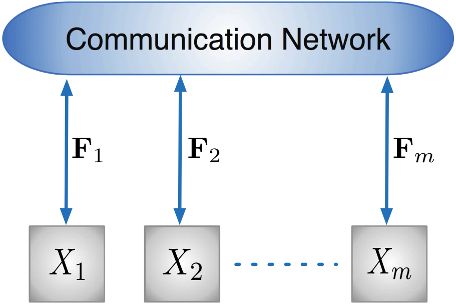
Strong Converse for a Degraded Wiretap Channel via Active Hypothesis Testing

Masahito Hayashi   Himanshu Tyagi   Shun Watanabe

Part 1: Interactive Communication

Interactive Communication



- In communication round i terminal j sends:

$$F_{ij} = F_{ij}(X_j,\ \text{prior communication})$$

- Overall communication: $\mathbf{F} = \mathbf{F}_1, ..., \mathbf{F}_m$

Common Randomness Based Converses

Common randomness (CR) is simply shared information:

A random variable L is ϵ-CR for $\mathbf{F}$ if

$$\mathrm{P}\left(L = L_i(X_i, \mathbf{F}), \quad 1 \leq i \leq m\right) \geq 1 - \epsilon$$

Converse approach for problems with interactive communication:

Bound the number of bits of CR that can be generated

Common Randomness Based Converses

Common randomness (CR) is simply shared information:

A random variable L is ϵ-CR for $\mathbf{F}$ if

$$\mathrm{P}\left(L = L_i(X_i, \mathbf{F}), \quad 1 \leq i \leq m\right) \geq 1 - \epsilon$$

Converse approach for problems with interactive communication:

Bound the number of bits of CR that can be generated

[Ahlswede-Csiszár '93, '98] Two-terminal secret key agreement

[Csiszár-Narayan '04, '08] Multiterminal secret key agreement

[T-Narayan-Gupta '10, '11] Secure computing with trusted parties

[T-Watanabe '14] General converse for information theoretic secrecy

A Property of Interactive Communication

[Csiszár-Narayan '08] (also, [Madiman-Tetali '10])

Lemma

For an interactive communication $\mathbf{F}$*, it holds that*

$$H(\mathbf{F}) \geq \sum_{B \in \mathcal{B}} \lambda_B H\left(\mathbf{F} \mid X_{B^c}\right)$$

for every fractional partition λ *of* $[m] = \{1, ..., m\}$

Here a *fractional partition* of $[m]$ refers to a set of weights λ_B s.t.

$$\sum_{B : B \ni i} \lambda_B = 1, \quad \text{for all } 1 \leq i \leq m$$

- This property does not hold for a noninteractive function F

Consequences:

- Independent observations remain so when conditioned on an interactive $\mathbf{F}$
- Extrinsic information is not less than intrinsic information

The Csiszár-Narayan CR Converse

Lemma (Recoverability Lemma)

Let L be an ϵ-CR for $\mathbf{F}$ taking values in $\mathcal{L}$. Then,

$$H(L) \leq \left[H\left(X_{\mathcal{M}}\right) - \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right)\right] + I(L \wedge \mathbf{F}) + O(\epsilon \log |\mathcal{L}|)$$

for every fractional partition λ of $\mathcal{M}$

The Csiszár-Narayan CR Converse

Lemma (Recoverability Lemma)

Let L be an ϵ-CR for $\mathbf{F}$ taking values in $\mathcal{L}$. Then,

$$H(L) \leq \left[H\left(X_{\mathcal{M}}\right) - \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right) \right] + I(L \wedge \mathbf{F}) + O(\epsilon \log |\mathcal{L}|)$$

for every fractional partition λ of $\mathcal{M}$

Measure of correlation

Leakage parameter

The Csiszár-Narayan CR Converse

Lemma (Recoverability Lemma)

Let L be an ϵ-CR for $\mathbf{F}$ taking values in $\mathcal{L}$. Then,

$$H(L) \leq \underbrace{\left[H\left(X_{\mathcal{M}}\right) - \sum_{B \in \mathcal{B}} \lambda_B H\left(X_B \mid X_{B^c}\right) \right]}_{\text{Measure of correlation}} + \underbrace{I(L \wedge \mathbf{F})}_{\text{Leakage parameter}} + O(\epsilon \log |\mathcal{L}|)$$

for every fractional partition λ of $\mathcal{M}$

Shortcoming: ϵ shows up in multiplication with $\log |\mathcal{L}|$

Conditional Independence Testing Converse

Digression: Binary Hypothesis Testing

Consider the following binary hypothesis testing problem:

$$\begin{aligned} H0: &\quad X \sim P \\ vs.& \\ H1: &\quad X \sim Q \end{aligned}$$

Define

$$\beta_\epsilon(P,Q) \triangleq \inf \sum_{x \in \mathcal{X}} Q(x)T(0|x),$$

where the $\inf$ is over all random tests $T : \mathcal{X} \to \{0,1\}$ s.t.

$$\sum_{x \in \mathcal{X}} P(x)T(1|x) \leq \epsilon$$

Conditional Independence Testing Converse

[T-Watanabe '14]

Consider $\mathcal{L}$-valued random variables $L, L_1, ..., L_m, Z$ s.t.

$$\mathrm{P}\left(L_1 = ... = L_m = L\right) \geq 1 - \epsilon$$

and let

$$\delta = \left\|\mathrm{P}_{LZ} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_Z\right\|_1$$

Theorem (Conditional Independence Testing Bound)

For any distribution Q *such that*

$$\mathrm{Q}_{L_1 \dots L_m|Z} = \prod_{i=1}^{m} \mathrm{Q}_{L_i|Z}$$

and any $\eta < 1 - \epsilon - \delta$*, it holds that*

$$\log|\mathcal{L}| \leq -\frac{1}{m-1} \log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1 \dots L_m Z}, \mathrm{Q}_{L_1 \dots L_m Z}) + \frac{m}{m-1} \log \frac{1}{\eta}$$

Application to Interactive Communication

Suppose $\mathrm{Q}_{X_1 \ldots X_m | Z} = \prod_i \mathrm{Q}_{X_i|Z}$

Then, for every interactive communication $\mathbf{F}$

$$\mathrm{Q}_{X_1 \ldots X_m | Z\mathbf{F}} = \prod \mathrm{Q}_{X_i|Z\mathbf{F}}$$

Consequently, for any CR L for $\mathbf{F}$

$$\mathrm{Q}_{L_1 \ldots L_m | Z\mathbf{F}} = \prod_{i=1}^{m} \mathrm{Q}_{L_i|Z\mathbf{F}}$$

Application to Interactive Communication

Suppose $\mathrm{Q}_{X_1 \ldots X_m | Z} = \prod_i \mathrm{Q}_{X_i | Z}$

Then, for every interactive communication $\mathbf{F}$

$$\mathrm{Q}_{X_1 \ldots X_m | Z \mathbf{F}} = \prod \mathrm{Q}_{X_i | Z \mathbf{F}}$$

Consequently, for any CR L for $\mathbf{F}$

$$\mathrm{Q}_{L_1 \ldots L_m | Z \mathbf{F}} = \prod_{i=1}^{m} \mathrm{Q}_{L_i | Z \mathbf{F}}$$

Then, with *leakage parameter* $\delta = \| \mathrm{P}_{L \mathbf{F} Z} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F} Z} \|_1$

$$\begin{aligned} \log |\mathcal{L}| &\leq -\frac{1}{m-1} \log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1 \ldots L_m \mathbf{F} Z}, \mathrm{Q}_{L_1 \ldots L_m \mathbf{F} Z}) + \frac{m}{m-1} \log \frac{1}{\eta} \\ &\leq -\frac{1}{m-1} \log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{X_1 \ldots X_m Z}, \mathrm{Q}_{X_1 \ldots X_m Z}) + \frac{m}{m-1} \log \frac{1}{\eta} \end{aligned}$$

Application to Interactive Communication

Suppose $\mathrm{Q}_{X_1 \dots X_m | Z} = \prod_i \mathrm{Q}_{X_i|Z}$

Then, for every interactive communication $\mathbf{F}$

$$\mathrm{Q}_{X_1 \dots X_m | Z\mathbf{F}} = \prod \mathrm{Q}_{X_i|Z\mathbf{F}}$$

Consequently, for any CR L for $\mathbf{F}$

$$\mathrm{Q}_{L_1 \dots L_m | Z\mathbf{F}} = \prod_{i=1}^{m} \mathrm{Q}_{L_i|Z\mathbf{F}}$$

Then, with *leakage parameter* $\delta = \|\mathrm{P}_{L\mathbf{F}Z} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\|_1$

$$\begin{aligned}\log|\mathcal{L}| &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1\dots L_m\mathbf{F}Z}, \mathrm{Q}_{L_1\dots L_m\mathbf{F}Z}) + \frac{m}{m-1}\log\frac{1}{\eta} \\ &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{X_1\dots X_m Z}, \mathrm{Q}_{X_1\dots X_m Z}) + \frac{m}{m-1}\log\frac{1}{\eta}\end{aligned}$$

Bound is in the spirit of *meta-converse* of [Polyanskiy-Poor-Verdú '10]

Application to Interactive Communication

Suppose $\mathrm{Q}_{X_1 \dots X_m | Z} = \prod_i \mathrm{Q}_{X_i|Z}$

Then, for every interactive communication $\mathbf{F}$

$$\mathrm{Q}_{X_1 \dots X_m | Z\mathbf{F}} = \prod \mathrm{Q}_{X_i|Z\mathbf{F}}$$

Consequently, for any CR L for $\mathbf{F}$

$$\mathrm{Q}_{L_1 \dots L_m | Z\mathbf{F}} = \prod_{i=1}^{m} \mathrm{Q}_{L_i|Z\mathbf{F}}$$

Then, with *leakage parameter* $\delta = \|\mathrm{P}_{L\mathbf{F}Z} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\|_1$

$$\begin{aligned}\log|\mathcal{L}| &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1\dots L_m\mathbf{F}Z}, \mathrm{Q}_{L_1\dots L_m\mathbf{F}Z}) + \frac{m}{m-1}\log\frac{1}{\eta} \\ &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{X_1\dots X_m Z}, \mathrm{Q}_{X_1\dots X_m Z}) + \frac{m}{m-1}\log\frac{1}{\eta}\end{aligned}$$

Can be applied to any partition π of $\{1, \dots, m\}$

Application to Interactive Communication

Suppose $\mathrm{Q}_{X_1\ldots X_m|Z} = \prod_i \mathrm{Q}_{X_i|Z}$

Then, for every interactive communication $\mathbf{F}$

$$\mathrm{Q}_{X_1\ldots X_m|Z\mathbf{F}} = \prod \mathrm{Q}_{X_i|Z\mathbf{F}}$$

Consequently, for any CR L for $\mathbf{F}$

$$\mathrm{Q}_{L_1\ldots L_m|Z\mathbf{F}} = \prod_{i=1}^{m} \mathrm{Q}_{L_i|Z\mathbf{F}}$$

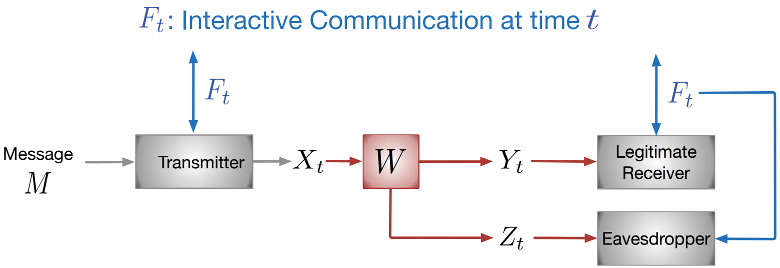
Then, with *leakage parameter* $\delta = \|\mathrm{P}_{L\mathbf{F}Z} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\|_1$

$$\begin{aligned}\log|\mathcal{L}| &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1\ldots L_m\mathbf{F}Z}, \mathrm{Q}_{L_1\ldots L_m\mathbf{F}Z}) + \frac{m}{m-1}\log\frac{1}{\eta} \\ &\leq -\frac{1}{m-1}\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{X_1\ldots X_m Z}, \mathrm{Q}_{X_1\ldots X_m Z}) + \frac{m}{m-1}\log\frac{1}{\eta}\end{aligned}$$

Implications for secure computing explored in [T-Watanabe '14]

Part 2: Application to the Wiretap Channel

Wiretap Channel with Interactive Communication



Transmitting M in n channel uses

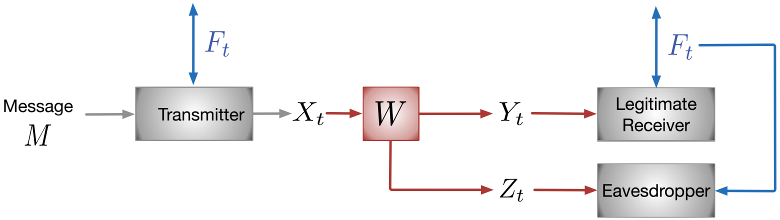
$$\text{Encoder } e_t : (M, F_1, ..., F_{t-1}) \mapsto X_t$$

$$\text{Decoder } d : (Y_1, ..., Y_n, \mathbf{F}) \mapsto \widehat{M}$$

- **Reliability:** $\mathrm{P}\left(M \neq \hat{M}\right) \leq \epsilon$
- **Secrecy:** $\|\mathrm{P}_{MZ^n\mathbf{F}} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{Z^n\mathbf{F}}\|_1 \leq \delta$

Wiretap Channel with Interactive Communication

F_t: Interactive Communication at time t



Transmitting M in n channel uses

$$\text{Encoder } e_t : (M, F_1, ..., F_{t-1}) \mapsto X_t$$

$$\text{Decoder } d : (Y_1, ..., Y_n, \mathbf{F}) \mapsto \widehat{M}$$

- **Reliability:** $\mathrm{P}\left(M \neq \hat{M}\right) \leq \epsilon$
- **Secrecy:** $\|\mathrm{P}_{MZ^n\mathbf{F}} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{Z^n\mathbf{F}}\|_1 \leq \delta$

What is the maximum rate $C_{\epsilon,\delta}$ of message M possible?

A Brief History of Wiretap Channel

[Wyner '75] Capacity of degraded wiretap channel

[Csiszár-Körner '78] Capacity of general wiretap channel

[L. Y. Cheong-Hellman '78] Capacity of Gaussian wiretap channel

[Mid 90's onward] Physical layer security based on wiretap models

[Morgan-Winters '14] A partial strong converse

[Tan-Bloch, this conference] Strong converse for all ϵ, if $\delta \approx 0$

A Brief History of Wiretap Channel

[Wyner '75] Capacity of degraded wiretap channel

[Csiszár-Körner '78] Capacity of general wiretap channel

[L. Y. Cheong-Hellman '78] Capacity of Gaussian wiretap channel

[Mid 90's onward] Physical layer security based on wiretap models

[Morgan-Winters '14] A partial strong converse

[Tan-Bloch, this conference] Strong converse for all ϵ, if $\delta \approx 0$

We prove strong converse for all ϵ, δ such that $\epsilon + \delta < 1$

Recall...

Consider a random variable L taking values in $\mathcal{L}$ s.t.

1. Estimates L_1, L_2 of L satisfy $\mathrm{P}\left(L_1 = L_2 = L\right) \geq 1 - \epsilon$
2. Let $\delta = \|\mathrm{P}_{LZ} - \mathrm{P}_{\texttt{unif}} \times \mathrm{P}_Z\|_1$

Lemma (Conditional Independence Testing Bound)

For any distribution Q *such that* $\mathrm{Q}_{L_1 L_2|Z} = \mathrm{Q}_{L_1|Z}\mathrm{Q}_{L_2|Z}$ *and any* $\eta < 1 - \epsilon - \delta$*, it holds that*

$$\log|\mathcal{L}| \leq -\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1 L_2 Z}, \mathrm{Q}_{L_1 L_2 Z}) + 2\log\frac{1}{\eta}$$

Recall...

Consider a random variable L taking values in $\mathcal{L}$ s.t.

1. Estimates L_1, L_2 of L satisfy $\mathrm{P}\left(L_1 = L_2 = L\right) \geq 1 - \epsilon$
2. Let $\delta = \|\mathrm{P}_{LZ} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_Z\|_1$

Lemma (Conditional Independence Testing Bound)

For any distribution Q *such that* $\mathrm{Q}_{L_1L_2|Z} = \mathrm{Q}_{L_1|Z}\mathrm{Q}_{L_2|Z}$ *and any* $\eta < 1 - \epsilon - \delta$*, it holds that*

$$\log|\mathcal{L}| \leq -\log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1L_2Z}, \mathrm{Q}_{L_1L_2Z}) + 2\log\frac{1}{\eta}$$

Do we have such L and Z?

Recall...

Consider a random variable L taking values in $\mathcal{L}$ s.t.

1. Estimates L_1, L_2 of L satisfy $\mathrm{P}\left(L_1 = L_2 = L\right) \geq 1 - \epsilon$
2. Let $\delta = \|\mathrm{P}_{LZ} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_Z\|_1$

Lemma (Conditional Independence Testing Bound)

For any distribution Q *such that* $\mathrm{Q}_{L_1 L_2|Z} = \mathrm{Q}_{L_1|Z}\mathrm{Q}_{L_2|Z}$ *and any* $\eta < 1 - \epsilon - \delta$, *it holds that*

$$\log |\mathcal{L}| \leq -\log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1 L_2 Z}, \mathrm{Q}_{L_1 L_2 Z}) + 2\log \frac{1}{\eta}$$

Do we have such L and Z?

Sure we do. Choose $L = M$ and $Z = Z^n, \mathbf{F}$

Recall...

Consider a random variable L taking values in $\mathcal{L}$ s.t.

1. Estimates L_1, L_2 of L satisfy $\mathrm{P}\left(L_1 = L_2 = L\right) \geq 1 - \epsilon$
2. Let $\delta = \|\mathrm{P}_{LZ} - \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_Z\|_1$

Lemma (Conditional Independence Testing Bound)

For any distribution Q *such that* $\mathrm{Q}_{L_1 L_2|Z} = \mathrm{Q}_{L_1|Z}\mathrm{Q}_{L_2|Z}$ *and any* $\eta < 1 - \epsilon - \delta$, *it holds that*

$$\log|\mathcal{L}| \leq -\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1 L_2 Z}, \mathrm{Q}_{L_1 L_2 Z}) + 2\log\frac{1}{\eta}$$

Do we have such L and Z?

Sure we do. Choose $L = M$ and $Z = Z^n, \mathbf{F}$

But how do we choose Q?

Recall...

Consider a random variable L taking values in $\mathcal{L}$ s.t.

1. Estimates L_1, L_2 of L satisfy $\mathrm{P}\left(L_1 = L_2 = L\right) \geq 1 - \epsilon$
2. Let $\delta = \|\mathrm{P}_{LZ} - \mathrm{P}_{\texttt{unif}} \times \mathrm{P}_Z\|_1$

Lemma (Conditional Independence Testing Bound)

For any distribution Q *such that* $\mathrm{Q}_{L_1L_2|Z} = \mathrm{Q}_{L_1|Z}\mathrm{Q}_{L_2|Z}$ *and any* $\eta < 1 - \epsilon - \delta$*, it holds that*

$$\log|\mathcal{L}| \leq -\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{L_1L_2Z}, \mathrm{Q}_{L_1L_2Z}) + 2\log\frac{1}{\eta}$$

Do we have such L and Z?

Sure we do. Choose $L = M$ and $Z = Z^n, \mathbf{F}$

But how do we choose Q?

Carefully!

Choosing Q for the Wiretap Channel

Lemma

For a wiretap channel $V : \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}$ such that

$$V(y, z|x) = V_2(z|x)V_1(y|z)$$

and any wiretap code, we get

$$\mathrm{Q}_{M\widehat{M}|Z^n\mathbf{F}} = \mathrm{Q}_{M|Z^n\mathbf{F}} \times \mathrm{Q}_{\widehat{M}|Z^n\mathbf{F}}$$

Choosing Q for the Wiretap Channel

Lemma

For a wiretap channel $V : \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}$ *such that*

$$V(y, z|x) = V_2(z|x)V_1(y|z)$$

and any wiretap code, we get

$$\mathrm{Q}_{M\widehat{M}|Z^n\mathbf{F}} = \mathrm{Q}_{M|Z^n\mathbf{F}} \times \mathrm{Q}_{\widehat{M}|Z^n\mathbf{F}}$$

Thus, by the *conditionally independence testing bound*

$$(\#\text{ of bits of message } M) \leq -\log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{M\widehat{M}Z}, \mathrm{Q}_{M\widehat{M}Z}) + 2\log\frac{1}{\eta}$$

Choosing Q for the Wiretap Channel

Lemma

For a wiretap channel $V : \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}$ such that

$$V(y, z|x) = V_2(z|x)V_1(y|z)$$

and any wiretap code, we get

$$\mathrm{Q}_{M\widehat{M}|Z^n\mathbf{F}} = \mathrm{Q}_{M|Z^n\mathbf{F}} \times \mathrm{Q}_{\widehat{M}|Z^n\mathbf{F}}$$

Thus, by the *conditionally independence testing bound*

$$(\text{\# of bits of message } M) \leq -\log \beta_{\epsilon+\delta+\eta}(\mathrm{P}_{M\widehat{M}Z}, \mathrm{Q}_{M\widehat{M}Z}) + 2\log\frac{1}{\eta}$$

We seek to distinguish W from V by observing $\mathbf{F}, X^n, Y^n, Z^n$

Let $\beta_{\epsilon+\delta+\eta}(W, V, n)$ be defined correspondingly

Active Hypothesis Testing

Theorem ([Hayashi '09])

For $0 < \epsilon < 1$,

$$\lim_n -\frac{1}{n} \log \beta_\epsilon(W, V, n) = \max_{\mathrm{P}_X} D(W \| V \mid \mathrm{P}_X)$$
$$= \max_x D(W_x \| V_x)$$

where W_x *and* V_x*, respectively, denote the* x*th row of* W *and* V

Active Hypothesis Testing

Theorem ([Hayashi '09])

For $0 < \epsilon < 1$,

$$\lim_n -\frac{1}{n} \log \beta_\epsilon(W, V, n) = \max_{\mathrm{P}_X} D(W \| V \mid \mathrm{P}_X)$$
$$= \max_x D(W_x \| V_x)$$

where W_x *and* V_x*, respectively, denote the* x*th row of* W *and* V

Thus, for every ϵ, δ such that $\epsilon + \delta < 1$

$$\lim_n \frac{1}{n} (\#\text{ of bits of message } M) \leq \max_{\mathrm{P}_X} D(W \| V \mid \mathrm{P}_X)$$

for every $V(y, z|x) = V_2(z|x) V_1(y|z)$

Strong Converse for a Degraded Wiretap Channel

Lemma

If the channel W is degraded, $i.e.$, $W(y,z|x) = W_2(z|y)W_1(y|x)$, *then*

$$\min_{V} \max_{\mathrm{P}_X} D(W \| V \mid \mathrm{P}_X) = \max_{\mathrm{P}_X} I(X \wedge Y \mid Z)$$

Theorem

For a degraded wiretap channel W

$$C_{\epsilon,\delta} = \begin{cases} \max_{\mathrm{P}_X} I(X \wedge Y \mid Z), & 0 < \epsilon < 1-\delta \\ \max_{\mathrm{P}_X} I(X \wedge Y), & 1-\delta \le \epsilon < 1 \end{cases}$$

In Closing...

The rate of a code for a degraded wiretap channel cannot be improved even if we don't ask for perfect reliability and secrecy

The Big Picture

Bounds on common randomness lead to converses for specific problems with interactive communication