8.5 Achievability of $R_I(D)$

How to Prove Achievability

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.
- For every random variable $\hat{X}$ taking values in $\hat{\mathcal{X}}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$, prove that the rate-distortion pair $(I(X; \hat{X}), D)$ is achievable by showing for large n the existence of a rate-distortion code such that

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.
- For every random variable $\hat{X}$ taking values in $\hat{\mathcal{X}}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$, prove that the rate-distortion pair $(I(X;\hat{X}), D)$ is achievable by showing for large n the existence of a rate-distortion code such that

    1. the rate of the code is not more than $I(X;\hat{X}) + \epsilon$;

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.

- For every random variable $\hat{X}$ taking values in $\hat{\mathcal{X}}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$, prove that the rate-distortion pair $(I(X; \hat{X}), D)$ is achievable by showing for large n the existence of a rate-distortion code such that

    1. the rate of the code is not more than $I(X; \hat{X}) + \epsilon$;
    2. $d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + \epsilon$ with probability almost 1.

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.

- For every random variable $\hat{X}$ taking values in $\hat{\mathcal{X}}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$, prove that the rate-distortion pair $(I(X; \hat{X}), D)$ is achievable by showing for large n the existence of a rate-distortion code such that

  1. the rate of the code is not more than $I(X; \hat{X}) + \epsilon$;
  2. $d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + \epsilon$ with probability almost 1.

- Minimize $I(X; \hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable.

How to Prove Achievability

Theorem 8.17 (The Rate-Distortion Theorem) $R(D) = R_I(D)$.

- An i.i.d. source $\{X_k : k \geq 1\}$ with generic random variable $X \sim p(x)$ is given.

- For every random variable $\hat{X}$ taking values in $\hat{\mathcal{X}}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$, prove that the rate-distortion pair $(I(X;\hat{X}), D)$ is achievable by showing for large n the existence of a rate-distortion code such that

  1. the rate of the code is not more than $I(X;\hat{X}) + \epsilon$;
  2. $d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + \epsilon$ with probability almost 1.

- Minimize $I(X;\hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable.

- This implies that $R_I(D) \geq R(D)$.

Random Coding Scheme

Random Coding Scheme

Parameter Settings

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

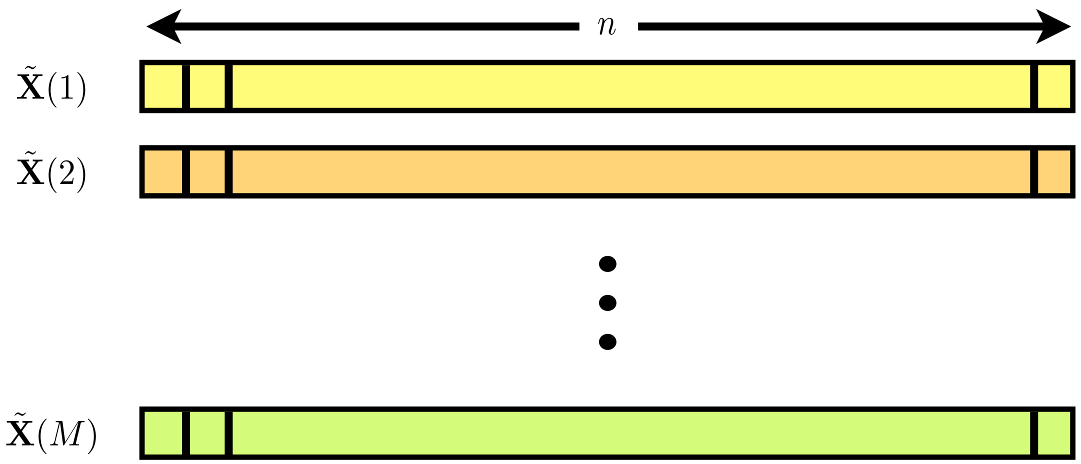
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

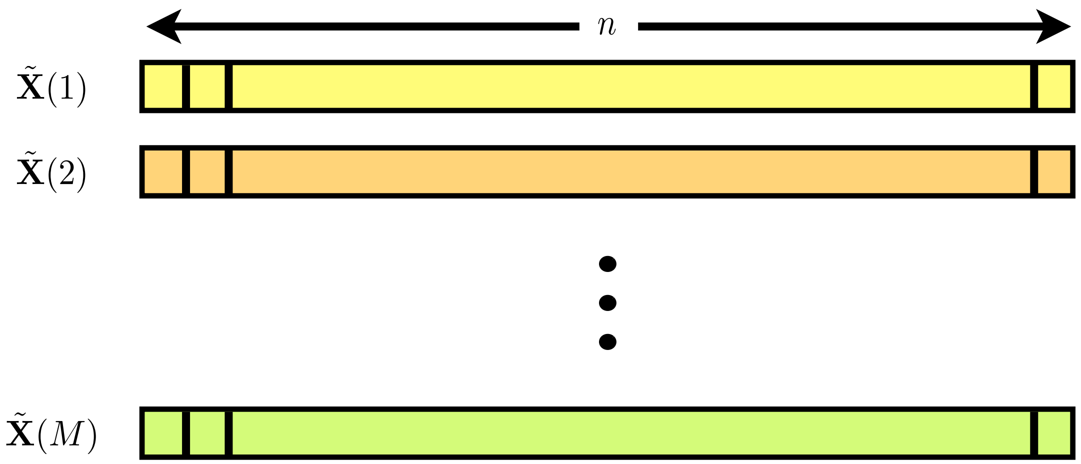
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

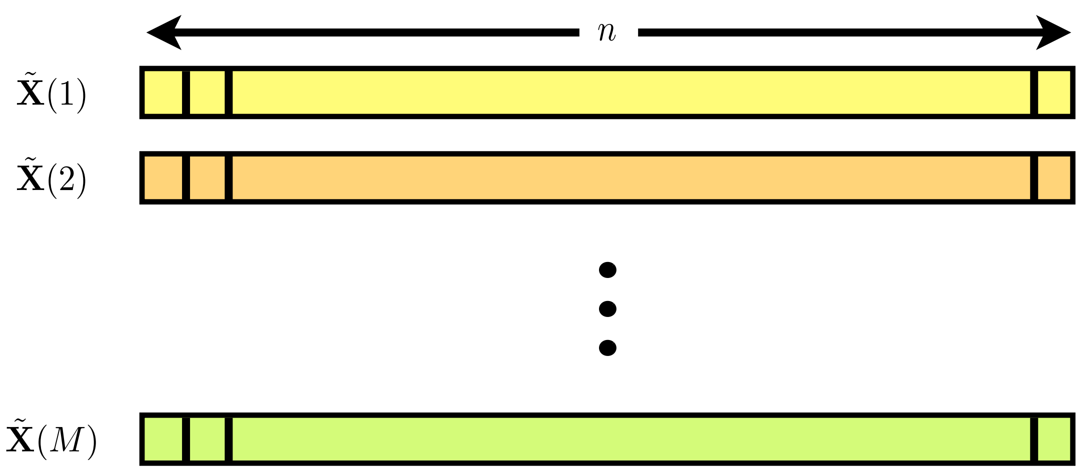
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

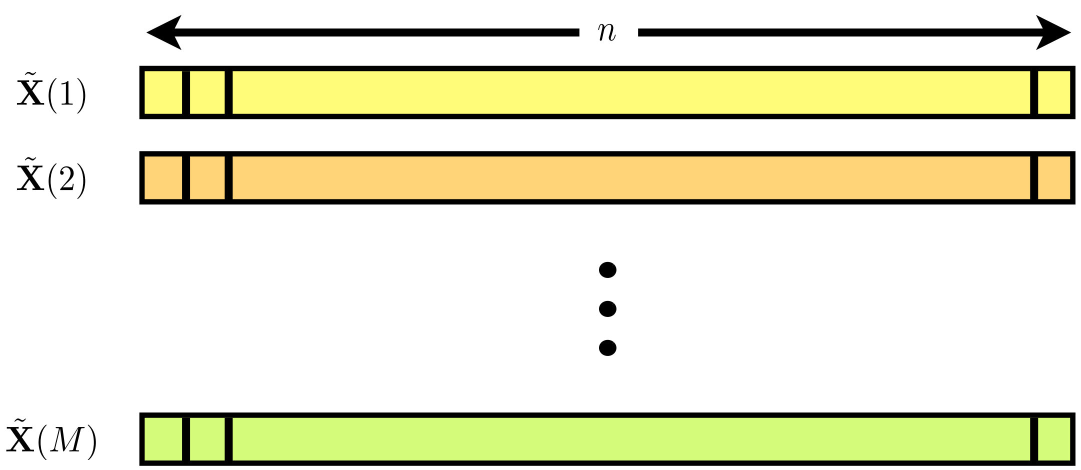
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

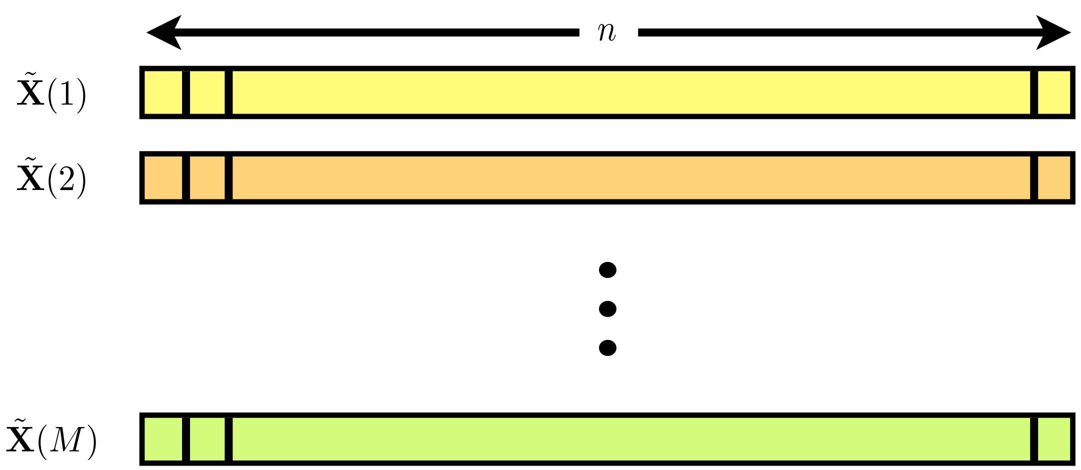
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

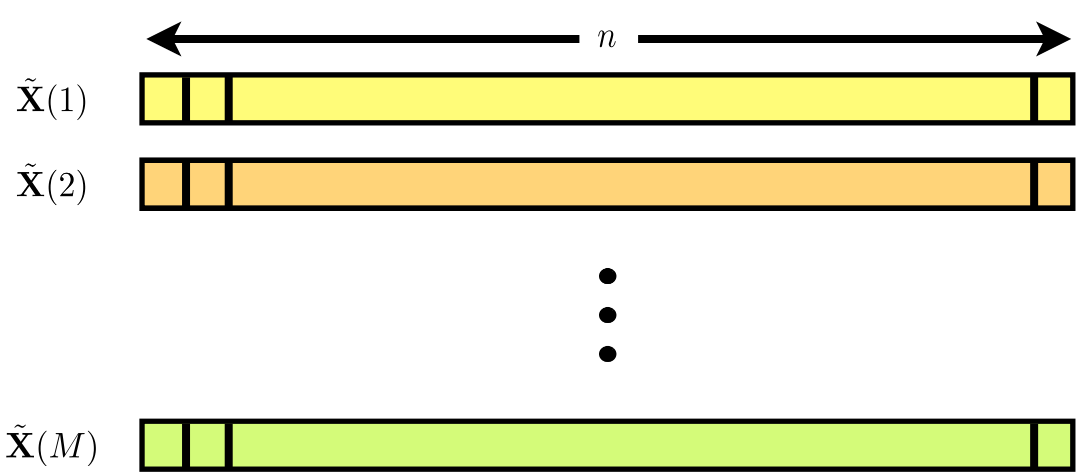
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

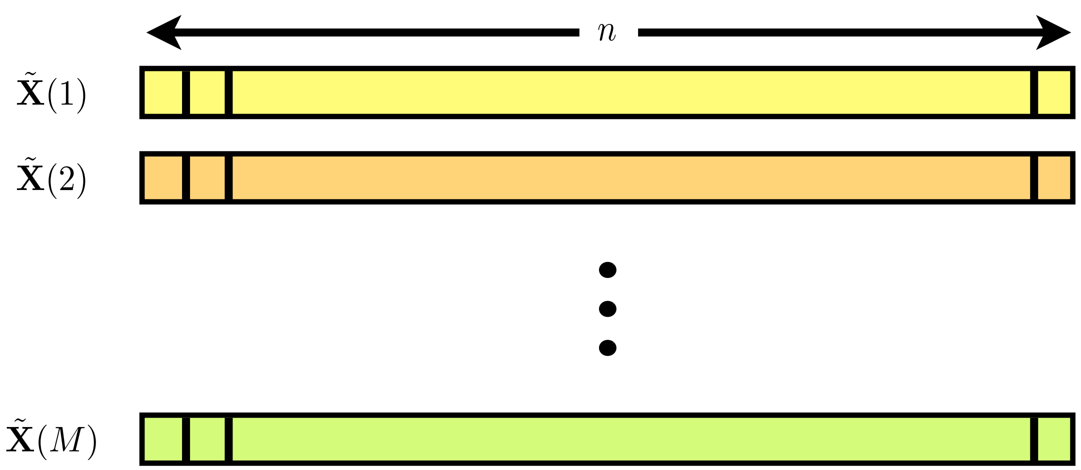
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \le D$, where $0 \le D \le D_{max}$. Let δ be specified later.

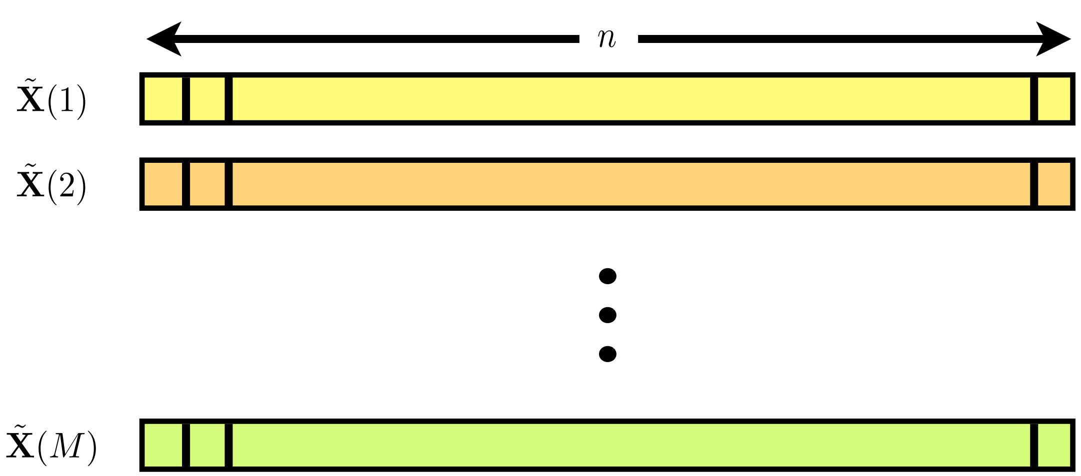
2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \le \frac{1}{n}\log M \le I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \le i$;

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

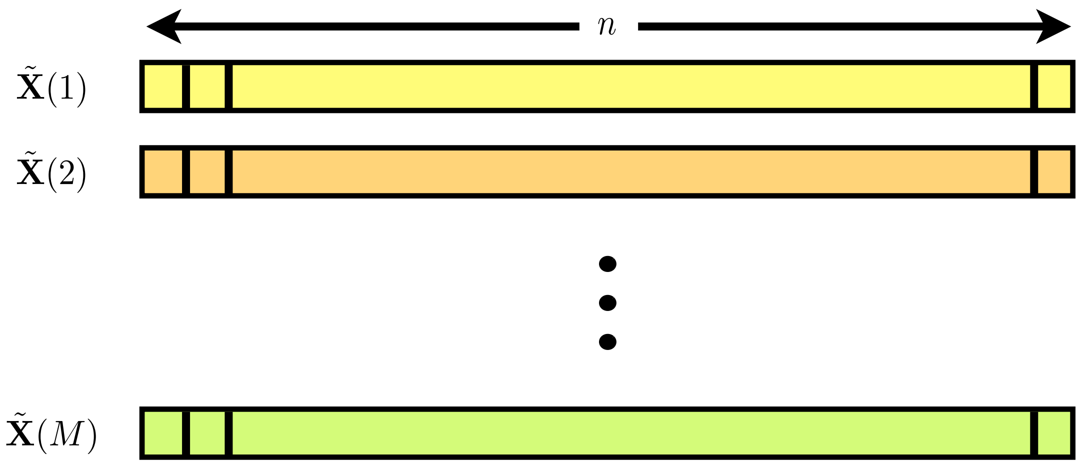
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

(a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

(b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

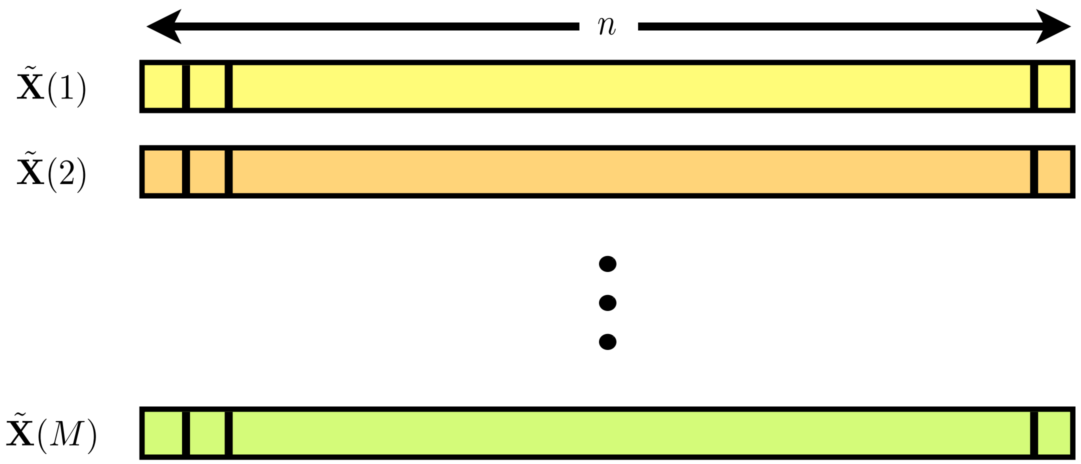
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

(a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

(b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

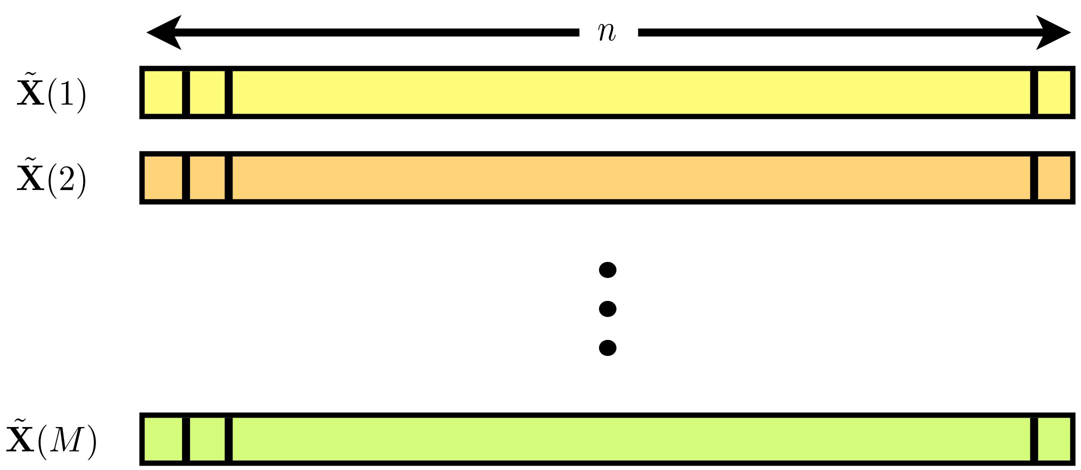
2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

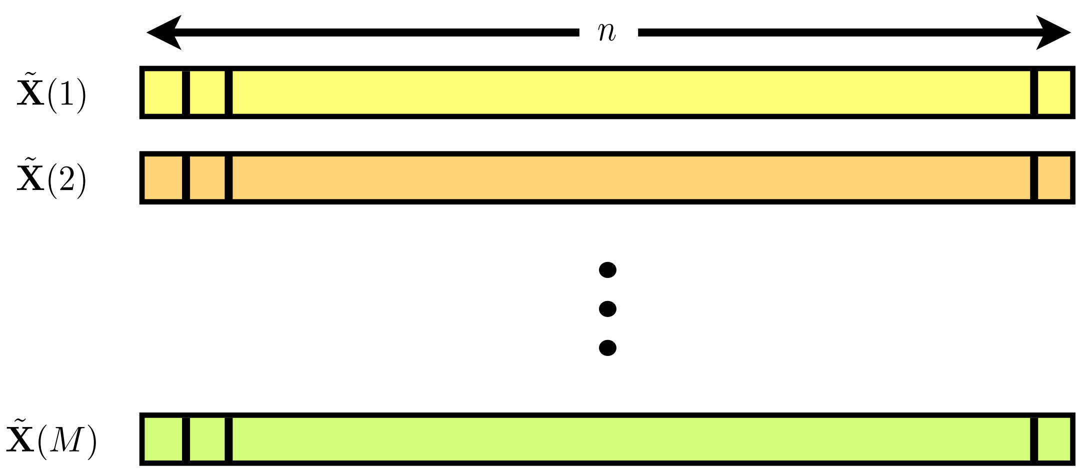
2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \le D$, where $0 \le D \le D_{max}$. Let δ be specified later.

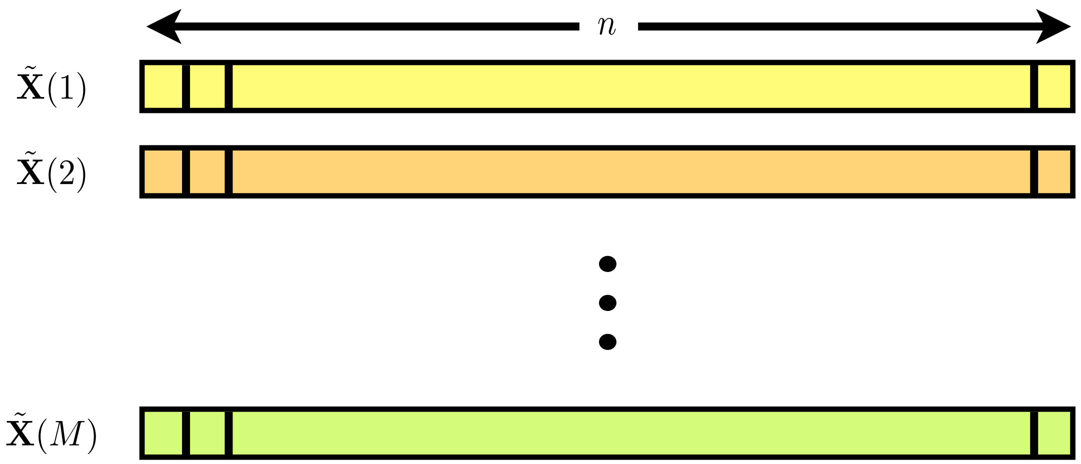
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \le \frac{1}{n} \log M \le I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \le i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

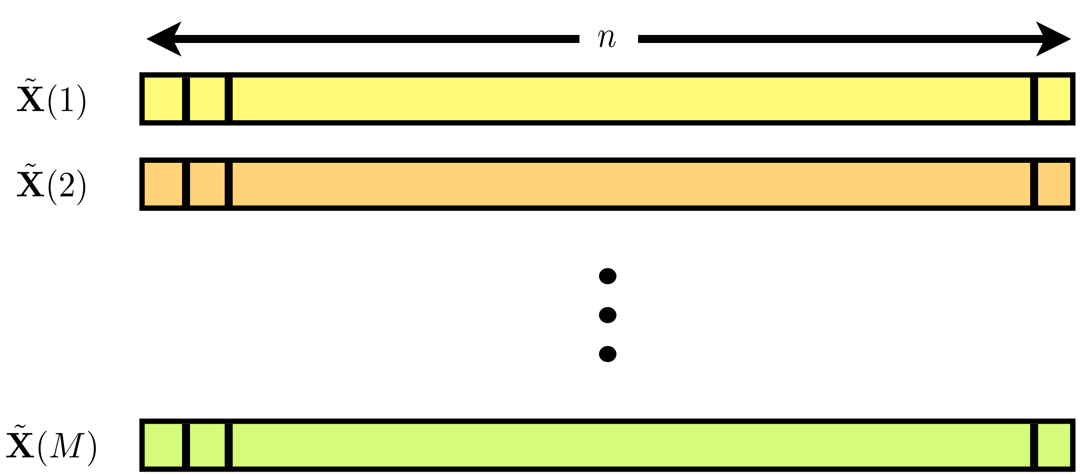
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Remarks

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \le D$, where $0 \le D \le D_{max}$. Let δ be specified later.

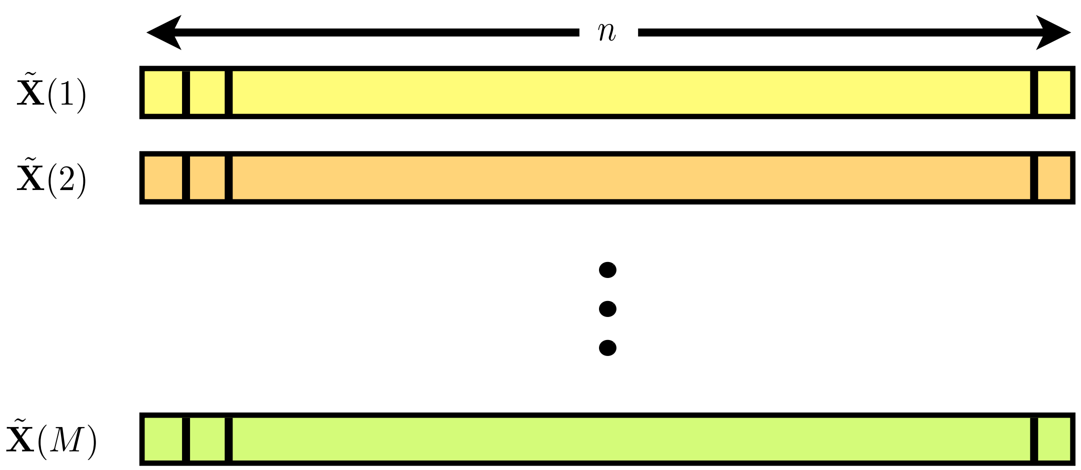
2. Let M be an integer satisfying

$$I(X;\hat{X}) + \frac{\epsilon}{2} \le \frac{1}{n} \log M \le I(X;\hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

   (a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

   (b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \le i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Remarks

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

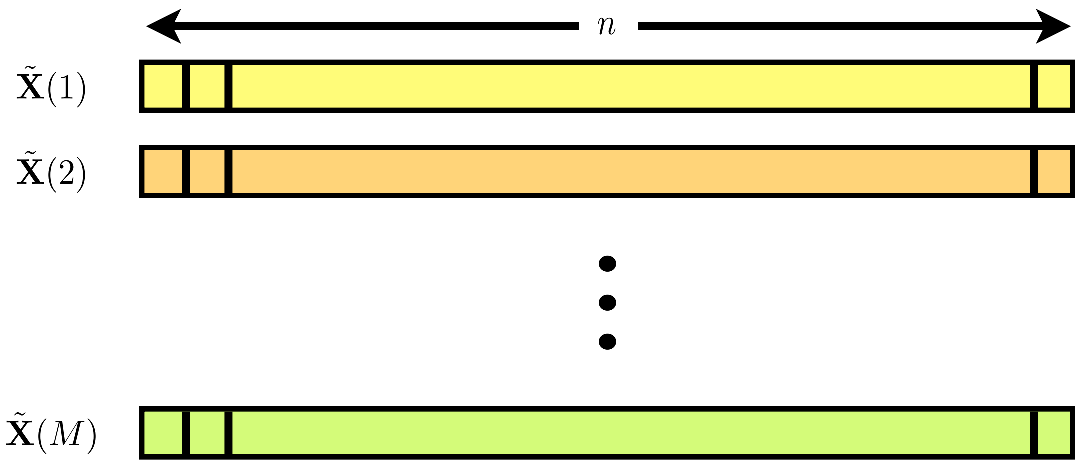
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

(a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

(b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Remarks

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \le D$, where $0 \le D \le D_{max}$. Let δ be specified later.

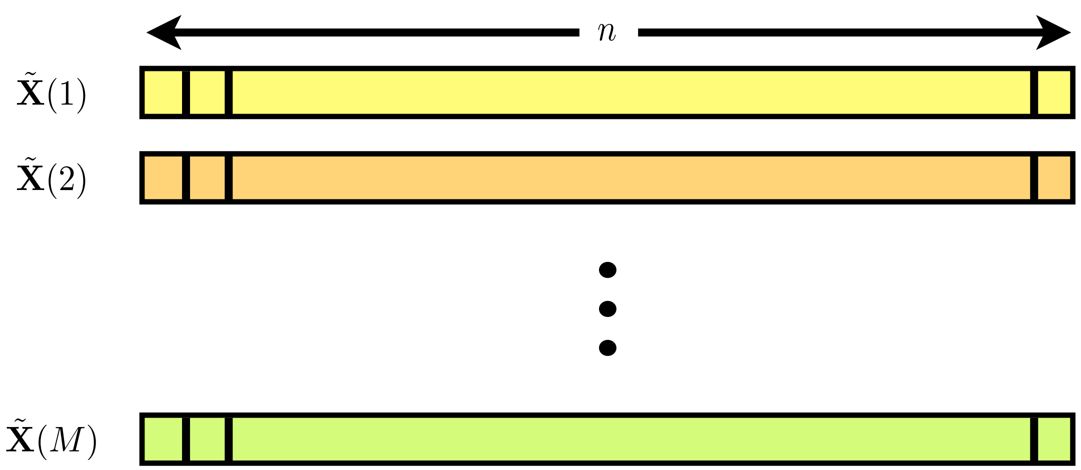
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \le \frac{1}{n} \log M \le I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

(a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

(b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \le i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Remarks

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and $\hat{X}$ with $Ed(X, \hat{X}) \leq D$, where $0 \leq D \leq D_{max}$. Let δ be specified later.

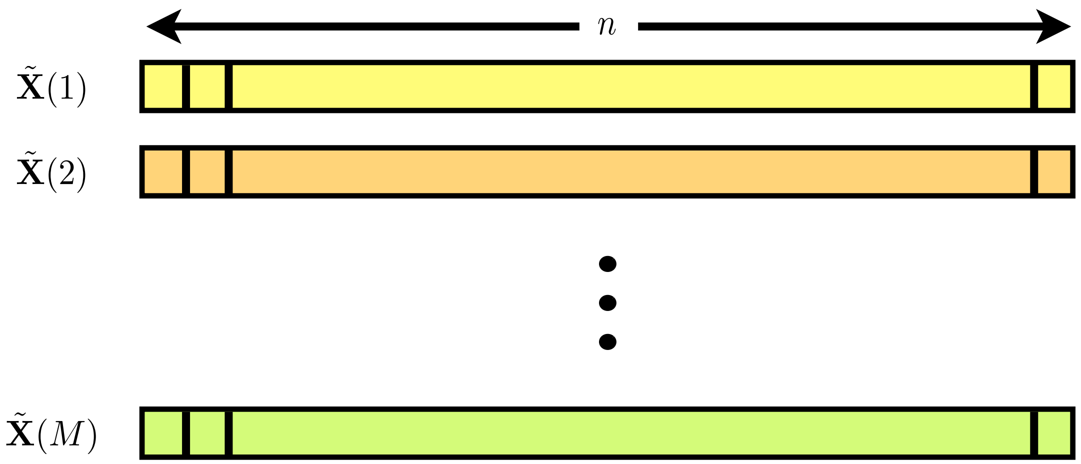
2. Let M be an integer satisfying

$$I(X; \hat{X}) + \frac{\epsilon}{2} \leq \frac{1}{n} \log M \leq I(X; \hat{X}) + \epsilon,$$

where n is sufficiently large.

The Random Coding Scheme

1. Construct a codebook $\mathcal{C}$ of an (n, M) code by randomly generating M codewords in $\hat{\mathcal{X}}^n$ independently and identically according to $p(\hat{x})^n$. Denote these codewords by $\hat{\mathbf{X}}(1), \hat{\mathbf{X}}(2), \cdots, \hat{\mathbf{X}}(M)$.



2. Reveal the codebook $\mathcal{C}$ to both the encoder and the decoder.

3. The source sequence $\mathbf{X}$ is generated according to $p(x)^n$.

4. The encoder encodes the source sequence $\mathbf{X}$ into an index K in the set $\mathcal{I} = \{1, 2, \cdots, M\}$. The index K takes the value i if

(a) $(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}$,

(b) for all $i' \in \mathcal{I}$, if $(\mathbf{X}, \hat{\mathbf{X}}(i')) \in T^n_{[X\hat{X}]\delta}$, then $i' \leq i$;

i.e., if there exists more than one i satisfying (a), let K be the largest one. Otherwise, K takes the constant value 1.

5. The index K is delivered to the decoder.

6. The decoder outputs $\hat{\mathbf{X}}(K)$ as the reproduction sequence $\hat{\mathbf{X}}$.

Remarks

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

2. If $K \neq 1$, then $\hat{\mathbf{X}}(K)$ is jointly typical with $\mathbf{X}$.

Performance Analysis

Performance Analysis

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\Pr\{K = 1 | \mathbf{X} = \mathbf{x}\}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \quad \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{x}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1 | \mathbf{X} = \mathbf{x}\} \\
&\quad \leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c | \mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c | \mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c | \mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1 | \mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

Proposition 6.13 With respect to a joint distribution $p(x, y)$ on $\mathcal{X} \times \mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} > 1 - \delta$$

for n sufficiently large.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\Pr\{E_1|\mathbf{X} = \mathbf{x}\} \quad = \quad \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned} &\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\ &\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\ &\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\ &\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\ &\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}. \end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned} \Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\ &= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}})
\end{aligned}$$

> **Theorem 6.7 (Consistency)** If $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, then $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)}
\end{aligned}$$

Theorem 6.10 (Conditional Strong AEP) If $|T^n_{[Y|X]\delta}(\mathbf{x})| \ge 1$, then

$$2^{n(H(Y|X)-\nu)} \le |T^n_{[Y|X]\delta}(\mathbf{x})| \le 2^{n(H(Y|X)+\nu)},$$

where $\nu \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)}
\end{aligned}$$

Theorem 6.10 (Conditional Strong AEP) If $|T^n_{[Y|X]\delta}(\mathbf{x})| \ge 1$, then

$$2^{n(H(Y|X)-\nu)} \le |T^n_{[Y|X]\delta}(\mathbf{x})| \le 2^{n(H(Y|X)+\nu)},$$

where $\nu \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)}
\end{aligned}$$

Theorem 6.10 (Conditional Strong AEP) If $|T^n_{[Y|X]\delta}(\mathbf{x})| \ge 1$, then

$$2^{n(H(Y|X)-\nu)} \le |T^n_{[Y|X]\delta}(\mathbf{x})| \le 2^{n(H(Y|X)+\nu)},$$

where $\nu \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned} &\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\ &\quad\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\ &\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\ &\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\ &\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}. \end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned} \Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\ &= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\ &\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\ &\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\ &= 2^{-n(\underline{H(\hat{X})-H(\hat{X}|X)}+\xi+\eta)} \end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1 | \mathbf{X} = \mathbf{x}\} \\
&\quad \le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c | \mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c | \mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c | \mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1 | \mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1 | \mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1 | \mathbf{X} = \mathbf{x}\} &= \Pr\left\{ (\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta} \right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad= (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad= (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\le (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta}\right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \le \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

Performance Analysis

1. The event $\{K = 1\}$ occurs in one of the following two scenarios:

- $\hat{X}(1)$ is the only codeword in $\mathcal{C}$ which is jointly typical with $\mathbf{X}$.
- No codeword in $\mathcal{C}$ is jointly typical with $\mathbf{X}$.

In other words, if $K = 1$, then $\mathbf{X}$ is jointly typical with none of the codewords $\hat{X}(2), \hat{X}(3), \cdots, \hat{X}(M)$. We will show that $\Pr\{K = 1\}$ can be made arbitrarily small.

2. Define the event

$$E_i = \left\{ (\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \right\}.$$

3. Then

$$\{K = 1\} \subset E_2^c \cap E_3^c \cap \cdots \cap E_M^c.$$

4. Since the codewords are generated i.i.d., conditioning on $\{\mathbf{X} = \mathbf{x}\}$ for any $\mathbf{x} \in \mathcal{X}^n$, the events E_i are mutually independent and have the same probability.

5. Then for any $\mathbf{x} \in \mathcal{X}^n$,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le \Pr\{E_2^c \cap E_3^c \cap \cdots \cap E_M^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = \prod_{i=2}^{M} \Pr\{E_i^c|\mathbf{X} = \mathbf{x}\} \\
&\quad = (\Pr\{E_1^c|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad = (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1}.
\end{aligned}$$

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \ge 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{ (\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta} \right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\ge \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\ge 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \le (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \le \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \left(1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\}\right)^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1) \underbrace{\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]}_{<\,0}
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \quad (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \quad \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \quad \Longleftrightarrow \quad M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \quad (M - 1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]
\end{aligned}$$

< 0

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M - 1) \ln \underbrace{\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]}_{<\,0} \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \left(1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\}\right)^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M - 1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \underline{\ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]}
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \underline{\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]}
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)}
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)}
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)}
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)}
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)}
\end{aligned}$$

$$\boxed{\ln a \leq a - 1}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \left(1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\}\right)^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - \underline{2^{-n(I(X;\hat{X})+\zeta)}}\right].
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

$$\ln a \leq a - 1$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \left(1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\}\right)^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \left(1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\}\right)^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[\underbrace{2^{n(\frac{\epsilon}{2}-\zeta)}}_{\infty} - \underbrace{2^{-n(I(X;\hat{X})+\zeta)}}_{0}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \quad \to -\infty \\
&\leq (M-1) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[\underbrace{2^{n(\frac{\epsilon}{2}-\zeta)}}_{\infty} - \underbrace{2^{-n(I(X;\hat{X})+\zeta)}}_{0}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to -\infty \\
&\quad \leq (M-1) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) \ln \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[\underbrace{2^{n(\frac{\epsilon}{2}-\zeta)}}_{\infty} - \underbrace{2^{-n(I(X;\hat{X})+\zeta)}}_{0}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad\leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to -\infty \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[\underbrace{2^{n(\frac{\epsilon}{2}-\zeta)}}_{\infty} - \underbrace{2^{-n(I(X;\hat{X})+\zeta)}}_{0}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

Performance Analysis

6. We will focus on $\mathbf{x} \in S^n_{[X]\delta}$ where

$$S^n_{[X]\delta} = \{\mathbf{x} \in T^n_{[X]\delta} : |T^n_{[\hat{X}|X]\delta}(\mathbf{x})| \geq 1\},$$

because $\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} \approx 1$ for large n.

7. For $\mathbf{x} \in S^n_{[X]\delta}$, obtain the following lower bound on $\Pr\{E_1|\mathbf{X} = \mathbf{x}\}$:

$$\begin{aligned}
\Pr\{E_1|\mathbf{X} = \mathbf{x}\} &= \Pr\left\{(\mathbf{x}, \hat{\mathbf{X}}(1)) \in T^n_{[X\hat{X}]\delta}\right\} \\
&= \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} p(\hat{\mathbf{x}}) \\
&\geq \sum_{\hat{\mathbf{x}} \in T^n_{[\hat{X}|X]\delta}(\mathbf{x})} 2^{-n(H(\hat{X})+\eta)} \\
&\geq 2^{n(H(\hat{X}|X)-\xi)} 2^{-n(H(\hat{X})+\eta)} \\
&= 2^{-n(H(\hat{X})-H(\hat{X}|X)+\xi+\eta)} \\
&= 2^{-n(I(X;\hat{X})+\zeta)},
\end{aligned}$$

where $\zeta = \xi + \eta \to 0$ as $n \to \infty$ and $\delta \to 0$.

8. Therefore,

$$\begin{aligned}
&\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (1 - \Pr\{E_1|\mathbf{X} = \mathbf{x}\})^{M-1} \\
&\quad \leq \left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right]^{M-1}.
\end{aligned}$$

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right) 2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

9. Now

$$\frac{1}{n} \log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \ge I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \ge 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\le (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\le \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\le -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \le \frac{\epsilon}{2}.$$

12. It follows that

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$\Pr\{K = 1\}$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad+ \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
\Pr\{K = 1\} &= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \underline{\Pr\{K = 1|\mathbf{X} = \mathbf{x}\}}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \underline{\Pr\{K=1|\mathbf{X}=\mathbf{x}\}}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \underline{\Pr\{K=1|\mathbf{X}=\mathbf{x}\}}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \underline{1}\cdot\Pr\{\mathbf{X}=\mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\qquad+ \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot \Pr\{\mathbf{X} = \mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot \Pr\{\mathbf{X} = \mathbf{x}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right]\\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right]\\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)}\\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\}\\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\}\\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\}\\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\}\\
&= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot \Pr\{\mathbf{X} = \mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1 \cdot \Pr\{\mathbf{X} = \mathbf{x}\} \\
&= \frac{\epsilon}{2}\cdot \Pr\{\mathbf{X} \in S^n_{[X]\delta}\} + \Pr\{\mathbf{X} \notin S^n_{[X]\delta}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&\quad = \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} = \mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad = \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} + \Pr\{\mathbf{X} \notin S^n_{[X]\delta}\} \\
&\quad \leq \frac{\epsilon}{2}\cdot 1 + \underline{\Pr\{\mathbf{X} \notin S^n_{[X]\delta}\}}
\end{aligned}$$

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} = \mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} + \Pr\{\mathbf{X} \notin S^n_{[X]\delta}\} \\
&\quad\leq \frac{\epsilon}{2}\cdot 1 + \underline{\Pr\{\mathbf{X} \notin S^n_{[X]\delta}\}}
\end{aligned}$$

Proposition 6.13 With respect to a joint distribution $p(x, y)$ on $\mathcal{X} \times \mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} > 1 - \delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \qquad (1)$$

Then the upper bound on $\ln\Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad\leq \frac{\epsilon}{2}\cdot 1 + \underline{\Pr\{\mathbf{X}\notin S^n_{[X]\delta}\}} \\
&\quad< \frac{\epsilon}{2} + \underline{\delta}.
\end{aligned}$$

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta>0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \qquad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta>0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \ge I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \ge 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\le (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\le \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\le -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \le \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\le \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad\le \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta>0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

13. By letting n be sufficiently large and δ be sufficiently small so that both (1) and $\delta < \frac{\epsilon}{2}$ are satisfied, we obtain

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \quad (1)$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

13. By letting n be sufficiently large and δ be sufficiently small so that both (1) and $\delta < \frac{\epsilon}{2}$ are satisfied, we obtain

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta>0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln\Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n\to\infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\}\to 0$ as $n\to\infty$.

11. This implies for $\mathbf{x}\in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

13. By letting n be sufficiently large and δ be sufficiently small so that both (1) and $\delta < \frac{\epsilon}{2}$ are satisfied, we obtain

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\} \\
&\quad\leq (M-1)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)\ln\left[1-2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad\leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}-1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad= -\left[2^{n(\frac{\epsilon}{2}-\zeta)}-2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2}-\zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K=1|\mathbf{X}=\mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K=1|\mathbf{X}=\mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K=1\} \\
&\quad= \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K=1|\mathbf{X}=\mathbf{x}\}\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad\leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}=\mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X}=\mathbf{x}\} \\
&\quad= \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad\leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X}\notin S^n_{[X]\delta}\} \\
&\quad< \frac{\epsilon}{2}+\delta.
\end{aligned}$$

$< \frac{\epsilon}{2}$

13. By letting n be sufficiently large and δ be sufficiently small so that both (1) and $\delta < \frac{\epsilon}{2}$ are satisfied, we obtain

Proposition 6.13 With respect to a joint distribution $p(x,y)$ on $\mathcal{X}\times\mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X}\in S^n_{[X]\delta}\} > 1-\delta$$

for n sufficiently large.

Performance Analysis

9. Now

$$\frac{1}{n}\log M \geq I(X;\hat{X}) + \frac{\epsilon}{2} \iff M \geq 2^{n(I(X;\hat{X})+\frac{\epsilon}{2})}.$$

Then

$$\begin{aligned}
&\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \\
&\quad \leq (M-1)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq \left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)\ln\left[1 - 2^{-n(I(X;\hat{X})+\zeta)}\right] \\
&\quad \leq -\left(2^{n(I(X;\hat{X})+\frac{\epsilon}{2})} - 1\right)2^{-n(I(X;\hat{X})+\zeta)} \\
&\quad = -\left[2^{n(\frac{\epsilon}{2}-\zeta)} - 2^{-n(I(X;\hat{X})+\zeta)}\right].
\end{aligned}$$

10. Let n be sufficiently large and δ be sufficiently small so that

$$\frac{\epsilon}{2} - \zeta > 0. \tag{1}$$

Then the upper bound on $\ln \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}$ tends to $-\infty$ as $n \to \infty$, i.e., $\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \to 0$ as $n \to \infty$.

11. This implies for $\mathbf{x} \in S^n_{[X]\delta}$, for sufficiently large n,

$$\Pr\{K = 1|\mathbf{X} = \mathbf{x}\} \leq \frac{\epsilon}{2}.$$

12. It follows that

$$\begin{aligned}
&\Pr\{K = 1\} \\
&\quad = \sum_{\mathbf{x}\in S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\qquad + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} \Pr\{K = 1|\mathbf{X} = \mathbf{x}\}\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad \leq \sum_{\mathbf{x}\in S^n_{[X]\delta}} \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} = \mathbf{x}\} + \sum_{\mathbf{x}\notin S^n_{[X]\delta}} 1\cdot\Pr\{\mathbf{X} = \mathbf{x}\} \\
&\quad = \frac{\epsilon}{2}\cdot\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} + \Pr\{\mathbf{X} \notin S^n_{[X]\delta}\} \\
&\quad \leq \frac{\epsilon}{2}\cdot 1 + \Pr\{\mathbf{X} \notin S^n_{[X]\delta}\} \\
&\quad < \frac{\epsilon}{2} + \delta.
\end{aligned}$$

$< \frac{\epsilon}{2}$

13. By letting n be sufficiently large and δ be sufficiently small so that both (1) and $\delta < \frac{\epsilon}{2}$ are satisfied, we obtain

$$\Pr\{K = 1\} < \epsilon.$$

Proposition 6.13 With respect to a joint distribution $p(x, y)$ on $\mathcal{X} \times \mathcal{Y}$, for any $\delta > 0$,

$$\Pr\{\mathbf{X} \in S^n_{[X]\delta}\} > 1 - \delta$$

for n sufficiently large.

Main Idea

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.
- $\mathbf{X} \in S^n_{[X]\delta}$ with high probability.

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.
- $\mathbf{X} \in S^n_{[X]\delta}$ with high probability.
- For $\mathbf{x} \in S^n_{[X]\delta}$, by conditional strong AEP,

$$\Pr\left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \,|\, \mathbf{X} = \mathbf{x}\right\} \approx 2^{-nI(X;\hat{X})}.$$

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.
- $\mathbf{X} \in S^n_{[X]\delta}$ with high probability.
- For $\mathbf{x} \in S^n_{[X]\delta}$, by conditional strong AEP,

$$\Pr\left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \,|\, \mathbf{X} = \mathbf{x}\right\} \approx 2^{-nI(X;\hat{X})}.$$

- If M grows with n at a rate higher than $I(X;\hat{X})$, then the probability that there exists at least one $\hat{\mathbf{X}}(i)$ which is jointly typical with the source sequence $\mathbf{X}$ with respect to $p(x, \hat{x})$ is high.

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.
- $\mathbf{X} \in S^n_{[X]\delta}$ with high probability.
- For $\mathbf{x} \in S^n_{[X]\delta}$, by conditional strong AEP,

$$\Pr\left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \,|\, \mathbf{X} = \mathbf{x}\right\} \approx 2^{-nI(X;\hat{X})}.$$

- If M grows with n at a rate higher than $I(X;\hat{X})$, then the probability that there exists at least one $\hat{\mathbf{X}}(i)$ which is jointly typical with the source sequence $\mathbf{X}$ with respect to $p(x, \hat{x})$ is high.
- Such an $\hat{\mathbf{X}}(i)$, if exists, would have $d(\mathbf{X}, \hat{\mathbf{X}}) \approx Ed(X, \hat{X}) \leq D$, because the joint relative frequency of $(\mathbf{x}, \hat{\mathbf{X}}(i)) \approx p(x, \hat{x})$. See the next proposition.

Main Idea

- Randomly generate M codewords in $\hat{\mathcal{X}}^n$ according to $p(\hat{x})^n$, where n is large.
- $\mathbf{X} \in S^n_{[X]\delta}$ with high probability.
- For $\mathbf{x} \in S^n_{[X]\delta}$, by conditional strong AEP,

$$\Pr\left\{(\mathbf{X}, \hat{\mathbf{X}}(i)) \in T^n_{[X\hat{X}]\delta} \,|\, \mathbf{X} = \mathbf{x}\right\} \approx 2^{-nI(X;\hat{X})}.$$

- If M grows with n at a rate higher than $I(X;\hat{X})$, then the probability that there exists at least one $\hat{\mathbf{X}}(i)$ which is jointly typical with the source sequence $\mathbf{X}$ with respect to $p(x, \hat{x})$ is high.
- Such an $\hat{\mathbf{X}}(i)$, if exists, would have $d(\mathbf{X}, \hat{\mathbf{X}}) \approx Ed(X, \hat{X}) \leq D$, because the joint relative frequency of $(\mathbf{x}, \hat{\mathbf{X}}(i)) \approx p(x, \hat{x})$. See the next proposition.
- Use this $\hat{\mathbf{X}}(i)$ to represent $\mathbf{X}$ to satisfy the distortion constraint.

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k)$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{k=1}^{n} d(x_k, \hat{x}_k)$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n} \sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n} \sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n} \sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n} \sum_{\underline{x, \hat{x}}} d(x, \hat{x}) \underline{N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}})} \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \le D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \le D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n} \sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n} \sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\ &= \frac{1}{n} \sum_{x,\hat{x}} d(x, \hat{x}) (\underline{np(x, \hat{x})} + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) \left(\underline{np(x, \hat{x})} + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - \underline{np(x, \hat{x})}\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\ &= \frac{1}{n}\sum_{x,\hat{x}} \underline{d(x, \hat{x})}(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\ &= \frac{1}{n}\sum_{x,\hat{x}} \underline{d(x, \hat{x})}\,(\underline{np(x, \hat{x})} + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned} d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\ &= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\ &= \frac{1}{n}\sum_{x,\hat{x}} \underline{d(x,\hat{x})}(\underline{np(x,\hat{x})} + N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - np(x,\hat{x})) \\ &= \underline{\sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x})} + \sum_{x,\hat{x}} d(x,\hat{x})\left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right) \end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} \underline{d(x,\hat{x})}(\underline{\not{n}p(x,\hat{x})} + N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - np(x,\hat{x})) \\
&= \underline{\sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x})} + \sum_{x,\hat{x}} d(x,\hat{x})\left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \le D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \le D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + \underline{N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}})} - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})(np(x,\hat{x}) + \underline{N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}})} - np(x,\hat{x})) \\
&= \sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x}) + \sum_{x,\hat{x}} d(x,\hat{x})\left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})(np(x,\hat{x}) + N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - np(x,\hat{x})) \\
&= \sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x}) + \sum_{x,\hat{x}} d(x,\hat{x})\left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - \underline{np(x, \hat{x})}) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - \underline{np(x, \hat{x})}) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - \underline{p(x, \hat{x})}\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})(np(x,\hat{x}) + N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - np(x,\hat{x})) \\
&= \sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x}) + \sum_{x,\hat{x}} d(x,\hat{x})\left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \underline{\sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x})} + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \underline{\sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x})} + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= \underline{Ed(X, \hat{X})} + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x}) \right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x}) \right)
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x})d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n}N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n}N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left|\frac{1}{n}N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right|
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x},\hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k,\hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x}) N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x,\hat{x})(np(x,\hat{x}) + N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - np(x,\hat{x})) \\
&= \sum_{x,\hat{x}} p(x,\hat{x})d(x,\hat{x}) + \sum_{x,\hat{x}} d(x,\hat{x}) \left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right) \\
&= Ed(X,\hat{X}) + \sum_{x,\hat{x}} d(x,\hat{x}) \left(\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right) \\
&\leq Ed(X,\hat{X}) + \sum_{x,\hat{x}} \underline{d(x,\hat{x})} \left|\frac{1}{n}N(x,\hat{x}|\mathbf{x},\hat{\mathbf{x}}) - p(x,\hat{x})\right|
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} \underline{d(x, \hat{x})} \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + \underline{d_{max}} \sum_{x,\hat{x}} \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right|
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max} \sum_{x,\hat{x}} \underline{\left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right|}
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max} \sum_{x,\hat{x}} \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max}\delta
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max} \sum_{x,\hat{x}} \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq \underline{Ed(X, \hat{X})} + d_{max}\delta
\end{aligned}$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

Proof For $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, consider

$$\begin{aligned}
d(\mathbf{x}, \hat{\mathbf{x}}) &= \frac{1}{n}\sum_{k=1}^{n} d(x_k, \hat{x}_k) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x}) N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) \\
&= \frac{1}{n}\sum_{x,\hat{x}} d(x, \hat{x})(np(x, \hat{x}) + N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - np(x, \hat{x})) \\
&= \sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&= Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left(\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right) \\
&\leq Ed(X, \hat{X}) + \sum_{x,\hat{x}} d(x, \hat{x}) \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max} \sum_{x,\hat{x}} \left|\frac{1}{n} N(x, \hat{x}|\mathbf{x}, \hat{\mathbf{x}}) - p(x, \hat{x})\right| \\
&\leq Ed(X, \hat{X}) + d_{max}\delta \\
&\leq D + d_{max}\delta.
\end{aligned}$$

The Remaining Details

The Remaining Details

The Remaining Details

1. For sufficiently large n, consider

The Remaining Details

1. For sufficiently large n, consider

$\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\}$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\}
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\} \Pr\{K = 1\}
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | \underline{K = 1}\} \Pr\{\underline{K = 1}\}
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\}
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | \underline{K \neq 1}\}\Pr\{K \neq 1\}
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\}
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\}
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \underline{\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}}\Pr\{K = 1\} \\
&\qquad\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad \underline{1} \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\underline{\Pr\{K = 1\}} \\
&\quad\quad\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\underline{\Pr\{K = 1\}} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq 1 \cdot \underline{\epsilon} + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\} \Pr\{K = 1\} \\
&\quad \qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \underline{\Pr\{K \neq 1\}} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1
\end{aligned}
$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\underline{\Pr\{K \neq 1\}} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot \underline{1}
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad\quad\quad\quad +\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \quad \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad\quad\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \quad \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad \qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \quad \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \quad \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}
$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&\quad = \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\qquad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\quad \leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&\quad = \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned} & \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\ = \quad & \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\ & + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\ \leq \quad & 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\ = \quad & \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}. \end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&\quad = \quad \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\qquad\quad +\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\quad \leq \quad 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&\quad = \quad \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \leq D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \leq \epsilon$.

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&\quad = \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\qquad + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\quad \leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&\quad = \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}. \quad 0
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K \neq 1\}\cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K \neq 1\}.
\end{aligned}$$

0

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K \neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} &= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K=1\}\Pr\{K=1\} \\ &\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}\Pr\{K\neq 1\} \\ &\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}\cdot 1 \\ &= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}. \quad 0\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} &= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\ &\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\ &\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\ &= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.\end{aligned}$$

0

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} &= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K=1\}\Pr\{K=1\} \\
&\quad + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}\Pr\{K\neq 1\} \\
&\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}\cdot 1 \\
&= \epsilon + \underbrace{\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\}}_{0}.
\end{aligned}
$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon \mid K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, $(I(X;\hat{X}), D)$ is achievable.

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$
\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}
$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

> **Proposition** For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then
>
> $$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, $(I(X;\hat{X}), D)$ is achievable.

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&\quad = \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\qquad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\quad \leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&\quad = \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

> **Proposition** For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then
>
> $$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq \underline{I(X;\hat{X})} + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > \underline{D} + \epsilon\} \leq \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \leq \underline{D}$, $(\underline{I(X;\hat{X})}, \underline{D})$ is achievable.

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \\
&= \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K = 1\}\Pr\{K = 1\} \\
&\quad + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}\Pr\{K \neq 1\} \\
&\le 1 \cdot \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} \cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X}, \hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X}, \hat{\mathbf{X}}) \le D + d_{max}\delta.$$

3. By taking $\delta \le \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X}, \hat{\mathbf{X}}) \le D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D + \epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X, \hat{X}) \le D$, if $(\mathbf{x}, \hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x}, \hat{\mathbf{x}}) \le D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon | K \neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \le \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \le I(X; \hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X}, \hat{\mathbf{X}}) > D + \epsilon\} \le \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X, \hat{X}) \le D$, $(I(X; \hat{X}), D)$ is achievable.

6. Finally, minimize $I(X; \hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable, i.e.,

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\le 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K\neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \le D + d_{max}\delta.$$

3. By taking $\delta \le \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \le D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \le \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \le I(X;\hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \le \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \le D$, $(I(X;\hat{X}), D)$ is achievable.

6. Finally, minimize $I(X;\hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable, i.e.,

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \le D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \le D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned} &\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\ &= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\ &\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\ &\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\ &= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}. \end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, $(I(X;\hat{X}), D)$ is achievable.

6. Finally, minimize $I(X;\hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable, i.e.,

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

The Remaining Details

1. For sufficiently large n, consider

$$\begin{aligned}
&\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \\
&= \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K=1\}\Pr\{K=1\} \\
&\quad +\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\Pr\{K\neq 1\} \\
&\leq 1\cdot\epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}\cdot 1 \\
&= \epsilon + \Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\}.
\end{aligned}$$

2. Recall that conditioning on $\{K \neq 1\}$, $(\mathbf{X},\hat{\mathbf{X}}) \in T^n_{[X\hat{X}]\delta}$. Then

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\delta.$$

3. By taking $\delta \leq \frac{\epsilon}{d_{max}}$, we obtain

$$d(\mathbf{X},\hat{\mathbf{X}}) \leq D + d_{max}\left(\frac{\epsilon}{d_{max}}\right) = D+\epsilon.$$

Proposition For $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, if $(\mathbf{x},\hat{\mathbf{x}}) \in T^n_{[X\hat{X}]\delta}$, then

$$d(\mathbf{x},\hat{\mathbf{x}}) \leq D + d_{max}\delta.$$

4. Therefore, $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon | K\neq 1\} = 0$, which implies $\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon$.

5. Thus we have shown the existence of a code such that

$$\frac{1}{n}\log M \leq I(X;\hat{X}) + \epsilon$$

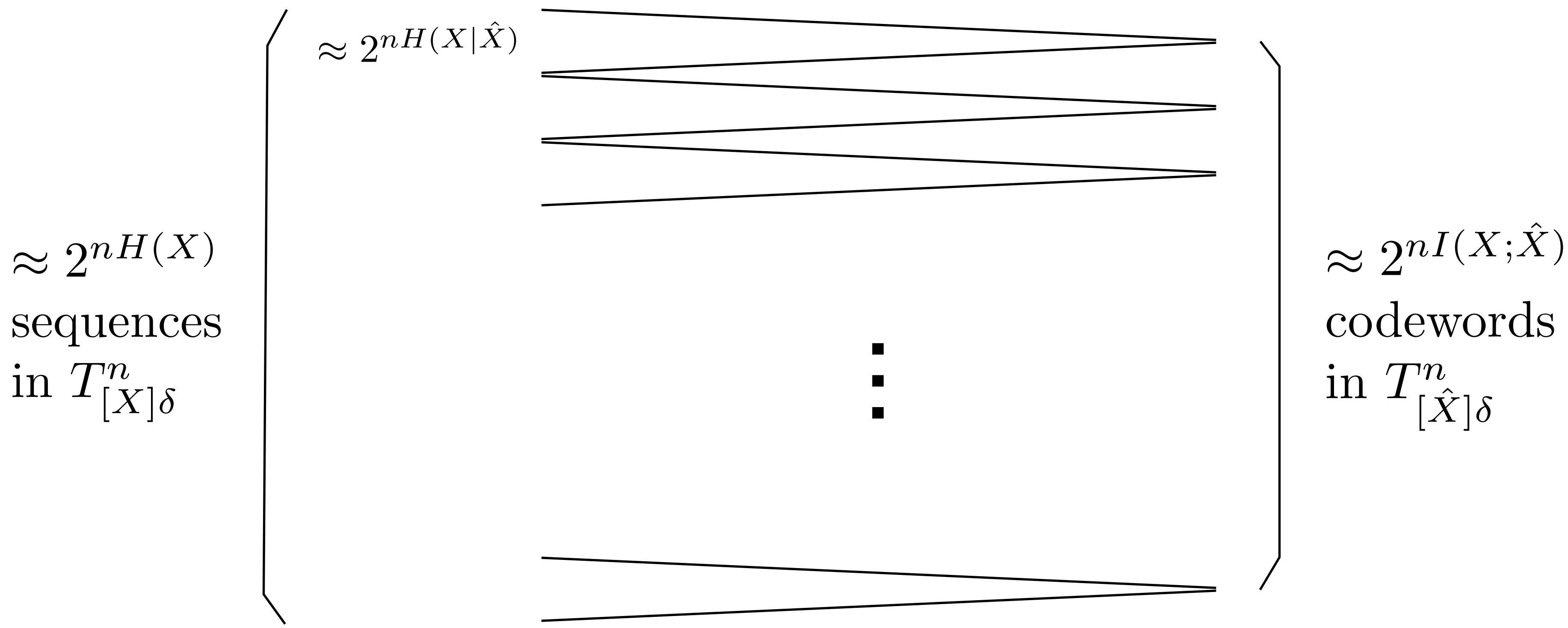
and

$$\Pr\{d(\mathbf{X},\hat{\mathbf{X}}) > D+\epsilon\} \leq \epsilon.$$

Hence, for any $\hat{X}$ such that $Ed(X,\hat{X}) \leq D$, $(I(X;\hat{X}), D)$ is achievable.

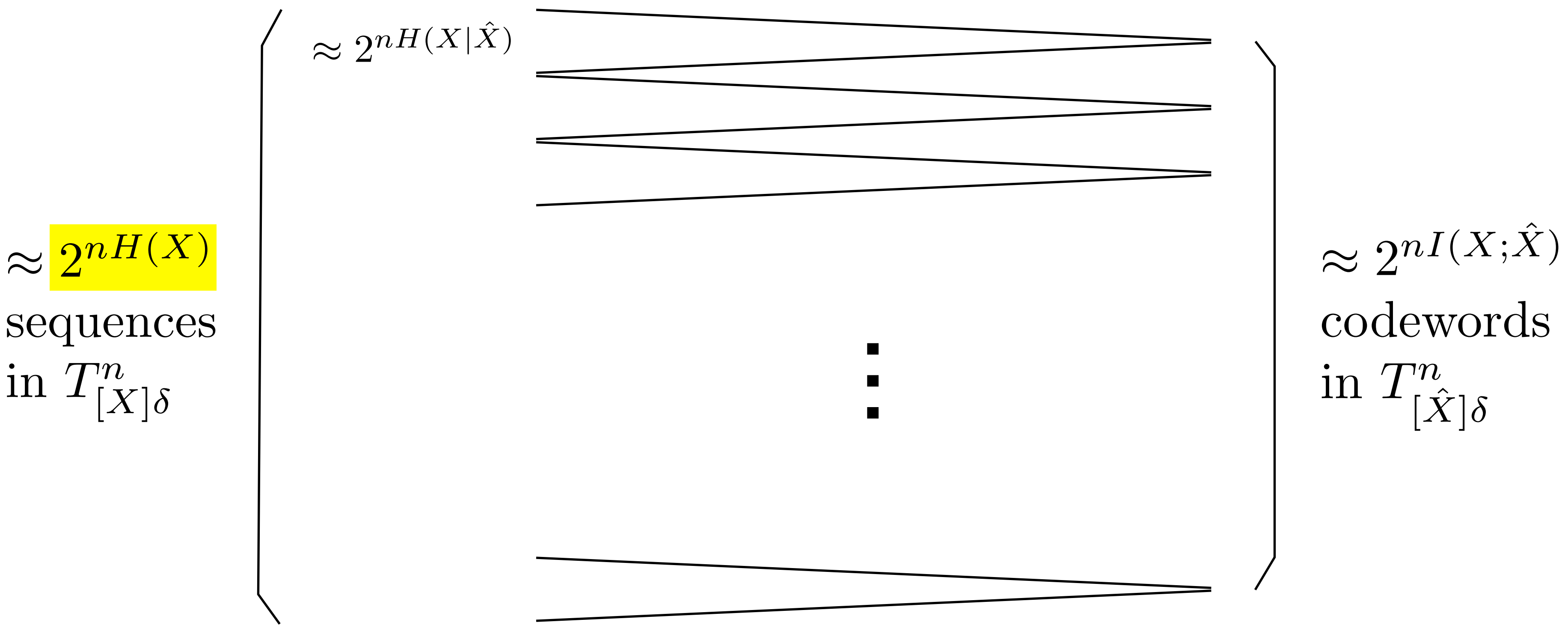
6. Finally, minimize $I(X;\hat{X})$ over all such $\hat{X}$ to conclude that $(R_I(D), D)$ is achievable, i.e.,

$$R_I(D) \geq R(D).$$



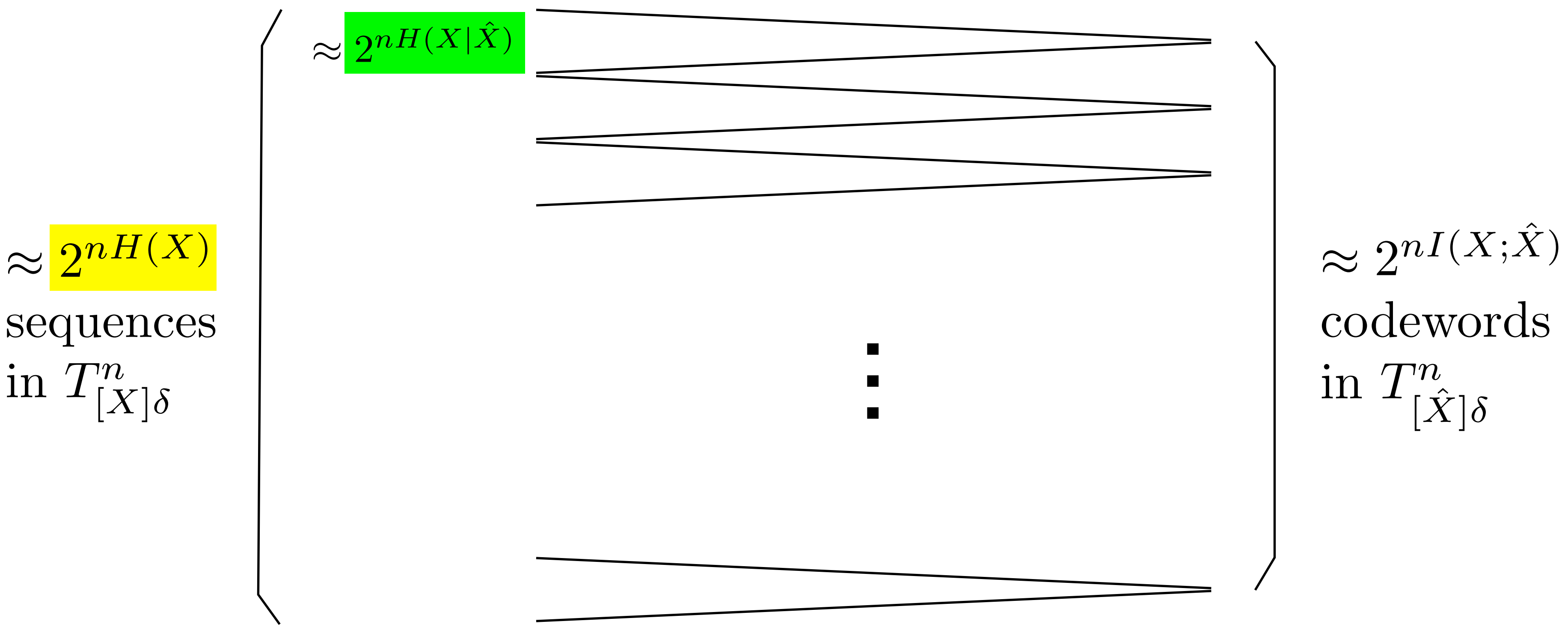
The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$



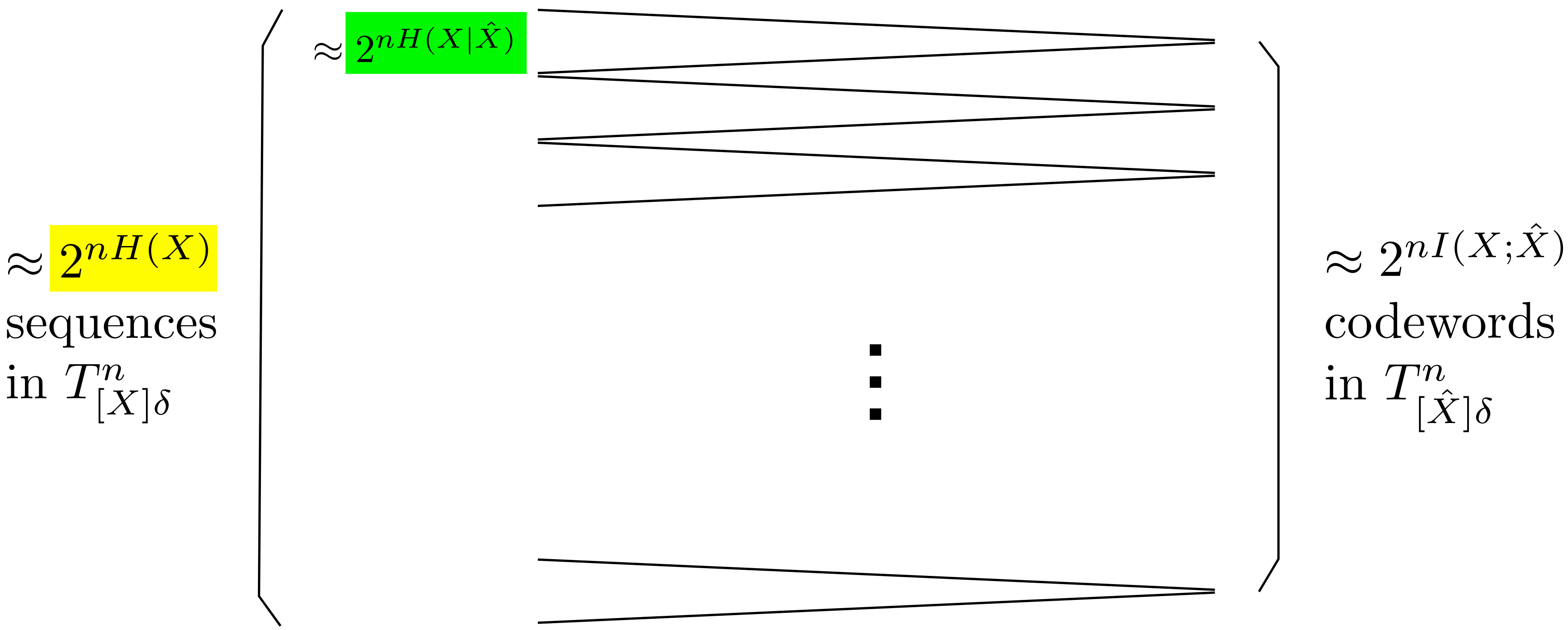
The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$



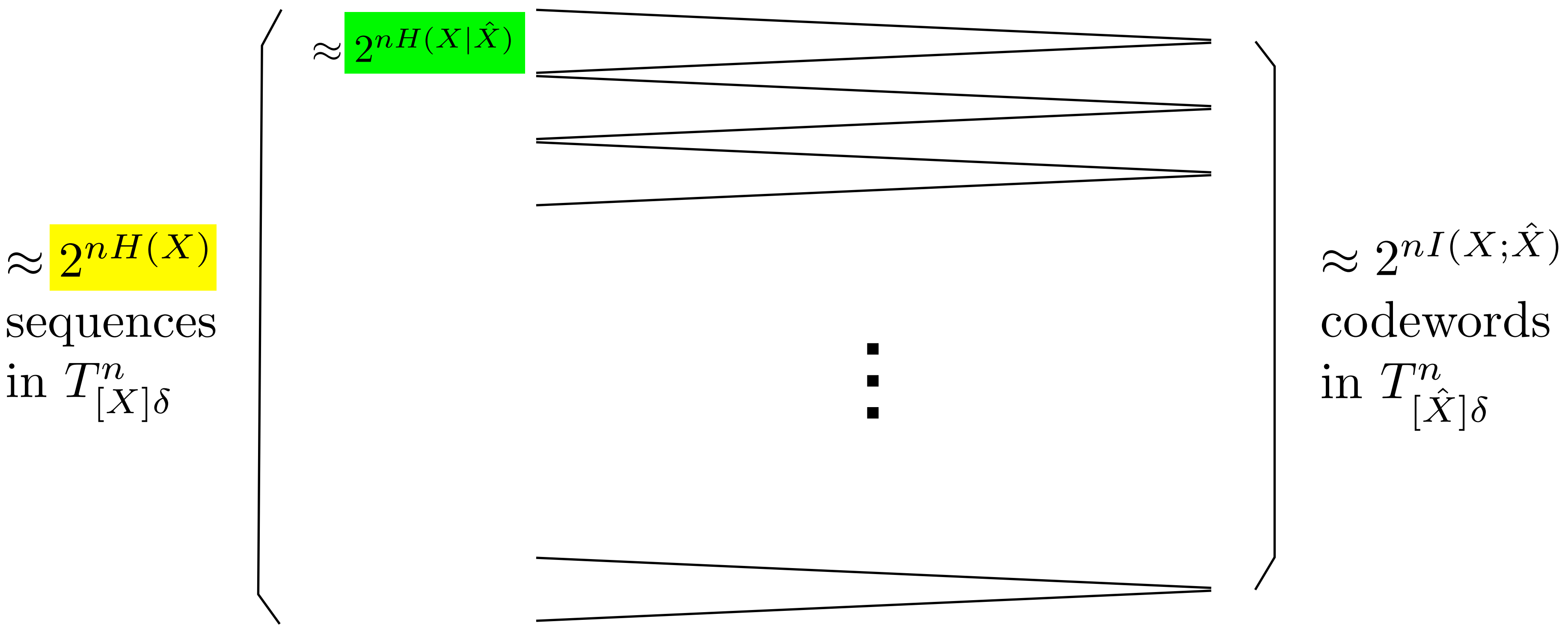
The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$



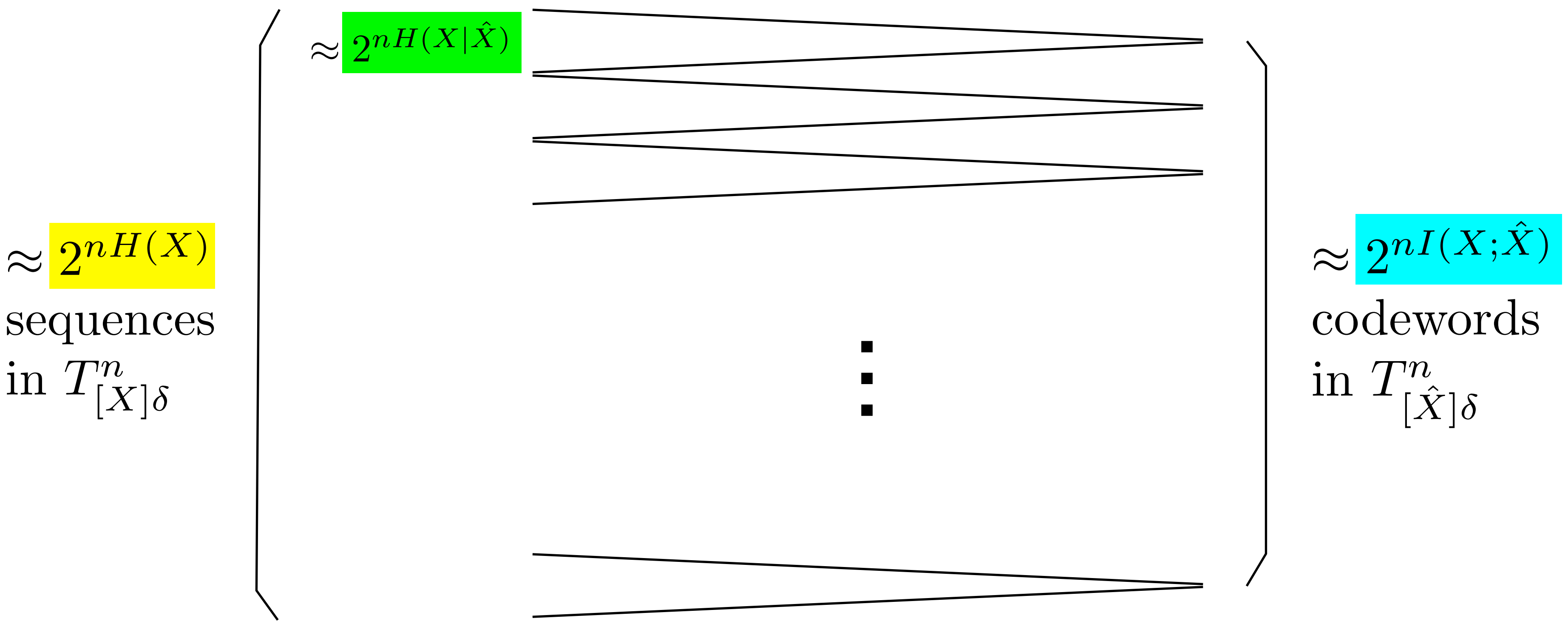
The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$



The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$



The number of codewords must be at least

$$\frac{2^{nH(X)}}{2^{nH(X|\hat{X})}} \approx 2^{nI(X;\hat{X})} \geq 2^{nR_I(D)}.$$