Coupled LDPC Codes: Complexity Aspects of Threshold Saturation

Lentmaier, Michael; Fettweis, Gerhard

Published in: [Host publication title missing]

DOI: 10.1109/ITW.2011.6089581

2011

Citation for published version (APA):
Lentmaier, M., & Fettweis, G. (2011). Coupled LDPC Codes: Complexity Aspects of Threshold Saturation. In [Host publication title missing] (pp. 668-672). IEEE--Institute of Electrical and Electronics Engineers Inc.. DOI: 10.1109/ITW.2011.6089581
Abstract—We analyze the convergence behavior of iteratively decoded coupled LDPC codes from a complexity point of view. It can be observed that the thresholds of coupled regular LDPC codes approach capacity as the node degrees and the number \( L \) of coupled blocks tend to infinity. The absence of degree two variable nodes in these capacity achieving ensembles implies for any fixed \( L \) a doubly exponential decrease of the error probability with the number of decoding iterations \( I \), which guarantees a vanishing block error probability as the overall length \( n \) of the coupled codes tends to infinity at a complexity of \( O(n \log n) \). On the other hand, an initial number of iterations \( I_{0} \) is required until this doubly exponential decrease can be guaranteed, which for the standard flooding schedule increases linearly with \( L \). This dependence of the decoding complexity on \( L \) can be avoided by means of efficient message passing schedules that account for the special structure of the coupled ensembles.

I. INTRODUCTION

For the binary erasure channel (BEC), capacity achieving sequences of irregular low-density parity-check (LDPC) codes can be found, for which a vanishing block error probability as the overall length \( n \) of the block codes tends to infinity at a complexity of \( O(n \log n) \). On the other hand, ensembles with thresholds close to capacity usually have poor minimum distance properties compared to Gallager’s original regular LDPC ensembles [2], resulting in comparatively high error floors.

It can be observed that the coupling of several regular LDPC code blocks together to a terminated convolutional code [3] leads to a slight irregularity at the ends of the Tanner graph, resulting in substantially better belief propagation (BP) decoding thresholds compared to their tail-biting version or the block codes they are constructed from [4] [5] [6]. In particular, as the number \( L \) of coupled blocks increases the BP decoding threshold of the coupled ensembles converges to the optimal maximum a-posteriori probability (MAP) decoding threshold of the underlying block ensemble. At the same time, it can be shown that the minimum distance of the coupled ensemble grows linearly with the block length as the block length tends to infinity, i.e., they are asymptotically good [7] [8].

In this paper we analyze the convergence behavior of BP decoded coupled LDPC codes. We show that for any \( L \) a vanishing block error probability is achievable at a complexity of \( O(n \log n) \). In the low signal-to-noise ratio (SNR) region, between the thresholds of the coupled ensemble and its underlying block ensemble, a dramatic increase in the decoding complexity with \( L \) can be avoided by efficient message passing schedules. An attractive, practical implementation of such a schedule is the sliding window decoder considered in [9], [10] and [4], for which both latency and complexity are independent of \( L \).

II. COUPLED LDPC CODE ENSEMBLES

A. From Block to Convolutional Codes

Consider the transmission of a sequence of codewords \( \mathbf{v}_t, t = 1, \ldots, L \), using an LDPC block code. Let the structure of this code be defined by a protograph with \( n_c \) check nodes and \( n_v \) variable nodes. The bi-adjacency matrix \( \mathbf{B} \) of the protograph is called its base matrix. The parity-check matrix \( \mathbf{H} \) of an individual code of length \( n_t = n_c N \) is then constructed by replacing each 1 in \( \mathbf{B} \) by an \( N \times N \) permutation matrix and each 0 by an \( N \times N \) all-zero matrix. Also multiple edges between a pair of nodes are allowed in a protograph, which are represented by integer entries larger than one in the base matrix and are replaced by a sum of permutation matrices. The performance of the code depends on the lifting factor \( N \), the chosen set of permutation matrices, and the structure of the protograph. A chain of protographs of a \((3,6)\)-regular LDPC code with \( n_c = 2 \), \( n_v = 1 \) and base matrix \( \mathbf{B} = [3, 3] \) is depicted in Fig. 1(a) for \( L = 6 \).

An essential feature of LDPC convolutional (LDPCC) codes [3] is that the blocks of different time instants are interconnected. Instead of encoding all codewords independently,
the blocks \( v_t \) are coupled by the encoder over various time instants \( t \). The maximal distance between a pair of coupled blocks defines the memory \( m_{cc} \) of the convolutional code. The coupling of consecutive blocks can be achieved by an edge spreading procedure [11] that divides the edges from variable nodes at time \( t \) among equivalent check nodes at times \( t+i, i=0,\ldots, m_{cc} \). The resulting ensemble can be described by means of a convolutional protograph with base matrix

\[
B_{[1,L]} = \begin{bmatrix}
B_0 & \cdots & \cdots & B_{m_{cc}} & \cdots & \cdots
\end{bmatrix}.
\] (1)

The corresponding sequence of coupled code blocks forms a codeword \( v = [v_1, v_2, \ldots, v_t \ldots v_L] \) of a terminated LDPC convolutional code. Note that \( m_{ce} n_c \) additional check nodes result in a rate loss due to termination. The block coding ensemble with disconnected protographs corresponds to the special case \( m_{cc} = 0 \) and \( B_0 = B \).

In general, we can apply this procedure to any given regular or irregular protograph. In order to maintain the degree distribution and the structure of the original ensemble, a valid edge spreading should satisfy the condition

\[
\sum_{i=0}^{m_{cc}} B_i = B.
\] (2)

This condition ensures that the entries of \( B \) are divided among the matrices \( B_i \) in such a way that the sums over the columns and rows of \( B_{[1,L]} \) are equal to those of \( B \). The only exception are the first and last \( m_{ce} n_c \) rows of \( B_{[1,L]} \), whose weights are reduced as a result of the termination at the ends of the convolutional ensemble.

Example 1: Two different variants of spreading the edges of the (3,6)-regular LDPC ensemble in Fig. 1(a) are illustrated in Fig. 1(b) and the corresponding component base matrices within \( B_{[1,L]} \) (see (1)) are equal to

\[
\begin{align*}
\text{Ens. (3,6)-A:} & \quad B_0 = B_1 = B_2 = [1,1], \quad m_{cc} = 2, \quad \text{(3)} \\
\text{Ens. (3,6)-B:} & \quad B_0 = [2,2], \quad B_1 = [1,1], \quad m_{cc} = 1.
\end{align*}
\]

Analogously, starting from a (4,8)-regular LDPC ensemble with base matrix \( B = [4,4] \) we can define different convolutional ensembles with

\[
\begin{align*}
\text{Ens. (4,8)-A:} & \quad B_0 = B_1 = B_2 = B_3 = [1,1], \quad m_{cc} = 3, \quad \text{(4)} \\
\text{Ens. (4,8)-B:} & \quad B_0 = B_1 = [2,2], \quad m_{cc} = 1, \\
\text{Ens. (4,8)-C:} & \quad B_0 = [3,3], B_1 = [1,1], \quad m_{cc} = 1, \\
\text{Ens. (4,8)-D:} & \quad B_0 = [2,2], B_1 = B_2 = [1,1], \quad m_{cc} = 2.
\end{align*}
\]

Note that all the given examples satisfy (2).

B. Threshold Saturation

It follows from the construction that the check nodes at the start and end of the coupled protograph have lower degrees (see also Fig. 1), resulting in a slight irregularity with stronger protection of the symbols associated with the connected variable nodes. As \( L \to \infty \), the fraction of lower degree check nodes vanishes and the degree distribution of the coupled ensemble \( B_{[1,L]} \) converges to that of the original block code ensemble \( B \). As a consequence, each of the convolutional protographs ensembles considered in Example 1 defines a sequence of asymptotically regular LDPC codes. Despite of the vanishing fraction of stronger check nodes, it turns out that the coupled ensembles have a substantially better BP decoding threshold than the block ensembles they are constructed from.

In particular, as \( L \to \infty \) the BP decoding threshold of the coupled ensembles converges to the optimal MAP decoding threshold of the underlying block ensemble. For regular LDPC codes this threshold saturation phenomenon has been proven analytically for the BEC in [5], and it can be observed empirically for the additive white Gaussian noise (AWGN) channel as well [6] [8]. The slight structured irregularity of the coupled ensembles leads to BP decoding thresholds that approach the Shannon limit as the node degrees increase.

Table I shows the AWGN channel thresholds \( E_b/N_0^* \) of the ensembles defined in (3) and (4) together with thresholds \( E_b/N_0^{\text{blk}} \) of their uncoupled block counterparts. The values have been computed by the discretized density evolution technique [12]. It can be observed that the thresholds may slightly change for different variations of edge spreading. Also shown are the thresholds \( E_b/N_0^{\text{win}} \) (\( W = 50 \)) corresponding to the window decoder considered in Section IV-C [9] [10]. Some degradation from \( E_b/N_0^* \) to \( E_b/N_0^{\text{win}} \) can be observed for the asymmetric protographs of Ensembles (3,6)-B, (4,8)-C and (4,8)-D.

<table>
<thead>
<tr>
<th></th>
<th>( E_b/N_0^* )</th>
<th>( E_b/N_0^{\text{win}} )</th>
<th>( E_b/N_0^{\text{blk}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3,6)-A</td>
<td>0.46 dB</td>
<td>0.46 dB</td>
<td>1.11 dB</td>
</tr>
<tr>
<td>(4,8)-A</td>
<td>0.26 dB</td>
<td>0.26 dB</td>
<td>1.55 dB</td>
</tr>
<tr>
<td>(4,8)-B</td>
<td>0.32 dB</td>
<td>0.32 dB</td>
<td>1.11 dB</td>
</tr>
<tr>
<td>(4,8)-C</td>
<td>0.27 dB</td>
<td>0.79 dB</td>
<td></td>
</tr>
<tr>
<td>(4,8)-D</td>
<td>0.26 dB</td>
<td>0.29 dB</td>
<td></td>
</tr>
</tbody>
</table>

Table I: AWGN channel thresholds: different coupled ensembles are compared to their uncoupled block counterparts.

III. CONVERGENCE BEHAVIOR OF BP DECODING

A. An upper bound on the error probability

Consider an LDPC block code with minimal variable node degree \( J_{\text{min}} > 2 \) and maximal check node degree \( K_{\text{max}} \). Let \( P_{\text{max}} \) denote the maximum hard decision error probability of variable-to-check node messages after \( I \) iterations of BP decoding, where the maximum is taken over all edges in the Tanner graph of the code. Assume now that we can show by density evolution that

\[
P^{(J_{\text{min}})}_{\text{max}} < B_{\text{gr}}^2/4,
\] (5)
where $B_{br}$ denotes the *breakout value*, defined as

$$B_{br} = A^{-1/(J_{min}-2)}(K_{max}-1)^{-(J_{min}-1)/(J_{min}-2)} \ ,$$  \hspace{1cm} (6)

and $A = \exp(-RE_b/N_0)$ denotes the Bhattacharyya parameter of the AWGN channel. Then, after a total number of $I$ iterations, $I > I_{br}$, the bit error probability can be upper bounded by [13]

$$P_b < \exp\left(-a(J_{min} - 1)^{I-I_{br}}\right) \ ,$$  \hspace{1cm} (7)

where $a$ is some positive constant. From $J_{min} > 2$ follows a doubly exponential decrease of the bit error probability with the number of decoding iterations.

An assumption in the derivation of (7) is that all messages exchanged in the $I$ decoding iterations are statistically independent, which puts some restrictions on the Tanner graph structure and the length of the code. For protograph based ensembles, by proper selection of the code, it is possible to perform $I > \alpha \log N - \beta$ independent iterations, where $N$ denotes the required lifting factor and $\alpha, \beta$ are some positive constants. Equation (8) follows either from Gallager’s deterministic construction algorithm in [2] or from a protograph-based modification of the random ensemble approach in [13]. The latter was applied to LDPC convolutional code ensembles in [14]. Combining (8) and (7) we can relate the achievable bit error probability to $N$, i.e.,

$$P_b < \exp\left(-a'N^{\gamma}\right) , \ a' > 0 \ ,$$  \hspace{1cm} (9)

and establish an upper bound on the block error probability

$$P_B < n_tP_b < n_vN \exp\left(-a'N^{\gamma}\right) , \ a' > 0 \ .$$  \hspace{1cm} (10)

It follows that the block error probability $P_B$ tends to zero whenever $P_b$ tends to zero as $N$ tends to infinity. Since the complexity per iteration scales linearly with $N$ it follows that a vanishing block error probability can be achieved with an overall decoding complexity of $O(N \log N)$.

**B. Application to Coupled Ensembles**

A common feature of other capacity approaching LDPC ensembles is the presence of degree two variable nodes, resulting in $J_{min} = 2$. The fraction of these nodes increases when the thresholds are optimized [1], which on the other hand reduces the convergence speed and lowers the error floor.

For coupled ensembles, defined by $B_{[1,L]}$ according to (1) and (2), all variable node degrees are equal to those of the underlying ensemble defined by $B$. Starting from a block ensemble with $J_{min} \geq 3$ we can construct terminated convolutional ensembles for arbitrary numbers of sub-blocks $L$. For large $L$ the termination rate loss decreases and the BP threshold of the coupled ensemble approaches the MAP threshold corresponding to $B$ [5] [6]. We can then apply the bounds from Section III-A to the convolutional protograph corresponding to $B_{[1,L]}$, defining an ensemble of specially structured LDPC block codes with overall length $n = n_tL = n_vNL$. After an initial number of iterations $I_{br}$, determined by the breakout value condition (5), a doubly exponential decrease of $P_b$ can be guaranteed for any channel parameter $\sigma$ below the threshold $\sigma^*$ of the coupled ensemble, resulting in a vanishing block error probability $P_B$ for each $v_t$. Replacing $n_t$ by $n = n_tL = n_vNL$ in (10) we see that for the overall sequence $v$ the block error probability bound increases linearly with $L$, but still it vanishes for any $L$ as $N$ tends to infinity.

**IV. COMPLEXITY ASPECTS OF THRESHOLD SATURATION**

**A. The Two Convergence Regions of Couple Ensembles**

The required number of initial iterations $I_{br}$ in general depends on the structure of the ensemble and on the gap to the threshold. For Ensemble (3,6)-A and Ensemble (4,8)-A the behavior of $I_{br}$ as function of $L$ is shown in Fig. 2 for different standard deviations $\sigma$ of the AWGN channel. Due to the special structure of the coupled ensembles it can be observed that the threshold $\sigma^*_b$ of the underlying block ensemble separates the principle behavior of $I_{br}$ into two different convergence regions:

- **Low SNR region** ($\sigma^*_b < \sigma < \sigma^*$): $I_{br}$ increases linearly with $L$.
- **High SNR region** ($\sigma < \sigma^*_b < \sigma^*$): $I_{br}$ approaches a constant value as $L$ increases.

In order to understand the behavior in the low SNR region, recall that the threshold saturation effect stems from the lower check node degrees at the ends of the convolutional protograph, which are enabled by the additional $m_{cc}n_c$ check nodes due to termination. As the number of iterations increases, the improved performance from these low-degree check nodes propagates through toward the center of the graph until, eventually, a low bit error probability can be observed for all $t$ [15]. The price for this progression, however, is an increase in the required number of iterations with $L$.

In general, $I_{br}$ is upper-bounded by the value corresponding to the underlying block ensemble. Since the graph structure of
MPS-I: standard (parallel) flooding schedule
MPS-II: as MPS-I, with deactivation of nodes for which a target probability $P_b^{\text{max}}$ is reached
MPS-III: as MPS-II, with deactivation of nodes with insignificant improvement
MPS-IV: as MPS-I within a sliding window of size $W$

Table II
Overview of message passing schedules considered in [15].

![Figure 3](image-url) Effective number of iterations $I_{\text{eff}}$ for $L = 100$ as function of the signal-to-noise ratio for MPS-I, MPS-III, and MPS-IV. A target probability of $P_b^{\text{max}} = 10^{-5}$ is considered.

this ensemble is preserved in the center region of the coupled graph, all messages exchanged during the iterations will be at least as reliable as those in the graph of the block ensemble. And in the high SNR region the block ensemble will satisfy the breakout value condition (5) within a finite number $I_{\text{br}}$.

B. Reduced Complexity Schedules for Coupled Ensembles

Up to this point a standard BP decoder with flooding schedule was assumed in our analysis. It can be shown, however, that the decoding complexity of coupled ensembles can be dramatically decreased by means of efficient message passing schedules (MPSs) that take advantage of the convolutional protograph structure by omitting superfluous node updates during the decoding iterations [15]. The considered schedules are summarized in Table II.

Let $I_t$ denote the number of times the variable and check nodes at position $t$ are updated during the iteration procedure. While for the standard decoder with MPS-I $I_t$ is constant and equal to the total number of iterations $I$ for all $t$, for the other schedules the value $I_t$ depends on the position in the protograph. The average decoding complexity per symbol is then proportional to the effective number of iterations, defined as

$$I_{\text{eff}} = \frac{1}{L} \sum_{t=1}^{L} I_t.$$  

Figure 3 shows $I_{\text{eff}}$ as function of the signal-to-noise ratio $E_b/N_0$ for Ensembles (3,6)-B and (4,8)-B and $L = 100$. For MPS-I, which was the schedule used in Fig. 2, we can again see the dramatic increase in complexity in the low SNR region beyond the threshold $E_b/N_0^*_{\text{blk}}$ of the block ensembles. The complexity of MPS-III, on the other hand, is lower than that of MPS-I throughout the entire channel parameter range and shows a substantial increase only close to the threshold $E_b/N_0^*$. An even more important property of MPS-III is the fact that $I_{\text{eff}}$ does no longer increase with $L$ [15].

Note that the chosen target error probability $P_b^{\text{max}} = 10^{-5}$ in Fig. 3 is always below the value required by (5). As a consequence, if we assume that the decoder switches to MPS-I after $I_{\text{eff}}$ initial iterations with MPS-III, (7) and (10) can be applied with $I_{\text{br}} = I_{\text{eff}}$. With this approach we can take advantage of the threshold saturation phenomenon for arbitrary $L$ with vanishing block error probability without a dramatic increase in the decoding complexity.

C. Low-Complexity and Low-Latency Window Decoding

The MPS-III decoding rule adapts the set of updated variable nodes and check nodes in each decoding iteration according to the improvement of the message reliabilities. It is interesting to observe in Fig. 3 the resulting transition of $I_{\text{eff}}$ between the two SNR regions in the vicinity of the block ensemble threshold $E_b/N_0^*_{\text{blk}}$. In fact, in the high SNR region the MPS-III decoding rule effectively implements the MPS-II rule, which in turn becomes equivalent to the MPS-I rule as $E_b/N_0$ increases. In the low SNR region, due to the structure of the coupled ensemble, the set of updated nodes moves from both ends of the protograph toward the center during the decoding process.

As $E_b/N_0$ decreases, a natural approximation of MPS-III can be easily implemented by means of the sliding window decoder (MPS-IV) [9] [10], whose effective number of iterations is also shown in Fig. 3. At each position $p = 1, \ldots, L$ of the window, $t_p^{\text{win}}$ iterations are performed for nodes at $t = p, \ldots, p + W - 1$ until the target error probability is achieved at $t = p$. The effective number of iterations is shown in Table III for the ensembles introduced in (3) and (4). For large $L$, since $t_p^{\text{win}}$ is almost constant for $p > 1$, the complexity increases linearly with the window size $W$. On the other hand, for small values $W$ an increase in $t_p^{\text{win}}$ can be observed, so that an optimal window size in terms of complexity can be identified. Observe that Ensembles (3,6)-B with $W = 6$ and (4,8)-D with $W = 4$ offer a good complexity-latency trade-off [9] despite of their degraded thresholds shown in Table I, which motivates their choice in our analysis. An important feature of the window decoder compared to MPS-III is that not only the complexity but also the latency and size of the decoder is determined by the window size $W$ and independent of $L$. 

<table>
<thead>
<tr>
<th>$E_b/N_0$</th>
<th>$I_{\text{eff}}$ (MPS-I)</th>
<th>$I_{\text{eff}}$ (MPS-III)</th>
<th>$I_{\text{eff}}$ (MPS-IV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>50</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>0.6</td>
<td>100</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>0.8</td>
<td>150</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>1.0</td>
<td>200</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>1.2</td>
<td>250</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>1.4</td>
<td>300</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>1.6</td>
<td>350</td>
<td>70</td>
<td>35</td>
</tr>
<tr>
<td>1.8</td>
<td>400</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>2.0</td>
<td>450</td>
<td>90</td>
<td>45</td>
</tr>
</tbody>
</table>

Table III
Effective number of iterations $I_{\text{eff}}$ for $L = 100$ as function of the signal-to-noise ratio for MPS-I, MPS-III, and MPS-IV. A target probability of $P_b^{\text{max}} = 10^{-5}$ is considered.
for lifting factors that are suitable for practical applications.

The pipeline decoder and the window decoder with MPS-IV convolutional codes [3] [16] [17], which is characterized by implemented by the traditional pipeline decoder for LDPC block ensemble. This decoding rule can efficiently be in each of the processors. Investigating the trade-offs between processing with a small number of operations per time unit a larger initial delay but allows for highly parallel high-speed performance of Ens (3,6)-B (MPS-IV) with MPS-III and the (3,6) block ensemble (MPS-I).

Fig. 4. Medium to high SNR region: comparison of \(I_{\text{eff}}\) for window decoding of Ens (3,6)-B (MPS-IV) with MPS-III and the (3,6) block ensemble (MPS-I).

A closer look at high SNR values in Fig. 4 reveals that after some \(E_b/N_0\) the window decoder complexity is slightly higher than that of MPS-III or the block ensemble with MPS-I. This can be prescribed to the overhead due to the overlap regions of the sliding window, whose effect could be reduced in the high SNR region by a larger step-width of the window position \(p\). Note, however, that this would require an increase in the overall window size \(W\), since an overlap size below the values of \(W\) in Table III would again increase the amount of iterations \(I_p\). In this region, a decoder of the coupled ensemble with standard MPS-I would have the same \(I_{\text{eff}}\) as the block ensemble. This decoding rule can efficiently be implemented by the traditional pipeline decoder for LDPC convolutional codes [3] [16] [17], which is characterized by a larger initial delay but allows for highly parallel high-speed processing with a small number of operations per time unit in each of the processors. Investigating the trade-offs between the pipeline decoder and the window decoder with MPS-IV for lifting factors that are suitable for practical applications form an interesting area for further research.

\[\text{Table III} \]

| Ens (3,6)-A | \(W\) | 11 | 14 | 15 | 16 |
| Ens (3,6)-B | \(W\) | 4 | 5 | 6 | 7 |
| Ens (4,8)-A | \(W\) | 9 | 10 | 12 | 13 |
| Ens (4,8)-B | \(W\) | 3 | 4 | 6 | 8 |
| Ens (4,8)-D | \(W\) | 4 | 5 | 6 | 8 |

\[\text{Number of required effective iterations } I_{\text{eff}} \text{ with MPS-IV and window sizes } W. \text{ An AWGN channel with } \sigma = 0.923 \text{ (} E_b/N_0 = 0.7dB \text{ as } L \to \infty \) \text{ and a target probability of } P_b^{\text{max}} = 10^{-5} \text{ is considered}. \]

\[\text{Fig. 4.} \text{ Medium to high SNR region: comparison of } I_{\text{eff}} \text{ for window decoding of Ens (3,6)-B (MPS-IV) with MPS-III and the (3,6) block ensemble (MPS-I).} \]

1 Ensemble (4,8)-C does not appear in the table because its threshold \(E_b/N_0_{\text{win}}\) is above the considered \(E_b/N_0\).