Polarization Theorem

Channel polarization:
Theorem 1: For any B-DMC W, the channels $\{W_N^{(i)}\}$ polarize in the sense that, for any fixed $\delta \in (0,1)$, as N goes to infinity through powers of two, the fraction of indices $i \in \{1, ..., N\}$ for which $I(W_N^{(i)}) \in (1-\delta, 1]$ goes to $I(W)$ and the fraction for which $I(W_N^{(i)}) \in [0, \delta)$ goes to $1 - I(W)$.

For fixed $\delta > 0$ where $\delta \in [0, \frac{1}{2}]$

$$\lim_{N\to\infty} \frac{|\{i \in [1,N] : \delta < I(U_i; Y_N | U^{i-1}) < 1-\delta\}|}{N} \to 0$$

true with $\delta = O(2^{-\sqrt{N}})$ also
Similarly,

$$\lim_{N\to\infty} \frac{|\{i \in [1,N] : I(U_i; Y_N | U^{i-1}) > 1-\delta\}|}{N} \to I(W)$$

and similarly

$$\lim_{N\to\infty} \frac{|\{i \in [1,N] : I(U_i; Y_N | U^{i-1}) < \delta\}|}{N} \to 1 - I(W)$$

Lemma: For all $\delta \in [0, \frac{1}{2}]$ for any symmetric capacity $I(W) \in [\delta, 1-\delta]$

$$\Delta W \triangleq \tfrac{1}{2}(I(W^+) - I(W^-))$$

Satisfying,

$$\Delta W \geq \kappa(\delta)$$
$$\kappa(\delta) \triangleq \min_{h^{-1}(\delta) \leq p \leq h^{-1}(1-\delta)} (h(2p(1-p)) - h(p))^2$$

Proof of the lemma:
If X_1, X_2 are binary and $I(W) = I(X_1; Y_1)$

$$= I(X_2; Y_2)$$
$$= 1 - h(p)$$

then MS Gerber's Lemma Proves

$$I(W^-) = 1 - H(X_1 \oplus X_2 | Y_1, Y_2) \leq 1 - h(2p(1-p))$$

Since, $I(W^+) = 2I(W) - I(W^-)$

$$\begin{aligned}\Delta W &= \tfrac{1}{2}(I(W^+) - (I(W^-)) \\ &= \tfrac{1}{2}(2I(W) - I(W^-) - I(W^-)) \\ &= I(W) - I(W^-) \\ &\geq 1 - h(p) - (1 - h(2p(1-p)))\end{aligned}$$

$$\Delta W \geq h(2p(1-p)) - h(p)$$

Since $0 < p < 2p(1-p) \leq \frac{1}{2}$,For $p \in (0, \frac{1}{2})$ and $h(p)$ as strictly increasing on the set, so $\Delta W > 0$ aslong as $p \in (0, \frac{1}{2})$ so it follows that for $I(W) \in [\delta, 1-\delta]$ one can minimize this bound over this range to see $\Delta W \geq \kappa(\delta)$.

Also we note that $\kappa(\delta) > 0$ for $\delta \in (0, \frac{1}{2}]$ because $\kappa(\delta) > 0$ as long as $0 < h^{-1}(\delta)$ and $h^{-1}(1-\delta) < \frac{1}{2}$.

Where $h(p)$ is the binary entropy function

Let the average mutual information after n steps of splitting be denoted by,

$$\mu_{n+1} \triangleq \tfrac{1}{2^{(n+1)}} \textstyle\sum_{i=1}^{2^{(n+1)}} I(W_{2^{(n+1)}}^{(i)}) = \tfrac{1}{2^n} \sum_{i=1}^{2^n} (\tfrac{1}{2} I(W_{2^{(n+1)}}^{(2i-1)}) + \tfrac{1}{2} I(W_{2^{(n+1)}}^{(2i)}))$$

i.e. split into odd-even channels.

$$= \frac{1}{2^n} \sum_{i=1}^{2^n} (\frac{1}{2} I(W_{2^{(n)}}^{(i)}) + \frac{1}{2} I(W_{2^n}^{(i)}))$$

follows from $W_N^{(2i-1)} = W_{N/2}^{(i)-}$ and $W_N^{(2i)} = W_{N/2}^{(i)+}$

$$= \frac{1}{2^{(n)}} \sum_{i=1}^{2^n} I(W_{2^n}^i)$$

follows from, $I(W) = \frac{1}{2}(I(W^+)) + \frac{1}{2}(I(W^-))$

$$\mu_{n+1} = \mu_n$$

This proves that average mutual information is conserved after every step i.e., $\mu_{n+1} = \mu_n = \mu_0 = I(W)$. We can also observe that,

$$I(W)^2 + \Delta(W)^2 = \frac{1}{2}(I(W^+)^2 + I(W^-)^2) - eqn(3)$$

Now let the average squared mutual information after n steps as v_{n+1}

$$v_{n+1} \triangleq \frac{1}{2^{n+1}} \sum_{i=1}^{2^{n+1}} I(W^i_{2^{n+1}})^2$$

$$= \frac{1}{2^n} \sum_{i=1}^{2^n} \frac{1}{2}(I(W^{2i-1}_{2^{n+1}})^2 + I(W^{2i}_{2^{n+1}})^2)$$

$$= \frac{1}{2^n} \sum_{i=1}^{2^n} \frac{1}{2}(I(W^{i+}_{2^n})^2 + I(W^{i-}_{2^n})^2)$$

$$= \frac{1}{2^n} \sum_{i=1}^{2^n}(I(W^i_{2^n})^2 + \Delta(W^i_{2^n})^2) \qquad from(3)$$

$$v_{n+1} = \frac{1}{2^n} \sum_{i=1}^{2^n}(I(W^i_{2^n})^2) + \frac{1}{2^n} \sum_{i=1}^{2^n} \Delta(W^i_{2^n})^2$$

$$v_{n+1} = v_n + \sum_{i=1}^{2^n} \Delta(W^i_{2^n})^2$$

Now we define the fraction of δ-unpolarized channels after n steps

$$\theta_n(\delta) \triangleq \frac{1}{2^n}|i \in [2^n]|I(W^i_{2^n} \in [\delta, 1-\delta])|$$

We need to prove that $\theta_n(\delta) \to 0$ as $n \to \infty$
Using Lemma we can see that,

$$\sum_{i=1}^{2^n} \Delta(W^i_{2^n})^2 \geq \theta_n(\delta)\kappa(\delta)$$

as there are $\theta_n(\delta)$ such mediocre channels so now we have,

$$0 \leq \theta_n(\delta) \leq \frac{v_{n+1} - v_n}{\kappa(\delta)}$$

as we know $v_{n+1} - v_n \to 0$, we imply $\theta_n(\delta) \to 0$ for all $\delta \in [0, \frac{1}{2})$
In addition standard results from real analysis imply the existence of a sequence $S_n \to 0$ such that $\theta_n(\delta) \to 0$.
Since $\mu_n = I(W)$ this implies that a fraction of I(W) of the virtual channels W_N^i become perfect i.e. $I(W) \to 1$ and the fraction $1 - I(W)$ of the channels become useless i.e. $I(W) \to 0$.