General Properties of Optimal Decoding: EXIT function characterization of MAP decoding

Alexei Ashikhmin
Bell Labs

- ML and MAP decoding
- Mutual Information
- EXIT functions
- EXIT functions over Binary Erasure Channel
- EXIT functions over Gaussian Channel
- Bounds on EXIT functions

This talk is base on joint works with G. Kramer, S. Litsyn, S. ten Brink, E. Sharon

ML and MAP decodings

C is a binary code of the length n
We transmit $\underline{c} \in C$ and receive $\underline{y} = \underline{c} + \text{noise}$

Decoding for minimization of the word error rate (ML Decoding):

Find $\underline{c}' \in C$ such that $\Pr(\underline{c}'|\underline{y}) = \max_{\underline{c} \in C} \Pr(\underline{c}|\underline{y})$

Bitwise decoding (MAP or APP decoding):

For each code coordinate j we compute

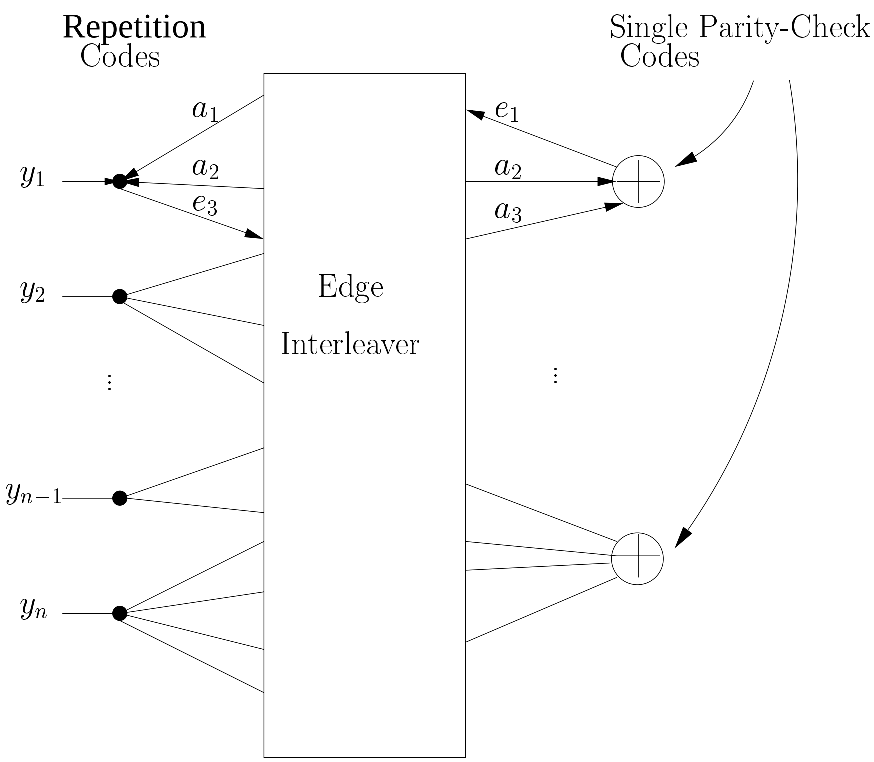
$$\ln \frac{\Pr(c_j = 0|\underline{y})}{\Pr(c_j = 1|\underline{y})}$$

Let $\underline{y}_{[j]} = (y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_n)$. Then

$$\ln \frac{\Pr(c_j = 0|\underline{y})}{\Pr(c_j = 1|\underline{y})} = \ln \frac{\Pr(c_j = 0|\underline{y}_{[j]})}{\Pr(c_j = 1|\underline{y}_{[j]})} + \ln \frac{\Pr(c_j = 0|y_j)}{\Pr(c_j = 1|y_j)}$$

MAP decoders are used as constituent decoders in iterative codes (TURBO, LDPC and so on codes)

LDPC codes:



Mutual Information

- X and Y are random variables

- Mutual Information between X and Y is defined as

$$I(X;Y) = \sum_{x,y} \Pr(x) \Pr(y|x) \log \frac{\Pr(x|y)}{\Pr(x)}$$

- If X and Y are independent then

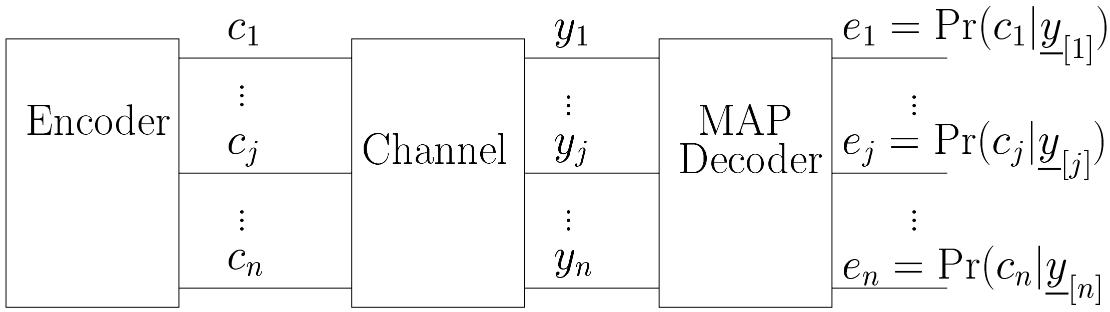
$$I(X;Y) = 0$$

- If Y is a function of X, and

$$\Pr(x=0) = \Pr(x=1) = 1/2$$

  then $I(X;Y) = 1$

Extrinsic Information Transfer (EXIT) function



The average input (apriori) information

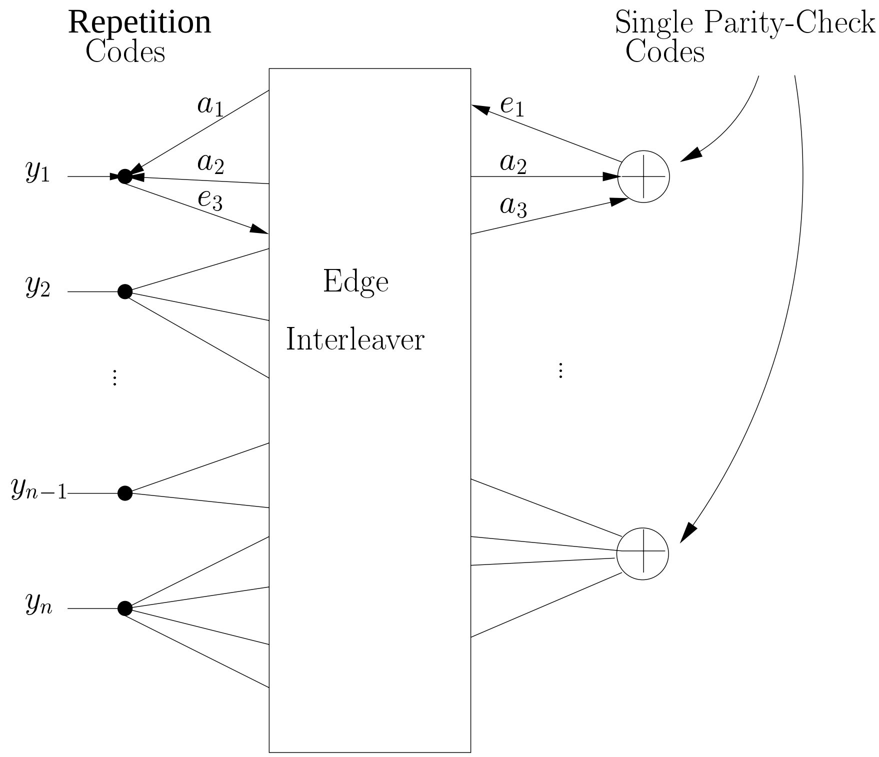
$$I_{in} = \frac{1}{n} \sum_j I(C_j; Y_j)$$

The average output (aposteriori) information

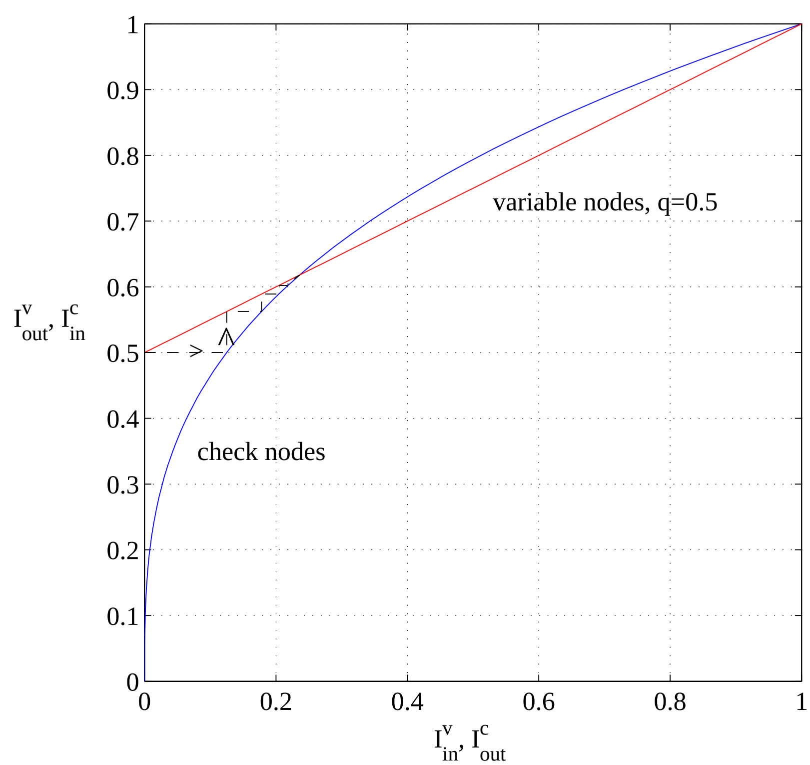
$$I_{out} = \frac{1}{n} \sum_j I(C_j; E_j)$$

$I_{out}(I_{in})$ is called the EXIT function of the code over a given channel

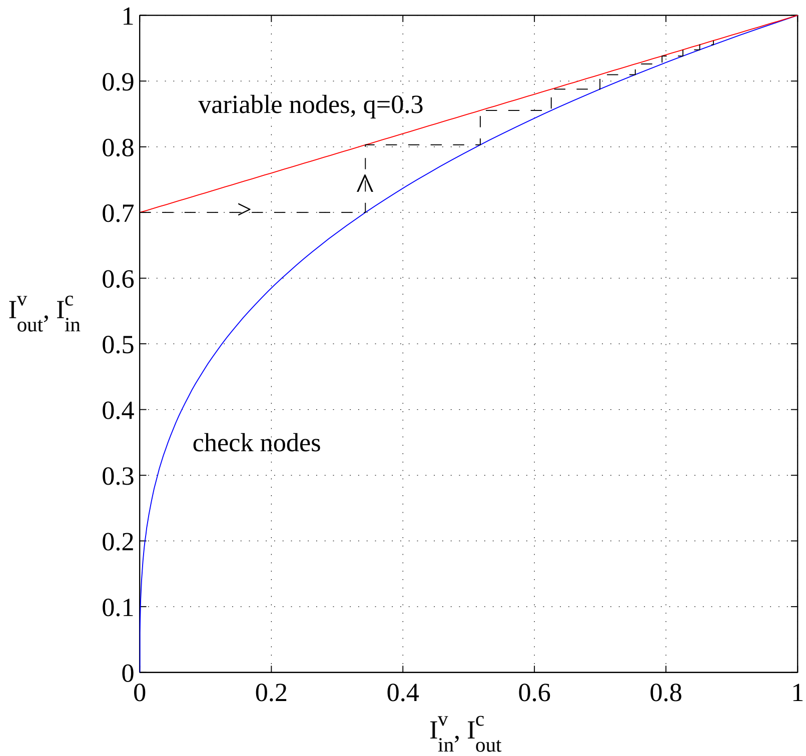
Stephan ten Brink intoroduced this notion; also suggested tracking evolution of mutual information during iterative decoding



Repetition Codes
Single Parity-Check Codes
a_1
a_2
e_3
e_1
a_2
a_3
y_1
y_2
y_{n-1}
y_n
Edge Interleaver

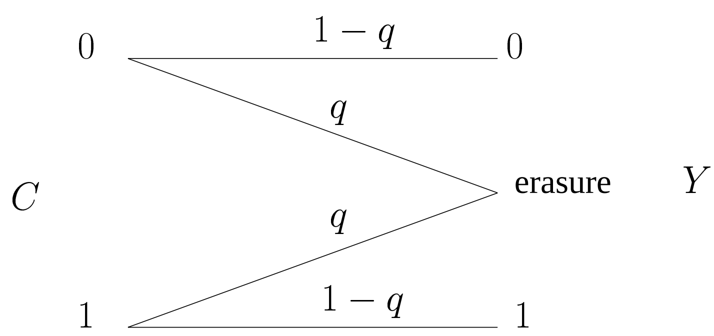


variable nodes, q=0.5
check nodes
I^v_{out}, I^c_{in}
I^v_{in}, I^c_{out}
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

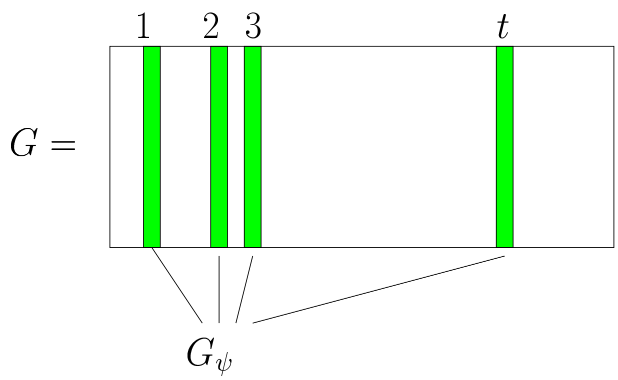


variable nodes, q=0.3
check nodes
I^v_{out}, I^c_{in}
I^v_{in}, I^c_{out}
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

Binary Erasure Channel



$$I(C;Y) = 1 - q$$



G is a generator matrix of a code C
G_ψ is a t columns submatrix of G
The function

$$r_t(C) = \sum_{\substack{\psi \in \{1,\ldots,n\} \\ |\psi| = t}} \text{rank}(G_\psi), \;\; t = 1, \ldots, n$$

is called the information function of C
(Helleseth, Kløve, Levenshtein, 1997)

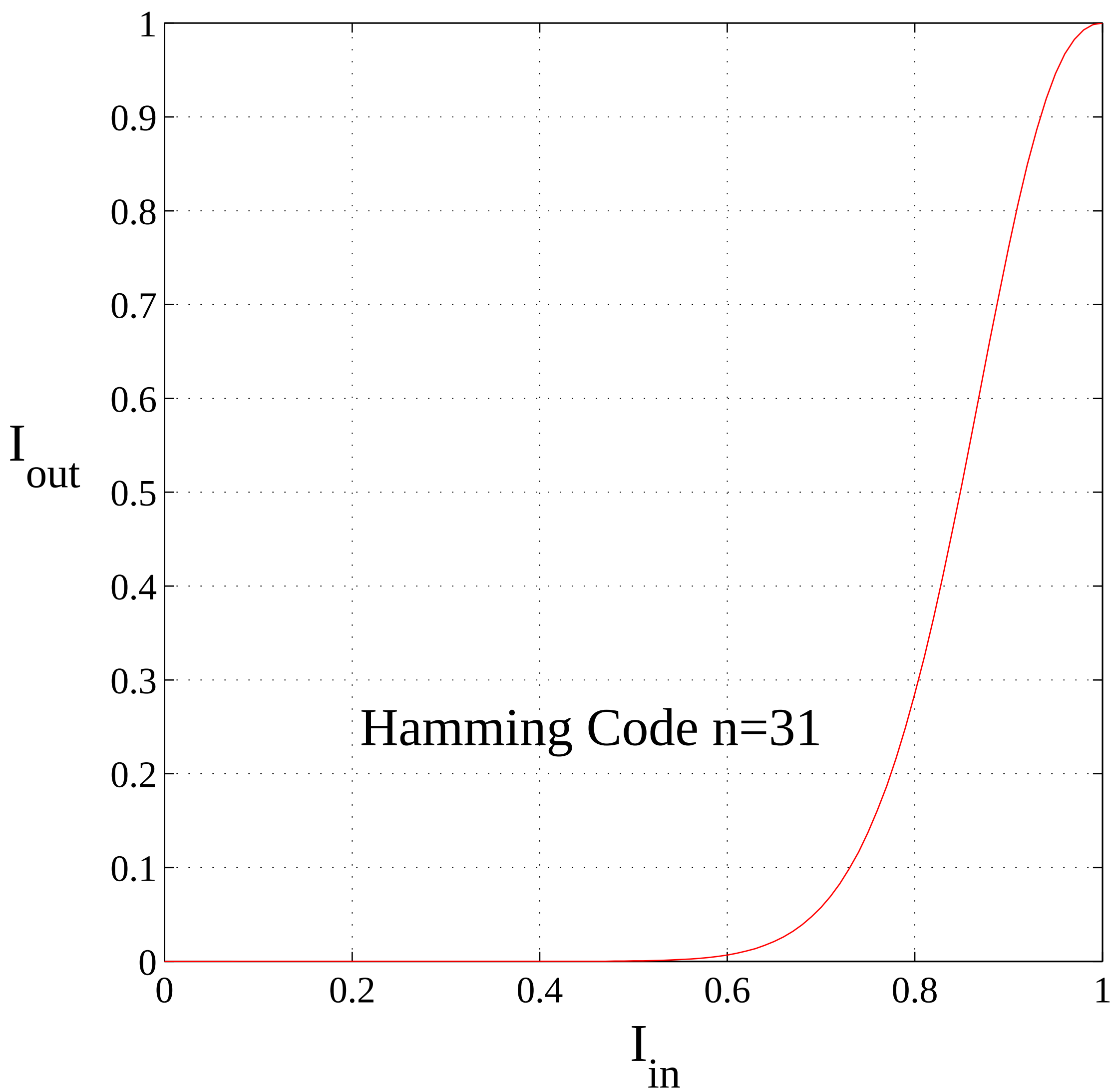
Theorem 1

$$I_{out}^{BEC}(I_{in})$$
$$= 1 - \frac{1}{n}\sum_{t=1}^{n} I_{in}^{i-1}(1 - I_{in})^{n-i}[i \cdot r_t(C) - (n - i + 1) \cdot r_t(C)].$$

Numbers r_t can be found with the help of generalized Hamming weight enumerator A_i^r:

$$r_t = \text{complicated function of}(A_i^r)$$

For example we know A_i^r for the Hamming code. Therefore we can compute $I_{out}(I_{in})$



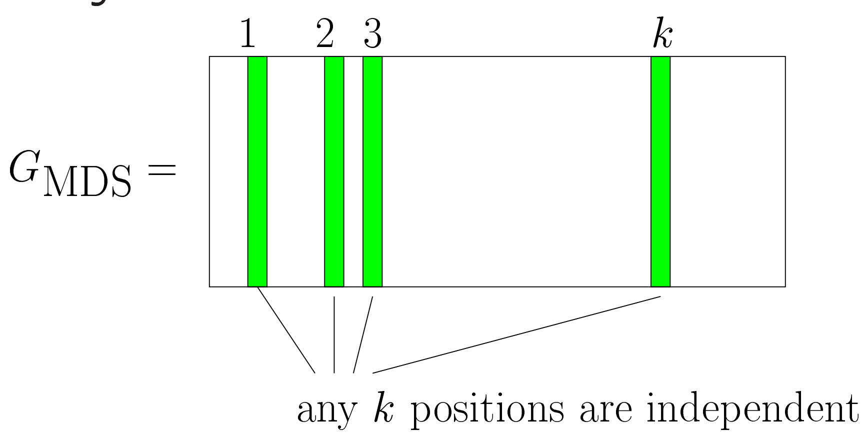
Hamming Code n=31
I_out
I_in
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

“Advantages” of MAP decoding over ML decoding

In the case of ML decoding there exist optimal codes:

Binary Symmetric Channel - Hamming codes are optimal

Binary Erasure Channel - MDS codes are optimal

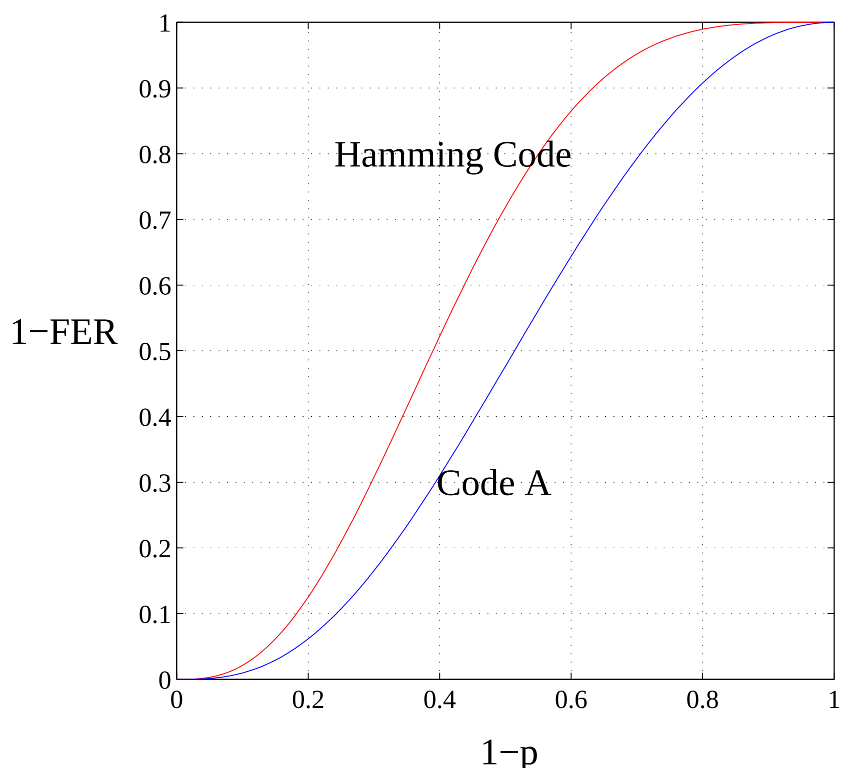


Simplex code

$$G = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Code A

$$G = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$



FER=frame (code word) error rate

p is the channel erasure probability

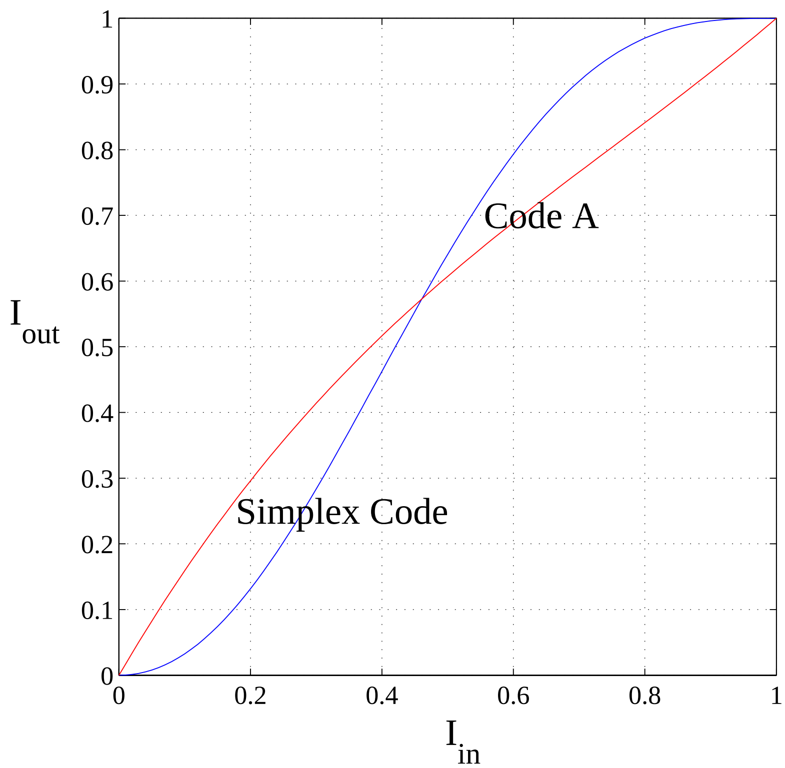
How about MAP decoding?

Area Property

Theorem 2 *In Binary Erasure Channel for a code (linear or nonlinear) of rate R we have*

$$\int_0^1 I_{out}(I_{in})dI_{in} = 1 - R$$

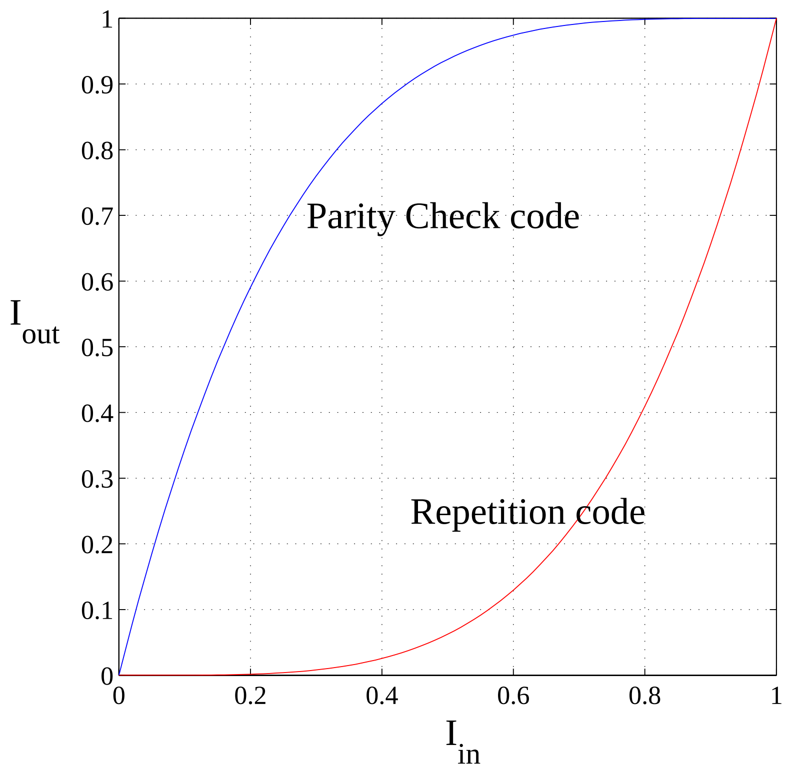
It means there is no discrimination; every code is the best one somewhere



Duality Property

Theorem 3 *In BEC the EXIT functions of a code and its dual code are connected:*

$$I_{out}(I_{in}) = 1 - I_{out}^{\perp}(1 - I_{in})$$



Gaussian and Other Channels

We transmit c_j and receive $y_j = c_j + z_j$

Def. 1 *A channel is called symmetric if*

$$p_{Y|C}(-y|-c) = p_{Y|C}(y|c).$$

Def. 2 *A density f is called consistent if it satisfies*

$$f(x) = f(-x)e^x. \tag{1}$$

Theorem 4 *If inputs of a MAP decoder of a linear (distance invariant) code are symmetric and consistent then its outputs are also symmetric and consistent.*

Repetition code of length n over AWGN channel $N(0, \sigma)$:

$$\begin{aligned} I_{in} &= \frac{\sigma}{\sqrt{8\pi}} \int e^{-(x-2/\sigma^2)\sigma^2/2} \log_2(1+e^{-x})dx \\ I_{out} &= \frac{n\sigma}{\sqrt{8\pi}} \int e^{-(x-2/n^2\sigma^2)n^2\sigma^2/2} \log_2(1+e^{-x})dx \end{aligned}$$

Theorem 5 *Single Parity Check Code of length n over symmetric and consistent channel has the following EXIT function*

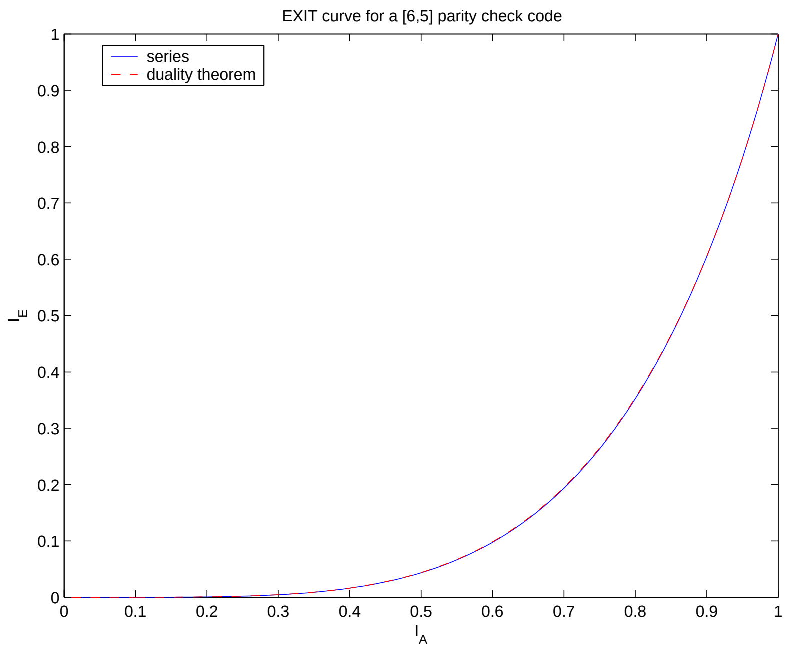
$$I_{in} = \frac{1}{ln2}\sum_{i=1}^{\infty}\frac{1}{(2i)(2i-1)}[E(T^{2i})]$$
$$I_{out} = \frac{1}{ln2}\sum_{i=1}^{\infty}\frac{1}{(2i)(2i-1)}[E(T^{2i})]^{n-1},$$

where

$$T = \Pr(c_j = 1|y_j) - \Pr(c_j = -1|y_j).$$

Duality for AWGN?

$$I^{AWGN}_{out,\text{parity check}}(I_{in}) = 1 - I^{AWGN}_{out,\text{repetition}}(1 - I_{in})$$



I_{in}	0.02	0.08	0.2	0.4	0.6	0.8	0.9	0.96
I_E exact	$1.105e-8$	$9.495e-6$	0.00070	0.0161	0.0975	0.3525	0.6046	0.8196
I_E from duality	0	$1.124e-5$	0.00073	0.0163	0.0984	0.3541	0.6052	0.8194

Area property for AWGN?

The area theorem also does not hold in AWGN channel, but at least for repetition and single parity check codes

$$\int_0^1 I_{out}(I_{in})dI_{in}$$

is very close to $1 - R$

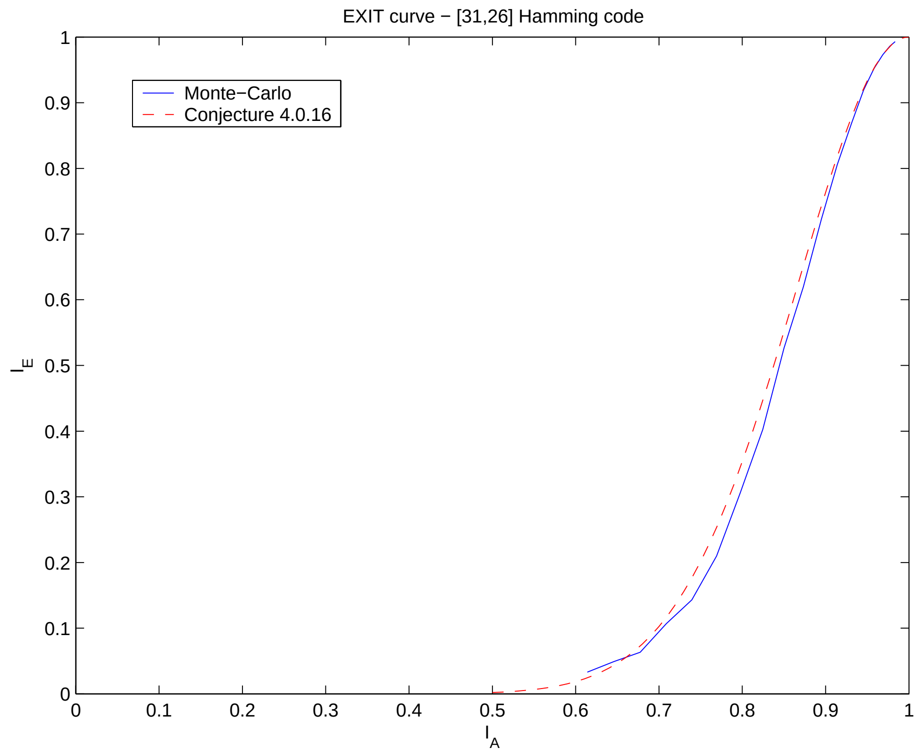
EXIT functions BEC ⇔ AWGN

Theorem 6 *For single parity check codes we have*

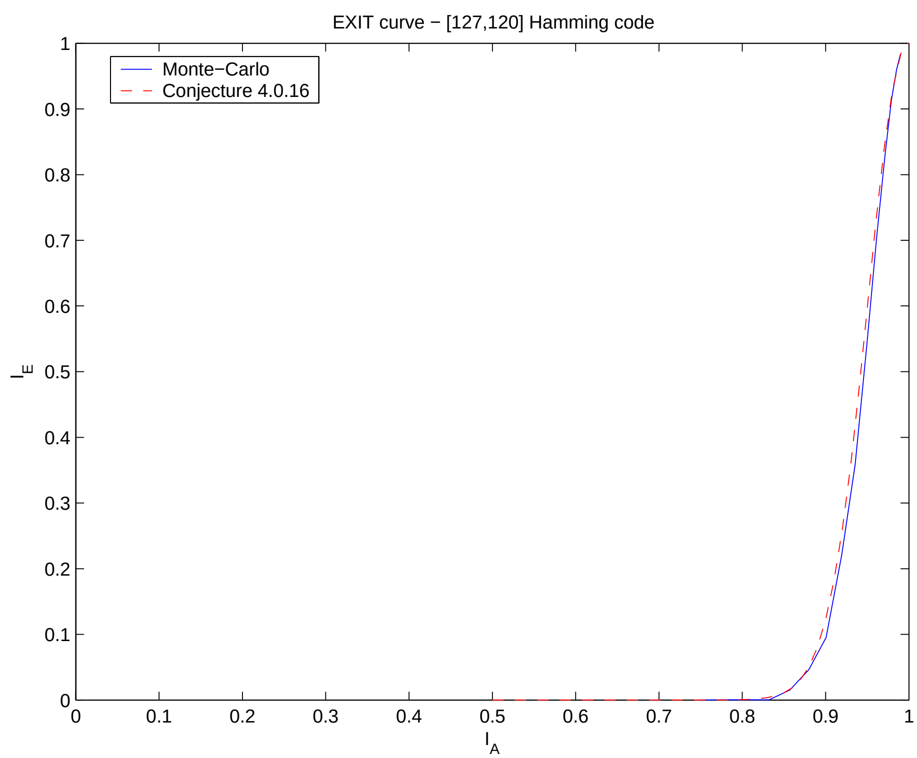
$$I_{out}^{AWGN}\left(\frac{E_b}{N_0}\right) = \sum_{i=1}^{\infty} \frac{1}{ln(2)(2i-1)(2i)} I_{out}^{BEC}(\epsilon_i),$$

where

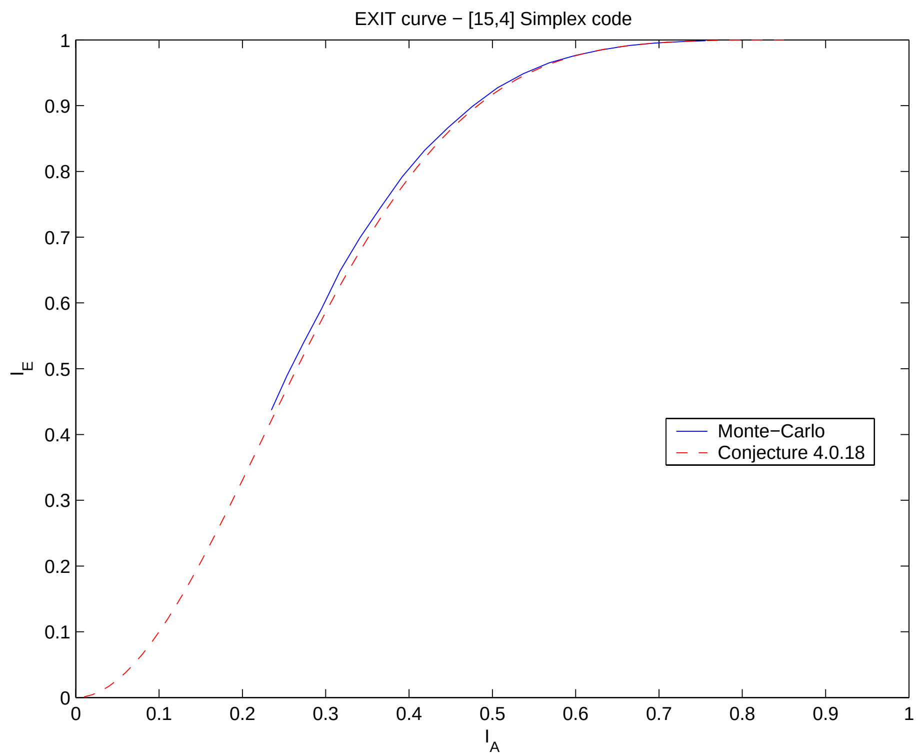
$$\epsilon_i = 1 - \int_{-1}^{+1} \frac{2t^{2i}}{(1-t^2)\sqrt{16\pi \frac{E_b}{N_0} R}} e^{-\frac{\left(ln\frac{1+t}{1-t} - 4\frac{E_b}{N_0}R\right)^2}{16\frac{E_b}{N_0}R}} dt$$

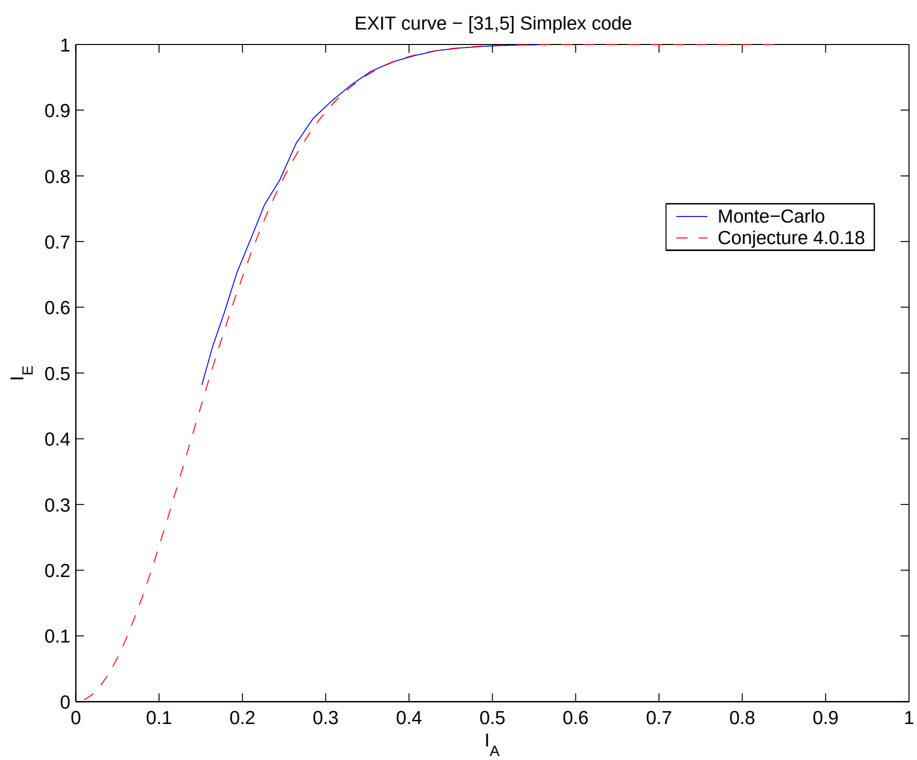


EXIT curve − [31,26] Hamming code
Monte−Carlo
Conjecture 4.0.16
I_E
I_A
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

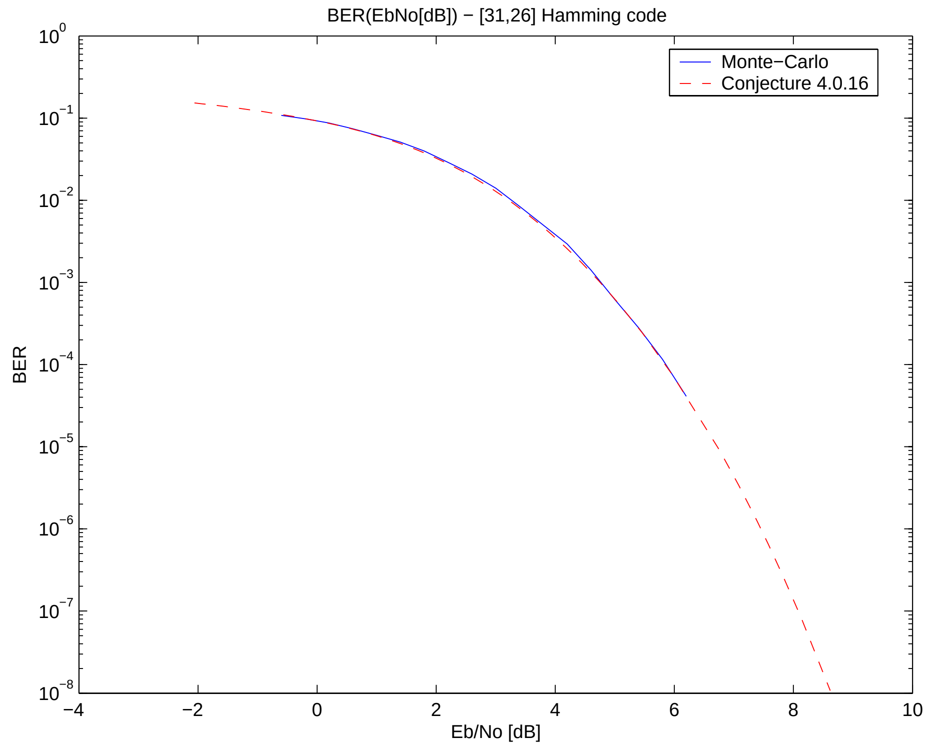


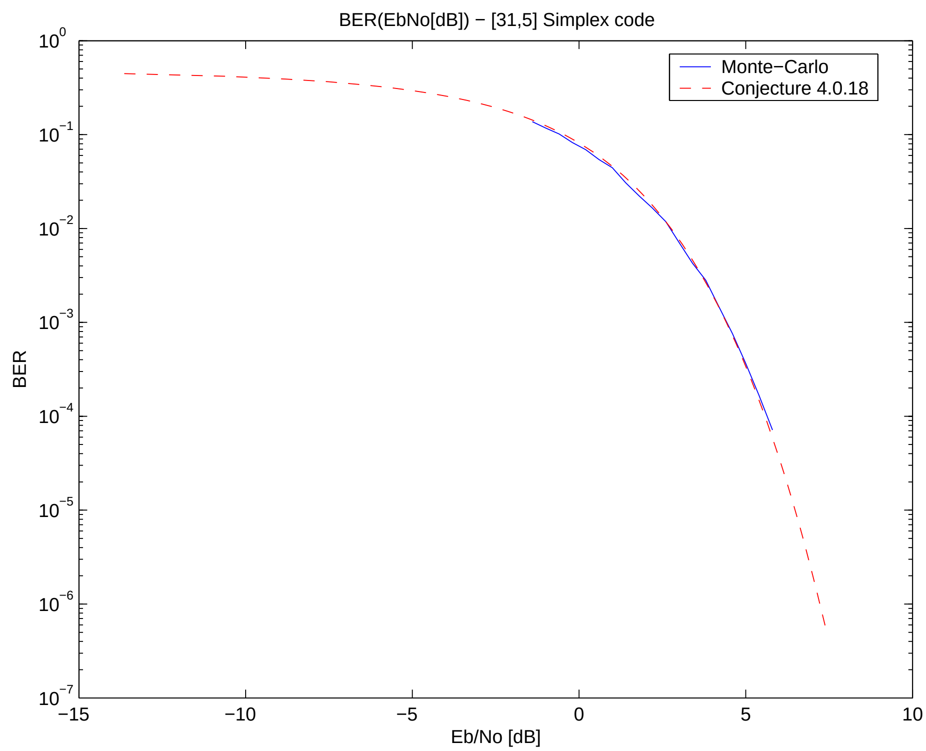
EXIT curve − [127,120] Hamming code
Monte−Carlo
Conjecture 4.0.16
I_E
I_A
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


EXIT curve − [15,4] Simplex code
Monte−Carlo
Conjecture 4.0.18
I_E
I_A
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

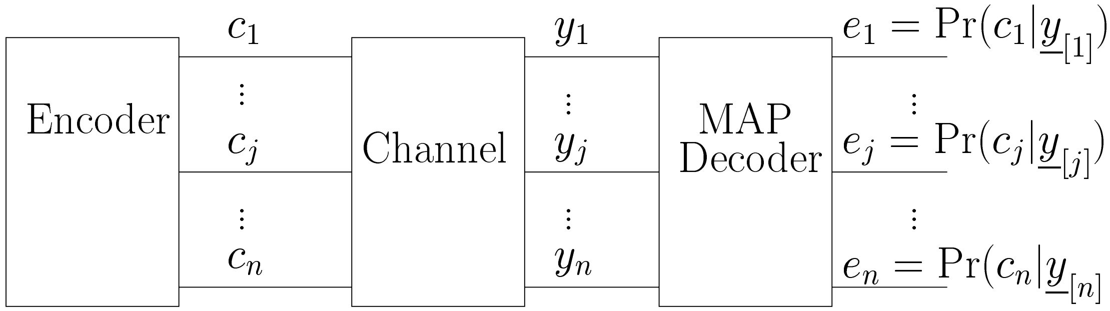

EXIT curve − [31,5] Simplex code
Monte−Carlo
Conjecture 4.0.18
I_E
I_A
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Probability of MAP Decoding Error





Bounds on EXIT functions



$$I_{in} = \frac{1}{n}\sum_j I(C_j; Y_j)$$

$$I_{out} = \frac{1}{n}\sum_j I(C_j; E_j)$$

We would like to find a channel that maximizes (minimizes) I_{out} for given I_{in}

I. Land, S. Huettinger, P. Hoeher, J. Huber; and I. Sutskover, S. Shamai proved recently the following theorems:

Theorem 7 *The Binary Symmetric Channel minimizes and Binary Erasure Channel maximizes I_{out} for given I_{in} in the case of repetition code.*

Theorem 8 *The Binary Symmetric Channel maximizes and Binary Erasure Channel minimizes I_{out} for given I_{in} in the case of single parity check code.*