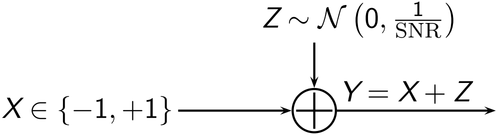
Information Distilling Quantizers

Or Ordentlich (Hebrew University of Jerusalem)
Joint work with Bobak Nazer (BU) and Yury Polyanskiy (MIT)

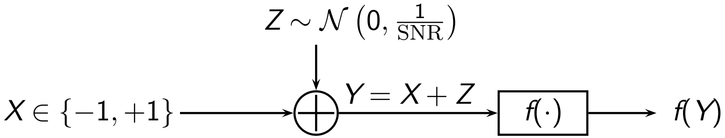
IZS,
Zurich,
February 21, 2018

Quantization for the BIAWGN Channel



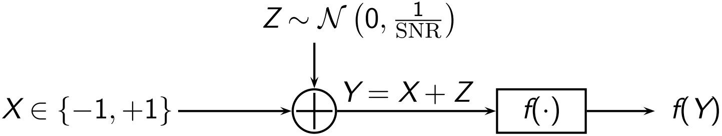
- X is BPSK, $Z \sim \mathcal{N}\left(0, \frac{1}{\mathrm{SNR}}\right)$, $Y = X + Z$

Quantization for the BIAWGN Channel



- X is BPSK, $Z \sim \mathcal{N}\left(0, \frac{1}{\mathrm{SNR}}\right)$, $Y = X + Z$
- $f : \mathbb{R} \mapsto \{0, 1\}$ is a 1-bit quantizer
  Goal: Choose $f(\cdot)$ that maximizes capacity

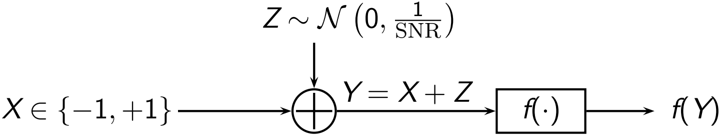
Quantization for the BIAWGN Channel



- X is BPSK, $Z \sim \mathcal{N}\left(0, \frac{1}{\text{SNR}}\right)$, $Y = X + Z$
- $f : \mathbb{R} \mapsto \{0, 1\}$ is a 1-bit quantizer
  Goal: Choose $f(\cdot)$ that maximizes capacity
- Tempting to take

$$
\begin{aligned}
f(y) &= f_{\mathsf{MAP}}(y) \\
&= \underset{x \in \{-1,+1\}}{\arg\max} \Pr(X = x | Y = y) \\
&= \operatorname{sign}(y)
\end{aligned}
$$

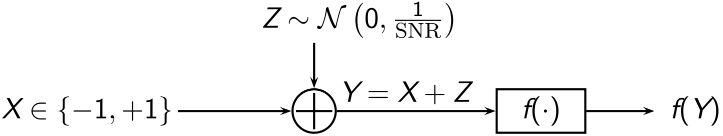
Quantization for the BIAWGN Channel



Viterbi-Omura

For $f(\cdot) = f_{\mathrm{MAP}}(y)$ we have $\frac{I(X;f(Y))}{I(X;Y)} \geq \frac{2}{\pi} \approx 0.6366$

Quantization for the BIAWGN Channel

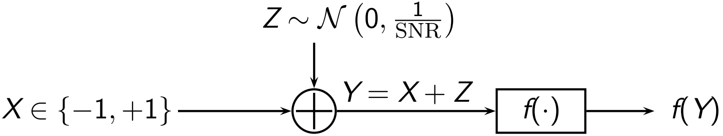


Viterbi-Omura

For $f(\cdot) = f_{\mathrm{MAP}}(y)$ we have $\frac{I(X;f(Y))}{I(X;Y)} \geq \frac{2}{\pi} \approx 0.6366$

To what extent this holds for general binary input channels?

Quantization for the BIAWGN Channel



Viterbi-Omura

For $f(\cdot) = f_{\text{MAP}}(y)$ we have $\frac{I(X;f(Y))}{I(X;Y)} \geq \frac{2}{\pi} \approx 0.6366$

- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\text{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\Pr(f_{\mathsf{MAP}}(Y) \neq X)$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\Pr(f_{\mathrm{MAP}}(Y) \neq X)$

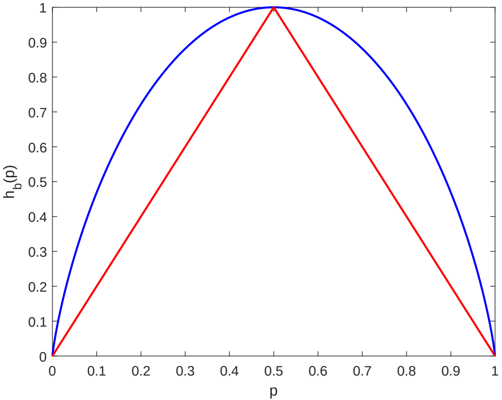
Large $I(X;Y) \Rightarrow$ small P_e

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\Pr(f_{\text{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small P_e

"Reverse Fano" [Feder-Merhav IT'94] : $h_b(p) > 2p$



Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\mathrm{Pr}(f_{\mathrm{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X;f_{\mathrm{MAP}}(Y))$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\Pr(f_{\mathsf{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X; f_{\mathsf{MAP}}(Y))$

"Fano" : $H(X|f_{\mathsf{MAP}}(Y)) = \mathbb{E} h_b(\Pr(f_{\mathsf{MAP}}(Y) \neq X)) \leq h_b(P_e)$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\mathsf{Pr}(f_{\mathsf{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X; f_{\mathsf{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$I(X; f_{\mathsf{MAP}}(Y)) \geq 1 - h_b\left(\frac{1 - I(X;Y)}{2}\right)$$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\Pr(f_{\mathsf{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X; f_{\mathsf{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$I(X; f_{\mathsf{MAP}}(Y)) \geq 1 - h_b\left(\frac{1 - I(X;Y)}{2}\right) \geq \frac{\log e}{2} I(X;Y)^2$$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\Pr(f_{\mathrm{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X;f_{\mathrm{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$\frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} \geq \frac{1-h_b\left(\frac{1-I(X;Y)}{2}\right)}{I(X;Y)} \geq \frac{\log e}{2} I(X;Y)$$

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\,\Pr(f_{\mathsf{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X; f_{\mathsf{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$\frac{I(X; f_{\mathsf{MAP}}(Y))}{I(X;Y)} \geq \frac{1 - h_b\left(\frac{1-I(X;Y)}{2}\right)}{I(X;Y)} \geq \frac{\log e}{2} I(X;Y)$$

- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\mathsf{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\Pr(f_{\mathsf{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X; f_{\mathsf{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$\frac{I(X; f_{\mathsf{MAP}}(Y))}{I(X;Y)} \geq \frac{1 - h_b\left(\frac{1-I(X;Y)}{2}\right)}{I(X;Y)} \geq \frac{\log e}{2} I(X;Y)$$

- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\mathsf{MAP}}(Y))}{I(X;Y)} > \text{const}$?

If $I(X;Y)$ not too small - YES

Performance of the MAP Quantizer - The Good

Let $P_e \triangleq \mathbb{E}\Pr(f_{\mathrm{MAP}}(Y) \neq X)$

Large $I(X;Y) \Rightarrow$ small $P_e \Rightarrow$ large $I(X;f_{\mathrm{MAP}}(Y))$

Easy to show that for $X \sim \mathrm{Bernoulli}(1/2)$

$$\frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} \geq \frac{1-h_b\left(\frac{1-I(X;Y)}{2}\right)}{I(X;Y)} \geq \frac{\log e}{2} I(X;Y)$$

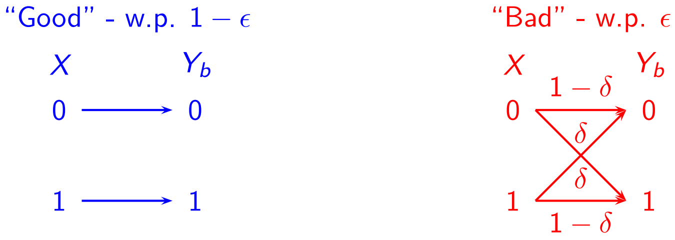
- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Thus, we restrict attention to $I(X;Y) \ll 1$

Performance of the MAP Quantizer - The Bad

Binary input : $X \sim \mathrm{Bernoulli}(1/2)$
Two outputs : $Y = (Y_b, S)$ - $Y_b \in \{0, 1\}$, $S \in \{$"Good", "Bad"$\}$

Performance of the MAP Quantizer - The Bad

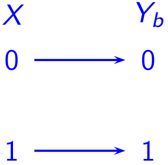


Binary input : $X \sim \text{Bernoulli}(1/2)$
Two outputs : $Y = (Y_b, S)$ - $Y_b \in \{0,1\}$, $S \in \{\text{"Good"}, \text{"Bad"}\}$

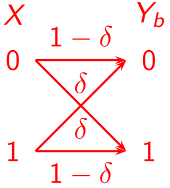
MAP quantizer: $f_{\text{MAP}}(y) = y_b$ (for $\delta \leq 1/2$)
$I(X; f_{\text{MAP}}(Y)) = 1 - h_b(\epsilon\delta)$

Performance of the MAP Quantizer - The Bad

"Good" - w.p. $1-\epsilon$



"Bad" - w.p. ϵ



Binary input : $X \sim \mathrm{Bernoulli}(1/2)$

Two outputs : $Y = (Y_b, S)$ - $Y_b \in \{0, 1\}$, $S \in \{$"Good", "Bad"$\}$

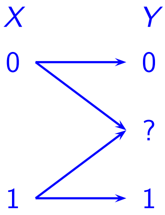
MAP quantizer: $f_{\mathrm{MAP}}(y) = y_b$ (for $\delta \leq 1/2$)

$$I(X; f_{\mathrm{MAP}}(Y)) = 1 - h_b(\epsilon\delta)$$

Asymmetric quantizer: $f_Z(y) = 1$ for $y = (1, \text{"Good"})$, $f_Z(y) = 0$ ow

$$I(X; f_Z(Y)) = 1 - \frac{1+\epsilon}{2} h_b\left(\frac{\epsilon}{1+\epsilon}\right)$$

Performance of the MAP Quantizer - The Bad



$\delta \to 1/2$: $\mathrm{BEC}(\epsilon)$; $I(X;Y) = 1 - \epsilon$

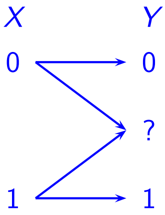
MAP quantizer: $f_{\mathsf{MAP}}(y) = y_b$

$$I(X; f_{\mathsf{MAP}}(Y)) = 1 - h_b\left(\frac{\epsilon}{2}\right)$$

Asymmetric quantizer: $f_Z(y) = 1$ for $y = (1, \text{“Good”})$, $f_Z(y) = 0$ ow

$$I(X; f_Z(Y)) = 1 - \frac{1+\epsilon}{2} h_b\left(\frac{\epsilon}{1+\epsilon}\right)$$

Performance of the MAP Quantizer - The Bad



$\delta \to 1/2$: $\mathrm{BEC}(\epsilon)$; $I(X;Y) = 1 - \epsilon$

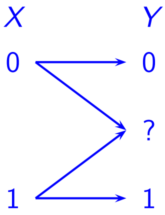
MAP quantizer: $f_{\mathsf{MAP}}(y) = y_b$

$$\frac{I(X;f_{\mathsf{MAP}}(Y))}{I(X;Y)} = \frac{1-h_b\left(\frac{\epsilon}{2}\right)}{1-\epsilon}$$

Asymmetric quantizer: $f_Z(y) = 1$ for $y = (1, \text{“Good”})$, $f_Z(y) = 0$ ow

$$\frac{I(X;f_Z(Y))}{I(X;Y)} = \frac{1-\frac{1+\epsilon}{2}h_b\left(\frac{\epsilon}{1+\epsilon}\right)}{1-\epsilon}$$

Performance of the MAP Quantizer - The Bad



$\delta \to 1/2$: $\mathrm{BEC}(\epsilon)$; $I(X;Y) = 1 - \epsilon$

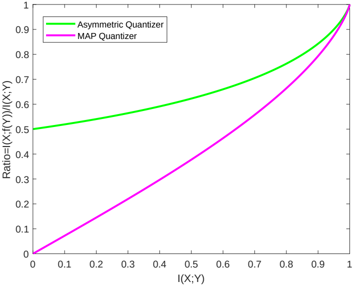
MAP quantizer: $f_{\mathsf{MAP}}(y) = y_b$

$$\lim_{I(X;Y)\to 0} \frac{I(X;f_{\mathsf{MAP}}(Y))}{I(X;Y)} = \lim_{I(X;Y)\to 0} \frac{\log e}{2} I(X;Y) = 0$$

Asymmetric quantizer: $f_Z(y) = 1$ for $y = (1, \text{“Good”})$, $f_Z(y) = 0$ ow

$$\lim_{I(X;Y)\to 0} \frac{I(X;f_Z(Y))}{I(X;Y)} = \frac{1}{2}$$

Performance of the MAP Quantizer - The Bad



$\delta \to 1/2$: $\mathrm{BEC}(\epsilon)$; $I(X;Y) = 1 - \epsilon$

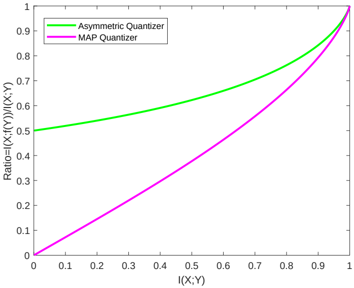
MAP quantizer: $f_{\mathrm{MAP}}(y) = y_b$

$$\lim_{I(X;Y)\to 0} \frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} = \lim_{I(X;Y)\to 0} \frac{\log e}{2} I(X;Y) = 0$$

Asymmetric quantizer: $f_Z(y) = 1$ for $y = (1, \text{"Good"})$, $f_Z(y) = 0$ ow

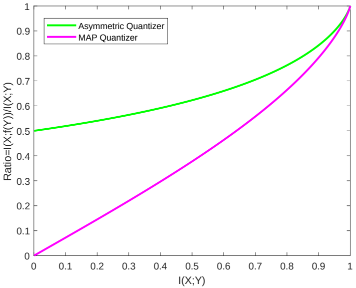
$$\lim_{I(X;Y)\to 0} \frac{I(X;f_Z(Y))}{I(X;Y)} = \frac{1}{2}$$

Performance of the MAP Quantizer - The Bad



- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\text{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Performance of the MAP Quantizer - The Bad



- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\text{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Q2 - NO

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \operatorname*{arg\,max}_{f:\mathcal{Y}\to\{1,\dots,M\}} I(X; f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max}\ I(X; f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Relation to quantization under log-loss:

- $X \in \mathcal{X}$, $\hat{X} \in \mathcal{P}^{|\mathcal{X}|}$ (simplex), $d(x, \hat{x}) = \log \frac{1}{\hat{x}(x)}$
- Remote setting - quantization is done as a function of y

$$d(y, \hat{x}) = \mathbb{E}\left[\log \frac{1}{\hat{x}(X)}\middle| Y = y\right]$$

- An M-"level" quantizer is equivalent to choosing $f : \mathcal{Y} \mapsto [M]$ and $\hat{x}(\cdot) = P_{X|f(Y)=f(y)}$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \arg\max_{f:\mathcal{Y}\to\{1,\ldots,M\}} I(X; f(Y)) = H(X) - \arg\min_{f:\mathcal{Y}\to\{1,\ldots,M\}} H(X|f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Relation to quantization under log-loss:

- $X \in \mathcal{X}$, $\hat{X} \in \mathcal{P}^{|\mathcal{X}|}$ (simplex), $d(x, \hat{x}) = \log \frac{1}{\hat{x}(x)}$
- Remote setting - quantization is done as a function of y

$$d(y, \hat{x}) = \mathbb{E}\left[\log \frac{1}{\hat{x}(X)} \middle| Y = y\right]$$

- An M-“level” quantizer is equivalent to choosing $f : \mathcal{Y} \mapsto [M]$ and $\hat{x}(\cdot) = P_{X|f(Y)=f(y)}$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \underset{f:\mathcal{Y}\to\{1,\dots,M\}}{\arg\max}\, I(X; f(Y)) = \underset{f:\mathcal{Y}\to\{1,\dots,M\}}{\arg\max}\, \mathbb{E}_Y D\left(P_{X|f(Y)} \| P_X\right),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Note:
$f^*(y)$ depends on y only through $P_{X|Y=y}$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) = \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max}\ \mathbb{E}_Y D\left(\mathbb{E}\left[P_{X|Y} | f(Y)\right] \| P_X\right),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Note:
$f^*(y)$ depends on y only through $P_{X|Y=y}$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) = \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max}\ \mathbb{E}_Y D\left(\mathbb{E}\left[P_{X|Y}|f(Y)\right] \| P_X\right),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Note:
$f^*(y)$ depends on y only through $P_{X|Y=y}$

Theorem [Burshtein-Pietra-Kanevsky-Nadas'92]

Suffices to consider $f(\cdot)$ that partitions the simples $\mathcal{P}^{|\mathcal{X}|}$ to M convex regions

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) = \underset{f:\mathcal{Y}\to\{1,\dots,M\}}{\arg\max}\ \mathbb{E}_Y D\left(\mathbb{E}\left[P_{X|Y}|f(Y)\right] \| P_X\right),$$

and

$$I(X;[Y]_M) \triangleq I(X; f^*(Y)).$$

Note:
$f^*(y)$ depends on y only through $P_{X|Y=y}$

Theorem [Burshtein-Pietra-Kanevsky-Nadas'92]

Suffices to consider $f(\cdot)$ that partitions the simples $\mathcal{P}^{|\mathcal{X}|}$ to M convex regions

For binary X, $\mathcal{P}^{|\mathcal{X}|} = [0,1]$ and it suffices to consider quantizers that partition it to M intervals

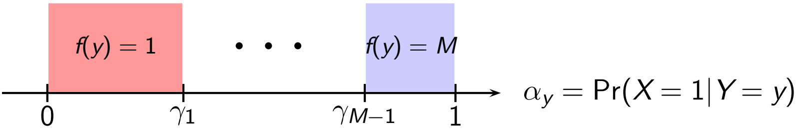
General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) = \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max}\ \mathbb{E}_Y D\left(\mathbb{E}\left[P_{X|Y}|f(Y)\right] \| P_X\right),$$

and

$$I(X;[Y]_M) \triangleq I(X;f^*(Y)).$$



General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max}\; I(X; f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Relation to Information Bottleneck:

- $$\mathrm{IB}_R(P_{XY}) \triangleq \max_{P_{T|Y}\,:\, I(Y;T)\le R} I(X; T)$$
- Take $(X^n, Y^n) \sim P_{XY}^{\otimes n}$
- [Gilad–Bachrach-Navot-Tishby'03, Courtade-Weissman '14]:

$$\lim_{n\to\infty} \frac{1}{n} I(X^n; [Y^n]_{M^n}) = \mathrm{IB}_{\log M}(P_{XY})$$

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \arg\max_{f:\mathcal{Y}\to\{1,\dots,M\}} I(X; f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Significance:

- Information-rate under scalar quantization
- Algorithms for Polar-Codes design
- Scalar version of various problems in IT - Wyner'75/Ahlswede-Korner'75, Erkip-Cover'98
- Required number of clusters for learning under log-loss

General Setup

For $(X, Y) \sim P_{XY}$ and $M \in \mathbb{N}$, define

$$f^*(Y) \triangleq \underset{f:\mathcal{Y}\to\{1,\ldots,M\}}{\arg\max} I(X; f(Y)),$$

and

$$I(X; [Y]_M) \triangleq I(X; f^*(Y)).$$

Simple properties:

- Suffices to consider deterministic $f(\cdot)$ in the arg max
- $$\frac{M-1}{|\mathcal{Y}|} I(X; Y) \leq I(X; [Y]_M) \leq \min\{I(X; Y), \log(M)\}.$$
- If $X - Y - V$, then $I(X; [V]_M) \leq I(X; [Y]_M)$
- $P_{Y|X} \mapsto I(X; [Y]_M)$ is convex;
  $P_X \mapsto I(X; [Y]_M)$ is not necessarily concave

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$, *and* $I(X;Y) = \beta$, *then*

$$I\big(X;[Y]_M\big) \geq \mathrm{constant} \times \frac{(M-1)\beta}{\log(1/\beta)}.$$

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$*, and* $I(X;Y) = \beta$*, then*

$$I(X;[Y]_M) \geq \mathrm{constant} \times \frac{(M-1)\beta}{\log(1/\beta)}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$, *and* $I(X;Y) = \beta$, *then*

$$I(X;[Y]_M) \geq \mathrm{constant} \times \frac{(M-1)\beta}{\log(1/\beta)}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

[Kartowsky-Tal'17, Tal'15]: Bound on additive gap

$$I(X;[Y]_M) \geq \beta - \mathrm{const} \cdot M^{-2}$$

and there is a sequence of channels for which this is **tight** (up to constants)

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$, *and* $I(X;Y) = \beta$, *then*

$$I(X;[Y]_M) \geq \mathrm{constant} \times \frac{(M-1)\beta}{\log(1/\beta)}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

[Kartowsky-Tal'17, Tal'15]: Bound on additive gap

$$I(X;[Y]_M) \geq \beta - \mathrm{const} \cdot M^{-2}$$

and there is a sequence of channels for which this is **tight** (up to constants)

Additive gap - meaningful for large β
Multiplicative gap - meaningful for small β

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$, *and* $I(X;Y) = \beta$, *then*

$$\frac{I(X;[Y]_M)}{I(X;Y)} \geq \mathrm{constant} \times \frac{(M-1)}{\log(1/I(X;Y))}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$*, and* $I(X;Y) = \beta$*, then*

$$\frac{I(X;[Y]_M)}{I(X;Y)} \geq \mathrm{constant} \times \frac{(M-1)}{\log(1/I(X;Y))}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \text{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} > \text{const}$?

Main Result

Theorem (Nazer-O.-Polyanskiy ISIT'17)

If $X \sim \mathrm{Bernoulli}(1/2)$*, and* $I(X;Y) = \beta$*, then*

$$\frac{I(X;[Y]_M)}{I(X;Y)} \geq \mathrm{constant} \times \frac{(M-1)}{\log(1/I(X;Y))}.$$

Also, there is a sequence of channels for which this is ***tight*** *(up to constants).*

- Q1: Does there always exist some 1-bit quantizer $f(\cdot)$ such that $\frac{I(X;f(Y))}{I(X;Y)} > \mathrm{const}$?
- Q2: Is it always true that $\frac{I(X;f_{\mathrm{MAP}}(Y))}{I(X;Y)} > \mathrm{const}$?

Q1 - NO

Upper Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $f(y)$ is defined by the intervals $\mathcal{I}_1, \ldots, \mathcal{I}_M$
- $d_b(p \| q) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$ is binary KL

$$
\begin{aligned}
I(X; f(Y)) &= \sum_{i=1}^{M} \Pr(\alpha_Y \in \mathcal{I}_i) d_b\left(\mathbb{E}[\alpha_Y | \alpha_Y \in \mathcal{I}_i] \,\|\, \frac{1}{2}\right) \\
&\leq M \max_{0 \leq a < b \leq 1} \Pr(a \leq \alpha_Y \leq b) d_b\left(\mathbb{E}[\alpha_Y | a \leq \alpha_Y \leq b] \,\|\, \frac{1}{2}\right)
\end{aligned}
$$

Upper Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $f(y)$ is defined by the intervals $\mathcal{I}_1, \ldots, \mathcal{I}_M$
- $d_b(p \| q) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$ is binary KL

$$I(X; f(Y)) = \sum_{i=1}^{M} \Pr(\alpha_Y \in \mathcal{I}_i) d_b\left(\mathbb{E}[\alpha_Y | \alpha_Y \in \mathcal{I}_i] \,\|\, \frac{1}{2}\right)$$

$$\leq M \max_{0 \leq a < b \leq 1} \Pr(a \leq \alpha_Y \leq b) d_b\left(\mathbb{E}[\alpha_Y | a \leq \alpha_Y \leq b] \,\|\, \frac{1}{2}\right)$$

Remains to find a channel $P_{Y|X}$ for which product is always small

Upper Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $f(y)$ is defined by the intervals $\mathcal{I}_1, \ldots, \mathcal{I}_M$
- $d_b(p \| q) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$ is binary KL

$$I(X; f(Y)) = \sum_{i=1}^{M} \Pr(\alpha_Y \in \mathcal{I}_i) d_b\left(\mathbb{E}[\alpha_Y | \alpha_Y \in \mathcal{I}_i] \,\|\, \frac{1}{2}\right)$$

$$\leq M \max_{0 \leq a < b \leq 1} \Pr(a \leq \alpha_Y \leq b) d_b\left(\mathbb{E}[\alpha_Y | a \leq \alpha_Y \leq b] \,\|\, \frac{1}{2}\right)$$

Remains to find a channel $P_{Y|X}$ for which product is always small

Take a BMS $Y = (X \oplus Z_T, T)$, where $Z_T \sim \mathrm{Bernoulli}(T)$ and

$$f_T(t) = \begin{cases} r\delta(t) + \frac{4r}{(1-2t)^3} & 0^- < t \leq \frac{1-\sqrt{r}}{2} \\ 0 & \text{otherwise} \end{cases}$$

Upper Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $f(y)$ is defined by the intervals $\mathcal{I}_1, \ldots, \mathcal{I}_M$
- $d_b(p \| q) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$ is binary KL

$$I(X; f(Y)) = \sum_{i=1}^{M} \Pr(\alpha_Y \in \mathcal{I}_i) d_b\left(\mathbb{E}[\alpha_Y | \alpha_Y \in \mathcal{I}_i] \,\|\, \frac{1}{2}\right)$$

$$\leq M \max_{0 \leq a < b \leq 1} \Pr(a \leq \alpha_Y \leq b) d_b\left(\mathbb{E}[\alpha_Y | a \leq \alpha_Y \leq b] \,\|\, \frac{1}{2}\right)$$

Theorem (Nazer-O.-Polyanskiy ISIT'17)

For $X \sim \text{Bernoulli}(1/2)$ and any $I(X;Y) \in [0,1]$, there exist a (BMS) $P_{Y|X}$ such that

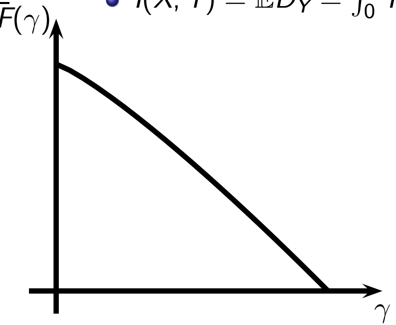
$$I(X; [Y]_M) \leq 2M \frac{I(X;Y)}{\ln\left(\frac{e \log(e)}{2I(X;Y)}\right)}$$

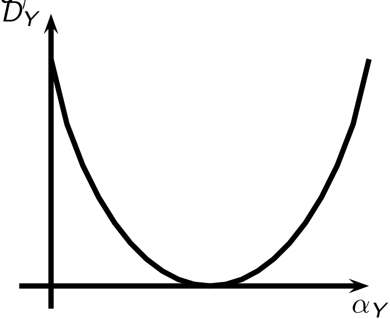
Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

Lower Bound - Sketch of Proof

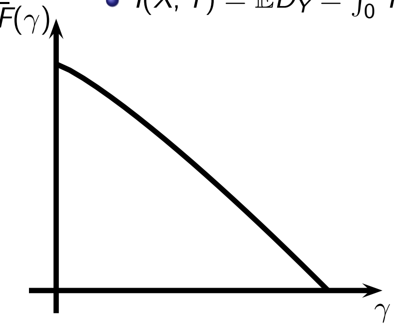
- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

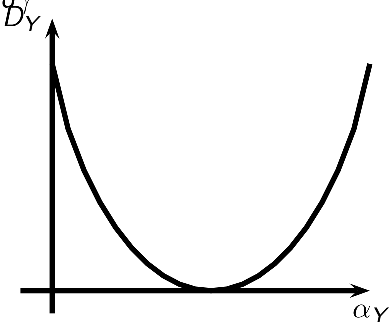




Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

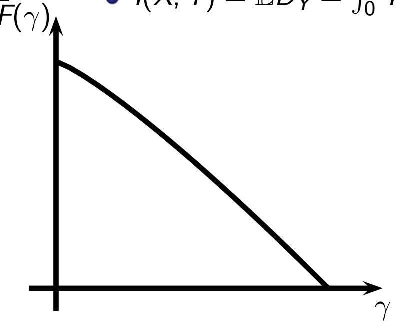


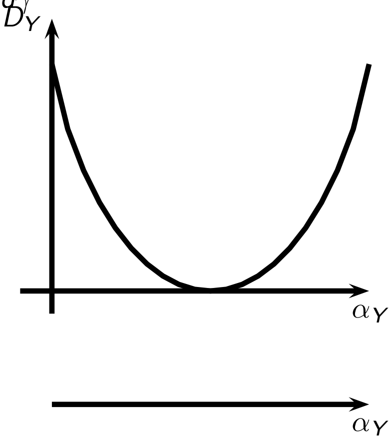


$2M + 1$ level quantizer design: choose $0 < \gamma_1 < \cdots < \gamma_{M-1} < 1$

Lower Bound - Sketch of Proof

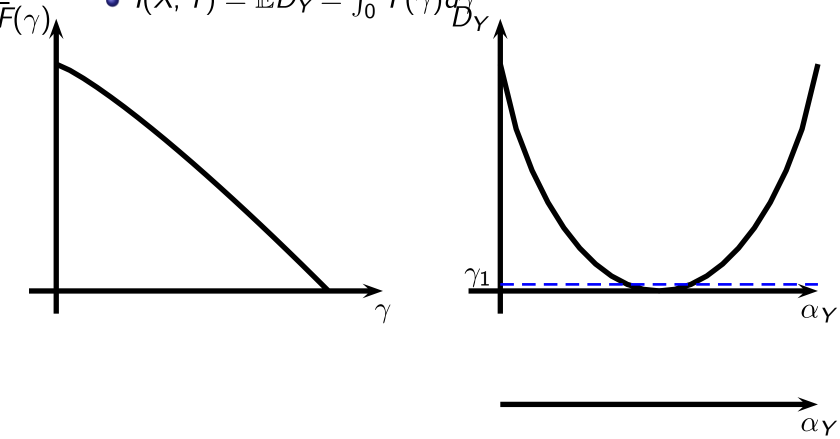
- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$





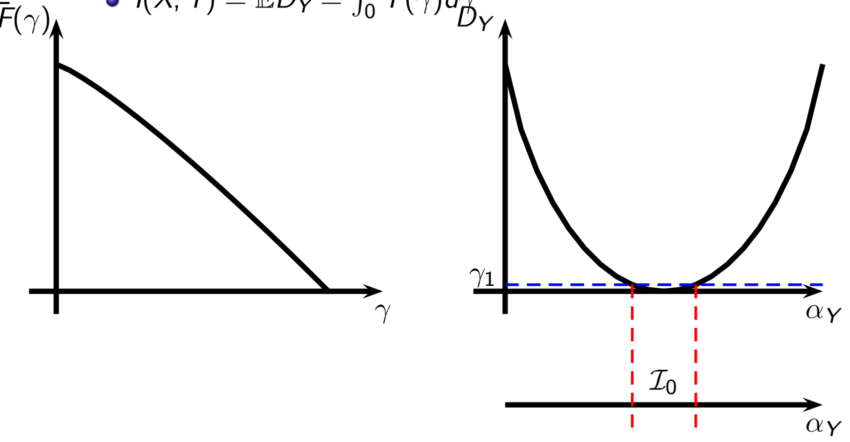
Lower Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E}D_Y = \int_0^1 \bar{F}(\gamma)d\gamma$



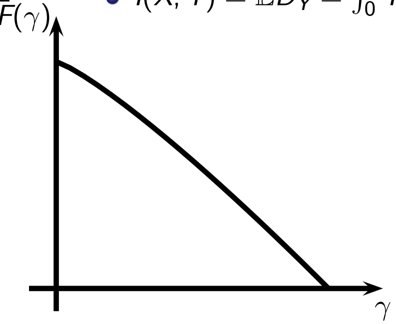
Lower Bound - Sketch of Proof

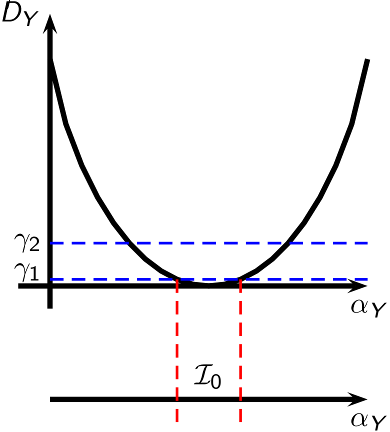
- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$



Lower Bound - Sketch of Proof

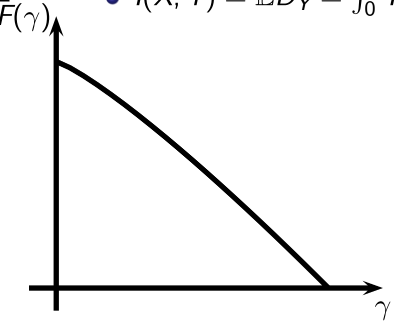
- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

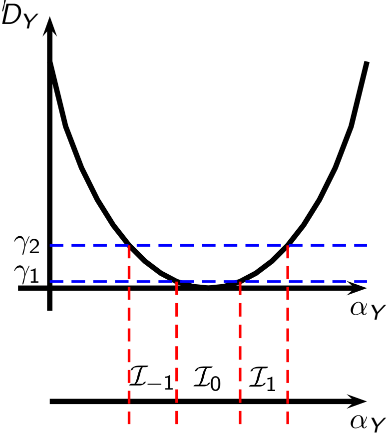




Lower Bound - Sketch of Proof

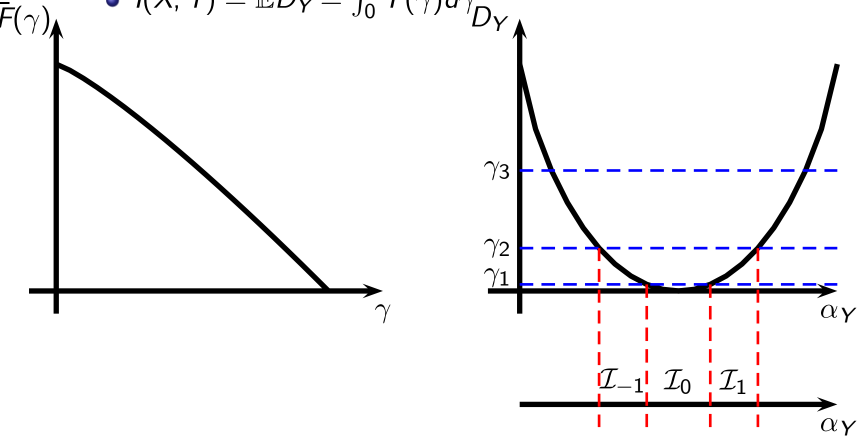
- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$





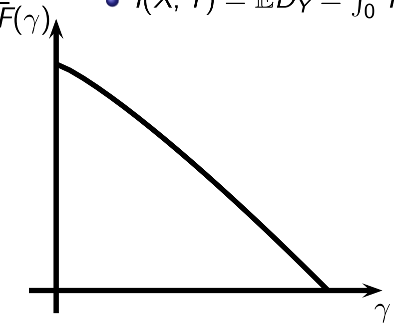
Lower Bound - Sketch of Proof

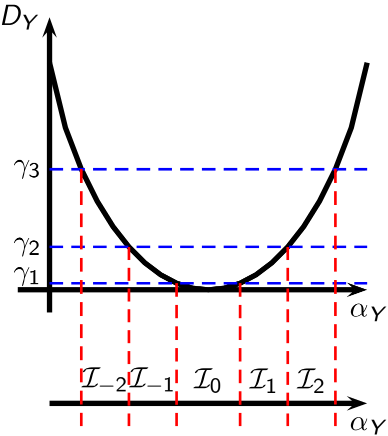
- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X=1|Y=y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E}D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$



Lower Bound - Sketch of Proof

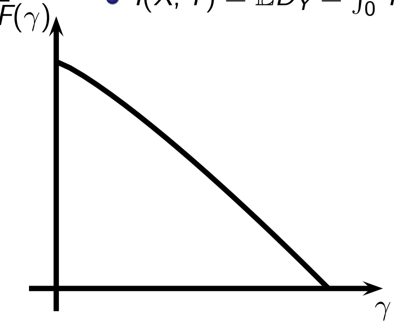
- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E}D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

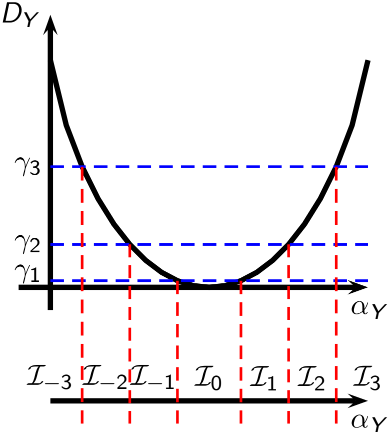




Lower Bound - Sketch of Proof

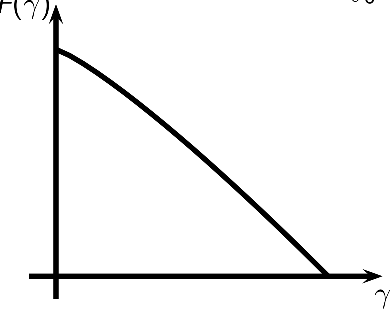
- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

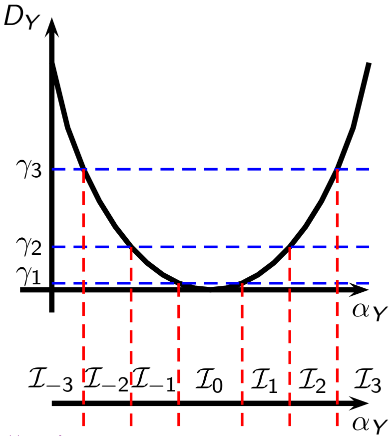




Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

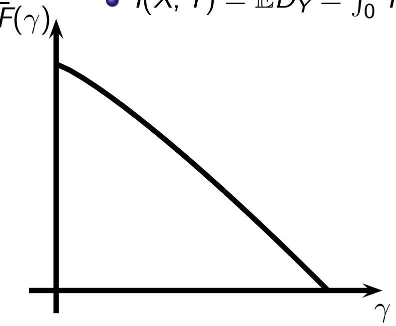


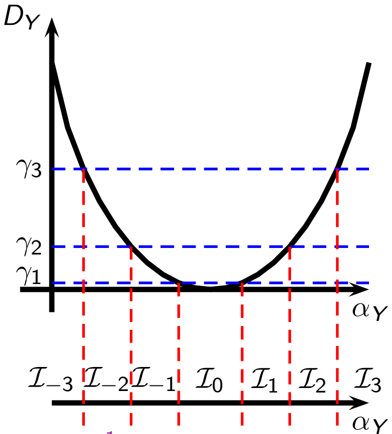


$I(X; f(Y))$
$= \sum_{\ell=-M}^{M} \Pr(f(Y) = \ell) D(P_{X|f(Y)=\ell} || P_X)$

Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

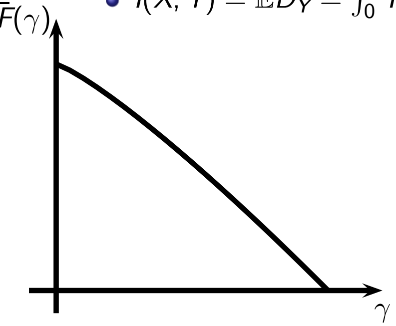


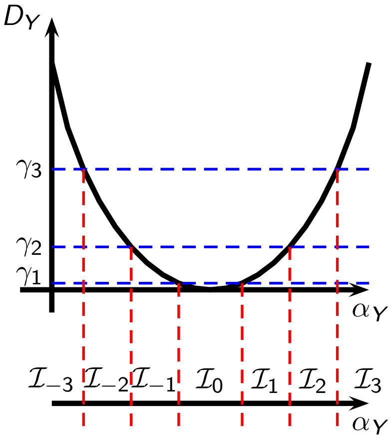


$$I(X; f(Y)) = \sum_{\ell=-M}^{M} \Pr(\alpha_Y \in \mathcal{I}_\ell) d_b(\mathbb{E}[\alpha_Y | \alpha_Y \in \mathcal{I}_\ell] \| \tfrac{1}{2})$$

Lower Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

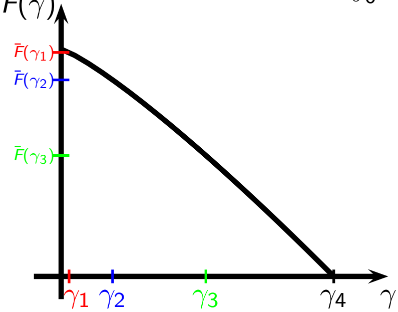


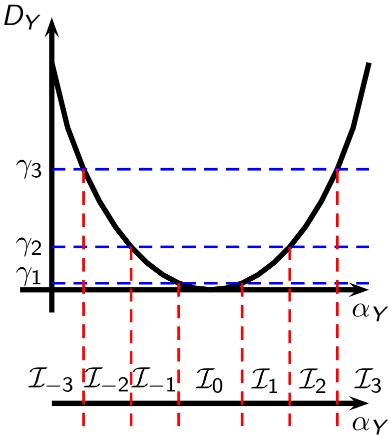


$$I(X; f(Y)) \geq \sum_{\ell=0}^{M} \left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1}) \right) \gamma_\ell$$

Lower Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

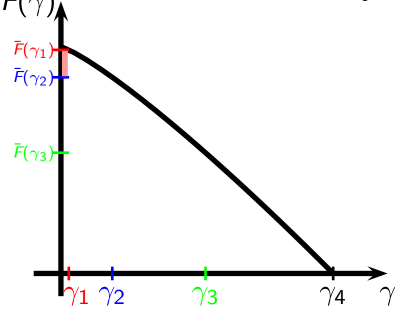


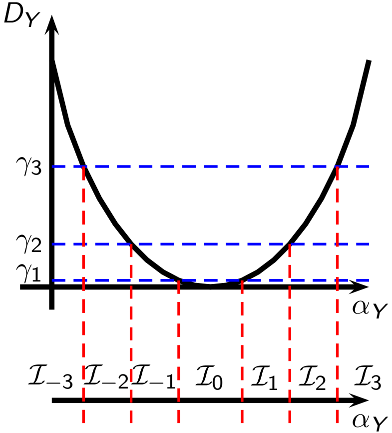


$$I(X; f(Y)) \geq \sum_{\ell=0}^{M} \left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1})\right) \gamma_\ell$$

Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

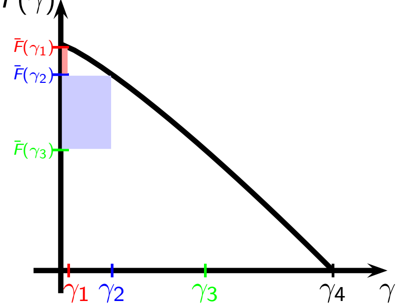


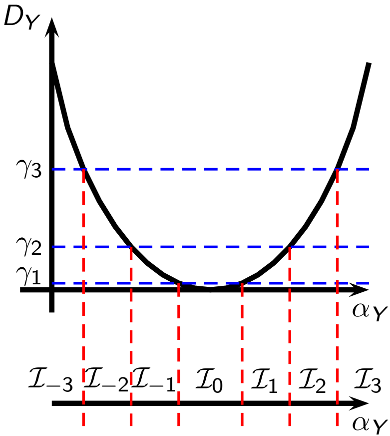


$$I(X; f(Y)) \geq \sum_{\ell=0}^{M} \left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1})\right) \gamma_\ell$$

Lower Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

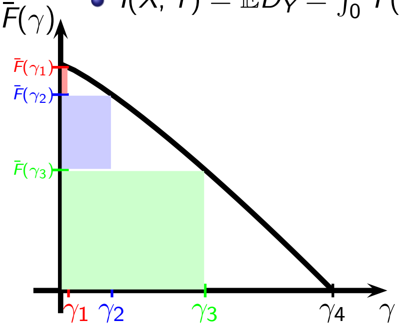


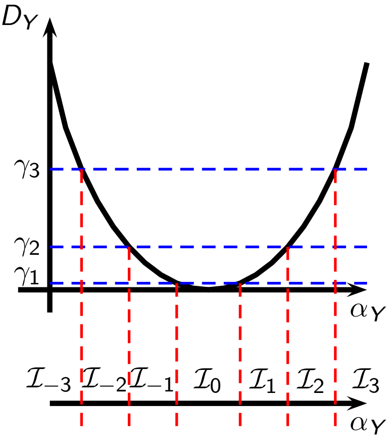


$$I(X; f(Y)) \geq \sum_{\ell=0}^{M} \left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1})\right) \gamma_\ell$$

Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

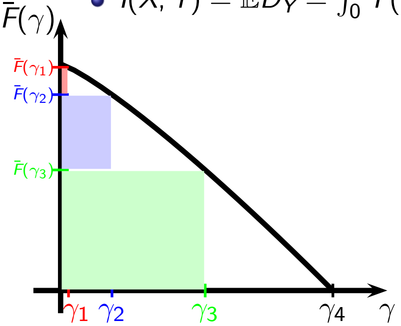


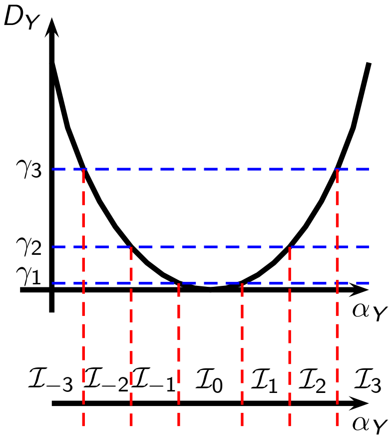


$$I(X; f(Y)) \geq \sum_{\ell=0}^{M} \left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1})\right) \gamma_\ell$$

Lower Bound - Sketch of Proof

- $X \sim \text{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1|Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X;Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$

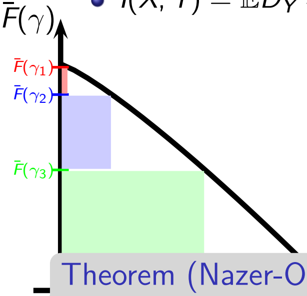


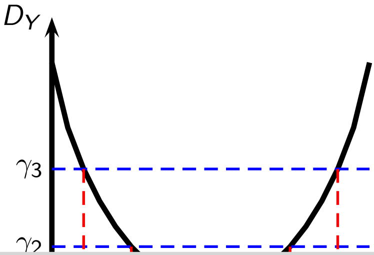


Quantizer design = approximating integral with a sum

Lower Bound - Sketch of Proof

- $X \sim \mathrm{Bernoulli}(1/2)$, $\alpha_y = \Pr(X = 1 | Y = y)$
- $D_y = D(P_{X|Y=y} \| P_X) = d_b(\alpha_y \| \frac{1}{2})$
- $\bar{F}(\gamma) = \Pr(D_Y \geq \gamma)$
- $I(X; Y) = \mathbb{E} D_Y = \int_0^1 \bar{F}(\gamma) d\gamma$





Theorem (Nazer-O.-Polyanskiy ISIT'17)

For $X \sim \mathrm{Bernoulli}(1/2)$ *and any* $P_{Y|X}$

$$I(X; [Y]_M) \geq const \cdot \frac{M-1}{\max\left\{\log\left(\frac{1}{I(X;Y)}\right), 1\right\}} \cdot I(X; Y)$$

Conclusions

- We studied the problem of quantizing the output of a channel with binary input, in the "low-SNR" regime
- For the worst-case channel

$$I(X; [Y]_M) = \Theta\left((M-1)\frac{I(X;Y)}{\log\left(\frac{1}{I(X;Y)}\right)}\right)$$

- Optimal quantizer can be highly asymmetric
- It is also possible to show that if we restrict $H(f(Y)) = \log M$ instead of $|f(Y)| = M$ then it is always possible to attain

$$I(X; f(Y)) > c\left((M-1)\frac{I(X;Y)}{\log\log\left(\frac{1}{I(X;Y)}\right)}\right)$$