibbs average with $l_a = 0$. Thus because of (8) we obtain the theorem.

D. BEC channel

Starting with the expression (3) we get a formula similar to (11) where

$$S_{\text{CORR}} \triangleq \ln 2\frac{\Lambda'(1)}{P'(1)}\mathbb{E}_C \sum_{b\in N_d(a)} \mathbb{E}_{l^m} \Bigg[\big(\langle\sigma_a\sigma_b\rangle^*_b - \langle\sigma_a\rangle^*_b\langle\sigma_b\rangle^*_b\big)\frac{\tanh l_b}{1+\langle\sigma_b\rangle^*_b\tanh l_b}\Bigg]$$

From corrolary 1 we get

$$|S_{\text{CORR}}| \le \frac{2r_{max}^2 K^d (K\delta_H(\epsilon))^d}{(1-K\delta_H(\epsilon))}$$

Choosing $H(\epsilon) < \frac{1}{4}\ln\big(1+\frac{1}{2K^2}\big)$ one deduces again $\lim_{d\to\infty}\lim_{n\to+\infty} S_{\text{CORR}} = 0$ as long as $\epsilon > 1 - \frac{1}{2K^2}$ and hence we get the theorem.

IV. Discussion

High noise regime. An obvious question that would require more investigations is the extension of theorem 1 to a wider class of channels specially in the case of unbounded likelihood ratios. Moreover our proof works in a regime where the level sets $|l| > H(\epsilon)$ do not percolate. It would be interesting to determine if there is a more deep connection with such a percolation problem and if this has an algorithmic significance.

Low noise regime. At low enough noise we conjecture that an exponential decay of correlation also holds. Indeed (9), (11) are valid for any noise regime so that if the MAP GEXIT function is given by the density evolution or replica formulas it should be the case that the sum over correlations between node a and $b \in \partial N_d(a)$ vanishes as $d \to +\infty$. For the Gaussian case the sum contains an exponential number of positive terms so each term should go to zero exponentially fast. That the MAP GEXIT function is given by the density evolution formula is supported by the following argument. In the gaussian case (11), (12) implies that $\lim_{n\to\infty}\frac{d\mathbb{E}_{\mathcal{C}}[h_n]}{d\epsilon}$ is upper bounded by the density evolution formula of section II-C. This inequality can also be deduced for other symmetric channels. A formal integration in a low noise interval then yields a bound of the form $\mathbb{E}_{\mathcal{C}}[h_n] \le \lim_{d\to\infty}\mathbb{E}_d[h_{RS}[v]]$. This is only a formal argument because the integration involves existence and smoothness issues of solutions of density evolution, a hard question in non linear analysis [13]. On the other hand we know that $\mathbb{E}_{\mathcal{C}}[h_n] \ge \max_{\text{distr of v}} h_{RS}[v]$ [2]. Therefore one should have that at low noise (namely for noise below the belief propagation threshold) $\mathbb{E}_{\mathcal{C}}[h_n] = \lim_{d\to\infty}\mathbb{E}_d[h_{RS}[v]]$ and therefore the equality in theorem 1 should hold.

A direct proof of correlation decay for low noise using an appropriate "cluster expansion" would be desirable and would lead to a direct proof of theorem 1 along the same lines than in section III.

Acknowledgment

We acknowledge discussions with Ruediger Urbanke on the decay of correlations. Shrinivas Kudekar acknowledges support from the Fonds National pour la Recherche Scientifique, grant no 200020-113412.

References

[1] T. Richardson, R. Urbanke "Modern Coding Theory," *Cambridge University Press*, in press.
[2] A. Montanari, "Tight Bounds for LDPC and LDGM Codes Under MAP Decoding," *IEEE Trans. Inf. Theory.*, **51**, no. 9, pp. 3221–3246, (2005).
[3] S. Kudekar, N. Macris, "Sharp Bounds for MAP Decoding of General Irregular LDPC Codes", *Proc ISIT, Seattle*, pp. 2259-2263 (2006)
[4] N. Macris, "Griffith-Kelly-Sherman Correlation Inequalities: A Useful Tool in the Theory of Error Correcting Codes," *IEEE Trans. Inf. Theory.*, **53**, No. 2, pp. 664-683 (2007).
[5] N. Macris, "Sharp Bounds on Generalized EXIT functions", *IEEE Trans. Inf. Theory.*, **53**, No. 7, pp. 2365-2375 (2007).
[6] C. Measson, A. Montanari, R. Urbanke, "Asymptotic rate versus Design Rate", *Proc ISIT, Nice* (2007); see also C. Measson, "Conservation Laws for Coding", *These EPFL no* 3485 (2006)
[7] S. Kudekar, S. Korada, N. Macris, "Exact Solution for the Conditional Entropy of Poissonian LDPC Codes over the Binary Erasure Channel", *Proc ISIT, Nice* (2007)
[8] H. O. Georgii, "Gibbs measures and phase transitions," *de Gruyter Studies in Mathematics*, 9. Walter de Gruyter & Co., Berlin, (1988).
[9] D. Ruelle, "Statistical Mechanics", *Benjamin, New York* (1969)
[10] L. Bassalygo, R. Dobrushin, "Uniqueness of a Gibbs field with random potential - an elementary approach," *Theory Probab. Appl.*, **31**, pp. 572-589 (1986).
[11] J. Froehlich, J. Imbrie, "Improved perturbation expansion for disordered systems: beating Griffiths singularities," *Commun. Math. Phys.*, **96**, pp. 145-180 (1984).
[12] H. Dreifus, A. Klein, J. Perez, "Taming Griffiths' singularities: Infinite differentiability of quenched correlation functions," *Communications in Mathematical Physics*, **170**, pp.21-39.
[13] V. Rathi, R. Urbanke, "Existence Proofs of Some EXIT Like Functions", *Proc ISIT, Nice* (2007)