Communication Efficient Data Exchange Among Multiple Nodes

Soumya Subhra Banerjee

Under guidance of,

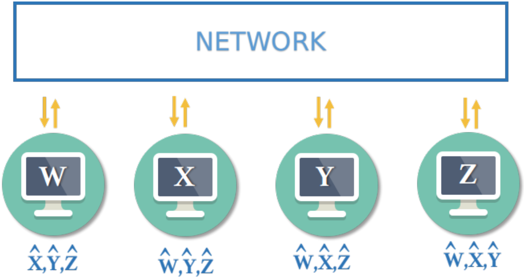
Himanshu Tyagi

Mid-Term Project presentation,
EP 299 Project, M.Tech, Communication and Networks, ECE

April 3, 2018

Motivation

The Data-Exchange problem

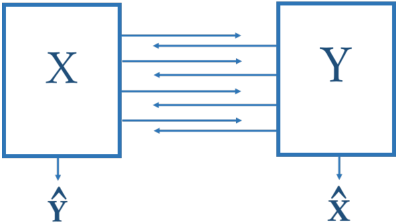


Multiple parties observing correlated data seek to recover each other's data. How can they accomplish this using minimum communication?

The Data-Exchange Problem

Two party case

- Random correlated data (X,Y) is distributed between two parties.
- The first observes X and second observes Y.
- They seek to recover each others data.
- The joint distribution of X and Y is unknown.



This project seeks to device a protocol which achieves this with minimal communication.

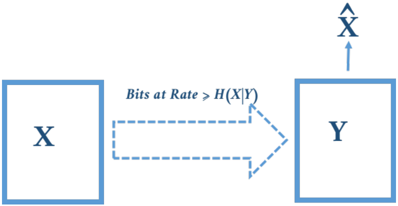
The order of communication is important in multiparty.

Working Solution

r-sync vs Slepian-Wolf compression

- In practice, algorithms like r-sync are used for data exchange.
  - Uses *one* guess.
  - Does not exploit the correlation between the data well.
  - Needs more communication.
  - Fast and low complexity.

- In theory, Slepian-Wolf compression is optimal.
  - $H(X|Y)$ is sufficient to estimate X from Y.



Implementation of SW compression.

Difficulties and suggested approach...

Difficulties in implementation of SW compression

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

Suggested Approach

- Implement SW Compression using Polar Codes.
- Achieve universality using *recursive data exchange* protocol (RDE).
- Realize RDE using Rateless Polar Codes with physical layer error detection.

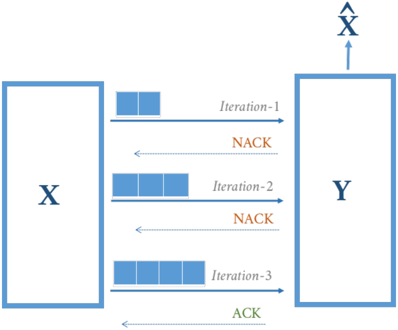
Outline

- Background
  - Recursive Data Exchange (RDE)
  - Brief introduction to Polar Codes
  - Slepian-Wolf compression with Polar Codes
  - Rateless Polar Codes
  - Rateless Polar Codes as HARQ
- Proposed implementation of RDE
  - Adaptation of Rateless Polar Code for RDE
  - PHY-Layer error detection
- Performance evaluation
- Conclusion and future work

Recursive Data Exchange (RDE)

The *recursive data exchange*[*] protocol is based on an interactive version of the SW protocol

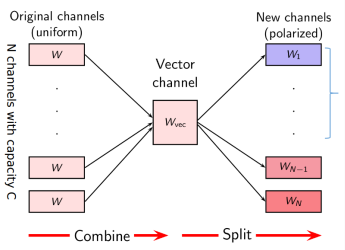
- This protocol is *universal* as it does not rely on knowledge of the joint distribution.
- The suggested decoders are theoretical constructs which are not implementable.

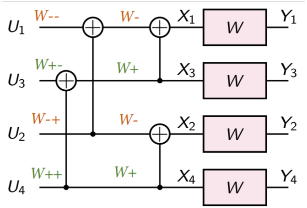


[*] H. Tyagi and S. Watanabe, Universal Multiparty Data Exchange and Secret Key arrangement, *ISIT*, 2016

Brief Introduction to Polar Codes

Channel polarization

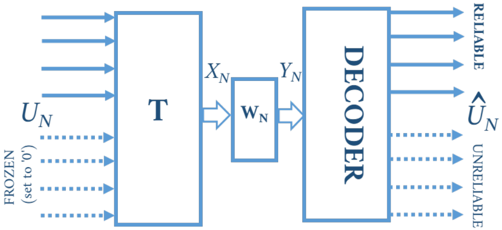




- N independent copies of a given B-DMC (W) are combined and split into a second set of N channels $\{W_N^{(i)} : 1 \leq i \leq N\}$.
- There symmetric capacity $I(W_N^{(i)})$ tend towards 0 or 1.
- The channels with Bhattacharya parameter $Z(W_N^{(i)}) = 0$ captures the capacity of W_{vec}.

Brief Introduction to Polar Codes

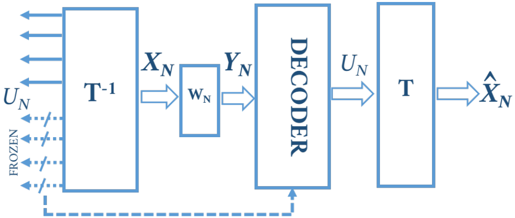
Encoding and decoding



- The encoding process[†] sends data on transformed channels with $Z(W_N^{(i)}) = 0$ (*good channels*) and treats the channels with $Z(W_N^{(i)}) = 1$ as *frozen*, sending no useful data on them.
- For our purpose, we shall be using Successive Cancellation (SC) decoding.

[†] U_N is a uniform message vector, T is a linear transform for the butterfly.

Slepian-Wolf compression with Polar Codes



- Y_N is a corrupted version of X_N by N BSC(p) channels.
- The bits that are to be sent for estimation of X_N from Y_N are the frozen bits in U_N. This is the data compression operation.
- These bits are communicated error free to the SC-decoder.
- $H(X_N) - I(W_N) = H(X_N/Y_N)$ bits are sent.

S. Onay.Polar Codes for Nonassymetric Slepian-Wolf Coding.arXiv.1208.3056v1[cs.IT], August 2012.

Rateless Polar Codes

Rateless code

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 rateless code is designed for a set of channels and judged for its performance for the entire set.
- In general rateless code design is based on Hybrid-ARQ techniques and uses code puncturing.
- Rateless Polar Codes can be constructed using nesting property of Polar Codes for degraded channels.

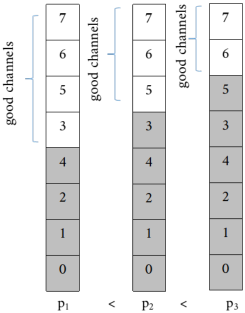
Rateless Polar Codes

Degraded channels and nesting property

Degraded channels

if X–Y–Z, and $W_1 = P_{Y|X}$, $W_2 = P_{Z|X}$ then $W_2 \preceq W_1$.

- The capacity of W_2 is lesser than that of W_1. W_2 has lesser number of good channels.
- e.g., $BSC(p_1) \preceq BSC(p_2)$ if $p_1 \geq p_2$.

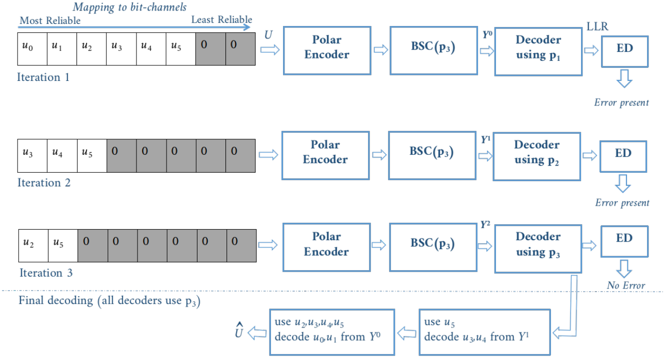


- The good bit indices of W_2 is a subset of the good bit indices of W_1.[‡]
- A more reliable bit-channel is always noiseless if a less reliable bit-channel is noiseless. This leads to *reliability ordering*.

[‡] E. Sasoglu and L. Wang. Universal Polarization. arXiv:1307.7495v2[cs.IT], Dec. 2013.

Incremental Freezing

Rateless Polar Code employing reliability ordering



- Initial transmission is done using a high rate Polar Code.
- If decoding fails then the comparatively lesser reliable channels are retransmitted.

M. Mondelli et al. Capacity achieving rate compatible polar codes for general channels, Jan. 2017.

more info...

Rateless Polar Codes as Hybrid ARQ

- 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.

Hybrid -ARQ

In Hybrid ARQ, the original data is encoded with a forward error correction (FEC) code, and the parity bits are only transmitted upon request when a receiver detects an erroneous message (Type-II).

- Construction of HARQ requires the following,
  - Rate Compatible Code
  - A choice of retransmission vector (RV).

Rateless Polar Codes as Hybrid ARQ

H-ARQ for Polar Codes

- Polar Codes for degraded channels are inherently rate compatible due to reliability ordering and nesting.
- The choice of RV may be based on,
  - Selective Repetition of unreliable bits.
  - Incremental Freezing.
  - Subset Polar codes.
- The ED code may be omitted using Reliability based H-ARQ.
  - Reliability based HARQ uses the soft outputs (LLR) of the decoder to perform ED.

Proposed Solution

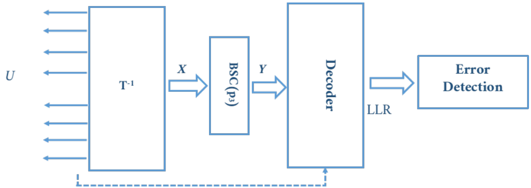
Our approach towards implementation of a solution to the Data-Exchange problem using RDE and Polar Codes ...

Proposed implementation of RDE

- Use Rateless Polar Codes with Incremental Freezing to implement RDE.
- Use a PHY layer error detection as retransmission criterion in Incremental Freezing.

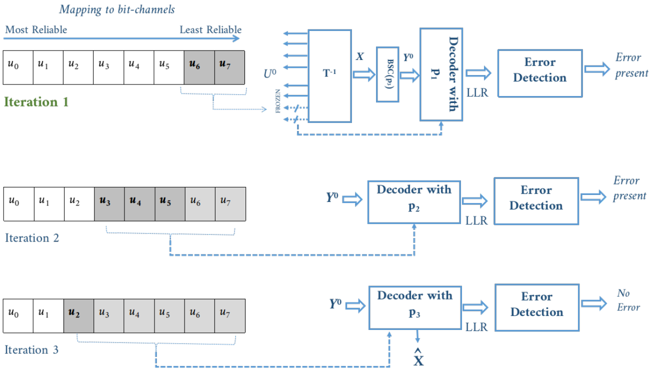
Adaptation of Rateless Polar Codes for RDE

Setting



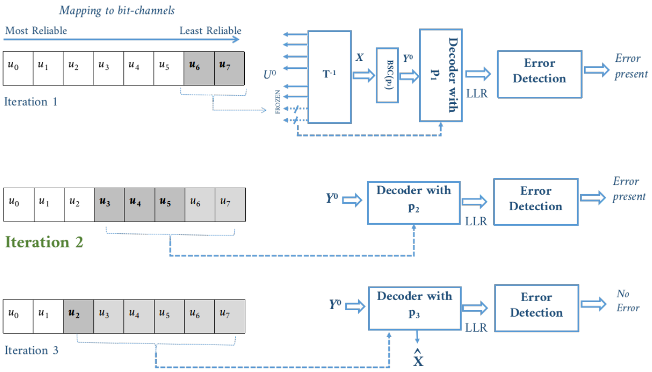
- Consider, a compound BSC channel $\mathcal{C} = \{p_1 \leq p_2 \leq p_3 \leq p_4\}$. where p_i is the flipover probability of channel i.
- The rates supported by the channels are $\{R_1 = R, R_2 = R/2, R_3 = R/3, R_4 = R/4\}$.
- The actual channel is BSC(p_3). We shall denote this as BSC($p_{channel}$).
- The vector channel is manufactured by polarization of N such BSC($p_{channel}$) channels.

Adaptation of Rateless Polar Codes for RDE



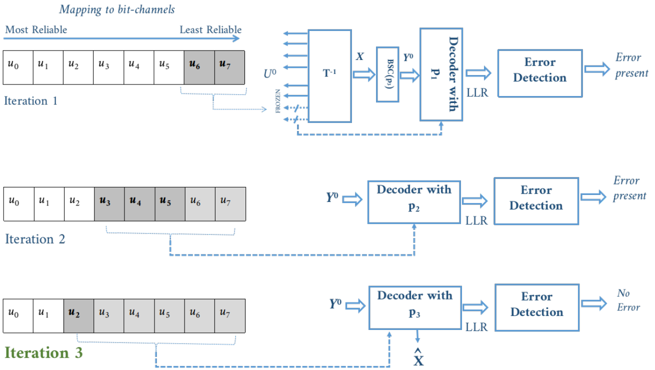
- In the first iteration the encoder and decoder *guesses* the channel to be BSC(p_1) greedily.
- The encoder sends the corresponding frozen bits to decoder error-free.
- The decoder computes $|LLR|$ using BSC(p_1) and performs ED.
- In case of error (as here) receiver replies with a NACK.

Adaptation of Rateless Polar Codes for RDE



- In the second iteration $p_{guess} = p_2$.
- The error-free transmission in this iteration consist of information bits which are reliable if $p_{channel} = p_1$ but unreliable if $p_{channel} = p_2$.
- The decoder replies with NACK.

Adaptation of Rateless Polar Codes for RDE



- In the third iteration $p_{guess} = p_3 = p_{channel}$.
- The ED declares no error and replies with ACK.
- The decoder now decodes the received channel output vector using p_3 and considering the bits transmitted error-free in *all* the iterations as frozen. Finally, with Arikan transform X is estimated.

PHY-Layer Error Detection

- In Polar Codes for error control CRC in U_N can be exploited for ED.
- In SW-compression with Polar Codes U_N is generated from X_N.
  - Hence, CRC cannot be used.
  - PHY-Layer ED is required as retransmission criterion.

PHY-ED as hypothesis test

Since Inc-Frz guesses the best channel first, we can say that at j^{th} iteration $p_{guess} = p_j$. Error detection may be seen as a hypothesis test in *each iteration* where,

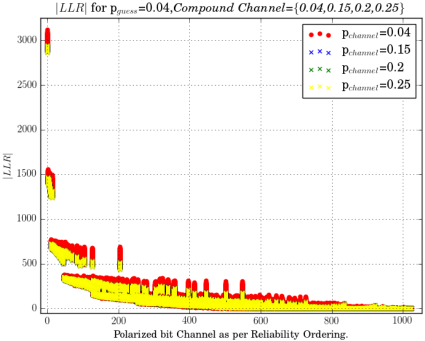
$$\mathcal{H}_0 : \text{The channel is the current guess, i.e., } j = i$$
$$\mathcal{H}_1 : \text{The channel is worse, i.e., } j < i$$

Note, $\mathcal{C} = \{p_1 \leq p_2 \leq p_3 \leq p_4\}$.

PHY-Layer Error Detection

Observable

Let $\Lambda_j^i(k)$ be the magnitude of the LLR of k^{th} bit-channel, $k \in 1, 2, \ldots N$ at the output of the decoder at the end of j^{th} iteration such that $p_{guess} = p_j$ for $p_{channel} = p_i$.
$\Lambda_j^i(k)$ serve as the observables of the test after j^{th} iteration.



PHY-Layer Error Detection

Proposed tests

Let there be K good bit-channels after polarization.

Test 1: All good channels are above a given threshold.

Initially, a test was considered where $p_{guess} = p_{channel}$ is declared if the $|LLR|$ of all the good channels clear a given threshold under the current guess. That is, after j^{th} iteration,

$$j = i, \begin{cases} \text{if, } \Lambda_j^i(k) > \lambda, \forall k \in 1, 2, 3 \ldots K \\ \text{alternatively, } \min_{k \in [K]} \Lambda_j^i(k) > \lambda \end{cases}$$

$$j < i, \text{ o.w.}$$

The support of the distributions of $\Lambda_{j=i}^i(k)$ and $\Lambda_{j \neq i}^i(k)$ overlap considerably and $\Lambda_{j=i}^i(k)$ has a higher variance. This gave rise to high missed detection probability P_M[§]

[§] P_M indicates the probability that the test declares a better channel as the true channel.

PHY-Layer Error Detection

Proposed tests

Test 2: A given fraction of good channels are above a threshold.

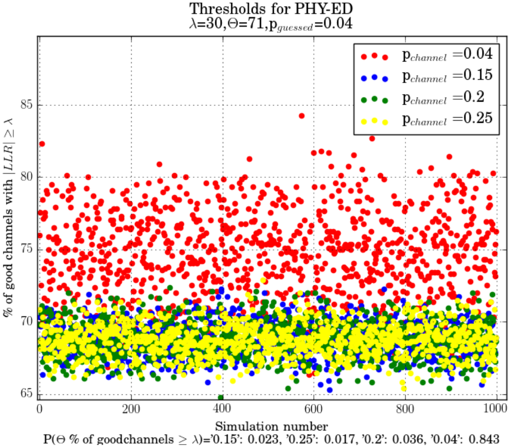
In this test $p_{guess} = p_{channel}$ is declared if the $|LLR|$ of a certain fraction of the good channels clear a threshold. The fraction is dependent on the iteration number. After j^{th} iteration,

$$j = i, \text{ if, } \frac{1}{K}\sum_{k=1}^{K} \mathbb{1}_{\{\Lambda_j^i(k) > \lambda\}} > \Theta_j$$

$$j < i, \text{ o.w.}$$

PHY-Layer Error Detection

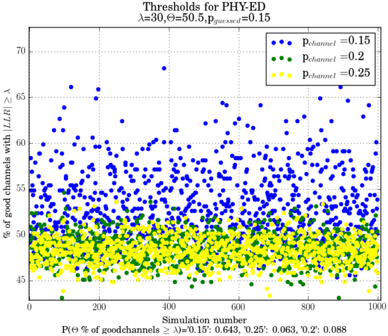
P_M and P_F for Test 2

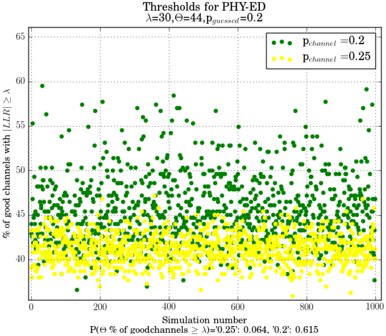


From figure, the detection error probabilities for the first iteration can be calculated as $P_F = 0.16$ and $P_M = 0.06$.

PHY-Layer Error Detection

- A higher value of P_M affects frame error probabilities adversely.
- A higher value of P_F does not affect the frame error probabilities but causes rate loss.

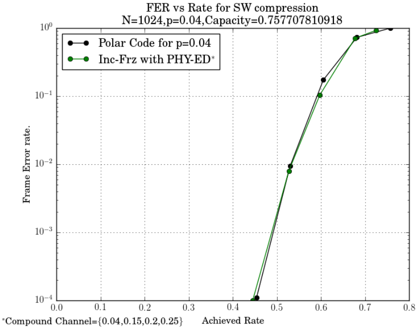


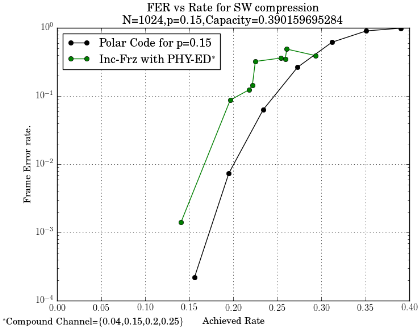


$P_F \approx 0.4$ for second and third iteration with $P_M \approx 0.13$.

Performance Evaluation

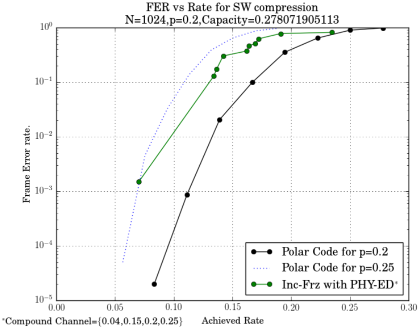
Let K be the number of bit-channels assumed to be good at the first iteration. For simulation, K is varied from 0 to K^* such that K^*/N is the capacity of the best channel. In case of SW compression the number of bits that remain unfrozen after the final iteration divided by N is viewed as the achieved rate.

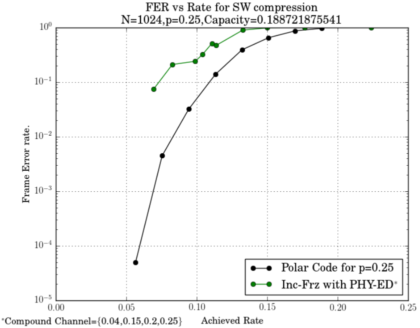




EC-performance

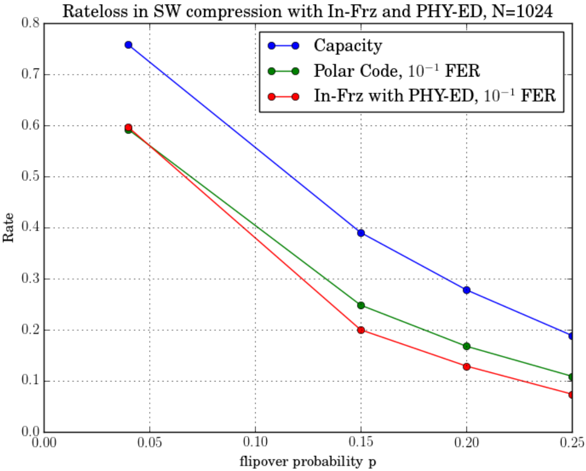
Performance Evaluation





EC-performance

Performance Evaluation



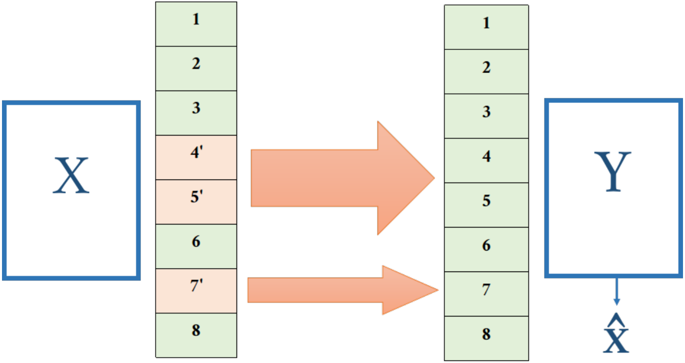
EC-performance

Conclusion and Future work

- The proposed scheme is an implementable solution to the *Data-Exchange* problem.
- It reduces the communication among nodes.
- The CRC-free universal polar code promises considerable rate gain for communication using short packet lengths.
- There are few channels which are good for use during the entire transmission. Communicating critical data over these channel ensure high reliability and availability .
- Future work.
  - Extensive performance analysis and theoretical analysis of proposed error detection scheme as a RB-HARQ for Polar Codes.
  - Implementation of the scheme for multiparty data exchange.

Thank You!

The r-sync protocol

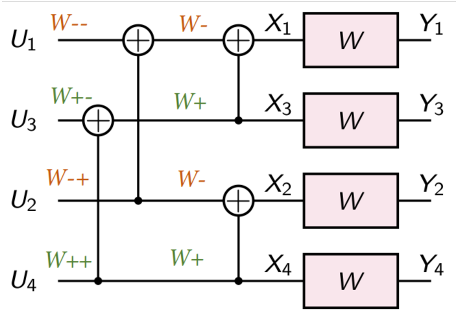


- Files are divided into blocks and there hashes are compared.
- The blocks with hash mismatch are transmitted.

For details see A. Tridgell and P. Mackerras. The r-sync algorithm. Joint Computer Science and Technical Report Series, TR-CS-96-05, 1996.

back

Polar Encoding



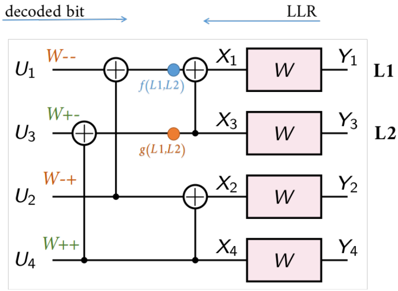
$$X_N = U_N * F^{\otimes n}, \text{where } N = 2^n$$

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

$$\text{set } U_1 = U_2 = 0$$

back

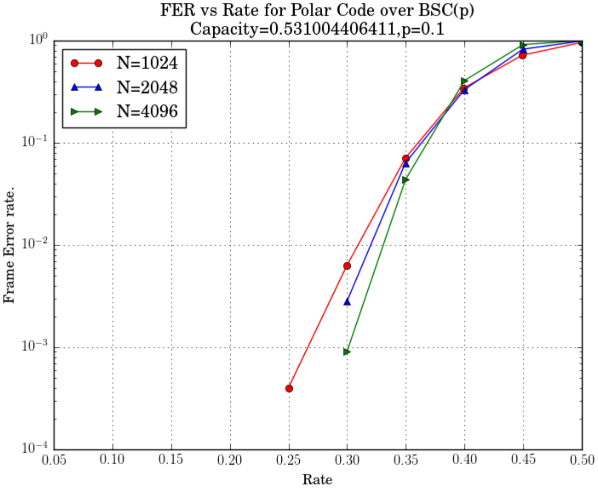
Succesive Cancellation Decoding



- Worse channels are decoded first.
- $f(L_1, L_2) = \frac{L_1 * L_2 + 1}{L_1 + L_2}$
- $g(L_1, L_2) = L_1 * L_2$, if bit at f is '1', else L_2/L_1
- if bit is frozen, frozen value is used, else decoding is based on f or g.

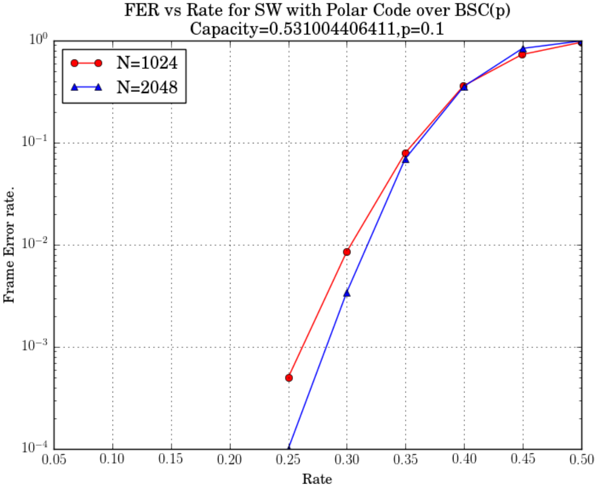
back

Performance of Polar Codes for error control



back

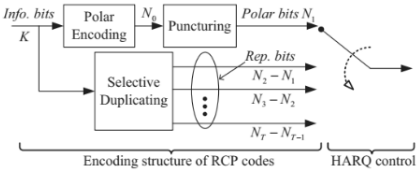
Performance of SW compression Polar Codes



back

Selective Repetition H-ARQ for polar codes

- Initially, 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.



- Retransmission process
  - On decoding failure receiver sends a NACK.
  - $N_2 - N_1$ of the information bits are retransmitted.
  - The receiver tries to perform decoding with all the N_2 received bits.
  - This process continues until the transmitter receives an ACK

The retransmitted bits (RV) are chosen one at a time as the most unreliable of the K bits transmitted, reliability is calculated after choosing one bit and the process is iterated. Note, $N_0 > \ldots N_3 > N_2 > N_1$.

back

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.
- This has the better performance compared to other HARQ methods.

back

Reliability based HARQ

Reliability based HARQ technique (RBHARQ) , eliminates the use of CRC by approximating bit and word error probability from likelihood ratios (LLR). The bit error probability for the k^{th} bit can be estimated from LLR $(\tilde{u}_k)$ as,

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

then word error probability becomes,

$$P_w = 1 - e^{log \bar{P}_w} \tag{2}$$

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 increases throughput, particularly evident in case of short packet lengths.

back

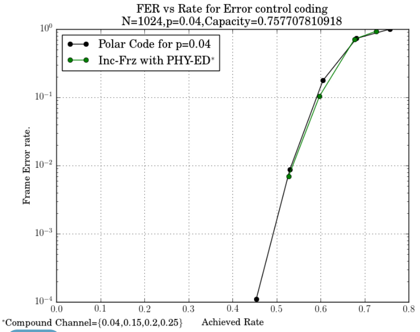
Rate Compatible Codes

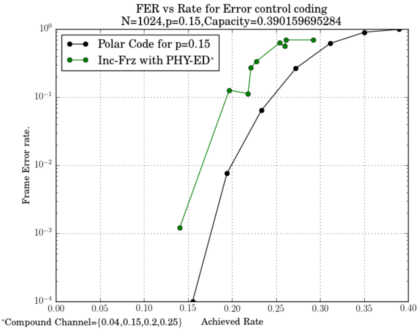
- Given a fixed number of information bits, consider a family of codes $\{C_1, C_2 ... C_n\}$ with rates $R_1 \geq R_2 \geq R_3 ... \geq R_n$, and block lengths $N_1 \leq N_2 \leq ... \leq N_n$. Then the set is rate compatible if codeword for C_i can be built by removing $N_j - N_i$ bits from codewords of code C_j , $j \geq i$, .
- Rate Compatible Codes can be constructed by puncturing low rate codes.

back

Performance Evaluation for Inc-Frz/PHY-ED error control coding

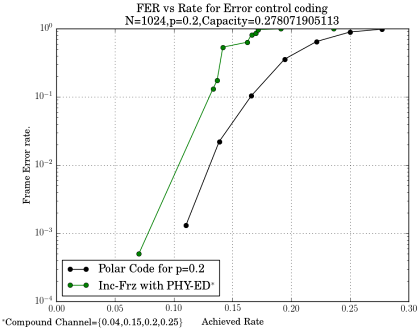
Let K be the number of bit-channels assumed to be good at the first iteration. For simulation, K is varied from 0 to K^* such that K^*/N is the capacity of the best channel. The scheme uses iterations to communicate these K bits, thus achieving some rate and corresponding frame error rate (FER).

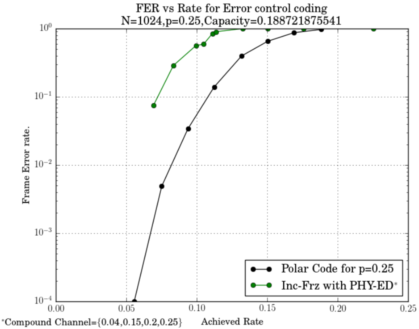




back

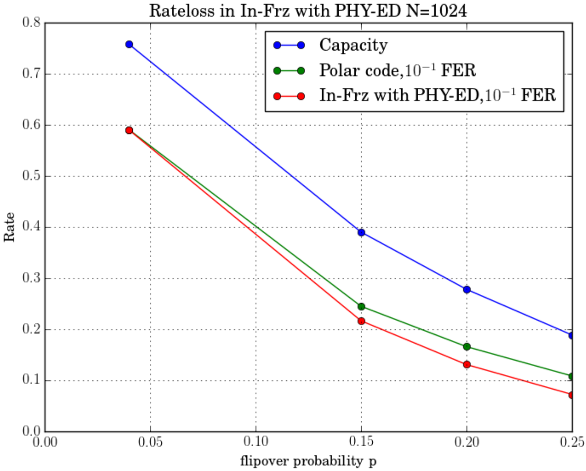
Performance Evaluation for Inc-Frz/PHY-ED error control coding





back

Performance Evaluation for Inc-Frz/PHY-ED error control coding



back

Incremental Freezing

continued...

- Features
  - By decoding the bits from future transmissions they effectively become frozen.
  - This scheme is capacity achieving in the sense that no rate has been wasted.¶
  - A certain number of channels in this scheme is "*always available*" guaranteeing a certain rate in each transmission.
  - n iterations of the scheme is almost equivalent in performance to a R/n fixed rate Polar Code.

¶, Figure illustrates the scheme for a set of channels with rates $\{R_1 = 6/8, R_2 = R_1/2 = 3/8, R_3 = R_1/3 = 1/4\}$. After the 3^{rd} transmission u_2 to u_5 have been incrementally frozen. The final rate achieved is, $R^* = \frac{6}{8*3} = \frac{1}{4} = R_3$