Iowa State University Lecturer: Aditya Ramamoorthy

EE 520: Topics in Communications: Network Coding

September 23, 2008 Scribe: Shurui Huang

Lecture 9

In this lecture, we will discuss the problem of lossless distributed source coding, also known as the Slepian Wolf Problem. let X and Y be two correlated discrete memoryless sources with joint distribution $P_{XY}(\cdot)$. It is obvious that a total rate $R = H(X,Y)$ is sufficient if we encode the sources together. However the result of Slepian and Wolf shows interestingly enough that a total rate $R = H(X,Y)$ is sufficient even when the sources are encoded separately (a block diagram is given in Fig. 1).

We begin with the proof of achievability. i.e. we find a set of rate pairs (R_1, R_2) such that $P((\hat{X}^n, \hat{Y}^n) \neq (X^n, Y^n)) \to 0$ as $n \to \infty$.

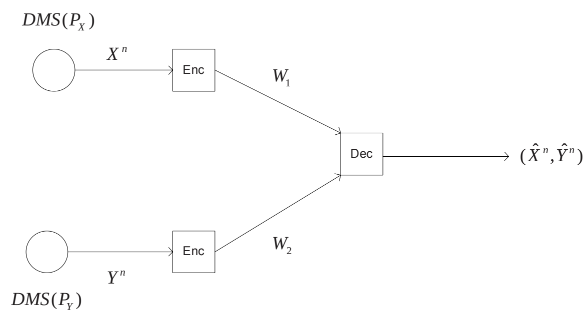

Figure 1: System Diagram

1 Achievability

The idea of the achievability scheme is to use random binning. Basically every typical x^n and y^n) sequence is assigned a bin and the encoders transmit the bin index. The decoder exploits the fact that there are are approximately $nH(X,Y)$ jointly typical sequences for performing unique decoding. This is illustrated in Fig. 2. Note that there will be approximately $2^{nH(X)}$ typical x^n sequences, $2^{nH(Y)}$ typical y^n sequences and $2^{nH(X,Y)}$ typical (x^n, y^n) sequences. If the bin sizes are chosen such that each bin pair contains at most one jointly typical (x^n, y^n), then the decoding will be correct.

1.1 Code Construction

Generate $2^{n(R_1+R_1')}$ codewords $x^n(w_1, v_1)$ where $w_1 = 1, 2 \cdots 2^{nR_1}$, $v_1 = 1, 2 \cdots 2^{nR_1'}$, by choosing $n2^{n(R_1+R_1')}$ symbols i.i.d. from $P_X(\cdot)$. Similarly generate $2^{n(R_2+R_2')}$ codewords $y^n(w_2, v_2)$ i.i.d. from $P_Y(\cdot)$.

1.2 Encoder

Given x^n, try to find (w_1, v_1) s.t. $x^n = X^n(w_1, v_1)$, if match is found, send w_1, otherwise, send 1. Given y^n, try to find (w_2, v_2) s.t. $y^n = Y^n(w_2, v_2)$, if match is found, send w_2, otherwise, send 1.

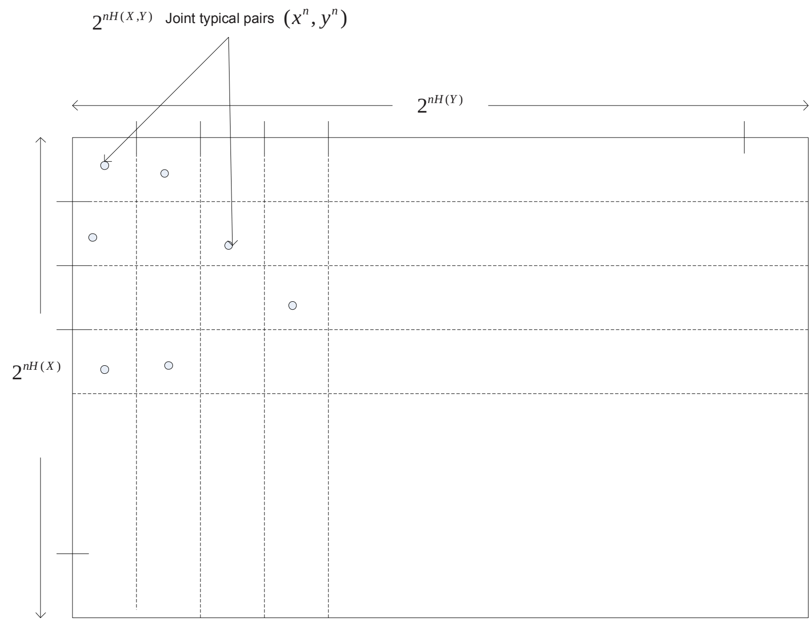

Figure 2: Random binning

1.3 Decoder

Given (w_1, w_2), try to find $(\hat{v}_1, \hat{v}_2)$, such that $(x^n(w_1, \hat{v}_1), y^n(w_2, \hat{v}_2)) \in T_\epsilon^n(P_{XY})$. If successful, then declare$(x^n(w_1, \hat{v}_1), y^n(w_2, \hat{v}_2))$ to be correct transmitted codeword. Otherwise, just put $(x^n(w_1, 1), y^n(w_2, 1))$.

1.4 Analysis

Let $0 < \epsilon_1 < \epsilon < \mu_{XY}$. We consider the following cases.

(i) *Case 1.* The pair (x^n, y^n) is not jointly typical. We know that the probability of this goes to zero. i.e.

$P((x^n, y^n) \notin T_{\epsilon_1}^n(P_{XY})) \leq \delta_{\epsilon_1}(n) \to 0, \ as \ n \to \infty$

(ii) *Case 2.* Suppose that $(x^n, y^n) \in T_{\epsilon_1}^n(P_{XY})$, but encoder 1 does not find (w_1, v_1) such that $x^n = X^n(w_1, v_1)$. The probability of this event can be upper bounded as

$(1 - P_X^n(x))^{2^{n(R_1+R_1')}} \leq \exp\{-2^{n(R_1+R_1')} P_X^n(x^n)\} \leq \exp\{-2^{n(R_1+R_1')} 2^{-nH(X)(1+\epsilon_1)}\}.$

A sufficient condition for the probability going to zero is that

$$R_1' = H(X) - R_1 + 2\epsilon_1 H(X, Y),$$
$$R_2' = H(Y) - R_2 + 2\epsilon_1 H(X, Y).$$

(iii) *Case 3.* The previous cases 1 & 2 do not happen, but there exists another pair of jointly typical typical sequences in the same pair of bins. i.e. suppose $x^n = X^n(w_1, v_1), y^n = Y^n(w_2, v_2)$ but there exists $\Rightarrow \exists(\widetilde{v}_1, \widetilde{v}_2) \neq (v_1, v_2)$ such that $(X^n(w_1, \widetilde{v}_1), Y^n(w_2, \widetilde{v}_2)) \in T_\epsilon^n(P_{XY})$. This will happen if the bins are too wide, or equivalently the number of bits used for R_1 and R_2 is not enough. We have,

$$P[\bigcup_{(\widetilde{v}_1,\widetilde{v}_2)\neq(v_1,v_2)}\{(X^n(w_1,\widetilde{v}_1),Y^n(w_2,\widetilde{v}_2))\in T_\epsilon^n(P_{XY})\}]\leq\sum_{\widetilde{v}_1\neq v_1}P\{(X^n(w_1,\widetilde{v}_1),Y^n(w_2,v_2))\in T_\epsilon^n(P_{XY})\}$$
$$+\sum_{\widetilde{v}_2\neq v_2}P\{(X^n(w_1,v_1),Y^n(w_2,\widetilde{v}_2))\in T_\epsilon^n(P_{XY})\}+\sum_{\widetilde{v}_1\neq v_1}\sum_{\widetilde{v}_2\neq v_2}P((X^n,Y^n)\in T_\epsilon^n(P_{XY})).$$

We now upper bound each of three terms on the RHS of the above expression.

$$\sum_{\widetilde{v}_1\neq v_1}P\{(X^n(w_1,\widetilde{v}_1),Y^n(w_2,v_2))\in T_\epsilon^n(P_{XY})\}\leq 2^{-n(I(X;Y)-2\epsilon H(X))}2^{nR_1'}. \tag{1}$$

Now substituting for R_1', we can conclude that if

$$R_1>H(X|Y)+4\epsilon H(X,Y)$$

the RHS of (1) goes to zero. Similarly, we can conclude that

$$R_2>H(Y|X)+4\epsilon H(X,Y)$$

makes the second term go to zero. Finally

$$\Rightarrow\sum_{\widetilde{v}_1\neq v_1}\sum_{\widetilde{v}_2\neq v_2}P((X^n,Y^n)\in T_\epsilon^n(P_{XY}))\leq -2^{n(R_1'+R_2')}2^{-n(I(X;Y)-3\epsilon H(X,Y))}$$

which implies that we need

$$\Rightarrow R_1+R_2>H(X,Y)+7\epsilon H(X,Y).$$

Therefore the achievable region becomes

$$R_1>H(X|Y),$$
$$R_2>H(Y|X),\ \text{and}$$
$$R_1+R_2>H(X,Y).$$

2 Converse

The converse of Slepian Wolf problem is defined as following: for any distributed source coding problem for the source (X,Y) drawn i.i.d $P_{XY}(\cdot)$, if the error probability $P((\hat{X}^n,\hat{Y}^n)\neq(X^n,Y^n))\to 0$ as $n\to\infty$, then we must have

$$R_1>H(X|Y),$$
$$R_2>H(Y|X),\ \text{and}$$
$$R_1+R_2>H(X,Y).$$

$$
\begin{aligned}
nR_1 &\overset{(a)}{\geq} H(W_1) \\
&\geq H(W_1|Y^n) \\
&\overset{(b)}{=} H(W_1|Y^n) - H(W_1|X^n, Y^n) \\
&= I(W_1; X^n|Y^n) \\
&= H(X^n|Y^n) - H(X^n|W_1, Y^n)
\end{aligned}
$$

We now focus on the second term in the RHS above

$$
\begin{aligned}
H(X^n|W_1, Y^n) &\overset{(c)}{=} H(X^n|W_1, Y^n, W_2) \\
&\overset{(d)}{=} H(X^n|W_1, W_2, Y^n, \hat{X}^n, \hat{Y}^n) \\
&= H(X^n, Y^n|W_1, W_2, Y^n, \hat{X}^n, \hat{Y}^n) \\
&\leq H(X^n, Y^n|\hat{X}^n, \hat{Y}^n)
\end{aligned}
$$

Thus we have,

$$
\begin{aligned}
nR_1 &\overset{(e)}{\geq} nH(X|Y) - H(X^n, Y^n|\hat{X}^n, \hat{Y}^n) \\
&\overset{(f)}{\geq} nH(X|Y) - n[P_e \log(|\mathcal{X}||\mathcal{Y}|)] - H_2(P_e) \\
&\overset{(g)}{=} nH(X, Y) - n\epsilon_n.
\end{aligned}
$$

(a) follows the fact that $W_1 \in \{1, 2, \ldots, 2^{nR_1}\}$

(b) follows the fact that W_1 is a function of X^n, the conditional entropy is 0.

(c) follows the fact that W_2 is a function of Y^n.

(d) follows the fact that $\hat{X}^n, \hat{Y}^n$ are the functions of W_1, W_2,respectivly.

(e) from the chain rule and the fact that (X_i, Y_i) are i.i.d.

(f)(g) from the Fano's inequality.

$$
H(X^n, Y^n|\hat{X}^n, \hat{Y}^n) \leq nP_e(\log|\mathcal{X}||\mathcal{Y}|) + H_2(P_e) = n\epsilon_n
$$

where $\epsilon_n \to 0$ as $n \to \infty$.

For arbitrarily small P_e, we need $R_1 > H(X|Y)$, in the same manner, $R_2 > H(Y|X)$
Similarly, by the same argument for the equalities and inequalities,

$$
R_1 + R_2 > H(X, Y)
$$

$$
\begin{aligned}
n(R_1 + R_2) &\geq H(W_1, W_2) \\
&= H(W_1, W_2) - H(W_1, W_2|X^n, Y^n) \\
&= I((W_1, W_2); (X^n, Y^n)) \\
&= H(X^n, Y^n) - H(X^n, Y^n|W_1, W_2)
\end{aligned}
$$

$$
\begin{aligned}
H(X^n, Y^n|W_1, W_2) &= H(X^n, Y^n|W_1, W_2, \hat{Y}^n, \hat{X}^n) \\
&\leq H(X^n, Y^n|\hat{X}^n, \hat{Y}^n)
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow n(R_1 + R_2) &\geq nH(X, Y) - H(X^n, Y^n|\hat{X}^n, \hat{Y}^n) \\
&\geq nH(X, Y) - n[P_e \log(|\mathcal{X}||\mathcal{Y}|)] - H_2(P_e) \\
&= nH(X, Y) - n\epsilon_n
\end{aligned}
$$

Dividing these inequalities by n, and taking the limit as $n \to \infty$, we have the desired converse.