ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 32 Principles of Digital Communications
Solutions to Homework 12 Dec. 21, 2020

Problem 1.

(a) Given D_1, D_2 and $0 \le \lambda \le 1$ we need to show that $\phi(D) \ge \lambda\phi(D_1) + (1-\lambda)\phi(D_2)$. Suppose $p_{Z_1^*}$ and $p_{Z_2^*}$ be the distributions on Z that achieve the maximization that define ϕ for D_1 and D_2, namely, $\phi(D_1) = H(Z_1^*)$ and $\phi(D_2) = H(Z_2^*)$ with $E[g(Z_1^*)] \le D_1$ and $E[g(Z_2^*)] \le D_2$. Consider now the distribution $p_{Z^*} = \lambda p_{Z_1^*} + (1-\lambda)p_{Z_2^*}$. For Z^* having this distribution

$$\begin{aligned} E[g(Z^*)] = \sum_z p_{Z*}(z)g(z) &= \lambda \sum_z p_{Z_1^*}(z)g(z) + (1-\lambda)\sum_z p_{Z_2^*}(z)g(z) \\ &= \lambda E[g(Z_1^*)] + (1-\lambda)E[g(Z_2^*)] \le \lambda D_1 + (1-\lambda)D_2 = D, \end{aligned}$$

and because of the concavity of H, $H(Z^*) \ge \lambda H(Z_1^*) + (1-\lambda)H(Z_2^*) = \lambda\phi(D_1) + (1-\lambda)\phi(D_2)$. As $\phi(D)$ is the maximum of $H(Z)$ over all Z with $E[g(Z)] \le D$, $\phi(D) \ge H(Z^*)$.

(b) In the (in)equalities

$$\begin{aligned} I(U;V) &\overset{(b1)}{=} H(U) - H(U|V) \\ &\overset{(b2)}{=} H(U) - H(U \ominus V|V) \\ &\overset{(b3)}{\ge} H(U) - H(U \ominus V) \\ &\overset{(b4)}{\ge} H(U) - \phi(D) \end{aligned}$$

(b1) is by definition of mutual information, (b2) because for a given V, U and $U \ominus V$ are in one-to-one correspondence, (b3) because conditioning reduces entropy and (b4) because $Z = U \ominus V$ has $E[g(Z)] \le D$.

(c) As $R(D) = \min\{I(U;V) : E[d(U,V)] \le D\}$, and by (b) for any U, V with $E[d(U,V)] \le D$ we have $I(U;V) \ge H(U) - \phi(D)$, the conclusion follows.

(d) Let Z be independent of U and have a distribution that achieves $\phi(D)$. Set $V = U \ominus Z$. Now,

$$p_{Z,V}(z,v) = p_{Z,U}(z, z \oplus v) = p_Z(z)p_U(z \oplus v) = p_Z(z)/|\mathcal{U}|.$$

By summing over z we see that V is uniformly distributed, and also that V is independent of $Z = U \ominus V$. Observe that the only inequalities in (b) were in (b3) and (b4), but in this case they are both equalities: (b3) because of the independence of $Z = U \ominus V$ and V, and (b4) because $H(Z) = \phi(D)$.

Problem 2. Suppose U, V satisfy $E[(U-V)^2] \le D$, and set $Z = U - V$. As $E[Z^2] \le D$, we know $h(Z) \le \frac{1}{2}\log(2\pi e D)$. Also,

$$I(U;V) = h(U) - h(U|V) = h(U) - h(Z|U) \ge h(U) - h(Z) \ge h(U) - \tfrac{1}{2}\log(2\pi e D),$$

and consequently $R(D) \geq h(U) - \frac{1}{2}\log(2\pi e D)$. We now turn to the upper bound on $R(D)$. Assume without loss of generality that $E[U] = 0$ so that $\sigma^2 = E[U^2]$. If $D \geq \sigma^2$, we can take $V = 0$ for which $E[(U-V)^2] = \sigma^2 \leq D$, and $I(U;V) = 0$, so that $R(D) = 0$. So, we need to only consider the case $D < \sigma^2$. For such D, let Z be a zero mean Gaussian independent of U with variance $D(1 - D/\sigma^2)$ and set $V = (1 - D/\sigma^2)U + Z$. We will show that for this choice of V we have $E[(V-U)^2] = D$ and $I(U;V) \leq \frac{1}{2}\log(\sigma^2/D)$, which will then establish that $R(D) \leq \frac{1}{2}\log(\sigma^2/D)$. To that end observe that $V - U = -(D/\sigma^2)U + Z$ and thus $E[(V-U)^2] = (D/\sigma^2)^2 E[U^2] + E[Z^2] = D$. Turning our attention now to $I(U;V)$, first compute $E[V^2] = (1 - D/\sigma^2)^2 E[U^2] + E[Z^2] = \sigma^2 - D$, so $h(V) \leq \frac{1}{2}\log(2\pi e(\sigma^2 - D))$. Furthermore,

$$\begin{aligned} h(V|U) &= h(V - (1 - D/\sigma^2)U \mid U) = h(Z|U) = h(Z) \\ &= \frac{1}{2}\log(2\pi e \operatorname{Var}(Z)) = \frac{1}{2}\log(2\pi e D(1 - D/\sigma^2)). \end{aligned}$$

Thus

$$I(U;V) = h(V) - h(V|U) \leq \frac{1}{2}\log\frac{\sigma^2 - D}{D(1 - D/\sigma^2)} = \frac{1}{2}\log(\sigma^2/D).$$

Problem 3.

(a) Since the channel is memoryless and feedback-free transmission is assumed, from code construction, it is obvious that $(\mathrm{enc}_1(m_1), \mathrm{enc}_2(m_2), Y^n)$ is an i.i.d. length-n sequence of (X_1, X_2, Y)'s drawn from distribution $p(x_1, x_2, y) = p_1(x_1)p_2(x_2)p(y|x_1, x_2)$. Therefore, for sufficiently large n, the probability of this sequence being ϵ-typical is as high as desired.

(b) Now, $(\mathrm{enc}_1(\tilde{m}_1), \mathrm{enc}_2(m_2), Y^n)$ is an i.i.d. sequence (of length n) whose components are distributed according to $p_1(x_1)p(y, x_2)$ where $p(y, x_2) = \sum_{x_1'} p_1(x_1')p_2(x_2)p(y|x_1', x_2)$.

(c) $\Pr\{(\mathrm{enc}_1(\tilde{m}_1), \mathrm{enc}_2(m_2), Y^n) \in T\}$ is the probability of a length n i.i.d. sequence X_1^n whose elements have distribution p_1 being jointly ϵ-typical (with respect to the distribution $p_1(x_1)p(y, x_2|x_1)$ where $p(y, x_2|x_1) = p(x_2)p(y|x_1, x_2)$) with an independent length n sequence of $(X_2, Y)^n$ whose elements have distribution $p(y, x_2)$ (defined in (b)). Thus, as we have seen in the course,

$$\Pr\{(\mathrm{enc}_1(\tilde{m}_1), \mathrm{enc}_2(m_2), Y^n) \in T\} \doteq 2^{-nI(X_1, X_2 Y)}.$$

(In the course we have seen this result for two random variables X and Y; it is obvious that we can replace X by X_1 and Y by $(X_2 Y)$ to derive the desired result).

(d) From (a) we know that the probability of the correct message m_1 not being on the list of typical m_1's at decoder 2 is small, say at most $\epsilon/2$.

From (c), the probability of each incorrect $\tilde{m}_1$ being on that list (at decoder 2) is equal (up to sub-exponential factors) to $2^{-nI(X_1;X_2Y)}$. Since there are $M - 1 \leq 2^{nR_1}$ such $\tilde{m}_1$'s, the probability of having *an* incorrect message on the list is, by the union bound, at most $2^{n[R_1 - I(X_1;X_2Y)]}$ which is exponentially small in n provided that $R_1 < I(X_1;X_2Y)$. Thus, for large enough n, this probability is also smaller than $\epsilon/2$.

Consequently, the average probability of decoding error at decoder 2 is at most ϵ provided that $R_1 < I(X_1, X_2 Y)$.

By symmetry, the average probability of decoding error at decoder 1 is smaller than ϵ if $R_2 < I(X_2, X_1, Y)$.

Since the average probability of error (over the generation of codebooks) is small (for rate pairs (R_1, R_2) satisfying $R_1 < I(X_1; Y, X_2)$ and $R_2 < I(X_2; Y, X_1)$), there exist a pair of codebooks of rates (R_1, R_2) in the ensemble for which the average error probability is small, thus such (R_1, R_2)'s are achievable.

(e) Firstly note that since X_1 and X_2 are independent, $I(X_1; YX_2) = I(X_1; Y|X_2)$ (similarly $I(X_2; YX_1) = I(X_2; Y|X_1)$).

Since $Y = X_1 \times X_2$, conditioned on $\{X_2 = 0\}$, Y contains no information about X_1, whereas conditioned on $\{X_2 = 1\}$, $Y = X_1$. Assuming $\Pr\{X_1 = 1\} = p_1$ and $\Pr\{X_2 = 1\} = p_2$,

$$\begin{aligned} I(X_1; Y|X_2) &= \Pr\{X_2 = 0\} I(X_1; Y|X_2 = 0) + \Pr\{X_2 = 1\} I(X_1; Y|X_2 = 1) \\ &= 0 + p_2 h_2(p_1) \end{aligned}$$

where $h_2(\cdot)$ is the binary entropy function. Similarly it follows that $I(X_2; Y|X_1) = p_1 h_2(p_2)$.

Suppose $p_1 = p_2 = p$, then all rates (R_1, R_2) satisfying

$$R_1 < p h_2(p) \qquad R_2 < p h_2(p)$$

are achievable. In particular, $p h_2(p) \geq \frac{1}{2}$ for some $p \geq \frac{1}{2}$ (it evaluates to $\frac{1}{2}$ at $p = \frac{1}{2}$ but it is increasing, so it will go above $\frac{1}{2}$ as p increases). The set of achievable rate pairs corresponding to such p's violate $R_1 + R_2 < 1$.

PROBLEM 4. (a) The Slepian-Wolf region for U and V is given as the set of rate pairs (R_u, R_v) satisfying

$$\begin{aligned} R_u &> H(U|V) \\ R_v &> H(V|U) \\ R_u + R_v &> H(UV) \end{aligned}$$

The joint distribution of (U, V) is given as

$$P(U = u, V = v) = \begin{cases} p^2 & (u, v) = (1, 2) \\ 2pq & (u, v) = (0, 1) \\ q^2 & (u, v) = (0, 0) \end{cases}.$$

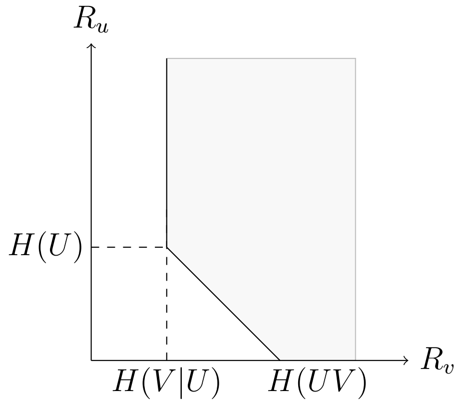
Therefore,

$$\begin{aligned} H(UV) &= H(2pq, p^2, q^2) \\ H(V) &= H(2pq, p^2, q^2) \\ H(U) &= H(p^2, 2pq + q^2) \end{aligned}$$

and

$$H(U|V) = H(UV) - H(V) = 0$$
$$H(V|U) = H(UV) - H(U) = H(p^2, 2pq, q^2) - H(p^2, 2pq + q^2) = (2pq + q^2) h_2\left(\frac{2pq}{2pq + q^2}\right)$$

where $h_2(.)$ is the binary entropy function. The rate region can be depicted as follows.



(b) $H(Z_1^n Z_2^n | U^n V^n) = nH(Z_1 Z_2 | UV) = n \sum_{u,v} H(Z_1 Z_2 | U = u, V = v) P(U = u, V = v)$. Knowing the (u, v) pair, the only uncertainty in (Z_1, Z_2) pair occurs when $u = 0$ and $v = 1$. Moreover $P(Z_1 = 1, Z_2 = 0 | U = 0, V = 1) = P(Z_1 = 0, Z_2 = 1 | U = 0, V = 1) = 1/2$. Thus,

$$H(Z_1^n Z_2^n | U^n V^n) = nH(Z_1 Z_2 | U = 0, V = 1) P(U = 0, V = 1) = 2npq$$

PROBLEM 5. (a) Given a code $\mathcal{C}$ with M codewords and blocklength n, and $0 \le k \le n$, partition the codewords into 2^k groups according to their first k bits. The group with the largest number of codewords will contain at least $M' = \lceil M/2^k \rceil$ codewords. The minimum distance within that group is upper bounded by $d_0(M', n-k)$ since all codewords in the group agree in their first k bits. Thus the minimum distance of the code $\mathcal{C}$ is upper bounded by $d_0(\lceil M/2^k \rceil, n-k)$. Since this is true for each $k \in \{0, \dots, n\}$ we conclude that $d_{\min} \le d_1(M, n)$.

(b) With $d_0(M, n) = \begin{cases} n & M \le 2 \\ \infty & M \le 1 \end{cases}$ the minimum over k is obtained by choosing k as large as possible while keeping $M/2^k > 1$. Thus the bound d_1 says "$d_{\min} \le n - k$ when $M > 2^{kn}$" which is the Singleton bound we derived in class.

(c) Each pair (m, m') contributes 1 to the sum when $a_m = 0$ and $a_{m'} = 1$ or when $a_m = 1$ and $a_{m'} = 0$. There are $M_0 M_1$ pairs of the first type and $M_1 M_0$ pairs of the second type. Thus the sum equals $2M_0M_1$. As $M_0 + M_1 = M$, we have $M_0 M_1 \le M^2/4$, from which the final inequality follows.

(d) As $d_H(\mathbf{x}_m, \mathbf{x}_{m'}) \ge d_{\min}$ for every $m \ne m'$, the first inequality follows by summing both sides. For the second write $d_H(\mathbf{x}_m, \mathbf{x}_{m'}) = \sum_{i=1}^n d_H(x_{mi}, x_{m'i})$ to obtain

$$\sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(\mathbf{x}_m, \mathbf{x}_{m'}) = \sum_{i=1}^{n} \sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(x_{mi}, x_{m'i}).$$

By (c) for each i the inner double-sum is upper bounded by $M^2/2$ and the conclusion follows.