

Asymptotic MAP upper bounds for LDPC codes

David Matas, Meritxell Lamarca

Universitat Politècnica de Catalunya Barcelona, Spain, Email: (david.matas, meritxell.lamarca)@upc.edu

Abstract—This paper aims at computing tight upper bounds for the maximum a posteriori threshold of low-density parity check codes in the asymptotic blocklength regime for the transmission over binary-input memoryless symmetric-output channels. While these bounds are already known, we propose a novel derivation based on a completely different approach: based solely on the concept of the chain rule and the conditional entropy, resorting to the concentration theorem for the code ensemble to compute the syndrome entropy with low complexity employing density evolution.

I. INTRODUCTION

Channel polarization and coupling schemes have shown how the performance of the maximum a posteriori (MAP) decoder can be asymptotically achieved by suboptimum decoders. This reactivated the interest on the analysis of bounds for the MAP decoder and the study of its link with the suboptimum ones.

On the basis of the Fano inequality, the MAP error performance is bounded in terms of the equivocation introduced by the channel. Namely, for the error to be able to approach zero, the equivocation should also approach zero. For this reason, research has focused on obtaining bounds on this metric. Gallager [1] was the first to introduce a lower bound on the equivocation for LDPC codes. Much more recently, this bound was improved and generalized for any binary-input memoryless output symmetric channel (BMS) [2], [3]. We follow the same approach and introduce a procedure to bound the entropy of the syndrome associated to the received sequence based on two simple concepts: the chain rule and the confinement of dependencies within a tree. The latter is essentially the principal assumption of a belief propagation (BP) decoder, while the former seems to have been overlooked in the asymptotic analysis of LDPC code ensembles.

The proposed method results in the computation of the area under an entropy function and allows us to obtain asymptotic upper bounds for the MAP threshold for LDPC codes over any BMS channel, such as the binary erasure channel (BEC), the binary symmetric channel (BSC) and the binary-input Gaussian channel (BAWGN).

A related approach was developed in [4], but was mostly justified as a mathematical tool derived from physics. More insightful results were presented in [5], [6]. As we will see later, the generalized extrinsic entropy (EXIT) area theorem presented in the cited references is strongly related to the computation proposed in this paper. They lead to identical

results and MAP bounds and have equivalent computational complexity. However, we believe that the proposed method is more straightforward conceptually.

Let us introduce some definitions to simplify the notation. For any integer i , let $[i]$ be the set of all integers from 0 to $i - 1$, i.e. $[i] = \{0, \dots, i - 1\}$. Let $V = \{v_0, \dots, v_{|V|-1}\}$ be a subset of distinct integers from $[i]$. Without loss of generality, we will assume $v_0 < v_1 < \dots < v_{|V|-1}$. Let \mathbf{s} be a vector of length at least i . We define the mapping operation $\mathbf{s}(V)$ as: $\mathbf{s}(V) = (s_{v_0}, \dots, s_{v_{|V|-1}})^T$, which corresponds to the vector with the entries of \mathbf{s} indexed by the elements of V .

II. MAP BOUND AND SYNDROME ENTROPY

Let us consider a LDPC code described by a parity check matrix \mathbf{H} of dimensions $N - K \times N$. We denote $M = N - K$. The rate of the code is $R = K/N$. Let \mathbf{x} and \mathbf{y} be the transmitted and the received sequences respectively. Let us assume a binary-input memoryless symmetric-output (BMS) channel parametrized by a scalar parameter θ such that the capacity of the channel C is monotonically decreasing with θ .

The MAP threshold is defined as the maximum channel parameter value for which the MAP decoder error probability converges asymptotically to 0. Thanks to the Fano inequality it can be written in terms of the system equivocation as:

$$\theta_{MAP} = \min_{\theta} \{ \theta \mid \lim_{N \rightarrow \infty} \frac{1}{N} H(\mathbf{x}|\mathbf{y}) > 0 \} \quad (1)$$

Thus, upper bounds on the MAP threshold can be obtained by lower bounding the equivocation. In order to do this, we follow the procedure presented in [3] and express it in terms of the entropy of a generalized syndrome \mathbf{s} :

$$H(\mathbf{x}|\mathbf{y}) \geq N(1 - C) - H(\mathbf{s}|\omega) \quad (2)$$

The generalized syndrome is defined as the syndrome for a BSC channel that introduces an error sequence equivalent to the hard decision of the received symbols in the original channel when the all zeros codeword was transmitted. The vector ω in the equation corresponds to the magnitudes of the log-likelihood ratios (LLR) for those symbols. The bound in (2) applies to the BEC, BSC and BAWGN channels among others.

III. BOUNDING THE SYNDROME ENTROPY

Establishing an order among the syndrome bits, their joint entropy can be expressed following the chain rule as:

$$H(\mathbf{s}|\omega) = \sum_{i=0}^{M-1} H(s_i | \mathbf{s}([i]), \omega) \quad (3)$$

In general, the computation of the statistics of s_i has an exponentially growing complexity since we are conditioning it to the value of all preceding syndrome bits. However, we will crop the dependency to a limited subset of precedent bits such that we can assume that the resulting subgraph is a tree. Then, the conditional probabilities of the syndrome bits can be computed by means of BP and their corresponding entropy by means of density evolution (DE).

Let us assume a regular $(1, r)$ LDPC code represented by a Tanner graph \mathcal{G} . According to the syndrome generation, the variable nodes are associated to the errors introduced by the equivalent BSC channel and the check nodes to the syndrome bits. This corresponds to having an additional variable node of degree one for each check node. To avoid more variable definitions, we will refer to the check nodes as the syndrome bits themselves $\{s_i\}$.

Let us define the distance $d(s_i, s_j)$ between two check nodes s_i and s_j as the minimum number of bit nodes that we cross in any path between them. For example, two checks nodes connected to the same bit are at distance 1. Then, we define the dependency set of depth D of the i -th check node, V_i^D , as the indexes of the check nodes that are at distance lower or equal than D of the node s_i and are precedent to it¹, i.e.

$$V_i^D = \{n \in [i] \mid d(s_i, s_n) \leq D\} \quad (4)$$

For $N \rightarrow \infty$ and fixed D , the subgraph spanned by the checks s_i and $\mathbf{s}(V_i^D)$ is a tree with probability going to one. We will call it the D -depth subtree of the i -th check node. It is rooted on the check node s_i and contains the check nodes $\mathbf{s}(V_i^D)$ and all bits in the code graph \mathcal{G} connected to them. Note that the topology of the subtree may not correspond to the one of the ensemble computation tree since the degree of the bit nodes depends on which adjacent check nodes are precedent to the i -th one in the chain. Then, the same node in two different orderings may have different subtrees. For the i -th syndrome bit, we compute $H(s_i|\mathbf{s}(V_i^D), \omega)$ instead of $H(s_i|\mathbf{s}([i]), \omega)$. Adding all contributions, we obtain an upper bound on the syndrome entropy:

$$\sum_{i=0}^{M-1} H(s_i|\mathbf{s}(V_i^D), \omega) \geq H(\mathbf{s}|\omega) \quad (5)$$

where the inequality holds because conditioning can only reduce the entropy.

IV. AVERAGE ENSEMBLE BOUNDS

According to the concentration theorem, the performance of BP for a LDPC code converges with probability going to one to the performance described by density evolution over the code ensemble average when the blocklength goes to infinity. Since the probabilities $p(s_i|\mathbf{s}(V_i^D), \omega)$ correspond to the LLRs of the syndrome bits computed by BP over the D -depth subtrees defined before, the same kind of convergence is expected for the syndrome entropy bound in equation (5).

¹Notice that, as for the definition of distance, we are having into account only check nodes when we talk about depth.

The proof of the concentration of the entropy here would be equivalent to the one for the extrinsic entropy of an LDPC code ensemble [9].

Let us assume a fixed order among check nodes that does not change from one code realization to another. Let us consider the i -th check node of this sequence. For a given code realization \mathcal{C} , it roots a D -depth subtree with a certain topology and produces an entropy contribution $H_C(s_i|\mathbf{s}(V_i^D), \omega)$. Its average over the code ensemble, $\mathcal{E}_C\{H_C^D(s_i|\mathbf{s}(V_i^D), \omega)\}$, can be computed, as done for LDPC code ensembles for $N \rightarrow \infty$, by means of DE over the ensemble average subtree. Then, instead of analyzing the subtrees with different topologies that can happen for a specific code realization and averaging over them, we analyze the average topology, assuming a constant degree profile but introducing the absence of check nodes depending on the position of the root check node within the computation chain. The conditioning on ω is also absorbed by the average ensemble computation when we introduce the distribution corresponding to the channel for the received LLRs.

Let us assume a D -depth computation tree, whose topology is described by the degree profile of the LDPC ensemble. As the i -th check node is preceded by $i-1$ others, the probability of a check node to be actually present in the average ensemble subtree is $\nu_i \triangleq \frac{i-1}{M-1}$. Then, we can write the evolution of the p.d.f.'s of the ratios computed by BP in the bit nodes in an ensemble tree as done in the usual DE but assuming that a bit node of a given layer receives through every edge the message coming from a check node in the previous layer with probability $1 - \nu_i$ and nothing with probability ν_i (when no information about the bit is provided, representing the absence of the check node).

Let us denote this p.d.f. as $p_{i,n}(l)$, where the subindex $n = 0, \dots, D$ indicates the layer in the subtree, which is also the temporal index in the DE process. Then:

$$p_{i,n}(l) = p(l) \otimes^{1-n} (\nu_i \boxtimes^{r-1} p_{i,n-1}(l) + (1 - \nu_i)\delta(l)) \quad (6)$$

where $p(l)$ is the channel LLR p.d.f., $\delta(l)$ is the Dirac delta function and \otimes and \boxtimes the convolution operations for the bit and check node updates described in [9].

Instead of obtaining the entropy for this p.d.f. at $n = D$ as we would do to characterize the extrinsic entropy function (EXIT) of the code after D iterations, we do a final update for the root check node:

$$p_i^D(l) = \boxtimes^r p_{i,D}(l)$$

The D -depth ensemble subtree entropy for the i -th syndrome bit can be computed from it as:

$$H^D(\nu_i) \triangleq \int_{-\infty}^{\infty} p_i^D(l) h_2\left(\frac{e^l}{1+e^l}\right) dl \quad (7)$$

We have expressed the entropy parametrized by ν_i with a purpose. For the blocklength going to infinity, this fraction of preceding check nodes tends to a continuous variable in the interval $[0, 1)$, that we will call ν . The entropy of the average ensemble subtrees for the i -th check node becomes a function

ϵ	0.42	0.44	0.47	0.49
$H_{s \omega}^\infty$	0.9035	0.9527	0.9694	0.9769

Table I

AVERAGE SYNDROME BIT ENTROPY FOR THE CODE ENSEMBLE (3, 6) AND DIFFERENT BEC CHANNELS.

of this parameter, $H^D(\nu_i) \rightarrow H^D(\nu)$ and then the sum for all the chain converges to the area under this function.

From the concentration theorem, the syndrome entropy bound (5) for a given code realization converges to the ensemble average asymptotically and, thus, can be expressed -normalized per syndrome bit- as the area:

$$H_{s|\omega}^D \triangleq \lim_{N \rightarrow \infty} \frac{1}{M} \sum_{i=0}^{M-1} H(s_i | s(V_i^D), \omega) = \int_0^1 H^D(\nu) d\nu \quad (8)$$

This area computation can be employed to obtain a MAP upper bound according to equations (1) and (2).

V. MAP BOUNDS

In this section, we present some results for the BEC, BSC and BAWGN channels. The values presented here are in all cases identical to those obtained by means of the generalized area theorem [6], which are slightly tighter than those of [3] and [4].

A. BEC channel

Let us consider a BEC with erasure probability ϵ . The p.d.f.'s in DE in this case are always a single delta function, so it is usually expressed as a single variable update, the erasure probability, that propagates along the graph. For a fraction of precedent check nodes ν , this probability becomes:

$$z_n(\nu) = \epsilon(1 - (1 - z_{n-1}(\nu))^{r-1}\nu)^{l-1} \quad (9)$$

where z is the erasure probability at the bit nodes propagated along the tree, which we parametrize as a function of the layer index ($n \in (0, D)$) and ν . We omit the trivial dependence on ϵ for the sake of clarity. The root check node entropy (of the corresponding syndrome bit) is then obtained from $z_D(\nu)$ as

$$H^D(\nu) = 1 - (1 - z_D(\nu))^r$$

and replacing it in the area equation (8), the average entropy bound per syndrome bit can be expressed as

$$H_{s|\omega}^D = \int_0^1 1 - (1 - z_D(\nu))^r d\nu \quad (10)$$

which remains as function of ϵ . Since we are assuming the blocklength to go to infinity, we can bring also the depth to infinity to obtain the tightest of these bounds ($H_{s|\omega}^\infty$).

Figure 1 represents the average subtree entropy function $H^\infty(\nu)$ and the corresponding area for a (3, 6) code and several erasure probabilities. The entropy function is always decreasing with respect to ν . For erasure probabilities below the BP threshold (e.g. $\epsilon = 0.42$), there is a threshold also on ν for which the entropy falls to zero, which means that the last

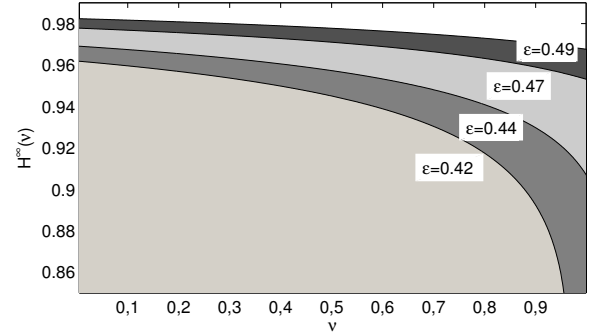


Figure 1. Average ∞ -depth subtree entropy for a check node as function of the fraction of precedent check nodes ν for a (3,6) LDPC code and several BEC channels. The shaded area corresponds the average syndrome entropy.

l	r	$D = 0$	$D = 1$	$D = 2$	$D = \infty$	BP	Sh
3	6	0.491	0.489	0.4885	0.4881	0.4294	0.5
3	5	0.5933	0.5917	0.5912	0.5910	0.5176	0.6
4	6	0.6658	0.6657	0.6656	0.6656	0.5061	0.667
4	8	0.498	0.4978	0.4977	0.4977	0.3834	0.5

Table II

MAP UPPER BOUNDS ON THE ERASURE PROBABILITY OF A BEC FOR THE TRANSMISSION EMPLOYING DIFFERENT REGULAR LDPC CODES OF DEGREES (l, r).

syndrome bits of the chain provide no information. Above the BP threshold, however, the entropy stays always above zero and approaches one when ϵ increases, so the area grows too. Table I lists the obtained areas which are shaded in the figure.

Table II lists MAP bounds for LDPC codes with different degrees and rates and for different subtree depths $\{0, 1, 2, \infty\}$. It can be seen again how it decreases with respect to the depth, being the tightest bound for $D = \infty$. The BP threshold and the Shannon limit are also shown for comparison purposes.

B. BSC channel

For a BSC with probability of error p , density evolution is initialized with the channel p.d.f.

$$p_{BSC}(l) = p \cdot \delta(l + L_p) + (1 - p) \cdot \delta(l - L_p)$$

whith $L_p = \log((1 - p)/p)$.

Figure 2 represents the curves $H^D(\nu)$ for the (3, 6) regular ensemble for several depths for a channel with $p = 0.1$. The entropy values (in decreasing order) for the syndrome bits obtained for a specific matrix realization are plotted for comparison [10]. The latter has a noticeable staircase shape for $D = 1$ since the number of different subtree topologies is very limited. For low ν values, some check nodes have higher entropy bound than the average ensemble one while the opposite happens for higher values. However, if we observe the average over an interval of syndrome bits within the chain, it coincides with the ensemble average for the corresponding fraction of precedent check nodes, validating the application of the concentration theorem. It can be observed also that the curve converges rapidly with D , which happens because BP converges fast to a fixed point for this channel error.

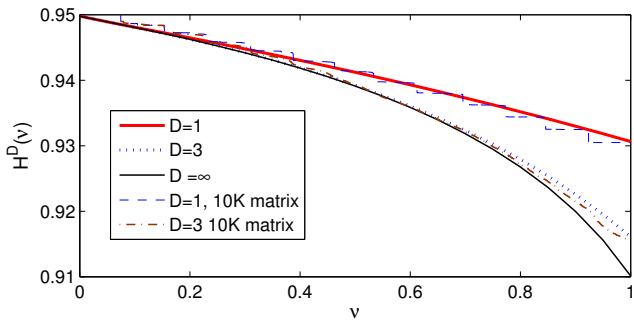


Figure 2. Average D -depth subtree entropy for a check node as function of the fraction of precedent check nodes ν for a (3,6) LDPC code and a BSC channel with $p = 0.1$ for $D = \{1, 3, \infty\}$. The D -depth subtree entropy of check nodes versus the normalized node index for a matrix realization is also depicted.

If we plotted the curves for infinite depth and different channel error probabilities, we would observe a behavior completely analogous than the one for the BEC, with the entropy decreasing with respect to ν and falling to zero if we are below the BP threshold.

Table III contains the MAP bounds for different LDPC codes in a BSC. As it can be seen, a depth of just $D = 2$ produces already a bound very close to the tightest one ($D = \infty$), identical for codes (4, 6) and (4, 8).

C. BAWGN channel

For a BAWGN channel with noise variance σ^2 and transmitted symbols ± 1 , the channel LLR p.d.f. is

$$p_{BAWGN}(l) = \sqrt{\frac{\sigma^2}{8\pi}} e^{-\frac{\sigma^2}{8}(l - \frac{\sigma^2}{2})^2}$$

Table IV shows the MAP bounds for the BAWGN channel for the same regular LDPC codes as before. We present it in terms of Eb/No , so they become lower bounds in this case. The dependency on ν of the average subtree entropy function $H^D(\nu)$ is analogous to the described one for the BEC and BSC.

It is well known that the BEC and BSC correspond to the two extremes of the capacity of all BMS channels and the MAP performance of linear codes transmitting through them [8]. We can observe that this holds for the BAWGN channel comparing its corresponding entropy function $H^\infty(\nu)$ with the one for the BEC and BSC. Figure 3 depicts $H^\infty(\nu)$ for the three channels, all them with a fixed capacity of 0.53

l	r	$D = 0$	$D = 1$	$D = 2$	$D = \infty$	BP	Sh
3	6	0.1024	0.1006	0.1001	0.0997	0.084	0.11
3	5	0.1397	0.1382	0.1376	0.1373	0.113	0.146
4	6	0.1726	0.1725	0.1724	0.1724	0.116	0.174
4	8	0.1076	0.1072	0.1071	0.1071	0.076	0.11

Table III

MAP UPPER BOUNDS ON THE ERROR PROBABILITY OF A BSC FOR THE TRANSMISSION EMPLOYING DIFFERENT REGULAR LDPC CODES OF DEGREES (l, r) .

l	r	$D = 0$	$D = 1$	$D = 2$	$D = \infty$	BP	Sh
3	6	0.372	0.426	0.446	0.458	1.1	0.184
3	5	-0.074	-0.026	-0.020	-0.013	0.89	-0.237
4	6	-0.466	-0.461	0.459	-0.458	1.67	-0.495
4	8	0.240	0.247	0.250	0.250	1.53	0.184

Table IV

MAP LOWER BOUNDS ON THE Eb/No (IN dB) OF A BAWGN FOR THE TRANSMISSION EMPLOYING DIFFERENT REGULAR LDPC CODES OF DEGREES (l, r) .

bits/channel use, and a (3, 6) LDPC code. For this same code, the MAP bounds from tables II, III and IV correspond to capacities of 0.512, 0.532 and 0.52 respectively. In both cases, the BAWGN performance lays between that one of the BEC and of the BSC.

VI. ALTERNATIVE AREA EXPRESSIONS

The bound expressed in terms of an area in equation (8) arises as the limit of the chain rule decomposition of the entropy of the syndrome bits. Since the result is the integration of an entropy function, it reminds of the area theorem for the EXIT function of a code. The area theorem was known since the introduction of EXIT charts [7] for the BEC and was generalized for other channels in [6]. Even if the form of the area equations looks similar and the numerical results match exactly, the connection between them is not obvious, as they follow a completely different approach. The integration for the proposed computation depends on the parameter ν , which indicates the position of the syndrome bits in the computation chain, instead of the channel parameter. Moreover, it employs the syndrome entropy (with the check nodes as root of the computation trees) and not the EXIT function or its derivative.

The approach proposed in this paper, based on the chain rule decomposition, provides a straightforward interpretation of the computation and a methodology that can also be applied to different entropy metrics other than the syndrome one. In this section, we apply it to obtain an alternative area expression for the BEC that coincides with the EXIT area theorem.

Starting from the mutual information of the system expressed by the chain rule as:

$$I(\mathbf{x}; \mathbf{y}) = \sum_{j=0}^{N-1} H(y_j | \mathbf{y}([j])) - \sum_{j=0}^{N-1} H(y_j | x_j)$$

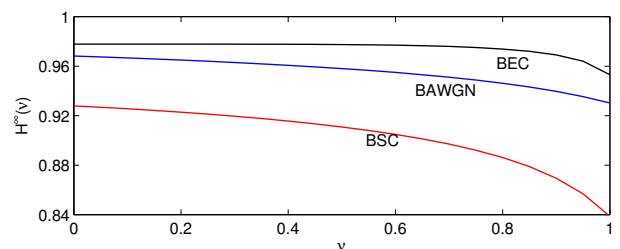


Figure 3. Average ∞ -depth subtree entropy for the check nodes as function of the fraction of precedent check nodes ν for a (3,6) LDPC code and BEC, BSC and BAWGN channel, all with capacity of 0.53 bits/channel use.

we arrive, after some manipulations, to ²:

$$I(\mathbf{x}; \mathbf{y}) = \sum_{j=0}^{N-1} H(x_j | \mathbf{y}([j])) - \sum_{j=0}^{N-1} H(x_j | \mathbf{y}([j+1])) \quad (11)$$

where it is expressed as a difference of conditional entropies: for each coded bit ($j \in [N]$), the entropy conditioned to the precedent received symbols ($\mathbf{y}([j])$) minus the entropy conditioned to the precedent and the corresponding j -th received symbol ($\mathbf{y}([j+1])$).

Let us particularize this equation for a BEC channel with erasure probability ϵ . In this case, the conditional entropy $H(x_j | \mathbf{y}([j+1]))$ can be expressed as a function of $H(x_j | \mathbf{y}([j]))$ as:

$$H(x_j | \mathbf{y}([j+1])) = \epsilon H(x_j | \mathbf{y}([j]))$$

and, replacing it in equation (11), we obtain:

$$I(\mathbf{x}; \mathbf{y}) = (1 - \epsilon) \sum_{j=0}^{N-1} H(x_j | \mathbf{y}([j])) \quad (12)$$

We follow the same asymptotic method proposed before for the syndrome entropy and analyze the entropy terms in the summation averaged over the code ensemble. The average ensemble subgraph rooted in the j -th bit has a tree structure up to a certain depth described by the code degree profiles but with the bits being present with probability $\frac{j}{N}$, the average fraction of precedent bits. The analysis of this subtree is equivalent to the one of the ensemble computation tree but having an additional BEC with erasure probability equal to $\mu = 1 - \frac{j}{N}$ concatenated in series with the original channel. In the asymptotic regime for $N \rightarrow \infty$, this parameter becomes continuous $\mu \in [0, 1]$, while the depth of the subtree can be assumed to go also to infinity.

Let us assume that we observe the output of the fictitious concatenated channel, $\hat{\mathbf{y}}$. Let us call the BP extrinsic entropy of the coded bits for this channel as $H_{\mathbf{x}|\hat{\mathbf{y}}\setminus}(\epsilon, \mu)$, where the symbol \setminus indicates that the symbol corresponding to the analyzed bit is removed from the received sequence and we indicate the dependence with respect to the original and the additional BEC erasure probabilities. Applying the concentration theorem over the subtrees seen by every received symbol in the ordered chain, we can introduce the entropy function of the fictitious channel in (12) and obtain the bound:

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}) \leq (1 - \epsilon) \int_0^1 H_{\mathbf{x}|\hat{\mathbf{y}}\setminus}(\epsilon, \mu) d\mu \quad (13)$$

This is defined over the equivalent concatenated channel, so the entropy function depends on both μ and ϵ . However, the series concatenation of the two BECs is equivalent to another BEC with overall erasure probability equal to $\epsilon' = 1 - (1 - \epsilon)(1 - \mu)$. Let us call $H_{\mathbf{x}|\mathbf{y}\setminus}(\epsilon')$ to the BP EXIT function

corresponding to the transmission through a BEC with erasure probability ϵ' . Then

$$H_{\mathbf{x}|\hat{\mathbf{y}}\setminus}(\epsilon, \mu) = H_{\mathbf{x}|\mathbf{y}\setminus}(1 - (1 - \epsilon)(1 - \mu))$$

and, introducing ϵ' in the integration (13) by a change of variables, we obtain:

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{x}; \mathbf{y}) \leq \int_{\epsilon}^1 H_{\mathbf{x}|\mathbf{y}\setminus}(\epsilon') d\epsilon' \quad (14)$$

which is equivalent to the area theorem for the EXIT function [6].

VII. CONCLUSIONS

In this paper we have proposed the computation of upper bounds for the MAP threshold of LDPC codes over BMS channels in the asymptotic blocklength regime. It is based on the chain rule decomposition of the joint entropy of the syndrome bits, which becomes an integration in the limit -an area computation- and the cropping of dependencies such that the conditional entropy function can be computed by means of DE over subtrees of certain depth. The results are identical to the ones obtained from the generalized area theorem [6], which makes us suspect of their equivalence despite coming from different approaches.

We have also indicated how the proposed approach can be employed to obtain alternative expressions in terms of the area of entropy functions, which in the case of the BEC turns out to be the already known EXIT function area theorem.

REFERENCES

- [1] R.G. Gallager, "Low-Density Parity-Check Codes," *M.I.T. Press*, Cambridge, Massachusetts, 1963.
- [2] D. Burshtein, M. Krivelevich, S. L. Litsyn, and G. Miller, "Upper bounds on the rate of LDPC codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2437-2449, Sept. 2002.
- [3] G. Wiechman and I. Sason, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels: New bounds and applications," *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 550-579, February 2007.
- [4] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3221-3246, Sept. 2005.
- [5] C. Méasson, A. Montanari, and R. Urbanke, "Maxwell's construction: The hidden bridge between maximum-likelihood and iterative decoding," *ISIT*, Chicago, 2004.
- [6] C. Méasson, A. Montanari, T. Richardson, R. Urbanke, "The generalized area theorem and some of its consequences," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4793-4821, Nov. 2009.
- [7] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel property," *IEEE Trans. Inform. Theory*, vol. 50, no. 11, pp. 2657-2673, Nov. 2004.
- [8] I. Land, P. A. Hoeher, S. Huettinger, and J. Huber, "Bounds on information combining," *Int. Symp. Turbo Codes and Related Topics*, Brest, France, Sep. 1-5, 2003, pp. 39-42.
- [9] T. Richardson and R. Urbanke, "Modern Coding Theory," *Cambridge University Press*, 2008.
- [10] D. Matas, M. Lamarca, "Analysis of LDPC code syndrome entropy based on subgraphs," *to be published*.

²For the proof, we need to introduce the equality:

$$H(y_j | \mathbf{y}([j])) = H(y_j | \mathbf{y}([j]), x_j) + H(x_j | \mathbf{y}([j])) - H(x_j | \mathbf{y}([j+1]))$$