6.441 Transmission of Information April 27, 2006

Lecture 19

Lecturer: Madhu Sudan *Scribe: Mehmet Akçakaya*

1 Administrative Issues

- Project presentations in approximately 2 weeks from today.
- Report due in around 12 days.

2 Today

- Multiple Access Channels
- "Correlated Source Coding" a.k.a. Slepian-Wolf Theorem

3 Structure For Report/Presentation

- "Problem in English"
- Motivation - Why is this problem considered?
- Formal Model
- Theorem - Result - without going into the rigour at this point.
  At this point we've surpassed the attention span of most people in the audience.
- How? - Construction and Analysis (for the few who are still listening)

4 Multiple Access Channels

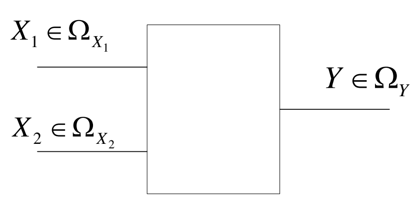

Figure 1: The Model

The multiple access channel is characterized by the input alphabets, Ω_{X_1} and Ω_{X_2}, the output alphabet, Ω_Y, and the transition probabilities, $p_{Y|(X_1,X_2)}$. We studied some specific channels in the last lecture, e.g. $Y = X_1 + X_2 + Z \bmod 2$, where all the alphabets were $\Omega = \{0, 1\}$.

Def (*Operational*): The rates (R_1, R_2) is achievable if $\exists$ encoding functions $X_1 : \{1, \ldots, 2^{nR_1}\} \to (\Omega_{X_1})^n$, $X_2 : \{1, \ldots, 2^{nR_2}\} \to (\Omega_{X_2})^n$ and decoding function $D : (\Omega_Y)^n \to \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\}$ such that $P_{\text{error}} \to 0$ as $n \to \infty$, meaning when the messages $W_1 \in_u \{1, \ldots, 2^{nR_1}\}$ and $W_2 \in_u \{1, \ldots, 2^{nR_2}\}$ are chosen independently, and we have $(W_1, W_2) \to (X_1(W_1), X_2(W_2)) \to Y \to (\widehat{W}_1, \widehat{W}_2)$, then $\mathbb{P}[(W_1, W_2) \neq (\widehat{W}_1, \widehat{W}_2)] \to 0$ as $n \to \infty$.

Def (*Basic Achievable*): The rates (R_1, R_2) is basic achievable if $\exists$ distributions p_{X_1}, p_{X_2} with $(X_1, X_2) \sim p_{X_1} p_{X_2}$, such that

$$R_1 \leq I(X_1;\, Y | X_2) \tag{1}$$
$$R_2 \leq I(X_2;\, Y | X_2) \tag{2}$$
$$R_1 + R_2 \leq I(X_1, X_2;\, Y) \tag{3}$$

Thm(*Capacity*): (R_1, R_2) is achievable if and only if it lies in the convex hull of the basic achievable rates $(\tilde{R}_1, \tilde{R}_2)$.

Def (*Convex Hull*): Given $(R_1^{(1)}, R_2^{(1)}), \ldots, (R_1^{(k)}, R_2^{(k)})$, the convex hull of these points are the points, (R_1, R_2) that can be written as:

$$R_1 = \sum_{i=1}^{k} \lambda_i R_1^{(i)}$$
$$R_2 = \sum_{i=1}^{k} \lambda_i R_2^{(i)}$$

where $\{\lambda_1, \ldots, \lambda_k : \lambda_j \geq 0,\ \sum_j \lambda_j = 1\}$. Examples can be seen in Fig. 2

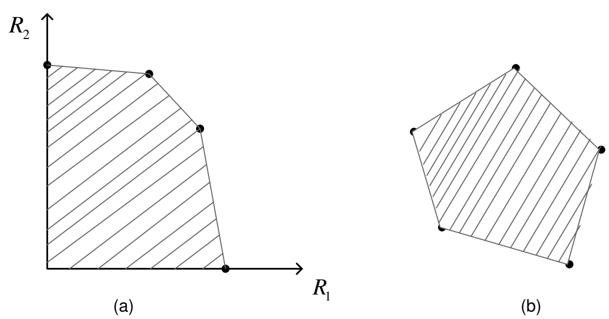

Figure 2: Examples of Convex Hulls: (a) with rates (b) in the plane

In other words, the theorem says (R_1, R_2) is achievable if and only if $\exists$ $(R_1^{(1)}, R_2^{(1)}), \ldots, (R_1^{(k)}, R_2^{(k)})$ basic achievable rates such that $R_1 = \sum_{i=1}^{k} \lambda_i R_1^{(i)}$ and $R_2 = \sum_{i=1}^{k} \lambda_i R_2^{(i)}$ with $\{\lambda_1, \ldots, \lambda_k : \lambda_j \geq 0,\ \sum_j \lambda_j = 1\}$.

Proof:
Achievability: We need

- Basic achievable pairs are achievable (shown via random coding and typical set decoding)
- Convex combinations are achievable (follows from a time-sharing argument)

Let $X_1(W_1)_i \sim p_{X_1}$ i.i.d. over W_1, i and $X_2(W_2)_i \sim p_{X_2}$ i.i.d. over W_2, i. Decoding function $D(Y)$ outputs (W_1, W_2) if $\exists!(W_1, W_2)$ such that $(X_1(W_1), X_2(W_2), Y)$ are jointly typical, else it outputs error.

When transmitting (W_1, W_2) a decoding error occurs when $(W_1', W_2') \neq (W_1, W_2)$ or the decoder outputs error:
- $(X_1(W_1), X_2(W_2), Y)$ is not jointly typical (by AEP the probability of this event $\to 0$ as $n \to \infty$).
- For $W_1' = W_1$, $W_2' \neq W_2$ (for fixed W_1, W_2), by joint AEP methods

$$\mathbb{P}[(X_1(W_1), X_2(W_2'), Y) \text{ is jointly typical}] \leq 2^{-nI(X_2;\,(X_1,Y))}$$

Thus the transmission will work if $R_2 \leq I(X_2; (X_1, Y)) = I(X_2; X_1) + I(X_2; Y|X_1) = I(X_2; Y|X_1)$. The last step follows since X_1 and X_2 are independent.

- Similar cases (i.e. $W_1' \neq W_1$, $W_2' = W_2$ and $W_1' \neq W_1$, $W_2' \neq W_2$) use similar inequalities.

Converse: Rigorous proof is omitted. But this follows from looking at the MAC in different ways: Looking at the MAC (Fig. 1) as a classical channel, i.e. point-to-point, we get $R_1 + R_2 \leq I(X_1, X_2; Y)$.

Alternatively we can look at it the other way (Fig. 3):

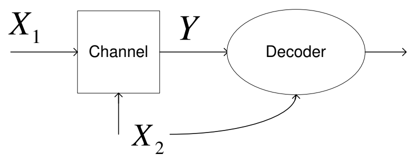

Figure 3: MAC viewed another way

In this case the decoder is more powerful than in the regular MAC case, since X_2 is available to it. We can view this as a point-to-point channel with additive noise X_2. Thus it follows reliable communication is possible only when $R_1 \leq I(X_1; Y|X_2)$. Since the decoder is more powerful than the MAC decoder, this will be an upper bound on the rate of communication possible with MAC.

5 Correlated Sources

The basic model is given in Fig. 4. Note that what makes this problem interesting is the fact that (X_1, X_2) are possibly dependent.

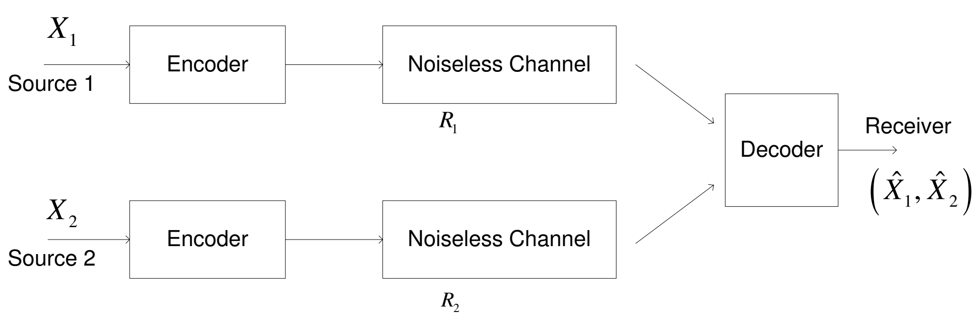

Figure 4: Correlated Sources Model

Ex: Let Z_0, Z_1, Z_2 be independent random variables with entropy H_0, H_1, H_2 respectively. Let $X_1 = (Z_0, Z_1)$ and $X_2 = (Z_0, Z_2)$ be the sources of interest. Note that $H(X_1) = H_0 + H_1$ and $H(X_2) = H_0 + H_2$.

It's easy to see that we can transmit at rates $R_1 = H_1$ and $R_2 = H_0 + H_2$, if we push all of Z_0 information through channel 2. Symmetrically we can transmit at $R_1 = H_0 + H_1$ and $R_2 = H_2$, if we push all of Z_0 information through cahhnel 1.

It follows naturally via time-sharing that we can transmit at the $R_1 = \alpha H_0 + H_1$ and $R_2 = (1 - \alpha)H_0 + H_2$, for $0 \leq \alpha \leq 1$, by proportionately transmitting Z_0 information through channel 1 and channel 2.

Based on this example, we can hope (conjecture) that rates (R_1, R_2) are achievable if $R_1 \geq H(X_1|X_2)$, $R_2 \geq H(X_2|X_1)$ and $R_1 + R_2 \geq H(X_1, X_2)$. In fact this turns out to be the statement of our main theorem.

Thm(*Slepian-Wolf*): In the correlated sources model, rates (R_1, R_2) are achievable if and only if

$$R_1 \geq H(X_1|X_2) \tag{4}$$
$$R_2 \geq H(X_2|X_1) \tag{5}$$
$$R_1 + R_2 \geq H(X_1, X_2) \tag{6}$$

The idea is to transmit only the jointly typical sequences (X_1, X_2). This idea is illustrated in Fig. 5. Note that $H_1 = H(X_1)$, $H_2 = H(X_2)$, $I = I(X_1; X_2)$.

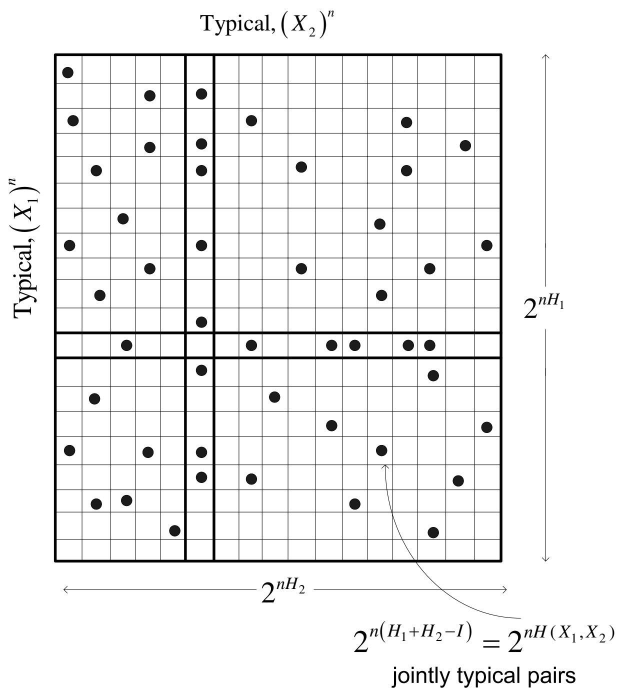

Figure 5: Slepian-Wolf Encoding

From the figure, we can infer the following:

$$\text{\# dots per row} = \frac{\text{\# dots}}{\text{\# rows}} = \frac{2^{n(H_1+H_2-I)}}{2^{nH_1}} = 2^{n(H_2-I)} = 2^{nH(X_2|X_1)} \tag{7}$$

Similarly we'll have # dots per column $= 2^{n(H_1-I)} = 2^{nH(X_1|X_2)}$. The random coding argument goes as follows: We need to assign indices to each row, but we don't have 2^{nH_1} indices. Thus for each row, we pick an index randomly from $\{1, \ldots, 2^{nR_1}\}$. We do the same thing for the columns. Decoder will get "boxes" defined by the indices. If there's only one typical element in the box, then we output that element. If there's zero or more than one, then we declare an error. Formal proof will be given in the next lecture.