COMMUNICATION EFFICIENT DATA EXCHANGE AMONG MULTIPLE NODES

EP 299 :Project for M.Tech,Communication and Networks,ECE

MID-TERM PROJECT REPORT

A report by

SOUMYA SUBHRA BANERJEE
SR No. 04-02-04-37-42-16-1-14191
DEPARTMENT OF ECE,
INDIAN INSTITUTE OF SCIENCE.

Under guidance of

HIMANSHU TYAGI
ASSISSTANT PROFFESSOR
DEPARTMENT OF ECE,
INDIAN INSTITUTE OF SCIENCE.

January 9, 2018

CONTENTS

[1]

1 This Project was supported by Robert Bosch Center for Cyber Physical Systems

ABSTRACT

Efficiently decodable deterministic coding schemes which achieve channel capacity provably have been elusive until the advent of polar codes[1] in the last decade.Further,the recent results by Urbanke et al.[**?**] show that doubly transitive codes achieve capacity on erasure channel under MAP decoding.Urbanke and his group use threshold phenomenon observed in EXIT functions (which capture the error probability) ,to prove the same.These results were applied to Reed-Muller codes [**?**].Alternative proof of the fact Polar codes achieve capacity was suggested in [**?**].This report is a comprehensive study of threshold phenomenon in EXIT function and its applications as indicated above.

1 INTRODUCTION

Random correlated data (X,Y) is distributed between two parties with first observing X and second observing Y. The two parties seek to recover each others data. *The Data-Exchange problem* essentially encompasses this scenario, as depicted in figure 1. The project seeks to device a practical protocol which achieves this with minimal communication.

A working solution for this problem is r-sync protocol, as described in [2].The algorithm identifies parts of the source file which are identical to some part of the destination file, and only sends those parts which cannot be matched in this way.Though this protocol is fast and low complexity, it does not exploit the correlation of the data to the best extent possible.In fact, we can view r-sync as an algorithm which uses only one guess, and thus ends up using more communication.
In [3], David Slepian and Jack Wolf had shown that the optimal solution to this problem is Slepian-Wolf compression.

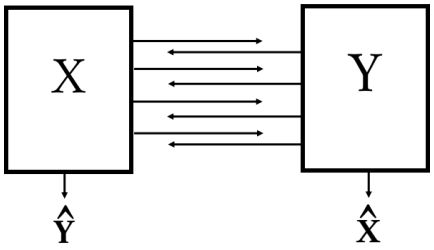

Figure 1: The Data-Exchange Problem

SLEPIAN–WOLF CODING THEOREM: states under joint decoding of X and Y a total rate H(X,Y) is sufficient.

Consider first the problem where X and Y are correlated discrete-alphabet memoryless sources, we have to compress X losslessly, with Y (side information) being known at the decoder and *not* at the encoder. If Y were known at both ends one can compress X at a theoritical rate of H(X|Y). But if Y were known only at decoder the same can be achieved by just knowing $P_{X|Y}$ at encoder without explicit information of Y, this has been depicted in figure2 [4].

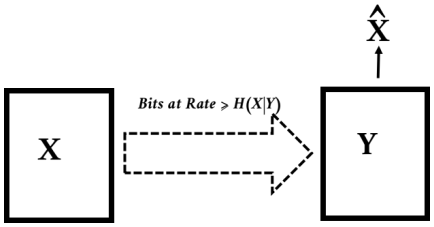

Figure 2: The Slepian-Wolf Compression

A practical implementation of Slepian Wolf compression faces the following difficulties.

- Search is over an exponential list in decoding.
- Knowledge of $P_{X|Y}$ is required.

Using structured channel codes as indicated in [4], particularly Polar Codes as shown in [5], alongwith *recursive data exchange protocol* mentioned in [6] for Slepian-Wolf compression eases the aforementioned implementation.

1.1 Suggested Approach for Solving The Data-Exchange Problem

In accordance with the above discussion the suggested approach towards solving *the Data-Exchange Problem* may be briefed as follows.

- Implement Slepian-Wolf Compression using Polar Codes.
- Achieve universality using *Recursive Data Exchange* protocol (RDE).
- Realise RDE using Rateless Polar Codes with Physical layer Error Detection.

The following sub-sections discuss the artefacts needed for this implementation in brief. Section 2, consolidates and elaborates the proposed scheme.

1.2 Interactive Communication for Data Exchange

The data exchange protocol is based on an interactive version of the Slepian-Wolf protocol where the length of communication is increased in steps until the second party decodes the data of the first. After each transmission second party sends ACK-NACK feedback signal, the protocol stops when ACK is recieved or l_{max} bits have been transmitted [6]. Note, this protocol is universal as it does not rely on knowledge of the joint distribution, instead uses an iterative variable length approach to reach rate optimality universally. The decoders suggested in [6] are theoritical constructs which use type classes to form a list of guesses for data of other parties and thus has exponential complexity.
In [5], Slepian Wolf compression is approached with structered (Polar) codes, in this work we use Rateless Polar Codes to implement RDE with a similar ideology.

1.3 Brief Introduction to Polar Codes

In 2008, E. Arikan in his paper [1] introduced Polar Codes, which provably achieves capacity on symmetric channels. Polar Codes rely on the phenomenon of channel polarization which can be described as follows.

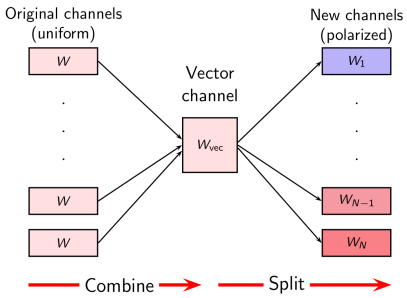

Figure 3: Channel combining and splitting

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 $\{W_N^{(i)} : 1 \leqslant i \leqslant N\}$ that show a polarization effect in the sense that, as N becomes large,the symmetric capacity terms $I(W_N^{(i)})$ tend towards 0 or 1,for all but

a vanishing fraction of indices i. Alternatively, Bhattacharya parameter $Z(W_N^{(i)})$ tend to 1 or 0 respectively.

This operation consists of a channel combining and a channel splitting phase, as shown in figure 3. The channel transformation for two independent channels is shown in 4,this is used recursively.

The encoding process sends data on transformed channels with $Z(W_N^{(i)}) = 0$ (*good channels*) and treats the channels with $Z(W_N^{(i)}) = 0$ as *bad or frozen*, thus sending no useful data on them. This scheme of error control coding with polar codes is illustrated in figure 5.

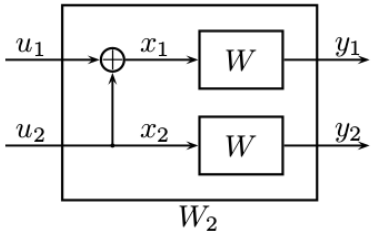

Figure 4: Arikan transformation butterfly

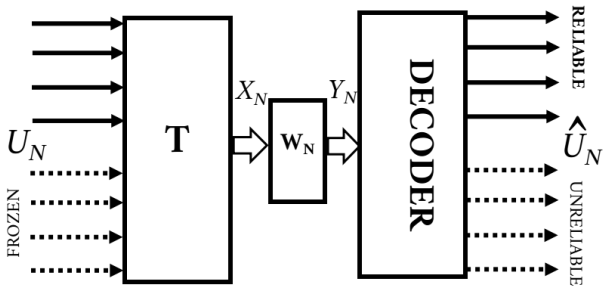

Figure 5: Polar Coding

In figure 5, U_N is a uniform message vector, T is a linear transform equivalent to the butterfly in figure 4 for a block length of N. It is useful to note here that $T = T^{-1}$.

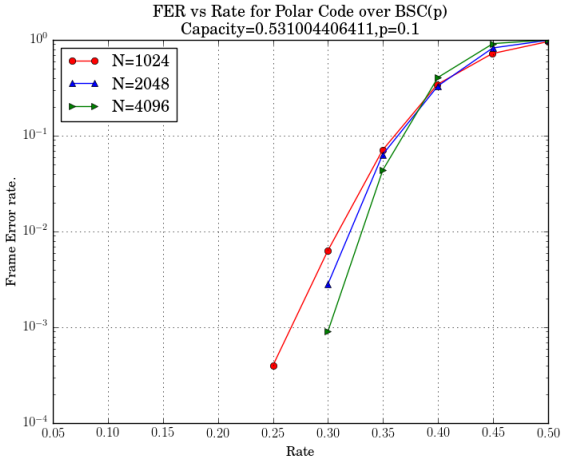

Figure 6: FER vs Rate for Polar Coding with SC

For our purpose, we shall be using Succesive Cancellation (SC) decoding. Each independent channel will considered as BSC(p). The Polar Code construction indicated in [7] has been employed for simplicity. A performance analysis of implementation of the scheme in figure 5 is presented in figure 6

1.4 Implementation of SW Compression using Polar Codes

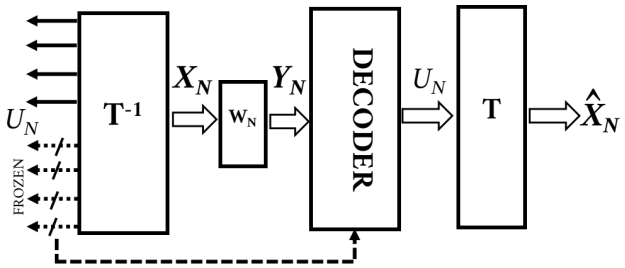

Figure 7: Polar Coding for SW Compression

In [4], the use of structured codes for Slepian-Wolf compression has been discussed.Further, in [5] a specific scheme using Polar Codes has been illustrated, as shown in figure 7.
Consider the setting where X_N and Y_N are uniform. Y_N is a corrupted version of X_N by a BSC(p). The bits that are to be sent for estimation of X_N from Y_N are the frozen bits in U_N. In other words applying T^{-1} to X_N and choosing the bits at frozen positions (the syndrome) essentially describes the compression operation. These bits are communicated error free to the SC-Decoder. On applying T to the output of SC-decoder the estimate $\hat{X}_N$ is received. It is to be noted here that in general the frozen positions are set to 0 in channel coding but in Slepian-Wolf compression they are non-zero by the nature of the scheme. Performance analysis of this scheme is presented in figure 8.

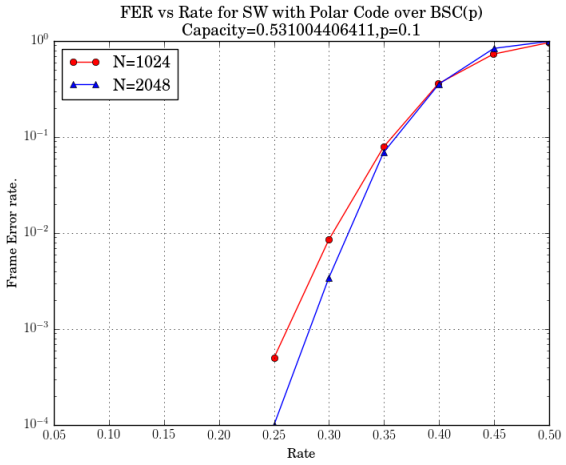

Figure 8: FER vs Rate for SW with Polar Code

Intuitively, a method to generate the syndrome incrementally will lead us to an implementation of RDE. Rateless Polar Codes will be instrumental to accomplish this.

1.5 Rateless Polar codes

RATELESS CODE: A rateless coding scheme transmits incrementally more and more coded bits over an unknown channel until all the information bits are decoded reliably by the receiver. A fixed rate code is designed for a specific channel. In contrast, a rateless code is designed for a set of channels and judged for its performance for the entire set (the compound channel). In general a rateless code design is based on hybrid-ARQ techniques and uses code puncturing.
For Polar Codes, puncturing is not straightforward. Hybrid-ARQ schemes and puncturing of Polar Codes has been proposed and compared in [8],[9],and [10].
In [11], the authors have proposed a provably capacity achieving rateless coding scheme based on Polar Codes. This scheme is useful for broad class of channels as long as they are ordered by degradation.[2]. The scheme stems from the inherent nesting property and degradedness of Polar codes, this makes puncturing a relatively simple affair.

1.5.1 *Degradedness and Nesting Property*

DEGRADED CHANNELS: A symmetric binary input channel W_2 is said to be degraded with respect to a channel W_1 if there exists random variables X,Y,Z such that X–Y–Z forms a Markov chain and $W_1 = P_{Y|X}, W_2 = P_{Z|X}$. By Data Processing Inequality it is evident that the capacity of W_2 is lesser than that of W_1. This is denoted by $W_2 \preceq W_1$. For example, $BSC(p_1) \preceq BSC(p_2)$ if $p_1 \geqslant p_2$.

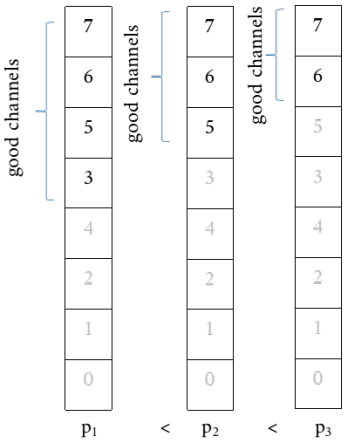

Figure 9: Nesting in BSC(p) channels

NESTING PROPERTY: Polarization suggests that $W_2 \preceq W_1$ will reflect as lesser number of *good channels* for W_2. As the polarization operation preserves degradedness, the good bit indices of W_2 must be a subset of the good bit indices of W_1. This is Nesting Property. This leads to a *reliability ordering* of the channels, such that a more reliable channel is always noiseless if a less reliable channel is noiseless regardless of the underlying channel.

2 using construction methods in [12],[13] this method can be extended to a broader class of *less noisy* ordered channels

1.5.2 *Incremental freezing*

These Rateless Polar Coding coding scheme described in [11] can be described as follows. Given the reliability ordering, a rateless scheme can be designed as follows. The initial transmission can be done using a high rate polar code with many information bits less frozen bits. If this transmission cannot be decoded[3] then among the informations bits sent, the ones on comparatively lesser reliable channels are retransmitted. By decoding these bits from future transmissions they effectively become frozen, allowing the rest of the information bits sent on the first transmission to be decoded. Thus, this scheme can be called *incremental freezing*, as future transmissions successively freeaze more and more bits sent in earlier transmission.

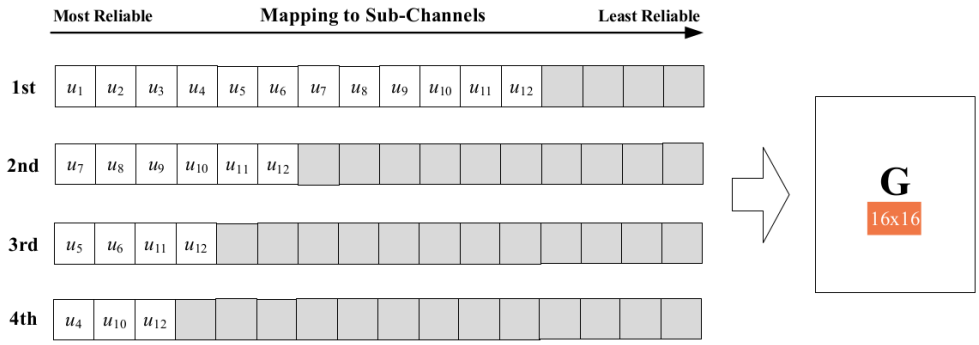

Figure 10: Incremental freezing for N=16, K=12 and 4 iterations

Figure 10 illustrates the scheme for a set of channels with rates $\{R = 12/16, R_1 = R/2, R_2 = R/3, R_3 = R/4\}$. Here u_i are the message bits. Note that with the 4^{th} transmission u_4 to u_{12} has been incrementally frozen. This scheme is capacity achieving in the sense that no rate has been wasted, the final rate achieved is,

$$R^* = \frac{12}{16x4} = \frac{3}{16} = R_3$$

. In case the real channel has capacity $\geqslant R_3$ the scheme stops at appropriate iteration to achieve that capacity.
Though this scheme is not truly rateless as it can achieve only R, R/2, R/3,...rather than a set of arbitrary rates, adding new information bits in future transmissions allow us to rectify this. In similar fashion this can be extended to parallel channels. It is useful to note that a certain number of channels in this scheme is "*always available*" guaranteeing a certain rate in each transmission. A specific application of this will be discussed in section 3.
In [11], The performance evaluation of the scheme establishes that n iterations of the scheme is almost equivalent in performance to a R/n fixed rate Polar Code.
Adaptation of this scheme for RDE will be indicated in section 2.

3 Decodability is checked by CRC in higher layers.

The following subsections briefly discuss Rateless Polar Coding from the perspective of Hybrid-ARQ.

1.5.3 *H-ARQ for Polar codes*

HYBRID ARQ: Hybrid automatic repeat request (hybrid ARQ or HARQ) is a combination of high-rate forward error-correcting coding and ARQ error-control. In standard ARQ, redundant bits are added to data to be transmitted using an error-detecting (ED) code such as a cyclic redundancy check (CRC). Receivers detecting a corrupted message will request a new message from the sender.
In Hybrid ARQ, the original data is encoded with a forward error correction (FEC) code, and the parity bits are either immediately sent along with the message (Type-I) or only transmitted upon request when a receiver detects an erroneous message (Type-II). The retransmission vector is also called the *Redundancy Vector*(RV).
The ED code may be omitted when a code is used that can perform both forward error correction (FEC) in addition to error detection, such as a Reed-Solomon code or Turbo Product Code (TPC) [14].
The FEC code is chosen to correct an expected subset of all errors that may occur, while the ARQ method is used as a fall-back to correct errors that are uncorrectable using only the redundancy sent in the initial transmission.
Construction of H-ARQ requires a Rate Compatible Code and a choice of RV (retransmission scheme).

RATE COMPATIBLE CODES: Given a fixed number of information bits, consider a family of codes $\{C_1, C_2 ... C_n\}$ with rates $R_1 \geqslant R_2 \geqslant R_3 ... \geqslant R_n$, and block lengths $N_1 \leqslant N_2 \leqslant ... \leqslant N_n$. Then the set is rate compatible if codeword for C_i can be built by removing $N_j - N_i$ bits from codewords of code C_j ,$j \geqslant i$, [13]. Rate Compatible Codes can be constructed by puncturing low rate codes.
Polar codes for degraded channels is inherently rate compatible due to *reliability ordering* and *nesting*.

H-ARQ schemes for Polar Codes

In [10], a H-ARQ scheme for Polar codes based on selective repetition has been studied. The incremental freezing scheme described in [11] and [13] can be classified as a H-ARQ scheme. In [8] another H-ARQ scheme based on Subset Polar Codes is described along with performance comparison with the other schemes. Here these schemes are briefly discussed.

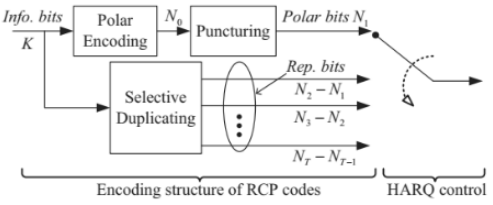

Figure 11: H-ARQ for Polar Codes with selective repetition

H-ARQ FOR POLAR CODES WITH SELECTIVE REPETITION: As the initial transmission, an information block of K bits is fed into a polar encoder. The output

codeword of N_0 bits is punctured into N_1 bits and sent over the channel. If the receiver fails to decode the codeword, an NACK (negative acknowledgement) is sent to the transmitter through the feedback channel. And then, $N_2 - N_1$ of the information bits are retransmitted. This time, the receiver tries to perform decoding with all the N_2 received bits. If the decoding is failed again, another $N_3 - N_2$ bits are transmitted. This process continues until the transmitter receives an ACK (acknowledgement), or a maximum number of transmissions T is achieved. The scheme has been illustrated in figure 11. The retransmitted bits (RV) are chosen one at a time as the most unreliable of the K bits transmitted, reliability is re-calculated after choosing one bit and the process is iterated.

INCREMENTAL FREEZING: this scheme is an improvement over the above scheme, it uses the reliability ordering to choose the RV as discussed in 1.5.2.

H-ARQ FOR POLAR CODES BASED ON SUBSET POLAR CODES: A Subset Polar Code can be created by greedily puncturing a low-rate mother code without re-optimizing the information bits.The scheme uses equivalent subset Polar Codes as RV. In [8], the author has claimed from simulation results that this scheme performs better among the ones discussed. Nevertheless, we have used incremental freezing for its simplicity.

RELIABILITY BASED H-ARQ: It is to be noted that the schemes described above use CRC for checking decodability. Reliability based HARQ technique (RBHARQ) [15], eliminates the use of CRC by approximating bit and word error probability from likelihood ratios (LLR). As the magnitude of a log-likelihood value is directly connected to the error probability of the corresponding bit, it can be used to determine which bits most likely caused a word error. The bit error probability for the k^{th} bit can be estimated form LLR ($\tilde{u}_k$) as,

$$P_{b,k} = P(\hat{u_k} \neq u_k) = \frac{1}{1 + e^{|\tilde{u}_k|}}$$

then word error probability becomes,

$$P_w = 1 - e^{\log \bar{P}_w}$$

where,

$$\log \bar{P}_w = \log \prod_{k=1}^{K} (1 - P_{b,k})$$

if the word error probability does not meet the requirements the bits with higher bit error probability may be retransmitted. This retransmission criterion results in increase of throughput, particularly evident in case of short packet lengths.

2 PROPOSED IMPLEMENTATION OF RECURSIVE DATA EXCHANGE

Let $\mathcal{C}$ denote an (N, K) proper binary linear code with length N ,dimension K ,minimum distance (d_{min}) atleast 2,and rate defined by $r \triangleq K/N$. We assume that a random codeword is chosen uniformly from this code and transmitted over a memoryless Binary Erasure Channel (BEC).A BEC with erasure probability p is denoted by BEC(p),or BEC($\underline{p}$) in case the erasure probability is different for each bit where $\underline{p} = (p_1, \ldots, p_n)$ and p_i indicates the erasure probability for bit i. The input and output alphabets of the BEC are denoted by $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, e\}$, respectively. Let $\underline{X} = (X_1, \ldots, X_N) \in \mathcal{X}_N$ be a uniform random codeword and $\underline{Y} = (Y_1, \ldots, Y_N) \in \mathcal{Y}_N$ be the received sequence obtained by transmitting $\underline{X}$ through a BEC(p).For a vector $a = (a_1, a_2, \ldots, a_N)$, the shorthand $\underline{a}_{\sim i}$ denotes $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N)$. Let $\underline{a}$,$\underline{b}$ denote the indicator vectors of the sets $A \subseteq [N]$,$B \subseteq [N]$.We say that A *covers* B if $B \subseteq A$, equivalently $\underline{a} \leqslant \underline{b}$.For linear codes and erasure channels, it is possible to recover the transmitted codeword if and only if the erasure pattern does not cover any codeword.Similarly, it is possible to recover bit i if and only if the erasure pattern does not cover any codeword where bit i is non-zero.

2.1 Adaptation of Rateless Polar Codes for RDE

BIT ERROR PROBABILITY: Let $D_i : \mathcal{Y}^N \to \mathcal{X} \cup \{e\}$ denote the bit-MAP decoder for bit i of $\mathcal{C}$. For a received sequence $\underline{Y}$, if X_i can be recovered uniquely, then $D_i(\underline{Y}) = X_i$. Otherwise, D_i declares an erasure and returns e. Let the erasure probability for bit $i \in [N]$ be

$$P_{b,i} \triangleq \mathbb{P}[D_i(\underline{Y}) \neq X_i].$$

and the average bit erasure probability be

$$P_b \triangleq \frac{1}{N} \sum_{i=1}^{N} P_{b,i}.$$

Whenever bit i can be recovered from a received sequence $\underline{Y} = \underline{y}$, $H(X_i|\underline{Y} = \underline{y}) = 0$. Otherwise, the uniform codeword assumption implies that the posterior marginal of bit i given the observations is $\mathbb{P}(X_i = x|\underline{Y} = \underline{y}) = 1/2$ and $H(X_i|\underline{Y} = \underline{y}) = 1$. This immediately implies that

$$P_{b,i} = H(X_i|\underline{Y})$$

and,

$$P_b = \frac{1}{N} \sum_{i=1}^{N} H(X_i|\underline{Y}).$$

2.2 PHY-Layer error Detection

The vector EXIT function associated with bit i of the (uniformly randomly chosen) codeword is

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

The average vector EXIT function is defined by

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

Scalar EXIT functions are defined by choosing $\underline{p} = (p, p, \ldots, p)$.

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

Therefore,

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

and

$$P_b(p) = ph(p).$$

Proposition 1. The MAP EXIT function for the ith bit satisfies $h_i(p) = \frac{\partial H(\underline{X}/\underline{Y}(\underline{p}))}{\partial p_i}$

Proof.

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

We note the second term is independent of p_i,the proposition follows on differentiation. □

INDIRECT RECOVERY Consider a code $\mathcal{C}$ 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

Definition 2. $\Omega_i \triangleq \{A \subseteq [N]\backslash\{i\} : \exists B \subseteq [N]\backslash\{i\}, B \cup \{i\} \in \mathcal{C}, B \subseteq A\}$.

For distinct $i, j \in [N]$, the set of erasure patterns where the j-th bit is *pivotal* for the indirect recovery of X_i is given by

Definition 3. $\partial_j\Omega_i \triangleq \{A \subseteq [N]\backslash\{i\} : A\backslash\{j\} \notin \Omega_i, A \cup \{j\} \in \Omega_i\}$

These are the erasure patterns where X_i can be recovered from $\underline{Y}_{\sim i}$ if and only if $Y_j \neq e$.Note $\partial_j\Omega_i$ includes patterns from both Ω_i and Ω_i^c.

Proposition 4. For a code $\mathcal{C}$ and transmission over a BEC, we have the following properties for the EXIT functions.

(a) The EXIT function associated with bit i satisfies

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

(b) For $j \in [N] \backslash \{i\}$, the partial derivative satisfies

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

(c) The average EXIT function satisfies the *area theorem*

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

Where $\mu_p(\Omega)$ is the measure of the set of erasure patterns Ω. Here (a) and (b) follow from definition of conditional entropy and the fact that $H(X_i/\underline{Y}_{\sim i} = \underline{y}_{\sim i}) = 1$ when $A \cup \{i\}$ covers some codeword and decoding fails, and 0 otherwise.(c) is a direct consequence of proposition 1 From the above discussion it is clear that the measure of the set Ω_i is equal to the probability of error for bit i , which in turn is equal to the ith EXIT function due to the uniform input assumption.

2.2.1 *The Error Detection Test*

Let S_N be the symmetric group on N elements. The permutation group of a code is defined as the subgroup of S_N whose group action on the bit ordering preserves the set of codewords.

Definition 5. The permutation group $\mathcal{G}$ of a code $\mathcal{C}$ is defined to be

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

Definition 6. - Suppose $\mathcal{G}$ is a permutation group. Then,

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

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

Proposition 7. *All EXIT functions are equal.* Suppose the permutation group $\mathcal{G}$ of a code $\mathcal{C}$ is transitive. Then, for any $i \in [N]$,

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

Proof. claim:if $\mathcal{G}$ is transitive, then so is Ω_i. As $A \in \Omega_i$,by definition2 $\exists B$, 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 proposition4 (a) □

Proposition 8. Suppose the permutation group $\mathcal{G}$ of a code $\mathcal{C}$ is doubly transitive. Then, for distinct $i, j, k \in [N]$, and any $0 \leqslant p \leqslant 1$,

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

Proof. Similar to the proof of 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. □

SUMMARY OF PROPERTIES OF EXIT FUNCTION.

1 $h_i(p)$ captures the bit error probability of MAP decoder.

2 $h_i(p)$ is measure of Ω_i.

3 All EXIT functions are equal to average EXIT function h(p).

4 $h_i(p)$ is strictly increasing and invertible.(follows from proposition 4 (b))

5 The area under the h vs p curve is the rate(by *Area Theorem*).

We may notice here, that is $A \in \Omega_i$, it is a bit erasure pattern that causes error at position i, then $B \supset A$ will surely cause errors, and $B \in \Omega_i$.Thus Ω_i is monotone.Proving Ω_i is symmetric and has a sharp threshold at $p = 1 - r$, will establish that 2-transitive codes achieve Capacity.We will formalize this in the following subsections.

2.3 Proposed tests

Definition 9. Suppose $\{\mathcal{C}_n\}$ is a sequence of codes with rates $\{r_n\}$ where $r_n \to r$ for $r \in (0, 1)$. a) $\{\mathcal{C}_n\}$ 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 theorem bridges capacity achieving codes, average EXIT functions, and the sharp transition framework that allows us to show that the transition width of certain functions goes to 0. The average EXIT functions of some rate-1/2 Reed-Muller codes are shown in Figure **??**. Observe that as the blocklength increases, the transition width of the average EXIT function decreases. According to the following proposition, if this width converges to 0, then Reed-Muller codes achieve capacity on the BEC under bit-MAP decoding.

Proposition 10. Let $\{\mathcal{C}_n\}$ be a seq. of codes with rates $\{r_n\}$, $r_n \to r$ for $r \in (0, 1)$. The following are equivalent -

S1: $\{\mathcal{C}_n\}$ 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 $0 < \epsilon \leqslant 1/2$,

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

where $h^{(n)}(p_{\epsilon}^{(n)}) = \epsilon$.
In short S1 $\Rightarrow$ S2, due to close relationship between bit error probability and average EXIT function pointed out in propostion 4 .S2 $\Rightarrow$ S3,and S3 $\Rightarrow$ S1 by area theorem.Hence, proving S3 suffices to complete the proof.

2.4 Performance Evaluation

3 CONCLUSION AND FUTURE WORK

proposed scheme shortpacket implementation

Definition 11. We can redefine Ω_i as a set of indicator vectors of A.Let,

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

$$\begin{aligned} \Omega_i' &\triangleq \{\phi_i(A) \in \{0,1\}^{N-1} : A \in \Omega_i\} \\ \partial_j\Omega_i' &\triangleq \{\phi_i(A) \in \{0,1\}^{N-1} : A \in \partial_j\Omega_i\}. \\ &= \{\underline{x} \in \{0,1\}^{N-1} | \mathbb{1}_{\Omega_i}(\underline{x}) \neq \mathbb{1}_{\Omega_i}(\underline{x}^{(j)})\} \end{aligned}$$

Here the last equality follows from definition of $\partial_j\Omega_i$.
Consider the space $\{0,1\}^M$,we can redefine measure μ_p such that

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

where the weight $|\underline{x}| = x_1 + x_2 + \ldots + x_M$ is the number of 1's in $\underline{x}$.

Definition 12. For a monotone set Ω.The influence of bit $j \in [N]$,is defined by,

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

The total influence in defined by,

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

Using proposition 4(a) and proposition 7, we have,

$$h(p) = h_i(p) = \mu_p(\Omega_i')$$

Further,from proposition 4(b), we get,

$$I_j^p(\Omega_i') = \mu_p(\partial_j\Omega_i') = \left.\frac{\partial h_i(\underline{p})}{\partial p_j'}\right|_{\underline{p}=(p,p,\ldots,p)}$$

where j' is given by

$$j' = \begin{cases} j \text{ if } j < i \\ j+1 \text{ if } j \geqslant i. \end{cases}$$

Since $\mathcal{G}$ is doubly transitive, from propostion 8,

$$I_j^p(\Omega_i') = I_k^p(\Omega_i') \text{ for all } j,k \in [N-1].$$

Hence,Ω_i is a *symmetric monotone set.*

The following theorem could be seen as a consequence of the result by Talagrand [**?**].

Theorem 1. *Let Ω be a monotone set and suppose that, for all $0 \leqslant p \leqslant 1$, the influences of all bits are equal $I_1^{(p)}(\Omega) = \ldots = I_M^{(p)}(\Omega)$. 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 $p_t = \inf\{p \in [0,1] : \mu_p(\Omega) \geqslant t\}$ is well defined because $\mu_p(\Omega)$ is strictly increasing in p with $\mu_0(\Omega) = 0$ and $\mu_1(\Omega) = 1$.

Proof. Using Russo's lemma [**?**]. □

We see that Ω_i satisfies the conditions of Theorem 1. Hence,

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

Further , using proposition 10, we state,
$\{\mathcal{C}_n\}$ is capacity achieving on the BEC under bit-MAP decoding.

REFERENCES

[1] E. Arıkan. Channel polarization: A method for constructing capacity- achieving codes for symmetric binary-input memoryless channels. *IEEE Transactions on Information Theory*, 55, no. 7:3051–3073, July 2009.

[2] A. Tridgell and P. Mackerras. The r-sync algorithm. *Joint Computer Science and Technical Report Series*, TR-CS-96-05, 1996.

[3] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 1973.

[4] S. Sandeep Pradhan and K. Ramachanderan. Distributed source coding using syndromes (discus):design and construction. *IEEE Transactions on Information Theory*, 49,No. 3, 2003.

[5] S. Onay. Polar codes for nonassymetric slepian-wolf coding. *arXiv.1208.3056v1[cs.IT]*, August 2012.

[6] H. Tyagi and S. Watanabe. Universal multiparty data exchange and secret key arrangement. *ISIT*, 2016.

[7] Y. Zhang et al. A practical construction method for polar codes. *IEEE Communication letters*, 18, no. 11, Nov. 2014.

[8] S. Tavildar. A h-arq scheme for polar codes. *arXiv.1606.08545v1[cs.IT]*, Jun. 2016.

[9] J. Cheng. Coding performance of hybrid arq schemes. *IEEE TRANSACTIONS ON COMMUNICATIONS*, VOL. 54, NO. 6, Jun. 2006.

[10] J. Lin K. Chen, K. Niu. A hybrid arq scheme based on polar codes. *IEEE TRANSACTIONS ON COMMUNICATIONS*, VOL. 17, NO. 10, Oct. 2013.

[11] K. Chen B. Li, D. Tse and H. Shen. Capacity achieving rateless polar codes. *arXiv.1508.0311v1[cs.IT]*, Aug. 2015.

[12] E. Sasoglu and L. Wang. Universal polarization. *arXiv:1307.7495v2[cs.IT]*, Dec. 2013.

[13] M. Mondelli et al. Capacity achieving rate compatible polar codes for general channels. *arXiv.1611.01199v2[cs.IT]*, Jan. 2017.

[14] M. Al-Mualla H. Mukhtar, A. Al-Dweik. Crc-free hybrid arq using turbo product codes. *IEEE TRANSACTIONS ON COMMUNICATIONS*, VOL. 62, NO. 12, Dec. 2014.

[15] H. Schoeneich J. Fricke and P. A. Hoeher. Reliability based retransmission criterion for harq. *IEEE TRANSACTIONS ON COMMUNICATIONS*, VOL. 57, NO. 8, Aug. 2009.