In class we have defined the virtual channels (conditional distributions) \( W^-(y_1, y_2 | u_1) \) and \( W^+(y_1, y_2, u_1 | u_2) \).

(a) Prove that \( W^-(y_1, y_2 | u_1) = \frac{1}{2} \sum_{u_2 = 0} W(y_1 | u_1 + u_2) W(y_2 | u_2) \).

(b) Prove that \( W^+(y_1, y_2, u_1 | u_2) = \frac{1}{2} W(y_1 | u_1 + u_2) W(y_2 | u_2) \).

(c) Prove that \( Z(W^+) = Z(W)^2 \).

(d) Prove that for any positive \( \alpha, i = 1, \ldots, 4 \) we have
\[
\sqrt{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4} \leq \sqrt{\alpha_1} + \sqrt{\alpha_2} + \sqrt{\alpha_3} + \sqrt{\alpha_4}.
\]

(e) Prove that
\[
Z(W^-) = \sum_{y_1, y_2 \in Y} \left( \frac{1}{2} \sum_{y_1, y_2 \in Y} \alpha(y_1) \alpha(y_2) + \beta(y_1) \beta(y_2) \right) \alpha(y_1) \beta(y_2) + \beta(y_1) \alpha(y_2),
\]
where \( \alpha(y) = W(y | 0) \), \( \beta(y) = W(y | 1) \).

(f) Prove that
\[
Z(W^-) \leq \frac{1}{2} \left( \sum_{y_1, y_2 \in Y} \alpha(y_1) \gamma(y_2) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in Y} \alpha(y_2) \gamma(y_1) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in Y} \beta(y_1) \gamma(y_2) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in Y} \beta(y_2) \gamma(y_1) \right),
\]
where \( \gamma(y) = \sqrt{W(y | 0) W(y | 1)} \).

(g) Using (f) prove that \( Z(W^-) \leq 2Z(W) \).

[Thus, in step \( n \) of the polarization recursion, we have \( Z(W s^n) = \left\{ \begin{array}{ll} Z(W s^{n-1} )^2 & \text{if } s_n = + \forall y \in \{+, -\}^n. \\
\leq 2 Z(W s^{n-1}) & \text{if } s_n = - \forall y \in \{+, -\}^n. 
\end{array} \right. \]

2. We will show that the Gaussian channel has the smallest capacity among all channels with additive noise of a given noise variance.

Consider a channel with input \( X \in \mathbb{R} \) and output \( Y = X + Z \), where \( Z \) is a real RV independent of \( X \), \( EZ = 0 \), \( \text{Var}(Z) = \sigma^2 \). Let \( X \sim N(0, P) \) be a Gaussian RV.

(a) Prove that \( I(X; Y) = h(x) - h(x - \alpha Y | Y) \) for any \( \alpha \in \mathbb{R} \).

(b) Prove that \( h(x - \alpha Y) \leq \frac{1}{2} \log 2 \pi e E((X - \alpha Y)^2) \) for any \( \alpha \in \mathbb{R} \).

(c) Show that Parts (a) and (b) imply that \( I(X; Y) \geq h(x) - \frac{1}{2} \log 2 \pi e E((X - \alpha Y)^2) \).

(d) Show that \( E((X - \alpha Y)^2) \geq \frac{\sigma^2}{\sigma^2 + P} \) with equality iff \( \alpha = \frac{P}{\sigma^2 + P} \). (Hint: compute the derivative on \( \alpha \).

(e) Show that (c) and (d) imply that
\[
I(X; Y) \geq \frac{1}{2} \log \left( 1 + \frac{P}{\sigma^2} \right).
\]

Conclude that the choice of \( Z \sim (0, \sigma^2) \) yields the smallest capacity value.

3. We know from class that among real-valued RVs, the Gaussian yields the largest differential entropy \( h(\cdot) \). Prove that, among all non-negative random variables with mean \( \lambda \), the largest \( h(X) \) is attained for \( X \sim \exp(1/\lambda) \), i.e., \( p_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \). (Hint: Let \( q(x), x \geq 0 \) be another pdf, and compute the divergence \( D(q || p) \).)