香港中文大學
The Chinese University of Hong Kong

7.4 Achievability

How to Prove Achievability

How to Prove Achievability

- Consider a DMC $p(y|x)$.

How to Prove Achievability

- Consider a DMC $p(y|x)$.
- For every input distribution $p(x)$, prove that the rate $I(X;Y)$ is achievable by showing for large n the existence of a channel code such that
  1. the rate of the code is arbitrarily close to $I(X;Y)$;
  2. the maximal probability of error λ_{max} is arbitrarily small.

How to Prove Achievability

- Consider a DMC $p(y|x)$.
- For every input distribution $p(x)$, prove that the rate $I(X;Y)$ is achievable by showing for large n the existence of a channel code such that
  1. the rate of the code is arbitrarily close to $I(X;Y)$;
  2. the maximal probability of error λ_{max} is arbitrarily small.
- Choose the input distribution $p(x)$ to be one that achieves the channel capacity, i.e., $I(X;Y) = C$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y})\in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \le 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \le 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \le 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y})$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned} &\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\ &= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned} \Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} &= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\ &\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y})\in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \le 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \le 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \le 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} &= \sum_{(\mathbf{x},\mathbf{y})\in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\le 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&= 2^{-n(H(X)+H(Y)-H(X,Y)-\xi-\eta-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&= 2^{-n(\underline{H(X)+H(Y)-H(X,Y)}-\xi-\eta-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&\quad = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\quad \leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&\quad = 2^{-n(\underline{H(X)+H(Y)-H(X,Y)}-\xi-\eta-\zeta)} \\
&\quad = 2^{-n(\underline{I(X;Y)}-\xi-\eta-\zeta)}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&\quad = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\quad \leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&\quad = 2^{-n(H(X)+H(Y)-H(X,Y)-\xi-\eta-\zeta)} \\
&\quad = 2^{-n(I(X;Y)\underline{-\xi-\eta-\zeta})}
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \le 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \le 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \le 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\le 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&= 2^{-n(H(X)+H(Y)-H(X,Y)-\xi-\eta-\zeta)} \\
&= 2^{-n(I(X;Y)\underline{-\xi-\eta-\zeta})} \\
&= 2^{-n(I(X;Y)\underline{-\tau})},
\end{aligned}$$

Lemma 7.17 Let $(\mathbf{X}', \mathbf{Y}')$ be n i.i.d. copies of a pair of generic random variables (X', Y'), where $X' \sim X$, $Y' \sim Y$, and X' and Y' are independent. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)},$$

where $\tau \to 0$ as $\delta \to 0$.

Proof

1. Consider

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} = \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}).$$

2. By the consistency of strong typicality, for $(\mathbf{x}, \mathbf{y}) \in T^n_{[XY]\delta}$, $\mathbf{x} \in T^n_{[X]\delta}$ and $\mathbf{y} \in T^n_{[Y]\delta}$.

3. By the strong AEP, all the $p(\mathbf{x})$ and $p(\mathbf{y})$ in the above summation satisfy

$$p(\mathbf{x}) \leq 2^{-n(H(X)-\eta)}$$

and

$$p(\mathbf{y}) \leq 2^{-n(H(Y)-\zeta)},$$

where $\eta, \zeta \to 0$ as $\delta \to 0$.

4. By the strong JAEP,

$$|T^n_{[XY]\delta}| \leq 2^{n(H(X,Y)+\xi)},$$

where $\xi \to 0$ as $\delta \to 0$.

5. Then we have

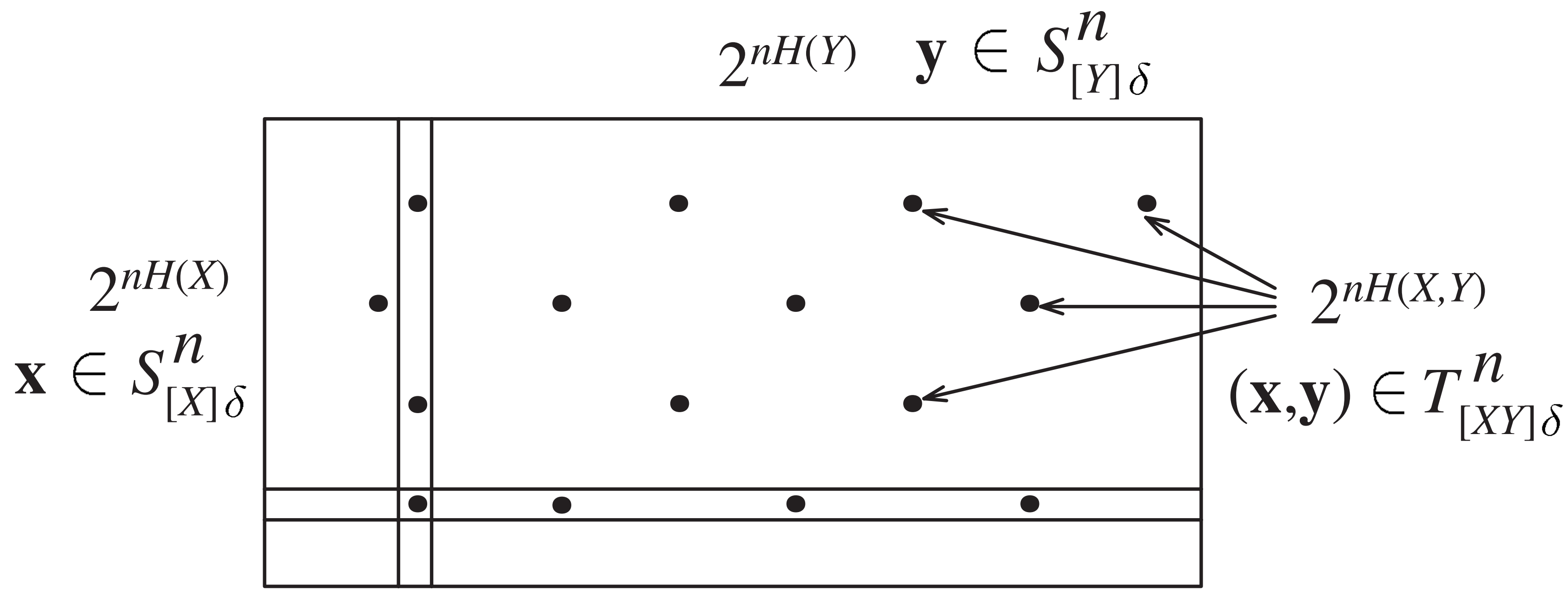
$$\begin{aligned}
&\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x})p(\mathbf{y}) \\
&\leq 2^{n(H(X,Y)+\xi)} \cdot 2^{-n(H(X)-\eta)} \cdot 2^{-n(H(Y)-\zeta)} \\
&= 2^{-n(H(X)+H(Y)-H(X,Y)-\xi-\eta-\zeta)} \\
&= 2^{-n(I(X;Y)\underline{-\xi-\eta-\zeta})} \\
&= 2^{-n(I(X;Y)\underline{-\tau})},
\end{aligned}$$

where

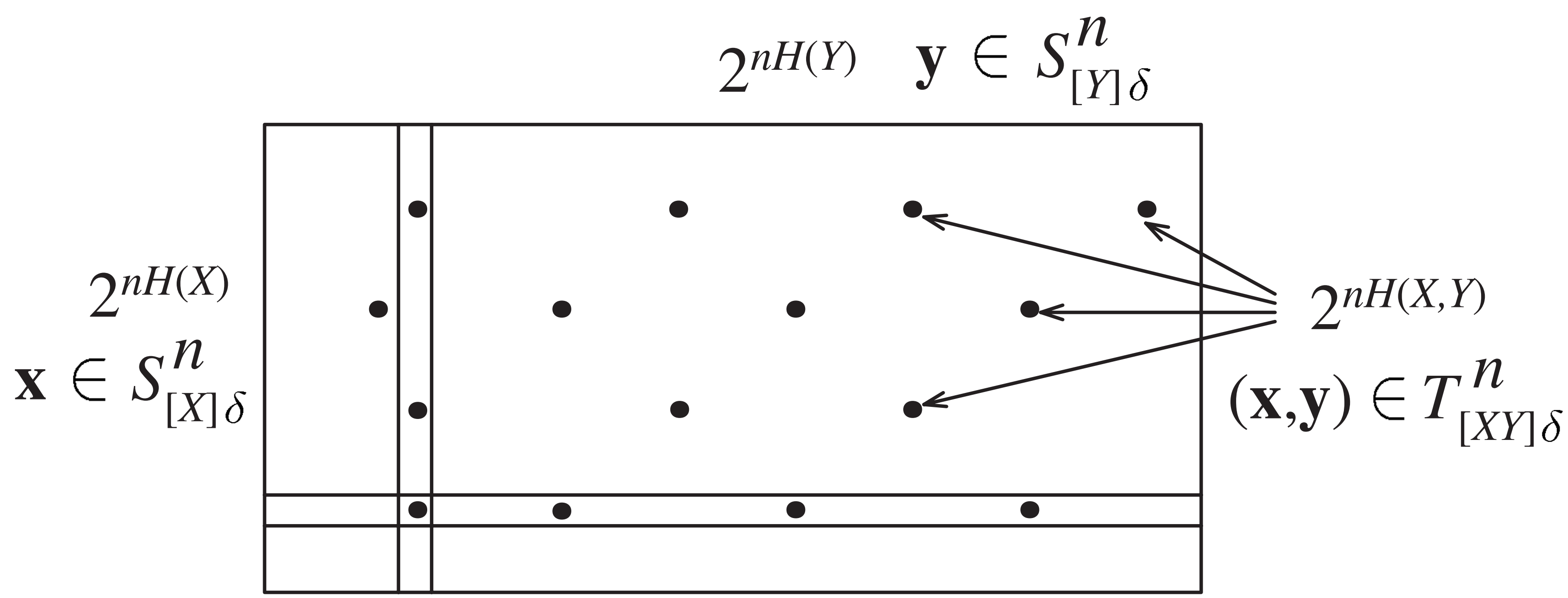
$$\tau = \xi + \eta + \zeta \to 0$$

as $\delta \to 0$. The lemma is proved.

An Interpretation of Lemma 7.17

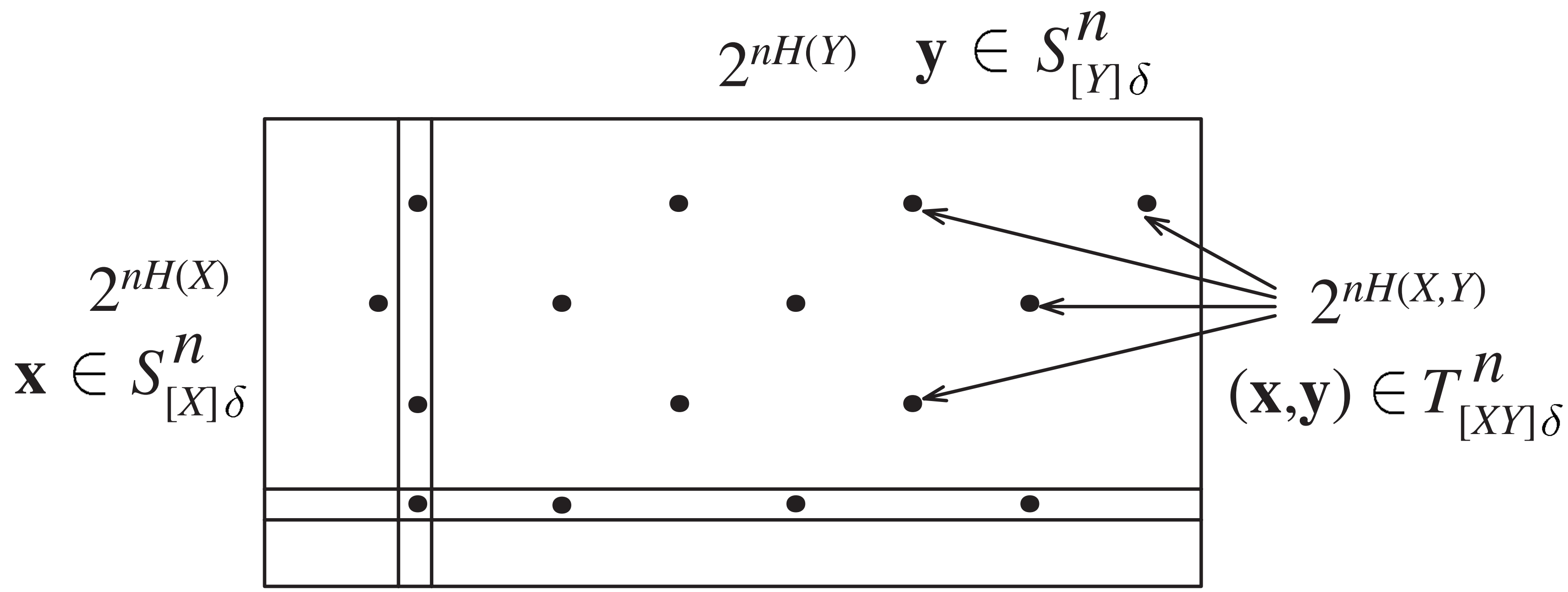


An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.

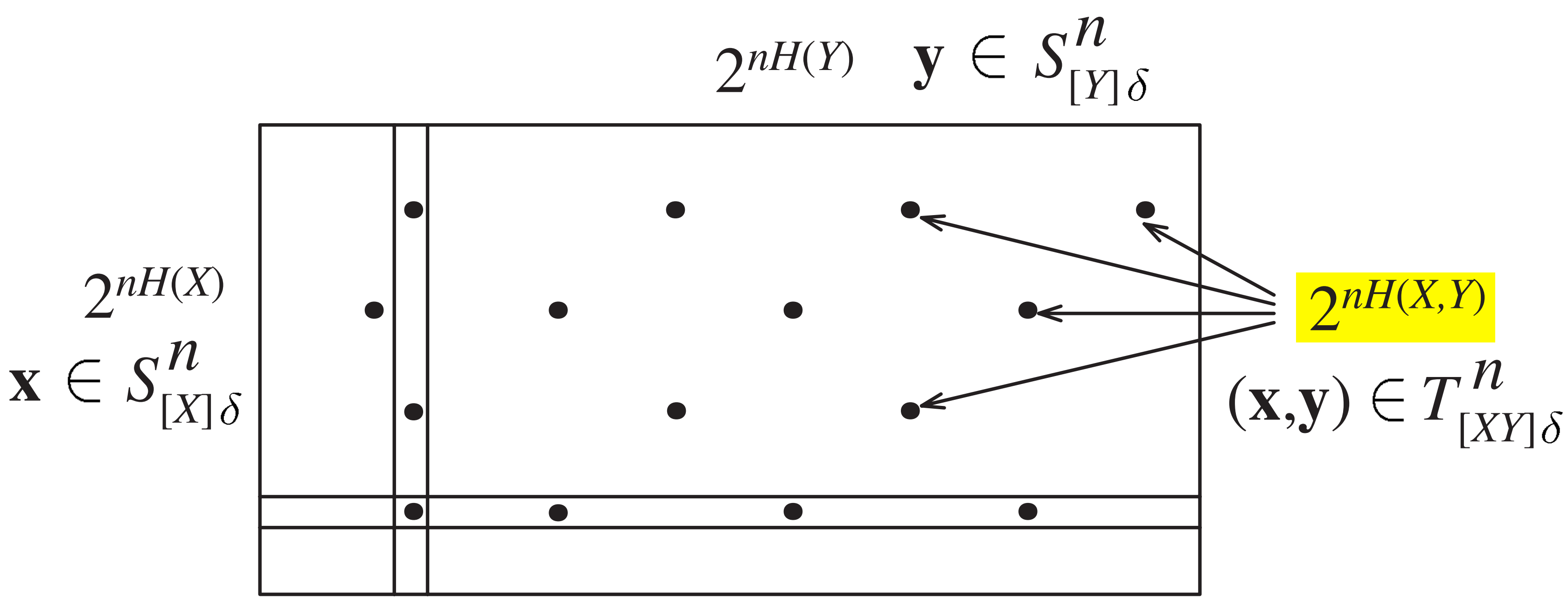
An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.
- Then

$$\Pr\{\text{obtaining a jointly typical pair}\} \approx \frac{2^{nH(X,Y)}}{2^{nH(X)}2^{nH(Y)}} = 2^{-nI(X;Y)}$$

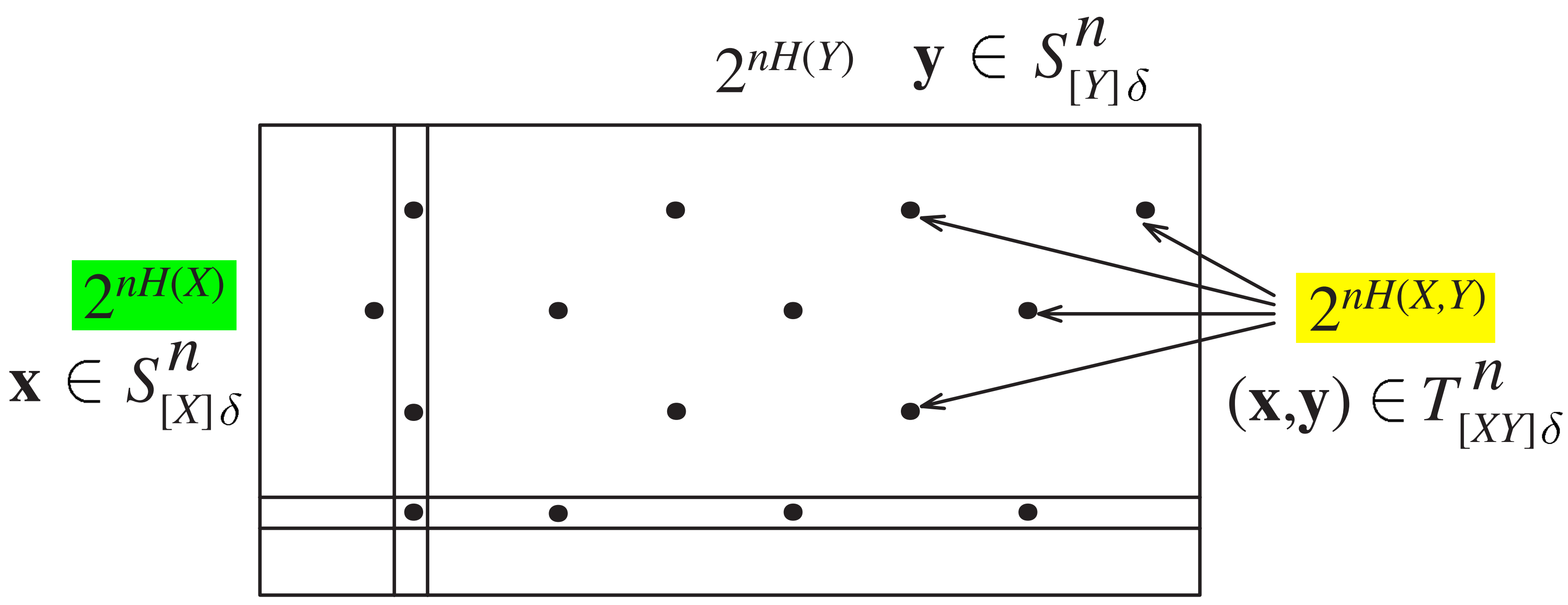
An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.
- Then

$$\Pr\{\text{obtaining a jointly typical pair}\} \approx \frac{2^{nH(X,Y)}}{2^{nH(X)}2^{nH(Y)}} = 2^{-nI(X;Y)}$$

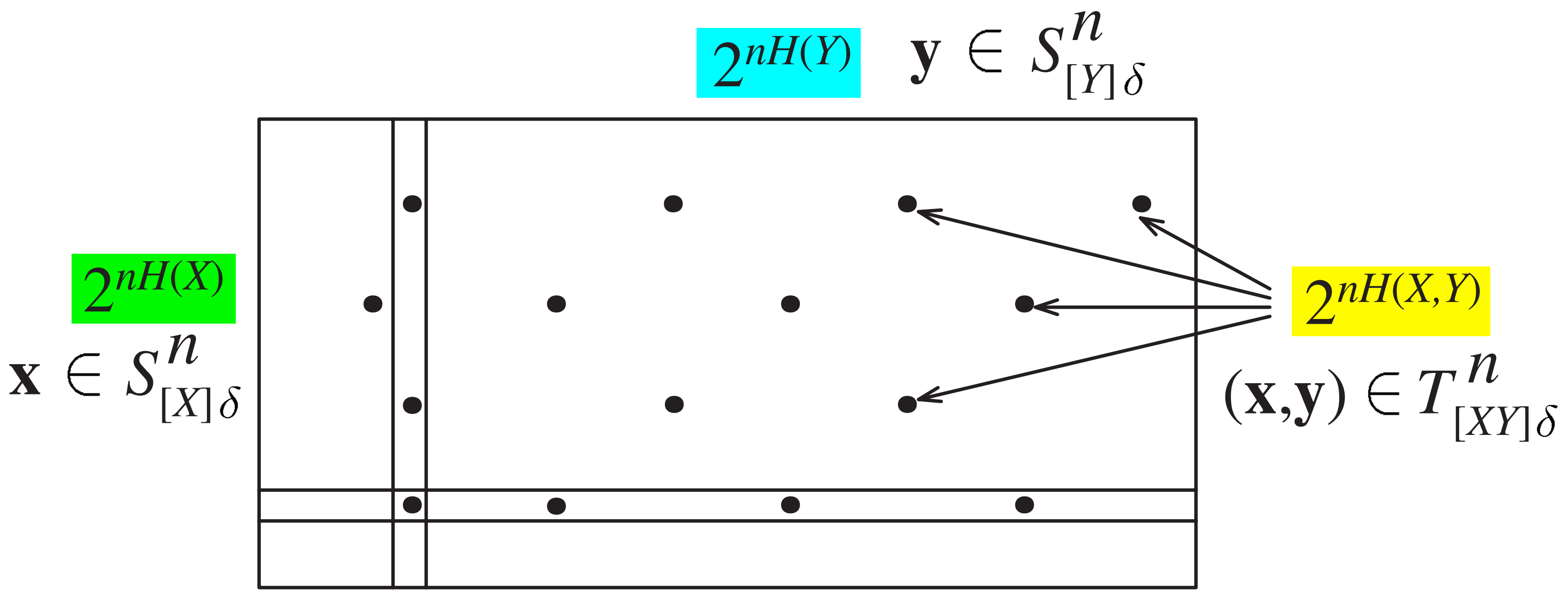
An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.
- Then

$$\Pr\{\text{obtaining a jointly typical pair}\} \approx \frac{2^{nH(X,Y)}}{2^{nH(X)}2^{nH(Y)}} = 2^{-nI(X;Y)}$$

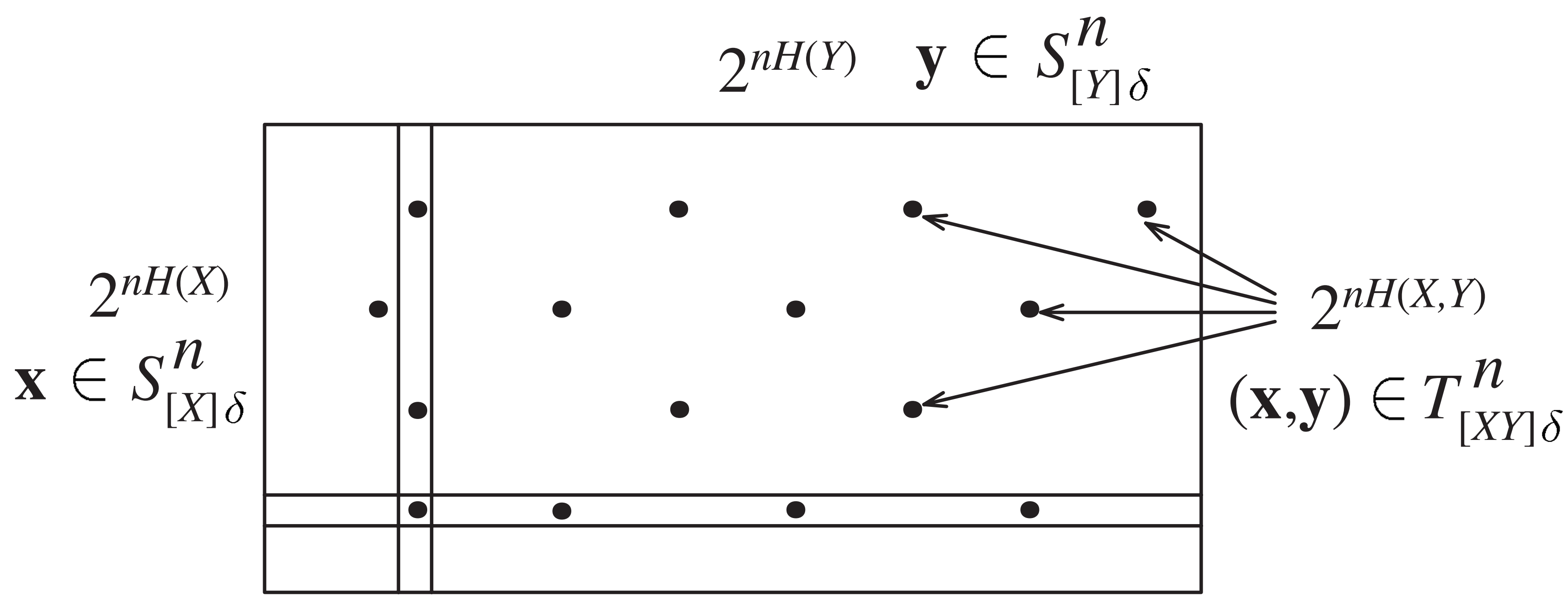
An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.
- Then

$$\Pr\{\text{obtaining a jointly typical pair}\} \approx \frac{2^{nH(X,Y)}}{2^{nH(X)}2^{nH(Y)}} = 2^{-nI(X;Y)}$$

An Interpretation of Lemma 7.17



- Randomly choose a row with uniform distribution and randomly choose a column with uniform distribution.
- Then

$$\Pr\{\text{obtaining a jointly typical pair}\} \approx \frac{2^{nH(X,Y)}}{2^{nH(X)}2^{nH(Y)}} = 2^{-nI(X;Y)}$$

Random Coding Scheme

Random Coding Scheme

Parameter Settings

Random Coding Scheme

Parameter Settings

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

Random Coding Scheme

Parameter Settings

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

2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

Random Coding Scheme

Parameter Settings

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

2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

Random Coding Scheme

Parameter Settings

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

2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code by generating M codewords in $\mathcal{X}^n$ independently and identically according to $p(x)^n$. 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 $p(x)$. Let δ to be specified later.

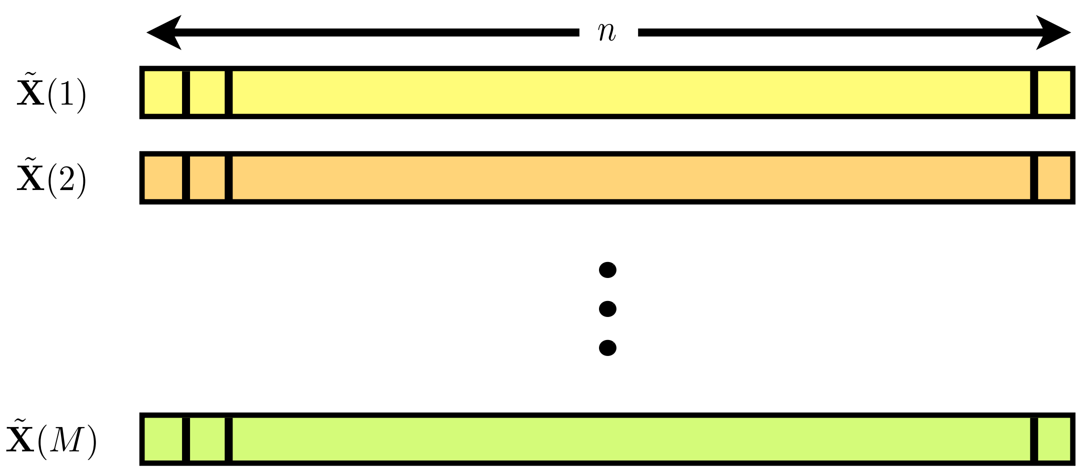
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code by generating M codewords in $\mathcal{X}^n$ independently and identically according to $p(x)^n$. 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 $p(x)$. Let δ to be specified later.

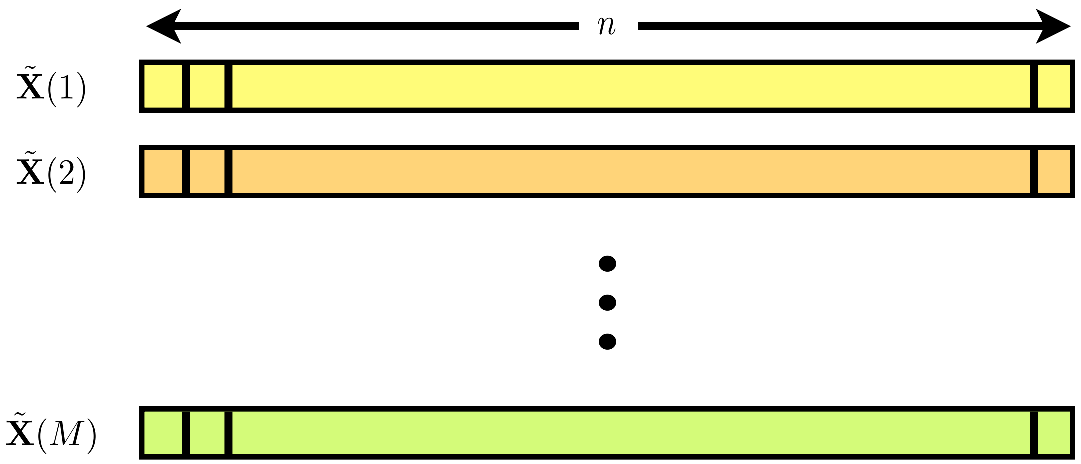
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

1. Construct the codebook $\mathcal{C}$ of an (n, M) code by generating M codewords in $\mathcal{X}^n$ independently and identically according to $p(x)^n$. 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 $p(x)$. Let δ to be specified later.

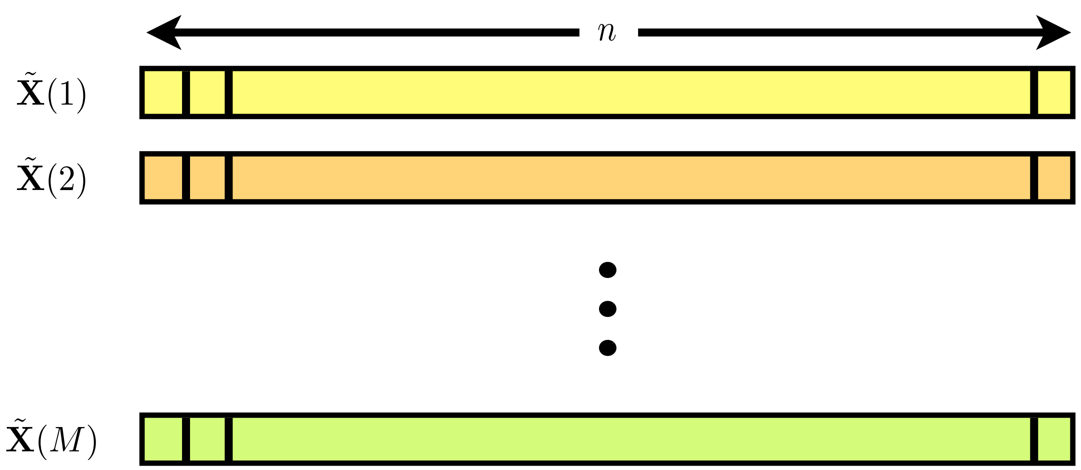
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.

Random Coding Scheme

Parameter Settings

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

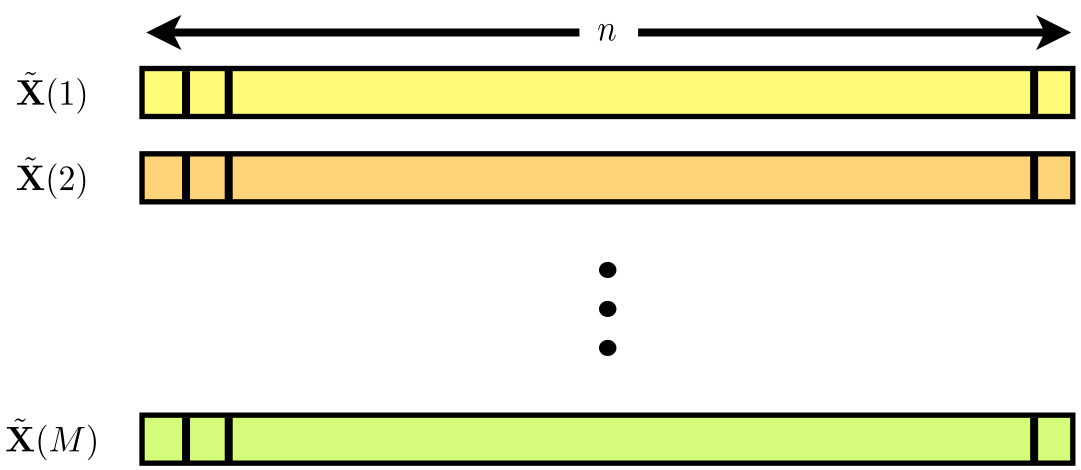
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.

Random Coding Scheme

Parameter Settings

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

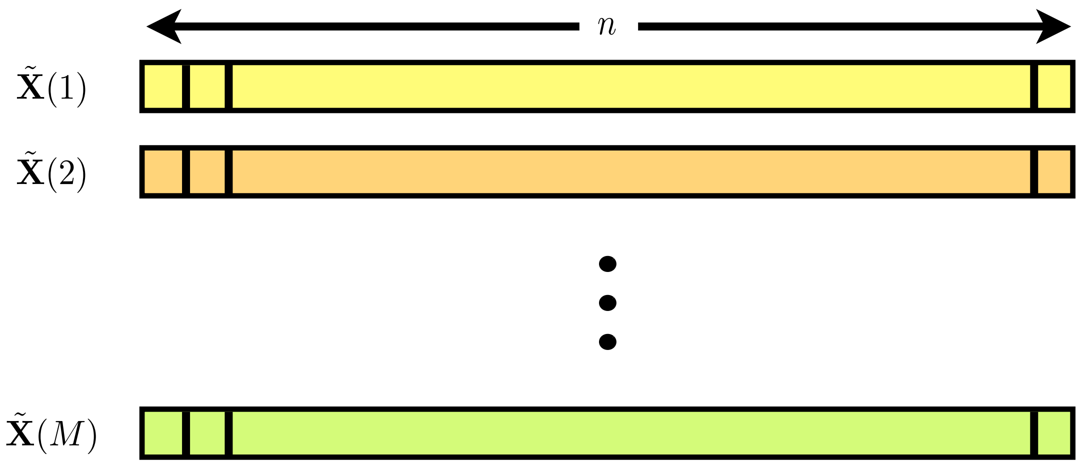
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.

Random Coding Scheme

Parameter Settings

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

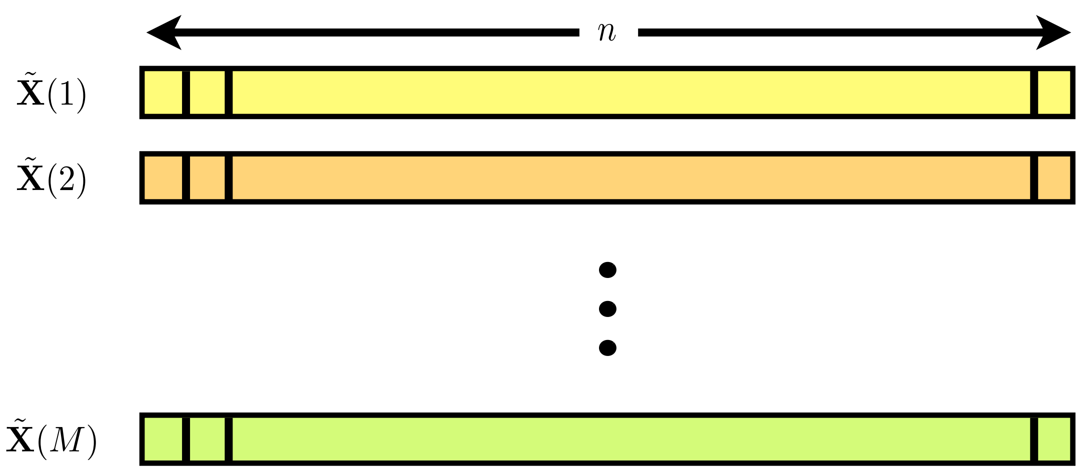
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.

Random Coding Scheme

Parameter Settings

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

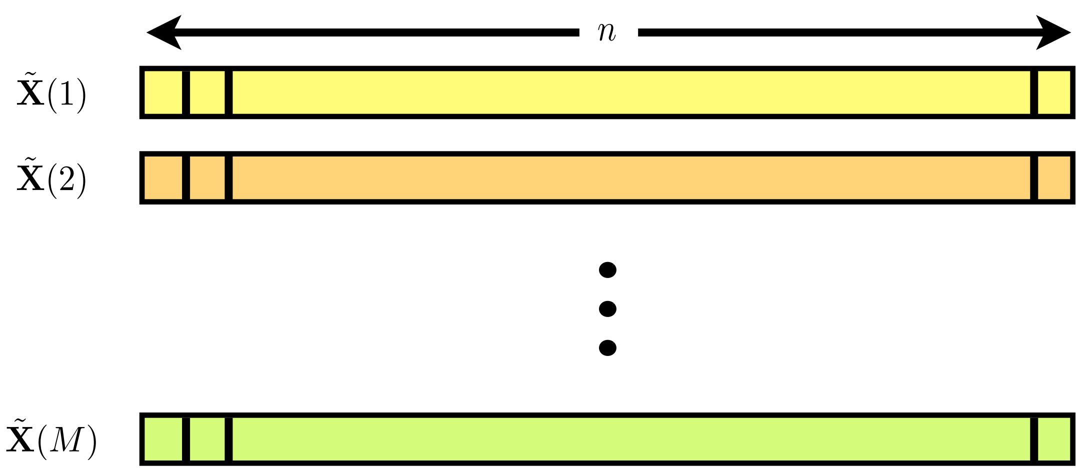
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

Random Coding Scheme

Parameter Settings

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

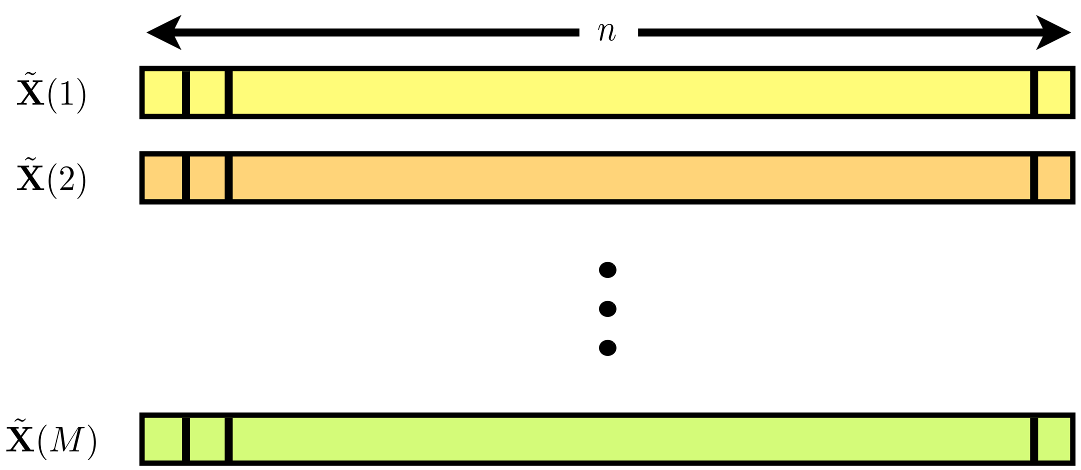
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

Random Coding Scheme

Parameter Settings

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

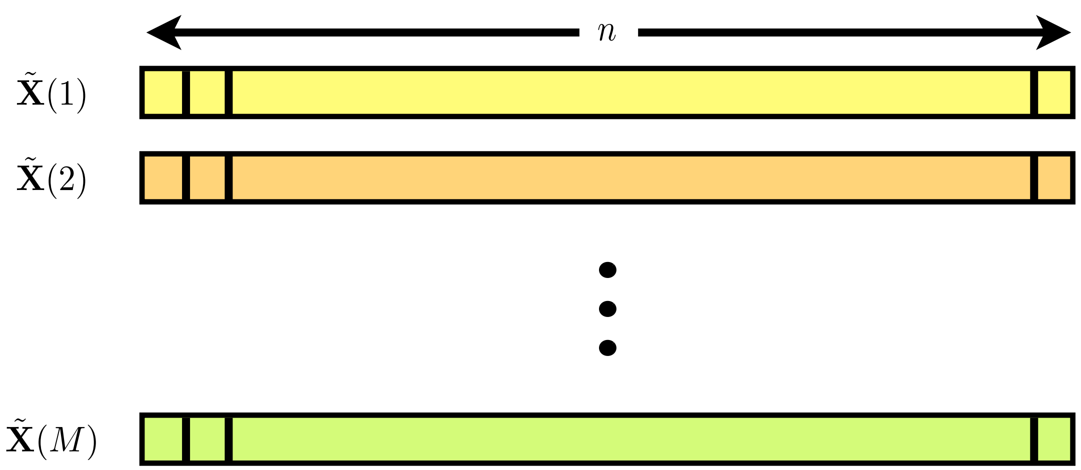
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

Random Coding Scheme

Parameter Settings

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

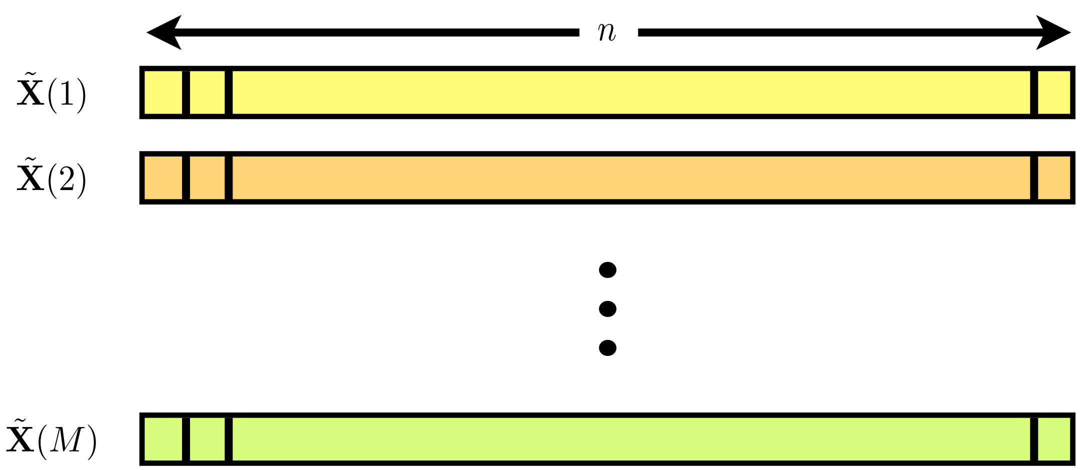
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

Random Coding Scheme

Parameter Settings

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

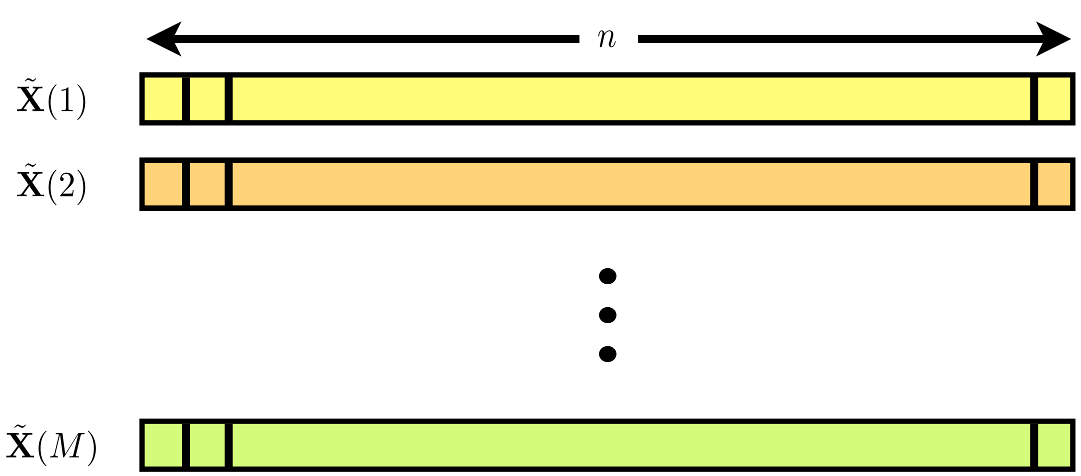
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

Random Coding Scheme

Parameter Settings

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

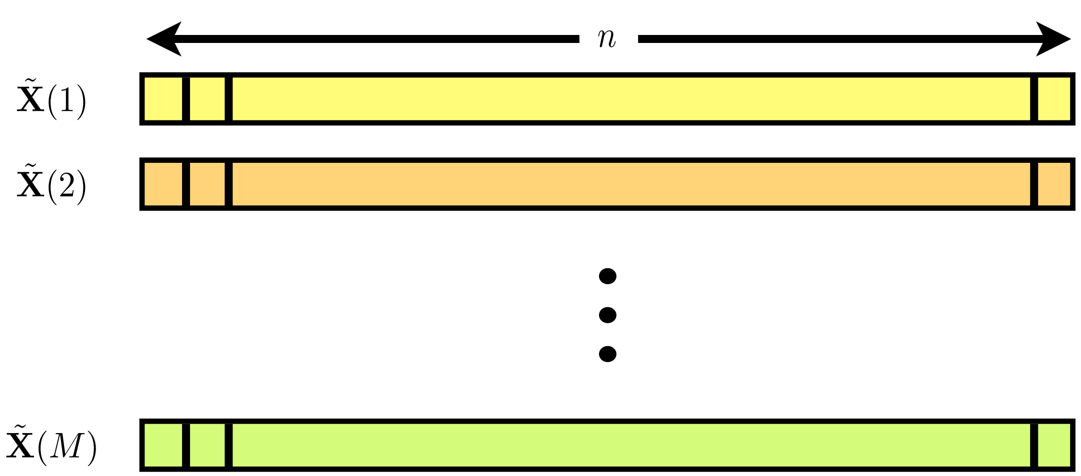
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

Random Coding Scheme

Parameter Settings

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

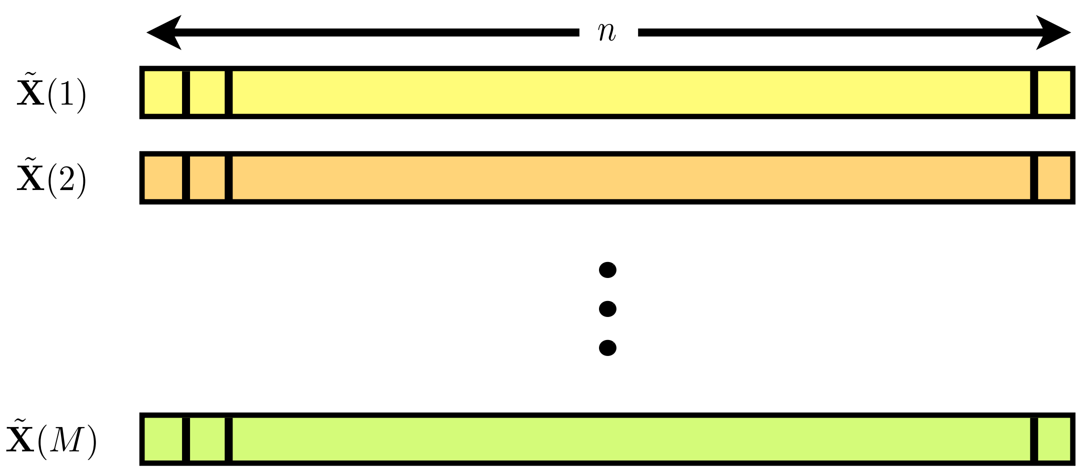
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

4. Transmit $\mathbf{X} = \tilde{\mathbf{X}}(W)$ through the channel.

Random Coding Scheme

Parameter Settings

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

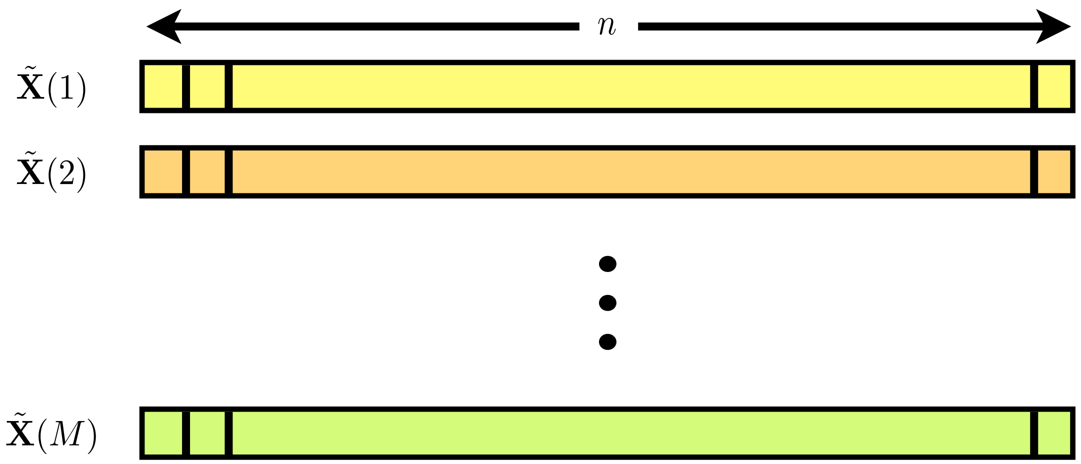
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



• Generate each component according to $p(x)$.

• There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.

• Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

4. Transmit $\mathbf{X} = \tilde{\mathbf{X}}(W)$ through the channel.

Random Coding Scheme

Parameter Settings

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

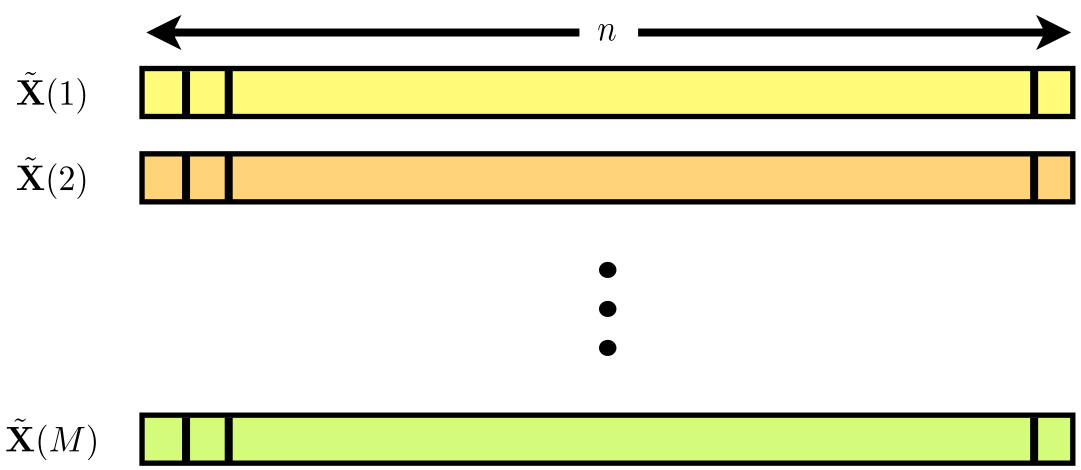
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



• Generate each component according to $p(x)$.

• There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.

• Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

4. Transmit $\mathbf{X} = \tilde{\mathbf{X}}(W)$ through the channel.

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

$$\Pr\{\mathbf{Y} = \mathbf{y} | \tilde{\mathbf{X}}(W) = \mathbf{x}\} = \prod_{i=1}^{n} p(y_i|x_i).$$

Random Coding Scheme

Parameter Settings

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

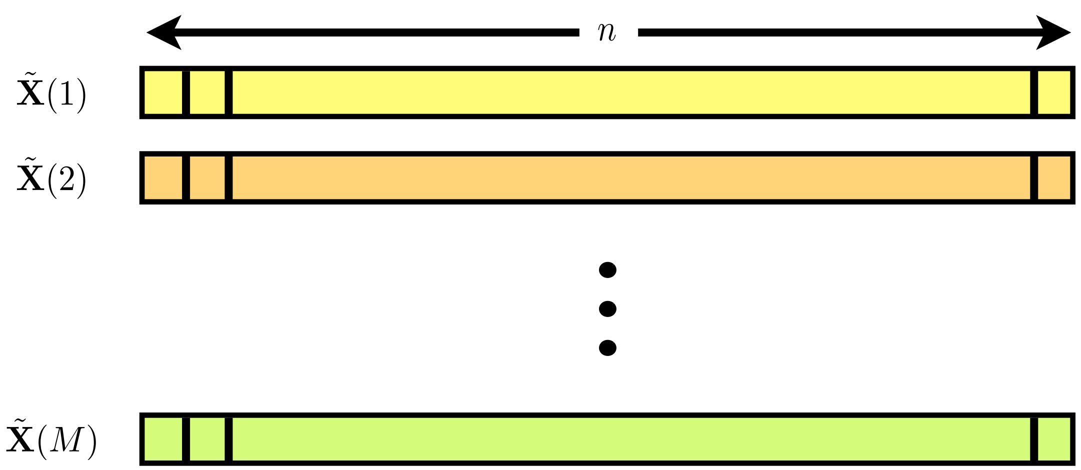
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

4. Transmit $\mathbf{X} = \tilde{\mathbf{X}}(W)$ through the channel.

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

$$\Pr\{\mathbf{Y} = \mathbf{y} | \tilde{\mathbf{X}}(W) = \mathbf{x}\} = \prod_{i=1}^{n} p(y_i|x_i).$$

Random Coding Scheme

Parameter Settings

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

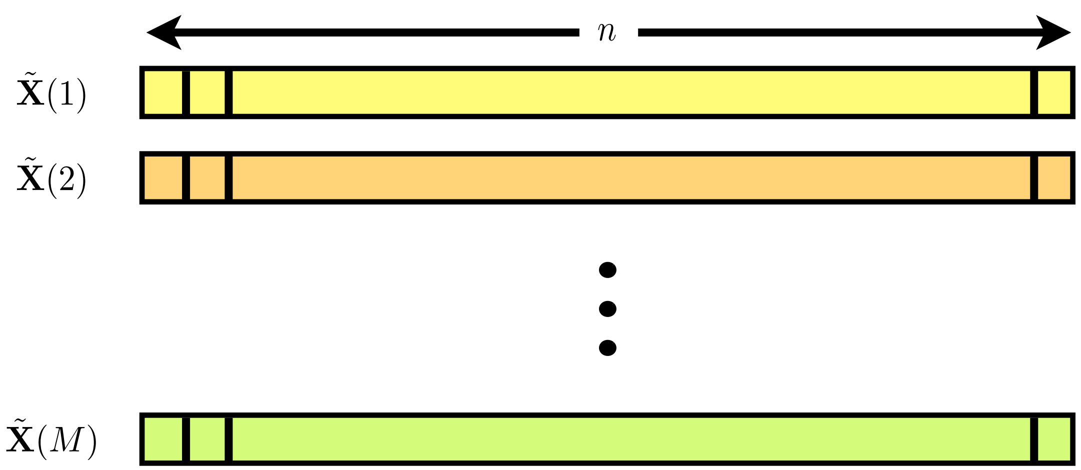
2. Let M be an even integer satisfying

$$I(X;Y) - \frac{\epsilon}{2} < \frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4},$$

where n is sufficiently large, i.e., $M \approx 2^{nI(X;Y)}$.

The Random Coding Scheme

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



- Generate each component according to $p(x)$.
- There are a total of $|\mathcal{X}|^{Mn}$ possible codebooks that can be constructed.
- Regard two codebooks whose sets of codewords are permutations of each other as two different codebooks.

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

3. A message W is chosen from $\mathcal{W}$ according to the uniform distribution.

4. Transmit $\mathbf{X} = \tilde{\mathbf{X}}(W)$ through the channel.

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

$$\Pr\{\mathbf{Y} = \mathbf{y} | \tilde{\mathbf{X}}(W) = \mathbf{x}\} = \prod_{i=1}^{n} p(y_i|x_i).$$

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

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

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

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\Pr\{Err\} \quad = \quad \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\}$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\Pr\{Err\} = \sum_{w=1}^{M} \underline{\Pr\{Err|W = w\}}\Pr\{W = w\}$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$
\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \underline{\Pr\{Err|W = w\}}\Pr\{W = w\} \\
&= \underline{\Pr\{Err|W = 1\}} \sum_{w=1}^{M} \Pr\{W = w\}
\end{aligned}
$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \end{aligned}$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$
\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\}
\end{aligned}
$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$
\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}
$$

Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

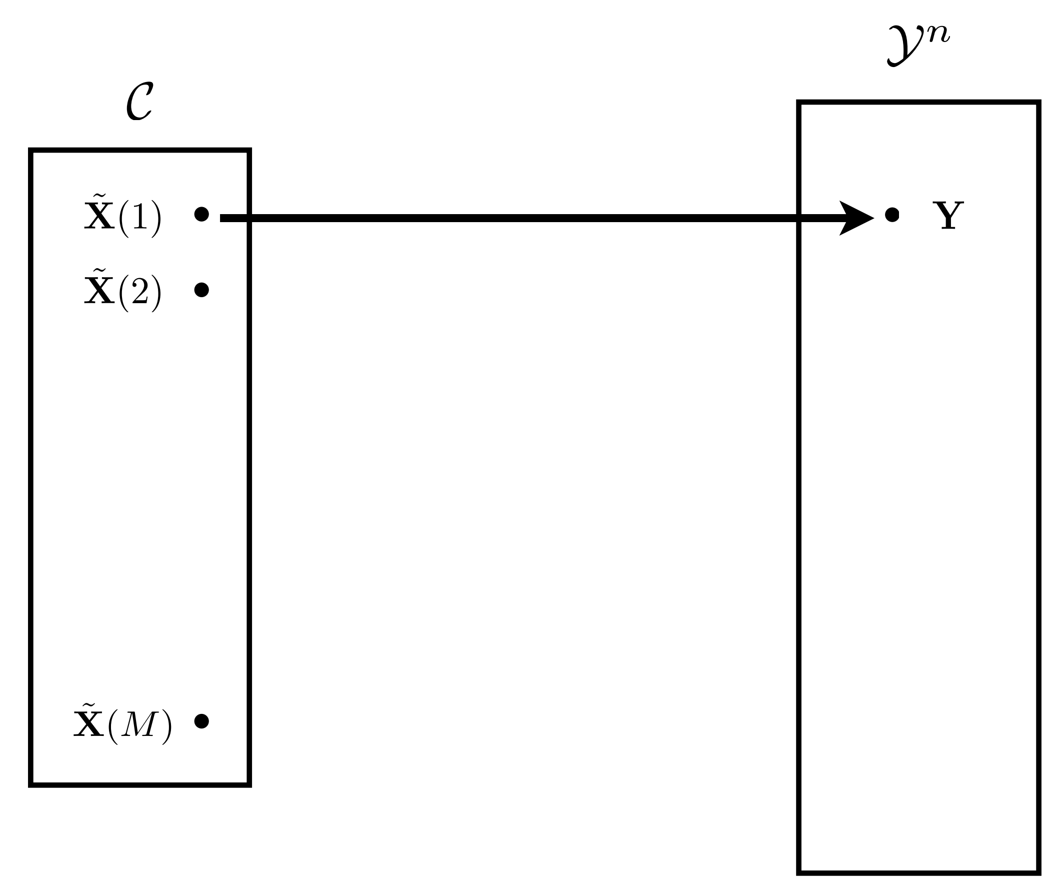
Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.



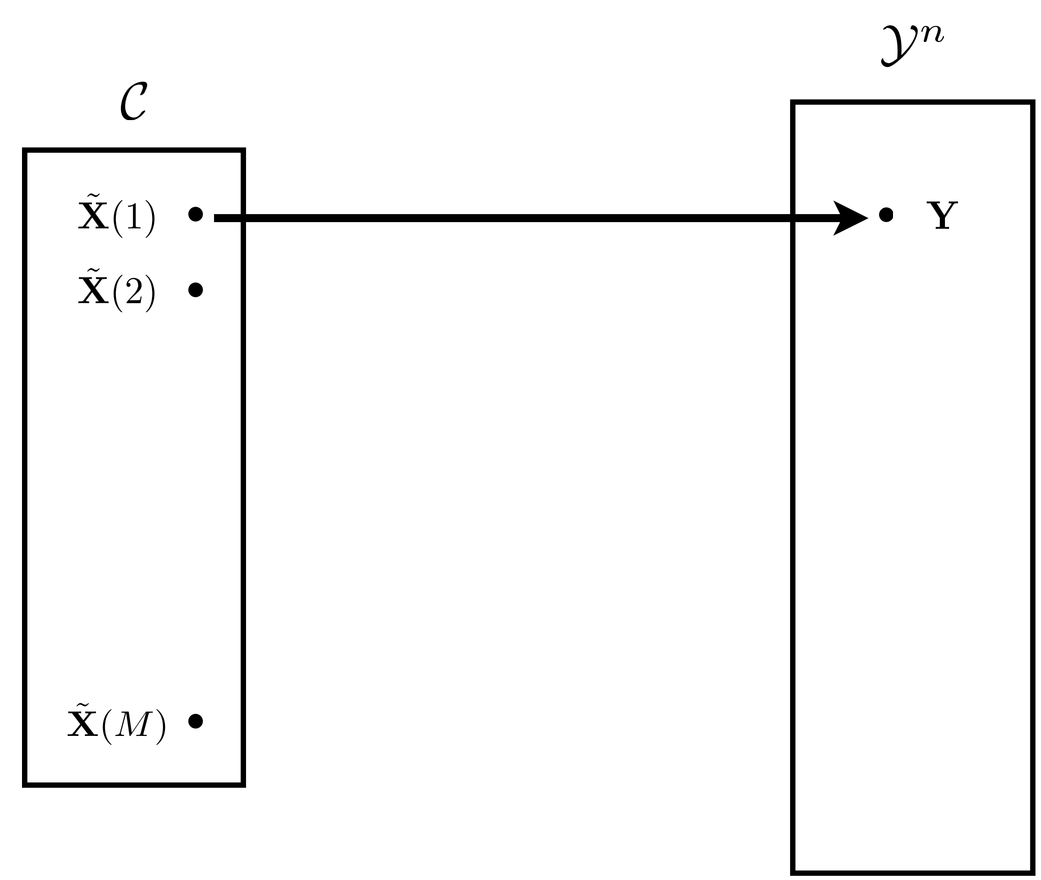
Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

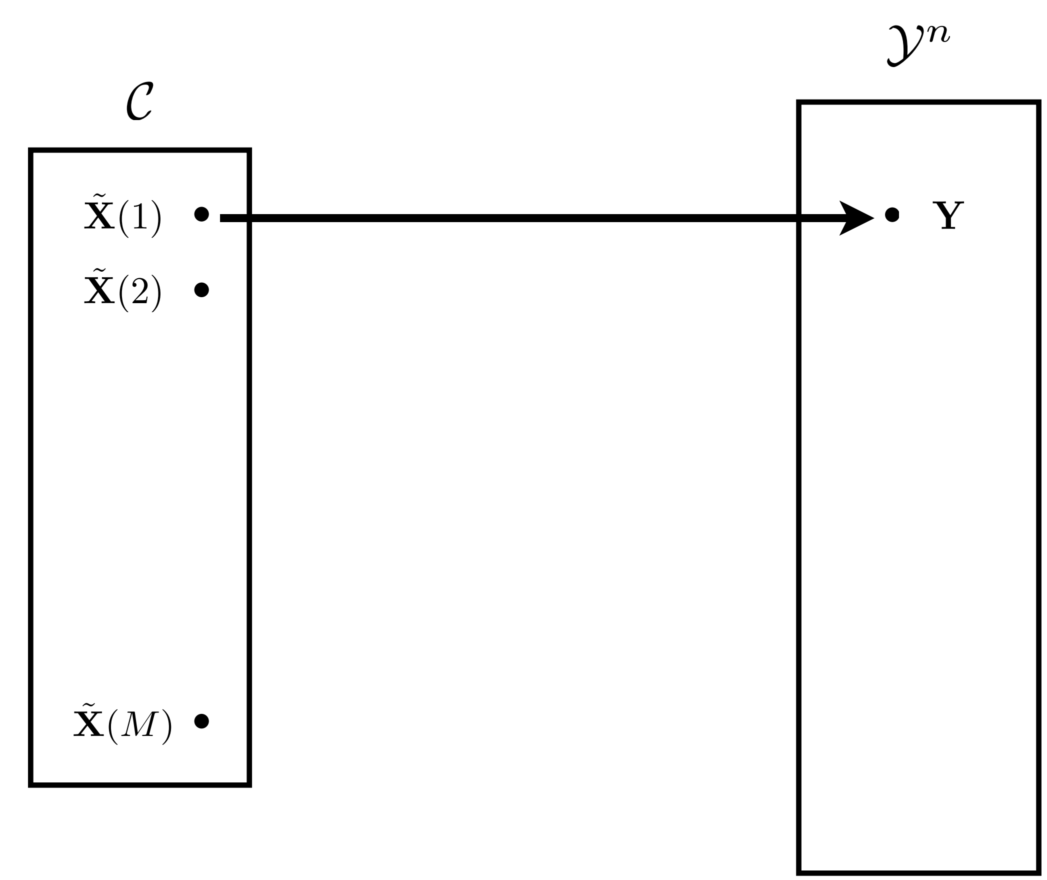
2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

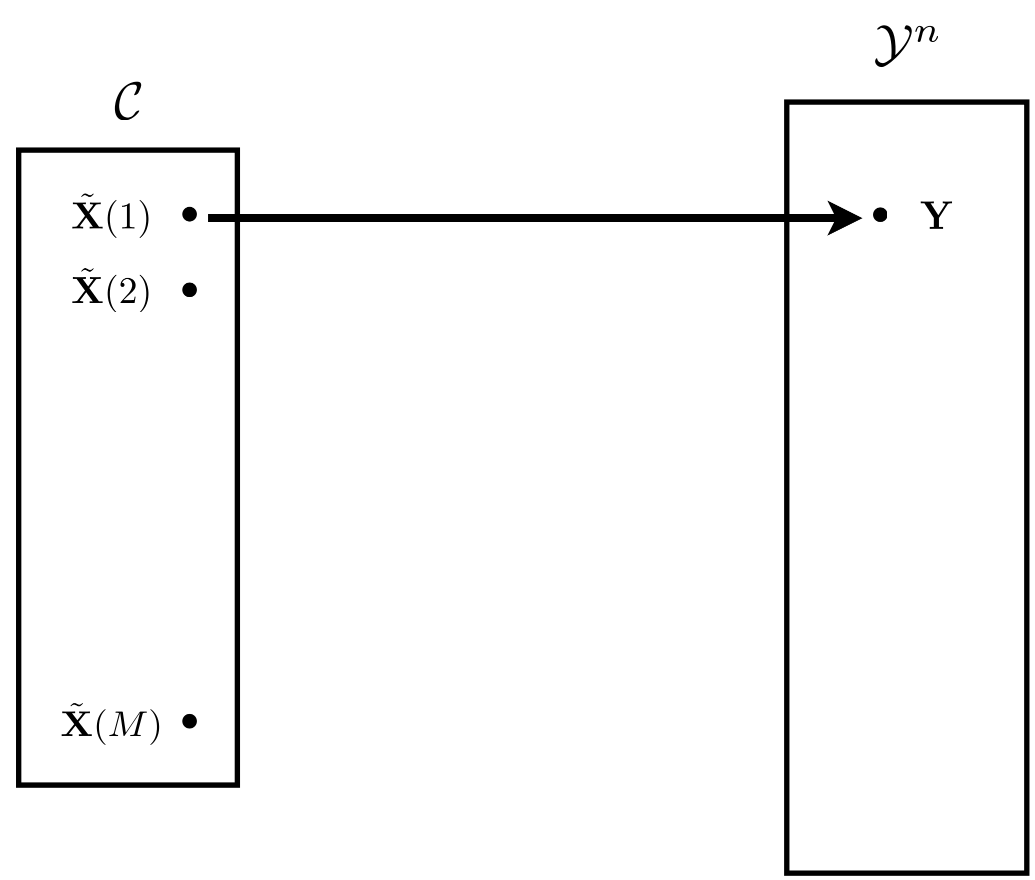
2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

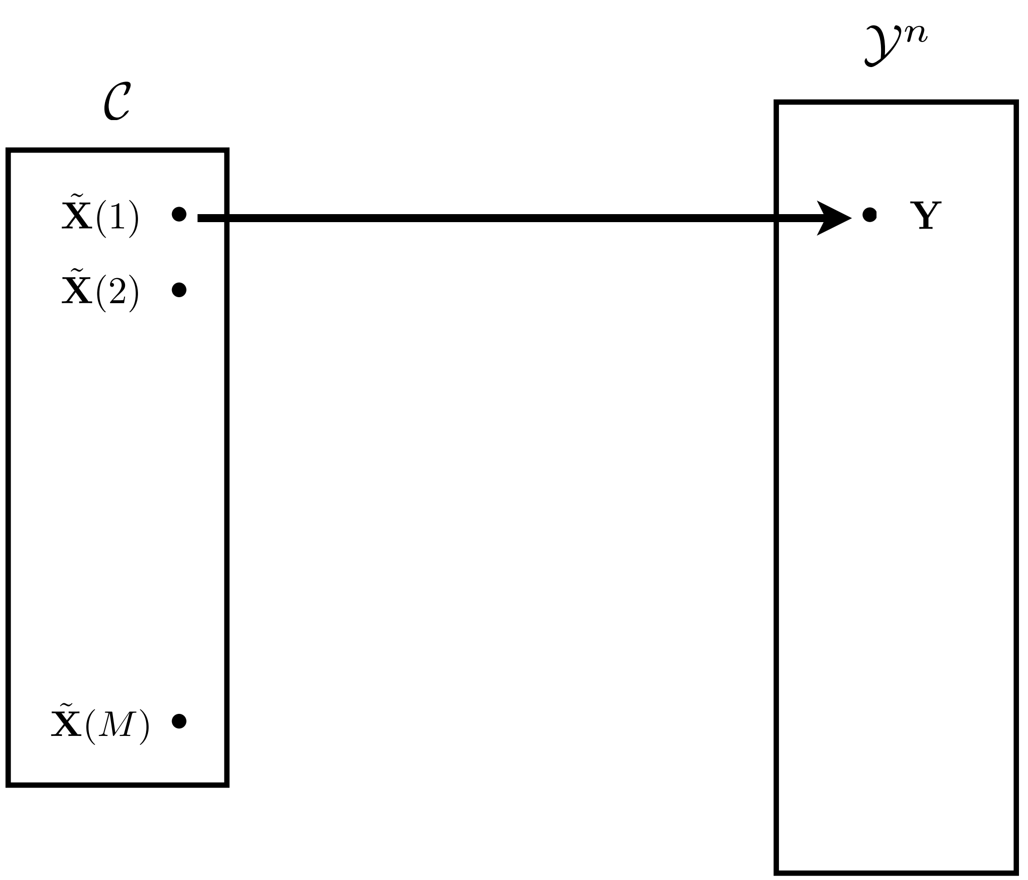
$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

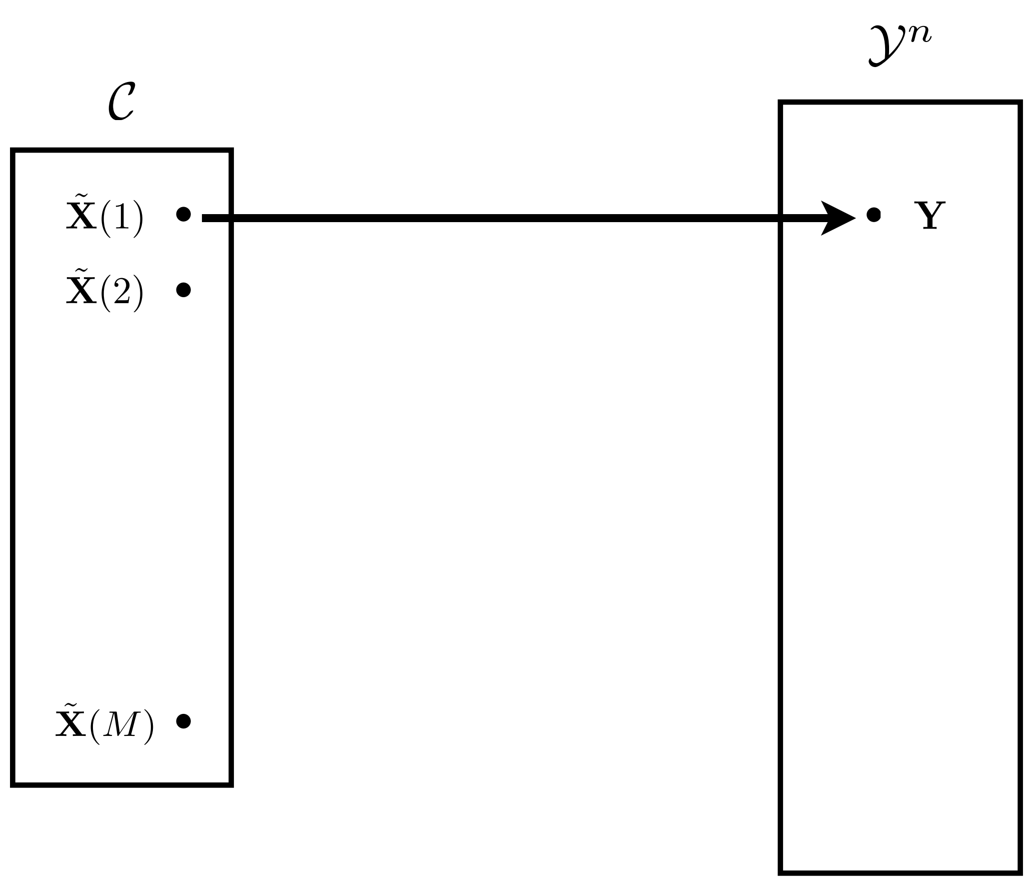
Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

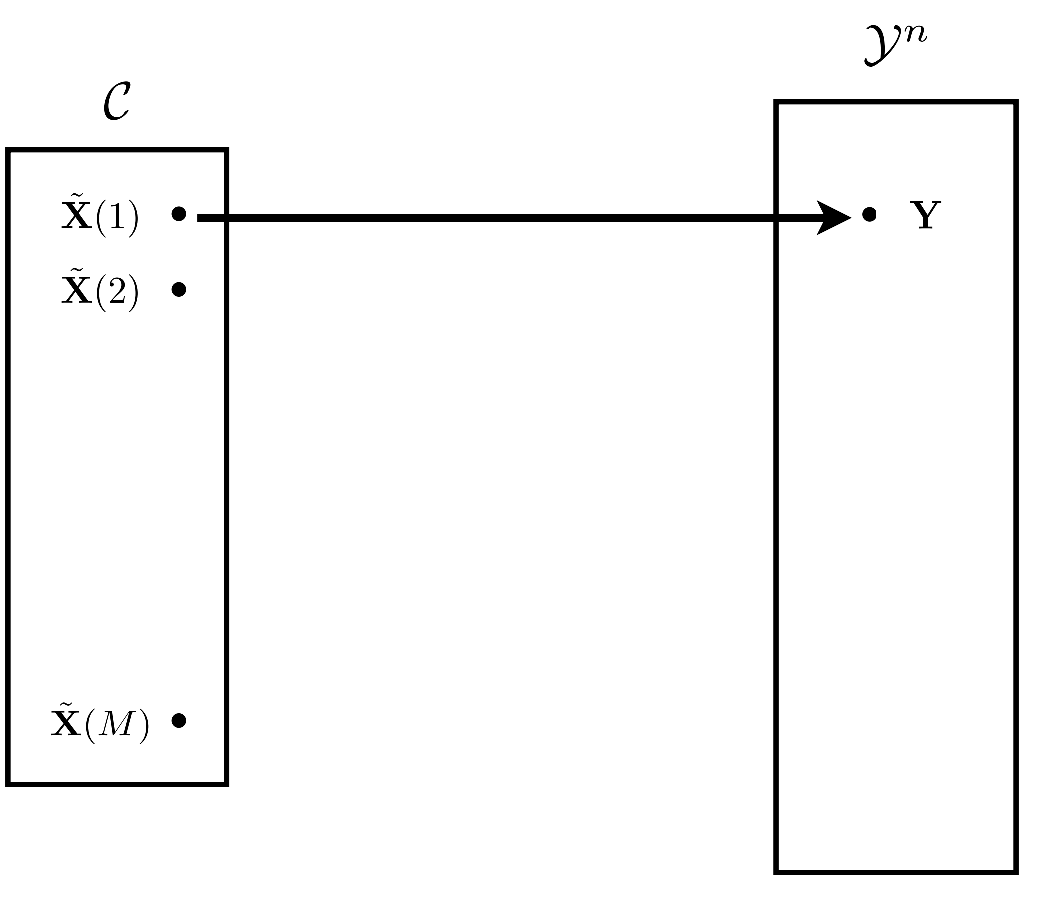
Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$
\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W=w\}\Pr\{W=w\} \\
&= \Pr\{Err|W=1\} \sum_{w=1}^{M} \Pr\{W=w\} \\
&= \Pr\{Err|W=1\}.
\end{aligned}
$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

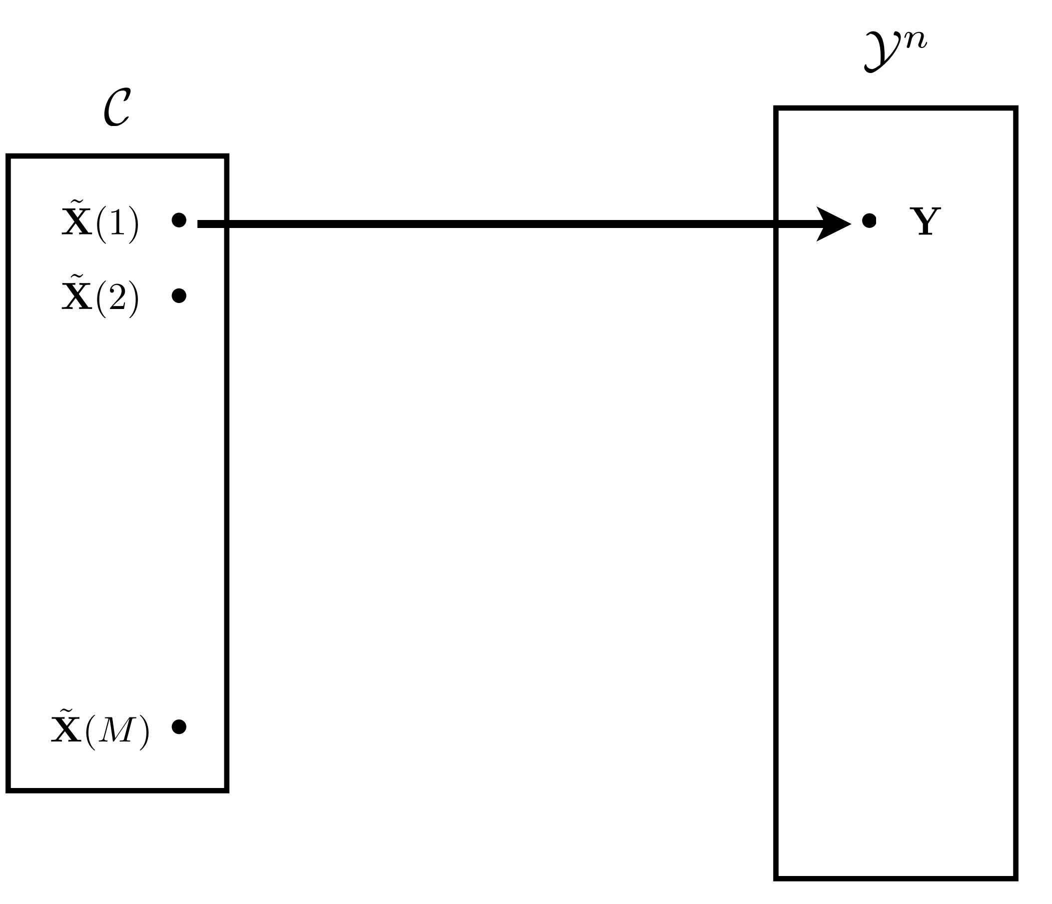
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W=1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\}.$$

5. Consider

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



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

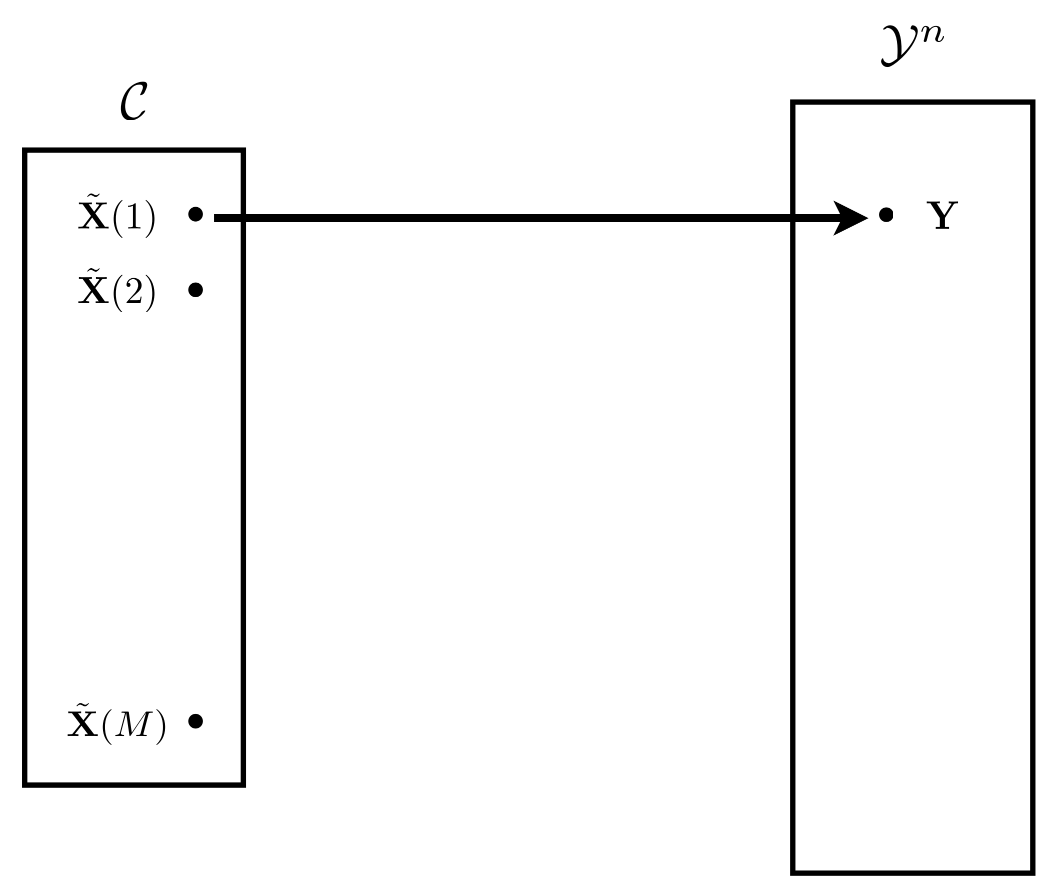
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}
&\Pr\{Err|W = 1\} \\
&\quad = 1 - \Pr\{Err^c|W = 1\}
\end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$
\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}
$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

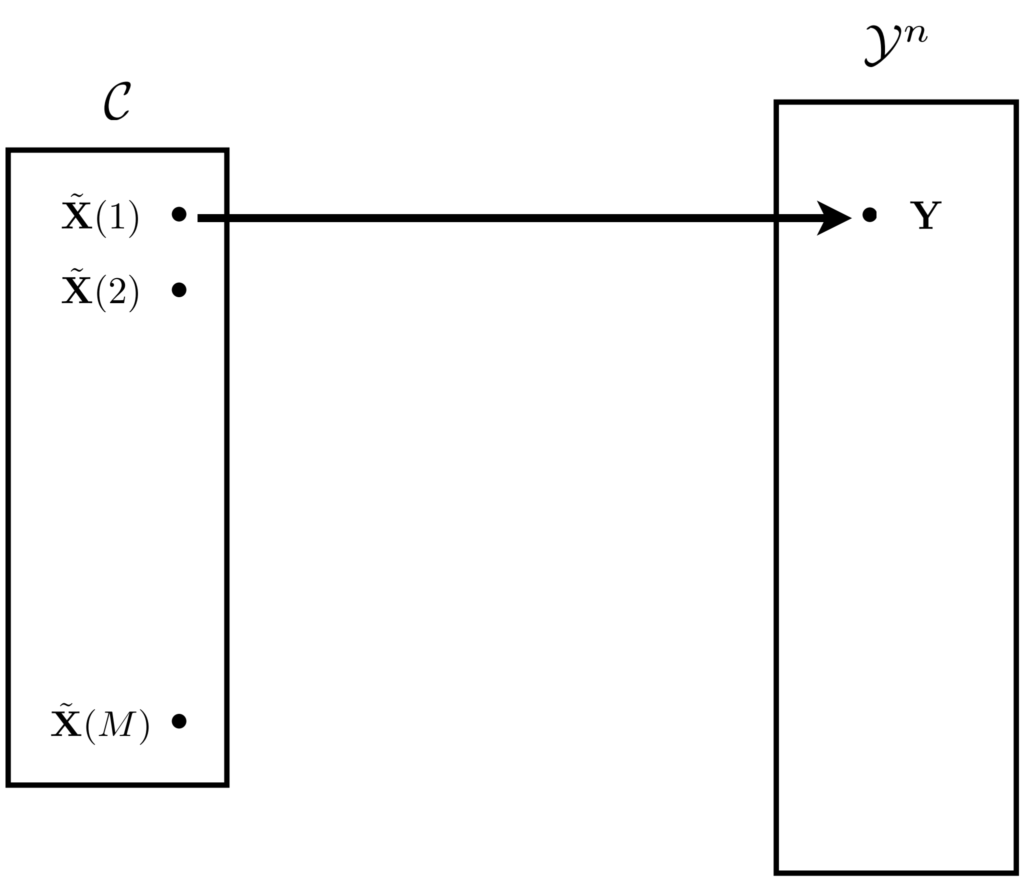
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$
\begin{aligned}
&\Pr\{Err|W = 1\} \\
&\quad = 1 - \Pr\{Err^c|W = 1\}
\end{aligned}
$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

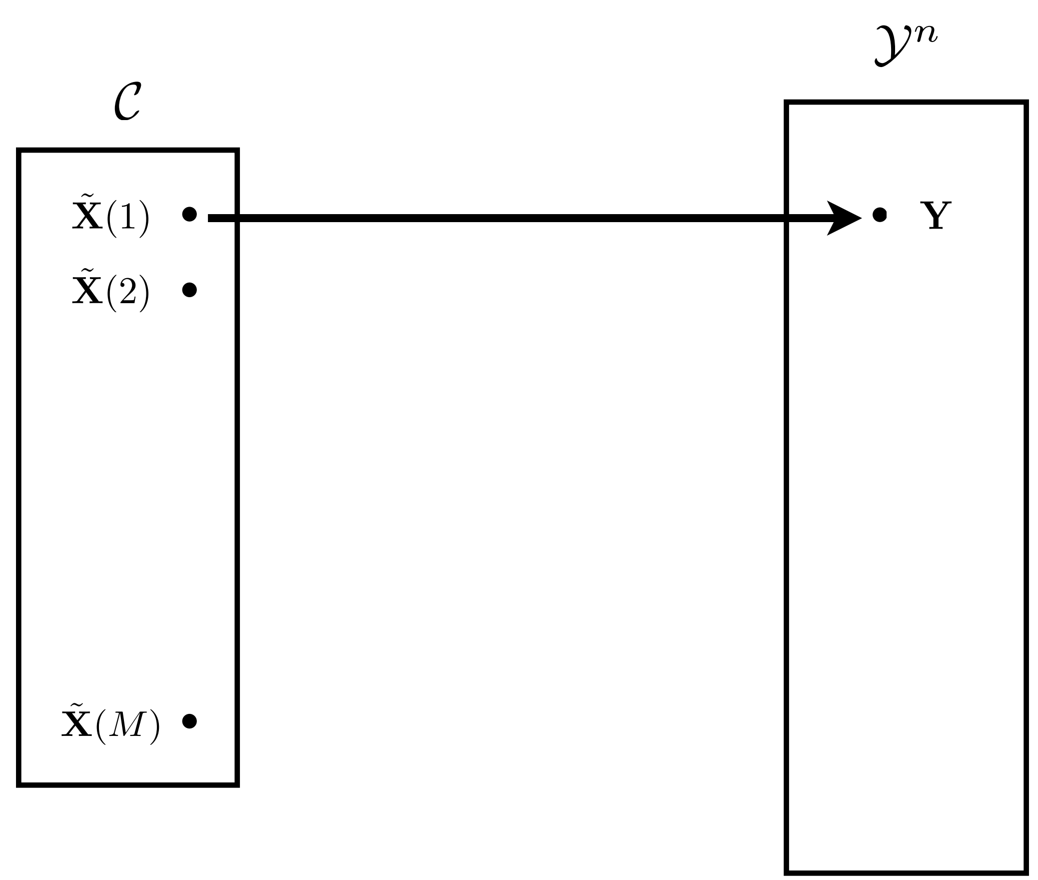
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned} &\Pr\{Err|W = 1\} \\ &\quad= 1 - \Pr\{Err^c|W = 1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W=w\}\Pr\{W=w\} \\ &= \Pr\{Err|W=1\} \sum_{w=1}^{M} \Pr\{W=w\} \\ &= \Pr\{Err|W=1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

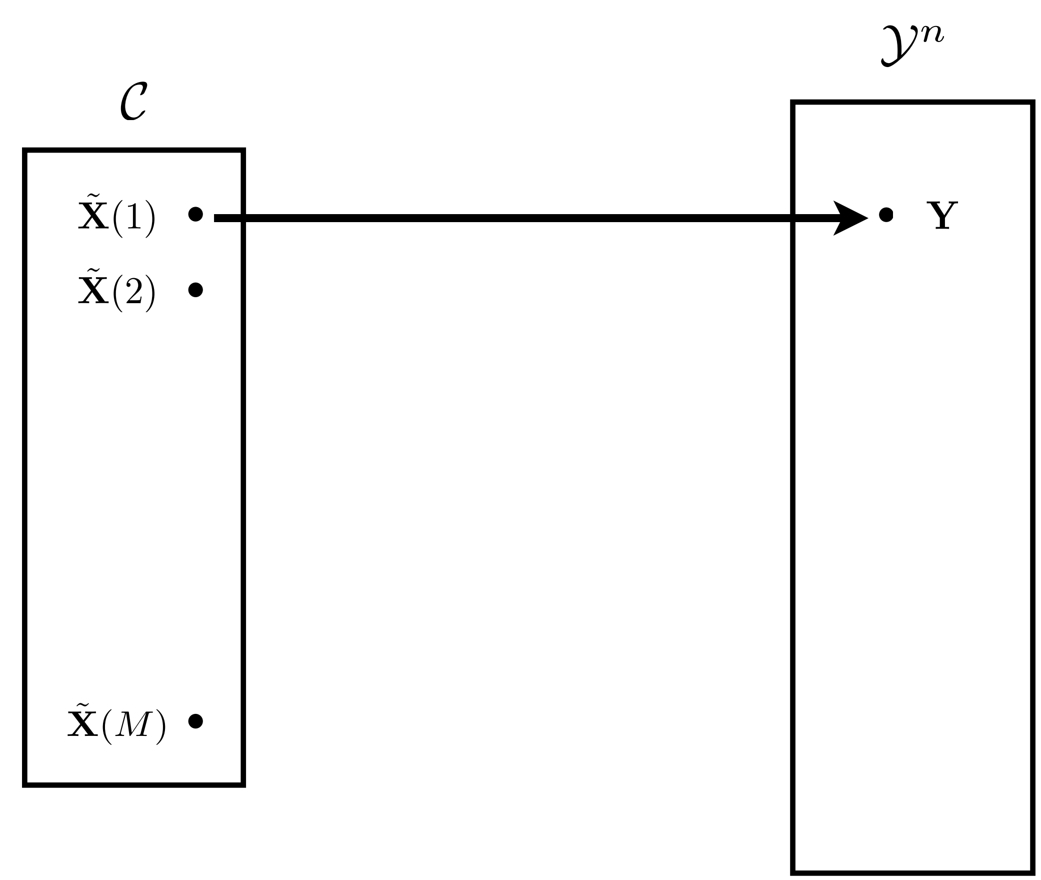
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W=1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\}.$$

5. Consider

$$\begin{aligned} &\Pr\{Err|W=1\} \\ &\quad= 1 - \Pr\{Err^c|W=1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\} \\ &\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W=1\} \end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W=w\}\Pr\{W=w\} \\ &= \Pr\{Err|W=1\} \sum_{w=1}^{M} \Pr\{W=w\} \\ &= \Pr\{Err|W=1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

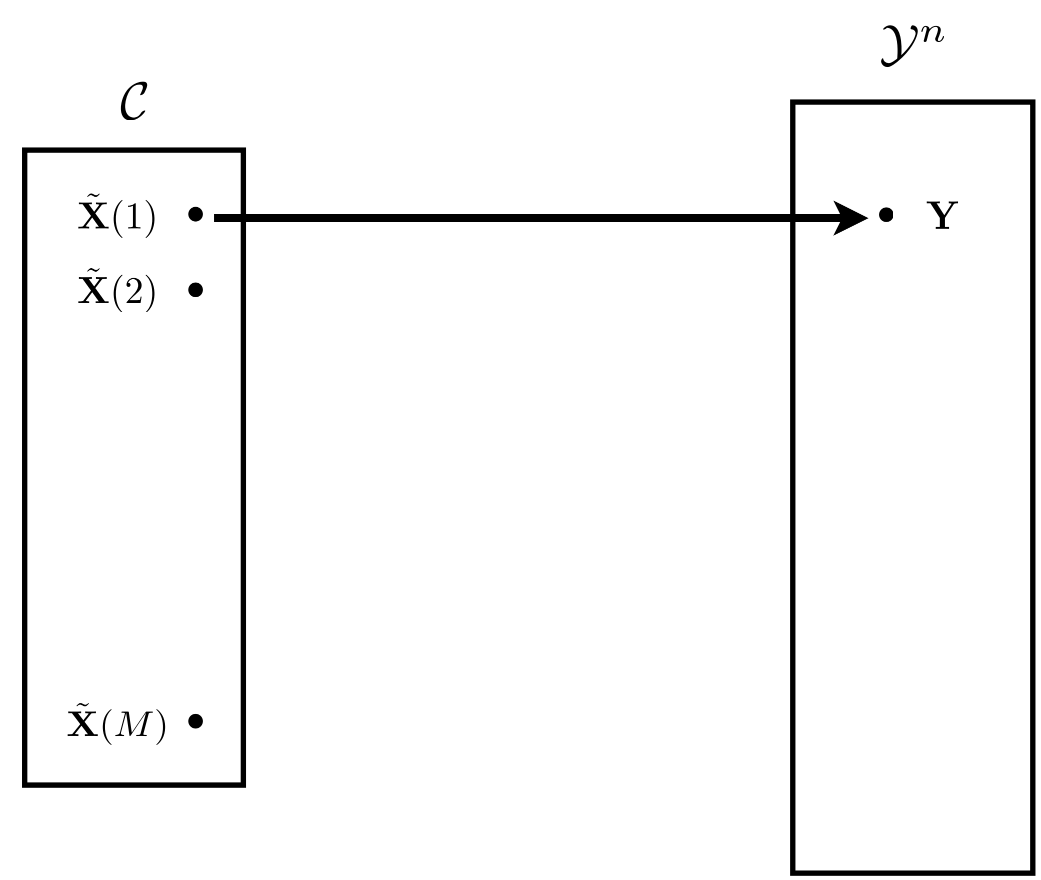
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W=1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W=1\} \\ &= 1 - \Pr\{Err^c|W=1\} \\ &\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\} \\ &= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W=1\}\end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}
\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\
&= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\
&= \Pr\{Err|W = 1\}.
\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

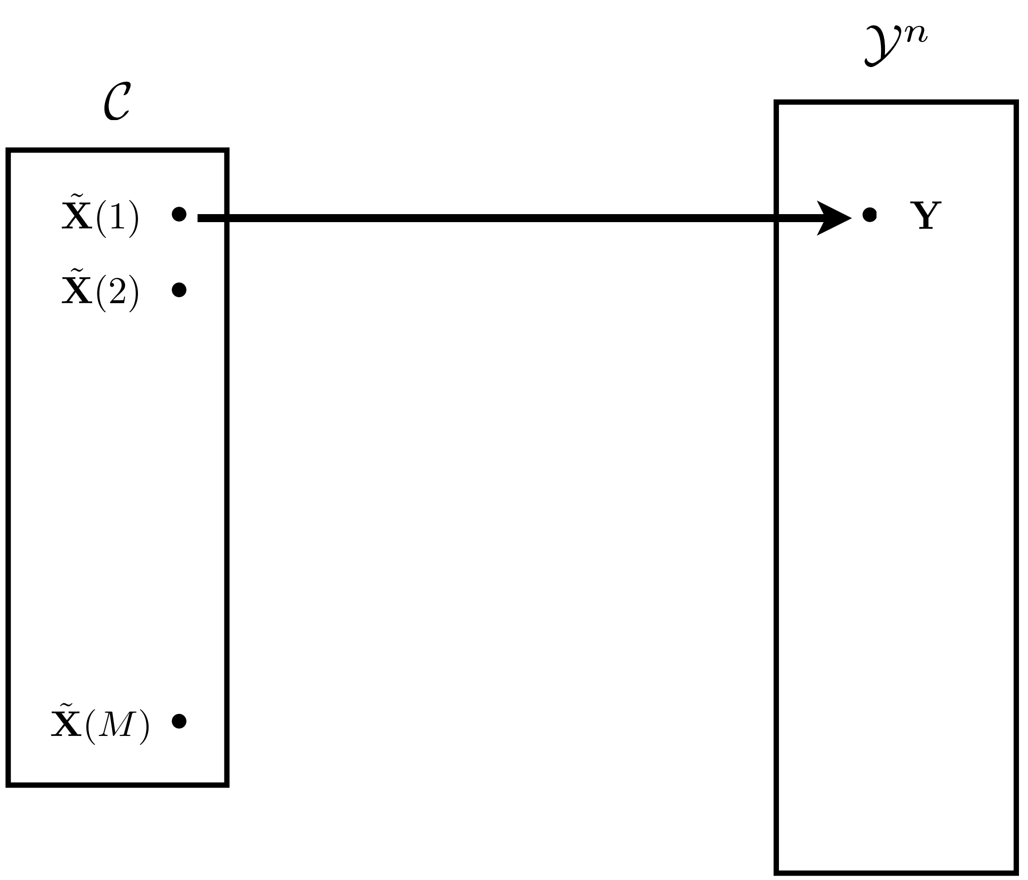
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}
&\Pr\{Err|W = 1\} \\
&\quad= 1 - \Pr\{Err^c|W = 1\} \\
&\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\
&\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\
&\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.
\end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W=w\}\Pr\{W=w\} \\ &= \Pr\{Err|W=1\} \sum_{w=1}^{M} \Pr\{W=w\} \\ &= \Pr\{Err|W=1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

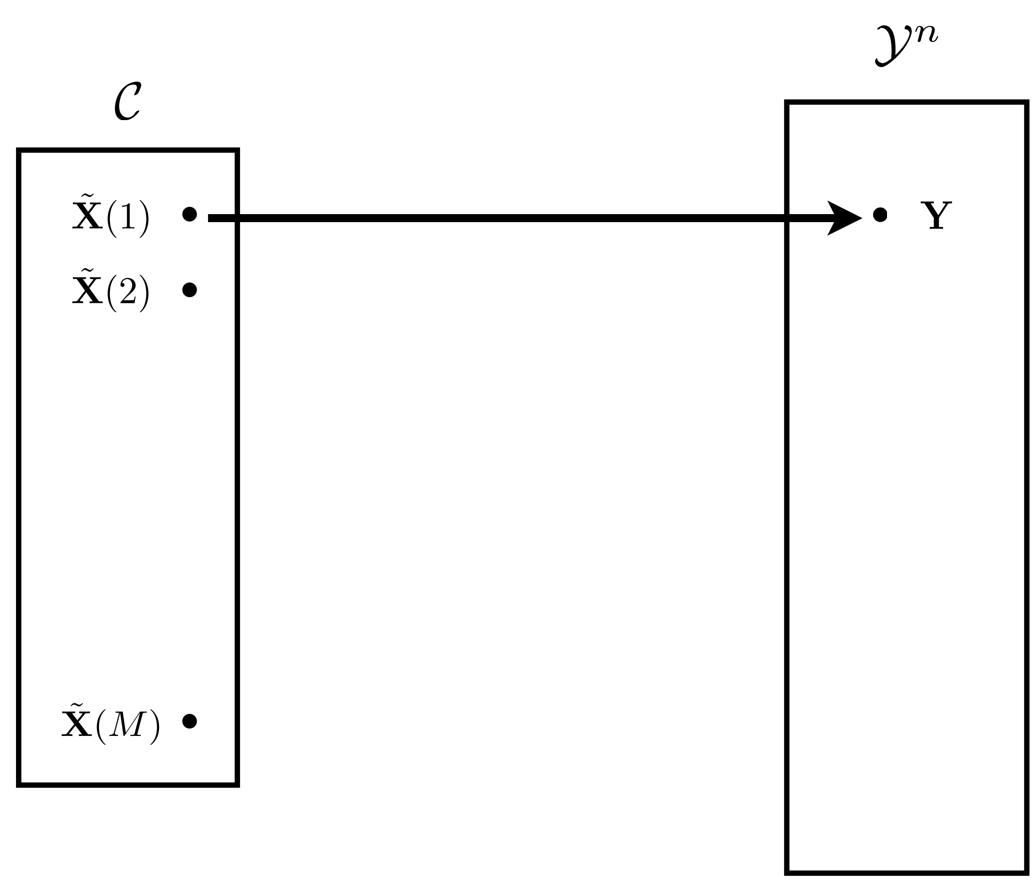
$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W=1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W=1\} \\ &\quad= 1 - \Pr\{Err^c|W=1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\} \\ &\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W=1\} \\ &\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W=1\}.\end{aligned}$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W=w\}\Pr\{W=w\} \\ &= \Pr\{Err|W=1\}\sum_{w=1}^{M}\Pr\{W=w\} \\ &= \Pr\{Err|W=1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

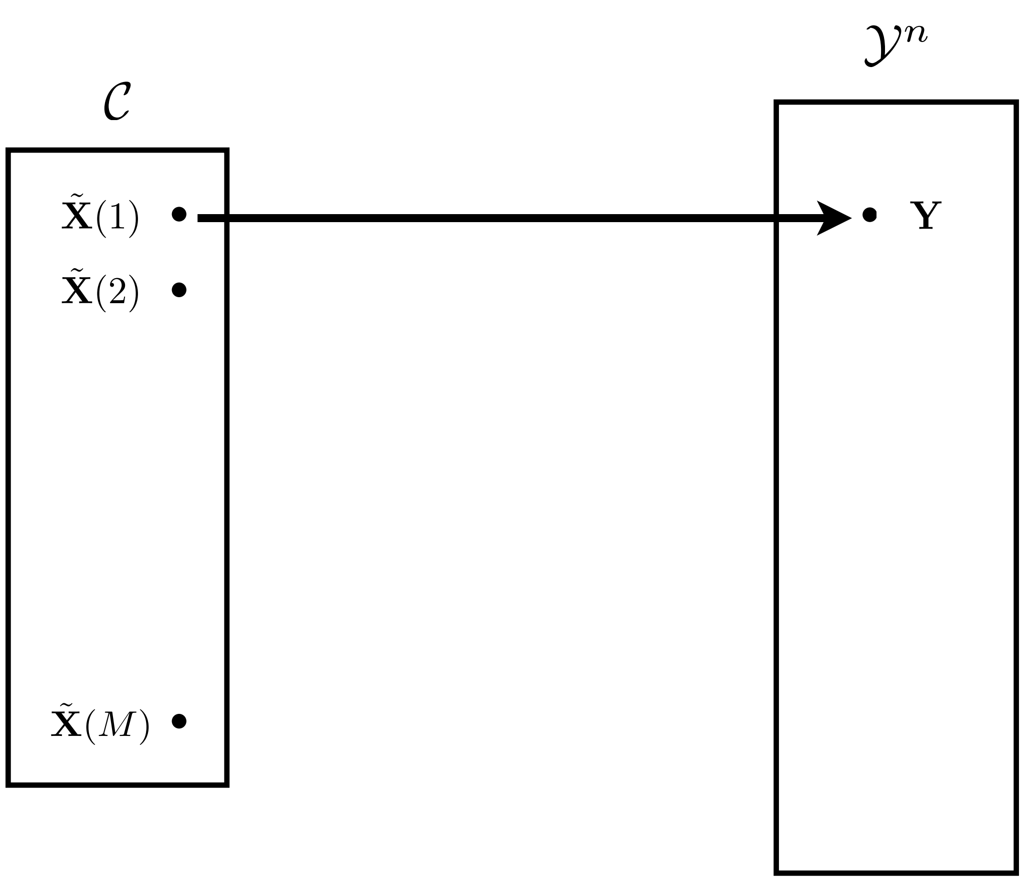
$$\Pr\{Err^c|W=1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W=1\} \\ &\quad= 1 - \Pr\{Err^c|W=1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W=1\} \\ &\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W=1\} \\ &\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W=1\}.\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M}\Pr\{E_w|W=1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

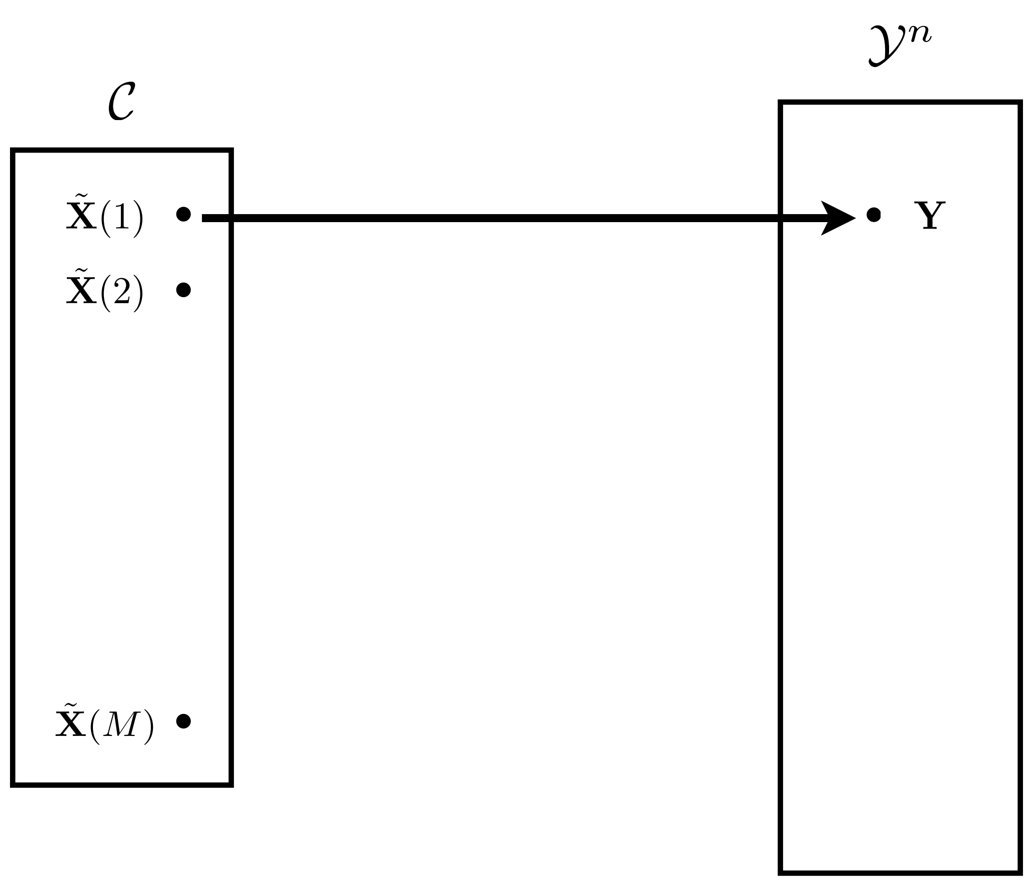
$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W = 1\} \\ &\quad = 1 - \Pr\{Err^c|W = 1\} \\ &\quad \leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad = \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad = \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

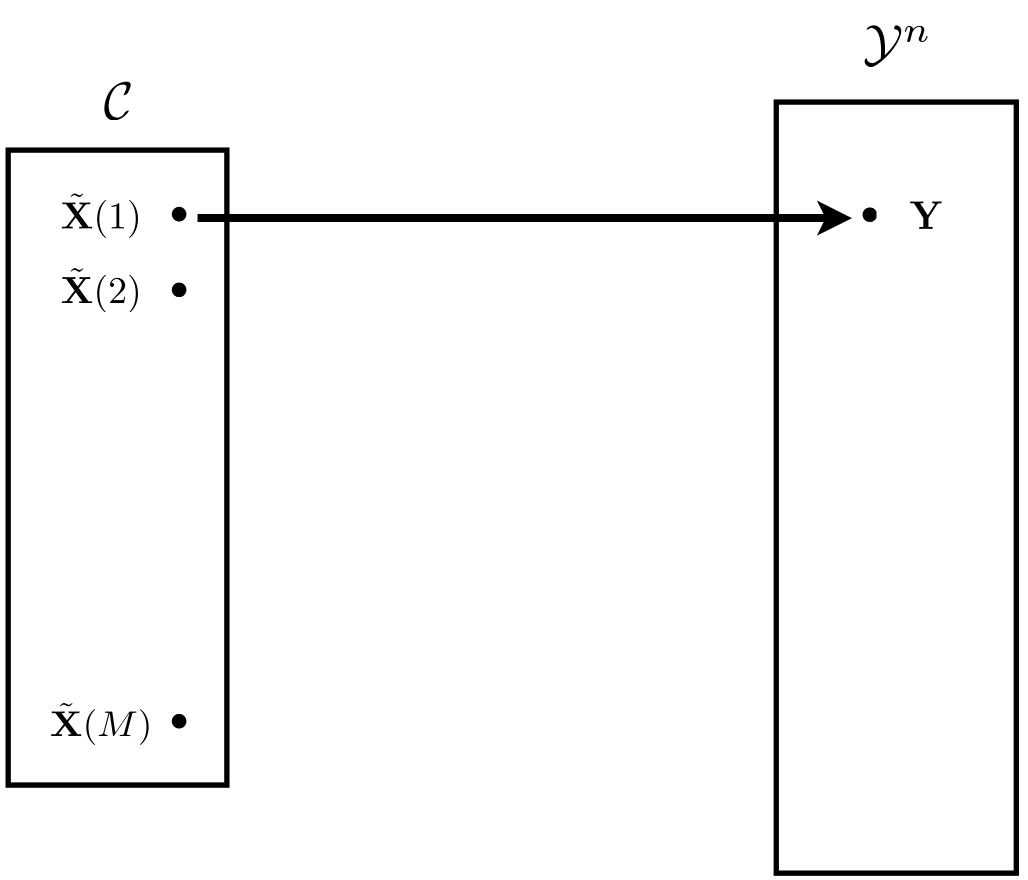
$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W = 1\} \\ &\quad = 1 - \Pr\{Err^c|W = 1\} \\ &\quad \leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad = \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad = \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

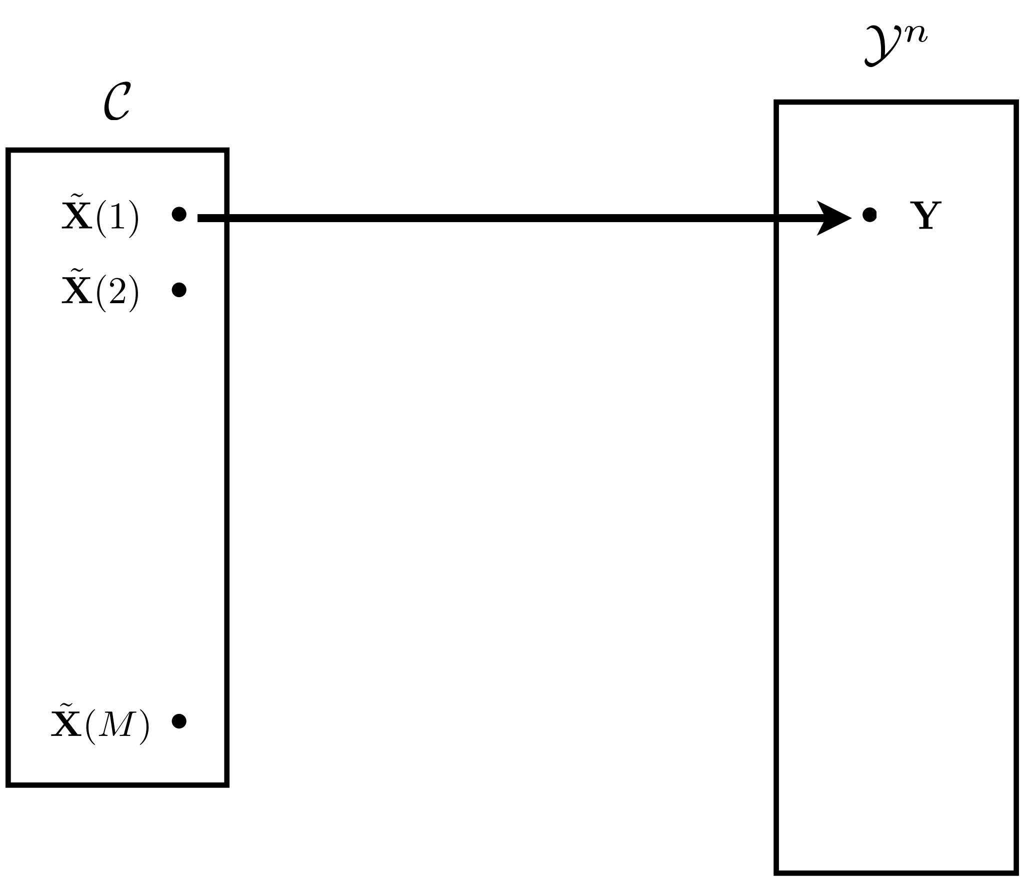
$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W = 1\} \\ &\quad = 1 - \Pr\{Err^c|W = 1\} \\ &\quad \leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad = \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad = \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

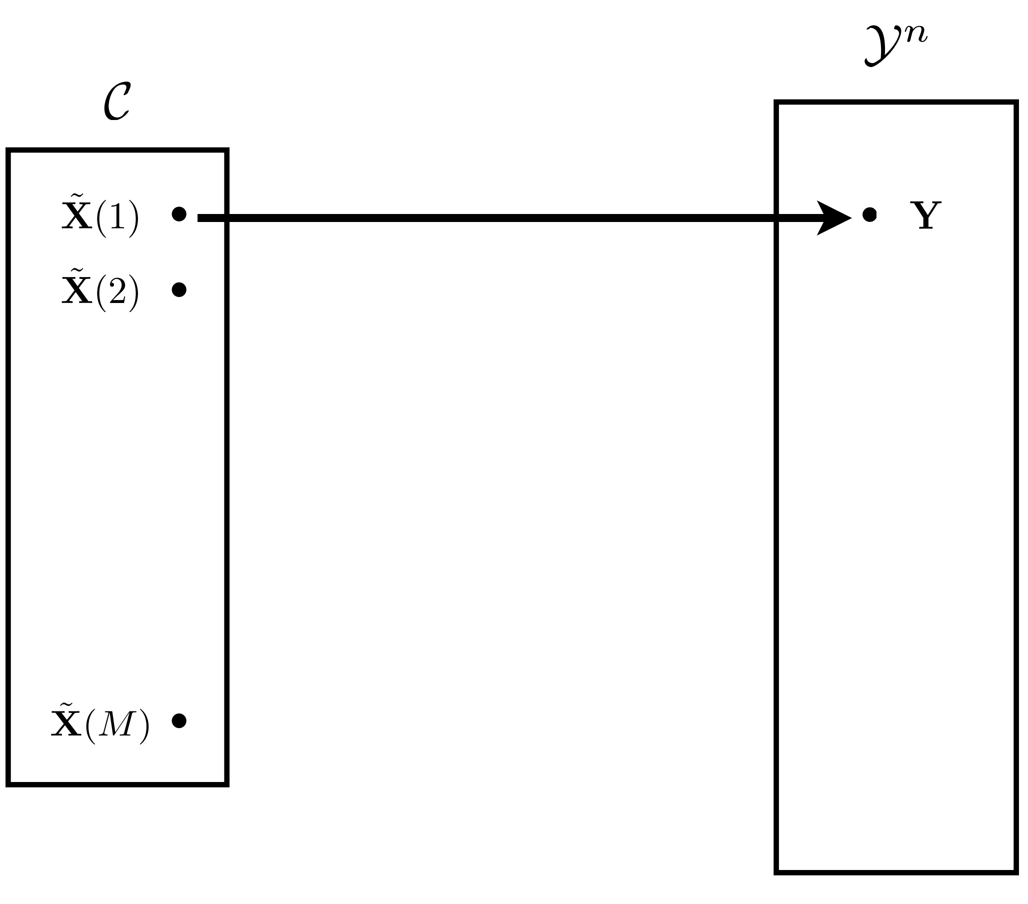
5. Consider

$$\begin{aligned}&\Pr\{Err|W = 1\} \\ &= 1 - \Pr\{Err^c|W = 1\} \\ &\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.\end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

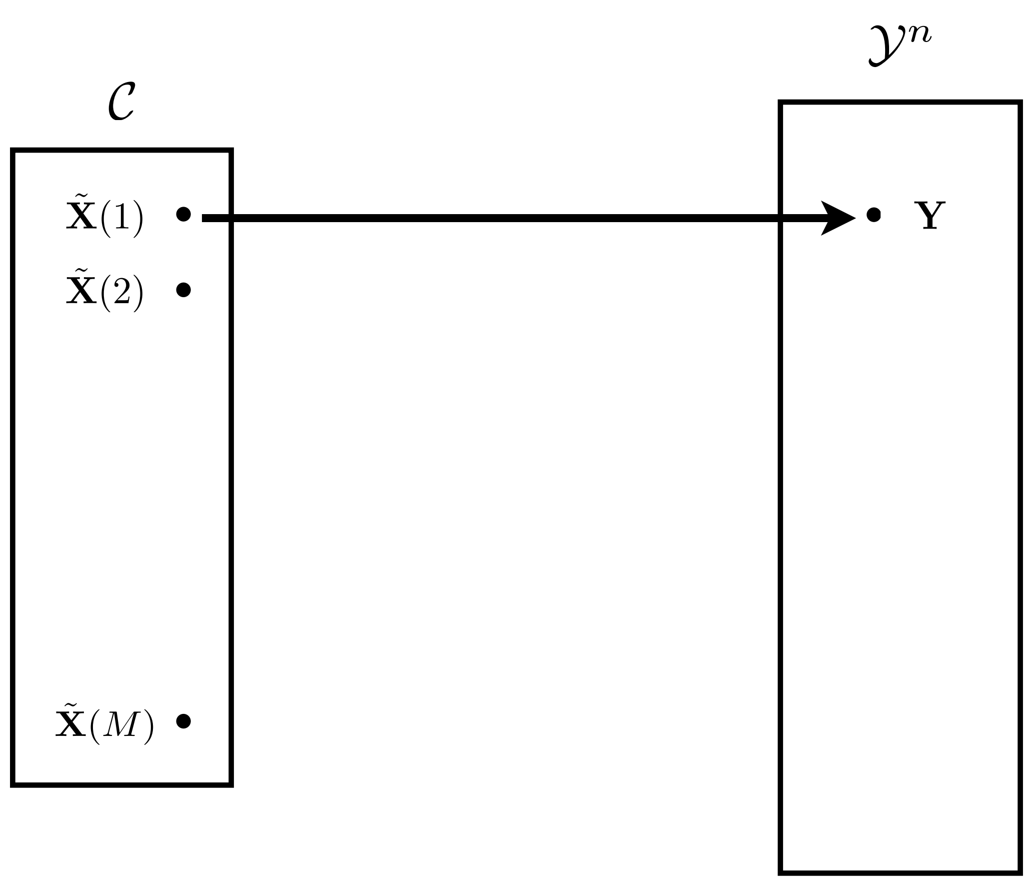
$$\begin{aligned} &\Pr\{Err|W = 1\} \\ &\quad = 1 - \Pr\{Err^c|W = 1\} \\ &\quad \leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad = \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad = \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}. \end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

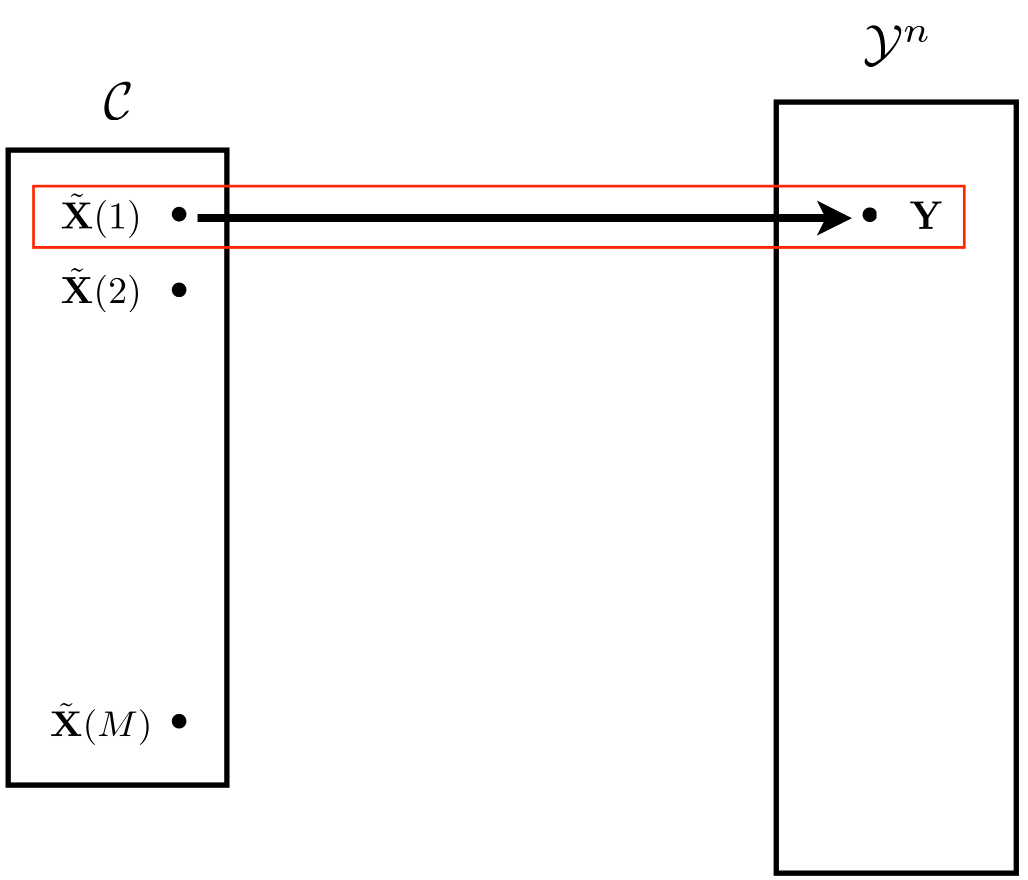
$$\begin{aligned} &\Pr\{Err|W = 1\} \\ &\quad = 1 - \Pr\{Err^c|W = 1\} \\ &\quad \leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad = \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad = \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}. \end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

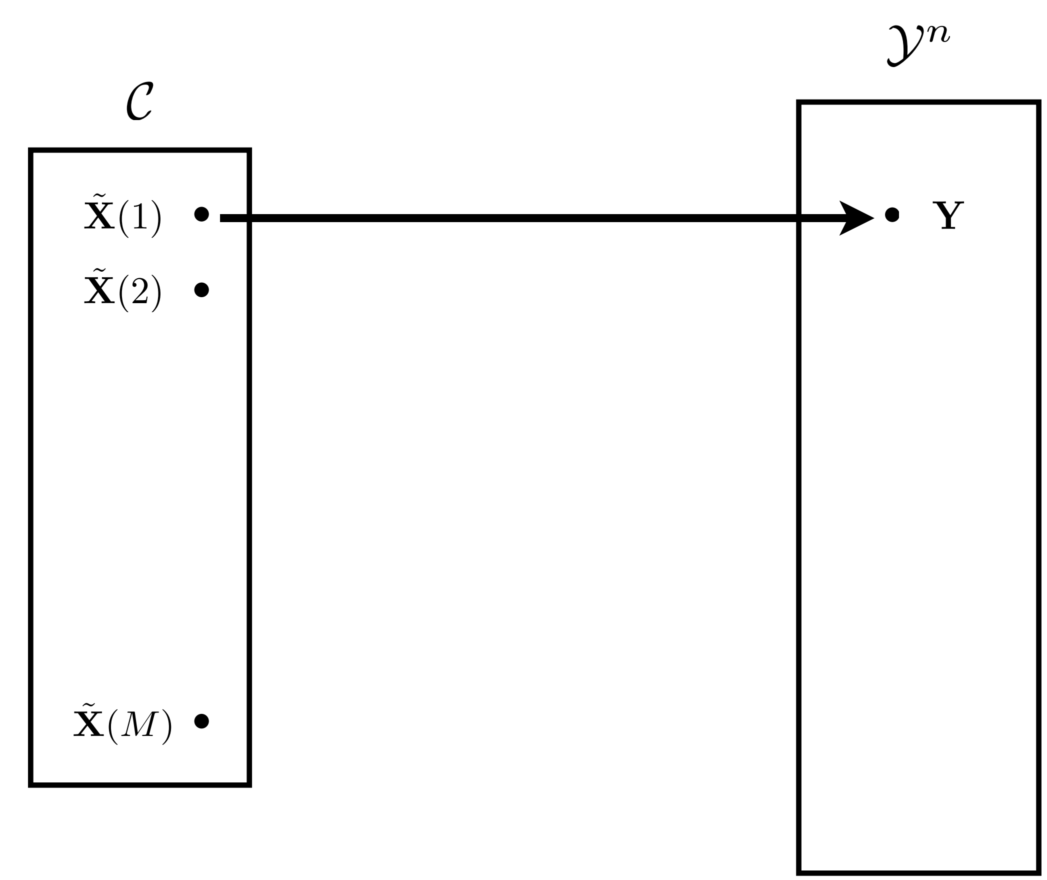
$$\begin{aligned} &\Pr\{Err|W = 1\} \\ &\quad= 1 - \Pr\{Err^c|W = 1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}. \end{aligned}$$

6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned}\Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}.\end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned}&\Pr\{Err|W = 1\} \\ &\quad= 1 - \Pr\{Err^c|W = 1\} \\ &\quad\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &\quad= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &\quad= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}.\end{aligned}$$

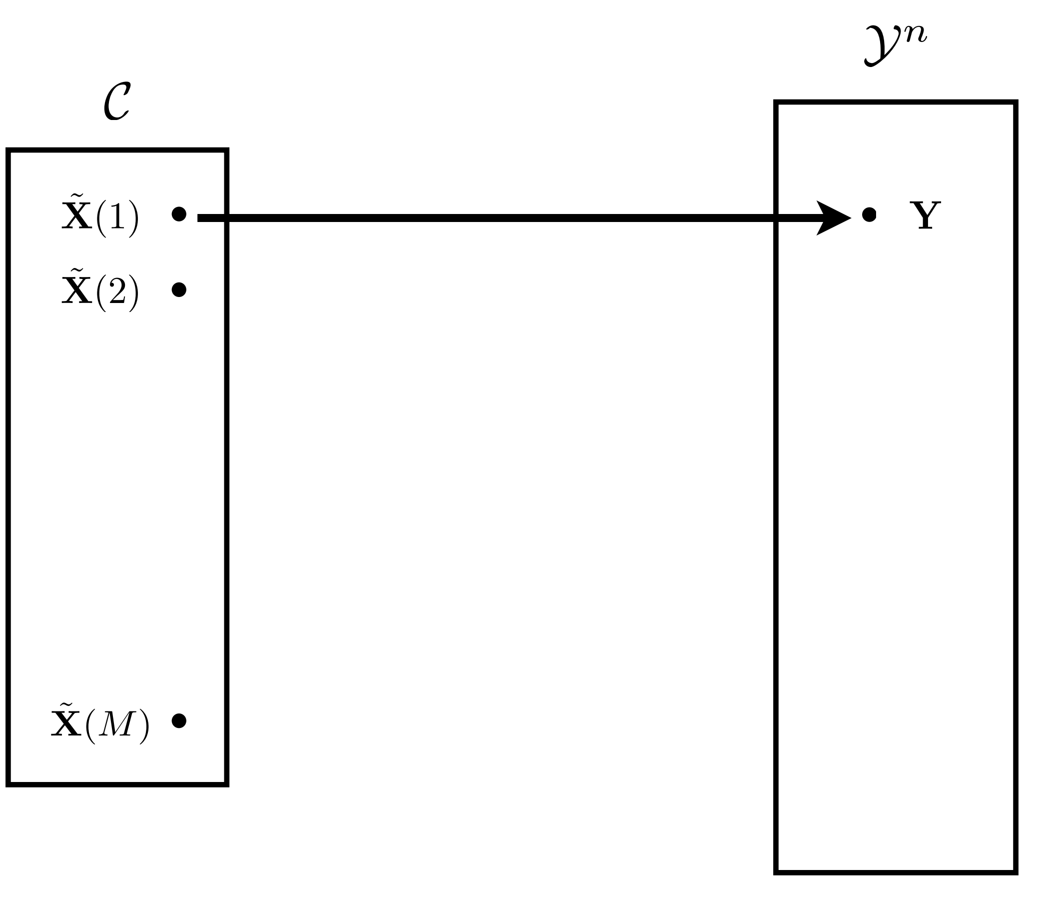
6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



Performance Analysis

1. We need to show that $\Pr\{Err\} = \Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

2. Consider

$$\begin{aligned} \Pr\{Err\} &= \sum_{w=1}^{M} \Pr\{Err|W = w\}\Pr\{W = w\} \\ &= \Pr\{Err|W = 1\} \sum_{w=1}^{M} \Pr\{W = w\} \\ &= \Pr\{Err|W = 1\}. \end{aligned}$$

Assume without loss of generality that the message 1 is chosen.

3. For $1 \leq w \leq M$, define the event

$$E_w = \{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}\}.$$

4. If E_1 occurs but E_w does not occur for all $2 \leq w \leq M$, then no decoding error. Therefore,

$$\Pr\{Err^c|W = 1\} \geq \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\}.$$

5. Consider

$$\begin{aligned} &\Pr\{Err|W = 1\} \\ &= 1 - \Pr\{Err^c|W = 1\} \\ &\leq 1 - \Pr\{E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c|W = 1\} \\ &= \Pr\{(E_1 \cap E_2^c \cap E_3^c \cap \cdots \cap E_M^c)^c|W = 1\} \\ &= \Pr\{E_1^c \cup E_2 \cup E_3 \cup \cdots \cup E_M|W = 1\}. \end{aligned}$$

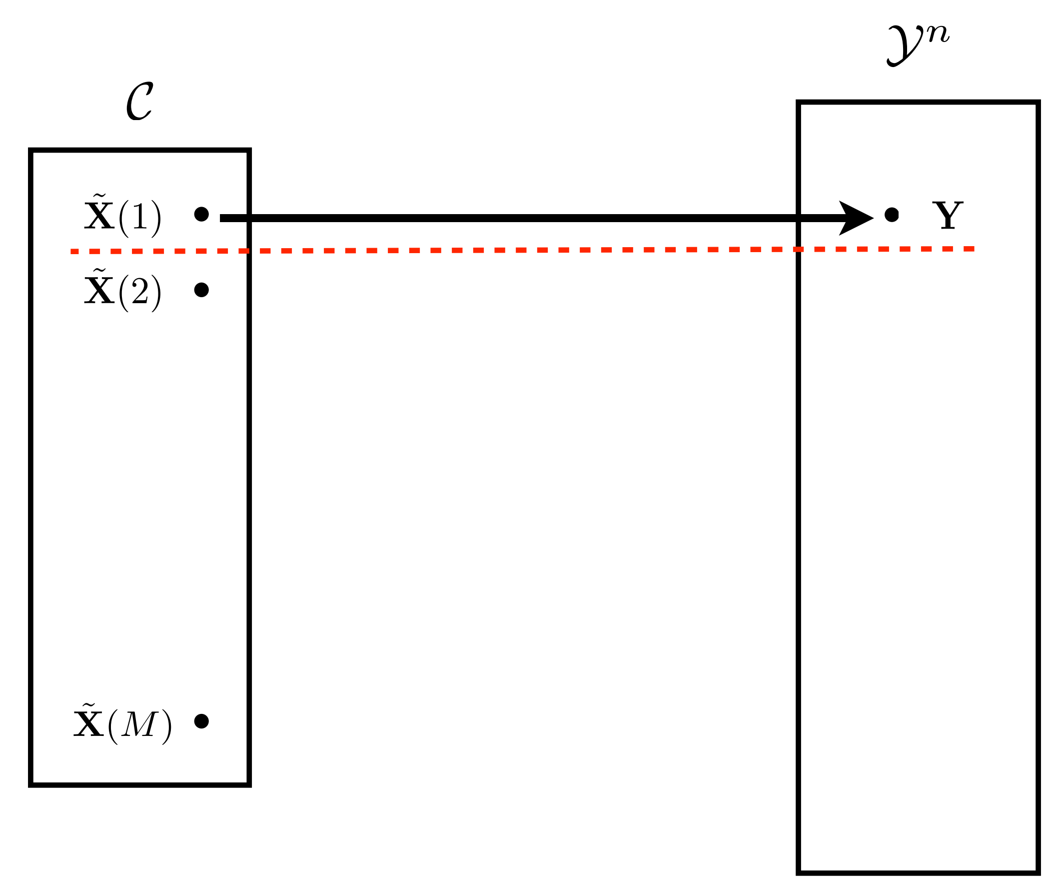
6. By the union bound,

$$\Pr\{Err|W = 1\} \leq \Pr\{E_1^c|W = 1\} + \sum_{w=2}^{M} \Pr\{E_w|W = 1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W = 1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W = 1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



Performance Analysis

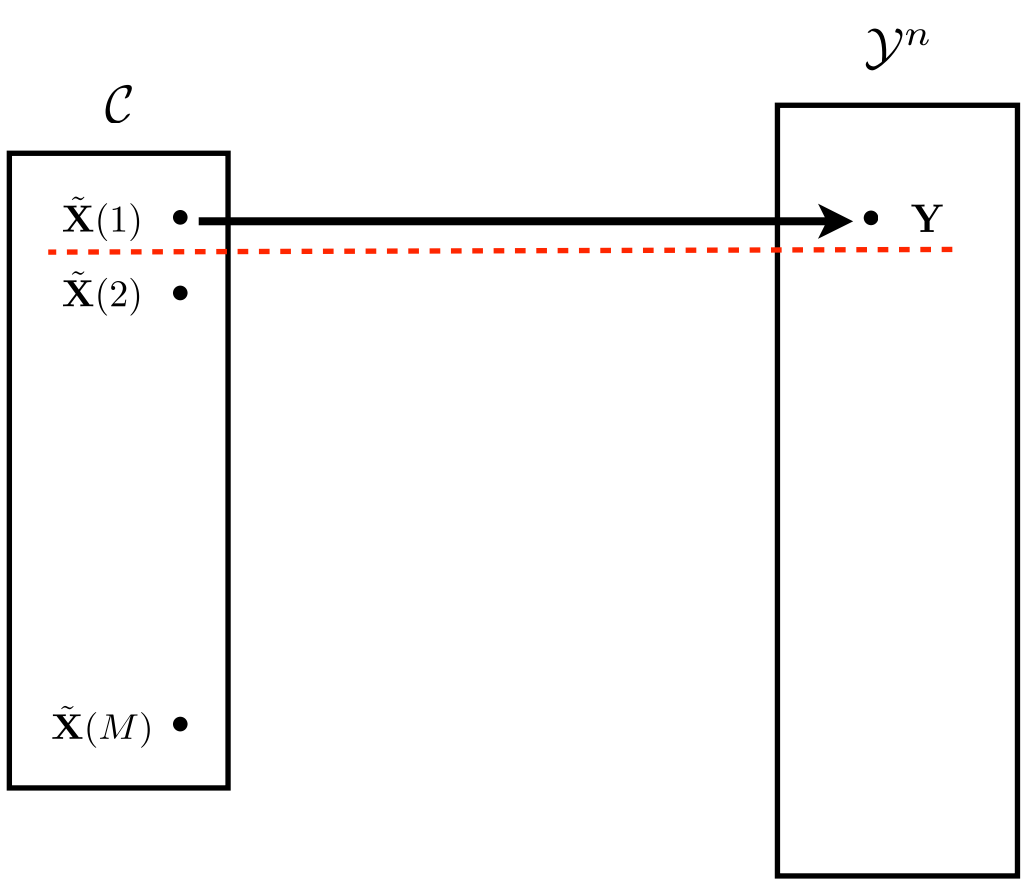
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



Performance Analysis

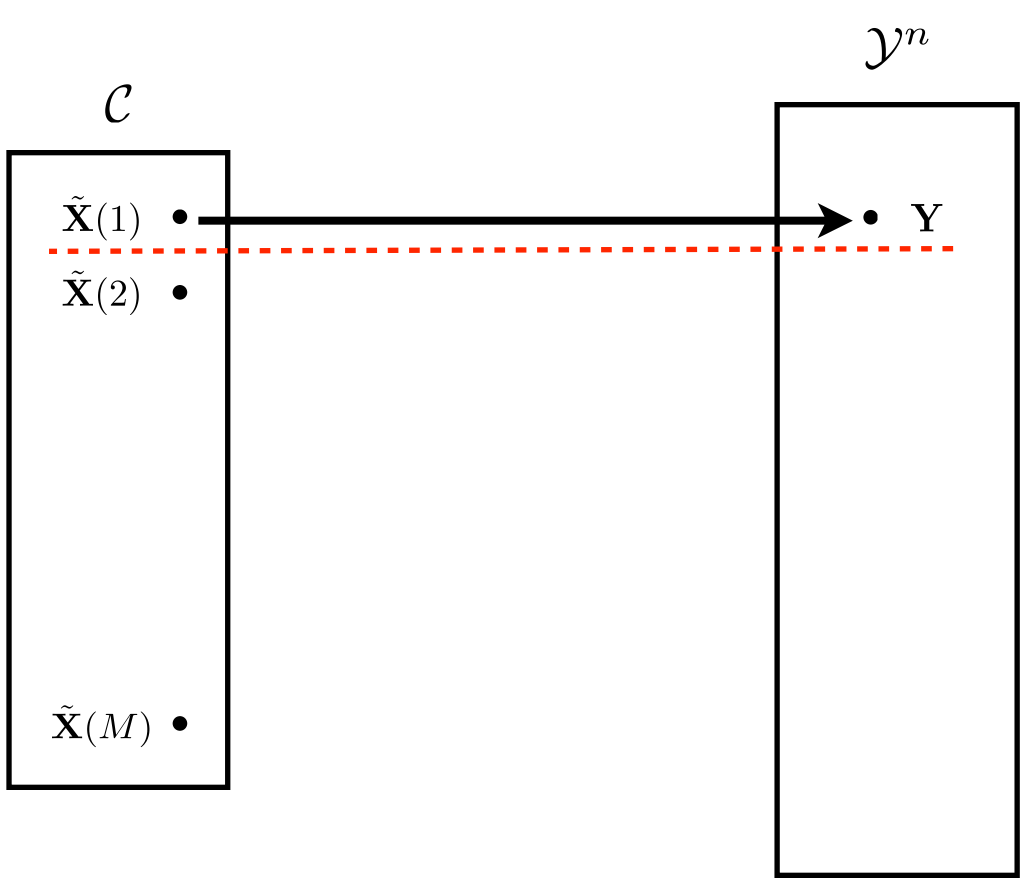
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



Performance Analysis

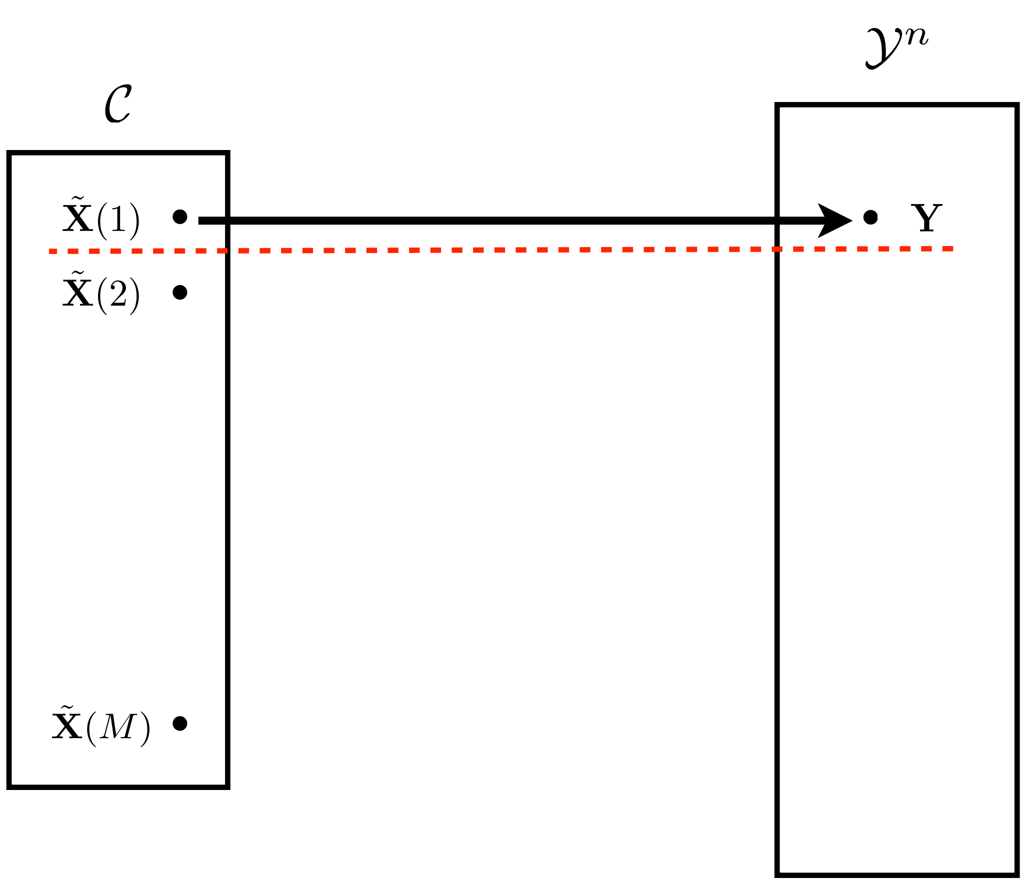
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

Performance Analysis

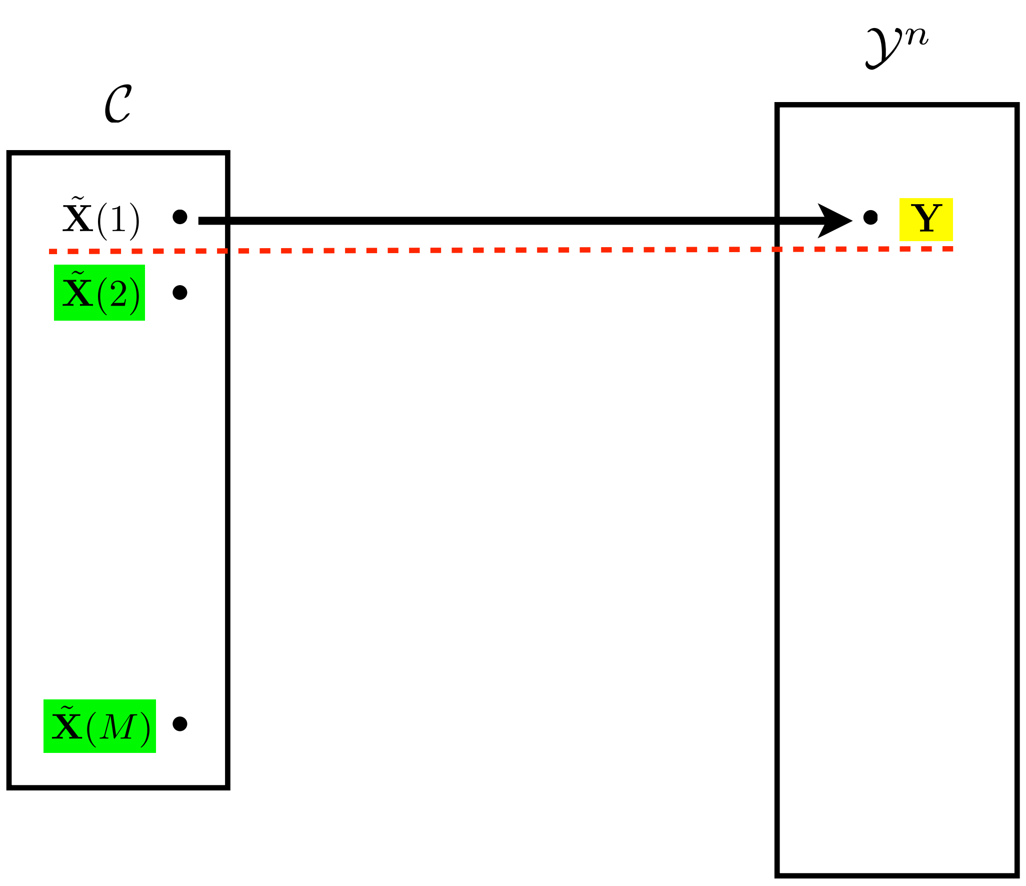
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

Performance Analysis

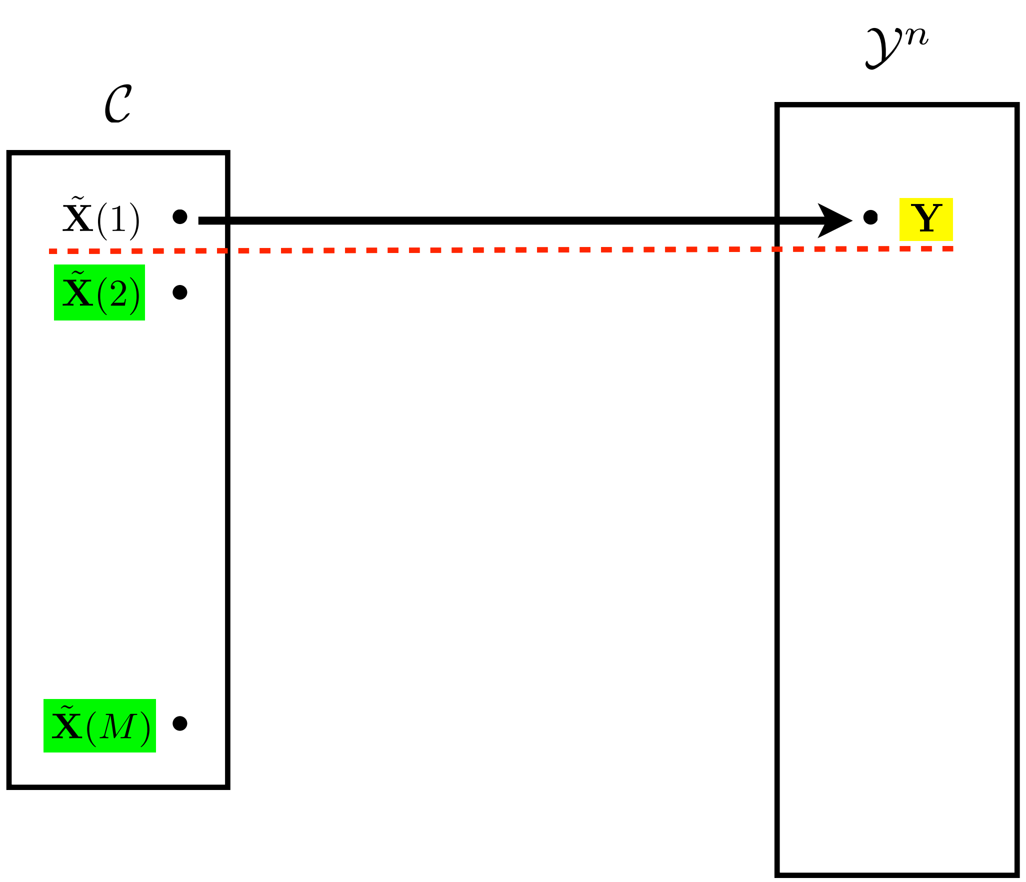
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

Performance Analysis

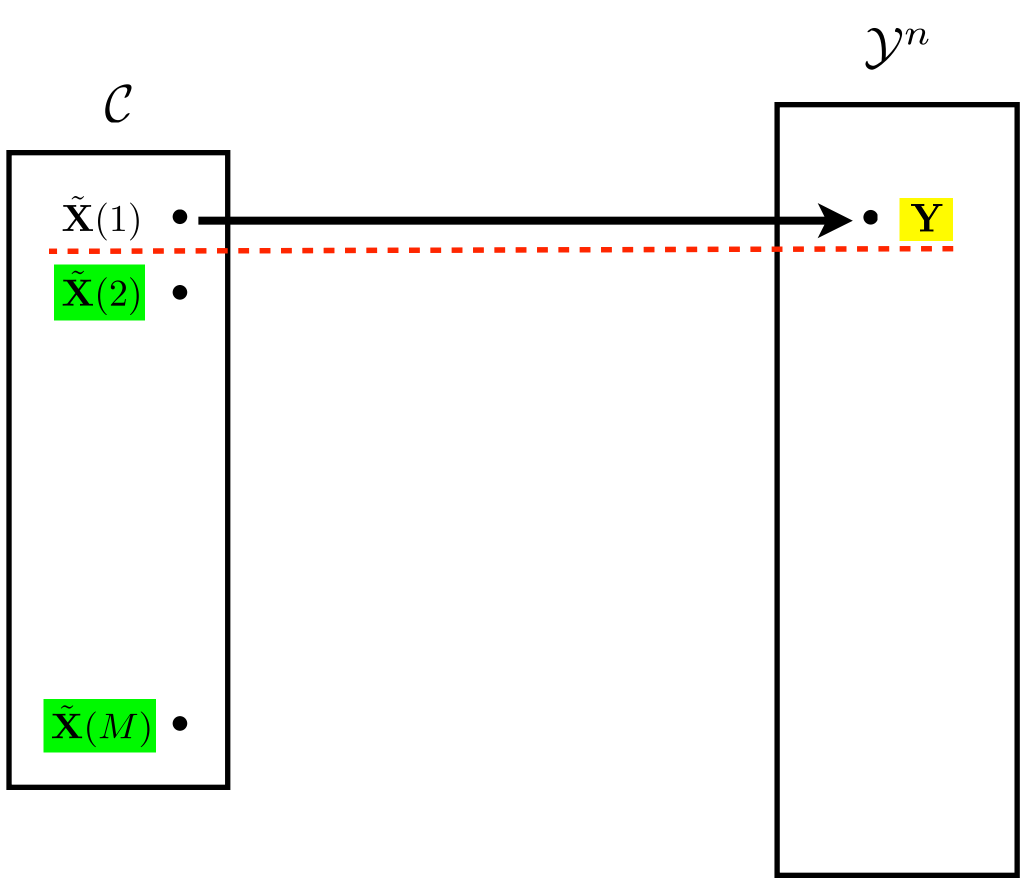
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

Performance Analysis

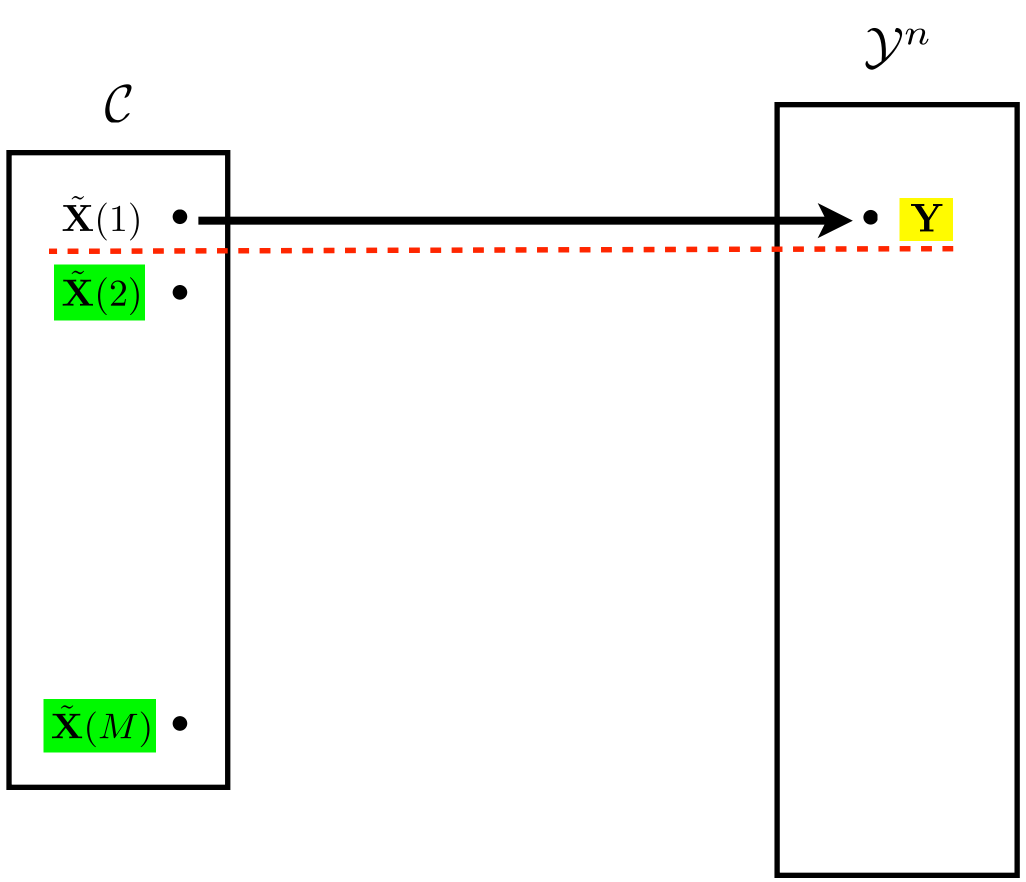
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\Pr\{E_w|W=1\} = \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\}$$

Performance Analysis

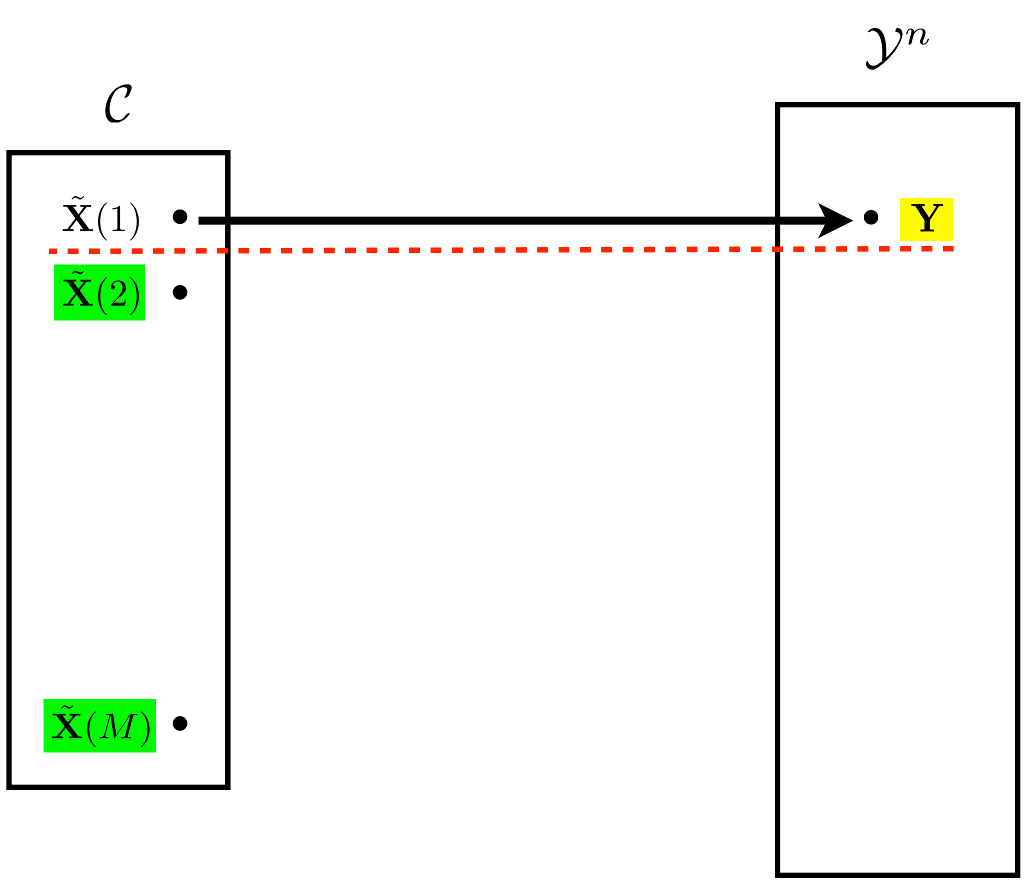
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\Pr\{E_w|W=1\} = \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

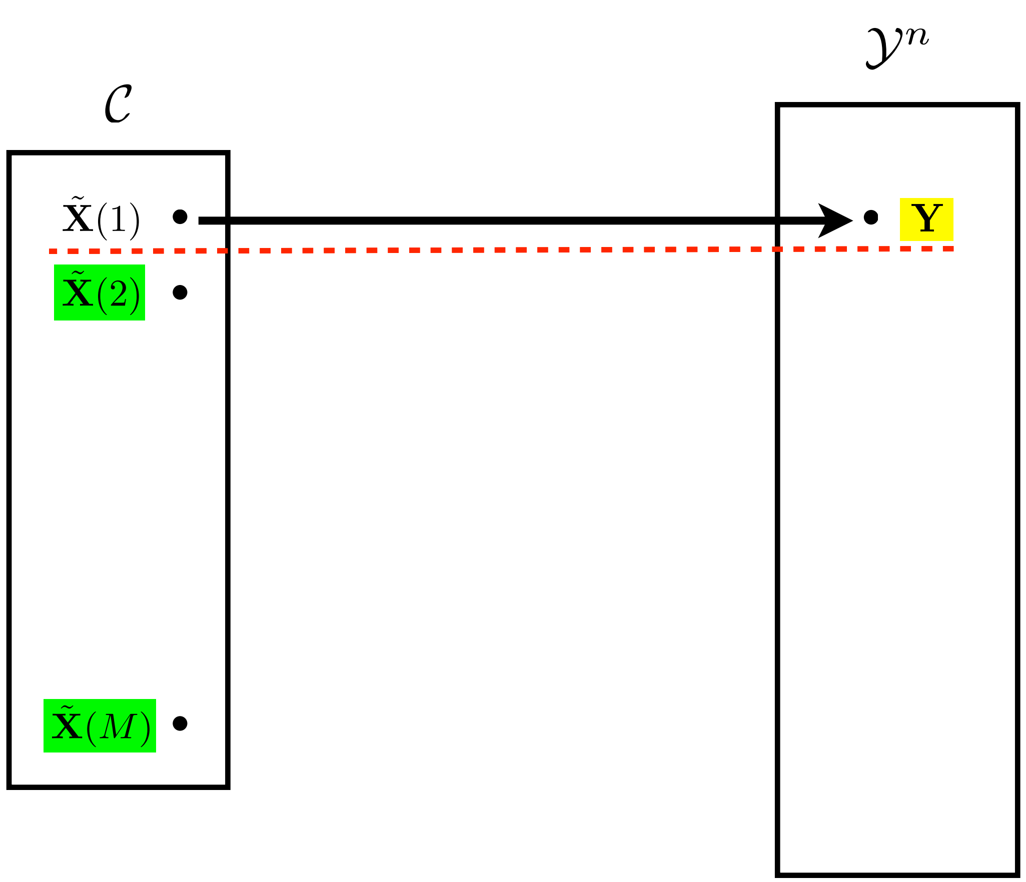
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W = 1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

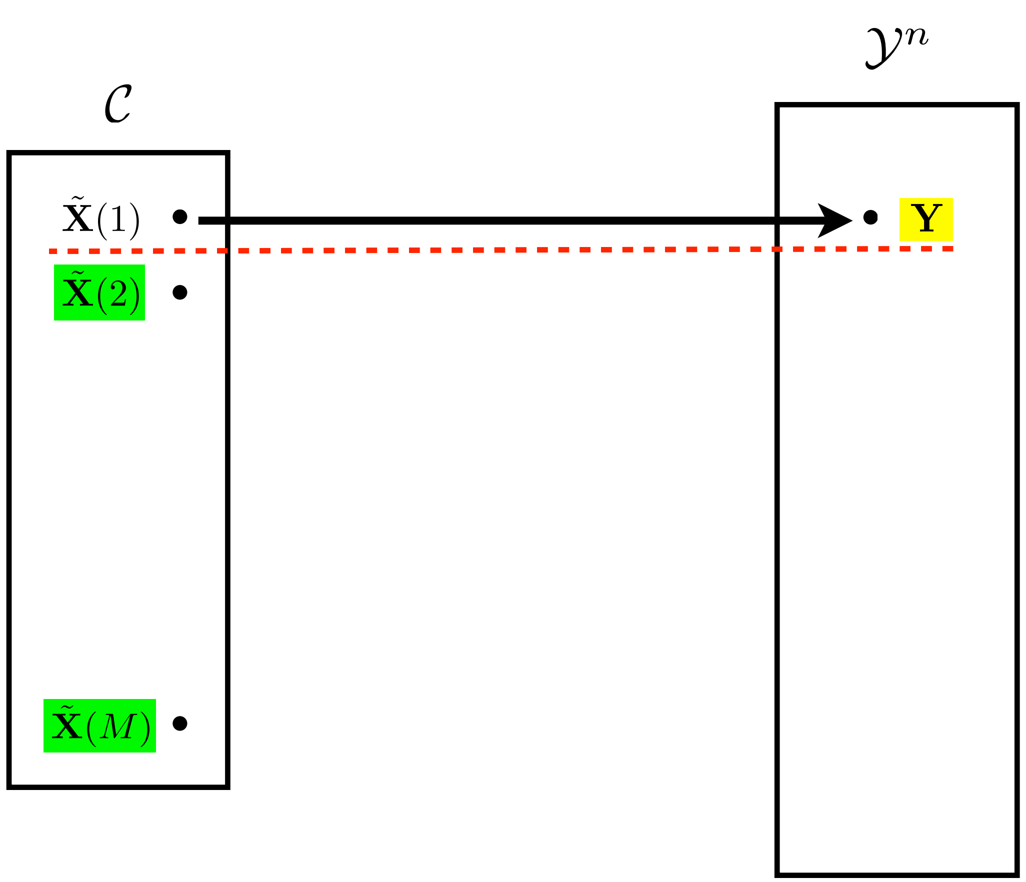
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

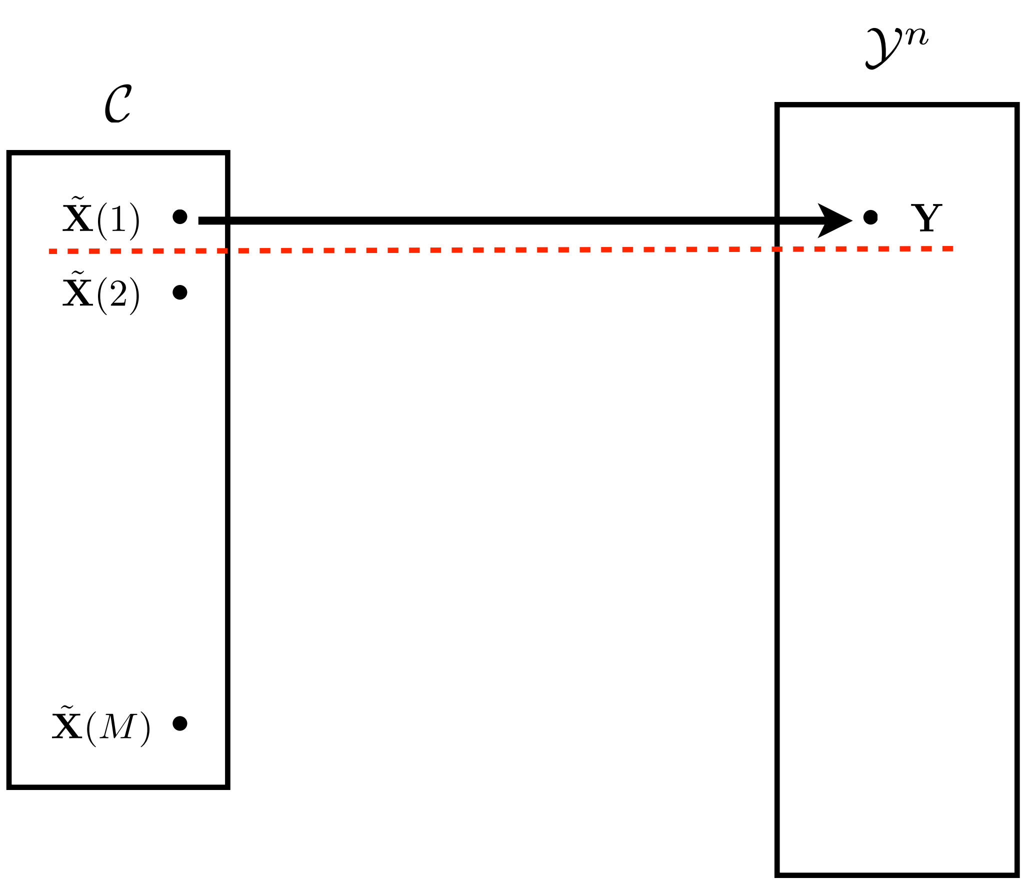
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

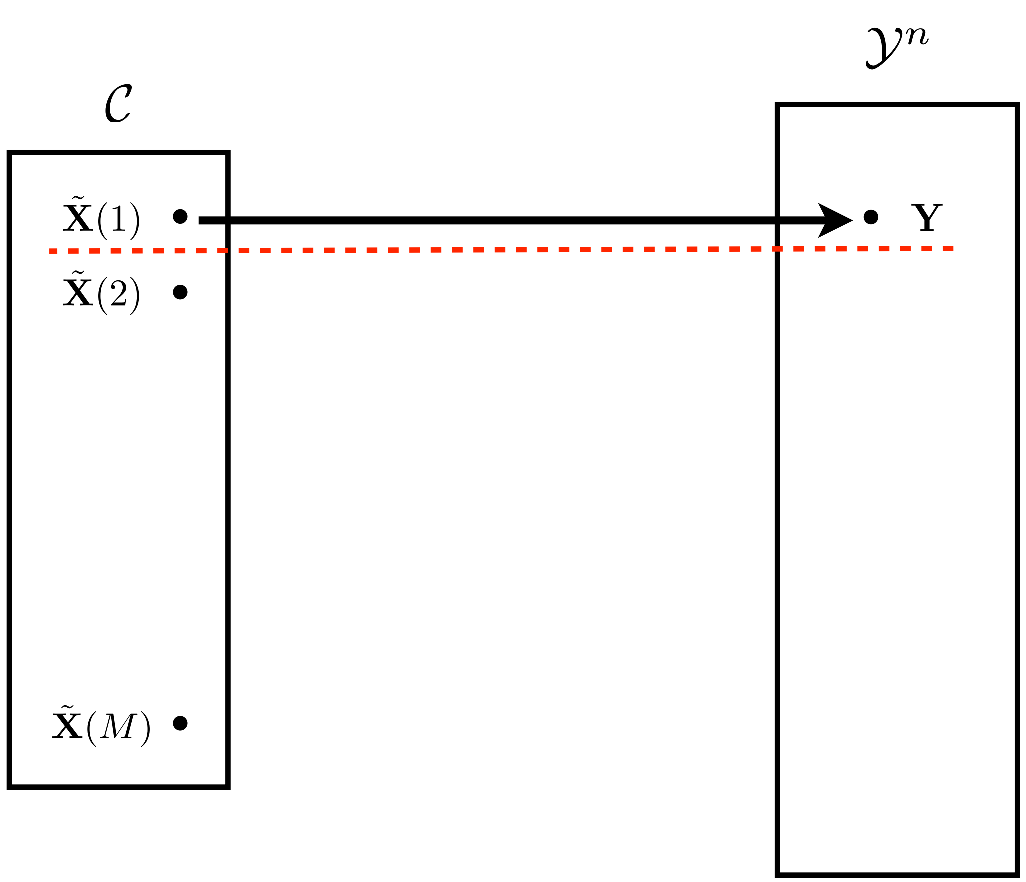
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

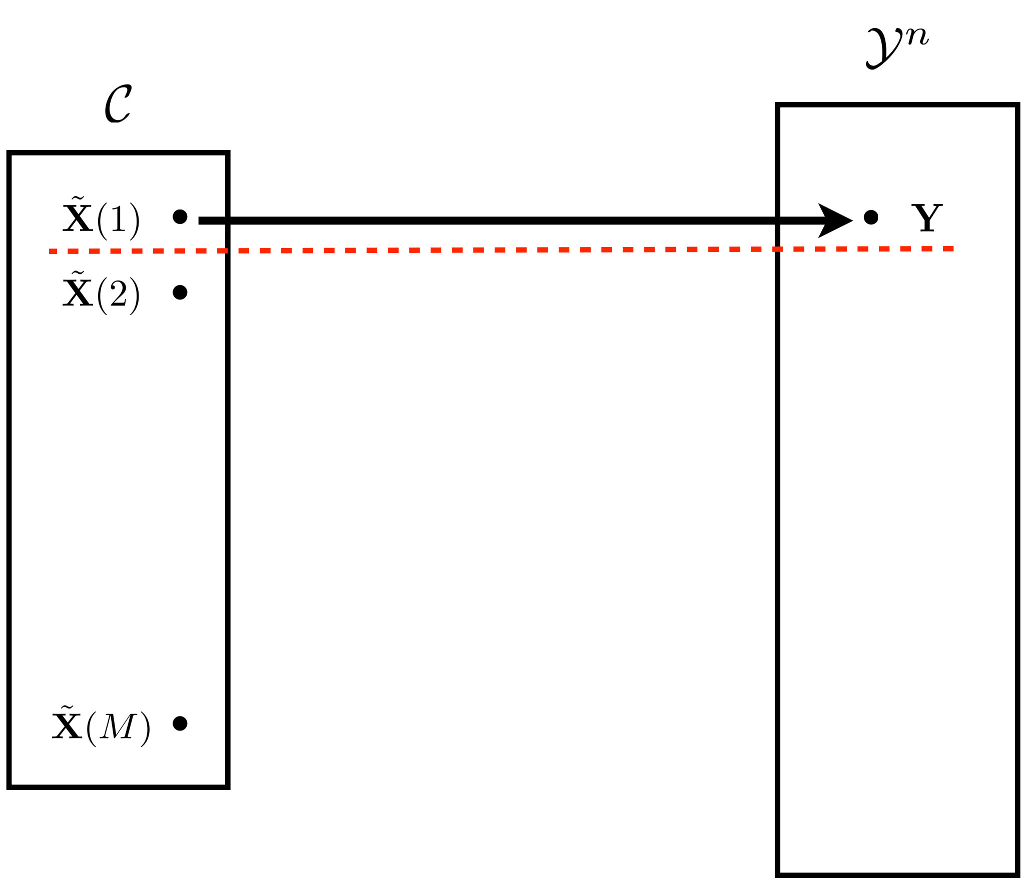
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

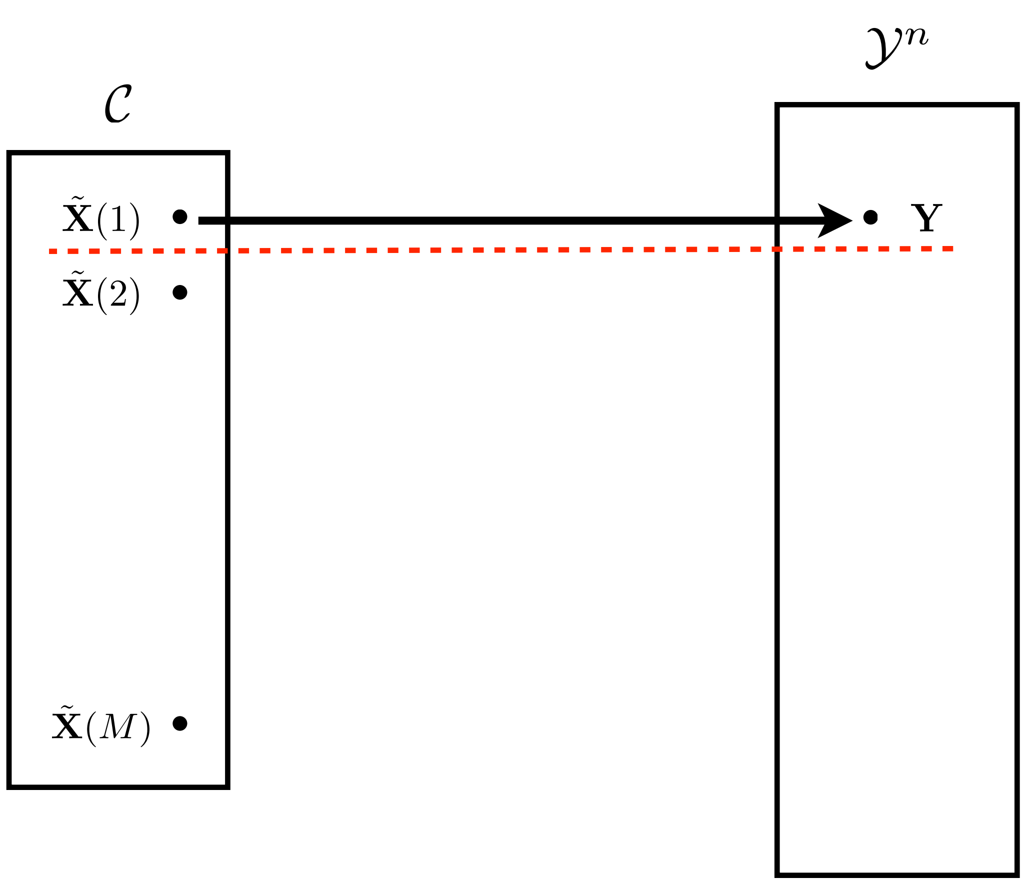
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

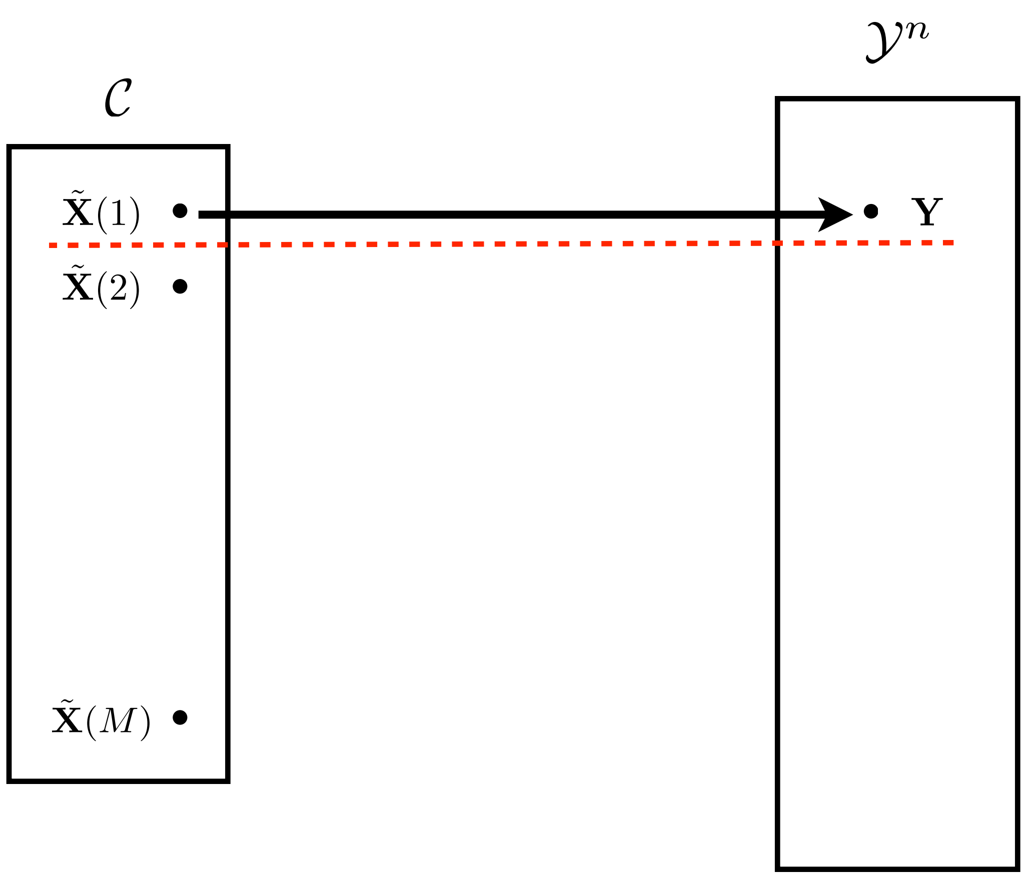
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

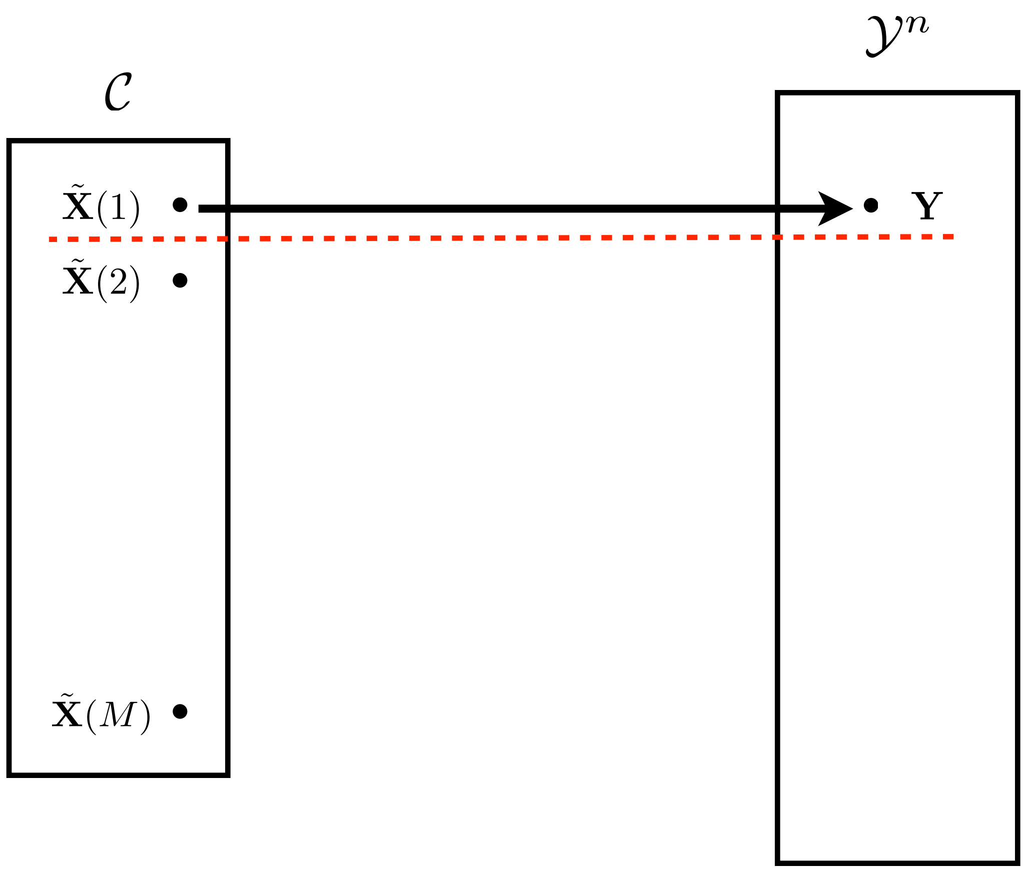
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

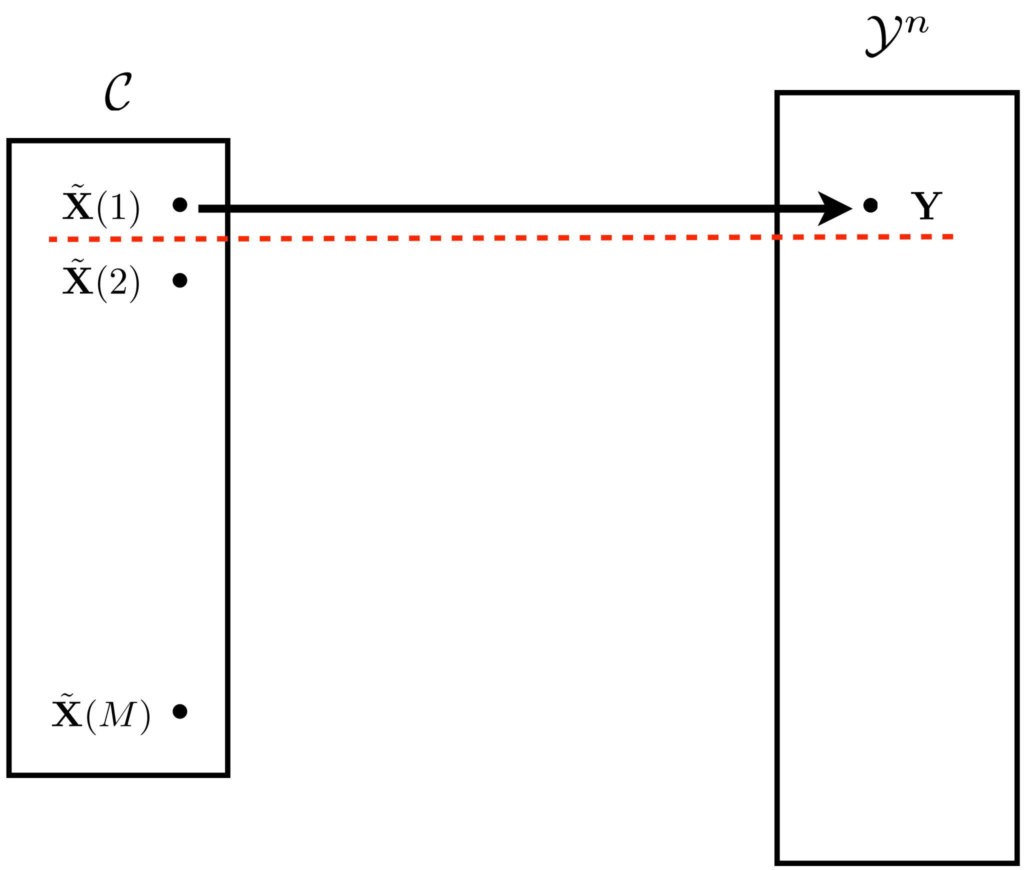
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

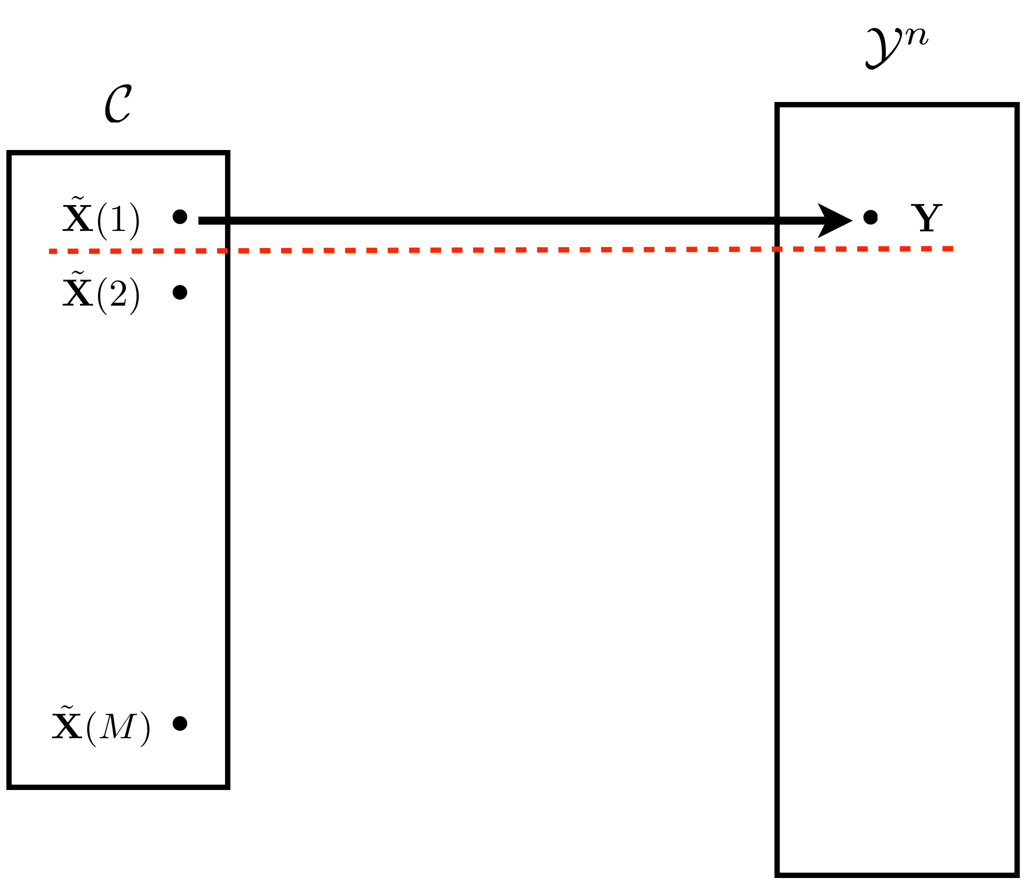
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

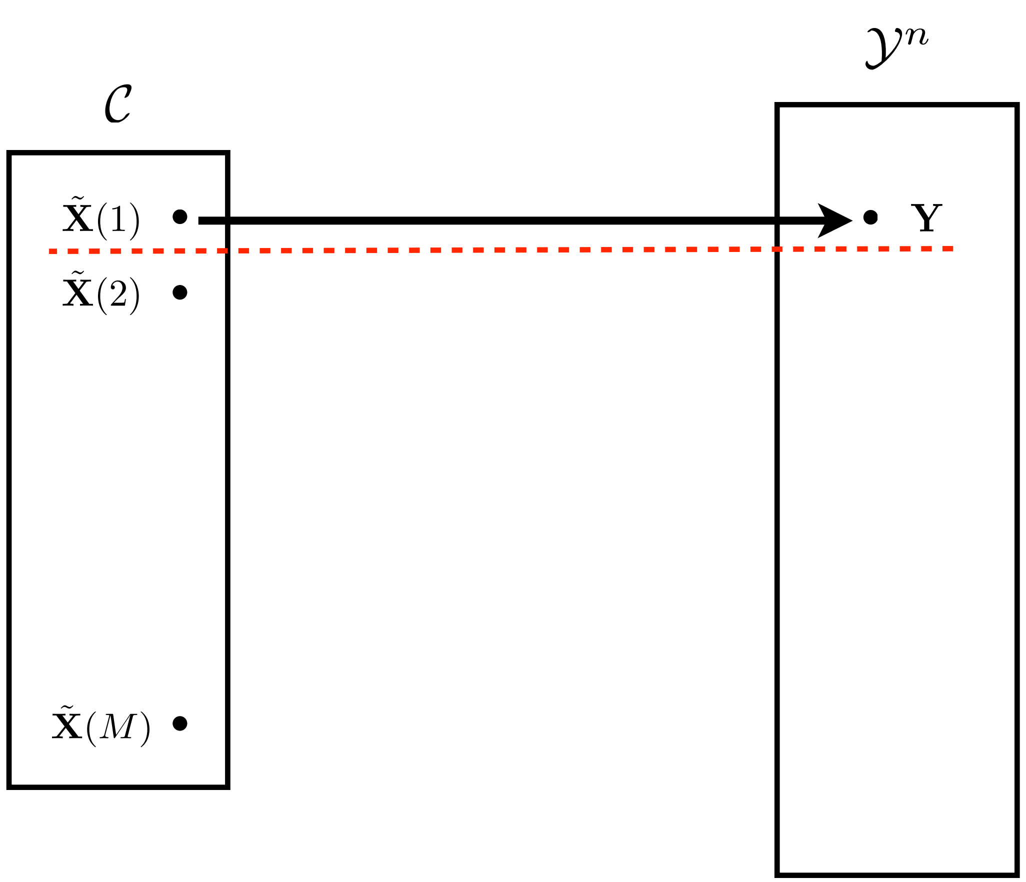
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

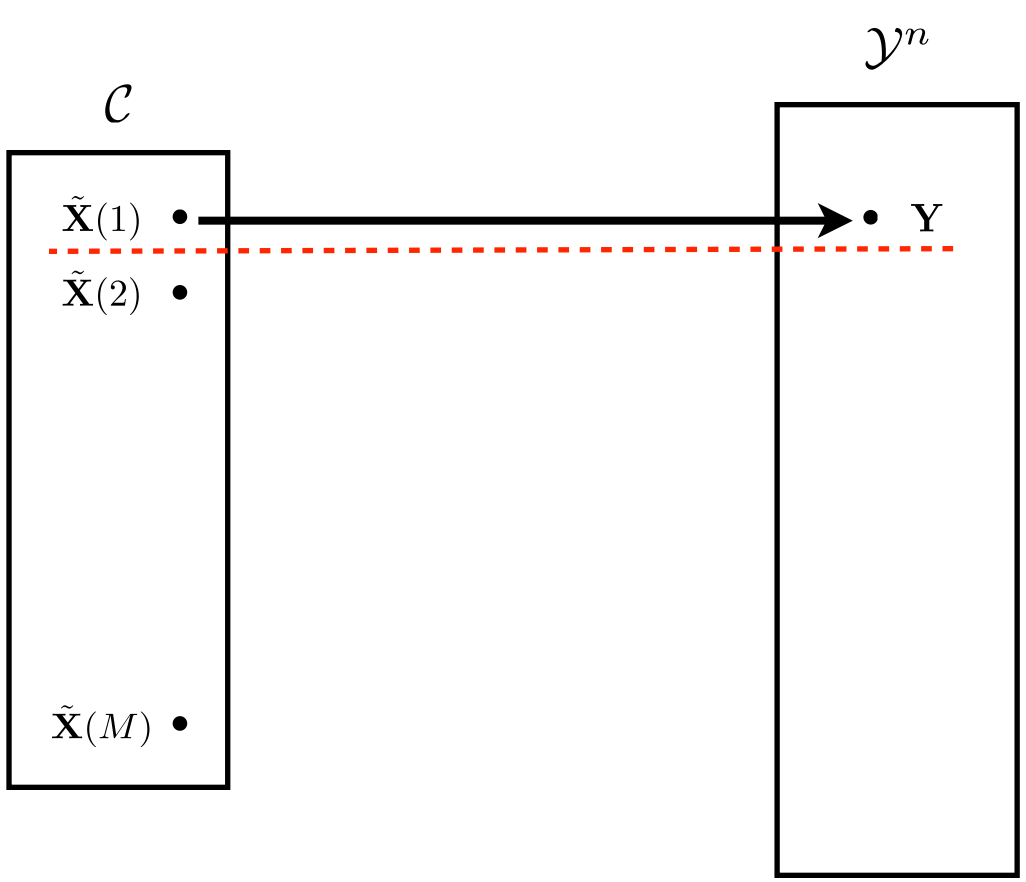
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

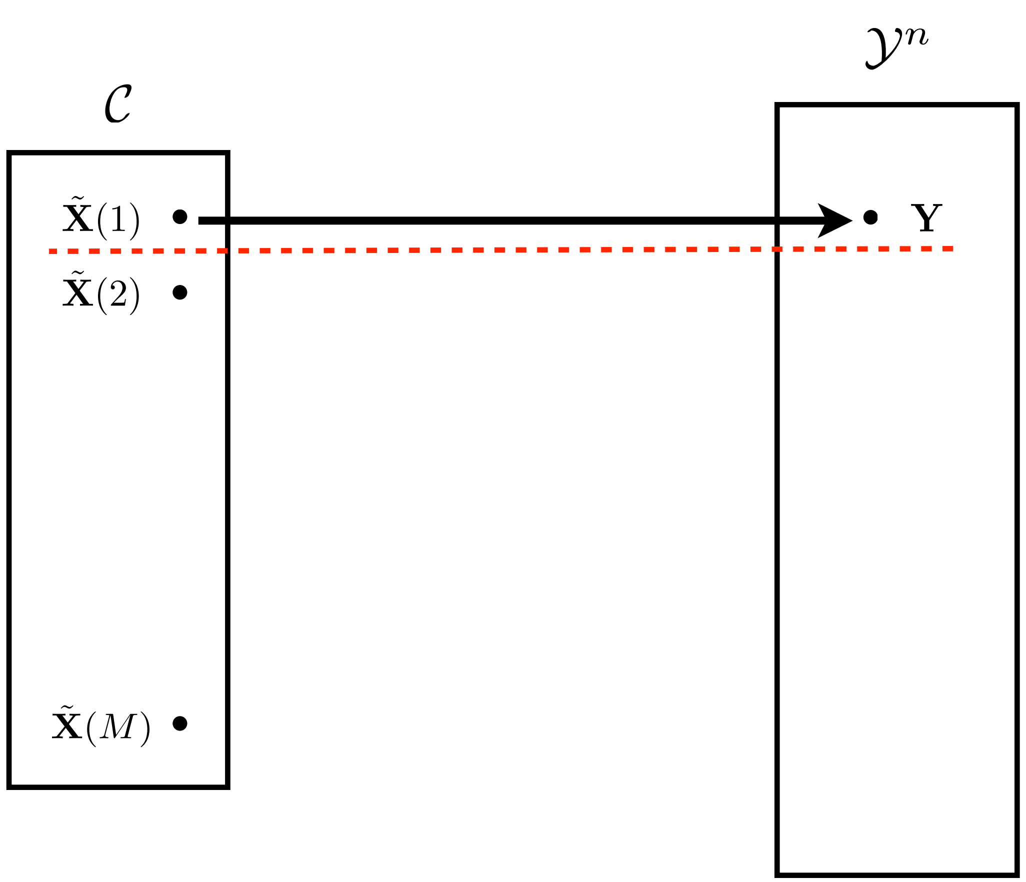
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\Pr\{E_w|W=1\} = \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \le 2^{-n(I(X;Y)-\tau)}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

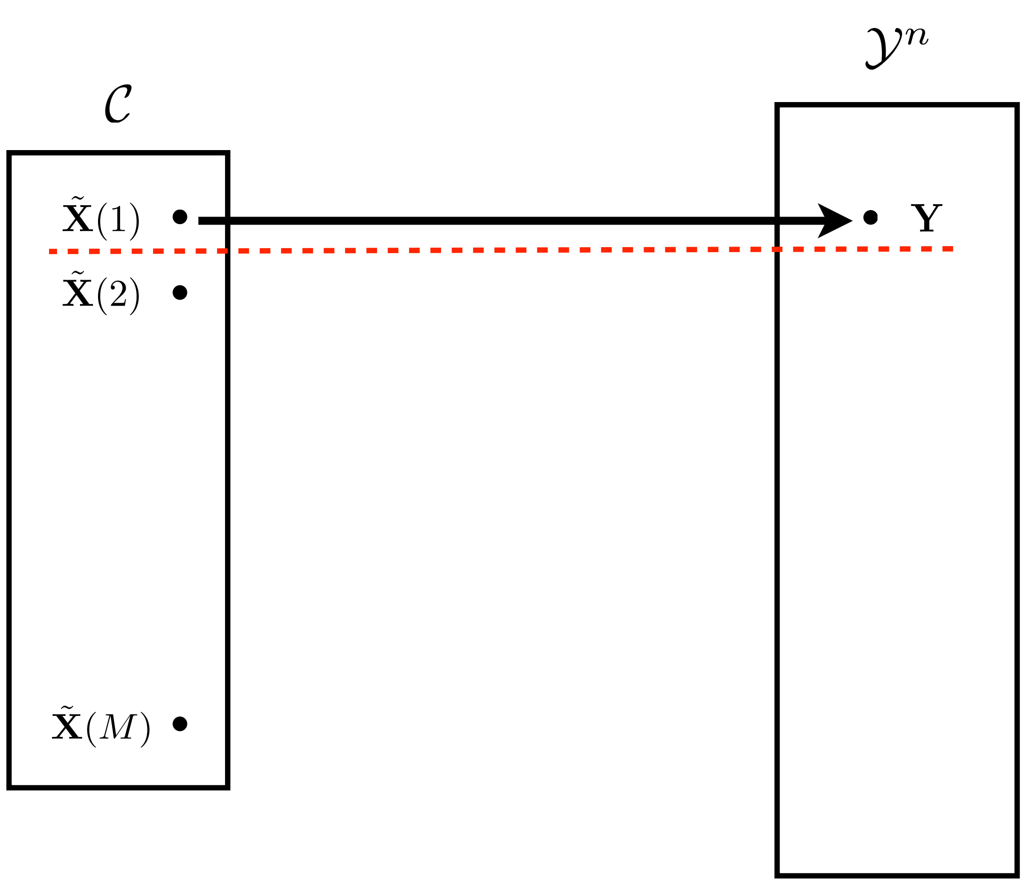
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

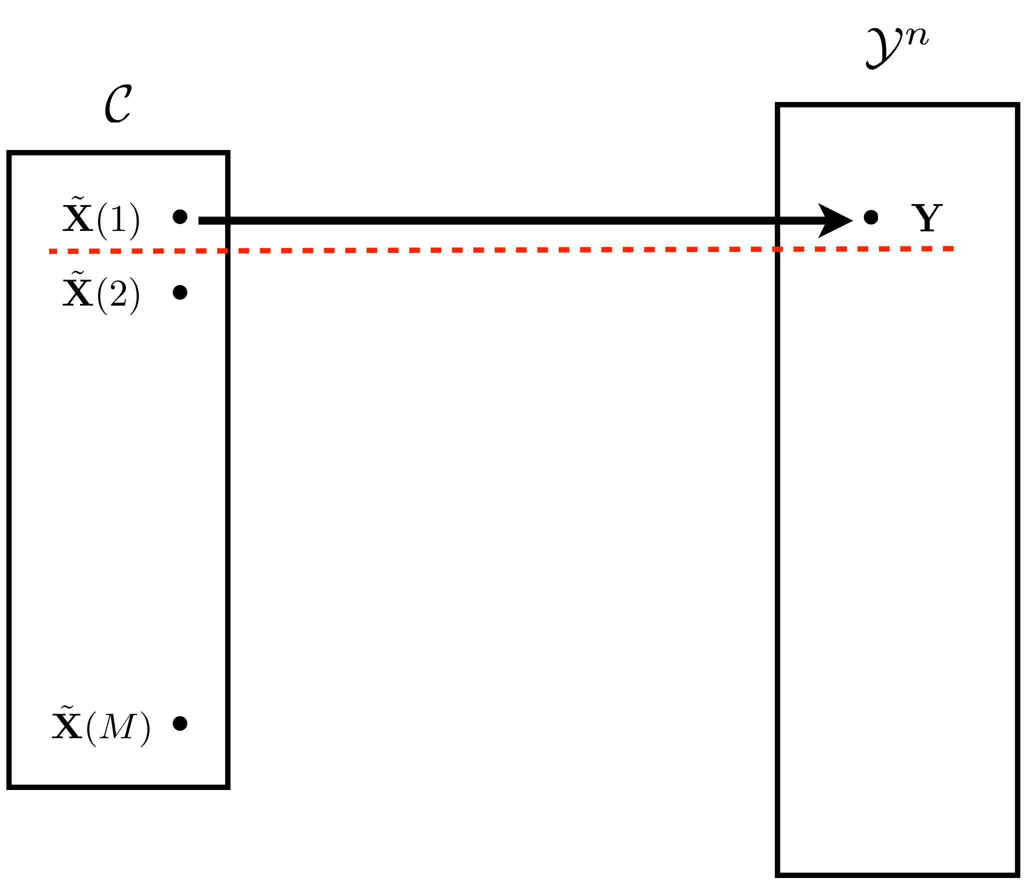
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

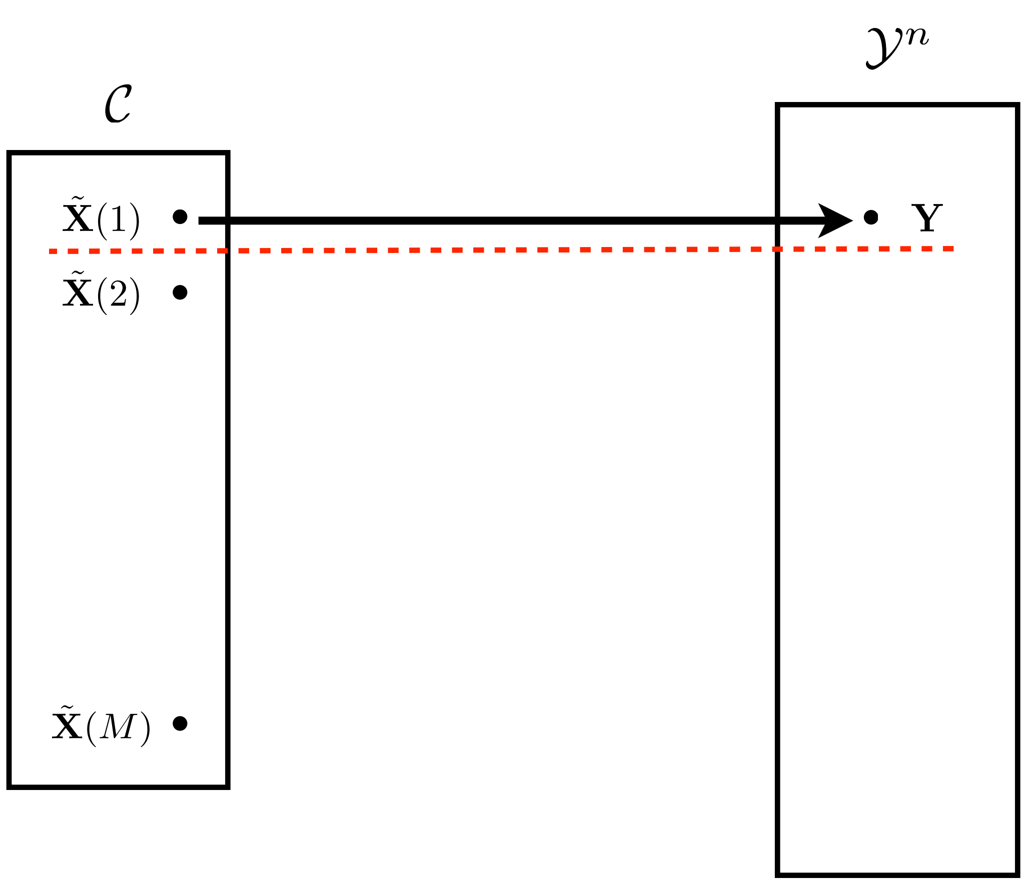
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1), \mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w), \mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X', Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

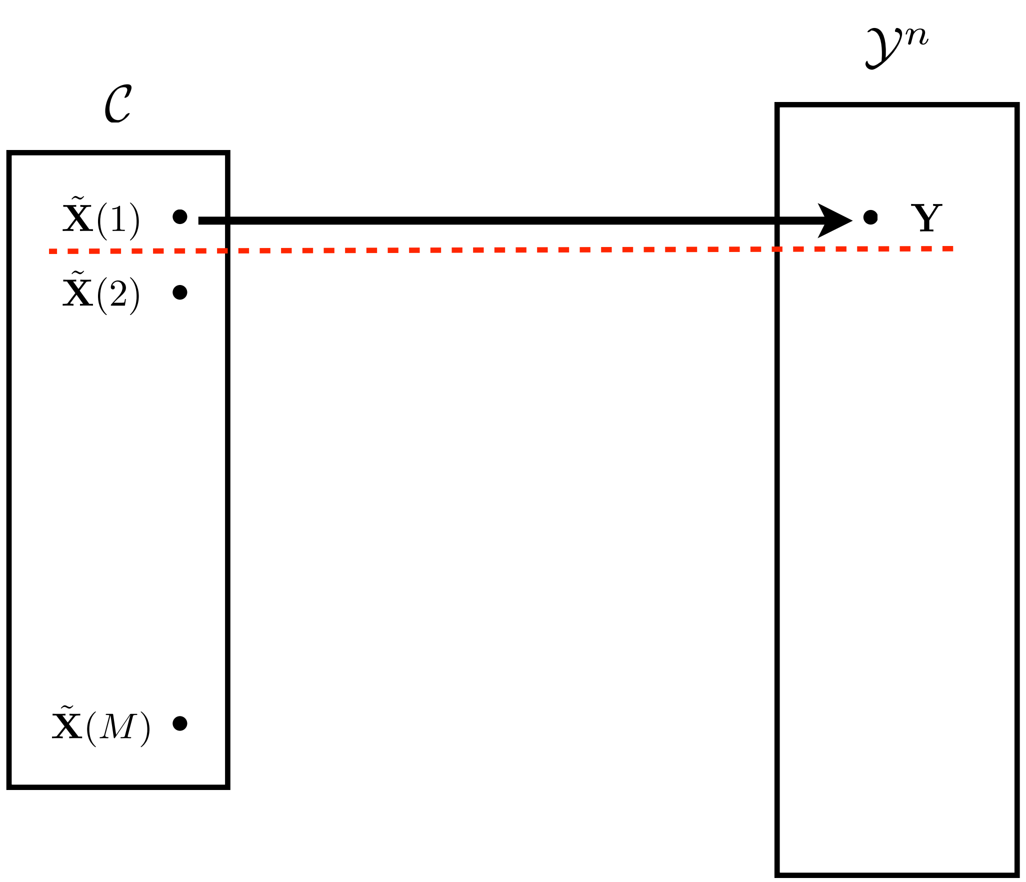
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

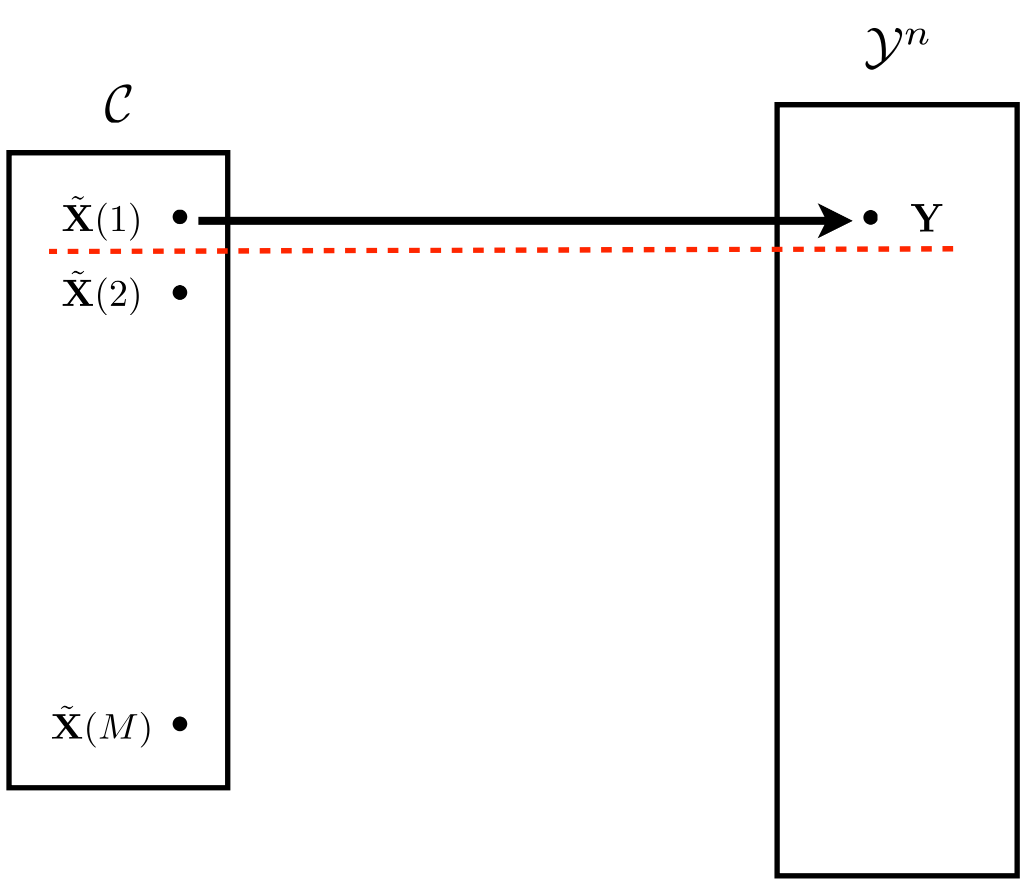
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

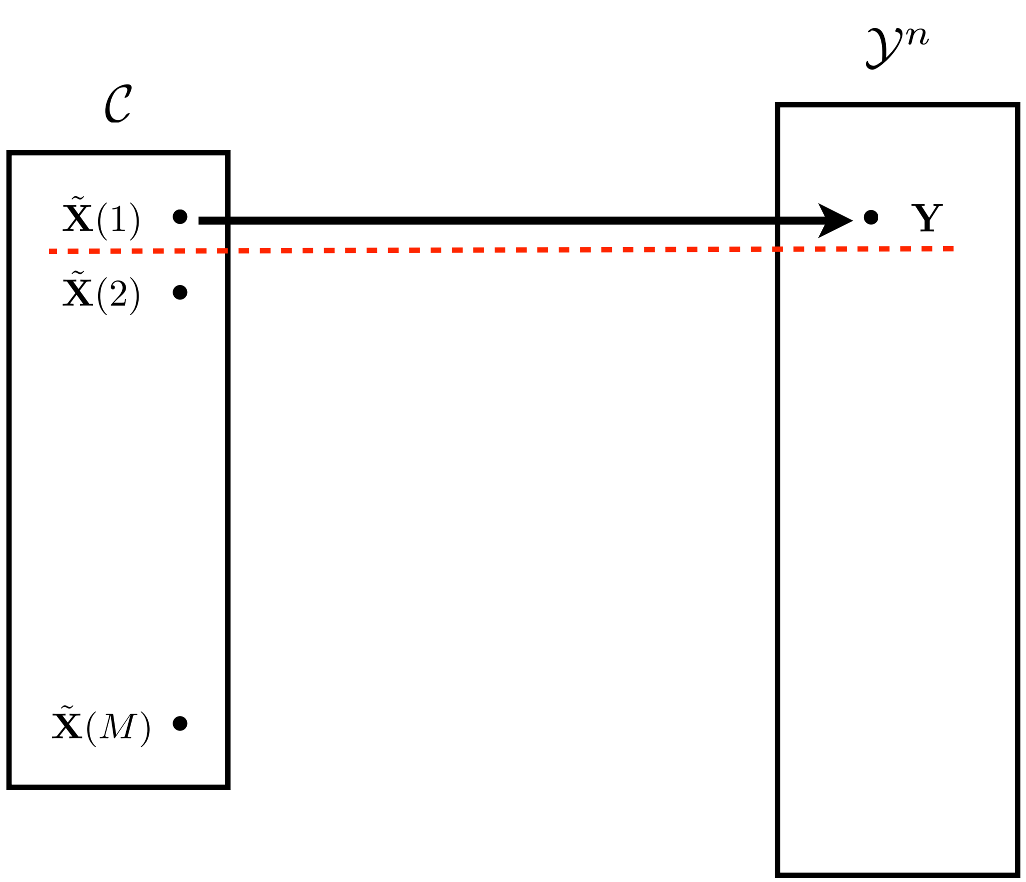
6. By the union bound,

$$\Pr\{Err|W=1\} \leq \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \leq w \leq M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

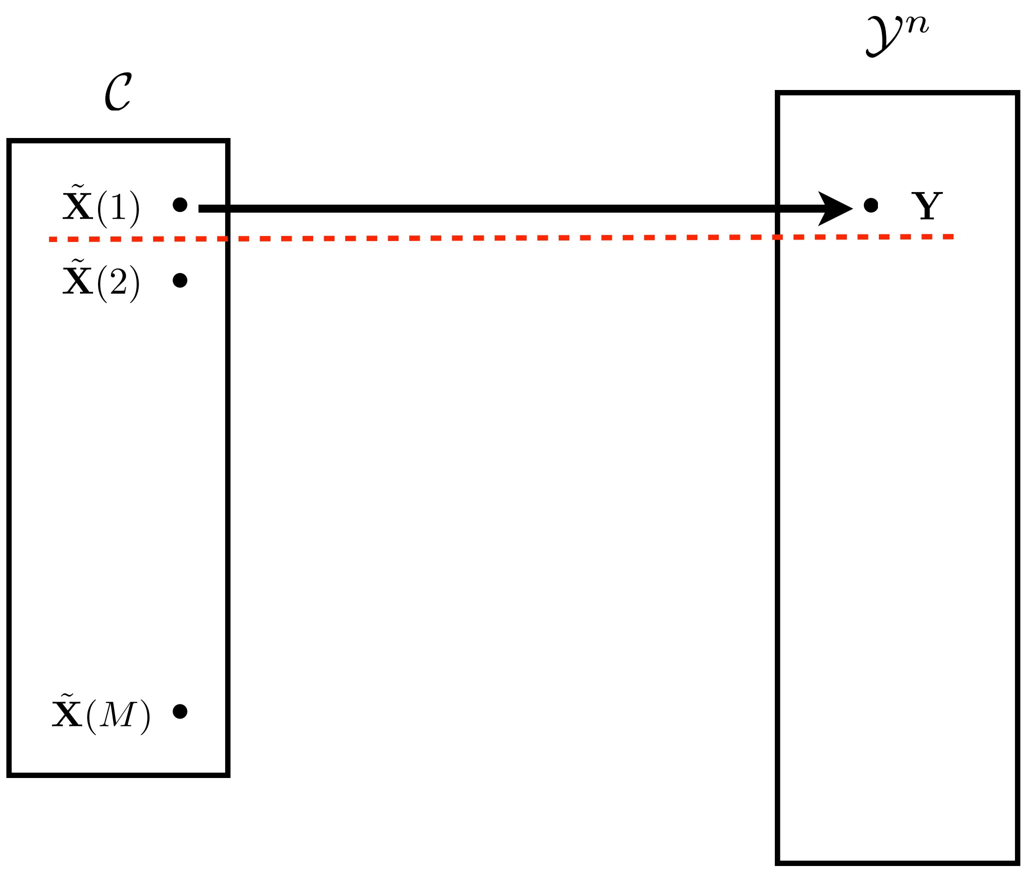
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\Pr\{Err\} < \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

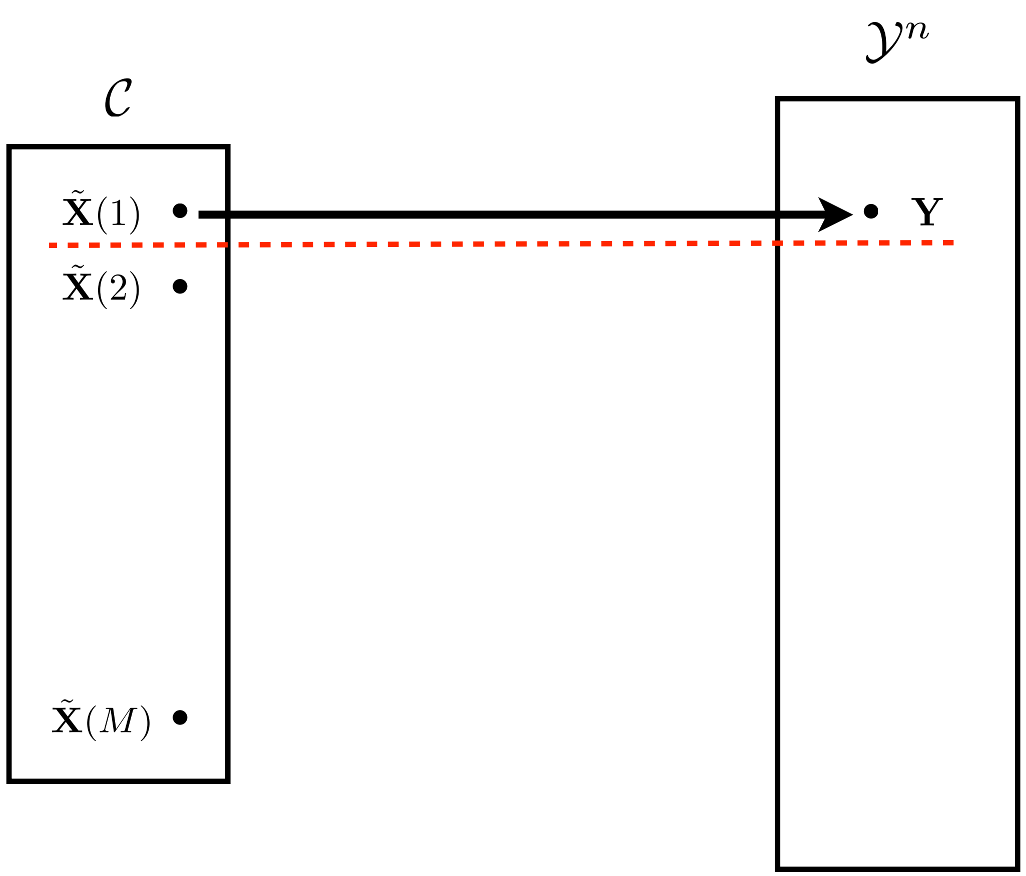
6. By the union bound,

$$\Pr\{Err|W=1\} \le \Pr\{E_1^c|W=1\} + \sum_{w=2}^{M} \Pr\{E_w|W=1\}.$$

7. By strong JAEP,

$$\Pr\{E_1^c|W=1\} = \Pr\{(\tilde{\mathbf{X}}(1),\mathbf{Y}) \notin T^n_{[XY]\delta}|W=1\} < \nu.$$

8. Conditioning on $\{W=1\}$, for $2 \le w \le M$, $(\tilde{\mathbf{X}}(w),\mathbf{Y})$ are n i.i.d. copies of the pair of generic random variables (X',Y'), where $X' \sim X$ and $Y' \sim Y$.



9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned}\Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w),\mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\le 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned}\Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}.\end{aligned}$$

Lemma 7.17

$$\Pr\{(\mathbf{X}',\mathbf{Y}') \in T^n_{[XY]\delta}\} \le 2^{-n(I(X;Y)-\tau)}.$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w \mid W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} \mid W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned}\Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}.\end{aligned}$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned}\Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)}\end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned}\Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}.\end{aligned}$$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w \mid W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} \mid W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w|W=1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta}|W=1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n}\log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

15. Let $\nu < \frac{\epsilon}{3}$ to obtain

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \leq w \leq M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\leq 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y) - \frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

$< \frac{\epsilon}{3}$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

15. Let $\nu < \frac{\epsilon}{3}$ to obtain

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\le 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

$< \frac{\epsilon}{3}$ $< \frac{\epsilon}{6}$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

15. Let $\nu < \frac{\epsilon}{3}$ to obtain

Performance Analysis

9. Since a DMC is memoryless, X' and Y' are independent because $\tilde{\mathbf{X}}(1)$ and $\tilde{\mathbf{X}}(w)$ are independent and the generation of $\mathbf{Y}$ depends only on $\tilde{\mathbf{X}}(1)$. See textbook for a formal proof.

10. For $2 \le w \le M$,

$$\begin{aligned} \Pr\{E_w | W = 1\} &= \Pr\{(\tilde{\mathbf{X}}(w), \mathbf{Y}) \in T^n_{[XY]\delta} | W = 1\} \\ &\le 2^{-n(I(X;Y)-\tau)} \end{aligned}$$

where $\tau \to 0$ as $\delta \to 0$.

11. Note that

$$\frac{1}{n} \log M < I(X;Y) - \frac{\epsilon}{4} \iff M < 2^{n(I(X;Y)-\frac{\epsilon}{4})}.$$

12. Therefore,

$$\begin{aligned} \Pr\{Err\} &< \nu + 2^{n(I(X;Y)-\frac{\epsilon}{4})} \cdot 2^{-n(I(X;Y)-\tau)} \\ &= \nu + 2^{-n(\frac{\epsilon}{4}-\tau)}. \end{aligned}$$

$< \frac{\epsilon}{3}$ $< \frac{\epsilon}{6}$

13. Recall that ϵ is fixed. Since $\tau \to 0$ as $\delta \to 0$, we can choose δ to be sufficiently small so that

$$\frac{\epsilon}{4} - \tau > 0.$$

14. Then

$$2^{-n(\frac{\epsilon}{4}-\tau)} \to 0$$

as $n \to \infty$.

15. Let $\nu < \frac{\epsilon}{3}$ to obtain

$$\Pr\{Err\} < \frac{\epsilon}{2}$$

for sufficiently large n.

Idea of Analysis

Idea of Analysis

- Let n be large.

Idea of Analysis

- Let n be large.
- $\Pr\{\tilde{\mathbf{X}}(1)$ jointly typical with $\mathbf{Y}\} \to 1$.

Idea of Analysis

- Let n be large.
- $\Pr\{\tilde{\mathbf{X}}(1)$ jointly typical with $\mathbf{Y}\} \to 1$.
- For $w \neq 1$, $\Pr\{\tilde{\mathbf{X}}(w)$ jointly typical with $\mathbf{Y}\} \approx 2^{-nI(X;Y)}$.

Idea of Analysis

- Let n be large.
- $\Pr\{\tilde{\mathbf{X}}(1)$ jointly typical with $\mathbf{Y}\} \to 1$.
- For $w \neq 1$, $\Pr\{\tilde{\mathbf{X}}(w)$ jointly typical with $\mathbf{Y}\} \approx 2^{-nI(X;Y)}$.
- If $|\mathcal{C}| = M$ grows at a rate $< I(X;Y)$, then

$$\Pr\{\tilde{\mathbf{X}}(w) \text{ jointly typical with } \mathbf{Y} \text{ for some } w \neq 1\ \}$$

can be made arbitrarily small.

Idea of Analysis

- Let n be large.
- $\Pr\{\tilde{\mathbf{X}}(1) \text{ jointly typical with } \mathbf{Y}\} \to 1$.
- For $w \neq 1$, $\Pr\{\tilde{\mathbf{X}}(w) \text{ jointly typical with } \mathbf{Y}\} \approx 2^{-nI(X;Y)}$.
- If $|\mathcal{C}| = M$ grows at a rate $< I(X;Y)$, then

  $$\Pr\{\tilde{\mathbf{X}}(w) \text{ jointly typical with } \mathbf{Y} \text{ for some } w \neq 1\ \}$$

  can be made arbitrarily small.
- Then $\Pr\{\hat{W} \neq W\}$ can be made arbitrarily small.

Existence of Deterministic Code

Existence of Deterministic Code

- According to the random coding scheme,

$$\Pr\{Err\} = \sum_{\mathcal{C}} \Pr\{\mathcal{C}\}\Pr\{Err|\mathcal{C}\}.$$

Existence of Deterministic Code

- According to the random coding scheme,

$$\Pr\{Err\} = \sum_{\mathcal{C}} \Pr\{\mathcal{C}\}\Pr\{Err|\mathcal{C}\}.$$

- Then there exists at least one codebook $\mathcal{C}^*$ such that

$$P_e = \Pr\{Err|\mathcal{C}^*\} \le \Pr\{Err\} < \frac{\epsilon}{2}.$$

Existence of Deterministic Code

- According to the random coding scheme,

$$\Pr\{Err\} = \sum_{\mathcal{C}} \Pr\{\mathcal{C}\} \Pr\{Err|\mathcal{C}\}.$$

- Then there exists at least one codebook $\mathcal{C}^*$ such that

$$P_e = \Pr\{Err|\mathcal{C}^*\} \leq \Pr\{Err\} < \frac{\epsilon}{2}.$$

- By construction, this codebook has rate

$$\frac{1}{n} \log M > I(X;Y) - \frac{\epsilon}{2}.$$

Code with $\lambda_{max} < \varepsilon$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2}$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w \leq \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M} \lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M} \lambda_w \leq \sum_{w=1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M} \lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M} \lambda_w \leq \sum_{w=1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w \leq \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M}\lambda_w < \epsilon.$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M}\lambda_w < \epsilon.$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M}\lambda_w < \epsilon.$$

- Hence,

$$\lambda_{M/2+1} < \epsilon$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M}\lambda_w < \epsilon.$$

- Hence,

$$\lambda_{M/2+1} < \epsilon \implies \lambda_{M/2} < \epsilon.$$

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M}\lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M}\lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M}\lambda_w < \epsilon.$$

- Hence,

$$\lambda_{M/2+1} < \epsilon \implies \lambda_{M/2} < \epsilon.$$

- Conclusion: If $P_e < \epsilon/2$, then $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{M/2} < \epsilon$.

Code with $\lambda_{max} < \varepsilon$

- We want a code with $\lambda_{max} < \epsilon$, not just $P_e < \epsilon/2$.
- Without loss of generality, assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$. Consider

$$\frac{1}{M}\sum_{w=1}^{M} \lambda_w < \frac{\epsilon}{2} \iff \sum_{w=1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon.$$

- Since M is even, $M/2$ is an integer. Then

$$\sum_{w=M/2+1}^{M} \lambda_w < \left(\frac{M}{2}\right)\epsilon \implies \frac{1}{M/2}\sum_{w=M/2+1}^{M} \lambda_w < \epsilon.$$

- Hence,

$$\lambda_{M/2+1} < \epsilon \implies \lambda_{M/2} < \epsilon.$$

- Conclusion: If $P_e < \epsilon/2$, then $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{M/2} < \epsilon$.
- Discard the worst half of the codewords in $\mathcal{C}^*$ to achieve $\lambda_{max} < \epsilon$.

- After discarding the worse half of $\mathcal{C}^*$, the rate of the code becomes

$$\begin{aligned}\frac{1}{n}\log\frac{M}{2} &= \frac{1}{n}\log M - \frac{1}{n}\\ &> \left(I(X;Y) - \frac{\epsilon}{2}\right) - \frac{1}{n}\\ &> I(X;Y) - \epsilon\end{aligned}$$

for sufficiently large n.

- After discarding the worse half of $\mathcal{C}^*$, the rate of the code becomes

$$\begin{aligned}\frac{1}{n}\log\frac{M}{2} &= \frac{1}{n}\log M - \frac{1}{n} \\ &> \left(I(X;Y) - \frac{\epsilon}{2}\right) - \frac{1}{n} \\ &> I(X;Y) - \epsilon\end{aligned}$$

  for sufficiently large n.

- Here we assume that the decoding function is unchanged, so that deletion of worst half of the codewords does not affect the conditional probabilities of error of the remaining codewords.