

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W; \mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W; \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^n H(Y_i | \mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^n H(Y_i | \mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^n H(Y_i | X_i) \\ &\leq \sum_{i=1}^n H(Y_i) - \sum_{i=1}^n H(Y_i | X_i) \\ &= \sum_{i=1}^n I(X_i; Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W; \mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

5. Formally, apply Fano's inequality to upper bound $H(W|\hat{W})$. Upon filling in the ϵ 's and δ 's as in the proof of the converse of the channel coding theorem, we conclude that

$$R \leq C$$

for any rate R achievable with complete feedback.