E2–301 Topics in Multiuser Communication October 18, 2007

Lecture 17 : Multi–terminal Distributed Source Coding

Instructor: Rajesh Sundaresan *Scribe: Premkumar K.*

1 Example:

- Temperatures measured at Palace road meteorological centre and at IISc. Call them X, Y.
- Temperatures X and Y to be sent to New Delhi.
- IISc and Palace road centre, being geographically separated, have to do separate encoding.
- The simplest method is $R_1 = H(X)$ and $R_2 = H(Y)$.
- When X and Y are independent, one can not do any better than the above, since even if they cooperate, to obtain $\widehat{X}\widehat{Y}$ with high reliability $R_1 + R_2 \geqslant H(XY) = H(X) + H(Y)$.
- $R_1 = H(X), R_2 = H(Y)$ achieves it with no need for cooperation.
- Suppose

$$Y = \begin{cases} X & w.p. 1/2 \\ X - 1 & w.p. 1/2 \end{cases}$$

  IISc can send just one bit; odd or even. 1 bit $= H(Y|X)$, achievable even if X is unknown.
- Key: universal source compression.

2 Definitions

Definition 1 (DMS). *A (two user) <u>discrete memoryless source</u> (DMS) denoted by* $\left(\mathbb{X}_1, \mathbb{X}_2\right)$, *consists of two finite sets* $\mathbb{X}_1$ *and* $\mathbb{X}_2$, *with the interpretation that* X_k *is the input to encoder* k, $k = 1, 2$ *and for* $n \in \mathbb{N}$, *with* $X_k^n = (X_{k1}, X_{k2}, \cdots, X_{kn})$, $k = 1, 2$ *has a pmf*

$$p_{X_{[2]}^n}(x_{[2]}^n) = \prod_{i=1}^{n} p_{X_{[2]}}(x_{[2]i}) \qquad (1)$$

Note: The successive output symbols of the source are independently drawn from $\mathbb{X}_{[2]}$ where as there could be a correlation between the components of $X_{[2],i}$.

Definition 2 (Code). *An* (n, M_1, M_2) *<u>distributed source code</u> for the DMS* $(\mathbb{X}_1, \mathbb{X}_2)$ *consists of the following:*

1. *An index set of messages for each user k, $\mathbb{W}_k = [M_k]$.*

2. *An encoder f_k for each user k, $f_k : \mathbb{X}_k^n \to \mathbb{W}_k, k = 1, 2$. Note that $\mathbb{X}_k^n \ni X_k^n \mapsto f_k(X_k^n) \in \mathbb{W}_k$.*

3. *A decoding rule, $g : \mathbb{W}_1 \times \mathbb{W}_2 \to \phi \cup \left(\mathbb{X}_1^n \times \mathbb{X}_2^n\right)$, i.e., $W_{[2]} \mapsto g(W_{[2]}) = \widehat{X}_{[2]}^n \in \phi \cup \left(\mathbb{X}_1^n \times \mathbb{X}_2^n\right)$.*

Definition 3 (Probability of error). *Let $W_{[2]}^n$ be the message transmitted. The probability of error for the distributed source code c (when the symbols $X_{[2]}^n$ come from a DMS) is given by*

$$P_e^{(n)}(c) \quad = \quad Pr\left\{g\left(f_1(X_1^n), f_2(X_2^n)\right) \neq X_1^n X_2^n\right\}.$$

Definition 4 (Achievability). *The rate pair (R_1, R_2) is achievable, if for every $\eta > 0, \lambda \in (0, 1)$, there exists a sequence of (n, M_1, M_2) distributed source codes that satisfy*

1. *$P_e^{(n)} \underset{n}{\leqslant} \lambda$, and*

2. *$\frac{\log_2 M_k}{n} \underset{n}{\leqslant} R_k + \eta$.*

Definition 5 (Achievable rate region). *The achievable rate region is the set of all achievable rate pairs.*

Lemma 1. *The achievable rate region is a closed convex set.*

Proof. Exercise. □

Theorem 2. *(Slepian–Wolf) The achievable rate region is*

$$\left\{(R_1, R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{rcl} R_1 & \geqslant & H(X_1|X_2) \\ R_2 & \geqslant & H(X_2|X_1) \\ R_1 + R_2 & \geqslant & H(X_1 X_2) \end{array}\right\}$$

Remark 1.

- *Joint decoding with X_1^n supplied to encoder 2 and vice versa will also yield this rate region.*
- *SW theorem says we can do this without the knowledge of encoder 2's observation (and similarly encoder 1's observation is not known at encoder 2).*
- *Consider Example 1. Palace road compresses $H(X)$ bits. IISc compresses to $H(Y|X) = 1$ bit. Code: IISc indicates odd/even.*
- *Since $H(X_2|X_1)$ is known, we have a non–stationary but independent (over time) source $\prod_{i=1}^{n} p(y_i|x_{1i})$, given x_1^n. Of course x_1^n is not observed. We have a universal code to compress this independent non-stationary source at its average entropy $H(X_2|X_1)$. Someone who knows X_1^n now reconstructs X_2^n.*
- *Not too surprising, since we know of the existence of universal codes for stationary and ergodic sources (Lempel-Ziv, fixed rate universal code, refer ITC-2 course notes).*
- *Universal compression at rate $R = H(X)$. Consider $\lfloor 2^{n(R+\eta)} \rfloor = M$ bins.*
- *For each $x^n \in \mathbb{X}^n$, assign a bin among $[M]$ uniformly. $f(x^n) = bin\#$. Reveal f to both encoder and decoder.*
- *Encoding is transmission of index $f(x^n)$ with $\frac{\log M}{n} < R + \eta$ bits/sample. Decoding: Look for a unique typical $\widehat{x}^n$ in bin $f(x^n)$. If none or more than 1, map to ϕ.*

- **Error analysis:** $\mathbb{E}_c Pr\{g(f(X^n) \neq X^n)\} \leqslant Pr\{X^n \notin T_\delta^n(X)\} + Pr\{E_{21}\}$, *where* E_{21} *is the event that another of the* $T_\delta^n(X) - 1$ *elements in* $T_\delta^n(X)$ *falls in the bin* $f(x^n)$.

$$\begin{aligned} Pr\{E_{21}\} &\leqslant \left(T_\delta^n(X) - 1\right)\frac{1}{M} \underset{n}{\leqslant} 2^{nH(X)+2n\delta} \cdot 2^{-nR-n\eta} \cdot 2^{n\eta/2} \\ &= 2^{-n(\eta/2-2\delta)} \downarrow 0 \quad \text{if } \eta > 4\delta. \end{aligned}$$

- *Any source with* $H(X) \leqslant R$ *can b compressed to rate* R *without knowledge of source. Of course, this i.i.d. property is lost in SW problem.*

  We now extend this to multi–terminal systems.

Proof. (Achievability.)

$$M_k = \lfloor 2^{n(R_k+\eta)} \rfloor, \ k = 1, 2$$

$(R_1, R_2) \in$ region in the SW theorem.
Random code: Assign x_k^n to one of bins $1, 2, \cdots, M_k$, independent of the sequence chosen and uniformly in the bins.

- $f_k(x_k^n) = \text{bin}\#$.
- Reveal f_1, f_2 to both encoders and decoder.
- Encoding is clear. Send $f_k(x_k^n)$ using $\frac{\log M_k}{n} < R_k + \eta$ bits/sample.

Decoding: Given bins $f_k(x_k^n), k = 1, 2$, look for a jointly typical $\widehat{x}_1^n \widehat{x}_2^n$ in the joint bin. Moreover, they should satisfy $\widehat{x}_k^n \in T_{\frac{\delta}{2}}^{(n)}(X_k)$.

- $\mathbb{E}_c P_e^{(n)}$:

$$\begin{aligned} \text{Error} \ \Leftrightarrow \ & E_0 && X_{[2]}^n \notin T_\delta^{(n)}(X_{[2]}) \text{ or } X_k^n \notin T_{\frac{\delta}{2}}^{(n)}(X_k) \\ & \cup E_1 && \exists \widehat{x}_1^n \neq x_1^n \text{s.t.} (\widehat{x}_1^n, x_2^n) \in T_\delta^{(n)}(X_{[2]}) \\ & \cup E_2 && \exists \widehat{x}_2^n \neq x_2^n \text{s.t.} (x_1^n, \widehat{x}_2^n) \in T_\delta^{(n)}(X_{[2]}) \\ & \cup E_{12} && \exists \widehat{x}_1^n \neq x_1^n, \widehat{x}_2^n \neq x_2^n, \text{s.t.} (\widehat{x}_1^n, \widehat{x}_2^n) \in T_\delta^{(n)}(X_{[2]}) \end{aligned}$$

$$\begin{aligned} \mathbb{E} P_e^{(n)} &\leqslant Pr\{E_0\} + Pr\{E_1 | E_0^c\} + Pr\{E_2 | E_0^c\} + Pr\{E_{12} | E_0^c\} \\ Pr\{E_0\} &\underset{n}{\leqslant} 3\delta \\ Pr\{E_1 | E_0^c\} &\leqslant \left| T_\delta^{(n)}(X_1 | x_2^n) \right| \frac{1}{M_1} \\ &\leqslant 2^{nH(X_1|X_2)+n\delta} \cdot 2^{-nR_1 - n\eta} \cdot 2^{n\eta/2} \\ &\leqslant 2^{-n(\eta/2-\delta)} \downarrow 0, \ \text{if} \eta > 2\delta. \\ \text{similarly}, Pr\{E_2 | E_0^c\} &\leqslant 2^{-n(\eta/2-\delta)} \downarrow 0. \\ Pr\{E_{12} | E_0^c\} = Pr\{E_{12}\} &\leqslant \left| T_\delta^{(n)}(X_1 X_2) \right| \frac{1}{M_1 M_2} \leqslant 2^{nH(X_1 X_2)+n\delta} \cdot 2^{-n(R_1+R_2+\eta+\eta)} \cdot 2^{n\eta/2} \\ &\leqslant 2^{-n(3\eta/2-\delta)} \downarrow 0 \quad \text{if } \eta > 2\delta/3. \end{aligned}$$

□