ECE 788: Network Information Theory
Final, Fall 2009

1. Prove the following statement: If any code $x^n(w) \in \mathcal{X}^n$, $w \in [1, 2^{nR}]$, over alphabet $\mathcal{X}$, of length n and rate R, exists that has probability of error P_e^n, then there must exist another code $\tilde{x}^n(w) \in \mathcal{X}^n$, $w \in [1, 2^{n\tilde{R}}]$ with codewords *all of the same type* with rate $\tilde{R} > R - \frac{|\mathcal{X}|\log_2(n+1)}{n}$ and probability of error $\tilde{P}_e^n \leq P_e^n$. In other words, all the codewords $\tilde{x}^n(w)$ of the new code satisfy $\tilde{x}^n(w) \in T_0^n(P_X)$ for some type P_X, the rate is almost the same as the original code (as n increases) and the probability of error improved.
(Hint: Given the original code $x^n(w)$, the number of codewords with the same type must be at least...)

Sol.: The total number of types is upper bounded by $(n+1)^{|\mathcal{X}|}$. Therefore, there must exist a type P_X such that the number of codewords $x^n(w)$ in the original code that satisfy $x^n(w) \in T_0^n(P_X)$ is at least $2^{nR}/((n+1)^{|\mathcal{X}|})$. In fact, this number corresponds to the worst case in which all types have an equal share of sequences (in any other case, one type will have more sequences than that). Restrict your code to such sequences and define this code as $\tilde{x}^n(w)$. The rate of this code is then $\geq 1/n \cdot \log_2\left[2^{nR}/((n+1)^{|\mathcal{X}|})\right] = R - \frac{|\mathcal{X}|\log_2(n+1)}{n}$. Moreover, using the same decoder as for the original code, the probability of error can only decrease (or stay equal), since there are less codewords that should be distinguished.

2.1. Consider two correlated sequences (U^n, V^n), which are binary (i.e., $U_i, V_i \in \{0,1\}$) and generated i.i.d. according to the joint distribution $P_{UV}(u,v)$, where $P_{UV}(u,v) = 0$ if $(u,v) = (1,0)$ and $P_{UV}(u,v) = 1/3$ otherwise. Assuming that the two sources are measured by two distinct encoders, say encoder 1 and 2, respectively, what are the conditions on the corresponding rates (R_1, R_2) necessary to convey (U^n, V^n) losslessly to a single destination (which receives noiselessly at rates R_1, R_2 from the two encoders)?
2.2. Consider now the multiple access channel $Y = X_1 + X_2$, where X_1, X_2 are binary (i.e., $X_1, X_2 \in \{0,1\}$) and "+” represents real addition. What are the pairs of rates (R_1, R_2) that the encoders modulating X_1 and X_2, respectively, can achieve towards the destination (which measures Y)? (You can use without loss of optimality uniform input distributions)
2.3. Compare the rate regions obtained at points 2.1. and 2.2. Based on this observation, what do you think would happen if the two encoders of point 2.1. communicated to the receiver through the multiple access channel of point 2.2. (rather than via noiseless channels as per point 2.1)? What would happen if only encoder 1 were present in the system?
2.4. Following the previous point, try now to set $X_1 = U$ and $X_2 = V$, how does this encoding scheme perform?

Sol.: **2.1.** From the Slepian-Wolf theorem, we have

$$\begin{aligned} R_1 &\geq H(U|V) = P_V(0)H(U|V=0) + P_V(1)H(U|V=1) \\ &= \frac{1}{3}\cdot 0 + \frac{2}{3}\cdot 1 = \frac{2}{3} \\ R_2 &\geq H(V|U) = \frac{2}{3} \\ R_1 + R_2 &\geq H(U,V) = H(U) + H(V|U) = H(1/3) + \frac{2}{3} \\ &= 0.91 + \frac{2}{3} = 1.58. \end{aligned}$$

2.2. It is immediate to see that all the mutual information terms in the capacity region are maximized by uniform input distributions, so that

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2) = H(X_1) = 1 \\ R_2 &\leq I(X_2;Y|X_1) = H(X_2) = 1 \\ R_1 + R_2 &\leq I(X_1X_2;Y) = H(X_1+X_2) = 1.5 \end{aligned}$$

2.3. Since the two regions do not share any rate point, it would be impossible for the Slepian-Wolf encoders to convey reliably the compression indices to the destination. Therefore, the two sources cannot be both reliably communicated to the destination. If instead only encoder 1 were present in the system, then reliable communications of U^n could take place since $H(U) < I(X_1;Y|X_2)$.
2.4. We would have $Y = U + V$, from which it is easy to see that one can always recover both U and V (since $P_{UV}(1,0) = 0$).

3. Consider the channel

$$Y_i = X_i \oplus S_i$$

where "$\oplus$" represents binary addition, all alphabets are binary, and the state sequence S^n is i.i.d. with probability mass function $P_S(1) = q$.
3.1. Assuming that S^n is known at the decoder only, what is the capacity?
3.2. Assuming that S^n is known at both encoder and decoder, what is the capacity?
3.3. Assume now that S^n is known only at the encoder. Consider the performance of a scheme based on coset expansion where the auxiliary codebook is selected according to an auxiliary variable $U = X \oplus S$ with X independent of S and such that $P_X(1) = p$, where p should be designed. Find the achievable rate of this scheme and the optimal p. Can we conclude that this scheme is optimal?

Sol.: **3.1.**

$$C^D = 1,$$

since the decoder can cancel the sequence S^n.
3.2.

$$C^{E-D} = 1.$$

3.3. From the Gelfand-Pinsker theorem, we have that the following rate is achievable

$$\begin{aligned} & I(U;Y) - I(U;S) \\ = \; & I(X \oplus S; X \oplus S) - I(X \oplus S; S) \\ = \; & H(X \oplus S) - H(X \oplus S) + H(X) \\ = \; & H(X) = H(p) \leq 1, \end{aligned}$$

with equality if $p = 1/2$. This scheme is clearly optimal since $C^E \leq C^{E-D}$.

4. Consider a Gaussian multiple access channel with power constraints P_1, P_2 for the two encoders and noise power equal to 1. We want to design a scheme that achieves rates (R_1, R_2) such that $R_1 + R_2 = 1/2 \log_2(1 + P_1 + P_2)$ without using time-sharing between the corner points of the region or joint decoding, but simply simultaneous transmission with successive decoding. To do this, we proceed as follows.
4.1. Start from the lower corner (corresponding to decoding order 2→1) of the capacity region and define the corresponding rates as (R_1', R_2') (write the expressions). Now, consider a rate pair $(R_1' - \Delta, R_2' + \Delta)$ *on the capacity region boundary* where $R_1 + R_2 = 1/2 \log_2(1 + P_1 + P_2)$ for some $\Delta \geq 0$. Argue that you can always find a $0 \leq \alpha \leq 1$ such that

$$R_2' + \Delta = \frac{1}{2} \log_2 \left(1 + \frac{P_2}{1 + \alpha P_1}\right).$$

4.2. Calculate $R_1' - \Delta$ as a function of α by using the equation above. Based on this, show that the pair $(R_1' - \Delta, R_2' + \Delta)$ can be achieved as follows: Let user 1 transmit $X_1 = X_{11} + X_{12}$ with an appropriate power allocation between X_{11} and X_{12}; for decoding, treat X_{11} and X_{12} as the input of two virtual users and perform successive decoding. Find the power allocation and the decoding order at the receiver that achieves such rates (Hint: Write $R_1' - \Delta$ as the sum of two terms: $1/2 \log_2 (1 + \alpha P_1) + ...$).

Sol.: **4.1.** We have

$$\begin{aligned} R_1' &= \frac{1}{2} \log_2 (1 + P_1) \\ R_2' &= \frac{1}{2} \log_2 \left(1 + \frac{P_1}{1 + P_2}\right). \end{aligned}$$

Since

$$\frac{1}{2} \log_2 \left(1 + \frac{P_1}{1 + P_2}\right) \leq R_2' + \Delta \leq \frac{1}{2} \log_2 (1 + P_1),$$

we can always find the required α.
4.2. We obtain

$$\Delta = \frac{1}{2} \log_2 \left(1 + \frac{P_2}{1 + \alpha P_1}\right) - \frac{1}{2} \log_2 \left(1 + \frac{P_1}{1 + P_2}\right),$$

so that (after some simple manipulation)

$$R_1' - \Delta = \frac{1}{2} \log_2 (1 + \alpha P_1) + \frac{1}{2} \log_2 \left(1 + \frac{(1 - \alpha) P_1}{1 + \alpha P_1 + P_2}\right)$$

The pair of rates $(R_1' - \Delta, R_2' + \Delta)$ can be achieved by allocating αP_1 to X_{11} and $(1-\alpha)P_1$ to X_{12}. moreover, the decoding order is $X_{12} \to X_2 \to X_{11}$.

5. The Decode-and-Foward (DF) scheme seen in class is referred to as performing *regular* encoding since the rate of the source and relay codebooks are the same, say R. An alternative DF scheme employs *irregular* encoding, whereby in each block the source transmits a new message with rate R and the relay transmits (cooperatively with the source) with a rate $R' \leq R$ as follows: Having decoded the source codeword of the previous block w_{b-1}, the relay assigns a random index z_{b-1}, uniformly chosen from the set $[1, 2^{nR'}]$, to each possible decoded message. The index z_{b-1} is sent cooperatively by the relay and source (the codebook is as always known to all nodes), where the source uses the usual block-Markov coding scheme.
5.1. Assume a block-Markov coding transmission with $B+1=4$ blocks, where the source message is divided into $B=3$ submessages. Sketch the block diagram of the transmitted codewords for these four blocks. If the rate of each submessage sent by the source is R, what is the overall rate?
5.2. Propose a sliding window decoder, inspired by the one seen in class (i.e., propose a joint typicality-based decoding rule). Sketch the analysis of the probability of error of such decoder by isolating the main error events (notice that the index z_{b-1} is a function of w_{b-1}). Through this analysis, find conditions on R and R' that guarantee reliable communication.
5.3. What value of R' should be chosen in order to obtain the same performance as with regular encoding?

Sol.: **5.1.** Please see class notes. The difference here is that the relay transmits $x_{2b}^n(z_{b-1})$, where $z_{b-1} \in [1, 2^{nR'}]$ is the index obtained randomly from the decoded w_{b-1} and the source transmits $x_{1b}^n(z_{b-1}, w_b)$.
5.2. Consider the following simple successive decoding rule: First, decode z_b from block $b+1$ by finding the unique z_b such that

$$(x_{2b}^n(z_b), y_{3b+1}^n) \in T_\epsilon(P_{X_1X_2Y_3}).$$

This step can be made reliable if

$$R' < I(X_2; Y_3).$$

Then, look for a unique w_b such that

$$(x_{1b}^n(\hat{z}_{b-1}, w_b), x_{2b}^n(\hat{z}_{b-1}), y_{3b}^n) \in T_\epsilon(P_{X_1X_2Y_3})$$

and w_b is assigned index z_b ($\hat{z}_{b-1}$ is assumed to be correctly decoded from previous blocks). The probability of this step is easily seen to go to zero with n if

$$R < I(X_1; Y_3|X_2) + R'.$$

Alternatively, we can use a joint decoding rule (see Razaghi Yu 09).
5.3. If we choose $R' = I(X_2; Y_3) - \epsilon$, we obtain the performance of regular encoding. (With joint decoding, we could choose any R' such that $I(X_2; Y_3) \leq R' \leq R$, see Razaghi Yu 09).

6. Consider the network in figure with a single source s and destination d. Edges are labeled with their capacity.
6.1. Find the cut-set bound.

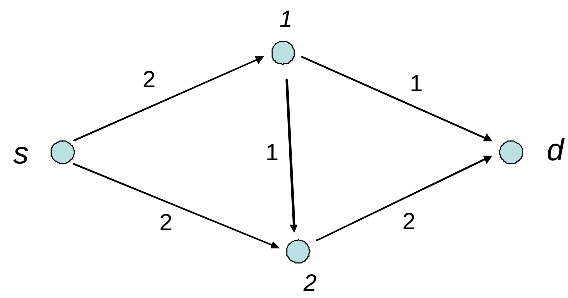

Figure 1:

6.2. Propose a scheme that achieves the cut-set bound.
6.3. Can the cut-set bound be achieved by network coding?

Sol.: **6.1**. The min-cut is 3.
6.2. We can use simple routing. For instance, one packet goes through s-1-d and two packets through s-2-d.
6.3. First of all, routing is a special case of network coding, so the answer is obviously yes. Moreover, we could use the (uselessly complicated) scheme $X_{s1} = b_1, X_{s2} = b_2 b_3, X_{12} = b_1, X_{1d} = b_1, X_{2d} = b_1 \oplus b_2, b_3$.