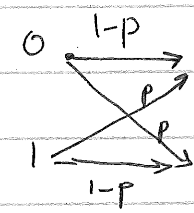
Tue Feb 19

Today: | Concatenated Codes
| Start Polar Codes.

Concatenated Codes

Let us consider BSC(p)

$C = 1 - h(p)$.



$0 \le p < 1/2$.

Shannon's Theorem implies: (after modifications)

Fix p, $\varepsilon > 0$, $\forall\, n \geq \Omega(1/\varepsilon^2)$,

$\exists$ linear code given by generator matrix

$G \in \{0,1\}^{n \times k}$ s.t. $k = (1 - h(p) - \varepsilon)n$,

$\forall u$, $\Pr_{\text{noise}}\left[G.u \text{ is closest to } BSC_p(Gu) \right] \geq 1 - 2^{-\Omega(\varepsilon^2 n)}$.

Concatenated Codes: Use Shannon's existential code for small lengths & Combine with explicit longer Codes.
"Outer"

We take $C_{in} \subseteq \{0,1\}^{n_0}$, also linear.

C_{in}: [ ]

↓ "adversarially" corrupt up to γn_0 bits.
but γ small.

$y = c + e$ [ ]

↓ Decoder

Correct c.

Such explicit codes are known. Assume we know it.
(i.e., deterministic poly(n_0) time construction).

The rate of C_w can be $1-\delta(\gamma)$, and it can correct worst-case γn_0 errors. ($\delta(\gamma) \to 0$ as $\gamma \to 0$).

We can take $\delta(\gamma) = \gamma^{1/3}$.

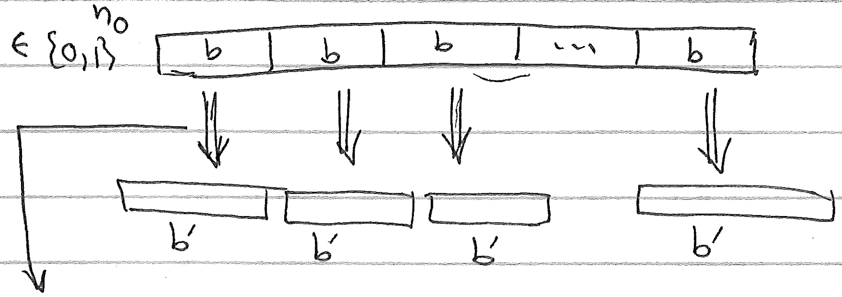
$$\underset{\text{message}}{m} \longrightarrow \boxed{\text{Encoder}} \longrightarrow c \in C_w$$

if $|m| = k \Rightarrow |c| \leq \dfrac{k}{1-\sqrt[3]{\delta}}$.

One possibility would be to re-encode c with a Shannon BSC(P) code. But that's too inefficient.

We decompose the task into blocks, and encode each block with Shannon code.

$c \in \{0,1\}^{n_0}$ $\qquad b \approx O\left(\frac{1}{\varepsilon^2} \log \frac{1}{\gamma}\right)$



rate = $1 - h(p) - \varepsilon/2$.

(3)

Find the "inner" Shannon code by brute force in time $2^{\text{poly}(1/\varepsilon)}$.

Overall rate $= (1-\sqrt[3]{\gamma})(1-h(p)-\varepsilon/2)$

$\geq 1-h(p)-\varepsilon$ for γ small enough.

* Decoding: (1) Decode inner codes first by brute force.

By Shannon's theorem, for each block of b' bits, the error probability is small.

(2) Decode the resulting string by the outer code decoder.

* Runtime $= \text{poly}(n_0 \cdot 2^b) + \text{poly}(n_0) = \text{poly}(n_0\, 2^{1/\varepsilon^2})$.

* Correctness: Need to make sure that the output of (1) has $\leq \gamma n_0$ errors. (w.h.p).

* For each block, $\Pr(\text{error}) \leq 2^{-\Omega(\varepsilon^2 b')} \leq \gamma^2$.

with constant in b large enough.

* The errors between blocks are independent $\Rightarrow$ We can use Chernoff's bound to make sure that the overall error fraction is $\leq \gamma$ w.h.p. $\square$

* Concatenated Codes is by Forney '1966.

Polar Codes.

Let us focus on BEC_α, $Cap = 1-\alpha$.

$k = R \cdot n \approx (1-\alpha) N$.

Say G is capacity achieving.

$$\begin{array}{cc} & k \\ N & \begin{bmatrix} & & \\ & G & \\ & & \end{bmatrix} \end{array} \cdot \begin{bmatrix} u_0 \\ | \\ u_{k-1} \end{bmatrix} = \begin{bmatrix} x_0 \\ | \\ x_{N-1} \end{bmatrix} \xrightarrow{BEC_\alpha} \begin{bmatrix} y_0 \\ | \\ y_{N-1} \end{bmatrix}$$

$U :=$ r.v. denoting $\underline{u}$.

$X :=$ " " $\underline{x}$.

Claim: Because G allows recovery of $\underline{u}$ from $\underline{y}$ whp, $H(U_0 - U_{k-1} \mid \underbrace{Y_0 - Y_{N-1}}_{=: Y_0^{N-1}}) \to 0$.

Chain rule:

$$H(U_0 \mid Y_0^{N-1}) + H(U_1 \mid U_0, Y_0^{N-1}) + \cdots + H(U_{k-1} \mid U_0^{k-2}, Y_0^{N-1}) \longrightarrow 0.$$

Now, think of an invertible matrix

$$G_N := \begin{array}{c} \\ N \end{array}\left[\begin{array}{c|c} \overset{N-k}{\longleftrightarrow} & \overset{k}{\longleftrightarrow} \\ A & G \end{array}\right]$$

$rk(G) = k$.

Say: Columns of A + " " G = basis of $\{0,1\}^N$.

Note: G_N is invertible. Now,

$$\begin{bmatrix} G_N \end{bmatrix} \begin{bmatrix} \underline{u} \end{bmatrix} \longrightarrow \begin{bmatrix} \underline{x} \end{bmatrix} \xrightarrow{BEC_\alpha} \begin{bmatrix} \underline{y} \end{bmatrix}.$$

$$H(u_0^{N-1} \mid y_0^{N-1}) = H(X_0^{N-1} \mid Y_0^{N-1}) \underset{\text{independence}}{=} N \cdot H(X_0 \mid Y_0)$$

$= N \cdot \alpha$. Chain rule again:

$$H(u_0 \mid y_0^{N-1}) + H(u_1 \mid u_0\, y_0^{N-1}) + \cdots + H(u_{N-1} \mid u_0^{N-2}\, y_0^{N-1})$$

$= \alpha \cdot N$.

We claim the way the sum adds up to αN is that an α fraction is close to 1 and the rest are close to 0.

6

Claim:

$$H\left(U_{N-k}^{N-1} \,\middle|\, Y_0^{N-1}, U_0^{N-k-1}\right) \longrightarrow 0.$$

Easy to see from the fact that G is a good code.

$\Rightarrow$ The last k terms in the summation are nearly 0.

$\Rightarrow$ The rest are nearly 1.

This is called Polarization!

This works for <u>any</u> memoryless channel and <u>any</u> good code G.

This is due to Arikan. (2009).
