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 8, 2018

CONTENTS

[1] 4This 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.

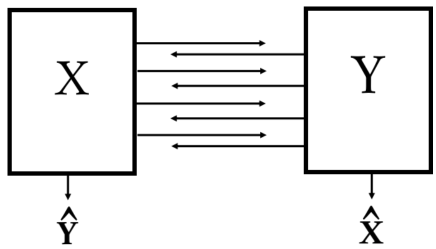

Figure 1: The Data-Exchange Problem

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].

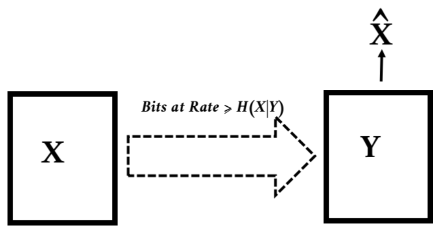

Figure 2: The Slepian-Wolf Compression

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 a 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 approach to

1.3 Brief Introduction to Polar Codes

Here we assume the reader has prior exposure to threshold phenomenon of boolean functions.

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 $1 - r$.Towards this,

(1) if the bit error probability under MAP-decoding can be captured by a function of erasure probability (p),

(2) which in turn is the measure of a symmetric monotone set ,then we shall observe threshold phenomenon,

(3) if the threshold is sharp occuring at $p = 1 - r$ under the settings of stated theorem ,it provides a proof for the theorem.

Extrinsic information transfer (EXIT) function [7] denoted by $h(p)$ in fig:**??**, and the area theorem for EXIT functions [8] occupy a central role in this work. For a given input bit, the EXIT function is defined to be the conditional entropy of the input bit i given the outputs associated with all other input bits.The value of the EXIT function at a particular erasure value is also directly related to the bit error probability under bit-MAP decoding,hence EXIT functions serve as (1).Furthermore, EXIT function can be associated with measure of set of erasure patterns Ω_i which are symmetric monotone sets for doubly transitive codes, solving(2).Finally application of area theorem solves (3). The rest of the report elaborates on (1),(2) and (3) with focus on bit error probability under bit-MAP decoding.

1.4 Implementation of SW Compression using Polar Codes

1.5 Rateless Polar codes

1.5.1 *Degradedness and Nesting Property*

1.5.2 *Incremental freezing*

H-ARQ for Polar codes

Rate Compatible Polar Codes

H-ARQ schemes

Reliability Based H-ARQ

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 $\underline{p} = (p_1, \dots, p_n)$ 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 $\underline{X} = (X_1, \dots, X_N) \in \mathcal{X}_N$ be a uniform random codeword and $\underline{Y} = (Y_1, \dots, Y_N) \in \mathcal{Y}_N$ be the received sequence obtained by transmitting $\underline{X}$ through a BEC(p).For a vector $a = (a_1, a_2, \dots, a_N)$, the shorthand $\underline{a}_{\sim i}$ denotes $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_N)$. 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 $0 < \epsilon \leqslant 1/2$,

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

where $h^{(n)}(p^{(n)}_{\epsilon}) = \epsilon$.
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 [9].

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 [10]. □

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] S. ten Brink. Convergence of iterative decoding. *Electronic Letters*, 35, no. 10:806–808, May 1999.

[8] A. Ashikhmin, G. Kramer, and S. ten Brink. Extrinsic information transfer functions: model and erasure channel properties. *IEEE Transactions on Information Theory*, 50, no. 11:2657–2674, Nov. 2004.

[9] M. Talagrand. On russo's approximate zero-one law. *The annals of probability*, pages 1576–1587, 1994.

[10] L. Russo. An approximate zero-one law. *Probability Theory and Related Fields*, 61, no. 1:129–139, 1982.