IERG 6154: Network Information Theory
Homework 4

Due: Feb 18, 2019

1. *Push-to-talk MAC*: Consider the following MAC: $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}$; the transition probability is defined according to $\mathrm{P}(Y = 1|X_1 = 0, X_2 = 0) = 0, \mathrm{P}(Y = 1|X_1 = 1, X_2 = 0) = 1, \mathrm{P}(Y = 1|X_1 = 0, X_2 = 1) = 1, \mathrm{P}(Y = 1|X_1 = 1, X_2 = 1) = 0.5$.

   (a) Find the capacity region for this MAC

   (b) Prove that you need convexification (i.e. Q) to achieve the capacity region. That is, the union of the pentagons will not yield the capacity region.

   (c) Plot (numerically) the region represented by the union of the pentagons.

2. *Two-sender two-receiver channel*: Consider the channel shown below. Assume the messages are independent and uniform. The two encoders do not interact with each other. The two decoders also do not interact with each other. Both receivers need to decode both the messages.

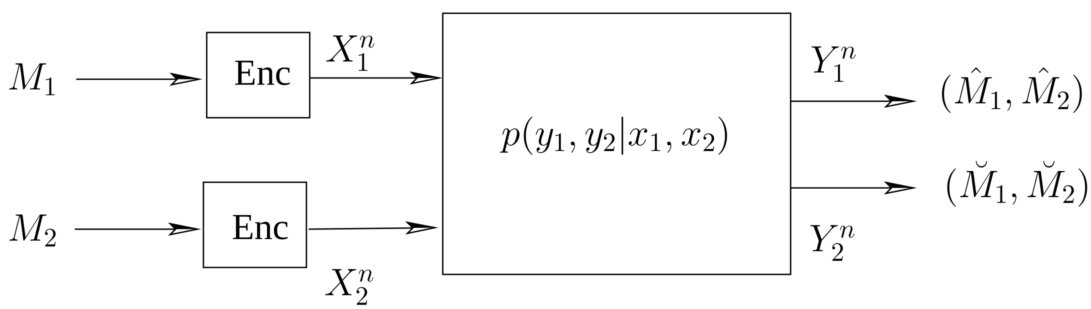

Figure 1: A two sender two receiver channel

   Compute the capacity region of this channel. Given both a sketch of the achievability and the converse.

3. *One-letter vs two-letter* Consider a generic discrete memoryless MAC channel characterized by $\mathfrak{q}(y|x_1, x_2)$. Let $\mathcal{R}_n(\mathfrak{q})$ denote the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \frac{1}{n} I(X_1^n; Y^n)$$
$$R_2 \leq \frac{1}{n} I(X_2^n; Y^n)$$

   over all distributions $p_1(x_1^n)p_2(x_2^n)$.

   (a) Show that $\mathcal{C}(\mathfrak{q}) = \overline{\cup_n \mathcal{R}_n(\mathfrak{q})}$, i.e. the capacity region is the closure of the union of all $\mathcal{R}_n$. (You need to show an achievability and a converse.)

   (b) Argue that if $\mathcal{R}_1(\mathfrak{q}) = \mathcal{R}_2(\mathfrak{q})$ for every $\mathfrak{q}$, then the capacity region of MAC would have been the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(X_1; Y)$$
$$R_2 \leq I(X_2; Y)$$

   over all distributions $p_1(x_1)p_2(x_2)$.

   (c) Consider the MAC Channel defined by $Y = X_1 + X_2$ with $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, 2\}$.

(i) Show that the capacity region is the rates $(R_1, R_2) \in [0,1] \times [0,1] \cap \{(R_1, R_2) : R_1 + R_2 \leq 1.5\}$.

(ii) Compute the maximum sum-rate in $\mathcal{R}_1(\mathfrak{q})$, i.e. compute

$$\max_{p_1(x_1)p_2(x_2)} I(X_1;Y) + I(X_2;Y),$$

and verify that it is strictly smaller than the true sum-capacity.

(iii) Numerically (using matlab or your favorite simulation platform) compute the maximum sum-rate in $\mathcal{R}_2$ and $\mathcal{R}_3$. (You should observe the lack of monotonicity in your simulations, i.e. check that maximum sum-rate is higher in $\mathcal{R}_2$ than in $\mathcal{R}_3$).

4. *Feedback increases the capacity of MAC* Consider the same MAC channel as in the previous question. Assume noiseless feedback to both transmitters, i.e. both senders receive the bit Y_i (through a noiseless feedback link) and can use the information to encode $X_{1,i+1}$ and $X_{2,i+1}$. This exercise is a result due to Gaarder and Wolf. The coding scheme is as follows. Both senders use the first nR bits in the block of length n to send the message bits uncoded. The receiver is clearly unable to distinguish between the message bits only when she receives a 1. Both senders cooperatively try to encode the message bits of the first encoder for these slots (the bits are complements of each other for these slots; hence second receiver can infer the message bits of the first encoder) by using only symbol-pairs $(0,0),(0,1),(1,1)$. Numerically calculate the best symmetric rate for the above scheme and hence conclude that $(R_1, R_2) = (0.76, 0.76)$ is achievable (i.e. probability of error goes to zero) for such a scheme.

5. *Maximal probability of error* In this problem we will show that the capacity region with the maximum probability of error criterion is same as the capacity region with the average probability of error criterion as long as we allow some randomization at the encoders. This proof is a modification of the one due to Ahlswede (1977).

   Consider a codebook of size $2^{nR_1} \times 2^{nR_2}$ with average probability of error ϵ. (I am dropping the dependence of ϵ on n with the implicit understanding that ϵ can be made arbitrarily small for large values of n. I am also assuming that both $R_1, R_2 > 0$, as otherwise the result is trivial.) Let $\lambda(m_1, m_2)$ denote the decoding error when the codeword corresponding to the pair $(M_1 = m_1, M_2 = m_2)$ is transmitted.

   Given the average probability of error criterion, it is clear that

$$\sum_{m_1 \in [2^{nR_1}], m_2 \in [2^{nR_2}]} \lambda(m_1, m_2) = 2^{nR_1} 2^{nR_2} \epsilon.$$

   Call a pair (m_1, m_2) as a *bad pair* if $\lambda(m_1, m_2) > \sqrt{\epsilon}$.

   (a) Show that the number of bad pairs, B, satisfies $B \leq 2^{nR_1} 2^{nR_2} \sqrt{\epsilon}$.

   Randomly permute the indices $\{1, .., 2^{nR_1}\}$ (uniformly across all permutations) and independently permute the indices $\{1, .., 2^{nR_2}\}$ (again uniformly across all permutations). Partition the permuted row (and column) numbers into bins of size n^2 each.

   We will now create a good code with low maximal probability of error performance as follows:

   - The new size of the codebook is $\frac{2^{nR_1}}{n^2} \times \frac{2^{nR_2}}{n^2}$.
   - For a given $(\tilde{m}_1, \tilde{m}_2)$ pair (from the new codebook size), encoder 1 picks the X_1^n sequence uniformly at random from the $\tilde{m}_1^{th}$ bin, and encoder 2 picks the X_2^n sequence uniformly at random from the $\tilde{m}_2^{th}$ bin.
   - The decoder uses the previous decoding rule to determine the transmitted message pair.

   It is clear that the probability of error corresponding to a given message pair $\tilde{m}_1, \tilde{m}_2$ is equal to the average of $\lambda(i,j)$ corresponding to the message pairs (i,j) (in the original codebook) mapped to the $\tilde{m}_1^{th}$ and $\tilde{m}_2^{th}$ bin respectively.

   We will call a pair $(\tilde{m}_1, \tilde{m}_2)$ in the new codebook as *inadmissible* if the average of $\lambda(i,j)$ corresponding to the message pairs (i,j) (in the original codebook) mapped to the $\tilde{m}_1^{th}$ and $\tilde{m}_2^{th}$ bin respectively is *greater* than $\epsilon^{1/4}$.

   Our scheme is to show that average over all the random permutations, the expected number of *inadmissible* pairs in the new codebook goes to zero; hence there must exists at least one pair of permutations which has no *bad* pair. This permutation will then give rise to the good maximal probability of error coding scheme. Observe that the encoders are still stochastic, i.e. the mapping of $\tilde{m}_1$ to $X_1^n(\tilde{m}_1)$ is not deterministic, and similarly for $\tilde{m}_2$.

   Let us consider $(\tilde{m}_1, \tilde{m}_2) = (1,1)$. Let $(i_1, j_1), \ldots, (i_{n^2}, j_{n^2})$ be the original message pairs that get mapped to the diagonal of the $(1,1)$ bin. (Note that this fixes all the other elements too as the row numbers and the column numbers that get mapped here are determined by the diagonal).

We wish to estimate the following probability

$$\mathrm{P}(\sum_{k=1}^{n^2} \lambda(i_k, j_k) \geq n^2 \epsilon^{1/4}).$$

To bound this, proceed as follows:

(b) If $\sum_{k=1}^{n^2} \lambda(i_k, j_k) \geq n^2\epsilon^{1/4}$, then show that the number of *bad pairs* among (m_{i_k}, m_{j_k}) is at least $n^2(\epsilon^{1/4} - \sqrt{\epsilon})$.

(c) Show that the probability that some fixed subset of cardinality r of these diagonal elements correspond to bad pairs while the rest are good is bounded from above by

$$\left(\frac{2^{nR_1}2^{nR_2}\sqrt{\epsilon}}{(2^{nR_1} - n^2)(2^{nR_2} - n^2)}\right)^r.$$

Remark: The above step is the key component of this proof, i.e. obtaining an exponential decay of the probability. To see this observe that, the first diagonal element is a uniformly distributed random entry of the matrix. Conditioned on the first diagonal element, the second diagonal element is a uniformly distributed random entry from the sub-matrix obtained after removing the row and column occupied by the first diagonal element, and so on. Therefore conditioned on previous entries, the probability that a particular diagonal element is a bad pair can be bounded from above by the maximum number of bad pairs possible in the matrix divided by the size of the sub-matrix.

(d) Show that for n large enough we have

$$\frac{2^{nR_1}2^{nR_2}\sqrt{\epsilon}}{(2^{nR_1} - n^2)(2^{nR_2} - n^2)} < 2\sqrt{\epsilon}.$$

(e) Therefore the probability that at least r_0 diagonal elements are bad pairs is bounded from above by (for n large enough)

$$\sum_{r=r_0}^{n^2} \binom{n^2}{r} \left(2\sqrt{\epsilon}\right)^r.$$

(f) Show that

$$\frac{\binom{k}{m+1}p^{m+1}}{\binom{k}{m}p^m} \leq \frac{1}{2} \iff m \geq \frac{2pk-1}{1+2p}.$$

(g) Use parts $(b), (e), (f)$ to argue that (for n large enough)

$$\mathrm{P}(\sum_{k=1}^{n^2} \lambda(i_k, j_k) \geq n^2\epsilon^{1/4}) \leq \binom{n^2}{n^2(\epsilon^{1/4} - \sqrt{\epsilon})} \left(2\sqrt{\epsilon}\right)^{n^2(\epsilon^{1/4}-\sqrt{\epsilon})} \cdot 2.$$

(h) Show using Stirling's approximation that for ϵ is sufficiently small, when n is large enough

$$\mathrm{P}(\sum_{k=1}^{n^2} \lambda(i_k, j_k) \geq n^2\epsilon^{1/4}) \leq e^{-a(\epsilon)n^2}$$

for some $a(\epsilon) > 0$, where $a(\epsilon) \to 0$ as $\epsilon \to 0$. (Essentially you have end up having to show is that when ϵ is sufficiently small $H(\epsilon^{1/4} - \sqrt{\epsilon}) + (\epsilon^{1/4} - \sqrt{\epsilon})\log(2\sqrt{\epsilon}) < 0$.)

(i) Use union bound to bound the probability that

$$\mathrm{P}(\sum_{k=1}^{n^2}\sum_{l=1}^{n^2} \lambda(i_k, j_l) \geq n^4\epsilon^{1/4}) \leq n^2 e^{-a(\epsilon)n^2}.$$

This captures the event that the message pair $(1, 1)$ in the reduced codebook (with stochastic encoding) has a probability of error $\geq \epsilon^{1/4}$. **Remark:** This idea is present in Ahlswede's argument. To see this bound, decompose the matrix into n^2 sums, along diagonal and its shifts. For instance the first shift may consist of elements $(1, 2), (2, 3), .., (n^2-1, n^2), (n^2, 1)$. Finally notice that the joint distributions of the random variables in every shift is identical.

(j) Show that one can bound the expected number of *inadmissible* pairs (expectation is over the random permutations of rows and columns) by

$$\frac{2^{nR_1}2^{nR_2}}{n^2}e^{-a(\epsilon)n^2}.$$

Since this expectation goes to zero as $n \to \infty$, we conclude that there exists a permutation of rows and columns so that *all pairs are admissible,* which then leads to a transmission scheme with small maximal probability of error.