Low-density Parity-Check Codes over the Binary Erasure Channel

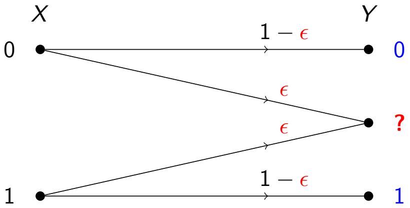
Pablo M. Olmos
Manuel A. Vázquez

March 6, 2024

Index

Index

The binary erasure channel (BEC)

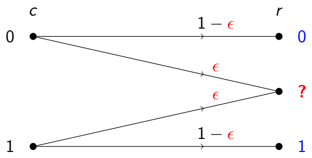


Capacity:
$C = 1 - \epsilon$

The model is very simple, but even so...

- quite surprisingly, **most properties and statements that we encounter in our investigation of LDPC codes over the BEC hold in much greater generality** (R. Urbanke and T. Richardson, Modern Coding Theory) and, moreover,
- erasure correcting codes are used in the link layer of some communications standards.

BEC: practical considerations



Uncoded transmission

Channel bit error probability $\equiv \epsilon$

Transmission of encoded bits

$$\underbrace{\mathbf{b}}_{k \text{ information bits}} \rightarrow \underbrace{\mathbf{c}}_{n \text{ encoded bits}} \rightarrow \underbrace{\mathbf{r}}_{n \text{ channel observations}} \rightarrow \underbrace{\hat{\mathbf{c}}}_{\text{decoded codeword}}$$

The rate of the code is still $R = \frac{k}{n}$

Channel coding theorem

We can attain a vanishing (codeword) error probability,

$$P(\hat{\mathbf{c}} \neq \mathbf{c}|\mathbf{r}) \to 0,$$

when $n \to \infty$ if the code rate is below the capacity, i.e.,

$$R < C.$$

You don't want that...

Using $n \to \infty$ is a waste of resources (time, energy)

Goal

...to design **feasible** encoding and decoding schemes that allow us to operate close to channel capacity.

Index

Linear block codes

- Generator matrix: $\mathbf{c} = \mathbf{b}\mathbf{G}$ where $\mathbf{b} \in \{0,1\}^k$.
- Parity check matrix: $\mathbf{c}\mathbf{H}^T = \mathbf{0} \quad \forall \mathbf{c} \in \mathcal{C}$.
  - $\mathcal{C}$ is the set of all codewords (*codebook*)
- Each row of the parity check matrix yields a linear constraint on the coded bits.

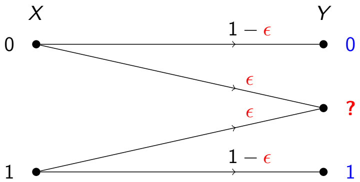
For a Hamming $(7,4)$ code

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Therefore...

$$c_1 \oplus c_3 \oplus c_5 \oplus c_7 = 0$$
$$c_2 \oplus c_3 \oplus c_6 \oplus c_7 = 0$$
$$c_4 \oplus c_5 \oplus c_6 \oplus c_7 = 0$$

Transmission over BEC



- Linear block code (n, k) with matrices $\mathbf{G}$ and $\mathbf{H}$.
- Codeword $\mathbf{c}$ is sent.
- Vector $\mathbf{r}$ is observed.

Some bits are *erased*, others are not:

- $\mathcal{E}$ is the set containing the indexes of the erased bits
- $\mathcal{R}$ is the set containing the indexes of the received bits.
- $\mathcal{E} \cup \mathcal{R} = \{1, \ldots, n\}$.

Thus, for the BEC

$$\mathbf{r}(\mathcal{E}) = ?, \qquad \mathbf{r}(\mathcal{R}) = \mathbf{c}(\mathcal{R})$$

Decoding over BEC: example

- Hamming $(7, 4)$ code.
- $\mathbf{c} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ is sent.
- $\mathbf{r} = \begin{bmatrix} 1 & ? & 1 & 0 & ? & ? & 0 \end{bmatrix}$ is received.
- $\mathcal{E} = \{2, 5, 6\}$ and $\mathcal{R} = \{1, 3, 4, 7\}$.

Thus, the system of equations can be simplified:

$$\left.\begin{aligned} c_1 \oplus c_3 \oplus c_5 \oplus c_7 &= 0 \\ c_2 \oplus c_3 \oplus c_6 \oplus c_7 &= 0 \\ c_4 \oplus c_5 \oplus c_6 \oplus c_7 &= 0 \end{aligned}\right\} \rightarrow \left.\begin{aligned} 1 \oplus 1 \oplus c_5 \oplus 0 &= 0 \\ c_2 \oplus 1 \oplus c_6 \oplus 0 &= 0 \\ 0 \oplus c_5 \oplus c_6 \oplus 0 &= 0 \end{aligned}\right\} \rightarrow \left.\begin{aligned} c_5 &= 0 \\ c_2 \oplus c_6 &= 1 \\ c_5 \oplus c_6 &= 0 \end{aligned}\right\}$$

By solving the system of binary equations we get a unique solution $\hat{\mathbf{c}} = [1110000] = \mathbf{c}$.

Decoding over BEC: general statement

- Linear block code (n, k) with matrices $\mathbf{G}$ and $\mathbf{H}$.
- Codeword $\mathbf{c}$ is sent.
- Vector $\mathbf{r}$ is observed.
- $\mathbf{H}_{\mathcal{E}}$ is the submatrix of $\mathbf{H}$ obtained by picking only those *columns* whose indexes are in $\mathcal{E}$ (and, analogously, $\mathbf{H}_{\mathcal{R}}$ is...).

Optimal maximum a posteriori decoding

Find $\mathbf{c}(\mathcal{E})$ by solving the following system of equations:

$$\mathbf{c}(\mathcal{E})\mathbf{H}_{\mathcal{E}}^T = \mathbf{c}(\mathcal{R})\mathbf{H}_{\mathcal{R}}^T$$

In the former example:

$$\begin{bmatrix} c_2 & c_5 & c_6 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^\top = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

System of linear equations for MAP decoding

$$
\begin{aligned}
\mathbf{c}\mathbf{H}^\top &= \begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_7 \end{bmatrix} = 0 \\
&= c_1\mathbf{h}_1 + c_2\mathbf{h}_2 + c_3\mathbf{h}_3 + c_4\mathbf{h}_4 + c_5\mathbf{h}_5 + c_6\mathbf{h}_6 + c_7\mathbf{h}_7 = 0 \\
&= c_2\mathbf{h}_2 + c_5\mathbf{h}_5 + c_6\mathbf{h}_6 + c_1\mathbf{h}_1 + c_3\mathbf{h}_3 + c_4\mathbf{h}_4 + c_7\mathbf{h}_7 = 0 \\
&= \begin{bmatrix} c_2 & c_5 & c_6 \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{h}_2 \\ \mathbf{h}_5 \\ \mathbf{h}_6 \end{bmatrix}}_{\mathbf{H}_{\mathcal{E}}{}^T} + \begin{bmatrix} c_1 & c_3 & c_4 & c_7 \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_3 \\ \mathbf{h}_4 \\ \mathbf{h}_7 \end{bmatrix}}_{\mathbf{H}_{\mathcal{R}}^T} = 0
\end{aligned}
$$

Hence,

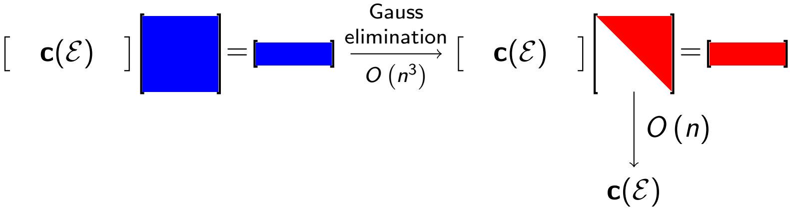
$$
\begin{bmatrix} c_2 & c_5 & c_6 \end{bmatrix} \mathbf{H}_{\mathcal{E}}{}^T = \begin{bmatrix} c_1 & c_3 & c_4 & c_7 \end{bmatrix} \mathbf{H}_{\mathcal{R}}^T
$$

$$
\mathbf{h}_j \equiv j\text{-th row of matrix } \mathbf{H}^\top = j\text{-th column of matrix } \mathbf{H}
$$

Optimal MAP decoding (classical approach)

When solving the system of linear equations, $\mathbf{c}(\mathcal{E})\mathbf{H}_{\mathcal{E}}^{T} = \mathbf{c}(\mathcal{R})\mathbf{H}_{\mathcal{R}}^{T}$ for $\mathbf{c}(\mathcal{E})$, there are two possible outcomes:

- the system has multiple solutions $\rightarrow$ all of them are equally likely, and we declare a decoding failure.
- the system has an unique solution $\rightarrow$ $\hat{\mathbf{c}} = \mathbf{c}$, and **no decoding error** is possible.



Computational complexity:

- Gaussian elimination requires $O\left(n^3\right)$ operations
- After Gaussian elimination, *backwards substitution* is $O\left(n\right)$

Index

Suboptimal decoding over the BEC: example I

Let us consider:

- Hamming code $(7,4)$
- $\mathbf{c} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ is sent.
- $\mathbf{r} = \begin{bmatrix} 1 & ? & 1 & 0 & ? & ? & 0 \end{bmatrix}$ is received.

Assuming the system is already triangularized and revealing as much information as possible...

$$c_5 = 0$$

$$c_5 + c_6 = 0 \rightarrow c_6 = 0$$

$$c_2 + c_6 = 1 \rightarrow c_2 = 1$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_2 \\ c_6 \\ c_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Decoding

Complexity is $O(n)$.

Suboptimal decoding over the BEC: example II

Another transmission:

- $(7,4)$ Hamming code
- $\mathbf{c} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ is sent.
- $\mathbf{r} = \begin{bmatrix} 0 & 1 & ? & 0 & 0 & ? & ? \end{bmatrix}$ is received.

Now,

$$\begin{aligned} c_3 \oplus c_7 &= 0 \\ c_3 \oplus c_6 \oplus c_7 &= 1 \\ c_6 \oplus c_7 &= 0 \end{aligned}$$

Decoding error

There are no equations with a single variable. No information can be revealed.

(if we were to use **optimal** decoding, c_3 is revealed ($c_3 = 1$) by adding the last two equations)

Classical v. modern coding theory

Classical

- **Optimal** decoding via ML/MAP rule with $O\left(n^3\right)$ operations. It restrains the coding schemes we can use in practice.
- **Small size ($\mathtt{n}$)** because otherwise decoding complexity becomes prohibitive. We **cannot operate very close to capacity** at vanishing error probability.
- Examples: Linear Block codes (BCH codes, Reed Solomon Codes), Convolutional codes...

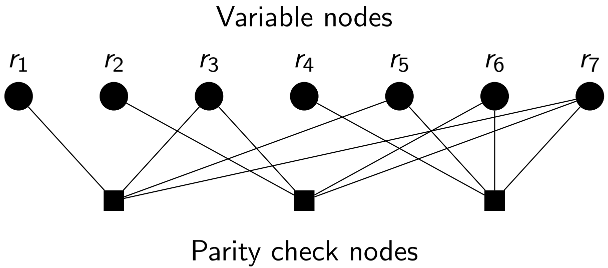
Modern

- **Approximate** decoding with worse performance for the sake of much less complexity ($O\left(n\right)$ operations).
- **Close to capacity** at vanishing error probability is achieved using **very long codes!** (large $\mathtt{n}$)
- Examples: Turbo Codes, LDPC codes, Polar Codes.

Tanner graph

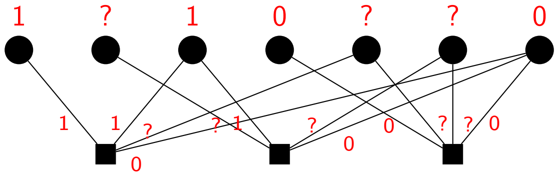
$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The constraints given by this matrix can be represented using a **Tanner graph**



Belief propagation

Initialization: variable nodes send the channel observation to the parity check nodes they are connected to:

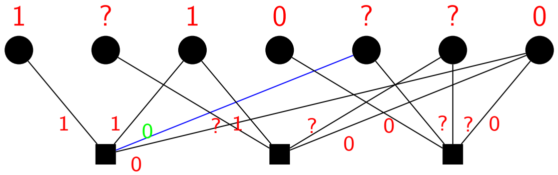


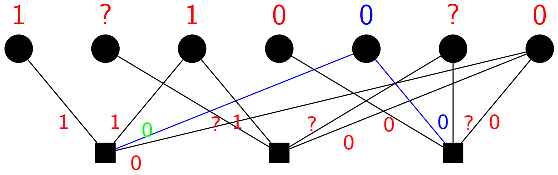
While there is any unsolved "?"

1. Using the received information, each *parity check node* tries to solve for the variable that sent a "?" message. If possible, they send the value obtained to the variable nodes. Otherwise they send a "?" message.
   - **Only parity-check nodes with a single unknown can solve a variable!**
2. *Variable nodes* send their new value to the parity check nodes...or they resend a "?" message.

Belief propagation

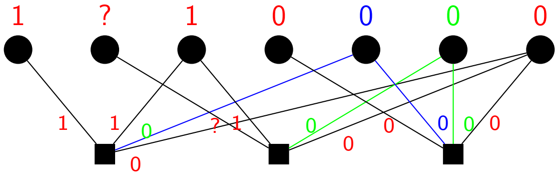
First iteration



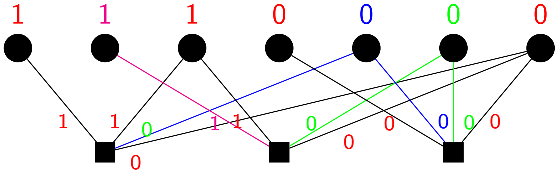


Belief propagation

Second iteration



Third iteration



Belief propagation

Some remarks:

- In general, the performance obtained with the suboptimal decoder is quite poor (lots of decoding errors).
- Given a parity check matrix **H** of dimensions $(n-k) \times n$, the number ones per row can be as high as n.
- If a row has αn ones, then the probability that $\alpha n - 1$ of the variables are correctly received and **only one** is unknown is

$$\alpha n \epsilon (1-\epsilon)^{(\alpha n - 1)}$$

which tends to 0 ($\Rightarrow$ decoding error) as $\mathtt{n} \to \infty$.

$(\alpha \in (0,1) \equiv \text{rate of 1s per bit})$

Index

Low-density Parity-Check codes

LDPC codes: linear block codes defined by sparse parity-check matrices.

$$\mathbf{H}_{(\mathtt{n}-k)\times\mathtt{n}}, \qquad \mathbf{c}\mathbf{H}^T = \mathbf{0} \;\; \forall \mathbf{c} \in \mathcal{C}$$

LDPC (3,6) with $\mathtt{n} = 20$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	1	0	0	0	1
2	0	0	0	0	0	0	1	1	0	0	1	1	0	1	0	1	0	0	0	0
3	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1
4	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	1	1	1	0
5	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
6	0	0	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	0	1
7	0	0	0	1	1	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0
8	1	0	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0
9	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
10	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	1	0

The density of ones in the matrix $\mathbf{H}$ is $6/\mathtt{n}$ and the rate is $R = 0.5$.

Imposing structure on H

If the number of ones per row is fixed to 6, e.g.,

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_{20} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{bmatrix} = 0 \Rightarrow c_5 + c_9 + c_{10} + c_{11} + c_{16} + c_{20} = 0$$

then the probability that each row in **H** yields a single unknown, e.g.,

$$c_5 + ? + c_{10} + c_{11} + c_{16} + c_{20} = 0$$

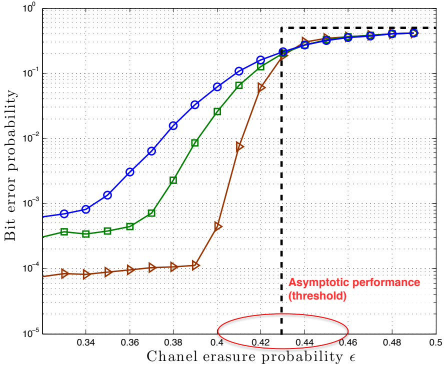
is

$$6\epsilon(1-\epsilon)^5,$$

which **does not depend on** n. An equation with a single unknown can be solved immediately...and once the variable is revealed, there is a **non-zero probability that a new row with a single unknown is created**. This probability does not depend on n either!!

BER over BEC

Bit-error rate of the $(3,6)$ ensemble over the BEC; $\mathtt{n} = 2^8$ ($\circ$), $\mathtt{n} = 2^9$ ($\square$), $\mathtt{n} = 2^{11}$ ($\triangleright$).



The threshold ε^* can be computed analytically. **It only depends on the connectivity pattern in matrix H!**.