# Analysis of Channel Polarization

A Theoretical Motivation towards Polar Codes

By Vallabh Ramakanth

# Background knowledge

- Probability Theory and Concentration Inequalities
- Information Theory
- Martingale Random Processes

# The Original Paper/Reference

Erdal Arıkan, *Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels,* IEEE Transactions On Information Theory, Vol. 55, No. 7, July 2009

# Definitions/Notation

$$W : \mathcal{X} \rightarrow \mathcal{Y} \quad W(y|x), x \in \mathcal{X}, y \in \mathcal{Y}.$$

$W$ is a transition probability map or **channel**.

$$I(W) \triangleq \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \frac{1}{2} W(y|x) \log \frac{W(y|x)}{\frac{1}{2}W(y|0) + \frac{1}{2}W(y|1)}$$

$I(W)$ is the symmetric capacity of the channel. Here, $x$ is uniformly distributed over $\{0, 1\}$. This parameter is strongly tied to the rate of transmission. $0 \leq I(W) \leq 1$.

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

$Z(W)$ is the Bhattacharya parameter of the channel. It measures the "reliability" of the channel. $0 \leq Z(W) \leq 1$.

$$W^N : \mathcal{X}^N \rightarrow \mathcal{Y}^N$$

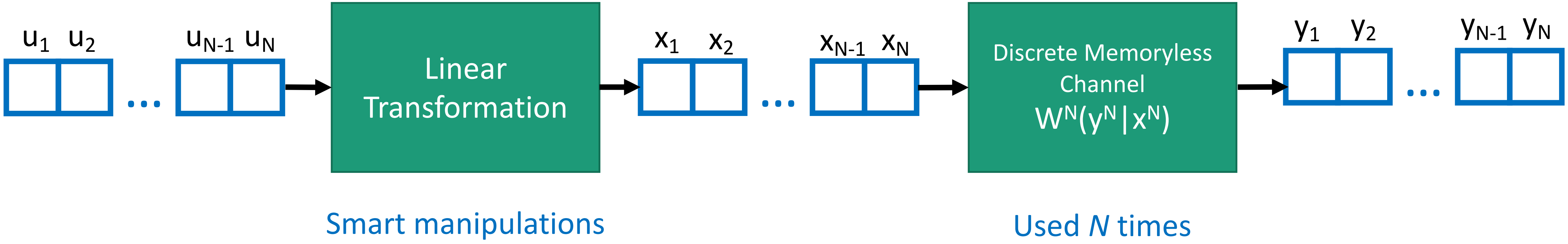
$$W^N(y^N|x^N) = \prod_{i=1}^{N} W(y_i|x_i)$$

$W^N$ is the equivalent channel when $W$ is used $N$ times independently.

$$y^N \coloneqq y_1^N \coloneqq (y_1, y_2, \dots, y_N),$$
$$x^N \coloneqq x_1^N \coloneqq (x_1, x_2, \dots, x_N)$$
$$x_a^b \coloneqq (x_a, \dots, x_b)$$

# Channel Polarization

- Use $W(y|x)$ independently $N$ times and **artificially manufacture** a new set of channels $W_N^{(i)}$ which are **polarized**, i.e., $I(W_N^{(i)})$ goes to either 0 or 1, $\forall i$ asymptotically.




- Channel polarization can visualized by breaking down the entire operation to two phases

1. Channel Combining
2. Channel Splitting

- For decoding, we look at

3. Successive Cancellation

# Channel Polarization

## 1. Channel Combining:

Consider the following definition of a channel $W_N : \mathcal{X}^N \to \mathcal{Y}^N$ , $W_N(y^N|x^N)$ by using independent copies of $W(y|x)$. $N = 2^k$.

$$W_1(y_1 \, |u_1) = W(y_1|u_1)$$

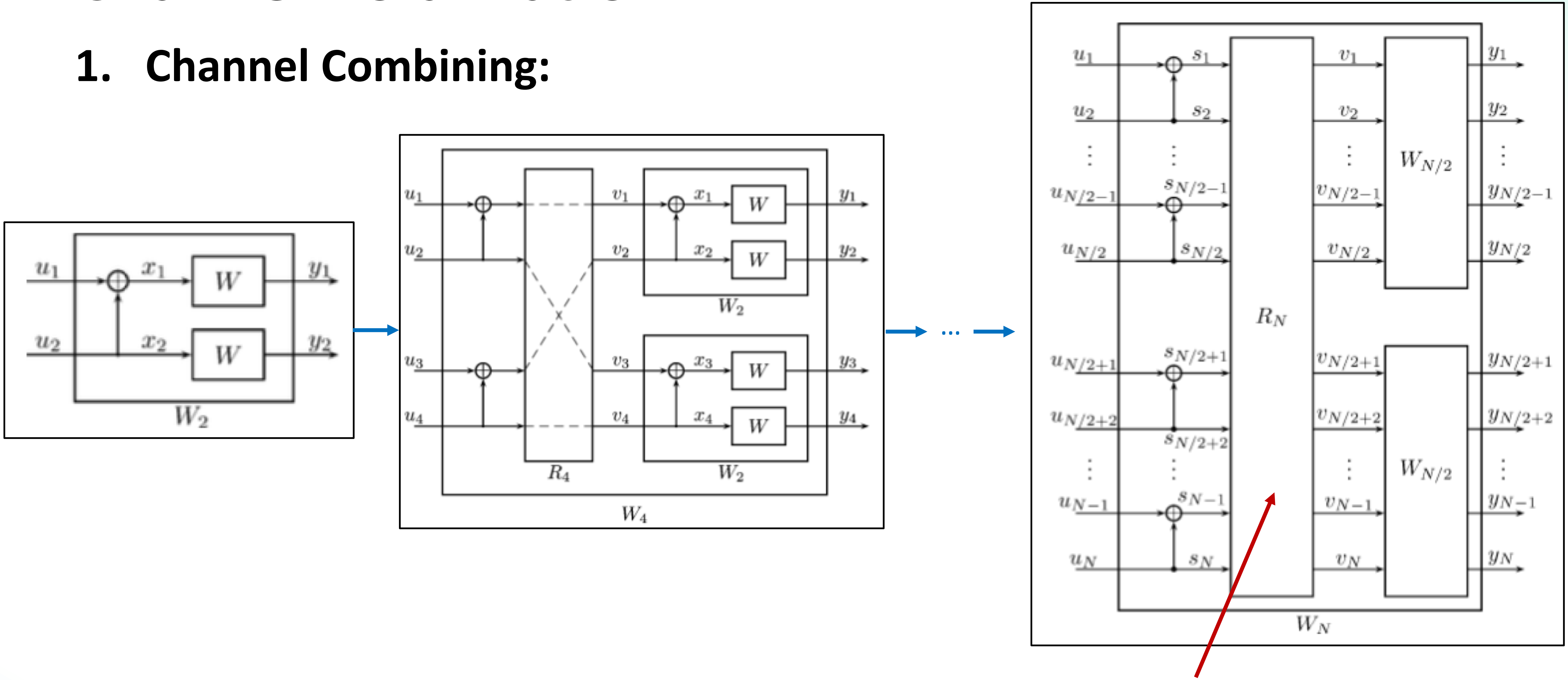
$$W_2(y_1, y_2|u_1, u_2) = W_1(y_1|u_1 \oplus u_2)W_1(y_2|u_2)$$

$$W_4(y_1, y_2, y_3, y_4|u_1, u_2, u_3, u_4) = W_2(y_2, y_1|u_1 \oplus u_2, u_3 \oplus u_4)W_2(y_3, y_4|u_2, u_4)$$

…

# Channel Polarization

## 1. Channel Combining:




**Reverse shuffle operation**: Group all odd terms in the first half and all even terms in the second half sequentially.

Recursive Construction of $W_N$

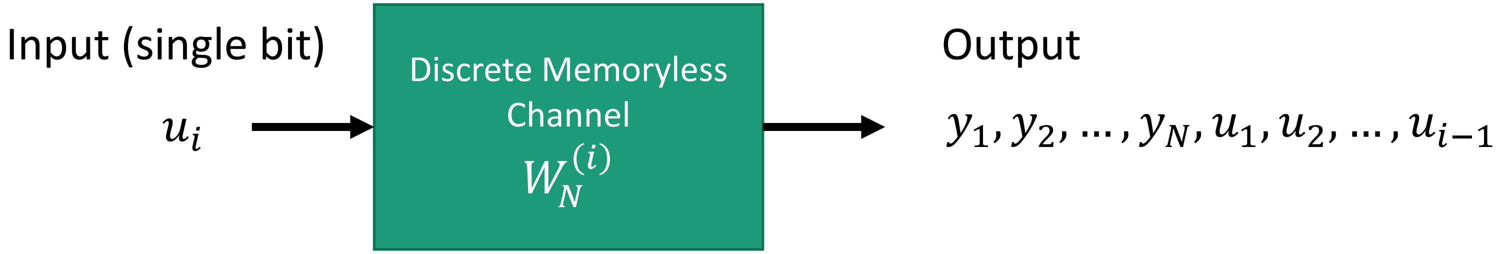
# Channel Polarization

## 2. Channel Splitting:

We now split $W_N$ channel into a set of individual single bit input channels $W_N^{(i)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^{i-1}, 1 \le i \le N$ defined as

$$W_N^{(i)}\left(y_1^N, u_1^{i-1} | u_i\right) \triangleq \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N\left(y_1^N | u_1^N\right)$$

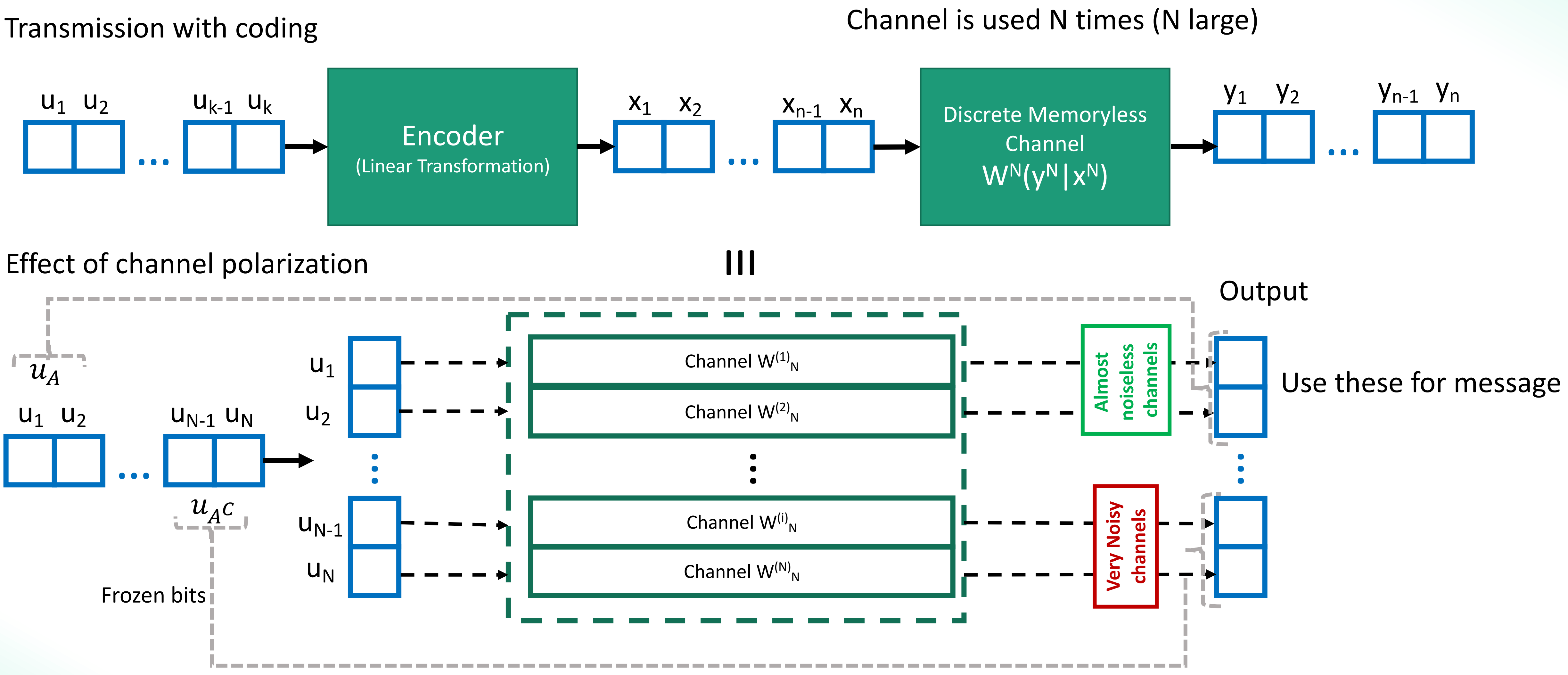
Summed over all "unknown" bits




**What we will try to do:**

Given $y_1, y_2, \dots, y_N, u_1, u_2, \dots, u_{i-1}$, try estimating $u_i$.
We will prove that this channel is polarized, i.e., its capacity is very close to either 1 or 0.

# **Polar codes**: Intuition




Asymptotically

# Decoding

## 3. Successive Cancellation Decoder

$$\hat{u}_i \triangleq \begin{cases} u_i, & \text{if } i \in \mathcal{A}^c \\ h_i\left(y_1^N, \hat{u}_1^{i-1}\right), & \text{if } i \in \mathcal{A} \end{cases}$$

$$h_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1} | 0\right)}{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1} | 1\right)} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

# Proofs

# Single step transformation of $(W, W)$ to $(W', W'')$

**Definition**:
Consider two independent copies of channel $W$. The following is a single step transformation

$$(W, W) \mapsto (W', W'')$$

Iff $\forall u_1, u_2 \in \mathcal{X}, y_1, y_2 \in \mathcal{Y}$

$$W'(y_1, y_2|u_1) = \frac{1}{2}\sum_{u_2'} W(y_1|u_1 + u_2{}')W(y_2|u_2{}')$$
$$W''(y_1, y_2, u_1|u_2) = \frac{1}{2}W(y_1|u_1 + u_2)W(y_2|u_2)$$

From this definition, we see $(W, W) \rightarrow \left(W_2^{(1)}, W_2^{(2)}\right)$

$$W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i\right) \triangleq \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N\left(y_1^N|u_1^N\right)$$

Can be generalized to $\left(W_N^{(i)}, W_N^{(i)}\right) \rightarrow \left(W_{2N}^{(2i-1)}, W_{2N}^{(2i)}\right)$ **(L1)**

# L1. Recursive Construction of $W_N^{(i)}$

$$W_{2N}^{(2i-1)}\left(y_1^{2N}, u_1^{2i-2} | u_{2i-1}\right) = \sum_{u_{2i}} \frac{1}{2} W_N^{(i)}\left(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2} | u_{2i-1} \oplus u_{2i}\right) \cdot W_N^{(i)}\left(y_{N+1}^{2N}, u_{1,e}^{2i-2} | u_{2i}\right)$$

$$W_{2N}^{(2i)}\left(y_1^{2N}, u_1^{2i-1} | u_{2i}\right) = \frac{1}{2} W_N^{(i)}\left(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2} | u_{2i-1} \oplus u_{2i}\right) \cdot W_N^{(i)}\left(y_{N+1}^{2N}, u_{1,e}^{2i-2} | u_{2i}\right)$$

i.e., $\left(W_N^{(i)}, W_N^{(i)}\right) \rightarrow \left(W_{2N}^{(2i-1)}, W_{2N}^{(2i)}\right)$

# L2. Channel Parameters after Single step transforms

If $(W, W\ ) \rightarrow (W', W'')$, then

$$I(W') + I(W'') = 2I(W)$$
$$I(W') \leq I(W'')$$

**Proof**:

Let $U_1, U_2$ be the inputs to the channels. Define $X_1 = U_1 + U_2$ and $X_2 = U_2$.
We send $(X_1, X_2)$, linear transformation of input vector $(U_1, U_2)$, through the channel $W^2$

$I(W') = I(Y_1, Y_2; U_1)$
$I(W'') = I(Y_1, Y_2, U_1; U_2) = I(U_1; U_2) + I(Y_1, Y_2; U_2|U_1) = I(Y_1, Y_2; U_2|U_1)$
$I(W') + I(W'') = I(Y_1, Y_2; U_1) + I(Y_1, Y_2; U_2|U_1) = I(Y_1, Y_2; U_1, U_2) = I(Y_1, Y_2; X_1, X_2)$
$\therefore, I(W') + I(W'') = 2I(W)$

Chain Rule of Mutual Information

Also, $I(W'') = I(Y_1, Y_2, U_1; U_2) = I(Y_2; U_2) + I(Y_1, U_1; U_2|Y_2) \geq I(W)$
$\Rightarrow I(W') \leq I(W)$

$$W'(y_1, y_2|u_1) = \frac{1}{2}\sum_{u_2'} W(y_1|u_1 + u_2{'})W(y_2|u_2{'})$$
$$W''(y_1, y_2, u_1|u_2) = \frac{1}{2}W(y_1|u_1 + u_2)W(y_2|u_2)$$

# L2. Channel Parameters after Single step transforms

If $(W, W\ ) \rightarrow (W', W'')$, then

$$Z(W'') = Z(W)^2$$
$$Z(W') \leq 2Z(W) - Z(W)^2$$
$$Z(W') \geq Z(W) \geq Z(W'').$$

**Proof**: First Equality

$$Z(W'') = \sum_{y_1,y_2,u_1} \sqrt{W''(y_1, y_2, u_1|0)W''(y_1, y_2, u_1|1)} = \frac{1}{2} \sum_{y_1,y_2,u_1} \sqrt{W(y_1\ |u_1)W(y_2|0)W(y_1|u_1 + 1)W(y_2|1)}$$

$$= \frac{1}{2} \sum_{y_1,u_1} \sqrt{W(y_1\ |u_1)W(y_1|u_1 + 1)} \sum_{y_2} \sqrt{W(y_2|0)W(y_2|1)}$$

$$= \frac{1}{2} \times (2Z(W)) \times (Z(W))$$

$$= Z(W)^2$$

$$\therefore Z(W'') \leq Z(W)$$

Second Equality can be shown to be true with simple algebraic identities.
Third inequality can be shown by exploiting the convex property of Z(W) and Minkowski's Inequality.

$$W'(y_1, y_2|u_1) = \frac{1}{2} \sum_{u_2'} W(y_1|u_1 + u_2{}')W(y_2|u_2{}')$$

$$W''(y_1, y_2, u_1|u_2) = \frac{1}{2} W(y_1|u_1 + u_2)W(y_2|u_2)$$

# Summary (Till Now)

$$W_{2N}^{(2i-1)}\left(y_1^{2N}, u_1^{2i-2} | u_{2i-1}\right) = \sum_{u_{2i}} \frac{1}{2} W_N^{(i)}\left(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2} | u_{2i-1} \oplus u_{2i}\right) \cdot W_N^{(i)}\left(y_{N+1}^{2N}, u_{1,e}^{2i-2} | u_{2i}\right)$$

$$W_{2N}^{(2i)}\left(y_1^{2N}, u_1^{2i-1} | u_{2i}\right) = \frac{1}{2} W_N^{(i)}\left(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2} | u_{2i-1} \oplus u_{2i}\right) \cdot W_N^{(i)}\left(y_{N+1}^{2N}, u_{1,e}^{2i-2} | u_{2i}\right)$$

Using **L1** and **L2**, we can summarize

**A**

$$\left(W_N^{(i)}, W_N^{(i)}\right) \to \left(W_{2N}^{(2i-1)}, W_{2N}^{(2i)}\right)$$

**B**

$$I\left(W_{2N}^{(2i-1)}\right) + I\left(W_{2N}^{(2i)}\right) = 2I\left(W_N^{(i)}\right)$$

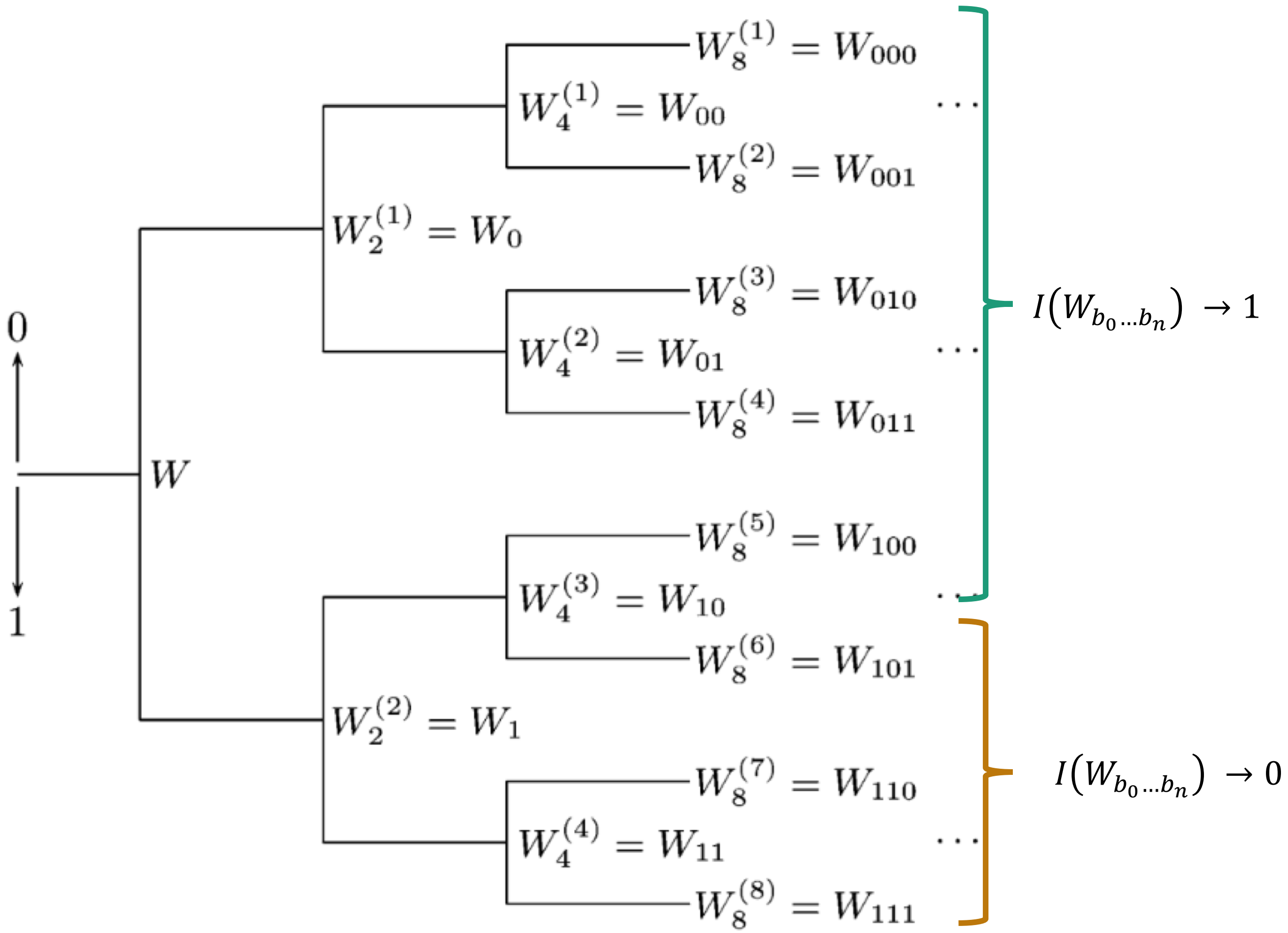
$$I\left(W_{2N}^{(2i-1)}\right) \le I\left(W_N^{(i)}\right) \le I\left(W_{2N}^{(2i-1)}\right)$$

**C**

$$Z\left(W_{2N}^{(2i)}\right) = Z\left(W_N^{(i)}\right)^2$$

$$Z\left(W_{2N}^{(2i-1)}\right) + Z\left(W_{2N}^{(2i)}\right) \le 2Z\left(W_N^{(i)}\right)$$

Binary/Power of 2 recursion




$$\left(W_N^{(i)}, W_N^{(i)}\right) \to \left(W_{2N}^{(2i-1)}, W_{2N}^{(2i)}\right)$$

Since $N = 2^n, W_N^{(i)} = W_{b_0,b_1,\dots,b_n}$ where

$$i = 1 + \sum_{r=0}^{n-1} b_r 2^r$$

# T1. Channel Polarization

For any B-DMC $W$, the channels $\left\{W_N^{(i)}\right\}$ polarize, i.e., for any $\delta > 0$, as $N = 2^n$ goes to $\infty$, a subset of indices $A \subset \{1, \dots, N\}$, $i \in A$, j $\in A^c$, $I\left(W_N^{(i)}\right) \in (1-\delta, 1]$ and $I\left(W_N^{(j)}\right) \in [0, \delta)$ with $\frac{|A|}{N} = I(W)$.

**Proof**:

Define $\{b_n\}_{n\geq 0}$ to be an i.i.d. Bernoulli random process, such that $b_n = 0$ or $1$ with equal probability $\frac{1}{2}$.
Then, $W_{b_0 b_1 \dots b_n}$ is a random process defined on the tree in the previous figure, with $W_0 = W$, the true B-DMC.
Moreover, $I_n \triangleq I(W_{b_0 b_1 \dots b_n})$ and $Z_n \triangleq Z(W_{b_0 b_1 \dots b_n})$ are defined random processes.

$$E[I_n | b^{n-1}] = E\left[I\left(W_{b_0 b_1 \dots b_n}\right) \middle| b^{n-1}\right] = \frac{1}{2} I\left(W_{b_0 b_1 \dots b_{n-1} 0}\right) + \frac{1}{2} I\left(W_{b_0 b_1 \dots b_{n-1} 1}\right) = I\left(W_{b_0 b_1 \dots b_{n-1}}\right)$$

because $I\left(W_{2N}^{(2i-1)}\right) + I\left(W_{2N}^{(2i)}\right) = 2I\left(W_N^{(i)}\right)$

Hence $I_n$ is a **bounded martingale process**, as $0 \leq I_n \leq 1$.
All moments of $I_n$ exist!

From Martingale convergence, we have that $I_\infty$ is a well defined random variable and $E|I_\infty - I_n| < \infty$.

# T1. Channel Polarization

For any B-DMC $W$, the channels $\left\{W_N^{(i)}\right\}$ polarize, i.e., for any $\delta > 0$, as $N = 2^n$ goes to $\infty$, a subset of indices $A \subset \{1, \dots, N\}$, $i \in A$, j $\in A^c$, $I\left(W_N^{(i)}\right) \in (1-\delta, 1]$ and $I\left(W_N^{(j)}\right) \in [0, \delta)$ with $\frac{|A|}{N} = I(W)$.

**Proof**:

$$E[Z_n|b^{n-1}] = E\left[Z\left(W_{b_0 b_1 \dots b_n}\right)\middle|b^{n-1}\right] = \frac{1}{2}Z\left(W_{b_0 b_1 \dots b_{n-1} 0}\right) + \frac{1}{2}Z\left(W_{b_0 b_1 \dots b_{n-1} 1}\right) \leq Z\left(W_{b_0 b_1 \dots b_{n-1}}\right)$$

because $Z\left(W_{2N}^{(2i-1)}\right) + Z\left(W_{2N}^{(2i)}\right) \leq 2Z\left(W_N^{(i)}\right)$.

Hence $Z_n$ is a **bounded supermartingale process**, as $0 \leq Z_n \leq 1$.
All moments of $Z_n$ exist!

From Martingale convergence, we have that $Z_\infty$ is a well defined random variable and $E|Z_\infty| < \infty$.
Since $Z_n = \sum_{i=1}^{n}(Z_i - Z_{i-1})$, and as $E|Z_n|$ converges, we have $E|Z_{n+1} - Z_n| \to 0$.

But $Z_{n+1}$ is $Z\left(W_{b_0 b_1 \dots b_n 0}\right) = Z_n^2$ with probability $\frac{1}{2}$ because $Z\left(W_{2N}^{(2i)}\right) = Z\left(W_N^{(i)}\right)^2$.
Hence, $E|Z_{n+1} - Z_n| \geq \frac{1}{2}E[Z_n^2 - Z_n] = \frac{1}{2}E[Z_n(1 - Z_n)]$.

# T1. Channel Polarization

For any B-DMC $W$, the channels $\left\{W_N^{(i)}\right\}$ polarize, i.e., for any $\delta > 0$, as $N = 2^n$ goes to $\infty$, a subset of indices $A \subset \{1, \dots, N\}$, $i \in A$, $\mathrm{j} \in A^c$, $I\left(W_N^{(i)}\right) \in (1-\delta, 1]$ and $I\left(W_N^{(j)}\right) \in [0, \delta)$ with $\frac{|A|}{N} = I(W)$.

**Proof**:

Hence, $E|Z_{n+1} - Z_n| \geq \frac{1}{2}E[|Z_n^2 - Z_n|] = \frac{1}{2}E[|Z_n(1-Z_n)|]$.

As $E|Z_{n+1} - Z_n| \to 0$, we also have $E[|Z_n(1-Z_n)|] \to 0$, which implies $\mathrm{Z}_n$ converges to either $0$ or $1$ almost surely!

As $Z_\infty = 0$ or $1$, we have $I_\infty = 1 - Z_\infty$. (Can see intuitively that $Z(W) = 0$ gives $I(W) = 1$ and vice-versa)

The above result is true whenever $Z(W) = 0$ or $1$.

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

But since $I_n$ is a martingale, $E[I_\infty] = I_0$ which immediately gives us

$$P(I_\infty = 1) = I_0 = I(W) \text{ and } P(I_\infty = 0) = 1 - I_0$$

■

# T2. Channel Polarization

For any B-DMC $W$, and any fixed rate $R < I(W)$, there exists a sequence of sets $A_N \subseteq \{1, \dots, N\}$, $N = 2^n$ such that $|A_N| \geq NR$ and $Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right)$ $\forall i \in A_N$.

**Proof**:

From the same setting as earlier, we have

$$\begin{cases} Z_{n+1} \leq Z_n^2, & b_n = 1 \left[\text{as } Z\left(W_{2N}^{(2i)}\right) = Z\left(W_N^{(i)}\right)^2\right] \\ Z_{n+1} \leq 2Z_n - Z_n^2 \leq 2Z_n, & b_n = 0 \left[\text{as } Z\left(W_{2N}^{(2i-1)}\right) \leq 2Z\left(W_N^{(i)}\right) - Z\left(W_N^{(i)}\right)^2\right] \end{cases}$$

For parameters $2 \geq \zeta \geq 0, m \geq 0$

$$T_m(\zeta) \triangleq \{\omega \in \Omega; Z_i \leq \zeta, \forall i \geq m\}$$

Then for $\omega \in T_m(\zeta)$ and $i \geq m$, we have

$$\frac{Z_{i+1}}{Z_i} \leq \begin{cases} 2, & b_n = 0 \\ \zeta, & b_n = 1 \end{cases}$$

This implies $Z_n \leq \zeta 2^{n-m} \prod_{i=m+1}^{n} \left(\frac{\zeta}{2}\right)^{b_i} = \zeta 2^{n-m} \left(\frac{\zeta}{2}\right)^{\sum_{i=m+1}^{n} b_i}$ for $n > m$

# T2. Channel Polarization

For any B-DMC $W$, and any fixed rate $R < I(W)$, there exists a sequence of sets $A_N \subseteq \{1, \dots, N\}$, $N = 2^n$ such that $|A_N| \geq NR$ and $Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right)$ $\forall i \in A_N$.

**Proof**:

For $n > m \geq 0$ and $0 < \eta < 1/2$, define,

$$U_{m,n}(\eta) \triangleq \left\{\omega \in \Omega: \sum_{i=m+1}^{n} b_i \geq \left(\frac{1}{2} - \eta\right)(n-m)\right\}$$

Then

$$Z_n(\omega) \leq \zeta \left[2^{\frac{1}{2}+\eta} \zeta^{\frac{1}{2}-\eta}\right]^{n-m}, \qquad \omega \in T_m(\zeta) \cap U_{m,n}(\eta)$$

Substitute $\eta_0 = 1/20$ and $\zeta_0 = 2^{-4}$ to get

$$Z_n(\omega) \leq 2^{-4 - \frac{5(n-m)}{4}}, \qquad \omega \in T_m(\zeta_0) \cap U_{m,n}(\eta_0)$$

We need to show that for a given $m, n$, $T_m(\zeta_0) \cap U_{m,n}(\eta_0)$ occurs with high probability.

# T2. Channel Polarization

For any B-DMC $W$, and any fixed rate $R < I(W)$, there exists a sequence of sets $A_N \subseteq \{1, \dots, N\}$, $N = 2^n$ such that $|A_N| \geq NR$ and $Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right)$ $\forall i \in A_N$.

**Proof**:
First consider $T_m(\zeta_0)$.
As seen from **T1**, $P(Z_\infty = 0) = I_0$, which implies that $S_n \coloneqq \{Z_n \leq \zeta_0\}$, $\{Z_\infty = 0\} \subseteq \cup_{n \geq m} S_n$ for large enough $m$, and hence $P(\cup_{n \geq m} S_n) \geq I_0 - \frac{\delta}{2}$.
$T_{m_0}(\zeta_0) = \cup_{n \geq m_0} S_n$, and from continuity of probability,

$$P\left(T_{m_0}(\zeta_0)\right) = P\left(\cup_{n \geq m_0} S_n\right) \geq I_0 - \frac{\delta}{2}, \qquad m_0 = m_0(\zeta_0, \delta)$$

Now consider $U_{m,n}(\eta)$.

$$P\left(U_{m,n}^c(\eta_0)\right) = P\left(\sum_{i=m+1}^{n} b_i < \left(\frac{1}{2} - \eta_0\right)(n-m)\right) = P\left(-t\sum_{i=m+1}^{n} b_i > -t\left(\frac{1}{2} - \eta_0\right)(n-m)\right)$$

$$\leq 2^{t\left(\frac{1}{2}-\eta_0\right)(n-m)} E[2^{-tb_i}]^{n-m} = \left[\left(\frac{2^{t\left(\frac{1}{2}-\eta_0\right)} + 2^{-t\left(\frac{1}{2}+\eta_0\right)}}{2}\right)\right]^{n-m}$$

Chernoff Bound

$$P\left(U_{m,n}^c(\eta_0)\right) \leq 2^{-(n-m)\left(1-H\left(\frac{1}{2}-\eta_0\right)\right)}$$

# T2. Channel Polarization

For any B-DMC $W$, and any fixed rate $R < I(W)$, there exists a sequence of sets $A_N \subseteq \{1, \dots, N\}$, $N = 2^n$ such that $|A_N| \geq NR$ and $Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right)$ $\forall i \in A_N$.

**Proof**:

$$P\left(T_{m_0}^c(\zeta_0)\right) \leq 1 - I_0 + \frac{\delta}{2}, \qquad m_0 = m_0(\zeta_0, \delta)$$

$$P\left(U_{m,n}^c(\eta_0)\right) \leq 2^{-(n-m)\left(1-H\left(\frac{1}{2}-\eta_0\right)\right)}$$

We can choose a finite $n_0(m_0, \eta_0, \delta)$ such that the RHS of the above inequality becomes at most $\frac{\delta}{2}$. Hence, from union bound, $\forall n > n_0$

$$P\left(T_{m_0}^c(\zeta_0) \cup U_{m_0,n}^c(\eta_0)\right) \leq 1 - I_0 + \delta$$

$$P\left(T_{m_0}(\zeta_0) \cap U_{m_0,n}(\eta_0)\right) \geq I_0 - \delta$$

Therefore

$$Z_n(\omega) \leq 2^{-4 - \frac{5(n-m_0)}{4}} = c2^{-\frac{5n}{4}}, \qquad \omega \in T_{m_0}(\zeta_0) \cap U_{m_0,n}(\eta_0), \forall n > n_0$$

# T2. Channel Polarization

For any B-DMC $W$, and any fixed rate $R < I(W)$, there exists a sequence of sets $A_N \subseteq \{1, \dots, N\}$, $N = 2^n$ such that $|A_N| \geq NR$ and $Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right)$ $\forall i \in A_N$.

**Proof**:

Define $V_n = \left\{\omega\colon Z_n \leq c2^{-\frac{5n}{4}}\right\}$, then immediately $T_{m_0}(\zeta_0) \cap U_{m_0,n}(\eta_0) \subseteq V_n$ $\forall n > n_0$

$$P(V_n) \geq I - \delta = R \ \forall n > 0$$

Notice that

$$P(V_n) = \sum_{b_0,b_1,\dots,b_n} \frac{1}{2^n} \mathbf{1}_{\left\{Z_n \leq c2^{-\frac{5n}{4}}\right\}} = \frac{|A_N|}{N}$$

Therefore,

$$|A_N| \geq NR \text{ and } Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{5}{4}}\right) \ \forall i \in A_N$$

■

# T3. Probability of Decoding error

$$\hat{u}_i \triangleq \begin{cases} u_i, & \text{if } i \in \mathcal{A}^c \\ h_i\left(y_1^N, \hat{u}_1^{i-1}\right), & \text{if } i \in \mathcal{A} \end{cases}$$

$$h_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|0\right)}{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|1\right)} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

$P(\mathcal{E}) \leq \sum_{i=1}^{n} Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{1}{4}}\right)$, where, $\mathcal{E} \triangleq \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N : \widehat{U}_A(u^N, y^N) \neq u_A\right\}$ and $R < I(W)$.

**Proof**:

Define $B_i = \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N: \ \widehat{U}_1^{i-1} = u_1^{i-1} \ \ \& \ \ \widehat{U}_i \neq u_i\right\}$. Then, $\mathcal{E} = \cup_{i\in A} B_i$.

$$\begin{aligned} B_i &= \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N: \ \widehat{U}_1^{i-1} = u_1^{i-1} \ \ and \ \ h_i\left(y_1^N, \widehat{U}_1^{i-1}\right) \neq u_i\right\} \\ &= \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N: \ \widehat{U}_1^{i-1} = u_1^{i-1} \ \ and \ \ h_i\left(y_1^N, u_1^{i-1}\right) \neq u_i\right\} \\ &\subseteq \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N: h_i\left(y_1^N, u_1^{i-1}\right) \neq u_i\right\} \subseteq \mathcal{E}_i \end{aligned}$$

where $\mathcal{E}_i = \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N: \ W_N^{(i)}\left(y_1^N, u_1^{i-1} \middle| u_i\right) \leq W_N^{(i)}\left(y_1^N, u_1^{i-1} \middle| u_i + 1\right)\right\}$

Therefore,

$$P(\mathcal{E}) \leq P(\cup_{i\in A} \mathcal{E}_i) \leq \sum_{i\in A} P(\mathcal{E}_i)$$

We now need an upper bound on $P(\mathcal{E}_i)$.

# T3. Probability of Decoding error

$$\hat{u}_i \triangleq \begin{cases} u_i, & \text{if } i \in \mathcal{A}^c \\ h_i\left(y_1^N, \hat{u}_1^{i-1}\right), & \text{if } i \in \mathcal{A} \end{cases}$$

$$h_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|0\right)}{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|1\right)} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

$P(\mathcal{E}) \leq \sum_{i=1}^{n} Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{1}{4}}\right)$, where, $\mathcal{E} \triangleq \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N : \widehat{U}_A(u^N, y^N) \neq u_A\right\}$ and $R < I(W)$.

**Proof**:

$$P(\mathcal{E}_i) = \sum_{u^N, y^N} P_{U^N, Y^N}(u^N, y^N) \mathbf{1}_{\{(u^N, y^N) \in \mathcal{E}_i\}} = \sum_{u^N, y^N} \frac{1}{2^N} W_N(y^N|u^N) \mathbf{1}_{\{(u^N, y^N) \in \mathcal{E}_i\}}$$

$$\leq \sum_{u^N, y^N} \frac{1}{2^N} W_N(y^N|u^N) \sqrt{\frac{W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i+1\right)}{W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i\right)}}$$

Since we know an error has occurred

$$= \sum_{u_1^{i-1}, y^N} \sum_{u_i} \frac{1}{2} \left( \sum_{u_{i+1}^N} \frac{1}{2^{N-1}} W_N(y^N|u^N) \right) \sqrt{\frac{W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i+1\right)}{W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i\right)}}$$

$$= \sum_{u_i} \frac{1}{2} \sum_{u_1^{i-1}, y^N} \sqrt{W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i\right) W_N^{(i)}\left(y_1^N, u_1^{i-1}|u_i+1\right)} = Z\left(W_N^{(i)}\right)$$

# T3. Probability of Decoding error

$$\hat{u}_i \triangleq \begin{cases} u_i, & \text{if } i \in \mathcal{A}^c \\ h_i\left(y_1^N, \hat{u}_1^{i-1}\right), & \text{if } i \in \mathcal{A} \end{cases}$$

$$h_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|0\right)}{W_N^{(i)}\left(y_1^N, \hat{u}_1^{i-1}|1\right)} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

$P(\mathcal{E}) \leq \sum_{i=1}^{n} Z\left(W_N^{(i)}\right) \leq O\left(N^{-\frac{1}{4}}\right)$, where, $\mathcal{E} \triangleq \left\{(u^N, y^N) \in \mathcal{X}^N \times \mathcal{Y}^N : \widehat{U}_A(u^N, y^N) \neq u_A\right\}$ and $R < I(W)$.

**Proof**:

$$P(\mathcal{E}_i) \leq Z\left(W_N^{(i)}\right)$$

Hence,

$$\begin{aligned} P(\mathcal{E}) &\leq \sum_{i \in A} P(\mathcal{E}_i) \leq \sum_{i \in A} Z\left(W_N^{(i)}\right) \\ &\leq |A| \max\left(Z\left(W_N^{(i)}\right)\right) \\ &\leq N \max\left(Z\left(W_N^{(i)}\right)\right) \\ &\leq O\left(N^{-\frac{1}{4}}\right) \end{aligned}$$

■

# Conclusions

# Conclusions

- Polar codes can be asymptotically rate achieving codes for B-DMCs.

- For a block length of $N = 2^n$ and transmission rate $R < I(W)$, we see $Z\left(W_N^{(i)}\right) \sim N^{-\frac{5}{4}}$, for at least $NR$ channels. This gives us a rate at which channels polarize.

  We can get better bounds and tighter bounds than this.

- For a block length of $N = 2^n$ and transmission rate $R < I(W)$, average probability of error goes to zero asymptotically, as, $P(error) \sim N^{-\frac{1}{4}}$.

  This is a loose bound, and we can strengthen the upper bound to an exponential function of $N$.
  The upper bound also does not explicitly depend on $R$, and one can try to obtain a sharper bound which tells us how the error probability degrades with increasing rate.