Min-sum decoding of irregular LDPC codes with adaptive scaling based on mutual information
Florence Alberge

To cite this version:
Florence Alberge. Min-sum decoding of irregular LDPC codes with adaptive scaling based on mutual information. International Symposium on Turbo Codes and Iterative Information Processing, Sep 2016, Brest, France. 2016, <10.1109/istc.2016.7593079>. <hal-01330905>

HAL Id: hal-01330905
https://hal.archives-ouvertes.fr/hal-01330905
Submitted on 13 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Min-sum decoding of irregular LDPC codes with adaptive scaling based on mutual information

Florence Alberge
Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506)
Univ. Paris-Sud - CNRS - CentraleSupélec - Université Paris-Saclay
3, rue Joliot-Curie, 91192 Gif-sur-Yvette
Tel: +33 1 69851757; fax: +33 1 69851765
e-mail:{florence.alberge@u-psud.fr}

Abstract—Min-Sum decoding (MS) is an alternative to belief propagation decoding with substantially lower complexity. MS often results in an overestimation of the log likelihood ratio (LLR) in particular in the early stage of the iterative process. A linear post-processing is usually performed as a compensation. With regular low density parity check codes (LDPC), a fixed scaling of the LLR yields sufficiently good results. In contrast, adaptive strategies are mandatory with irregular codes. It is well known that the scaling factor is an increasing function of the reliability of the LLR. In most of the publications, the scaling factor is envisioned as a function of both the iteration number and the signal-to-noise ratio. It is proposed here to use the mutual information between extrinsics as a measure of the reliability of the LLR. A practical implementation is derived with reasonable complexity. Compared to the literature, the proposed method yields slightly better results in terms of BER and a significant reduction in the number of iterations.

Index Terms—Parity check codes, iterative decoding, linear approximation, mutual information, adaptive LLR scaling.

I. INTRODUCTION

LDPC were introduced in the literature by Gallager [1] in the early sixties. Limited computing resources prevented Gallager from demonstrating the capabilities of message-passing decoders even for moderate code length. This work was ignored and rediscovered 30 years later by MacKay [2]. Combined with belief propagation (BP) decoding, LDPC codes exhibit performance near to the Shannon limit. The computational complexity of BP, especially at check-node, is not suitable for hardware implementation. Complexity reduction can be obtained through an approximation of the check-node calculation leading to min-sum decoding [3]. This efficient implementation suffers performance degradation which can be compensated by linear post-processing of the messages computed at the check nodes [4]. In this seminal work, a single correcting function is used during the whole iterative process and independently of the check-node. A density evolution analysis is often conducted in order to find the optimal value of the coefficients of the linear function [4]. This approach is not efficient enough with many irregular codes. It is observed in [5] that considering different post-processing depending on the degree of the nodes can improve the overall performance. Several authors suggest to consider in addition an adaptive rule for updating the linear function within the iterative process. It is proved in [6] that good results could be obtained with irregular codes by scaling the message at the check nodes with a scaling factor which is a function of both the degree of the node and of the mutual information. Since the mutual information is expected to increase during the iterative process, the value of the scaling factor should not be fixed. The practical implementation of the method is not solved in [6], a sub-optimal approach is derived in which the scaling factor is not anymore a function of two variables but is computed as a function of the sole iteration number. Recently, a variable scaling scheme based on the generalized mutual information was proposed in [7]. The method select the scaling factors as per iteration and as per check node degree independently overcoming the multi-dimensional issue faced by previous methods. We propose in this paper to make a step further. Here, the variable scaling factors are adjusted with respect to both check node degrees and mutual information with a particular focus on implementation aspects that are missing in [6]. It will be proved that the method can be implemented at low cost. Comparison with previously mentioned methods is also provided.

II. ITERATIVE DECODING

The iterative decoding of LDPC codes consists in an exchange of information between the variable and check nodes of the code. Let assume a \((N, K)\) LDPC code over an AWGN channel. Let \(R_i\) denote the log-likelihood ratio (LLR) of bit \(i\) derived from the observations. At iteration \(k\), let \(E_{ji}^{(k)}\) and \(M_{ji}^{(k)}\) be the LLR of bit \(i\) which are sent from check node \(j\) to variable node \(i\) and from variable node \(i\) to check node \(j\) respectively. The LLR \(E_{ji}^{(k)}\) and \(M_{ji}^{(k)}\) are defined if and only if the variable node \(i\) and the check node \(j\) are connected in the Tanner graph of the code. The standard BP algorithm proceeds in two steps [1]. The horizontal step reads:

\[
E_{ji}^{(k)} = 2 \tanh^{-1} \left( \prod_{j' \neq j} \tanh \left( \frac{M_{j'i}^{(k-1)}}{2} \right) \right) \tag{1}
\]

and is followed by a vertical step as :

\[
M_{ji}^{(k)} = R_i + \sum_{j' \neq j} E_{j'i}^{(k)} \tag{2}
\]
In min-sum (MS), the horizontal step is replaced by a simpler updating rule by modifying (1) into
\[ E^{(k)}_{ji} = \prod_{i' \neq i} \text{sgn}(M^{(k-1)}_{ji'}) \min_{i' \neq i} |M^{(k-1)}_{ji'}| \quad (3) \]

This approximation degrades the performance of the iterative process and yields over-estimated LLR especially in early iterations. From a statistical point of view, exchanged messages are usually modelled using Gaussian approximation [8] that can be seen as noisy estimates of the desired message \( X \). It is assumed here that \( X \in \{-1, 1\} \). Let \( L \) denote a LLR. Under Gaussian approximation \( L \) reads:
\[ L = \frac{\sigma^2}{2} X + \sigma n \quad (4) \]
where \( n \sim \mathcal{N}(0,1) \) and \( \sigma \in \mathbb{R^+} \setminus \{0\} \). In particular, the LLR that proceeds from the received data at the output of an AWGN channel can be modelled by (4). The validity of this assumption has been questioned in [9] for both regular and irregular LDPC codes and under AWGN channel. The main conclusion is that it is inappropriate to use a single parameter \( \sigma \) to model the statistical behaviour of \( L \). Alternatively, the mean and the variance of the density should be considered as independent variables leading to the following model:
\[ L = \alpha \frac{\sigma^2}{2} X + \sigma n \quad (5) \]
where \( \alpha \in \mathbb{R^+} \setminus \{0\} \). Let \( G_\alpha(X) \) denote the set of LLR\(^1\) following (5), then \( \alpha L \in G_{\frac{\alpha}{2}}(X) \) where \( G_{\frac{\alpha}{2}}(X) \) is the set of LLR following (4). LLR scaling is thus equivalent to propagating messages which are elements of \( G_{\frac{\alpha}{2}}(X) \). From theoretical point of view [5], [7], [6], the scaling factor should depend on the check node degree. The rationale behind that is the relationship between the check node degree and the reliability. It is well known that the scaling factor should be an increasing function of the reliability of the LLR and that the reliability is larger with low degrees. Putting all these facts together, the vertical step in min-sum decoding is replaced with
\[ M^{(k)}_{ji} = R_i + \sum_{d \in D(i)} \alpha(d) \sum_{j' \in M(d)} E^{(k)}_{j'i} \quad (6) \]
where \( D(i) \) is the set of indexes corresponding to parity-check equations in which bit \( i \) is involved and where \( M(d) \) is the set of indexes of parity-check equations with degree \( d \). In eq. (6), \( \sum_{j' \in M(d)} E^{(k)}_{j'i} \) is scaled in order to be in \( G_{\frac{\alpha}{2}}(X) \).

Since \( G_{\frac{\alpha}{2}}(X) \) is stable under addition, \( M^{(k)}_{ji} \in G_{\frac{\alpha}{2}}(X) \). As a consequence, the horizontal step remains unchanged since \( \{E^{(k)}_{j'i}\} \) is already scaled in (6) with \( \{\alpha(d)\} \). It has been previously mentioned that the numerical value of \( \{\alpha(d)\} \) depends on the reliability of the LLR. A natural measure of the reliability is the mutual information \( I(L, X) \). \( I(L, X) \) is unlikely to be obtained at the decoder. It was observed in [10] for both turbo-codes and regular LDPC codes that the scaling factor can be obtained as a function of a single parameter: the mutual information between extrinsics.

An efficient computation scheme for finite sequences is also provided leading to an estimation of the mutual information between extrinsics at the receiver side. In the next section, the results in [10] are briefly reviewed and extended to irregular LDPC codes. The method for computing \( \alpha(k)(d) \) from the mutual information between extrinsics is also detailed. The complexity of the proposed method is addressed in section IV in order to be compatible with a low-cost implementation. Numerical results are provided in section V.

### III. Mutual Information Between Extrinsics

Let \( L_y, L_z \in G_\alpha(X) \), the mutual information \( I(L_y, L_z) \) is defined as:
\[ I(L_y, L_z) = E_{p(L_y, L_z)} \left[ \log_2 \left( \frac{p(L_y, L_z)}{p(L_y)} \right) \right] \quad (7) \]

The mutual information between extrinsics is related to \( I(L_y, X) \) and \( I(L_z, X) \) through [10]:
\[ I(L_y, L_z) = I(L_y, X) + I(L_z, X) - I(L_y + L_z, X) \quad (8) \]
where \( I(L_y, X) = 1 - \int_{-\infty}^{+\infty} p_L(l\mid X = 1) \log_2(1 + e^{-\alpha l}) \, dl \) and \( L \) stands for \( L_y \) or \( L_z \). The evolution of \( I(L_y, L_z) \) for a given system can be obtained offline via an histogram method applied to each of the three terms involved in (8) as in EXIT Charts. Let \( L_y \) resp. \( L_z \) denote the random variable associated with \( \{\sum_j E_{j,i}\} \) resp. \( \{\sum_j M_{j,i}\} \). In min-sum decoding, \( E_{j,i} \) is the LLR of the probability that bit \( i \) causes parity check \( j \) to be satisfied. On the other side, \( M_{j,i} \) is the LLR of the probability of bit \( i \) based on the knowledge of the received sequence and of the probability of the other bits in parity-check \( j \). Thus, \( \sum_j E_{j,i} \) is the opinion of the parity-check nodes on bit \( i \) whereas \( \sum_j M_{j,i} \) is the opinion of the variable nodes on bit \( i \). Figure 1 depicts the extension of the single-curve EXIT-Chart presented in [10] to irregular LDPC codes. The principle is quite simple. All messages entering a vertical or a horizontal node belong to \( G_{\frac{\alpha}{2}}(X) \). The scaling factor \( \alpha(d) \) is computed easily from \( \{E_{j,i}\}_{j \in M(d)} \) since \( X \) is known. Then \( \{E_{j,i}\}_{j \in M(d)} \) is scaled and the vertical step is processed with (6). In parallel, \( I(L_y, L_z) \) is computed with (8) and an histogram method. From this scheme, the evolution of \( \alpha(d) \) can be plotted as a function of \( I(L_y, L_z) \). We consider as an example, an irregular code of rate \( \frac{1}{2} \) and length 10032 with degree distribution \( \rho(x) = 0.25x^4 + 0.75x^{14} \) and \( \lambda(x) = 0.1917x^2 + 0.0125x^3 + 0.0417x^4 + 0.0750x^5 + 0.2040x^6 + 0.0667x^7 + 0.225x^8 + 0.0833x^9 + 0.058x^{10} + 0.0541x^{12} \). This LDPC code was previously considered in [7] and is used through this paper for comparison purpose. This code has two check node degrees, 5 and 15. The evolution of \( \alpha(5) \) and \( \alpha(15) \) as a function of \( I(L_y, L_z) \) is given in Fig. 2. As expected, the scaling factor takes different values depending on the check-node degree. This fact was already established in [7] where the scaling factor is plotted as a function of

\(^1\)The term LLR is used indifferently for the log-ratio in (4) and (5) even if (5) does not strictly match the definition of a log-likelihood ratio.
The next section is devoted to the estimation of length $K$ the receiver side with finite length sequences. Here the value of the scaling factor is plotted as a function of similar to the ones in Fig. 5 of [7]. The main difference is that the iteration number. Let $\alpha$ obtained for a particular value of the SNR and are accurate locally. Let $\alpha$ and is valid for any signal to noise ratio (SNR) and $\alpha$ belongs to $\{1, +1\}$. An estimator $\tilde{I}_{yz}$ of $I(L_y, L_z)$ is given in [10] with expression:

$$\tilde{I}_{yz} = 1 + \frac{1}{K} \sum_{k=1}^{K} \log_2(s_k)$$

with $s_k = \sum_{x_k=1} e^{x_k(\ell_{y,k}+\ell_{z,k})}/(1+e^{\ell_{y,k}})(1+e^{\ell_{z,k}})^2$. An equivalent formulation is given below:

$$\tilde{I}_{yz} = 1 + \frac{1}{K} \sum_{k=1}^{K} \log_2 \left( \frac{1 + e^{\ell_{y,k}+\ell_{z,k}}}{(1+e^{\ell_{y,k}})(1+e^{\ell_{z,k}})} \right)$$

The accuracy of this estimator was evaluated in [10] for a serially concatenated turbo-code. Obviously, the distance between $\tilde{I}_{yz}$ and $I(L_y, L_z)$ depends on the system and also of the block-length $K$. A numerical evaluation is provided in Fig. 3 for the irregular LDPC code defined in the previous section. The mean square error (MSE) is plotted as a function of $I(L_y, L_z)$ where $I(L_y, L_z)$ is evaluated with the histogram method whereas $\tilde{I}_{yz}$ is obtained with (12). It can be observed that the histogram method is computationally demanding and necessitates the knowledge of X. This is an offline tool only. In contrast, $\tilde{I}_{yz}$ can be computed offline an online as well. From Fig. 3 we can conclude that, for the irregular code under consideration, the MSE is lower than $6.10^{-4}$ whatever the value of the mutual information is. On average, $|I(L_y, L_z) - \tilde{I}_{yz}| \leq 2.510^{-2}$ which gives an idea of the accuracy of the proposed estimator. The implementation of the method in min-sum decoding is considered below. Let $f(\ell) = \log \left( \frac{1 + e^{\ell}}{1 + e^{\ell}} \right)$ then

$$\tilde{I}_{yz} = 1 + \frac{1}{K\log(2)} \sum_{k=1}^{K} \left( f(\ell_{y,k} + \ell_{z,k}) - f(\ell_{y,k}) - f(\ell_{z,k}) \right)$$

Using the fact that $\log(e^a + e^b) = \max(a, b) + \log(1 + e^{-|a-b|})$, $f(\ell) = \max(0, \ell) + \log(1 + e^{-|\ell|}) \equiv \max(0, \ell) + g(\ell)$. It is
Min./Max. selection | $\sum_{j=1}^{M} r(i)(r(i)-2)$ | $0$ | $K-U$ \\

Table I: Complexity of $I_{yz}^0$, compared to MS

In this section, the performance of the proposed method (labelled “MS, $\alpha = f(IM)$” in the following) is evaluated in terms of bit error rate (BER) and number of iterations. The irregular LDPC code of rate $\frac{3}{4}$ and length 10032 with degree distribution $\rho(x) = 0.25x^4 + 0.75x^{14}$ and $\lambda(x) = 0.1917x + 0.0125x^2 + 0.0417x^3 + 0.0750x^5 + 0.2040x^6 + 0.0667x^7 + 0.225x^8 + 0.0833x^9 + 0.058x^{10} + 0.0541x^{12}$ is again considered. Comparison is provided with classical methods and also with the most recent literature. The following schemes are considered: (SP) iterative decoding with sum-product implementation, (MS, $\alpha = f(IM)$) min-sum implementation with scaling factor as in (6) in which the scaling factors are given by (9-10) and $I = I_{yz}^0$. (MS, $\alpha = f(EbN0_{threshold}; it)$) min-sum implementation with scaling factor acquired using training LLR and with a value of $EbN0_{threshold}$ corresponding to the practical threshold ($EbN0_{threshold}$) of the code [7].

The results are given in Fig. 4 for BPSK modulation and in Fig. 5 for 256-QAM modulation. For each implementation, the same scaling sequence (if apply) is used for both BPSK and 256QAM. The main gain obtained with our method compared to [7] lies in the reduction in the number of iterations. Indeed, as mentioned above, the scaling sequence $\alpha = f(EbN0_{threshold}; it)$ has been acquired for a particular value of $SNR EbN0_{threshold}$. When the true $SNR$ is significantly above $EbN0_{threshold}$, the reliability of the LLR is

$M \times K$, $r(i)$ denotes the number of ones in raw $i$ of the parity-check matrix and $U$ is the number of bit positions in which $sign(\ell_{y,k}) = sign(\ell_{z,k})$. The comparison is conducted for a single iteration and for the main operations (addition and minimum or maximum selection).
rapidly increasing and the mutual information between extrinsics as well. In that case, higher values of the scaling factors could be chosen even in first iterations. Using the rule $\alpha = f(EbN_0_{threshold}, it)$ slows down the convergence since the same rule is used whatever the value of $EbN_0$ is. The same problem arises with DNMS. With BPSK, the iteration gain with the proposed method compared to $\alpha = f(EbN_0_{threshold}, it)$ is approximately 45% with $EbN_0 = 2dB$ and up to 60% with higher $EbN_0$.

VI. CONCLUSION

In this paper, the practical implementation of adaptive LLR scaling has been investigated. The originality of the proposed method is the use of a mutual information driven rule for the scaling factor. This is an old idea but, up to our knowledge, this is the first time that an implementation is proposed with acceptable complexity. The efficiency of the proposed method has been demonstrated in terms of BER and convergence rate. It is worth mentioning that the estimate of the mutual information between extrinsics could also be used as an early stopping criterion or sent back to the transmitter via a feedback channel. The last point could be of interest in several applications including secure transmissions.

REFERENCES