Strong and epsilon-Dependent Converses for Source Coding, Channel Coding, and Hypothesis Testing

Michèle Wigger
Telecom Paris, IP Paris

Joint work with Sara Faour, Mustapha Hamad, Sadaf Salehkalaibar, and Mireille Sarkiss

ETH Zurich
17 November 2022

Strong and ϵ-Dependent Converses

- In IT we typically have constraints of the form:

  A performance criteria (an expectation) needs to be below $\epsilon > 0$ asymptotically as $n \to \infty$

- Strong converse: Show that ultimate limits of other performance measures (rate etc.) do not depend on $\epsilon > 0$.

- If other performances depend on ϵ, we call it an ϵ-dependent converse

Message and Outline of the Talk

Take-Away Message

Converse proofs based on a “typical” change of measure
(inspired by Guo-Effros ’09 and Tyagi-Watanabe ’19)
and asymptotic Markov chains

In this talk:

- Source Coding
- Distributed Hypothesis Testing
- Channel Coding

Change of Measure on the Typical Set and a Useful Lemma

- $\{(X_i, Y_i)\}$ i.i.d. $\sim P_{XY}$ and independent thereof $T \sim \mathcal{U}\{1, \ldots, n\}$

- Strongly typical set $\mathcal{T}^{(n)}(P_{XY})$

$$\left\{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n \colon |\pi_{x^n y^n}(a,b) - P_{XY}(a,b)| < n^{-1/3}\right\}.$$

- $\mathcal{D}_n \subseteq \mathcal{T}^{(n)}(P_{XY})$ so that $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}[\mathcal{D}_n] = 0.$

- Change of measure:

$$(\tilde{X}^n, \tilde{Y}^n) \sim P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n) = \frac{P_{XY}^{\otimes}(x^n, y^n)}{\mathbb{P}[\mathcal{D}_n]} \cdot \mathbb{1}\{(x^n, y^n) \in \mathcal{D}_n\}$$

Lemma

$$\left|P_{\tilde{X}_T \tilde{Y}_T} - P_{XY}\right| \to 0$$

$$\left|\frac{1}{n} H(\tilde{X}^n \tilde{Y}^n) - H(\tilde{X}_T \tilde{Y}_T)\right| \to 0, \quad \left|\frac{1}{n} H(\tilde{X}^n | \tilde{Y}^n) - H(\tilde{X}_T | \tilde{Y}_T)\right| \to 0$$

Proof of the Lemma

1. $P_{\tilde{X}_T \tilde{Y}_T}(x, y) - P_{XY}(x, y) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$

   Proof:

$$\begin{aligned} P_{\tilde{X}_T \tilde{Y}_T}(x, y) &= \frac{1}{n}\sum_{t=1}^{n} P_{\tilde{X}_t \tilde{Y}_t}(x, y) = \mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{\tilde{X}_t = x, \tilde{Y}_t = y\}\right] \\ &= \underbrace{\mathbb{E}[\pi_{\tilde{X}^n \tilde{Y}^n}(x, y)]}_{=P_{XY}(x,y) \pm n^{-1/3}} \end{aligned}$$

Proof of the Lemma

1. $P_{\tilde{X}_T \tilde{Y}_T}(x, y) - P_{XY}(x, y) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$
2. $D(P_{\tilde{X}_T \tilde{Y}_T} \| P_{XY}) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$

Proof of the Lemma

1. $P_{\tilde{X}_T \tilde{Y}_T}(x,y) - P_{XY}(x,y) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$

2. $D(P_{\tilde{X}_T \tilde{Y}_T} \| P_{XY}) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$

3. $\frac{1}{n} D(P_{\tilde{X}^n \tilde{Y}^n} \| P_{XY}^{\otimes n}) \to 0$ because $\frac{1}{n} \log \mathbb{P}[\mathcal{D}_n] \to 0$

   Proof:

$$
\begin{aligned}
0 &\leq \frac{1}{n} D(P_{\tilde{X}^n \tilde{Y}^n} \| P_{XY}^{\otimes n}) = \frac{1}{n} \sum_{(x^n, y^n) \in \mathcal{D}_n} P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n) \log \frac{P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n)}{P_{XY}^{\otimes n}(x, y)} \\
&= -\frac{1}{n} \sum_{(x^n, y^n) \in \mathcal{D}_n} P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n) \log \mathbb{P}[\mathcal{D}_n] = -\frac{1}{n} \log \mathbb{P}[\mathcal{D}_n] \to 0.
\end{aligned}
$$

Proof of the Lemma

1. $P_{\tilde{X}_T \tilde{Y}_T}(x, y) - P_{XY}(x, y) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$
2. $D(P_{\tilde{X}_T \tilde{Y}_T} \| P_{XY}) \to 0$ because $\mathcal{D}_n \subseteq \mathcal{T}(P_{XY})$
3. $\frac{1}{n} D(P_{\tilde{X}^n \tilde{Y}^n} \| P_{XY}^{\otimes n}) \to 0$ because $\frac{1}{n} \log \mathbb{P}[\mathcal{D}_n] \to 0$
4. $\frac{1}{n} H(\tilde{X}^n \tilde{Y}^n) + \frac{1}{n} D(P_{\tilde{X}^n \tilde{Y}^n} \| P_{XY}^{\otimes n}) = H(\tilde{X}_T \tilde{Y}_T) + D(P_{\tilde{X}_T \tilde{Y}_T} \| P_{XY})$

Back to the Useful Lemma

- (X_i, Y_i) i.i.d. $\sim P_{XY}$ and independent thereof $T \sim \mathcal{U}\{1, \ldots, n\}$

- Strongly typical set $\mathcal{T}^{(n)}(P_{XY})$
$$\left\{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n \colon |\pi_{x^n y^n}(a,b) - P_{XY}(a,b)| < n^{-1/3}\right\}.$$

- $\mathcal{D}_n \subseteq \mathcal{T}^{(n)}(P_{XY})$ so that $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}[\mathcal{D}_n] = 0.$

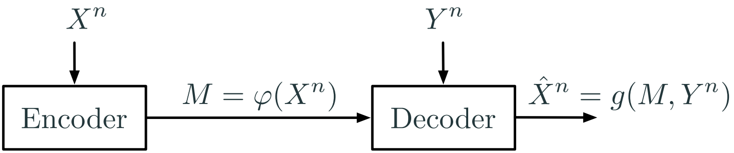
- Change of measure:
$$(\tilde{X}^n, \tilde{Y}^n) \sim P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n) = \frac{P_{XY}^{\otimes}(x^n, y^n)}{\mathbb{P}[\mathcal{D}_n]} \cdot \mathbb{1}\{(x^n, y^n) \in \mathcal{D}_n\}$$

Lemma

$$\frac{1}{n} H(\tilde{X}^n \tilde{Y}^n) \to H(\tilde{X}_T \tilde{Y}_T), \qquad \frac{1}{n} H(\tilde{X}^n | \tilde{Y}^n) \to H(\tilde{X}_T | \tilde{Y}_T), \quad P_{\tilde{X}_T \tilde{Y}_T} \to P_{XY}$$

Lossless Source Coding

Almost Lossless Source Coding with Side-Information



- $\{(X_i, Y_i)\}$ i.i.d. $\sim P_{XY}$
- $M \in \{1, \ldots, 2^{nR}\}$
- Rate $R > 0$ is $\epsilon > 0$-*achievable* if $\overline{\lim}_{n\to\infty} \mathbb{P}\left[X^n \neq \hat{X}^n\right] \leq \epsilon$.

Theorem (Slepian-Wolf '83, Oohama-Han '94)

Given $\epsilon \in [0, 1)$, rates $R > H(X|Y)$ are ϵ-achievable and rates $R < H(X|Y)$ not.

Strong Converse Proof Irrespective of $\epsilon \in [0, 1)$

- $\mathcal{D}_n := \left\{(x^n, y^n) \in \mathcal{T}^{(n)}(P_{XY}) : \quad g\left(\varphi(x^n), y^n\right) = x^n\right\}$ (no error!)

  with $\mathbb{P}[\mathcal{D}_n] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

- Change of measure:

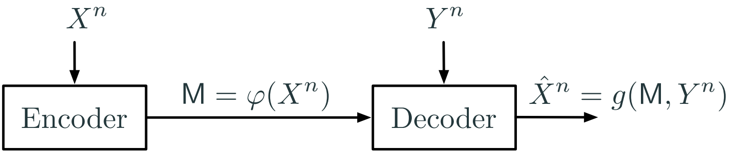
$$(\tilde{X}^n, \tilde{Y}^n) \sim P_{\tilde{X}^n \tilde{Y}^n}(x^n, y^n) = \frac{P_{XY}^{\otimes}(x^n, y^n)}{\mathbb{P}[\mathcal{D}_n]} \cdot \mathbb{1}\{(x^n, y^n) \in \mathcal{D}_n\}$$

  and $\tilde{M} = \varphi(\tilde{X}^n)$

- Rate bound:

$$R \geq \frac{1}{n} H(\tilde{\mathrm{M}}) \geq \frac{1}{n} H(\tilde{\mathrm{M}} | \tilde{Y}^n) \geq \frac{1}{n} H(\tilde{X}^n | \tilde{Y}^n) \to H(X|Y)$$

Lossless Source Coding under Variable-Length Coding



- (X^n, Y^n) i.i.d. $\sim P_{XY}$
- $\mathsf{M} \in \{0,1\}^\star$ and $\mathbb{E}[\mathrm{len}(\mathsf{M})] \leq nR$
- Rate $R > 0$ is $\epsilon > 0$-*achievable* if $\overline{\lim}_{n\to\infty} \mathbb{P}\left[X^n \neq \hat{X}^n\right] \leq \epsilon$.

Theorem

Given $\epsilon \in [0,1)$. Rates $R > (1-\epsilon)H(X|Y)$ are ϵ-achievable and rates $R < (1-\epsilon)H(X|Y)$ not.

- Achievability: With probability ϵ send a dummy bit for M

ϵ-Dependent Converse Proof

- $\mathcal{D}_n := \left\{(x^n, y^n) \in \mathcal{T}^{(n)}(P_{XY}) : \quad g\left(\varphi(x^n), y^n\right) = x^n\right\}$ (no error!)

  with $\mathbb{P}[\mathcal{D}] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

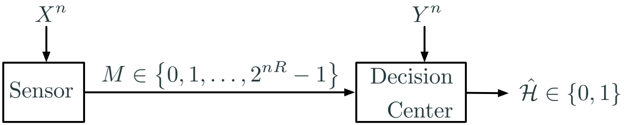
- Change of measure $(\tilde{X}^n, \tilde{Y}^n)$ on $\mathcal{D}_n$ and $\tilde{M} = \varphi(\tilde{X}^n)$

- Proof steps to incorporate expected rate constraint:

$$
\begin{aligned}
R &\geq \frac{1}{n}\mathbb{E}[\mathrm{len}(\mathrm{M})] \geq \frac{1}{n}\mathbb{E}[\mathrm{len}(\tilde{\mathrm{M}})] \cdot \mathbb{P}[\mathcal{D}_n] \geq \frac{1}{n}H(\tilde{\mathrm{M}}|\mathrm{len}(\tilde{\mathrm{M}})) \cdot \mathbb{P}[\mathcal{D}_n] \\
&= \frac{1}{n}H(\tilde{\mathrm{M}}) \cdot \mathbb{P}[\mathcal{D}_n] - \underbrace{\frac{1}{n}H(\mathrm{len}(\tilde{\mathrm{M}}))}_{\to 0}\mathbb{P}[\mathcal{D}_n] \\
&\geq \frac{1}{n}H(\tilde{X}^n|\tilde{Y}^n) \cdot \mathbb{P}[\mathcal{D}_n] - o(1) \to H(X|Y) \cdot (1 - \epsilon)
\end{aligned}
$$

Distributed Hypothesis Testing

Testing Against Independence



- $\mathcal{H} = 0 :\ (X^n, Y^n) \sim$ i.i.d. P_{XY}
- $\mathcal{H} = 1 :\ (X^n, Y^n) \sim$ i.i.d. $P_X P_Y$
- Type-I error $\overline{\lim_{n\to\infty}} \mathbb{P}[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0] \leq \epsilon$
- Type-II error exponent $\theta = - \overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1]$

Largest Exponent independent of $\epsilon \in [0, 1)$ (Ahlswede-Csiszár 86)

$$\theta^*_\epsilon(R) = \max_{\substack{P_{S|X}: \\ R \geq I(S;X)}} I(S; Y)$$

Strong Converse: Change of Measure And Rate Constraint

- $\mathcal{D}_n := \left\{(x^n, y^n) \in \mathcal{T}^{(n)}(P_{XY}) \colon \quad g\left(\varphi(x^n), y^n\right) = 0\right\} \qquad \to \hat{\mathcal{H}} = 0$

  $P_{XY}^{\otimes}[\mathcal{D}_n] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

- Change of measure: $(\tilde{X}^n, \tilde{Y}^n)$ on $\mathcal{D}_n$ and $\tilde{M} = \varphi(\tilde{X}^n)$

- Proof steps for the rate:

$$
\begin{aligned}
R &\geq \frac{1}{n} H(\tilde{M}) = \frac{1}{n} I(\tilde{M}; \tilde{X}^n, \tilde{Y}^n) = \frac{1}{n} H(\tilde{X}^n \tilde{Y}^n) - \frac{1}{n} H(\tilde{X}^n \tilde{Y}^n | \tilde{M}) \\
&= H(\tilde{X}_T \tilde{Y}_T) + o(1) - \frac{1}{n} \sum_{t=1}^{n} H(\tilde{X}_t \tilde{Y}_t | \tilde{X}^{t-1} \tilde{Y}^{t-1} \tilde{M}) \\
&= H(\tilde{X}_T \tilde{Y}_T) + o(1) - H(\tilde{X}_T \tilde{Y}_T | \tilde{X}^{T-1} \tilde{Y}^{T-1} \tilde{M} T) \\
&= I(\tilde{X}_T \tilde{Y}_T; \underbrace{\tilde{X}^{T-1} \tilde{Y}^{T-1} \tilde{M} T}_{=:S}) \geq I(\tilde{X}_T; S)
\end{aligned}
$$

Strong Converse: Exponent

- $\hat{\mathcal{H}} = \varphi(M, Y^n)$ and $\tilde{\mathcal{H}} = \varphi(\tilde{M}, \tilde{Y}^n) = 0$
- Interesting inequality

$$D(P_{\tilde{Y}^n\tilde{M}} \| P_{\tilde{Y}^n} P_{\tilde{M}}) \geq D(P_{\tilde{Y}^n\tilde{M}}(\tilde{\mathcal{H}}) \| P_{\tilde{Y}^n} P_{\tilde{M}}(\tilde{\mathcal{H}})) = 1 \cdot \log \frac{1}{P_{\tilde{Y}^n} P_{\tilde{M}}(\tilde{\mathcal{H}} = 0)}$$

- Proof steps for exponent

$$\begin{aligned}
& -\frac{1}{n} \log P_{Y^n} P_M(\hat{\mathcal{H}} = 0) \\
\leq \quad & -\frac{1}{n} \log P_{\tilde{Y}^n} P_{\tilde{M}}(\tilde{\mathcal{H}} = 0) - \frac{2}{n} \log P_{XY}^{\otimes n}[\mathcal{D}_n] \\
\leq \quad & \frac{1}{n} D(P_{\tilde{Y}^n\tilde{M}} \| P_{\tilde{Y}^n} P_{\tilde{M}}) + o(1) = \frac{1}{n} I(\tilde{M}; \tilde{Y}^n) + o(1) \\
\leq \quad & \frac{1}{n} H(\tilde{Y}^n) - H(\tilde{Y}_T | \underbrace{\tilde{M}\tilde{X}^{T-1}\tilde{Y}^{T-1}T}_{=S}) + o(1) \\
= \quad & H(\tilde{Y}_T) + o(1) - H(\tilde{Y}_T | S) \to I(\tilde{Y}_T; S)
\end{aligned}$$

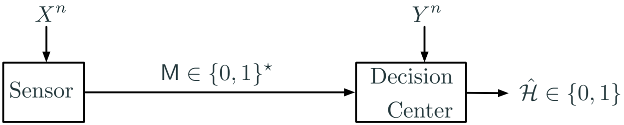
Asymptotic Markov Chain

- In the theorem we need asymptotic Markov chain $S \to \tilde{X}_T \to \tilde{Y}_T$
- For any n:

$$
\begin{aligned}
0 &= \frac{1}{n} I(\tilde{M}; \tilde{Y}^n | \tilde{X}^n) \\
&= \frac{1}{n} H(\tilde{Y}^n | \tilde{X}^n) - \frac{1}{n} H(\tilde{Y}^n | \tilde{X}^n \tilde{M}) \\
&\geq H(\tilde{Y}_T | \tilde{X}_T) + o(1) - H(\tilde{Y}_T | \tilde{X}_T \underbrace{\tilde{X}^{T-1} \tilde{Y}^{T-1} \tilde{M} T}_{S}) \\
&= I(\tilde{Y}_T; S | \tilde{X}_T) + o(1) \geq o(1).
\end{aligned}
$$

- So $I(\tilde{Y}_T; S | \tilde{X}_T) \to 0$ and the Markov chain holds asymptotically.

Testing Against Indep. Under Variable-Length Coding



- $\mathcal{H} = 0: \ (X^n, Y^n) \sim$ i.i.d. P_{XY}
- $\mathcal{H} = 1: \ (X^n, Y^n) \sim$ i.i.d. $P_X P_Y$
- Expected rate constraints $\mathbb{E}[\mathrm{len}(\mathsf{M})] \leq nR$

Largest Possible Error Exponent depends on ϵ

$$\theta^*_{\mathsf{VL},\epsilon}(R) = \max_{\substack{P_{S|X}: \\ R \geq (1-\epsilon) I(S;X)}} I(S;Y)$$

- Achievability: With probability ϵ send a dummy bit for M

Change of Measure And Rate Constraint as Before!

- $\mathcal{D}_n := \left\{(x^n, y^n) \in \mathcal{T}^{(n)}(P_{XY}) : \quad g\left(\varphi(x^n), y^n\right) = 0\right\}$
  no error under $\mathcal{H} = 0! \Rightarrow P_{XY}^{\otimes n}[\mathcal{D}_n] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

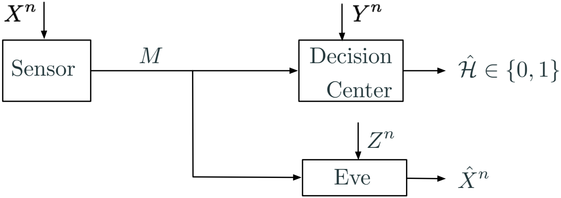
- Change of measure $(\tilde{X}^n, \tilde{Y}^n, \tilde{M})$

- Proof steps for the rate

$$
\begin{aligned}
R &\geq \frac{1}{n}\mathbb{E}[\mathrm{len}(\mathrm{M})] \geq \frac{1}{n}\mathbb{E}[\mathrm{len}(\tilde{\mathrm{M}})] \cdot P_{XY}^{\otimes n}[\mathcal{D}_n] \geq \frac{1}{n}H(\tilde{\mathrm{M}}|\mathrm{len}(\tilde{\mathrm{M}})) \cdot P_{XY}^{\otimes n}[\mathcal{D}_n] \\
&= \frac{1}{n}H(\tilde{\mathrm{M}}) \cdot P_{XY}^{\otimes n}[\mathcal{D}_n]] - \underbrace{\frac{1}{n}H(\mathrm{len}(\tilde{\mathrm{M}}))}_{\to 0} P_{XY}^{\otimes n}[\mathcal{D}_n] \\
&\vdots \\
&\geq I(\tilde{X}_T; S) \cdot P_{XY}^{\otimes n}[\mathcal{D}_n] + o(1) \to I(\tilde{X}_T; S)(1 - \epsilon)
\end{aligned}
$$

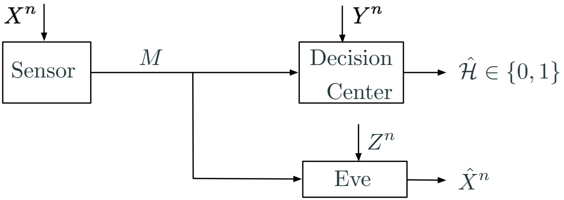
- Remaining steps as before!

Testing Against Independence with an Eavesdropper



- $\mathcal{H} = 0$: $(X^n, Y^n, Z^n) \sim$ i.i.d. P_{XYZ}
- $\mathcal{H} = 1$: $(X^n, Y^n, Z^n) \sim$ i.i.d. $P_X P_Y P_{Z|XY}$
- Equivocation constraint $\overline{\lim}_{n\to\infty} \frac{1}{n} H(X^n | M Z^n \mathcal{H} = 0) \geq \Delta$

Testing Against Independence with an Eavesdropper



- $\mathcal{H} = 0 :\ (X^n, Y^n, Z^n) \sim$ i.i.d. P_{XYZ}
- $\mathcal{H} = 1 :\ (X^n, Y^n, Z^n) \sim$ i.i.d. $P_X P_Y P_{Z|XY}$
- Equivocation constraint $\overline{\lim}_{n\to\infty} \frac{1}{n} H(X^n | M Z^n \mathcal{H} = 0) \geq \Delta$

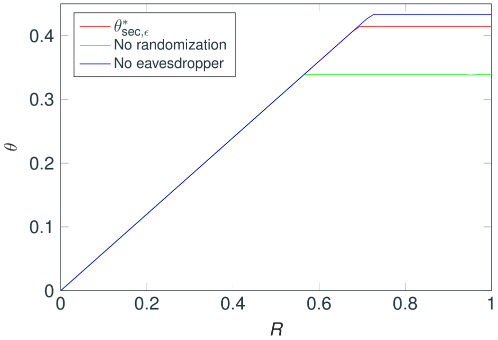
Largest Possible Error Exponent depends on ϵ

$$\theta^*_{\text{sec},\epsilon}(R) = \max_{\substack{P_{S|X}: \\ R \geq I(S;X) \\ \Delta \leq (1-\epsilon) H(X|ZS) + \epsilon H(X|Z)}} I(S;Y)$$

- Achievability: With probability ϵ send a dummy bit for M

An Example

- $X \sim \mathcal{B}(0.2)$
- $P_{Y|X}$ a BEC(0.4)
- $P_{Z|X}$ a BSC(0.2)
- $\epsilon = 0.2$
- $\Delta = 0.13$



Converse Proof

- $E = \mathbb{1}\{(X^n, Y^n) \in \mathcal{D}_n\}$

- Equivocation bound:

$$\frac{1}{n}H(X^n|M, W^n) = \sum_{t=1}^{n} H(X_t|X^{t-1}, M, W^n)$$

$$\begin{aligned}
&= \frac{1}{n}\sum_{t=1}^{n} H(X_t|X^{t-1}, Y^{t-1}, M, W^n) \\
&\leq \frac{1}{n}\sum_{t=1}^{n} H(X_t|X^{t-1}, Y^{t-1}, M, W^n, E) \\
&\quad + \underbrace{\frac{1}{n}\sum_{t=1}^{n} I(E; X_t|X^{t-1}, Y^{t-1}, M, W^n)}_{\leq 1} \\
&\leq H(\tilde{X}_T|S, \tilde{W}_T)P_E(1) + H(X_T|W_T, E = 0)P_E(0) + o(1)
\end{aligned}$$

Channel Coding

Capacity of Discrete Memoryless Channels

- M uniform over $\{1, \ldots, 2^{nR}\}$
- Memoryless channel $P_{Y^n|X^n} = P_{Y|X}^{\otimes n}$
- $R > 0$ is $\epsilon > 0$-*achievable* if $\overline{\lim}_{n \to \infty} \max_m \mathbb{P}\left[\hat{M} \neq M \middle| M = m\right] \leq \epsilon$.
- Capacity $C \triangleq \max_{P_X} I(X; Y)$

Theorem (Wolfowitz '78)

For any $\epsilon \in [0, 1)$, rates $R < C$ are ϵ-achievable and rates $R > C$ not.

Strong Converse Proof Irrespective of $\epsilon \in [0,1)$)

- $\mathcal{T}^{(n)}(P_{Y|X}, x^n) = \left\{y^n \colon |\pi_{x^n,y^n}(x,y) - \pi_{x^n}(x)P_{Y|X}(y)| < n^{-1/3}\right\}$
- $\mathcal{D}_m := \left\{y^n \in \mathcal{T}^{(n)}(P_{Y|X}, x^n(m)) \colon \quad g^{(n)}(y^n) = m\right\}$

  no decoding error when $M = m \Rightarrow \mathbb{P}[\mathcal{D}_m | M = m] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

- Change of measure:

  $$P_{\tilde{X}^n\tilde{Y}^n\tilde{M}}(x^n, y^n, m) = \frac{1}{2^{nR}} \cdot \mathbb{1}\{x^n = \varphi(m)\} \frac{P_{Y|X}^{\otimes n}(y^n|x^n)}{\mathbb{P}[\mathcal{D}_m | M = m]} \cdot \mathbb{1}\{y^n \in \mathcal{D}_m\}$$

Strong Converse Proof Irrespective of $\epsilon \in [0,1)$)

- $\mathcal{T}^{(n)}(P_{Y|X}, x^n) = \left\{ y^n \colon |\pi_{x^n,y^n}(x,y) - \pi_{x^n}(x)P_{Y|X}(y)| < n^{-1/3} \right\}$

- $\mathcal{D}_m := \left\{ y^n \in \mathcal{T}^{(n)}(P_{Y|X}, x^n(m)) \colon \quad g^{(n)}(y^n) = m \right\}$

  no decoding error when $M = m \Rightarrow \mathbb{P}[\mathcal{D}_m | M = m] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

- Change of measure:

  $$P_{\tilde{X}^n \tilde{Y}^n \tilde{M}}(x^n, y^n, m) = \frac{1}{2^{nR}} \cdot \mathbb{1}\{x^n = \varphi(m)\} \frac{P_{Y|X}^{\otimes n}(y^n|x^n)}{\mathbb{P}[\mathcal{D}_m | M = m]} \cdot \mathbb{1}\{y^n \in \mathcal{D}_m\}$$

- Rate bound: $\quad R = \frac{1}{n} H(\tilde{M}) = \frac{1}{n} I(\tilde{M}; \tilde{Y}^n) \leq H(\tilde{Y}_T) - \frac{1}{n} H(\tilde{Y}^n | \tilde{M})$

Strong Converse Proof Irrespective of $\epsilon \in [0,1))$

- $\mathcal{T}^{(n)}(P_{Y|X}, x^n) = \left\{y^n \colon |\pi_{x^n,y^n}(x,y) - \pi_{x^n}(x)P_{Y|X}(y)| < n^{-1/3}\right\}$

- $\mathcal{D}_m := \left\{y^n \in \mathcal{T}^{(n)}(P_{Y|X}, x^n(m)) \colon \quad g^{(n)}(y^n) = m\right\}$

  no decoding error when $M = m \Rightarrow \mathbb{P}[\mathcal{D}_m | M = m] \geq 1 - \epsilon - \frac{|\mathcal{X}||\mathcal{Y}|}{4n^{1/3}}$

- Change of measure:

  $$P_{\tilde{X}^n\tilde{Y}^n\tilde{M}}(x^n, y^n, m) = \frac{1}{2^{nR}} \cdot \mathbb{1}\{x^n = \varphi(m)\} \frac{P_{Y|X}^{\otimes n}(y^n|x^n)}{\mathbb{P}[\mathcal{D}_m|M=m]} \cdot \mathbb{1}\{y^n \in \mathcal{D}_m\}$$

- Rate bound: $\quad R = \frac{1}{n}H(\tilde{M}) = \frac{1}{n}I(\tilde{M}; \tilde{Y}^n) \leq H(\tilde{Y}_T) - \frac{1}{n}H(\tilde{Y}^n|\tilde{M})$

- To show:

  $$\frac{1}{n}H(\tilde{Y}^n|\tilde{M}) \quad \to \quad H_{P_X P_{Y|X}}(Y|X)$$

  $$P_{\tilde{Y}_T} \quad \to \quad \sum_x P_X(x) P_{Y|X}(y|x)$$

- Then: $R \leq H_{P_X P_{Y|X}}(Y) - H_{P_X P_{Y|X}}(Y|X) = I_{P_X P_{Y|X}}(X;Y) \leq C.$

Strong Converse for Channel Coding Continued I

$$
\begin{aligned}
P_{\tilde{Y}_T}(y) &= \frac{1}{n}\sum_{t=1}^{n} P_{\tilde{Y}_t}(y) = \mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n} \mathbb{1}\{\tilde{Y}_t = y\}\right] = \mathbb{E}\left[\pi_{\tilde{Y}^n}(y)\right] \\
&= \mathbb{E}\left[\sum_{x\in\mathcal{X}} \pi_{\tilde{Y}^n x^n(\tilde{M})}(y,x)\right] \\
&= \frac{1}{2^{nR}}\sum_{m=1}^{2^{nR}} \mathbb{E}\left[\sum_{x\in\mathcal{X}} \pi_{\tilde{Y}^n x^n(m)}(y,x)\middle|\tilde{M}=m\right] \\
&= \sum_{x\in\mathcal{X}} \underbrace{\frac{1}{2^{nR}}\sum_{m=1}^{2^{nR}} \pi_{x^n(m)}(x)}_{\to P_X(x)} \cdot \underbrace{\mathbb{E}\left[\pi_{\tilde{Y}^n|\tilde{x}^n(m)}(y|x)\middle|\tilde{M}=m\right]}_{=P_{Y|X}(y|x)\pm n^{-1/3}}.
\end{aligned}
$$

Strong Converse for Channel Coding Continued II

$$
\begin{aligned}
\frac{1}{n} H(\tilde{Y}^n | \tilde{M} = m) &= -\frac{1}{n} \sum_{y^n \in \mathcal{D}_m} P_{\tilde{Y}^n | \tilde{M} = m}(y^n) \log \frac{P_{Y|X}^{\otimes n}(y^n | x^n(m))}{\mathbb{P}[\mathcal{D}_m | M = m]} \\
&= -\frac{1}{n} \sum_{t=1}^{n} \sum_{y^n \in \mathcal{D}_m} P_{\tilde{Y}^n | \tilde{M} = m}(y^n) \log P_{Y|X}(y_t | x_t(m)) + \underbrace{\frac{1}{n} \log \mathbb{P}[\mathcal{D}_m | M = m]}_{\to 0} \\
&= -\frac{1}{n} \sum_{t=1}^{n} \sum_{y \in \mathcal{Y}} P_{\tilde{Y}_t | \tilde{M} = m}(y) \log P_{Y|X}(y | x_t(m)) + o(1) \\
&= -\sum_{x \in \mathcal{X}} \frac{n_x(m)}{n} \sum_{y \in \mathcal{Y}} \mathbb{E}\left[\frac{\sum_{t:\, x_t(m) = x} \mathbb{1}\{\tilde{Y}_t = y\}}{n_x(m)} \middle| \tilde{M} = m \right] \log P_{Y|X}(y|x) + o(1) \\
&= -\sum_{x \in \mathcal{X}} \pi_{x^n(m)}(x) \sum_{y \in \mathcal{Y}} \underbrace{\mathbb{E}\left[\pi_{\tilde{Y}^n | x^n(m)}(y|x) \middle| M = m \right]}_{= P_{Y|X}(y|x) \pm n^{-1/3}} \cdot \log P_{Y|X}(y|x) + o(1)
\end{aligned}
$$

- Average over $\tilde{M}$ and take $n \to \infty$.

Summary

- Strong converse proofs based on change of measure arguments and asymptotic Markov chains for source and channel coding and for hypothesis testing

- Methods well adapted also to prove ϵ-dependent converses under expectation constraints

References

- M. Hamad, M. W., M. Sarkiss, “Strong Converses using Change of Measure and Asymptotic Markov Chains,” ITW 2022.
- M. Hamad, M. Sarkiss, and M. W., “Multi-Hop Network with Multiple Decision Centers under Expected-Rate Constraints,” Arxiv.
- S. Faour, M. Hamad, M. Sarkiss, and M. Wigger, “Testing Against Independence with an Eavesdropper,” Arxiv.
- S. Salehkalaibar and M.W. “Distributed Hypothesis Testing With Variable-Length Coding,” JSAIT 2020.
- H. Tyagi and S. Watanabe, “Strong converse using change of measure arguments,” Trans. IT 2019.
- W. Gu and M. Effros, “A strong converse for a collection of network source coding problems,” ISIT 2009.

References continued

- J. C. Kieffer, “Strong converses in source coding relative to a fidelity criterion,” Trans. IT 2002.
- Y. Oohama and T. S. Han, “Universal coding for the Slepian-Wolf data compression system and the strong converse theorem,” Trans. IT 1994.
- Y. Oohama, “Exponential strong converse for source coding with side information at the decoder,” Entropy 2018.
- J. Wolfowitz, “Coding Theorems of Information Theory”. Springer 1978.
- R. Ahlswede and I. Csiszár, “Hypothesis testing with communication constraints,” Trans IT. 1986.

Proof of Equivalence of Entropy-Divergence Sums

$$
\begin{aligned}
&\frac{1}{n}H(\tilde{X}^n\tilde{Y}^n) + \frac{1}{n}D(P_{\tilde{X}^n\tilde{Y}^n}\|P_{XY}^{\otimes}) \\
&= -\frac{1}{n}\sum_{(x^n,y^n)\in\mathcal{D}_n} P_{\tilde{X}^n\tilde{Y}^n}(x^n,y^n)\log P_{XY}^{\otimes}(x^n,y^n) \\
&= -\frac{1}{n}\sum_{t=1}^{n}\sum_{(x^n,y^n)\in\mathcal{D}_n} P_{\tilde{X}^n\tilde{Y}^n}(x^n,y^n)\log P_{XY}(x_i,y_i) \\
&= -\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}}\frac{1}{n}\sum_{t=1}^{n} P_{\tilde{X}_t\tilde{Y}_t}(x,y)\log P_{XY}(x,y) \\
&= H(\tilde{X}_T\tilde{Y}_T) + D(P_{\tilde{X}_T\tilde{Y}_T}\|P_{XY})
\end{aligned}
$$