Lecture #3  Distributed Lossless Compression

(Reading: NIT 10.1–10.5, 4.4)

- Distributed lossless source coding
- Lossless source coding via random binning
- Time sharing
- Achievability proof of the Slepian–Wolf theorem
- Extension to more than two sources

Two-component DMS (2-DMS) \((\mathcal{X}_1 \times \mathcal{X}_2, p(x_1, x_2))\)

- A \((2^{nR_1}, 2^{nR_2}, n)\) code:
  - **Two encoders:** \(m_1(x_1^n) \in [1: 2^{nR_1}]\) and \(m_2(x_2^n) \in [1: 2^{nR_2}]\)
  - **Decoder** \((\hat{x}_1^n, \hat{x}_2^n)(m_1, m_2)\)

- **Probability of error:** \(P_e^{(n)} = P\{ (\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n) \} \)

- \((R_1, R_2)\) achievable if \(\exists (2^{nR_1}, 2^{nR_2}, n)\) codes such that \(\lim_{n \to \infty} P_e^{(n)} = 0\)

- **Optimal rate region** \(R^*\): Closure of the set of achievable \((R_1, R_2)\)
Bounds on the optimal rate region

- Sufficient condition for individual compression:
  \[ R_1 > H(X_1), \quad R_2 > H(X_2) \]

- Necessary condition for centralized compression:
  \[ R_1 + R_2 \geq H(X_1, X_2) \]

- Can also show that
  \[ R_1 \geq H(X_1|X_2), \quad R_2 \geq H(X_2|X_1) \] are necessary

Slepian–Wolf theorem

Theorem 10.1 (Slepian–Wolf 1973)
The optimal rate region \( \mathcal{R}^* \) is the set of \((R_1, R_2)\) such that

\[
\begin{align*}
R_1 &\geq H(X_1|X_2), \\
R_2 &\geq H(X_2|X_1), \\
R_1 + R_2 &\geq H(X_1, X_2)
\end{align*}
\]
Example

- **Doubly symmetric binary source** (DSBS(p)) \((X_1, X_2)\)

  ![Doubly symmetric binary source diagram]

- Let \(p = 0.01\)
- **Individual compression**: 2 bits/symbol-pair
- **Slepian–Wolf coding**: \(H(X_1, X_2) = 1.0808\) bits/symbol-pair

---

**Lossless source coding via random binning**

- **Codebook generation**:
  - Randomly assign an index \(m(x^n) \in [1:2^{nR}]\) to each sequence \(x^n \in \mathcal{X}^n\)
  - The set of sequences with the same index \(m\) form a bin \(B(m), m \in [1:2^{nR}]\)
  - Bin assignments are revealed to the encoder and decoder

- **Encoding**:
  - Upon observing \(x^n \in B(m)\), send the bin index \(m\)

- **Decoding**:
  - Find the unique typical sequence \(\hat{x}^n \in B(m)\)
Lossless source coding via random binning

• Codebook generation:
  - Randomly assign an index \( m(x^n) \in [1: 2^{nR}] \) to each sequence \( x^n \in \mathcal{X}^n \)
  - The set of sequences with the same index \( m \) form a bin \( B(m), m \in [1: 2^{nR}] \)
  - Bin assignments are revealed to the encoder and decoder

• Encoding:
  - Upon observing \( x^n \in B(m) \), send the bin index \( m \)

• Decoding:
  - Find the unique typical sequence \( \hat{x}^n \in B(m) \)
Analysis of the probability of error

- We bound $P_e^{(n)}$ averaged over random bin assignments
- Let $M$ denote the random bin index of $X^n$, i.e., $X^n \in B(M)$
- Note that $M \sim \text{Unif}[1 : 2^{nR}]$, independent of $X^n$
- Error events:
  \[ E_1 = \{ X^n \notin T_e^{(n)} \}, \text{ or } \]
  \[ E_2 = \{ \tilde{x}^n \in B(M) \text{ for some } \tilde{x}^n \neq X^n, \tilde{x}^n \in T_e^{(n)} \} \]

Thus
\[
P(E) \leq P(E_1) + P(E_2)
\]

- By the LLN, $P(E_1) \to 0$

Analysis of the probability of error

- Consider
\[
P(E_2 | X^n \in B(1)) = \sum_{x^n} P\{X^n = x^n | X^n \in B(1)\} \cdot P\{\tilde{x}^n \in B(1) \text{ for some } \tilde{x}^n \neq x^n, \tilde{x}^n \in T_e^{(n)} | x^n \in B(1), X^n = x^n\}
\]
\[
\leq \sum_{x^n} p(x^n) \sum_{\tilde{x}^n \in T_e^{(n)}} P\{\tilde{x}^n \in B(1) | x^n \in B(1), X^n = x^n\}
\]
\[
= \sum_{x^n} p(x^n) \sum_{\tilde{x}^n \in T_e^{(n)}} P\{\tilde{x}^n \in B(1)\}
\]
\[
\leq |T_e^{(n)}| \cdot 2^{-nR}
\]
\[
\leq 2^{n(H(X) + \delta(\varepsilon))} 2^{-nR}
\]

- Hence $P(E_2) \to 0$ as $n \to \infty$ if $R > H(X) + \delta(\varepsilon)$
Achievability via linear binning

- Let $X$ be a Bern($p$) source
- $R = H(X)$ achieved via linear binning (hashing)
  - Let $H$ be a randomly generated $nR \times n$ binary parity-check matrix
  - Encoder sends $HX^n$
  - Decoder recovers $X^n$ with high probability if $R > H(p)$ (why?)

Time sharing

Proposition 4.1

If $(R_1', R_2')$, $(R_1'', R_2'') \in \mathcal{R}^*$, then $(R_1, R_2) = (\alpha R_1' + \bar{\alpha} R_1'', \alpha R_2' + \bar{\alpha} R_2'') \in \mathcal{R}^*$ for $\alpha \in [0, 1]$

- The rate region $\mathcal{R}^*$ is convex
**Proposition 4.1**

If \((R'_1, R'_2), (R''_1, R''_2) \in \mathcal{R}^*\), then \((R_1, R_2) = (\alpha R'_1 + \bar{\alpha} R''_1, \alpha R'_2 + \bar{\alpha} R''_2) \in \mathcal{R}^*\) for \(\alpha \in [0, 1]\)

- The rate region \(\mathcal{R}^*\) is convex
- **Proof:** Time sharing argument
  - Let \(C'_k\) be a sequence of \((2^k R'_1, 2^k R'_2, k)\) codes with \(P^{(k)}_{e1} \to 0\)
  - Let \(C''_k\) be a sequence of \((2^k R''_1, 2^k R''_2, k)\) codes with \(P^{(k)}_{e2} \to 0\)
  - Construct a new sequence of codes by using \(C'_i\) for \(i \in [1 : an]\) and \(C''_i\) for \(i \in [an + 1 : n]\)

- By the union of events bound, \(P^{(n)}_e \leq P^{(an)}_{e1} + P^{(\bar{a}n)}_{e2} \to 0\)

- **Remarks:**
  - **Time division** is a special case of time sharing (between \((R_1, 0)\) and \((0, R_2)\))
  - The rate (capacity) region of any source (channel) coding problem is convex

**Achievability proof of the S–W theorem (Cover 1975)**

- We show that the corner point \((H(X_1), H(X_2|X_1))\) is achievable
- Achievability of the other corner point \((H(X_1|X_2), H(X_2))\) follows similarly
- The rest of the region is achieved using time sharing
We show that the corner point \((H(X_1), H(X_2|X_1))\) is achievable.

Achievability of the other corner point \((H(X_1|X_2), H(X_2))\) follows similarly.

The rest of the region is achieved using time sharing.
Achievability of \((H(X_1), H(X_2|X_1))\)

- **Codebook generation:**
  - Assign a distinct index \(m_1 \in [1 : 2^{nR_1}]\) to each \(x_1^n \in \mathcal{T}_e^{(n)}(X_1)\), and \(m_1 = 1\), otherwise
  - Randomly assign an index \(m_2(x_2^n) \in [1 : 2^{nR_2}]\) to each \(x_2^n \in \mathcal{X}_2^n\)
  - The sequences with the same index \(m_2\) form a bin \(B(m_2)\)

- **Encoding:**
  - Upon observing \(x_1^n\), encoder 1 sends the index \(m_1(x_1^n)\)
  - Upon observing \(x_2^n \in B(m_2)\), encoder 2 sends \(m_2\)
Achievability of \((H(X_1), H(X_2|X_1))\)

- **Decoding**: Sources recovered successively
  - Declare \(\hat{x}_1^n = x_1^n(m_1)\) for the unique \(x_1^n(m_1) \in \mathcal{T}_e(n)(X_1)\)
  - Find the unique \(\hat{x}_2^n \in \mathcal{B}(m_2) \cap \mathcal{T}_e(n)(X_2)|\hat{x}_1^n)\)
  - If there is none or more than one, the decoder declares an error

![](image.png)

Analysis of the probability of error

- Let \(M_1\) and \(M_2\) denote the random bin indices for \(X_1^n\) and \(X_2^n\)
- Error events:
  
  \[
  \mathcal{E}_1 = \{(X_1^n, X_2^n) \notin \mathcal{T}_e(n)\},
  \]
  
  \[
  \mathcal{E}_2 = \{\hat{x}_2^n \in \mathcal{B}(M_2) \text{ for some } \hat{x}_2^n \neq X_2^n, (X_1^n, \hat{x}_2^n) \in \mathcal{T}_e(n)\}\]

  Then,

  \[
P(\mathcal{E}) \leq P(\mathcal{E}_1) + P(\mathcal{E}_2)
  \]

- \(P(\mathcal{E}_1) \to 0\) by LLN
Analysis of the probability of error

- By symmetry,
  \[ P(E_2) = P(E_2 | X_2^n \in B(1)) \]
  \[ = \sum_{(x_1^n, x_2^n)} P\{ (X_1^n, X_2^n) = (x_1^n, x_2^n) | X_2^n \in B(1) \} \cdot P\{ \tilde{x}_2^n \in B(1) \text{ for some } \tilde{x}_2^n \neq x_2^n, (x_1^n, \tilde{x}_2^n) \in T_e^{(n)} | x_2^n \in B(1), (X_1^n, X_2^n) = (x_1^n, x_2^n) \} \]
  \[ \leq \sum_{(x_1^n, x_2^n)} p(x_1^n, x_2^n) \sum_{\tilde{x}_2^n \in T_e^{(n)}(x_2^n)} P\{ \tilde{x}_2^n \in B(1) \} \]
  \[ \leq 2^{n(H(X_2 | X_1) + \delta(\epsilon))} 2^{-nR_2} \]

- Hence, \( P(E_2) \to 0 \) as \( n \to \infty \) if \( R_2 > H(X_2 | X_1) + \delta(\epsilon) \)

- Remark: Achievability can be proved without time-sharing (NIT 10.3.2)

Extension to more than two sources

Theorem 10.3

The optimal rate region \( R^*(X_1, \ldots, X_k) \) for the \( k \)-DMS \( (X_1, \ldots, X_k) \) is the set of \( (R_1, \ldots, R_k) \) such that

\[ \sum_{j \in S} R_j \geq H(X(S) | X(S^c)) \quad \text{for all } S \subseteq [1 : k] \]

- For \( k = 3 \), \( R^*(X_1, X_2, X_3) \) is the set of \( (R_1, R_2, R_3) \) such that

  \[ R_1 \geq H(X_1 | X_2, X_3), \]
  \[ R_2 \geq H(X_2 | X_1, X_3), \]
  \[ R_3 \geq H(X_3 | X_1, X_2), \]
  \[ R_1 + R_2 \geq H(X_1, X_2 | X_3), \]
  \[ R_1 + R_3 \geq H(X_1, X_3 | X_2), \]
  \[ R_2 + R_3 \geq H(X_2, X_3 | X_1), \]
  \[ R_1 + R_2 + R_3 \geq H(X_1, X_2, X_3) \]
Summary

- $k$-Component discrete memoryless source ($k$-DMS)
- Distributed lossless source coding for a $k$-DMS:
  - Slepian–Wolf optimal rate region
  - Random binning
- Time sharing

References
