REED MULLER CODES ACHIEVE CAPACITY ON ERASURE CHANNELS

S.Kudekar,S.Kumar,M.Mondelli,H.D.Pfister,E.Sasoglu,R.Urbanke

E2:207:Concentration Inequalities
Course Project Report
by

Ramakrishnan
Soumya Subhra Banerjee
ECE,IISc

December 2,2017

Motivation

The channel capacity is the largest rate at which communication ***can be*** made with arbitrary small error.

Channel coding tells us **how** capacity can be achieved.
Turbo , SC-LDPC achieve capacity, but no proof is presented yet
Polar and Reed-Muller were proved to achieve Capacity…

- Another* answer to the search for PROVABLY capacity achieving channel codes.
- Perhaps one of the most important applications of ***threshold phenomenon.***
- Establishes that a certain class of codes achieve capacity, Which include Reed-Muller codes,Polar codes.

**That polar codes achieve capacity was proved by E.Arikan before this.*
This provides an alternate method of the same proof leveraging fewer properties of polar codes

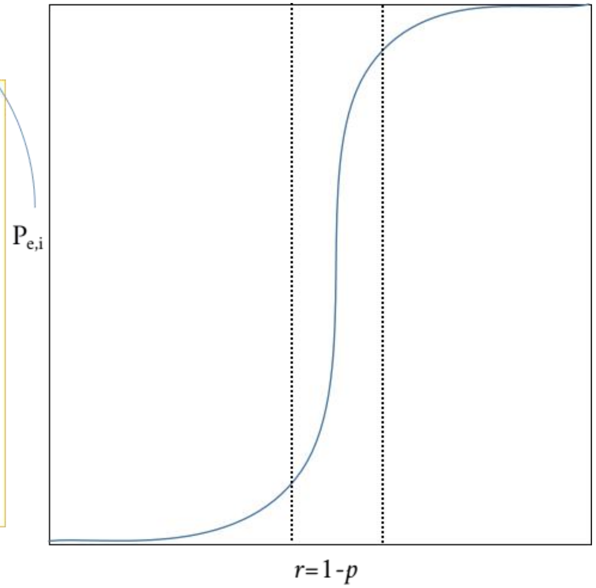
Channel Capacity from the perspective of Threshold Phenomenon...*proof idea*

For transmission over BEC(p), with Capacity r=1-p
For a sequence of binary linear codes with rate r to be capacity achieving, **the bit error probability,must converge to 0 for any erasure rate below 1 − r.**

if ,
1.Pe,i=h(p)=$P_p(\boldsymbol{\Omega}\mathbf{i})$

2.*where* $\boldsymbol{\Omega}$*i*
is a symmetric monotone , we can apply Russo's lemma.

3.*If the critical value is p=1-r, the proof is complete*



To prove this using Threshold Phenomenon.

1. *Bit Error Probability for a code , under MAP decoding needs to be captured by a function of erasure probability.We shall explore **MAP EXIT (Extrinsic Information Transfer) functions** for this.*
2. *MAP EXIT function should be expressed as measure of a **monotone set***
3. *The influences of the set must be equal , in other words ,the set must be **symmetric**.This happens when we assume the **code is 2-transitive** under automorphic permutation.*
4. *Hence we will apply Russo's lemma.*
5. *The critical value will occur at p=1-r, this follows, from **Area theorem***
6. *RM codes , Polar codes are 2-transitive.*

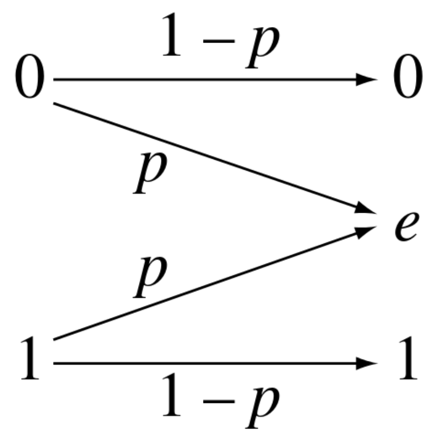
Bit Map Decoding under BEC

Decoding rule

$$D_i : \mathcal{Y}^N \to \mathcal{X} \cup \{e\}$$
$$D_i(\underline{Y}) = X_i.$$

Bit error Probability

$$P_{b,i} \triangleq \mathbb{P}[D_i(\underline{Y}) \neq X_i].$$
$$P_b \triangleq \frac{1}{N} \sum_{i=1}^{N} P_{b,i}.$$



$\mathbb{P}(Y_i = e) = p$

***Observation* 1, *Bit error probability = Transinformation*:**

$$P_{b,i} = H(X_i|\underline{Y}) \qquad \text{,if } X_i \text{ is } Ber(1/2).$$

***Proof*:**

if X_i can be recovered from $\underline{Y} = \underline{y}$, $H(X_i|\underline{Y} = \underline{y}) = 0$.

Else , $\mathbb{P}(X_i = x|\underline{Y} = \underline{y}) = 1/2$ and $H(X_i|\underline{Y} = \underline{y}) = 1$.

MAP EXIT functions...*to capture bit error probability*

Definition:

the vector EXIT function associated with bit i,

$$h_i(\underline{p}) \triangleq H(X_i|\underline{Y}_{\sim i}(\underline{p}_{\sim i})).$$

the average EXIT function,

$$h(\underline{p}) \triangleq \frac{1}{N}\sum_{i=1}^{N} h_i(\underline{p})$$

we can choose p_i=p for all i.

Implication:
It is the Transinformation associated with each bit 'i',which comes from all bits of received vector except bit 'i' (hence the name EXtrinsic Information Transfer)

***Observation* 2,*Bit error probability vs EXIT function*:**
By definition of conditional Entropy...

$$\begin{aligned} H(X_i|\underline{Y}) &= \mathbb{P}(Y_i = e)H(X_i|\underline{Y}_{\sim i}, Y_i = e) + \mathbb{P}(X_i = Y_i)H(X_i|\underline{Y}_{\sim i}, Y_i = X_i) \\ &= \mathbb{P}(Y_i = e)H(X_i|\underline{Y}_{\sim i}). \end{aligned}$$

Further using Observation 1,

$$P_{b,i}(p) = p h_i(p)$$

MAP EXIT function *...area theorem*

***Propostion* 1*.**

The MAP EXIT function satisfies ...

$$h_i(p) = \frac{\partial H(\underline{X}/\underline{Y}(\underline{p}))}{\partial p_i}$$

Proof:

$$\begin{aligned} H(\underline{X}/\underline{Y}(\underline{p})) &= H(X_i/\underline{Y}(\underline{p})) + H(X_{\sim i}/X_i, \underline{Y}(\underline{p}) \\ &= H(X_i/\underline{Y}(\underline{p})) + H(X_{\sim i}/X_i, Y_{\sim i}) \text{ ,by memorylessness} \\ &= p_i h_i(p) + H(X_{\sim i}/X_i, Y_{\sim i}) \end{aligned}$$

noting that the second term is independent of pi, and differentiating gives the above result.

** Note the propositions are numbered according to original paper*

MAP EXIT function *...area theorem*

Proposition 4(c), Area Theorem:

The area under average EXIT function is equal to the rate

$$\int_0^1 h(p)\,dp = \frac{K}{N}.$$

Implication:

This theorem will govern where the threshold will lie.

Proof:

from proposition 1, $h_i(p) = \frac{\partial H(\underline{X}/\underline{Y}(\underline{p}))}{\partial p_i}$

this implies, for a parameterized vector, $\underline{p}(t) = (p_1(t), \ldots, p_n(t))$, $t \in (0,1)$

$$H(\underline{X}|\underline{Y}(\underline{p}(1))) - H(\underline{X}|\underline{Y}(\underline{p}(0))) = \int_0^1 \left(\sum_{i=1}^{N} h_i(\underline{p}(t))p_i'(t) \right) dt.$$

consider $p_i(t) = t$

noting, $H(\underline{X}|\underline{Y}(\underline{1})) = H(\underline{X}) = K$, $H(\underline{X}|\underline{Y}(\underline{0})) = 0.$,*and using the definition of* $h(p)$ *completes the proof.*

MAP EXIT function *...as a measure of a set*

Consider **bit erasure pattern *a*** as a indicator vector s.t. $\boldsymbol{a}_i = \mathbf{1}_{\{Y_i = e\}}$

A binary vector $\boldsymbol{a}$ *covers another vector* $\boldsymbol{c}$ *if* $\boldsymbol{a}_i > \boldsymbol{c}_{i},$ *for all i.*

- *A transmitted codeword can be recovered iff the erasure pattern does not cover any codeword.*
- *A bit* ***i*** *can be recovered uniquely iff the error pattern does not cover any codeword where bit* ***i*** *is non-zero.*

	1	2	3	4	5	6	E *does not cover*
E	1	0	0	*1*	0	0	
C1	1	0	1	*1*	1	0	True
C2	1	1	0	*1*	0	1	True
C3	1	0	0	*1*	0	0	True
C4	1	0	0	*0*	0	0	X

MAP EXIT function *...as a measure of a set*

Indirect Recovery: *Consider a code C and the indirect recovery of X_i from the subvector $\underline{Y}_{\sim i}$ (i.e., the bit-MAP decoding of Y_i from $\underline{Y}$ when $Y_i = e$).*

For $i \in [N]$, the ***set of erasure patterns that prevent indirect recovery of X_i*** *under bit-MAP decoding is given by,*

$$\Omega_i \triangleq \{A \subseteq [N] \backslash \{i\} : \exists B \subseteq [N] \backslash \{i\}, B \cup \{i\} \in \mathcal{C}, B \subseteq A\}.$$

Proposition ***3(a),4(a)****, $h_i(p)$ is measure of $\boldsymbol{\Omega_i}$* **:**

$$h_i(p) = \mu_p(\Omega_i) = \sum_{A \in \Omega_i} p^{|A|}(1-p)^{N-1-|A|}.$$

Intuitively, it is the set of all erasure patterns, that cover some codeword with bit 'i' = 1, if 'i' is included.hence strengthens $P_{b,i}(p) = ph_i(p)$

MAP EXIT function *...as a measure of a set*

Proof:

by definition of conditional entropy,

$$\begin{aligned} h_i(\underline{p}) &= H(X_i|\underline{Y}_{\sim i}(p_{\sim i})) \\ &= \sum_{\underline{y}_{\sim i} \in \mathcal{Y}^{N-1}} \Pr(\underline{Y}_{\sim i} = \underline{y}_{\sim i}) H\left(X_i|\underline{Y}_{\sim i} = \underline{y}_{\sim i}\right). \end{aligned}$$

WOLOG assume all 0's *were sent.due to error pattern A , few bits got erased.*

thus,

$$\Pr(\underline{Y}_{\sim i} = \underline{y}_{\sim i}) = \prod_{\ell \in A} p_\ell \prod_{\ell \in A^c \setminus \{i\}} (1 - p_\ell).$$

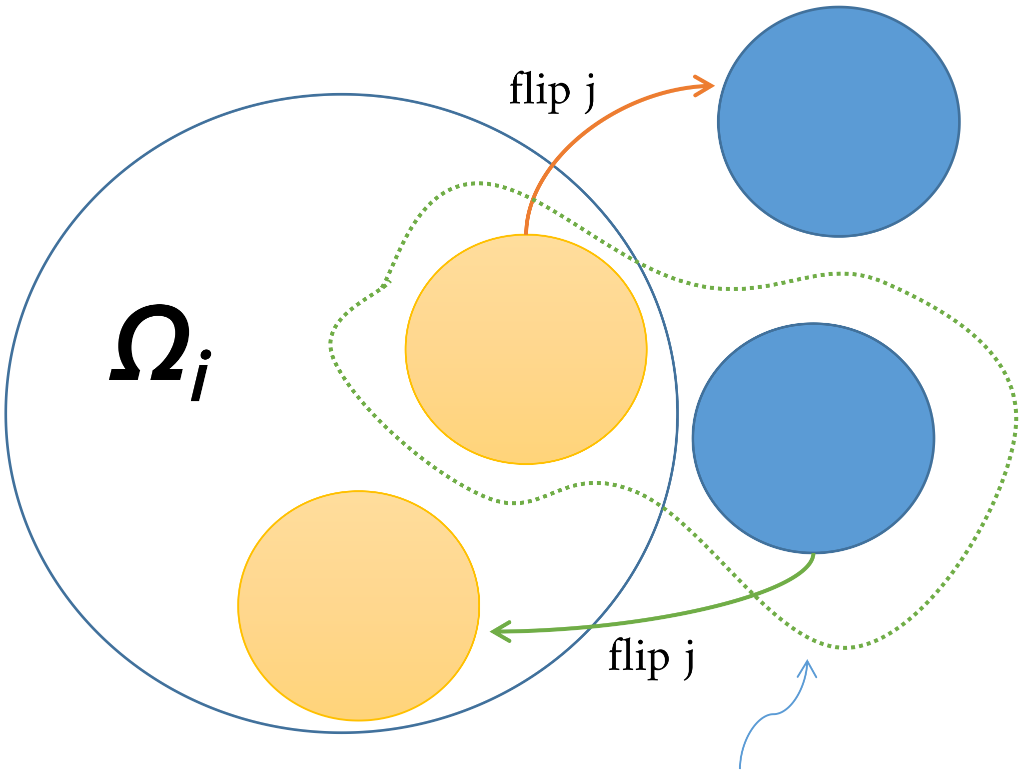
Now , if A $\in$ ***Ω**$_i$* ***, then** decoding fails hence ,* $H(X_i|\underline{Y}_{\sim i} = \underline{y}_{\sim i}) = 1.$ *,else* 0.

thus,

$$h_i(\underline{p}) = \sum_{A \in \Omega_i} \prod_{\ell \in A} p_\ell \prod_{\ell \in A^c \setminus \{i\}} (1 - p_\ell).$$

the result follows on considering, p_l=*p.*

MAP EXIT function *...Influences of $\boldsymbol{\Omega_i}$*



The measure of this set is the j-th influence of $\boldsymbol{\Omega_i}$

$$\partial_j \Omega_i \triangleq \{A \subseteq [N] \backslash \{i\} \mid A \backslash \{j\} \notin \Omega_i, A \cup \{j\} \in \Omega_i\}.$$

MAP EXIT function *…Influences of* $\boldsymbol{\Omega_i}$

Proposition *3(b),4(b), jth partial derivative of $h_i(p)$ is jth influence of* $\boldsymbol{\Omega_i}$ **:**

$$\frac{\partial^2 H(\underline{X}|\underline{Y}(\underline{p}))}{\partial p_j \partial p_i} = \frac{\partial h_i(\underline{p})}{\partial p_j} = \sum_{A \in \partial_j \Omega_i} \prod_{\ell \in A} p_\ell \prod_{\ell \in A^c \setminus \{i\}} (1 - p_\ell).$$

$$\left.\frac{\partial h_i(\underline{p})}{\partial p_j}\right|_{\underline{p}=(p,p,\ldots,p)} = \mu_p(\partial_j \Omega_i) = \sum_{A \in \partial_j \Omega_i} p^{|A|}(1-p)^{N-1-|A|}.$$

Proof: similar to proposition 3(a),4(a)

We need to show next :

- *all influences are equal, hence the set is symmetric …*
- *all hi(p) are equal , hence area theorem can be applied…*

these will stem from the assumption that the code is 2-transitive

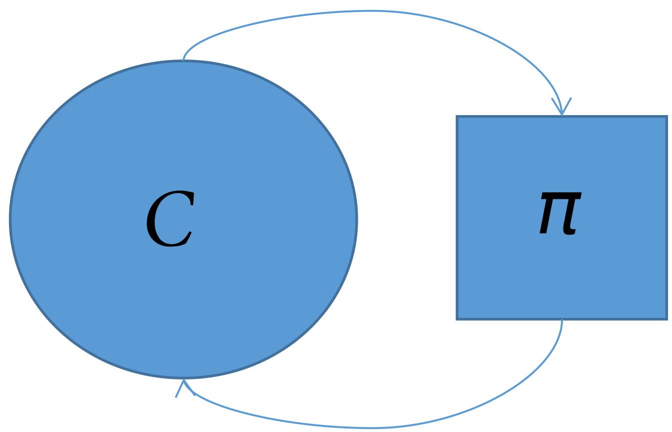
Permutation group...*briefing*

- ***Symmetric group S_N of $[N]$*:**

Is the set of all permutations of N possible.

$[N=3]$, SN={{1,2,3},{1,3,2},{2,1,3}...}

- ***Permutation group*** *is a subset of symmetric group.*



Considering $[N]$ denotes the set of indices of a code a permutation operation (π) is a rearrangement of the bits of a code.

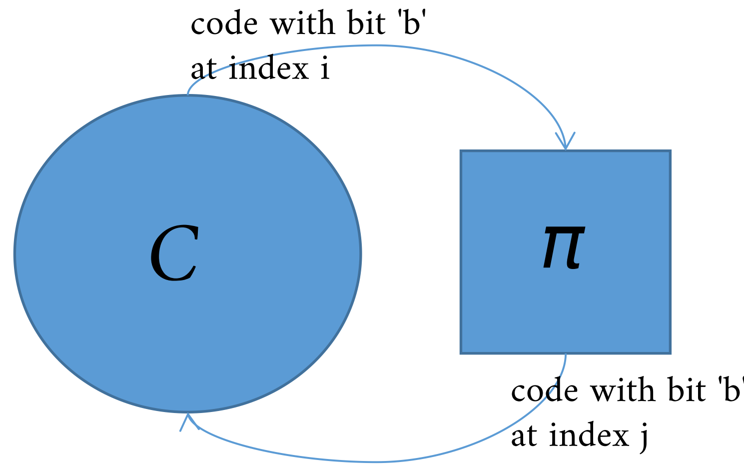
- *A **Permutation automorphic group $P_{Aut}(C)$** of a code is the permutation group , suchthat under $\pi \in P_{Aut}(C)$ the membership of the code is preserved, we denote $P_{Aut}(C)$, as*

$$\mathcal{G} = \{\pi \in S_N : \pi(A) \in \mathcal{C} \text{ for all } A \in \mathcal{C}\}.$$

Permutation group...*transitivity*

$\mathcal{G} = \{\pi \in S_N : \pi(A) \in \mathcal{C} \text{ for all } A \in \mathcal{C}\}.$

if for any $i, j \in [N]$, *there exists.*
$\pi \in \mathcal{G}$ *, such that* $\pi(i) = j$,
then $\mathcal{G}$ *is **transitive.***

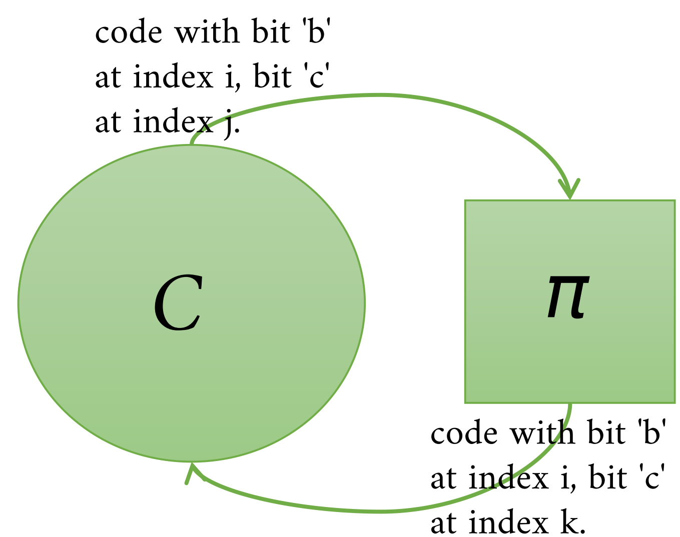


$\mathcal{G} = \{\pi \in S_N : \pi(A) \in \mathcal{C} \text{ for all } A \in \mathcal{C}\}.$

if for any $i, j, k \in [N]$, *there exists,*
$\pi \in \mathcal{G}$ *, such that,*

$$\pi(i) = i \text{ and } \pi(j) = k.$$

then $\mathcal{G}$ *is **doubly-transitive.***



MAP EXIT function *...for doubly transitive code*

Proposition *7, All EXIT functions are equal **,** for a transitive code***:**

$$h(p) = h_i(p) \text{ for } 0 \leqslant p \leqslant 1.$$

Proof. claim: *if $\mathcal{G}$ is transitive, then so is Ω_i.*
As $A \in \Omega_i$,by definition $\exists B$, s.t., $B \cup \{i\} \in \mathcal{C}$, but by transitivity of $\mathcal{G}$, $\pi(B \cup \{i\}) \in \mathcal{C}$.
Observe,

$$\pi(B \cup \{i\}) = \pi(B) \cup \pi(\{i\}) = \pi(B) \cup j$$

.

*Since $\pi(B) \subseteq \pi(A)$,it follows $\pi(A) \in \Omega_j$ .This indicates a bijection between Ω_i and Ω_j,i.e,$|\Omega_i| = |\Omega_j|$.Moreover since, $|A| = |\pi(A)|$, propostion follows from proposition*4 (a) □

i.e,using, $h_i(p) = \mu_p(\Omega_i) = \sum_{A \in \Omega_i} p^{|A|}(1-p)^{N-1-|A|}.$

MAP EXIT function *...symmetry*

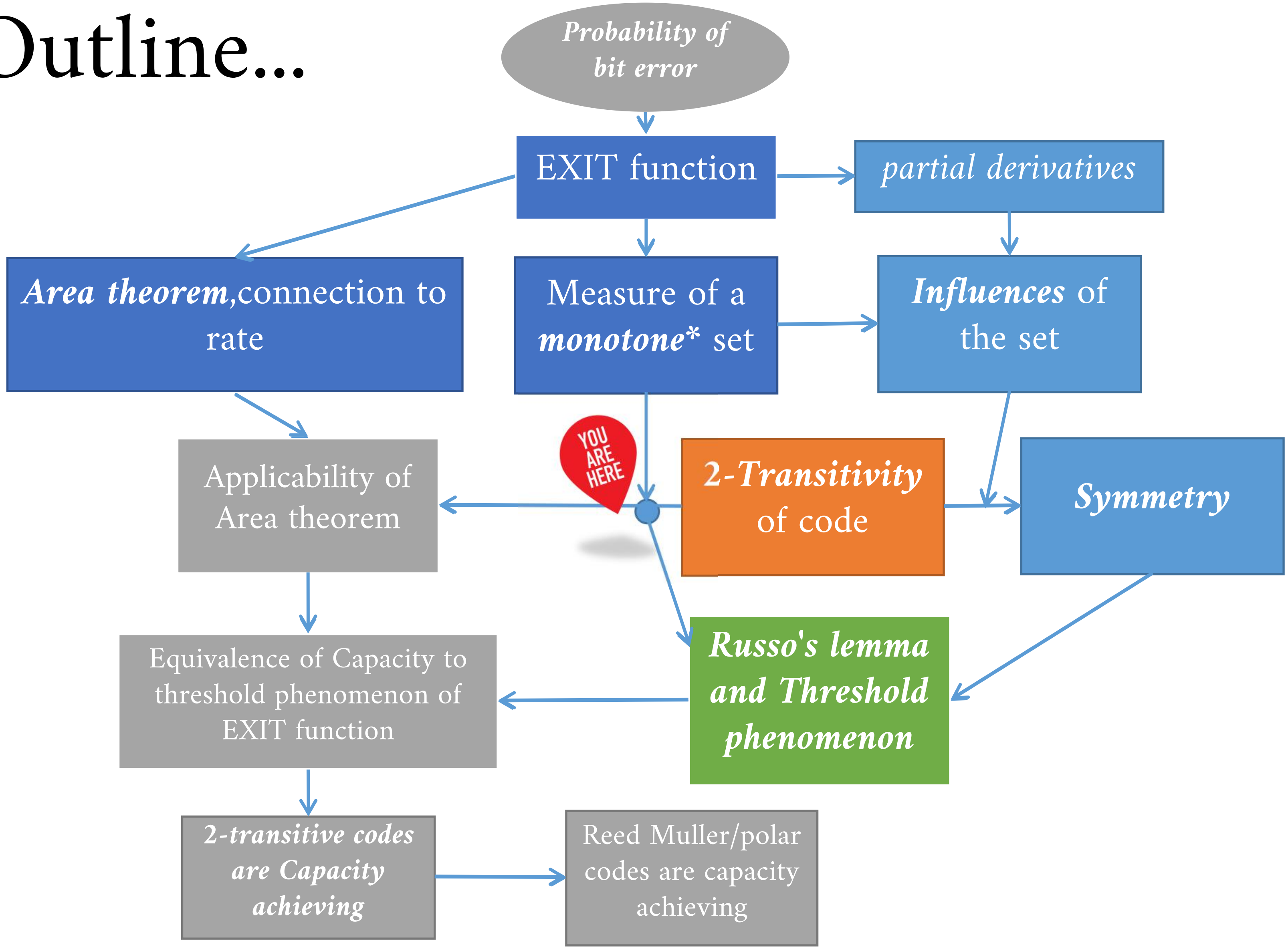
Proposition* 8**, *All Influences are equal **,** for a transitive code:**

$$\left.\frac{\partial h_i(\underline{p})}{\partial p'_j}\right|_{\underline{p}=(p,p,\ldots,p)} = \left.\frac{\partial h_i(\underline{p})}{\partial p'_k}\right|_{\underline{p}=(p,p,\ldots,p)}.$$

for distinct, $i, j, k \in [N]$,

Proof:
Similar to Proposition 7.Intuitively we expect that once we permute the locations , the bits which were pivotal must continue to remain so.Otherwise we could have decoded the concerned bit using simple permutations.

Outline...


Probability of bit error
EXIT function
partial derivatives
Area theorem,connection to rate
Measure of a monotone* set
Influences of the set
YOU ARE HERE
Applicability of Area theorem
2-Transitivity of code
Symmetry
Equivalence of Capacity to threshold phenomenon of EXIT function
Russo's lemma and Threshold phenomenon
2-transitive codes are Capacity achieving
Reed Muller/polar codes are capacity achieving

Capacity achieving codes...*equivalence*

Suppose $\{C_n\}$ is a sequence of codes with rates $\{r_n\}$ where $r_n \rightarrow r$ for $r \in (0, 1)$. a) $\{Cn\}$ is said to be ***capacity achieving on the BEC*** *under bit-MAP decoding, if for any $p \in [0, 1 - r)$, the average bit-erasure probabilities satisfy,*

$$\lim_{n\to\infty} P_b^{(n)}(p) = 0.$$

The following are equivalent,

- *S1: $\{Cn\}$ is capacity achieving on the BEC under bit-MAP decoding.*
- *S2: The sequence of average EXIT functions satisfies*

$$\lim_{n\to\infty} h^{(n)}(p) = \begin{cases} 0 \text{ if } 0 \leqslant p < 1 - r \\ 1 \text{ if } 1 - r < p \leqslant 1. \end{cases}$$

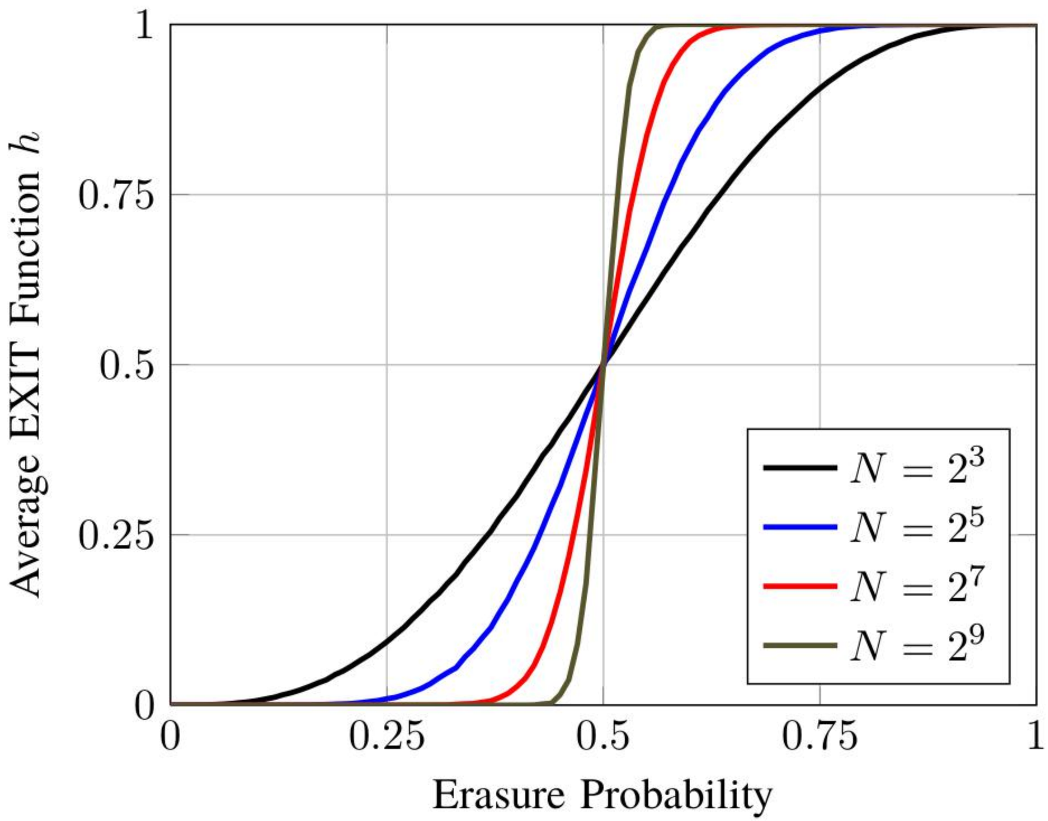
- *S3: For any,*

$$\lim_{n\to\infty} p_{1-\epsilon}^{(n)} - p_{\epsilon}^{(n)} = 0.$$

note, $h^{(n)}(p_{\epsilon}^{(n)}) = \epsilon.$

While S1 S2 , follows from relationship between $h_i(p)$ and probability of bit error ,the rest follow from area theorem

Capacity achieving codes...



*The average EXIT function of Rate-*1/2 *RM-code with,Block length N. Shows threshold phenomenon.*

Main result...*threshold phenomenon*

***Influence*:**
for a Monotone set Ω, the j-th influence is defined as.

$$I_j^{(p)}(\Omega) \triangleq \mu_p(\partial_j \Omega)$$

the total, Influence is defined as,

$$I^{(p)} \triangleq \sum_{j=1}^{N} I_j^{(p)}.$$

Theorem* 19:Let Ω be a monotone set and suppose that, for all $0 \leqslant p \leqslant 1$, the influences of all bits are equal ,i.e,*

$$I_1^{(p)}(\Omega) = \ldots = I_M^{(p)}(\Omega).$$

Then, for any $0 < \epsilon \leqslant 1/2$,

$$p_{1-\epsilon} - p_\epsilon \leqslant \frac{2 \log \frac{1-\epsilon}{\epsilon}}{C \log(N-1)},$$

where $p_t = \inf\{p \in [0, 1] : \mu p(\Omega) > t\}$ is well defined because $\mu p(\Omega)$ is strictly increasing in p with $\mu_0(\Omega) = 0$ and $\mu_1(\Omega) = 1$.
Proof: follows from Russo's lemma

Main result...*notations*

Redefinitions

We can redefine Ω_i *as a set of indicator vectors of A.*

$$[\phi_i(A)]_l = \begin{cases} \mathbf{1}_A(l) \text{ if } l < i \\ \mathbf{1}_A(l+1) \text{ if } l \geqslant i. \end{cases}$$

$$\Omega_i' \triangleq \{\phi_i(A) \in \{0,1\}^{N-1} : A \in \Omega_i\}$$

$$\partial_j \Omega_i' \triangleq \{\phi_i(A) \in \{0,1\}^{N-1} : A \in \partial_j \Omega_i\}.$$

$$= \{\underline{x} \in \{0,1\}^{N-1} | \mathbb{1}_{\Omega_i}(\underline{x}) \neq \mathbb{1}_{\Omega_i}(\underline{x}^{(j)})\}$$

Accordingly ,we can refine the definition of the measure...

$$\mu_p(\Omega) = \sum_{\underline{x} \in \Omega} p^{|\underline{x}|}(1-p)^{M-|\underline{x}|}, \text{ for } \Omega \subseteq \{0,1\}^M,$$

Main result...

Let $\{C_n\}$ be a sequence of codes where the blocklengths satisfy $N_n \to \infty$, the rates satisfy $r_n \to r$, and the permutation group $G_{(n)}$ (of C_n) is ***doubly transitive*** *for each n. If $r \in (0, 1)$,* ***then $\{C_n\}$ is capacity achieving on the BEC under bit-MAP decoding.***

Proof:

Note, $h(p) = h_i(p) = \mu_p(\Omega_i')$ *, where $h^{(n)}(p)$ is the EXIT function of C_n.*

Properties of Ω_i'

- *Ω_i' is the set of error patterns that prevent the detection of Xi.*
  Consider $A \in \Omega_i'$, any $B > A$ will surely cause error.Hence $B \in \Omega_i'$ and thus,
  Ω_i' is monotone.
- *Recall,* $I_j^p(\Omega_i') = \mu_p(\partial_j \Omega_i') = \left.\frac{\partial h_i(\underline{p})}{\partial p_j'}\right|_{\underline{p}=(p,p,\ldots,p)}$

  hence, $I_j^p(\Omega_i') = I_k^p(\Omega_i')$ for all $j, k \in [N-1]$. ***Ω_i' is symmetric.***

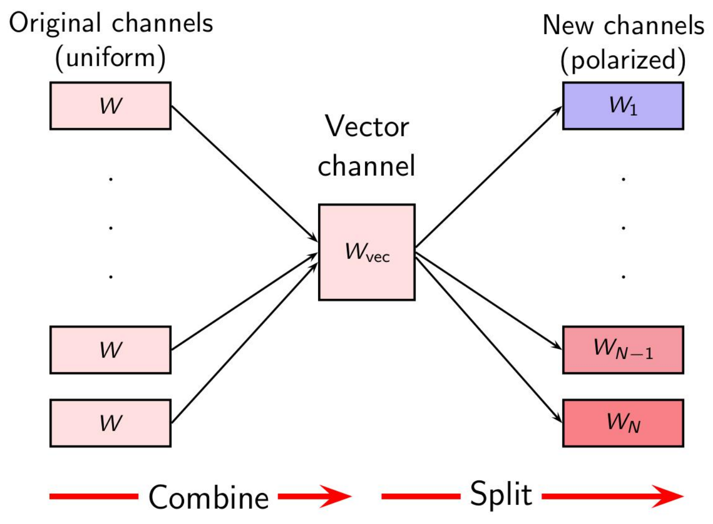
This implies, $p_{1-\epsilon} - p_\epsilon \leqslant \frac{2\log\frac{1-\epsilon}{\epsilon}}{C\log(N-1)},$

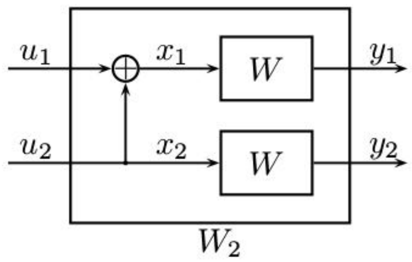
Hence $\lim_{n\to\infty}(p_{1-\epsilon} - p_\epsilon) = 0.$ *, and* ***C_n achieves Capacity.***

Polar codes are Capacity achieving…

Introduction to polar codes

Channel polarization is an operation by which one manufactures out of N independent copies of a given B-DMC W,a second set of N channels $\{W^{(i)}{}_N : 1 \le i \le N\}$ *that show a polarization effect in the sense that, as N becomes large,the symmetric capacity terms* $\{\boldsymbol{I(W^{(i)}{}_N)}\}$ ***tend towards 0 or 1****,for all but a vanishing fraction of indices i. This operation consists of a channel combining and a channel splitting phase.*





$$x_1^N = u_1^N G_N$$

Polar codes are Capacity achieving...

polar codes are doubly - transitive ,hence achieve capacity

Proof:

Assume we are given four distinct locations in the code a,b,c and d $\in [N]$ *,we need to give a permutation* $\pi : [N] \to [N]$ *such that,*

$\pi(a) = c$ *and* $\pi(b) = d$ *, and it is automorphic.*

Consider any permutation which satisfies the above constraint,proving it is automorphic suffices.

$$\text{Note , } \pi_{N\times N} * [u_1 \ldots u_n] = [u_{\pi(1)} \ldots u_{\pi(n)}]$$

$$\text{and, } \pi_{N\times N} * [x_1 \ldots x_n] = [x_{\pi(1)} \ldots x_{\pi(n)}]$$

Hence,upon premultiplying the encoding equation with π *,we get*

$$\pi * [x_1 \ldots x_n] = \pi * [u_1 \ldots u_n] * Gn$$

$$\Rightarrow [x_{\pi(1)} \ldots x_{\pi(n)}] = [u_{\pi(1)} \ldots u_{\pi(n)}] * Gn$$

As polar codes are linear , and $[u_{\pi(1)} \ldots u_{\pi(n)}]$ *is a valid message, so* $[x_{\pi(1)} \ldots x_{\pi(n)}]$ *is a valid codeword too.hence proved*


Probability of bit error
partial derivatives
EXIT function
Influences of the set
Measure of a monotone* set
Area theorem,connection to rate
Symmetry
2-Transitivity of code
Applicability of Area theorem
Russo's lemma and Threshold
Equivalence of Capacity to threshold phenomenon of EXIT function
YOU ARE HERE
2-transitive codes are Capacity achieving
Reed Muller/pol
codes are capacity
achieving

Thanks...