Finite-Length Scaling of Regular LDPC Codes over $GF(2^m)$

for the Binary Erasure Channel

Iryna Andriyanova

joint work with Kenta Kasai (Tokyo IT)

ETIS Group
ENSEA / University of Cergy-Pontoise / CNRS
Cergy, France

May 17, 2009
Non-binary LDPC Codes

We consider:
- Regular codes: symbol nodes of degree $l$, $\oplus$ of degree $r$
- over GF, $q = 2^m$
- $f$: linear bijective mappings (equiprobable subspaces of the same dimension)
Why Non-binary LDPC Codes Are Interesting?

Improved performance for short-lengths observed by Mackay, Hu, Shu Lin and others

[Hu2002] WER of (2, 4) LDPCCs of binary length $\approx 660$
Setup

But ...

Not much is known about the iterative decoding of non-binary codes, even over simple binary-input channels.

We assume the transmission over the binary erasure channel with probability \( \epsilon \) (\( BEC(\epsilon) \)):

- the transmitted bit is erased at the receiving end w.p. \( \epsilon \)
- or perfectly known

Our framework: regular LDPCCs over \( GF(2^m) \) of fixed binary length \( mn + BEC(\epsilon) \)

Our Goal

Decay of the Word-Erasure-Rate (WER): does it always improve with the alphabet size?
Finite-Length Scaling of Regular LDPC Codes over $GF(2^m)$

Note on Asymptotic Analysis

Iterative decoding threshold:

$$\epsilon^* = \min_\epsilon \{ \epsilon : P_b > 0 \}$$

A symbol from $GF(2^m)$ is seen as a binary vector of length $m$.

$\epsilon^*$ given by density evolution [Rathi&Urbanke’06] :

$$X^{(l)} = E_\epsilon \Box \left( (X^{(l-1)})^{\Box (d-1)} \right)^\Box (c-1)$$

$X^{(l)}$ - $(m + 1)$-length probability vector of messages $\bullet \rightarrow \oplus$ at iteration $l$

$E_\epsilon$ - $(m + 1)$-length probability vector of input messages

$\Box$ - operation at nodes $\bullet$, $\bigcirc$ - operation at nodes $\oplus$
Main Result

Basic Scaling Law for \((c, d)\) LDPCCs over \(GF(2^m)\)

Assume the transmission over \(BEC(\epsilon)\) with a \((c, d)\) LDPC code over \(GF(2^m)\) of length \(n\), \(c > 2\).
Let \(h^*\) be the fraction of not revealed bits when \(X = X^*\) and assume it to be Gaussian distributed. Then

\[
\mathbb{E}[P_W(\epsilon)] = Q \left( \frac{\sqrt{n}(\epsilon^* - \epsilon)}{\alpha} \right) + O(n^{-1/6})
\]

with \(Q(x) = 0.5\text{erfc}(x/\sqrt{2})\) and

\[
\alpha = \left. \frac{\partial^2 \epsilon(x)}{\partial x^2} \right|_{x=x^*} \lim_{\epsilon \to \epsilon^*} \frac{x - x^*}{1 - (c - 1)(d - 1)\lambda_2} \sqrt{\frac{\xi}{c}}.
\]

Here \(\epsilon(x)\) - an EXIT-like curve, \(\lambda_2\) - the second eigenvalue of a transition matrix \(M\), \(\xi\) is given in terms of \(M\).
Some Remarks

- **Concentration theorem**: for $\epsilon$ close to $\epsilon^*$, the performance of almost any code of the ensemble is close to the average.

- The result is an extension of the basic scaling law for binary LDPC codes [Amraoui’06], [Ezri et al’08]

- **Direct extension**: erased messages in the binary case become vector messages with $k$ erasures, $k = 1, \ldots, m$

- **Our approach**: using duality, the known messages are only tracked
Comparison with Simulations: 
(3, 6)-LDPC over $GF(2)$

Consistent with the results obtained in the binary case
Comparison with Simulations:
(3, 6)-LDPC over $GF(4)$
Comparison with Simulations:
(3, 6)-LDPC over $GF(8)$
Sketch of the Proof

1. Average WER vs $\epsilon$:

$$
\mathbb{E}[P_W(\epsilon)] \approx Q \left( \frac{\epsilon^* - \epsilon}{\sigma_h} \right)
$$

2. Link between $\sigma_h^2$ and the variance of the number of known messages $\mathcal{V}$ for $\epsilon \to \epsilon^*$

$$
\sigma_h^2 = \frac{1}{nc} \left( \frac{\partial^2 \epsilon(x)}{\partial x^2} \right) \lim_{\epsilon \to \epsilon^*} (x - x^*)^2 \mathcal{V}
$$

3. Computation of $\mathcal{V}$ close to $\epsilon^*$

$$
\mathcal{V} = \left( \frac{\xi}{1 + (c - 1)(d - 1)\lambda_2} \right)^2 + o \left( \frac{\xi}{1 + (c - 1)(d - 1)\lambda_2} \right)^2
$$
Step 1 : Variance $\sigma^2_h$

$h^*$ : fraction of not revealed bits when $X = X^*$

$h^*$ is a random variable with mean $\epsilon^*$ and some variance $\sigma^2_h$

$h^*$ is supposed to be Gaussian distributed, which implies

$$\mathbb{E}[P_W(\epsilon)] \approx Q\left(\frac{\epsilon^* - \epsilon}{\sigma_h}\right)$$

$\sigma^2_h$ ?

In the binary case : directly

In our case : consider the dual quantity - fraction $\bar{h}^*$ of revealed bits when $X = X^*$, $\bar{h}^* = 1 - h^*$

$\bar{h}^*$ is a random variable with mean $1 - \epsilon^*$ and variance $\sigma^2_h$
Step 1: Variance $\sigma_h^2$

$h^*$ : fraction of not revealed bits when $X = X^*$

$h^*$ is a random variable with mean $\epsilon^*$ and some variance $\sigma_h^2$

$h^*$ is supposed to be Gaussian distributed, which implies

$$\mathbb{E}[P_W(\epsilon)] \approx Q\left(\frac{\epsilon^* - \epsilon}{\sigma_h}\right)$$

$\sigma_h^2$?

In the binary case: directly

In our case: consider the dual quantity - fraction $\bar{h}^*$ of revealed bits when $X = X^*$, $\bar{h}^* = 1 - h^*$

$\bar{h}^*$ is a random variable with mean $1 - \epsilon^*$ and variance $\sigma_h^2$
Step 2: Link between $\sigma_h^2$ and $V$

$x$: fraction of revealed bits after iterative decoding for given $\epsilon$

EXIT-like curve $\epsilon(x)$ (for $n \to \infty$):

Expansion around $(\epsilon^*, x^*)$:

$$\Delta \epsilon = \frac{\partial^2 \epsilon}{\partial x^2} |^{x=x^*} \Delta x (x - x^*)$$

$$\sigma_h^2 = \mathbb{E}[(\Delta \epsilon)^2] = \left( \frac{\partial^2 \epsilon}{\partial x^2} |^{x=x^*} \right)^2 \lim_{\epsilon \to \epsilon^*} (x - x^*)^2 \mathbb{E}[(\Delta x)^2]$$
Step 2 : Link between $\sigma_h^2$ and $\mathcal{V}$

Define $\mathcal{V} = nc\mathbb{E}[(\Delta x)^2]$, then

$$
\sigma_h^2 = \frac{1}{nc} \left( \frac{\partial^2 \epsilon}{\partial x^2} \right)_*^2 \lim_{\epsilon \to \epsilon^*} (x - x^*)^2 \mathcal{V}
$$
Step 3 : Dominant term of $\mathcal{V}$

$N_x$ : number of known messages in the decoder

One can show

$$\mathcal{V} = \frac{\mathbb{E}[N_x - \mathbb{E}[N_x]]^2}{nc}$$

One should consider all the messages in the iterative decoder and compute correlations between them. Fortunately, there is a dominant term!
Step 3 : Dominant term of $\mathcal{V}$

2 transition matrices of dimensions $(2m + 2) \times (2m + 2)$
$A$ and $B$ (depending on $C$ and $V$)

$M = BA$

- $\lambda_1 = 1$ and is simple
- $\lambda_2$ is assumed to have multiplicity 2
- $\xi = c_1(c - 1)^2(d - 1)e_2^{(2)}$, $e_2^{(2)}$ - generalized eigenvector related to $\lambda_2$
- $\mathcal{V} \rightarrow \left( \frac{\xi}{1 -(c-1)(d-1)\lambda_2} \right)^2$ as $\epsilon \rightarrow \epsilon^*$
Step 3 : Dominant term of $\mathcal{V}$

2 transition matrices of dimensions $(2m + 2) \times (2m + 2)$

$A$ and $B$ (depending on $C$ and $V$)

$M = BA$

- $\lambda_1 = 1$ and is simple
- $\lambda_2$ is assumed to have multiplicity 2
- $\xi = c_1(c - 1)^2(d - 1)e_2^{(2)}$, $e_2^{(2)}$ - generalized eigenvector related to $\lambda_2$
- $\mathcal{V} \rightarrow \left(\frac{\xi}{1-(c-1)(d-1)\lambda_2}\right)^2$ as $\epsilon \rightarrow \epsilon^*$
Final Example of (3, 6) LDPC Codes

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\epsilon^*$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.42944</td>
<td>0.559044</td>
</tr>
<tr>
<td>2</td>
<td>0.423472</td>
<td>0.590753</td>
</tr>
<tr>
<td>3</td>
<td>0.4122024</td>
<td>0.598144</td>
</tr>
</tbody>
</table>
Conclusions and Discussion

1. Decay of the WER of non-binary LDPC codes: scaling law and scaling coefficient $\alpha$

2. Degradation of $\alpha$ with $m$ observed for all taken examples ($c > 2$)

3. More results on the second eigenvalue et related eigenvectors needed (Gerschgorin theory?)

4. Interesting case of 2-regular codes is not covered: under investigation