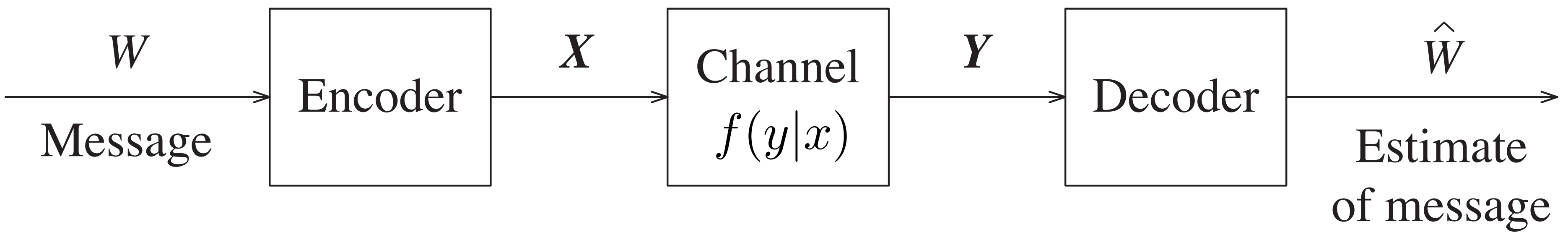
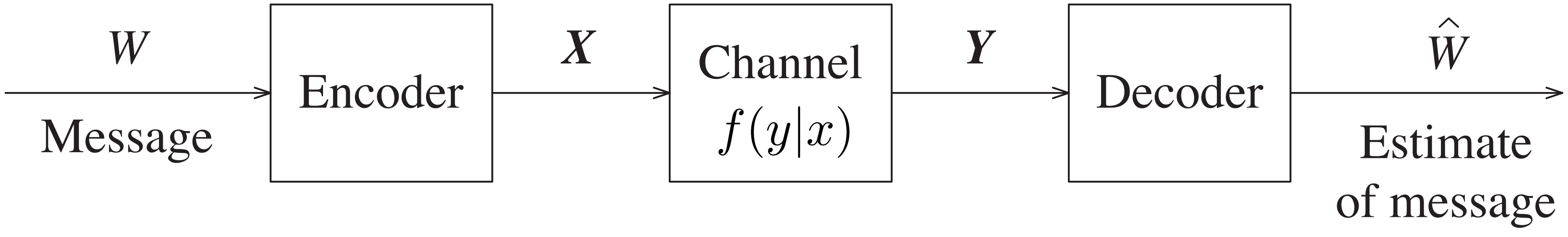
11.3.1 The Converse

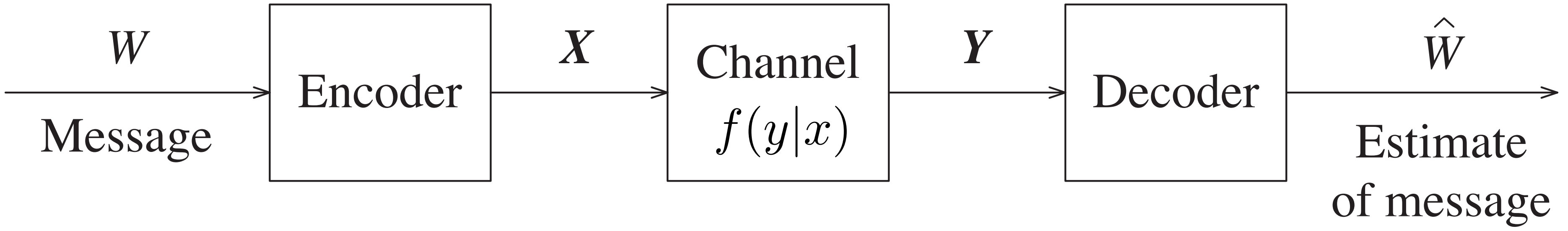

W
Message
Encoder
$\boldsymbol{X}$
Channel
$f(y|x)$
$\boldsymbol{Y}$
Decoder
$\widehat{W}$
Estimate
of message

Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

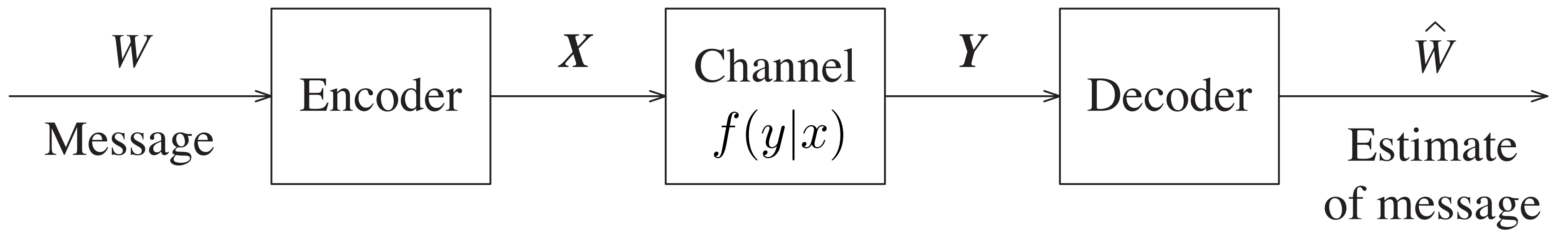
Sketch of Proof



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

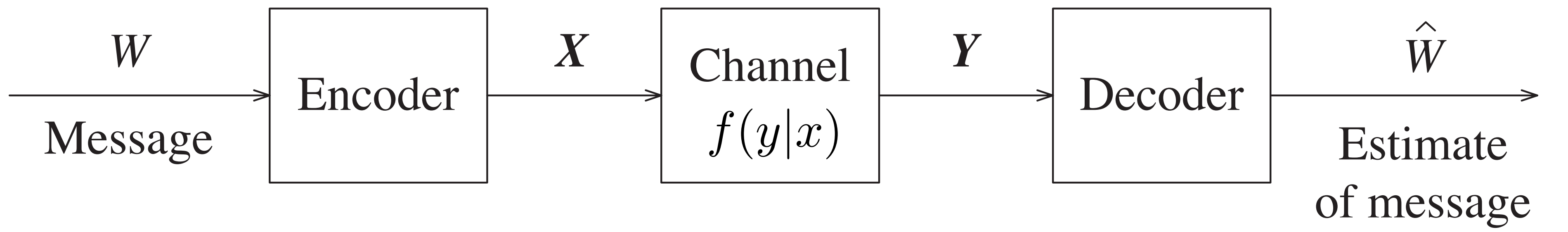
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

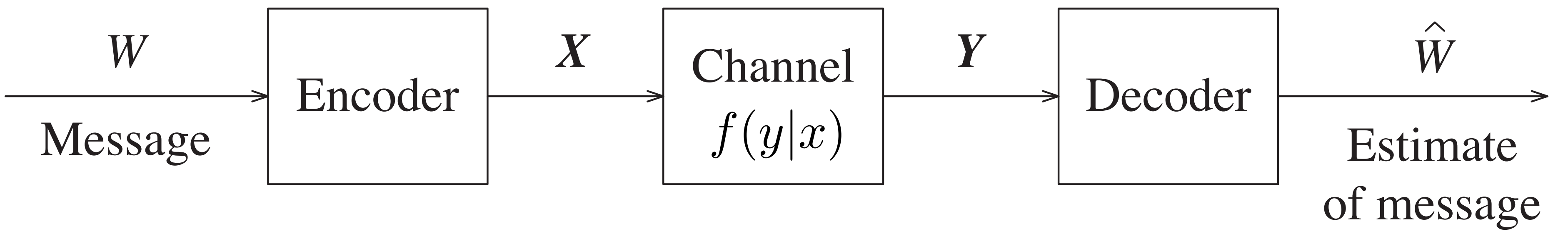
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

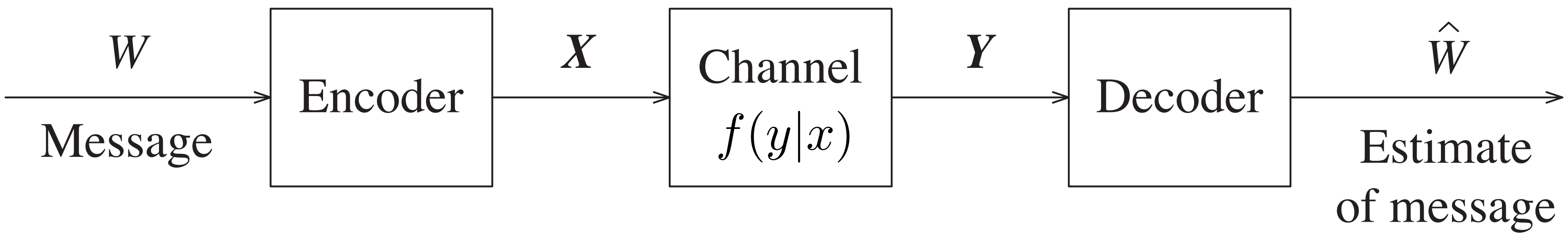
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete
- $\mathbf{X}$ – real but discrete



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

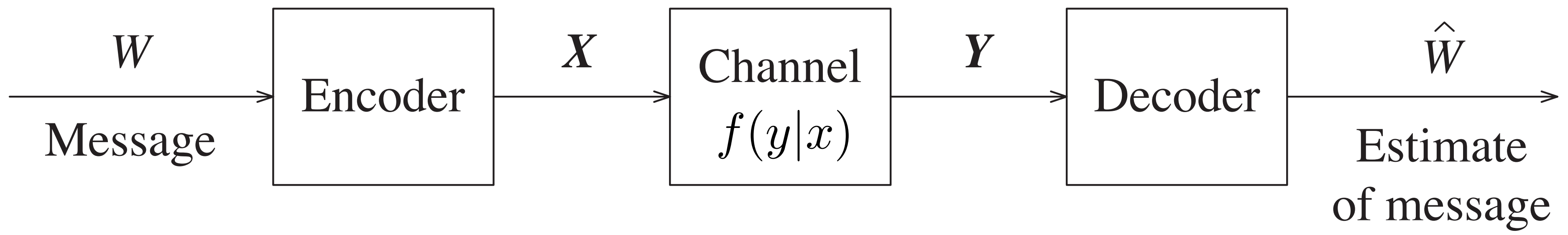
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete
- $\mathbf{X}$ – real but discrete
- $\mathbf{Y}$ – real and continuous



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

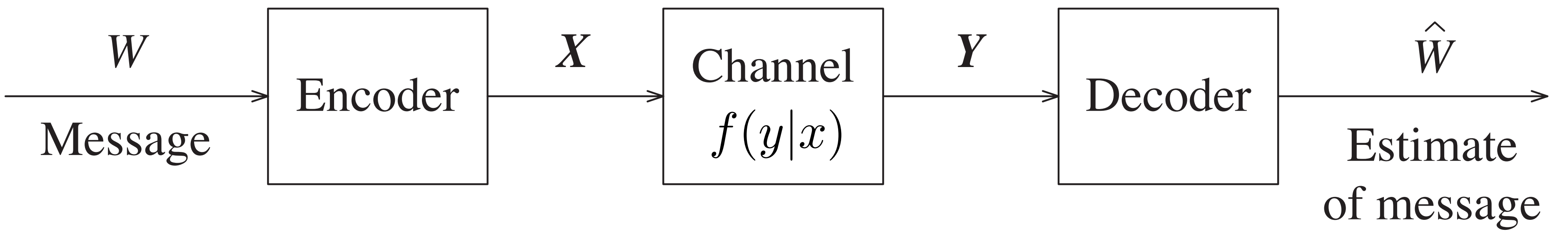
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete
- $\mathbf{X}$ – real but discrete
- $\mathbf{Y}$ – real and continuous
- $\mathbf{Y}$ needs to be handled with caution because it is continuous.



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

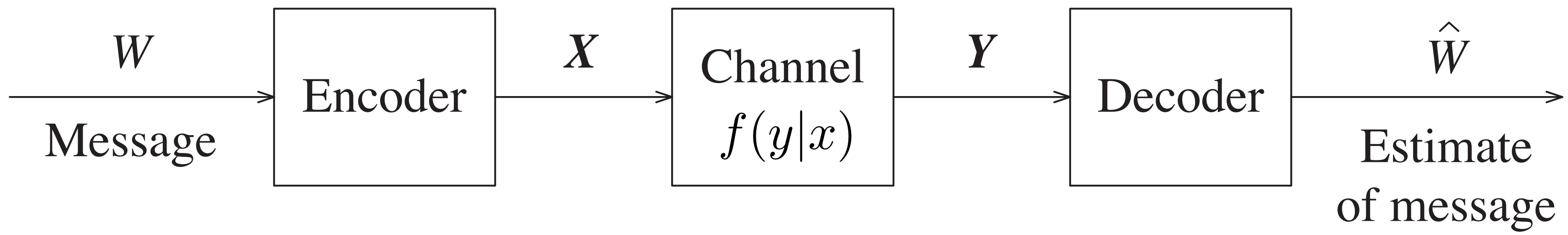
- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete
- $\mathbf{X}$ – real but discrete
- $\mathbf{Y}$ – real and continuous
- $\mathbf{Y}$ needs to be handled with caution because it is continuous.
- In particular, the existence of the conditional pdf $f(\mathbf{y}|\hat{w})$ needs to be established so that $I(\mathbf{X};\mathbf{Y}|\hat{W})$ can be defined.



Lemma 11.15 (Data Processing Theorem) $I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y})$.

Sketch of Proof

- $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$ can be established like the discrete case.
- W, $\hat{W}$ – discrete
- $\mathbf{X}$ – real but discrete
- $\mathbf{Y}$ – real and continuous
- $\mathbf{Y}$ needs to be handled with caution because it is continuous.
- In particular, the existence of the conditional pdf $f(\mathbf{y}|\hat{w})$ needs to be established so that $I(\mathbf{X};\mathbf{Y}|\hat{W})$ can be defined.
- Refer to the textbook for the technical details.



Proof of Converse

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\log M = H(W)$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + \underline{I(W;\hat{W})} \end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + \underline{I(W;\hat{W})} \\
&\leq H(W|\hat{W}) + \underline{I(\mathbf{X};\mathbf{Y})}
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + \underline{I(\mathbf{X};\mathbf{Y})} \\ &= H(W|\hat{W}) + \underline{h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X})} \end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + \underline{h(\mathbf{Y})} - h(\mathbf{Y}|\mathbf{X})
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + \underline{h(\mathbf{Y})} - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \underline{\sum_{i=1}^{n} h(Y_i)} - h(\mathbf{Y}|\mathbf{X})
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \underline{h(\mathbf{Y}|\mathbf{X})}
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \underline{h(\mathbf{Y}|\mathbf{X})} \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \underline{\sum_{i=1}^{n} h(Y_i|X_i)}
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} \underline{h(Y_i)} - \sum_{i=1}^{n} h(Y_i|X_i)
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} \underline{h(Y_i)} - \sum_{i=1}^{n} \underline{h(Y_i|X_i)}
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} \underline{h(Y_i)} - \sum_{i=1}^{n} \underline{h(Y_i|X_i)} \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} \underline{I(X_i;Y_i)}.
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

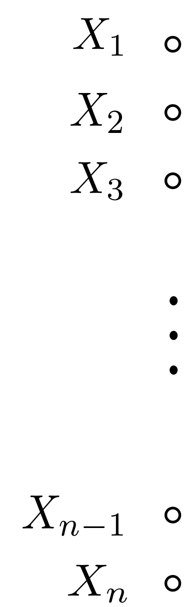
Proof of Converse

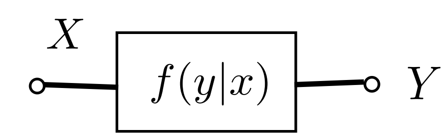
1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$





3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

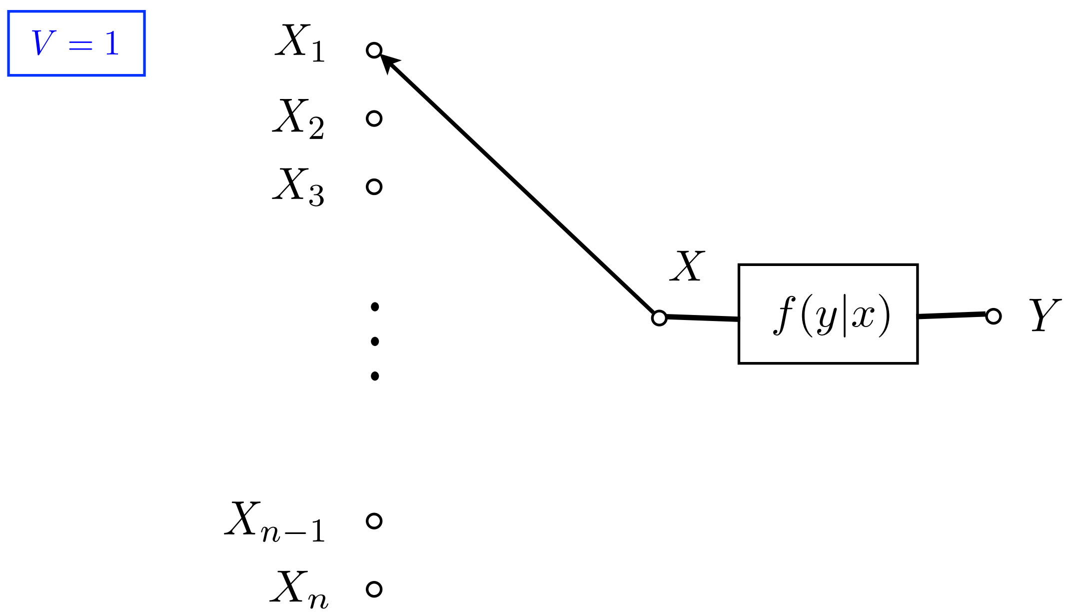
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



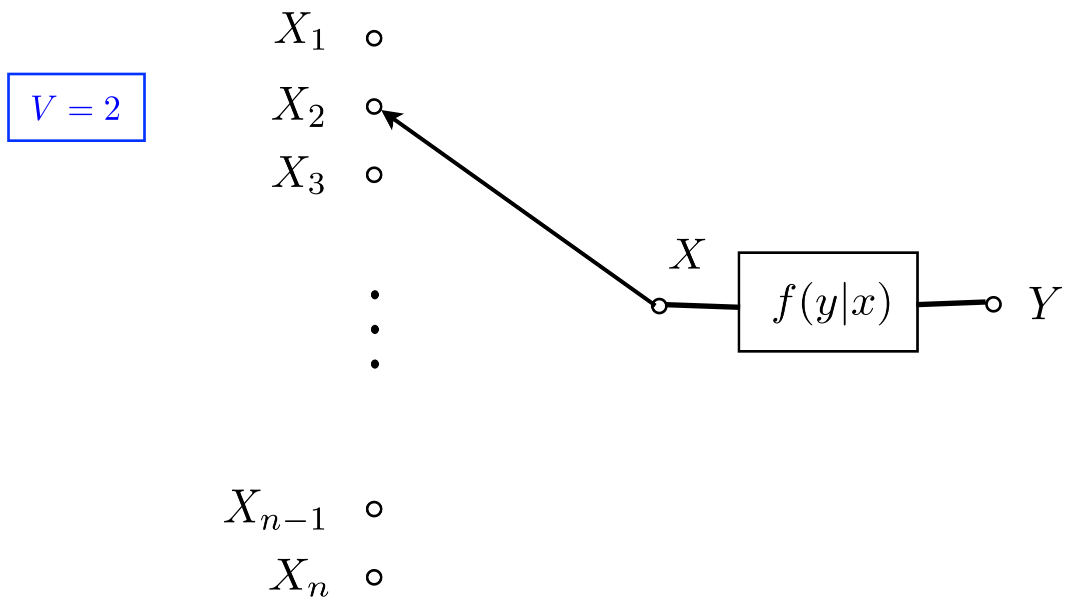
Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$



3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

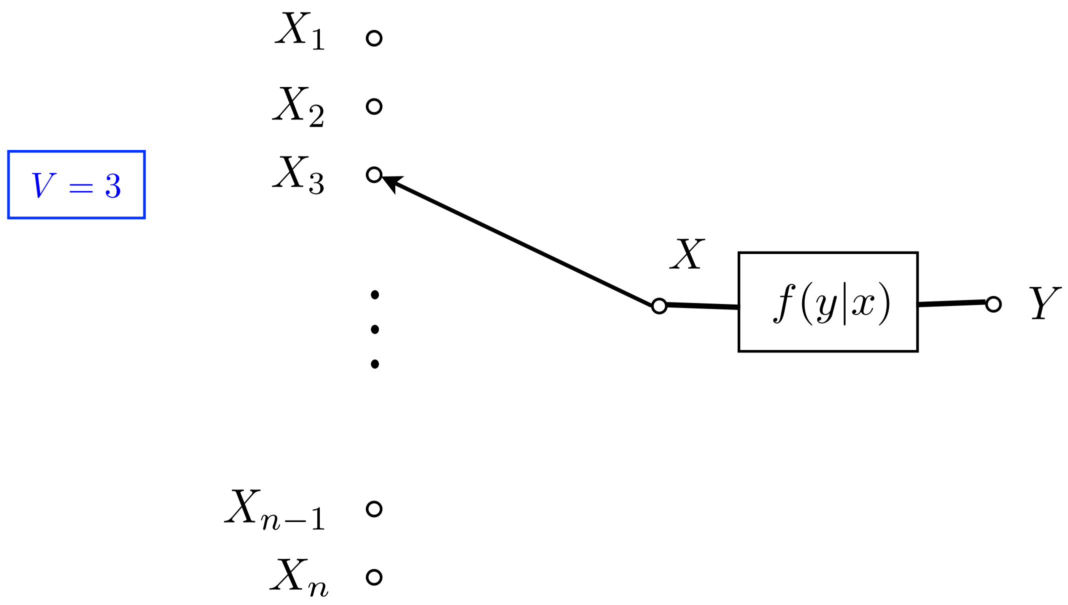
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



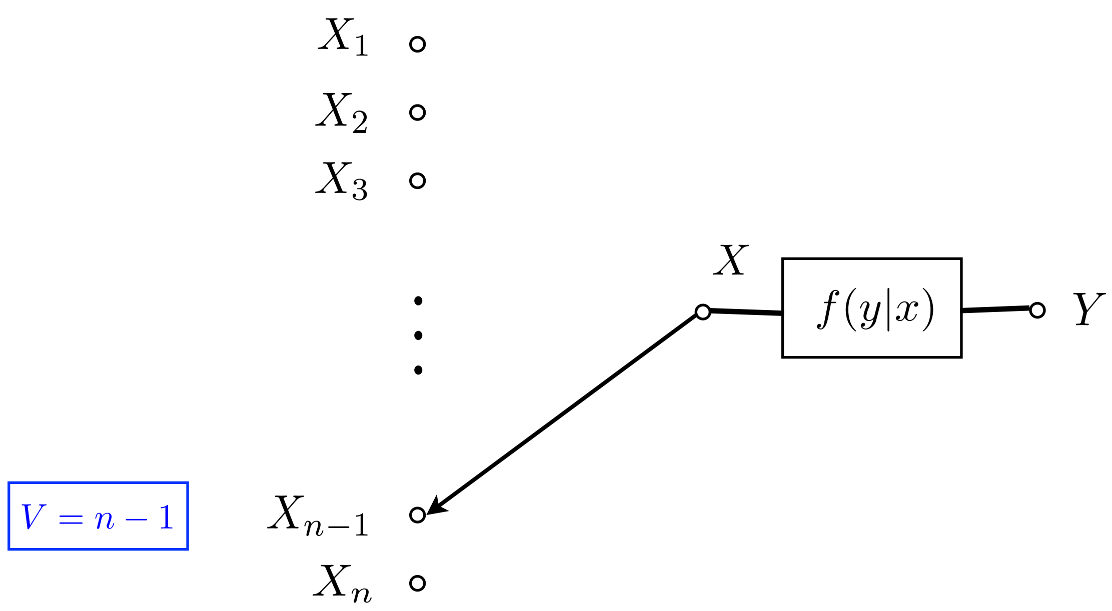
Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$



3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

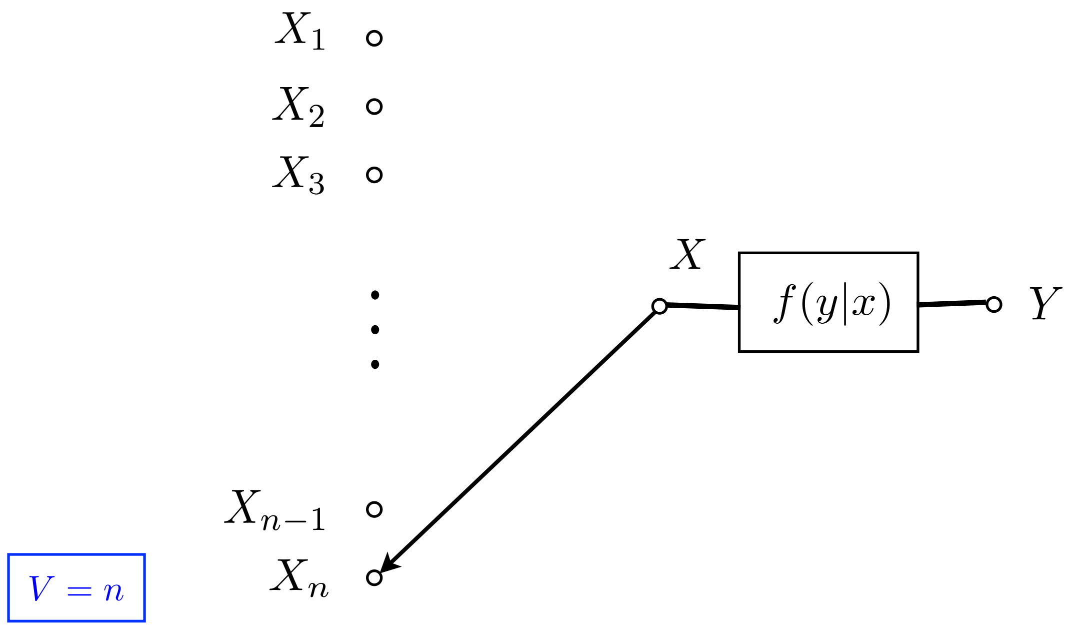
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$E\kappa(X) = EE[\kappa(X)|V]$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$E\kappa(X) = EE[\kappa(X)|V]$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i)
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i)
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|\cancel{V = i}] \\
&= \sum_{i=1}^{n} \frac{1}{n} \underline{E}\kappa(X_i) \\
&= \underline{E}\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right]
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right]
\end{aligned}$$

Each codeword satisfies the input constraint:

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(x_i(w)) \le P \quad \text{for } 1 \le w \le M.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i) \le P$$

with probability 1.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right]
\end{aligned}$$

Each codeword satisfies the input constraint:

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(x_i(w)) \leq P \quad \text{for } 1 \leq w \leq M.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i) \leq P$$

with probability 1.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

Each codeword satisfies the input constraint:

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(x_i(w)) \le P \quad \text{for } 1 \le w \le M.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i) \le P$$

with probability 1.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n}E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n}\kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

Each codeword satisfies the power constraint:

$$\frac{1}{n}\sum_{i=1}^{n}\kappa(x_i(w)) \leq P \quad \text{for } 1 \leq w \leq M.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^{n}\kappa(X_i) \leq P$$

with probability 1.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n}E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n}\kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

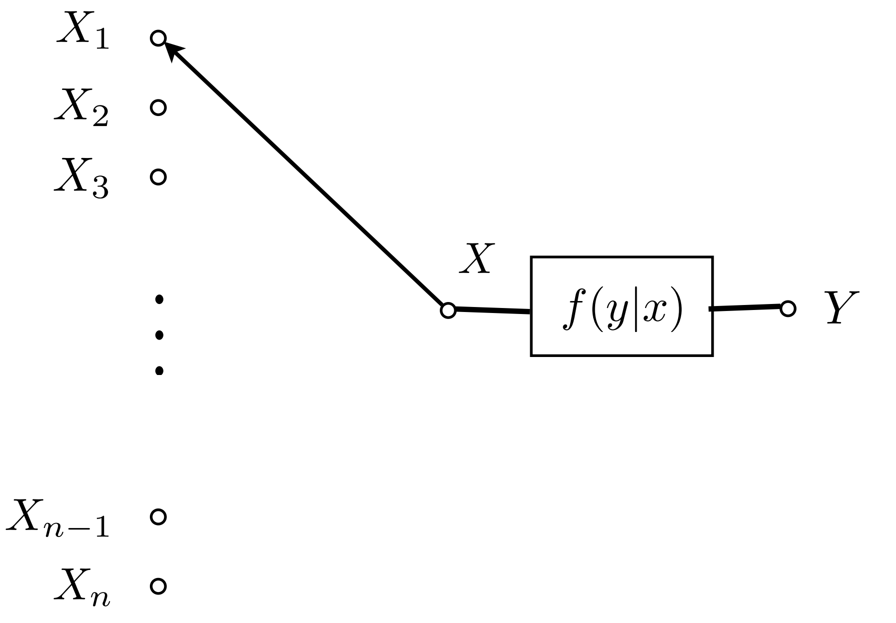
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

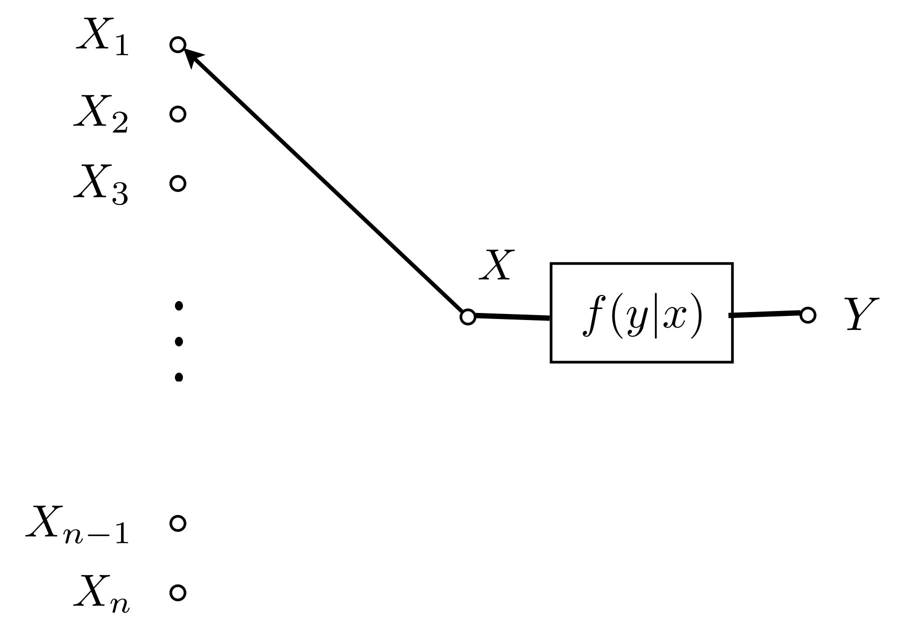
$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

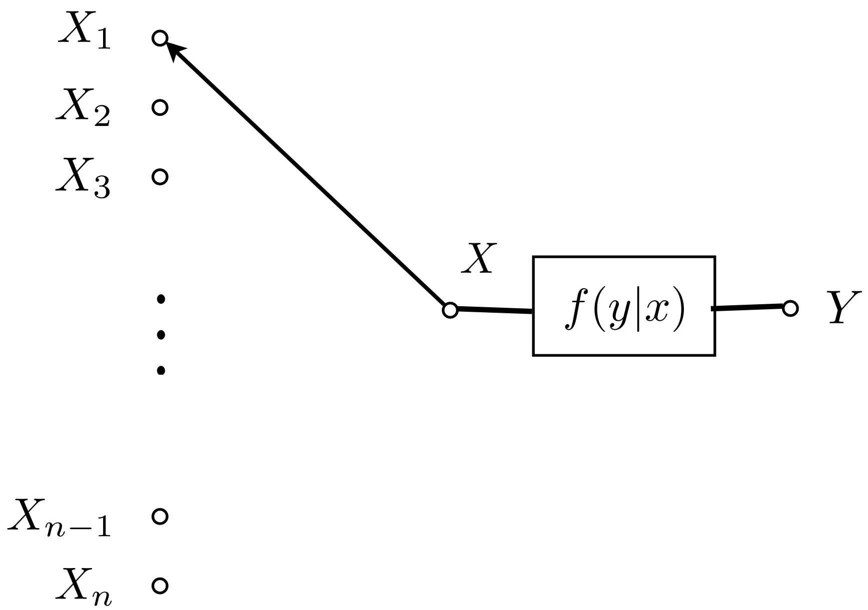
$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



6. By the concavity of mutual information with respect to the input distribution,

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

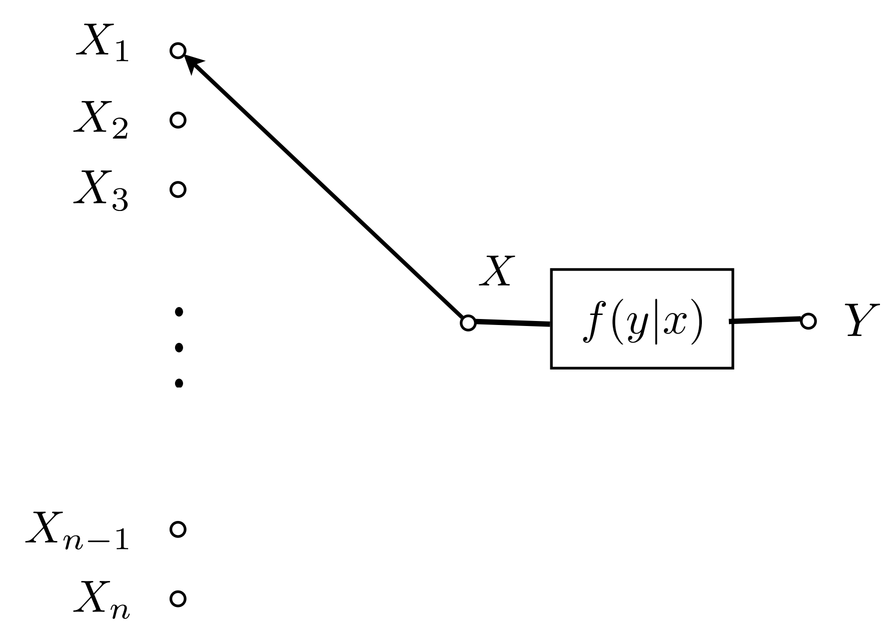
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

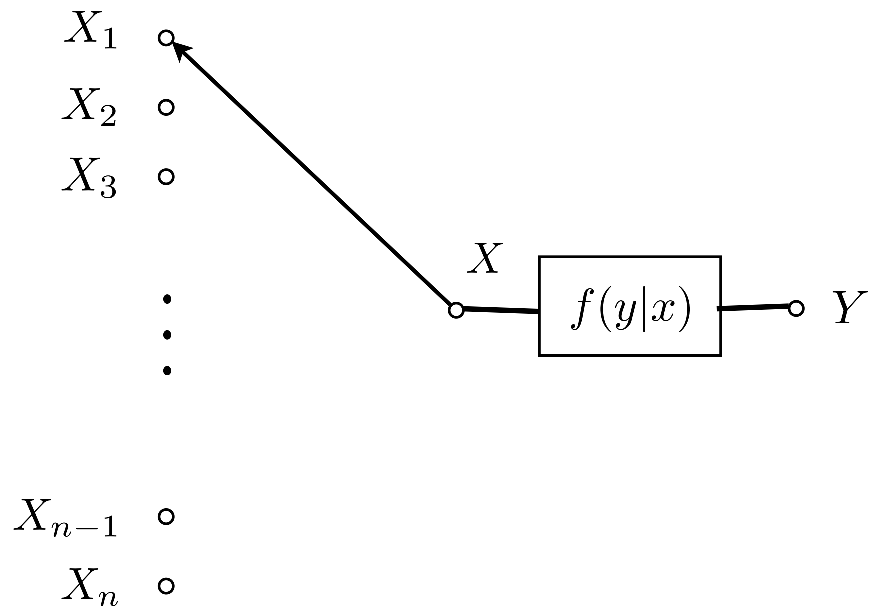
$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.



6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n} \sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n} \sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i).
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \le nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \leq nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

7. It follows that

$$n(R - \epsilon) < \log M \leq H(W|\hat{W}) + nC(P).$$

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \le nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n} \sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n} \sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

7. It follows that

$$n(R - \epsilon) < \underline{\log M \le H(W|\hat{W}) + nC(P)}.$$

8. Invoke Fano's inequality to conclude that $R \le C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \le nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

7. It follows that

$$n(R - \epsilon) < \log M \le H(W|\hat{W}) + nC(P).$$

8. Invoke Fano's inequality to conclude that $R \le C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \le nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

7. It follows that

$$n(R - \epsilon) < \log M \le H(W|\hat{W}) + nC(P).$$

8. Invoke Fano's inequality to conclude that $R \le C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \leq nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\}E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

7. It follows that

$n(R - \epsilon) <$ ~~$H(W|\hat{W})$~~ $+ nC(P).$

8. Invoke Fano's inequality to conclude that $R \leq C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \le nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \le i \le n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\le P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \le I(X;Y) \le C(P).$$

7. It follows that

$$n(R - \epsilon) < H(W|\hat{W}) + nC(P).$$

8. Invoke Fano's inequality to conclude that $R \le C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \leq nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n} \sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n} \sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

7. It follows that

$$n(R - \epsilon) < H(W|\hat{W}) + nC(P).$$

8. Invoke Fano's inequality to conclude that $R \leq C(P)$.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n and (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&= H(W|\hat{W}) + h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - h(\mathbf{Y}|\mathbf{X}) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Y_i|X_i) \\
&= H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i). \quad \leq nC(P)
\end{aligned}$$

3. Let V be a mixing random variable distributed uniformly on $\{1, 2, \cdots, n\}$ which is independent of X_i, $1 \leq i \leq n$.

4. Let $X = X_V$ and Y be the output of the channel with X being the input.

5. Then

$$\begin{aligned}
E\kappa(X) &= EE[\kappa(X)|V] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X)|V = i] \\
&= \sum_{i=1}^{n} \Pr\{V = i\} E[\kappa(X_i)|V = i] \\
&= \sum_{i=1}^{n} \frac{1}{n} E\kappa(X_i) \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} \kappa(X_i)\right] \\
&\leq P.
\end{aligned}$$

6. By the concavity of mutual information with respect to the input distribution,

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y) \leq C(P).$$

7. It follows that

$$n(R - \epsilon) < H(W|\hat{W}) + nC(P).$$

8. Invoke Fano's inequality to conclude that $R \leq C(P)$.

11.3.2 Achievability

Mutual Typicality

Mutual Typicality

- In the formula

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y),$$

Mutual Typicality

- In the formula

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y),$$

  X can be a mixed random variable, so it is difficult to consider sequences that are typical w.r.t. $F(x)$.

Mutual Typicality

- In the formula
$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y),$$
X can be a mixed random variable, so it is difficult to consider sequences that are typical w.r.t. $F(x)$.

- Need a new notion of joint typicality – mutual typicality.

Mutual Typicality

- In the formula
$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y),$$
X can be a mixed random variable, so it is difficult to consider sequences that are typical w.r.t. $F(x)$.

- Need a new notion of joint typicality – mutual typicality.

- Recall that for any input distribution $F(x)$, $f(y)$ exists as long as $f(y|x)$ exists. Hence

Mutual Typicality

- In the formula
$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y),$$
X can be a mixed random variable, so it is difficult to consider sequences that are typical w.r.t. $F(x)$.

- Need a new notion of joint typicality – mutual typicality.

- Recall that for any input distribution $F(x)$, $f(y)$ exists as long as $f(y|x)$ exists. Hence
$$I(X;Y) = E \log \frac{f(Y|X)}{f(Y)}.$$

Definition 11.16 The mutually typical set $\Psi^n_{[XY]\delta}$ with respect to $F(x, y)$ is the set of $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X; Y) \right| \leq \delta,$$

Definition 11.16 The mutually typical set $\Psi^n_{[XY]\delta}$ with respect to $F(x, y)$ is the set of $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta,$$

where

$$f(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} f(y_i|x_i) \quad \text{and} \quad f(\mathbf{y}) = \prod_{i=1}^{n} f(y_i),$$

Definition 11.16 The mutually typical set $\Psi^n_{[XY]\delta}$ with respect to $F(x, y)$ is the set of $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta,$$

where

$$f(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} f(y_i|x_i) \quad \text{and} \quad f(\mathbf{y}) = \prod_{i=1}^{n} f(y_i),$$

and δ is an arbitrarily small positive number. A pair of sequences $(\mathbf{x}, \mathbf{y})$ is called mutually δ-typical if it is in $\Psi^n_{[XY]\delta}$.

Lemma 11.17 For any $\delta > 0$, for sufficiently large n,

$$\Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \geq 1 - \delta.$$

Proof

Lemma 11.17 For any $\delta > 0$, for sufficiently large n,

$$\Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \geq 1 - \delta.$$

Proof

1. Consider

$$\frac{1}{n} \log \frac{f(\mathbf{Y}|\mathbf{X})}{f(\mathbf{Y})} = \frac{1}{n} \log \prod_{i=1}^{n} \frac{f(Y_i|X_i)}{f(Y_i)}$$

Lemma 11.17 For any $\delta > 0$, for sufficiently large n,

$$\Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \geq 1 - \delta.$$

Proof

1. Consider

$$\frac{1}{n} \log \frac{f(\mathbf{Y}|\mathbf{X})}{f(\mathbf{Y})} = \frac{1}{n} \log \prod_{i=1}^{n} \frac{f(Y_i|X_i)}{f(Y_i)} = \frac{1}{n} \sum_{i=1}^{n} \log \frac{f(Y_i|X_i)}{f(Y_i)}.$$

Lemma 11.17 For any $\delta > 0$, for sufficiently large n,

$$\Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \geq 1 - \delta.$$

Proof

1. Consider

$$\frac{1}{n} \log \frac{f(\mathbf{Y}|\mathbf{X})}{f(\mathbf{Y})} = \frac{1}{n} \log \prod_{i=1}^{n} \frac{f(Y_i|X_i)}{f(Y_i)} = \frac{1}{n} \sum_{i=1}^{n} \log \frac{f(Y_i|X_i)}{f(Y_i)}.$$

2. By WLLN,

$$\frac{1}{n} \sum_{i=1}^{n} \log \frac{f(Y_i|X_i)}{f(Y_i)} \;\rightarrow\; E \log \frac{f(Y|X)}{f(Y)} = I(X;Y)$$

   in probability.

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$1 \quad \geq \quad \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then
$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or
$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\geq \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \le \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \ge f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\ge \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \end{aligned}$$

$f(x)dx$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \le \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \ge f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\ge \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta}\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} \underline{f(\mathbf{y}|\mathbf{x})} dF(\mathbf{x})\, d\mathbf{y} \end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\geq \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta}\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \\ &\geq 2^{n(I(X;Y)-\delta)} \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}) dF(\mathbf{x}) d\mathbf{y} \end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \le \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \ge f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\ge \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \\ &\ge 2^{n(I(X;Y)-\delta)} \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}) dF(\mathbf{x}) d\mathbf{y} \end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned}
1 &\geq \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\
&= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \\
&\geq 2^{n(I(X;Y)-\delta)} \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}) dF(\mathbf{x}) d\mathbf{y} \\
&= 2^{n(I(X;Y)-\delta)} \underline{\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\}}.
\end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \le \delta.$$

2. Then

$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \ge 2^{n(I(X;Y)-\delta)}$$

or

$$f(\mathbf{y}|\mathbf{x}) \ge f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned}
1 &\ge \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\
&= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \\
&\ge 2^{n(I(X;Y)-\delta)} \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}) dF(\mathbf{x}) d\mathbf{y} \\
&= 2^{n(I(X;Y)-\delta)} \Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\}.
\end{aligned}$$

Lemma 11.18 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where X' and Y' are independent and have the same marginal distributions as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \Psi^n_{[XY]\delta}$, by definition,

$$\left| \frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} - I(X;Y) \right| \leq \delta.$$

2. Then
$$\frac{1}{n} \log \frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq I(X;Y) - \delta$$

which implies

$$\frac{f(\mathbf{y}|\mathbf{x})}{f(\mathbf{y})} \geq 2^{n(I(X;Y)-\delta)}$$

or
$$f(\mathbf{y}|\mathbf{x}) \geq f(\mathbf{y}) 2^{n(I(X;Y)-\delta)}.$$

3. Consider

$$\begin{aligned} 1 &\geq \Pr\{(\mathbf{X}, \mathbf{Y}) \in \Psi^n_{[XY]\delta})\} \\ &= \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}|\mathbf{x}) dF(\mathbf{x})\, d\mathbf{y} \\ &\geq 2^{n(I(X;Y)-\delta)} \int\int_{\Psi^n_{[XY]\delta}} f(\mathbf{y}) dF(\mathbf{x}) d\mathbf{y} \\ &= 2^{n(I(X;Y)-\delta)} \Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\}. \end{aligned}$$

4. Hence

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in \Psi^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\delta)}.$$

Random Coding Scheme

Random Coding Scheme

Parameter Settings

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P - \gamma) > C(P) - \frac{\epsilon}{6}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

$$C(P) = \sup_{F(x):E\kappa(X)\le P} I(X;Y)$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

$$\begin{aligned} C(P) &= \sup_{F(x): E\kappa(X) \le P} I(X;Y) \\ C(P-\gamma) &= \sup_{F(x): E\kappa(X) \le P-\gamma} I(X;Y) \end{aligned}$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$\begin{aligned} C(P) &= \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \\ C(P-\gamma) &= \sup_{F(x): E\kappa(X) \leq P-\gamma} I(X;Y) \end{aligned}$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$E\kappa(X) \le P - \gamma$ and

$$\begin{aligned} C(P) &= \sup_{F(x):E\kappa(X)\le P} I(X;Y) \\ C(P-\gamma) &= \sup_{F(x):E\kappa(X)\le P-\gamma} I(X;Y) \end{aligned}$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y)$$

$$\underline{C(P-\gamma)} = \sup_{F(x): E\kappa(X) \le P-\gamma} \underline{I(X;Y)}$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P - \gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P - \gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{8}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \underline{\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P - \gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P - \gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P - \gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P - \gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P - \gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P - \gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n} \log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P - \gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P - \gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P - \gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n} \log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

$$\Pr\{Y_i \leq y_i, 1 \leq i \leq n | \mathbf{X}(W) = \mathbf{x}\} = \prod_{i=1}^{n} \int_{-\infty}^{y_i} f(y|x_i)\,dy.$$

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

$$\Pr\{Y_i \leq y_i, 1 \leq i \leq n | \mathbf{X}(W) = \mathbf{x}\} = \prod_{i=1}^{n} \int_{-\infty}^{y_i} f(y|x_i)dy.$$

7. The sequence $\mathbf{Y}$ is decoded to the message w if

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

$$\Pr\{Y_i \le y_i, 1 \le i \le n | \mathbf{X}(W) = \mathbf{x}\} = \prod_{i=1}^{n} \int_{-\infty}^{y_i} f(y|x_i)dy.$$

7. The sequence $\mathbf{Y}$ is decoded to the message w if

- $(\mathbf{X}(w), \mathbf{Y}) \in \Psi^n_{[XY]\delta}$, and

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \ge C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

$$\Pr\{Y_i \le y_i, 1 \le i \le n | \mathbf{X}(W) = \mathbf{x}\} = \prod_{i=1}^{n} \int_{-\infty}^{y_i} f(y|x_i)dy.$$

7. The sequence $\mathbf{Y}$ is decoded to the message w if

- $(\mathbf{X}(w), \mathbf{Y}) \in \Psi^n_{[XY]\delta}$, and
- there does not exist $w' \neq w$ such that $(\mathbf{X}(w'), \mathbf{Y}) \in \Psi^n_{[XY]\delta}$.

Random Coding Scheme

Parameter Settings

1. Fix $\epsilon > 0$ and input distribution $F(x)$. Let δ to be specified later.

2. Since $C(P)$ is left-continuous, there exists $\gamma > 0$ such that

$$C(P-\gamma) > C(P) - \frac{\epsilon}{6}.$$

3. By the definition of $C(P-\gamma)$, there exists an input random variable X such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

4. Choose for a sufficiently large n an even integer M satisfying

$$I(X;Y) - \frac{\epsilon}{6} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{8}.$$

5. Then

$$\frac{1}{n}\log M > I(X;Y) - \frac{\epsilon}{6} \geq C(P-\gamma) - \frac{\epsilon}{3} > C(P) - \frac{\epsilon}{2}.$$

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code randomly by generating M codewords in $\Re^n$ independently and identically according to $F(x)^n$.

2. Denote these codewords by $\tilde{\mathbf{X}}(1), \tilde{\mathbf{X}}(2), \cdots, \tilde{\mathbf{X}}(M)$.

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

4. A message W is chosen from $\mathcal{W}$ uniformly.

5. The sequence $\mathbf{X} = \tilde{\mathbf{X}}(W)$ is transmitted through the channel.

6. The channel outputs a sequence $\mathbf{Y}$ according to

$$\Pr\{Y_i \leq y_i, 1 \leq i \leq n | \mathbf{X}(W) = \mathbf{x}\} = \prod_{i=1}^{n} \int_{-\infty}^{y_i} f(y|x_i)dy.$$

7. The sequence $\mathbf{Y}$ is decoded to the message w if

- $(\mathbf{X}(w), \mathbf{Y}) \in \Psi^n_{[XY]\delta}$, and
- there does not exist $w' \neq w$ such that $(\mathbf{X}(w'), \mathbf{Y}) \in \Psi^n_{[XY]\delta}$.

Otherwise, $\mathbf{Y}$ is decoded to a constant message in $\mathcal{W}$. Denote by $\hat{W}$ the message to which $\mathbf{Y}$ is decoded.

Performance Analysis

Performance Analysis

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \le P - \gamma \quad \text{and} \quad I(X;Y) \ge C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\Pr\{Err\} = \Pr\{Err|W = 1\}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \underline{\Pr\{E_d|W = 1\}}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \underline{\Pr\{E_d|W = 1\}}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \underline{\Pr\{E_d|W=1\}}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X; Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$\Pr\{E_e|W = 1\}$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \end{aligned}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma) + \gamma \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \underline{\gamma}) + \underline{\gamma} \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (\underline{P - \gamma}) + \gamma \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (\underline{P - \gamma}) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > \underline{E\kappa(X)} + \gamma \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \end{aligned}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (\underline{P - \gamma}) + \gamma \right\} \\ &\underline{\leq} \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > \underline{E\kappa(X)} + \gamma \right\} \end{aligned}$$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\underline{\tilde{X}_i(1)}) > E\kappa(X) + \gamma \right\} \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\underline{\tilde{X}_i(1))} > E\kappa(X) + \gamma \right\} \end{aligned}$$

$\tilde{X}_i(1) \sim X$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \underline{\kappa(\tilde{X}_i(1))} > E\kappa(X) + \gamma \right\} \end{aligned}$$

$\tilde{X}_i(1) \sim X$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \end{aligned}$$

$$\tilde{X}_i(1) \sim X$$
$$\kappa(\tilde{X}_i(1)) \sim \kappa(X)$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \underline{\frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1))} > E\kappa(X) + \gamma \right\} \end{aligned}$$

$$\tilde{X}_i(1) \sim X$$
$$\kappa(\tilde{X}_i(1)) \sim \kappa(X)$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > \underline{E\kappa(X)} + \gamma \right\} \end{aligned}$$

$\tilde{X}_i(1) \sim X$

$\kappa(\tilde{X}_i(1)) \sim \kappa(X)$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P-\gamma) - \frac{\epsilon}{6}.$$

5. By WLLN, for sufficiently large n,

$\Pr\{E_e|W=1\}$

$$\begin{aligned} &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) \underline{\geq E\kappa(X) + \gamma} \right\} \end{aligned}$$

$\tilde{X}_i(1) \sim X$

$\kappa(\tilde{X}_i(1)) \sim \kappa(X)$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err \mid W = 1\} \\ &\leq \Pr\{E_e \mid W = 1\} + \Pr\{E_d \mid W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d \mid W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d \mid W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e \mid W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) \underline{> E\kappa(X) + \gamma} \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

$\tilde{X}_i(1) \sim X$

$\kappa(\tilde{X}_i(1)) \sim \kappa(X)$

Code Construction: Fix input distribution $F(x)$ such that

$$E\kappa(X) \leq P - \gamma \quad \text{and} \quad I(X;Y) \geq C(P - \gamma) - \frac{\epsilon}{6}.$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.

2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\le \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \le \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\le \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \\ &\le \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\ &\quad \leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\ &\quad \leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W = 1\} \\ &\leq \Pr\{E_e|W = 1\} + \Pr\{E_d|W = 1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W = 1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W = 1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W = 1\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W = 1 \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &\quad = \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\ &\quad \leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \\ &\quad \leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

Performance Analysis

1. Let $\tilde{\mathbf{X}}(w) = (\tilde{X}_1(w), \tilde{X}_2(w), \cdots, \tilde{X}_n(w))$.
2. Define the error event $Err = E_e \cup E_d$, where

$$E_e = \left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(W)) > P \right\}$$

and

$$E_d = \{\hat{W} \neq W\}.$$

3. Consider

$$\begin{aligned} \Pr\{Err\} &= \Pr\{Err|W=1\} \\ &\leq \Pr\{E_e|W=1\} + \Pr\{E_d|W=1\}. \end{aligned}$$

4. The analysis of $\Pr\{E_d|W=1\}$ is exactly the same as the analysis of the decoding error in the discrete case. So we can choose δ to be small to make

$$\Pr\{E_d|W=1\} \leq \frac{\epsilon}{4}$$

for sufficiently large n.

5. By WLLN, for sufficiently large n,

$$\begin{aligned} &\Pr\{E_e|W=1\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\ &= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\ &\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\ &\leq \frac{\epsilon}{4}. \end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
&\Pr\{E_e|W=1\} \\
&\quad = \quad \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\
&\quad = \quad \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&\quad = \quad \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\quad \leq \quad \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\quad \leq \quad \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e | W = 1\} \\
&= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W = 1 \right\} \\
&= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P - \gamma) + \gamma \right\} \\
&\leq \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X) + \gamma \right\} \\
&\leq \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err | \mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e|W=1\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma\right\} \\
&\le \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma\right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\leq \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\leq \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \leq \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \leq \epsilon$ and the input constraint.

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e|W=1\} & \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \middle| W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e|W=1\} &= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma\right\} \\
&\le \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma\right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \le \epsilon$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \le \epsilon$$

and

$$\Pr\{E_e|\mathcal{C}^*, W=w\} \le \epsilon.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\le \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \le \epsilon$$

and

$$\Pr\{E_e|\mathcal{C}^*, W=w\} \le \epsilon.$$

11. Observe that conditioning on $\{\mathcal{C}^*, W=w\}$, the codeword $\tilde{\mathbf{X}}(w)$ is deterministic, so either

Performance Analysis

5. By WLLN, for sufficiently large n,

$$
\begin{aligned}
\Pr\{E_e|W=1\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma\right\} \\
&\le \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma\right\} \\
&\le \frac{\epsilon}{4}.
\end{aligned}
$$

6. Therefore,

$$\Pr\{Err\} \le \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \le \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \le \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \le \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \le \epsilon$$

and

$$\Pr\{E_e|\underline{\mathcal{C}^*, W=w}\} \le \epsilon.$$

11. Observe that conditioning on $\{\mathcal{C}^*, W=w\}$, the codeword $\tilde{\mathbf{X}}(w)$ is deterministic, so either

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 0$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma\right\} \\
&\leq \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma\right\} \\
&\leq \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \leq \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \leq \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \leq \epsilon$$

and

$$\Pr\{E_e|\underline{\mathcal{C}^*, W=w}\} \leq \epsilon.$$

11. Observe that conditioning on $\{\mathcal{C}^*, W=w\}$, the codeword $\tilde{\mathbf{X}}(w)$ is deterministic, so either

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 0$$

or

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 1.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1 \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > P \right\} \\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma \right\} \\
&\leq \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} \kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma \right\} \\
&\leq \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \leq \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \leq \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \leq \epsilon$$

and

$$\Pr\{E_e|\mathcal{C}^*, W=w\} \leq \epsilon.$$

11. Observe that conditioning on $\{\mathcal{C}^*, W=w\}$, the codeword $\tilde{\mathbf{X}}(w)$ is deterministic, so either

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 0$$

or

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 1.$$

Performance Analysis

5. By WLLN, for sufficiently large n,

$$\begin{aligned}
&\Pr\{E_e|W=1\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P \,\middle|\, W=1\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > P\right\} \\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > (P-\gamma)+\gamma\right\} \\
&\leq \Pr\left\{\frac{1}{n}\sum_{i=1}^{n}\kappa(\tilde{X}_i(1)) > E\kappa(X)+\gamma\right\} \\
&\leq \frac{\epsilon}{4}.
\end{aligned}$$

6. Therefore,

$$\Pr\{Err\} \leq \frac{\epsilon}{2}$$

which implies for some codebook $\mathcal{C}^*$,

$$\Pr\{Err|\mathcal{C}^*\} \leq \frac{\epsilon}{2}.$$

7. Rank the codewords in $\mathcal{C}^*$ in ascending order according to $\Pr\{Err|\mathcal{C}^*, W=w\}$.

8. After discarding the worst half of the codewords in $\mathcal{C}^*$, if a codeword $\tilde{\mathbf{X}}(w)$ remains in $\mathcal{C}^*$, then

$$\Pr\{Err|\mathcal{C}^*, W=w\} \leq \epsilon.$$

9. However, it is not clear whether $\tilde{\mathbf{X}}(w)$ satisfies both $\lambda_w \leq \epsilon$ and the input constraint.

10. Since $Err = E_e \cup E_d$, we have

$$\lambda_w = \Pr\{E_d|\mathcal{C}^*, W=w\} \leq \epsilon$$

and

$$\Pr\{E_e|\mathcal{C}^*, W=w\} \leq \epsilon.$$

11. Observe that conditioning on $\{\mathcal{C}^*, W=w\}$, the codeword $\tilde{\mathbf{X}}(w)$ is deterministic, so either

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 0$$

or

$$\Pr\{E_e|\mathcal{C}^*, W=w\} = 1.$$

Therefore, $\Pr\{E_e|\mathcal{C}^*, W=w\} = 0$.