ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 36 Information Theory and Coding

Final exam solutions Jan. 29, 2019

PROBLEM 1.

(a) $\Pr(\mathbf{X}(m) = \mathbf{X}(m')) = \sum_{\mathbf{x}} \Pr(\mathbf{X}(m) = x) \Pr(\mathbf{X}(m') = x) = \sum_{\mathbf{x}} \Pr(\mathbf{X}(m) = x)^2$. Since $\mathbf{X}(m)$ is uniformly distributed over $\{0,1\}^n$, we find $\Pr(\mathbf{X}(m) = \mathbf{X}(m')) = 2^{-n}$.

(b) Taking the hint, $\Pr(G_i = 1 \mid G_1 = \cdots = G_{i-1} = 1)$ is the probability that $\mathbf{X}(m)$ is different than $i-1$ values. Since $\mathbf{X}(m)$ equals each value with the probability found in (a), we see that $\Pr(G_i = 1 \mid G_1 = \cdots = G_{i-1} = 1) = 1 - (i-1)2^{-n}$.

(c) By the chain rule $\Pr(G_1 = \cdots = G_M = 1) = \prod_{i=1}^M \Pr(G_i = 1 \mid G_1 = \cdots = G_{i-1} = 1) = \prod_{i=1}^M (1 - (i-1)2^{-n})$.

(d) The value of q is already computed in (c). With the hint, $q \le \prod_{i=1}^M \exp(-(i-1)/2^n) = \exp(-\sum_{i=1}^M (i-1)/2^n)$

(e) By (d), $q \le \exp(-M(M-1)/2^{n+1})$. When $R > 1/2$, $M(M-1)$ grows faster than 2^n, thus $q \to 0$ as n gets large.

(f) With $p(\mathbf{x})$ denoting $\Pr(\mathbf{X}(m) = \mathbf{x})$, the probability in (a) is $\sum_{\mathbf{x}} p(\mathbf{x})^2$. By the Cauchy-Schwartz inequality, $\left[\sum_{\mathbf{x}} p(\mathbf{x})\right]^2 \le \sum_{\mathbf{x}} p(\mathbf{x})^2 \sum_{\mathbf{x}} 1$, thus we get that $\Pr(\mathbf{X}(m) = \mathbf{X}(m')) \ge 2^{-n}$. This then implies that $\Pr(G_i = 1 | G_1 = \cdots = G_{i-1} = 1) \le 1 - (i-1)2^{-n}$, and consequently, the value of q in (d) is an upper bound to $\Pr(G_1 = \cdots = G_M = 1)$.

Moral of the story: a randomly constructed binary code with rate larger than 1/2 will (with high probability) have two (or more) identical codewords, and thus its $P_{e,\max} \ge 1/2$, no matter on what channel it is used. This is the reason why we go through $P_{e,ave}$ and then expurgate to construct a code with small $P_{e,\max}$ rather than trying to prove the existence of codes with small $P_{e,\max}$ by random coding directly.

PROBLEM 2.

(a) Write $I(X^2; Y^2) = H(X^2) - H(X^2|Y^2)$. By the chain rule and that conditioning reduces entropy $H(X^2|Y^2) \le H(X_1|Y_1) + H(X_2|Y_2)$. Moreover when X_1 and X_2 are independent $H(X^2) = H(X_1) + H(X_2)$. The conclusion follows.

(b) The capacity of the effective channel is given by $C = \max p_{X^2} I(X^2; Y^2)$. By (a) $I(X^2; Y^2) \ge I(X_1; Y_1) + I(X_2; Y_2)$. Consequently, $C \ge \max_{p_{X^2}} I(X_1; Y_1) + I(X_2; Y_2) = C_1 + C_2$ where $C_i = \max_{p_{X_i}} I(X_i; Y_i)$ is the capacity of the i'th channel.

(c) The individual channels are BSC's with crossover probability 1/2, so $C_1 = C_2 = 0$. However $I(X^2; Y^2) = H(Y^2) - H(Y^2|X^2) = H(Y^2) - H(Z^2) = H(Y^2) - 1$. Since Y^2 can take only 4 possible values, $H(Y^2) \le 2$. On the other hand, choosing X_1 and X_2 to be independent and equally likely to be 0 or 1 makes Y^2 uniformly distributed on its four possible values, so the capacity of the effective channel is $C = 1$.

Moral of the story: memory in the channel noise increases capacity.

Problem 3.

(a) As H had four columns the blocklength $n = 4$. Observe that we can rearrange $H\mathbf{x} = \mathbf{0}$ to solve for x_1, x_2 in terms of x_3, x_4. As there are 3^2 possibilities for (x_3, x_4) the code has $M = 9$ codewords. The code rate is thus $\frac{1}{2}\log 3$.

(b) The receiver receives $\mathbf{y} = \mathbf{x} + \mathbf{z}$ where $\mathbf{z}$ is either the zero vector, or it has only a single nonzero component z_i which can take the value 1 or 2. With h_i denoting the ith column of H, $H\mathbf{y} = H\mathbf{z}$ is either zero, or takes on the value h_i (if $z_i = 1$) or $2h_i$ ($z_i = 2$). Since the collection of eight vectors $h_1, 2h_1, h_2, 2h_2, h_3, 2h_3, h_4, 2h_4$ are all distinct and different from zero, the receiver can identify if z is the zero vector or the i and the value of z_i from $H\mathbf{y}$

(c) This will increase the block length to 5 and the number of codewords to 3^3 yielding a new rate of $\frac{3}{5}\log 3$ which is larger than the rate found in (a).

(d) We need to ensure that the new column and its multiple by 2 is different from the zero and the collection of 8 vectors above. We see that this is not the case for any of the vectors listed.

(e) Now z_i can take on only the value 1 (but not 2). Thus to ensure detection and correction we only need h_i's to be distinct and different from zero. Now, all columns except the zero column in (d) can be added.

Problem 4.

(a) This was found in class to be the pentagon given by the constraints $R_1 \le 1$, $R_2 \le 1$, $R_1 + R_2 \le 3/2$. Note that the highest rate R for which (R, R) is in the capacity region is $R = 3/4$.

(b) At the end of phase 1, both the encoders know $Y^k = U_1^k + U_2^k$. Since each knows its own message each can discover the message of the other. Consequently, they can both compute Q.

The receiver knows the value of U_{1i} and U_{2i} for those i's for which Y_i is 0 or 2. For those i's for which $Y_i = 1$ (i.e., $i_1, \dots, i_T$) it knows that one of U_{1i} and U_{2i} is 0 and the other is 1, but does not know which. So, unless $T = 0$, it does not know Q.

(c) By (b) both encoders know Q and thus $v_1, \dots, v_{\lceil S \rceil}$. They can then set

$$(U_{1,k+i}, U_{2,k+i}) = \begin{cases} (0,0) & \text{if } v_i = 0 \\ (1,0) & \text{if } v_i = 1 \\ (1,1) & \text{if } v_i = 2, \end{cases} \qquad i = 1, \dots, \lceil S \rceil.$$

to ensure that the receiver receives $v_1, \dots, v_{\lceil S \rceil}$. Note that at the end of phase 2 the receiver can compute Q, and thus find U_1^k and U_2^k. The two phase scheme thus reliably sends k bits from each transmitter to the receiver.

(d) Note that during the first phase $\Pr(Y_i = 1) = \frac{1}{2}$. Thus, $E[T] = \frac{1}{2}k$, and $E[S] = \frac{1}{2}k\log_3 2$. Consequently $E[k + S] = \big(1 + \frac{1}{2}\log_3(2)\big)k$.

(e) Set $c = 1 + \frac{1}{2}\log_3 2$. Since $k + S \le N < k + S + 1$, we find $ck \le E[N] < ck + 1$. Thus $k/E[N] \to 1/c$.

(f) Note that in the first phase $Y_1, \dots, Y_k$ are i.i.d. Thus, by the law of large numbers, as k gets large, $T/k \to 1/2$ with probability 1. Consequently the rate $R = k/N \to 1/c$ with probability 1. As $\log_3 2 < 2/3$, $c \leq 4/3$ and thus $R > 3/4$ with probability 1.

Moral of the story: Feedback allows us to achieve the rate pair $(R, R) > (3/4, 3/4)$ which is outside of the region computed in (a). Thus, feedback may enlarge the capacity region of a memoryless multiple access channel. Recall that this was not the case for the single user channel — feedback does not increase the capacity of a single user memoryless channel.