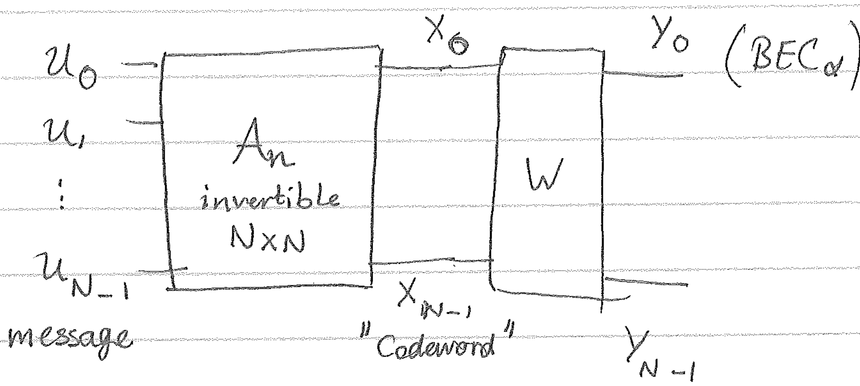
Tue 2/26

# Polar Codes, continued...

## Recap:




$$A_n = M_n^{-1} \text{ (from the last lecture)}$$

$$= G_2^{\otimes n} \cdot B_n$$

$$X_0^{N-1} = A_n U_0^{N-1}$$

Chain Rule: $\sum H(U_i \mid U_0^{i-1}, Y_0^{N-1})$

$$= \sum_{i=0}^{N-1} \underbrace{H(X_i \mid Y_i)}_{=\alpha} = N\alpha .$$

Now why does polarization occur for our matrix?

Last Class: $\alpha_n : \{0, \ldots, 2^n - 1\} \to [0,1]$

Uncertainty: $\alpha_n(i) \triangleq H(u_i \mid u_0^{i-1}, y_0^{N-1})$

$$= \mathrm{Prob}(u_i = ? \mid u_0^{i-1}, y_0^{N-1}).$$

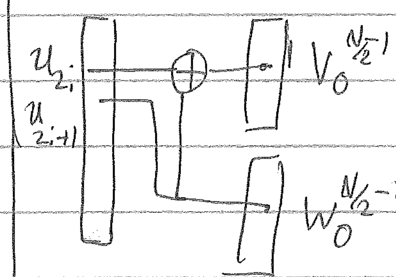
We proved that

$\alpha_n(\cdot)$ evolves according to recurrence:

$$\begin{cases} \alpha_n(i) = 2\,\alpha_{n-1}\left(\left\lfloor \tfrac{i}{2} \right\rfloor\right) - \alpha_{n-1}^2\left(\left\lfloor \tfrac{i}{2} \right\rfloor\right) & \text{if } i \text{ even} \\ \alpha_n(i) = \alpha_{n-1}^2\left(\left\lfloor \tfrac{i}{2} \right\rfloor\right) & \text{if } i \text{ odd} \end{cases}$$

$u_{2i} = V_i \oplus W_i$

$u_{2i+1} = W_i$




We wish to understand this recurrence.

Think of $\alpha_n$ as a random var.

on $\{0 \ldots 2^n - 1\}$, or, $\{0,1\}^n$ (induced by the uniform dist. on $\{0,1\}^n$).

$\alpha$
- $\alpha^2$
  - $\alpha^4$ — —
  - $2\alpha^2 - \alpha^4$ — — —
- $2\alpha - \alpha^2$
  - $(2\alpha - \alpha^2)^2$ — — —
  - $2(2\alpha - \alpha^2) - (2\alpha - \alpha^2)^2$ — — —

Define a r.v. $Z_n$ on $\{0,1\}^n$:

$$Z_n(b) \triangleq \alpha_n\Big(\underbrace{i(b)}_{\text{integer with } b \text{ as the binary representation.}}\Big) \in [0,1].$$

With uniform dist. on $\{0,1\}^n$.

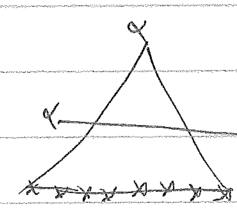
Suppose we know $Z_n$. Then,

$$Z_{n+1} = \begin{cases} 2Z_n - Z^2 & \text{w.p. } 1/2 \\ Z^2 & \text{w.p. } 1/2. \end{cases}$$

$$\Rightarrow \mathbb{E}(Z_{n+1} \mid Z_n) = Z_n.$$

$$\Rightarrow \mathbb{E}(Z_n) = \mathbb{E}(Z_0) = \alpha.$$

$\Rightarrow$ at every level, the avg. is $\alpha$.



$\Rightarrow$ The sequence $Z_0, Z_1, \ldots, Z_n, \ldots$ is a martingale.

Def (Martingale): A seq of r.v's $Z_0, Z_1, \ldots$ such that

④

$$\begin{cases} 1) & \mathbb{E}\left(|Z_n|\right) < \infty \quad \text{(trivial for us)} \\ 2) & \mathbb{E}\left(Z_{n+1} \mid Z_0 Z_1 \cdots Z_n\right) = Z_n \end{cases}$$

(also true for us)

$\Rightarrow$ Our $\{Z_i\}$ indeed form a martingale.

Since our r.v's are bounded in $[0,1]$, the $\{Z_n\}$ converge, by virtue of the so-called "Martingale Convergence Theorem".

The convergence is to a r.v. that is invariant under our operation. That is, constant 0 or constant 1.

That is, $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \left[ \Pr\left(\left|Z_{n+1}(\omega) - Z_n(\omega)\right| > \varepsilon\right]\right) = 0.$$

$$\Rightarrow \mathbb{E}\left[\,\left|Z_{n+1} - Z_n\right|\,\right] \xrightarrow{n\to\infty} 0.$$

$$
\begin{aligned}
&\mathbb{E}\left[\,|Z_{n+1}-Z_n|\,\right] \\
&= \tfrac{1}{2}\mathbb{E}\left[2Z_n - Z_n^2 - Z_n\right] + \tfrac{1}{2}\mathbb{E}\left[Z_n - Z_n^2\right] \\
&= \mathbb{E}\left[Z_n - Z_n^2\right] \xrightarrow[n\to\infty]{} 0 \\
&\Rightarrow \mathbb{E}\left(Z_n(1-Z_n)\right) \xrightarrow[n\to\infty]{} 0 \\
&\Rightarrow \forall \varepsilon>0,\ \lim_{n\to\infty} \Pr\left(Z_n \in (\varepsilon, 1-\varepsilon)\right) = 0. \\
&\Rightarrow Z_n \text{ is} \approx 0 \text{ or } 1 \text{ w.h.p.}
\end{aligned}
$$

$$
\& \ \mathbb{E}(Z_n) = \mathbb{E}(Z_0) = \alpha.
$$

$$
\Rightarrow \begin{cases} \Pr(Z_n \approx 0) = 1-\alpha \\ \Pr(Z_n \approx 1) = \alpha. \end{cases}
$$

In fact, $\forall \varepsilon > 0$, one can prove

$$
\lim_{n\to\infty} \Pr\Big(Z_n < \underbrace{2^{-2^{0.49n}}}_{2^{-N^{0.49}}}\Big) = 1-\alpha.
$$

<u>Comment:</u> The $0.49n \approx n/2$ is because we square r.v. about $n/2$ times.

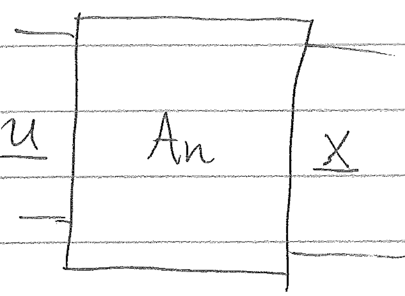
<u>Corollary:</u> $\forall$ large enough $n$,

$\exists\, F_n \subseteq \{0, \ldots, 2^n - 1\}$, $|F_n| \leq (\alpha + o(1))N$,

s.t. $\alpha_n(i) < \underbrace{2^{-2^{0.49n}}}_{= 2^{-N^{0.49}}}$ $\forall i \notin F_n$

("bad" bits)

* Obtaining a code from Polarization:




<u>Idea:</u> Freeze $u_i$, $i \in F_n$ to $0$.

$\Rightarrow$ We get a code of rate $1 - \alpha - o(1)$.

Formally, we get a Linear Code

$$C_n = \left\{ A_n \underline{u} \;\middle|\; \begin{array}{l} \underline{u} \in \{0,1\}^N, \\ u_i = 0,\ \forall i \in F_n \end{array} \right\}$$

⑦

**Decoding:** (Algorithm for $BEC_\alpha$)

For $i = 0, \ldots, 2^n - 1$

(i) If $i \in F_n$, $u_i = 0$.

(ii) If $i \notin F_n$, recover

$u_i$ from $u_0^{i-1}$, $y_0^{N-1}$ if possible.

If not, FAIL. (Code linear $\Rightarrow$ random $\Leftrightarrow$ fixed)

random

**Claim:** $\Pr($ FAIL for any ~~fixed~~ $\underline{u}$ under $BEC_\alpha)$

$$\leq \sum_{i \in \overline{F_n}} H(u_i \mid u_0^{i-1}, y_0^{N-1})$$

$$\leq \sum_{i \in \overline{F_n}} \alpha_n(i) \leq N \cdot 2^{-N^{0.49}} \xrightarrow[(n \to \infty)]{} 0.$$

$\square$

__Comment:__ Decoding is efficient (at least easy to see for $BEC_\alpha$).

* For $BEC_\alpha$, the whole $\alpha_n(i)$ tree is easy to compute $\Rightarrow$ $F_n$ is easy to compute.

## General Channels?

Recursive setup & $A_n$ remain the same. ($F_n$ will depend on the channel).

Recovery of $U_{2i}, U_{2i+1}$ from $V_0^{i-1}, W_0^{i-1}, Y_0^{N-1}$:

$$\beta := H(U_{2i} \mid \underbrace{V_0^{i-1}, W_0^{i-1}}_{U_0^{2i-1}}, Y_0^{N-1})$$

We know $\beta$ is small $\Rightarrow$ Can estimate each bit from prev. ones.

Similarly, $\gamma := H(U_{2i+1} \mid V_0^{i-1}, W_0^{i-1}, Y_0^{N-1})$

$H(V_i \mid V_0^{i-1}, Y_0^{\frac{N}{2}-1})$ these are iid r.v's.

$H(W_i \mid V_0^{i-1}, Y_{N/2}^{N-1})$ with entropies $\beta$ & $\gamma$.

By conservation of entropy,

$\beta + \gamma = 2\alpha$. always, one is $\geq \alpha$
other is $\leq \alpha$

$\Downarrow$

polarization occurs.

Lemma: (Pf omitted)

let $(B_1, D_1)$ & $(B_2, D_2)$ be iid pairs of discrete r.v's, $B_1, B_2 \in \{0,1\}$,

Then if $H(B_i \mid D_i) \in (\delta, 1-\delta)$ for some $\delta > 0$,

then $H(B_1 + B_2 \mid D_1, D_2) - H(B_1, D_1)$

$\geq \gamma(\delta) > \varepsilon$.

Assuming Lemma, $\beta > \alpha$, $\gamma < \alpha$.

But we no longer have explicit equations....

To get Quantitative bounds that imply correct decoding (whp), we compute not $H(V_i | \cdots)$ but the so-called "Bhattacharya parameters" of these channels.

$$Z(W) \triangleq \sum_{y \in \mathcal{Y}} \sqrt{p(y|0)\,p(y|1)}$$

$$= \langle v_0, v_1 \rangle \text{ where } v_{\substack{0\\1}} = \left(\sqrt{p(y|\substack{0\\1})}\right)_y$$

for $BEC_\alpha$, $Z(BEC_\alpha) = \alpha$.

$\square$