Source Coding with Lists and Rényi Entropy
or
The Honey-Do Problem

Amos Lapidoth

ETH Zurich

October 8, 2013

Joint work with Christoph Bunte.

A Task from your Spouse

Using a fixed number of bits, your spouse reminds you of one of the following tasks:

- Honey, don't forget to feed the cat.
- Honey, don't forget to go to the dry-cleaner.
- Honey, don't forget to pick-up my parents at the airport.
- Honey, don't forget the kids' violin concert.
-
-
-

A Task from your Spouse

Using a fixed number of bits, your spouse reminds you of one of the following tasks:

- Honey, don't forget to feed the cat.
- Honey, don't forget to go to the dry-cleaner.
- Honey, don't forget to pick-up my parents at the airport.
- Honey, don't forget the kids' violin concert.
-
-
-

The combinatorical approach requires

$$\# \text{ of bits} = \lceil \log_2 \# \text{ of tasks} \rceil .$$

It guarantees that you'll know what to do. . .

The Information-Theoretic Approach

- Model the tasks as elements of $\mathcal{X}^n$ generated IID P.
- Ignore the atypical sequences.
- Index the typical sequences using $\approx n\,H(X)$ bits.
- Send the index.
- Typical tasks will be communicated error-free.

The Information-Theoretic Approach

- Model the tasks as elements of $\mathcal{X}^n$ generated IID P.
- Ignore the atypical sequences.
- Index the typical sequences using $\approx n\,H(X)$ bits.
- Send the index.
- Typical tasks will be communicated error-free.

Any married person knows how ludicrous this is:

What if the task is atypical?

Yes, this is unlikely, but:

- You won't even know it!
- Are you ok with the consequences?

Improved Information-Theoretic Approach

- First bit indicates whether task is typical.
- You'll know when the task is lost in transmission.

Improved Information-Theoretic Approach

- First bit indicates whether task is typical.
- You'll know when the task is lost in transmission.

What are you going to do about it?

Improved Information-Theoretic Approach

- First bit indicates whether task is typical.
- You'll know when the task is lost in transmission.

What are you going to do about it?

- If I were you, I would perform them all.
- Yes, I know there are exponentially many of them.
- Are you beginning to worry about the expected number of tasks?

Improved Information-Theoretic Approach

- First bit indicates whether task is typical.
- You'll know when the task is lost in transmission.

What are you going to do about it?

- If I were you, I would perform them all.
- Yes, I know there are exponentially many of them.
- Are you beginning to worry about the expected number of tasks?

You could perform a subset of the tasks.

- You'll get extra points for effort.
- But what if the required task is not in the subset?
- Are you ok with the consequences?

Our Problem

- A source generates X^n in $\mathcal{X}^n$ IID P.
- The sequence is described using nR bits.
- Based on the description, a list is generated that is guaranteed to contain X^n.
- For which rates R can we find descriptions and corresponding lists with expected listsize arbitrarily close to 1?

More generally, we'll look at the ρ-th moment of the listsize.

What if you are not in a Relationship?

Should you tune out?

Rényi Entropy

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \left[\sum_{x \in \mathcal{X}} P(x)^\alpha \right]^{1/\alpha}$$

Alfréd Rényi
(1921–1970)

A Homework Problem

Show that

1. $\lim_{\alpha \to 1} H_\alpha(X) = H(X)$.
2. $\lim_{\alpha \to 0} H_\alpha(X) = \log |\text{supp} P|$.
3. $\lim_{\alpha \to \infty} H_\alpha(X) = -\log \max_{x \in \mathcal{X}} P(x)$.

Do not Tune Out

- Our problem gives an operational meaning to

$$H_{\frac{1}{1+\rho}}, \quad \rho > 0 \qquad \text{(i.e., } 0 < \alpha < 1\text{).}$$

- It reveals many of its properties.
- And it motivates the conditional Rényi entropy.

Lossless List Source Codes

- Rate-R blocklength-n source code with list decoder:

$$f_n \colon \mathcal{X}^n \to \{1, \ldots, 2^{nR}\}, \quad \lambda_n \colon \{1, \ldots, 2^{nR}\} \to 2^{\mathcal{X}^n}$$

- The code is lossless if

$$x^n \in \lambda_n(f_n(x^n)), \quad \forall x^n \in \mathcal{X}^n$$

- ρ-th listsize moment ($\rho > 0$):

$$\mathsf{E}[|\lambda_n(f_n(X^n))|^\rho] = \sum_{x^n \in \mathcal{X}^n} P^n(x^n)|\lambda_n(f_n(x^n))|^\rho$$

The Main Result on Lossless List Source Codes

Theorem

1. *If* $R > H_{\frac{1}{1+\rho}}(X)$*, then there exists* $(f_n, \lambda_n)_{n \geq 1}$ *such that*

$$\lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n))|^\rho] = 1.$$

2. *If* $R < H_{\frac{1}{1+\rho}}(X)$*, then*

$$\lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n))|^\rho] = \infty.$$

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ
   (Monotonicity of $\rho \mapsto a^{\rho}$ when $a \geq 1$.)

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ

2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log |\mathcal{X}|$

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ

2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log|\mathcal{X}|$
   ($R < H(X) \implies$ listsize ≥ 2 w.p. tending to one.
   And $R = \log|\mathcal{X}|$ can guarantee listsize $= 1$.)

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ
2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log|\mathcal{X}|$
3. $\lim_{\rho \to 0} H_{\frac{1}{1+\rho}}(X) = H(X)$

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ

2. $H(X) \le H_{\frac{1}{1+\rho}}(X) \le \log|\mathcal{X}|$

3. $\lim_{\rho\to 0} H_{\frac{1}{1+\rho}}(X) = H(X)$
   ($R > H(X) \implies$ prob(listsize ≥ 2) decays exponentially. For small ρ beats $|\lambda_n(f_n(X^n))|^{\rho}$, which cannot exceed $e^{n\rho\log|\mathcal{X}|}$.)

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ

2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log|\mathcal{X}|$

3. $\lim_{\rho \to 0} H_{\frac{1}{1+\rho}}(X) = H(X)$

4. $\lim_{\rho \to \infty} H_{\frac{1}{1+\rho}}(X) = \log|\text{supp}(P)|$

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ
2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log|\mathcal{X}|$
3. $\lim_{\rho\to 0} H_{\frac{1}{1+\rho}}(X) = H(X)$
4. $\lim_{\rho\to\infty} H_{\frac{1}{1+\rho}}(X) = \log|\mathrm{supp}(P)|$
   ($R < \log|\mathrm{supp}(P)| \implies \exists \mathbf{x}_0 \in \mathrm{supp}(P)^n$ for which $|\varphi_n(f_n(\mathbf{x}_0))| \geq e^{n(\log|\mathrm{supp}(P)|-R)}$. Since $P^n(\mathbf{x}_0) \geq p_{\min}^n$, where $p_{\min} = \min\{P(x) : x \in \mathrm{supp}(P)\}$

   $$\sum_{\mathbf{x}} P^n(\mathbf{x})|\varphi_n(f_n(\mathbf{x}))|^\rho \geq e^{n\rho(\log|\mathrm{supp}(P)|-R-\frac{1}{\rho}\log\frac{1}{p_{\min}})}.$$

   Hence R is not achievable if ρ is large.)

Some Properties of $H_{\frac{1}{1+\rho}}(X)$

1. Nondecreasing in ρ
2. $H(X) \leq H_{\frac{1}{1+\rho}}(X) \leq \log|\mathcal{X}|$
3. $\lim_{\rho\to 0} H_{\frac{1}{1+\rho}}(X) = H(X)$
4. $\lim_{\rho\to\infty} H_{\frac{1}{1+\rho}}(X) = \log|\mathrm{supp}(P)|$

Sketch of Direct Part

1. Partition each type-class T_Q into 2^{nR} lists of $\approx$ lengths

$$\lceil 2^{-nR}|T_Q|\rceil \approx 2^{n(H(Q)-R)}.$$

2. Describe the type of x^n using $o(n)$ bits.
3. Describe the list containing x^n using nR bits.
4. $\Pr(X^n \in T_Q) \approx 2^{-nD(Q||P)}$ and small number of types, so

$$\sum_Q \Pr(X^n \in T_Q)\lceil 2^{n(H(Q)-R)}\rceil^\rho$$
$$\leq 1 + 2^{-n\rho(R-\max_Q\{H(Q)-\rho^{-1}D(Q||P)\}-\delta_n)}$$

   where $\delta_n \to 0$.

5. By Arıkan'96,

$$\max_Q\{H(Q) - \rho^{-1}D(Q||P)\} = H_{\frac{1}{1+\rho}}(X). \quad \square$$

The Key to the Converse

Lemma

If

1. P *is a PMF on a finite nonempty set* $\mathcal{X}$,
2. $\mathcal{L}_1, \ldots, \mathcal{L}_M$ *is a partition of* $\mathcal{X}$,
3. $L(x) \triangleq |\mathcal{L}_j|$ *if* $x \in \mathcal{L}_j$.

Then

$$\sum_{x \in \mathcal{X}} P(x) L^{\rho}(x) \geq M^{-\rho} \left[\sum_{x \in \mathcal{X}} P(x)^{\frac{1}{1+\rho}} \right]^{1+\rho}.$$

A Simple Identity for the Proof of the Lemma

$$\sum_{x \in \mathcal{X}} \frac{1}{L(x)} = M.$$

A Simple Identity for the Proof of the Lemma

$$\sum_{x \in \mathcal{X}} \frac{1}{L(x)} = M.$$

Proof:

$$\begin{aligned}
\sum_{x \in \mathcal{X}} \frac{1}{L(x)} &= \sum_{j=1}^{M} \sum_{x \in \mathcal{L}_j} \frac{1}{L(x)} \\
&= \sum_{j=1}^{M} \sum_{x \in \mathcal{L}_j} \frac{1}{|\mathcal{L}_j|} \\
&= \sum_{j=1}^{M} 1 \\
&= M.
\end{aligned}$$

Proof of the Lemma

1. Recall Hölder's Inequality: If $p, q > 1$ and $1/p + 1/q = 1$, then

$$\sum_x a(x)b(x) \leq \left[\sum_x a(x)^p\right]^{\frac{1}{p}} \left[\sum_x b(x)^q\right]^{\frac{1}{q}}, \quad a(\cdot), b(\cdot) \geq 0.$$

2. Rearranging gives

$$\sum_x a(x)^p \geq \left[\sum_x b(x)^q\right]^{-\frac{p}{q}} \left[\sum_x a(x)b(x)\right]^p.$$

3. Choose $p = 1 + \rho$, $q = (1 + \rho)/\rho$, $a(x) = P(x)^{\frac{1}{1+\rho}} L(x)^{\frac{\rho}{1+\rho}}$ and $b(x) = L(x)^{-\frac{\rho}{1+\rho}}$, and note that

$$\sum_{x \in \mathcal{X}} \frac{1}{L(x)} = M.$$

Converse

1. WLOG assume $\lambda_n(m) = \{x^n \in \mathcal{X}^n : f_n(x^n) = m\}$.
2. $\Rightarrow$ The lists $\lambda_n(1), \ldots, \lambda_n(2^{nR})$ partition $\mathcal{X}^n$.
3. $\lambda_n(f_n(x^n))$ is the list containing x^n.
4. By the lemma:

$$\sum_{x^n \in \mathcal{X}^n} P_X^n(x^n) |\lambda_n(f_n(x^n))|^\rho \geq 2^{-n\rho R} \left[\sum_{x^n \in \mathcal{X}^n} P_X^n(x^n)^{\frac{1}{1+\rho}} \right]^{1+\rho}$$

$$= 2^{n\rho\left(H_{\frac{1}{1+\rho}}(X) - R\right)}.$$

□

Recall the lemma:

$$\sum_{x \in \mathcal{X}} P(x) L^\rho(x) \geq M^{-\rho} \left[\sum_{x \in \mathcal{X}} P(x)^{\frac{1}{1+\rho}} \right]^{1+\rho}.$$

How to Define Conditional Rényi Entropy?

Should it be defined as

$$\sum_{y \in \mathcal{Y}} P_Y(y)\, H_\alpha(X|Y = y) \quad ?$$

How to Define Conditional Rényi Entropy?

Should it be defined as

$$\sum_{y \in \mathcal{Y}} P_Y(y)\, H_\alpha(X|Y=y) \quad ?$$

Consider Y as side information to both encoder and decoder,

$$(X_i, Y_i) \sim \text{IID } P_{XY}.$$

You and your spouse hopefully have something in common...

Lossless List Source Codes with Side-Information

- $(X_1, Y_1), (X_2, Y_2), \ldots \sim$ IID $P_{X,Y}$
- Y^n is side-information.
- Rate-R blocklength-n source code with list decoder:

$$f_n \colon \mathcal{X}^n \times \mathcal{Y}^n \to \{1, \ldots, 2^{nR}\}, \quad \lambda_n \colon \{1, \ldots, 2^{nR}\} \times \mathcal{Y}^n \to 2^{\mathcal{X}^n}$$

- Lossless property:

$$x^n \in \lambda_n(f_n(x^n, y^n), y^n), \quad \forall (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$$

- ρ-th listsize moment:

$$\mathrm{E}[|\lambda_n(f_n(X^n, Y^n), Y^n)|^\rho]$$

Result for Lossless List Source Codes with Side-Information

Theorem

1. *If* $R > H_{\frac{1}{1+\rho}}(X|Y)$*, then there exists* $(f_n, \lambda_n)_{n\geq 1}$ *such that*

$$\lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n, Y^n), Y^n)|^\rho] = 1.$$

2. *If* $R < H_{\frac{1}{1+\rho}}(X|Y)$*, then*

$$\lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n, Y^n), Y^n)|^\rho] = \infty.$$

Here $H_{\frac{1}{1+\rho}}(X|Y)$ is defined to make this correct...

So $H_{\frac{1}{1+\rho}}(X|Y)$ is:

$$H_\alpha(X|Y) = \frac{\alpha}{1-\alpha} \log \sum_{y \in \mathcal{Y}} \left[\sum_{x \in \mathcal{X}} P_{X,Y}(x,y)^\alpha \right]^{1/\alpha}$$

Some Properties of $H_{\frac{1}{1+\rho}}(X|Y)$

1. Nondecreasing in $\rho > 0$
2. $\lim_{\rho \to 0} H_{\frac{1}{1+\rho}}(X|Y) = H(X|Y)$
3. $\lim_{\rho \to \infty} H_{\frac{1}{1+\rho}}(X|Y) = \max_y \log |\mathrm{supp}(P_{X|Y=y})|$
4. $H_{\frac{1}{1+\rho}}(X|Y) \leq H_{\frac{1}{1+\rho}}(X)$

Direct Part

1. Fix a side-information sequence y^n of type Q.
2. Partition each V-shell of y^n into 2^{nR} lists of lengths at most
$$\lceil 2^{-nR}|T_V(y^n)|\rceil \leq \lceil 2^{n(H(V|Q)-R)}\rceil.$$
3. Describe V and the list containing x^n using $nR + o(n)$ bits.
4. The ρ-th moment of the listsize can be upper-bounded by
$$\sum_{Q,V} \Pr\big((X^n, Y^n) \in T_{Q\circ V}\big)\lceil 2^{n(H(V|Q)-R)}\rceil^{\rho}$$
$$\leq 1 + 2^{-n\rho(R-\max_{Q,V}\{H(V|Q)-\rho^{-1}D(Q\circ V||P_{X,Y})\}-\delta_n)},$$
where $\delta_n \to 0$.
5. Complete the proof by showing that
$$H_{\frac{1}{1+\rho}}(X|Y) = \max_{Q,V}\{H(V|Q) - \rho^{-1}D(Q\circ V||P_{X,Y})\}. \quad \square$$

Conditional Rényi Entropy

$$H_\alpha(X|Y) = \frac{\alpha}{1-\alpha} \log \sum_{y\in\mathcal{Y}} \Big[\sum_{x\in\mathcal{X}} P_{X,Y}(x,y)^\alpha \Big]^{1/\alpha}$$

Suguru Arimoto

Arimoto's Motivation

- Define "capacity of order α" as

$$C_\alpha = \max_{P_X}\{H_\alpha(X) - H_\alpha(X|Y)\}$$

- Arimoto showed that

$$C_{\frac{1}{1+\rho}} = \frac{1}{\rho}\max_P E_0(\rho, P),$$

where $E_0(\rho, P)$ is Gallager's exponent function:

$$E_0(\rho, P) = -\log \sum_y \left[\sum_x P(x) W(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho},$$

- Gallager's random coding bound thus becomes

$$P_e \le \exp\Big(-n\rho\big(C_{\frac{1}{1+\rho}} - R\big)\Big), \quad 0 \le \rho \le 1.$$

List Source Coding with a Fidelity Criterion

1. Rate-R blocklength-n source code with list decoder:

$$f_n \colon \mathcal{X}^n \to \{1, \ldots, 2^{nR}\}, \quad \lambda_n \colon \{1, \ldots, 2^{nR}\} \to 2^{\hat{\mathcal{X}}^n}$$

2. Fidelity criterion:

$$d(f_n, \lambda_n) \triangleq \max_{x^n \in \mathcal{X}^n} \min_{\hat{x}^n \in \lambda_n(f_n(x^n))} d(x^n, \hat{x}^n) \leq D$$

3. ρ-th listsize moment:

$$\mathrm{E}[|\lambda_n(f_n(X^n))|^\rho]$$

A Rate-Distortion Theorem for List Source Codes

Theorem

1. *If* $R > R_\rho(D)$*, then there exists* $(f_n, \lambda_n)_{n \geq 1}$ *such that*

$$\sup_n d(f_n, \lambda_n) \leq D \quad \& \quad \lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n))|^\rho] = 1.$$

2. *If* $R < R_\rho(D)$ *and* $\limsup_{n\to\infty} d(f_n, \lambda_n) \leq D$*, then*

$$\lim_{n\to\infty} \mathsf{E}[|\lambda_n(f_n(X^n))|^\rho] = \infty.$$

But what is $R_\rho(D)$?

A Rényi Rate-Distortion Function

$$R_\rho(D) \triangleq \max_Q \{ R(Q, D) - \rho^{-1} D(Q||P) \},$$

where $R(Q, D)$ is the rate-distortion function of the source Q.

Direct Part

1. Type Covering Lemma: If $n \geq n_0(\delta)$, then for every type Q we can find $B_Q \subset \hat{\mathcal{X}}^n$ such that

$$|B_Q| \leq 2^{n(R(Q,D)+\delta)} \quad \text{and} \quad \max_{x^n \in T_Q} \min_{\hat{x}^n \in B_Q} d(x^n, \hat{x}^n) \leq D.$$

2. Partition each B_Q into 2^{nR} lists of lengths at most

$$\lceil 2^{n(R(Q,D)-R+\delta)} \rceil.$$

3. Use $nR + o(n)$ bits to describe the type Q of x^n and a list in the partition of B_Q that contains some $\hat{x}^n$ with $d(x^n, \hat{x}^n) \leq D$.
4. The ρ-th moment of the listsize can be upper-bounded by

$$\sum_Q \Pr(X^n \in T_Q) \lceil 2^{n(R(Q,D)-R+\delta)} \rceil^\rho$$
$$\leq 1 + 2^{-n\rho(R - \max_Q\{R(Q,D) - \rho^{-1} D(Q||P)\} - \delta - \delta_n)}.$$

□

Converse

1. WLOG assume $\lambda_n(m) \cap \lambda_n(m') = \emptyset$ if $m \neq m'$.
2. For each $\hat{x}^n \in \bigcup_{m=1}^{2^{nR}} \lambda_n(m)$ let $m(\hat{x}^n)$ be the unique index s.t. $\hat{x}^n \in \lambda_n(m(\hat{x}^n))$.
3. Define $g_n \colon \mathcal{X}^n \to \hat{\mathcal{X}}^n$ such that

$$g_n(x^n) \in \lambda_n(f_n(x^n)) \quad \text{and} \quad d(x^n, g_n(x^n)) \leq D, \quad \forall x.$$

4. Observe that

$$\begin{aligned}\sum_{x^n} P_X^n(x^n)|\lambda_n(f_n(x^n))|^\rho &= \sum_{\hat{x}^n} P_X^n\left(g_n^{-1}(\{\hat{x}^n\})\right)|\lambda_n(m(\hat{x}^n))|^\rho \\ &= \sum_{\hat{x}^n} \tilde{P}_n(\hat{x}^n)|\lambda_n(m(\hat{x}^n))|^\rho,\end{aligned}$$

where

$$\tilde{P}_n(\hat{x}^n) = P_X^n\left(g_n^{-1}(\{\hat{x}^n\})\right).$$

Converse contd.

5. Applying the lemma yields

$$\sum_{\hat{x}^n} \tilde{P}_n(\hat{x}^n)|\lambda_n(m(\hat{x}^n))|^\rho \geq 2^{-n\rho R}\, 2^{\rho H_{\frac{1}{1+\rho}}(\tilde{P}_n)}$$

6. It now suffices to show that

$$H_{\frac{1}{1+\rho}}(\tilde{P}_n) \geq nR_\rho(D).$$

7. The PMF $\tilde{P}_n$ can be written as

$$\tilde{P}_n = P_X^n \widetilde{W}_n,$$

   where

$$\widetilde{W}_n(\hat{x}^n|x^n) = 1\{\hat{x}^n = g_n(x^n)\}.$$

Converse contd.

8. Let $Q_\star$ achieve $R_\rho(D)$, i.e.,

$$R_\rho(D) = R(Q_\star, D) - \rho^{-1} D(Q_\star || P_X).$$

9. For every PMF Q on $\hat{\mathcal{X}}^n$

$$H_{\frac{1}{1+\rho}}(\tilde{P}_n) \geq H(Q) - \rho^{-1} D(Q || \tilde{P}_n).$$

10. Choosing $Q = Q_\star^n \widetilde{W}_n$ gives

$$\begin{aligned} H_{\frac{1}{1+\rho}}(\tilde{P}_n) &\geq H(Q_\star^n \widetilde{W}_n) - \rho^{-1} D(Q_\star^n \widetilde{W}_n || P_X^n \widetilde{W}_n) \\ &\geq H(Q_\star^n \widetilde{W}_n) - \rho^{-1} D(Q_\star^n || P_X^n) \quad \text{(Data processing)} \\ &= H(Q_\star^n \widetilde{W}_n) - n\rho^{-1} D(Q_\star || P_X). \end{aligned}$$

Converse contd.

11. Let $\tilde{X}^n$ be IID $\sim Q_\star$ and let $\hat{X}^n = g_n(\tilde{X}^n)$. Then

$$\begin{aligned} H(Q_\star^n \tilde{W}_n) &= H(\hat{X}^n) \\ &= I(\tilde{X}^n; \hat{X}^n). \end{aligned}$$

12. By construction of $g_n(\cdot)$

$$\mathrm{E}[d(\tilde{X}^n, \hat{X}^n)] \leq D.$$

13. From the converse to the Rate-Distortion Theorem it follows

$$I(\tilde{X}^n; \hat{X}^n) \geq nR(Q_\star, D).$$

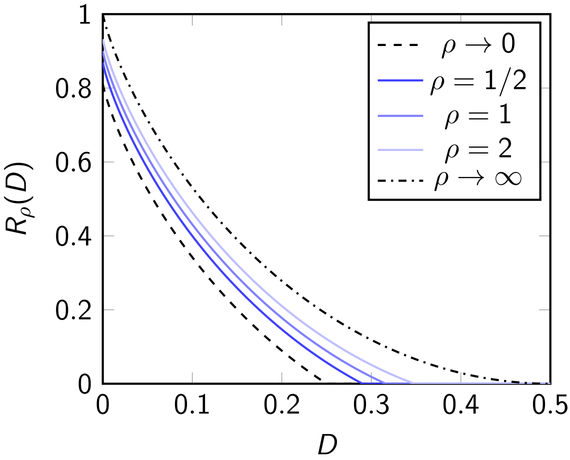
□

Example: Binary Source with Hamming Distortion

- $\mathcal{X} = \hat{\mathcal{X}} = \{0, 1\}$
- $\Pr(X_i = 1) = p$
- $d(x, \hat{x}) = 1\{x \neq \hat{x}\}$
- $R(D) = |h(p) - h(D)|^+$
- $R_\rho(D) = |H_{\frac{1}{1+\rho}}(p) - h(D)|^+$

where $|\xi|^+ = \max\{0, \xi\}$ and $h(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$.

Example: Binary Source with Hamming Distortion contd.



$R_\rho(D)$ plotted for binary source ($p = 1/4$) and Hamming distortion

This function Is also not New!

$$R_\rho(D) \triangleq \max_Q \{R(Q,D) - \rho^{-1} D(Q||P)\},$$

where $R(Q,D)$ is the rate-distortion function of the source Q.

Erdal Arıkan

Neri Merhav

Arıkan & Merhav's Motivation

- Let $\mathcal{G}_n = \{\hat{x}_1^n, \hat{x}_2^n, \ldots\}$ be an ordering of $\hat{\mathcal{X}}^n$.
- Define
$$G_n(x^n) = \min\{j : d(x^n, \hat{x}_j^n) \leq D\}.$$
- If $X_1, X_2, \ldots$ are IID $\sim P$, then
$$\lim_{n\to\infty} \frac{1}{n} \min_{\mathcal{G}_n} \log \mathrm{E}[G_n(X_1, \ldots, X_n)^\rho]^{1/\rho} = R_\rho(D).$$

To Recap

Replacing "messages" with "tasks" leads to new operational characterizations of

$$H_{\frac{1}{1+\rho}}(X) = \frac{1}{\rho}\log\Big[\sum_x P(x)^{\frac{1}{1+\rho}}\Big]^{1+\rho}$$

$$H_{\frac{1}{1+\rho}}(X|Y) = \frac{1}{\rho}\log\sum_y\Big[\sum_x P_{X,Y}(x,y)^{\frac{1}{1+\rho}}\Big]^{1+\rho}$$

$$R_\rho(D) = \max_Q\{R(Q,D) - \rho^{-1}D(Q||P)\}$$

for all $\rho > 0$.

To Recap

Replacing "messages" with "tasks" leads to new operational characterizations of

$$H_{\frac{1}{1+\rho}}(X) = \frac{1}{\rho}\log\Big[\sum_x P(x)^{\frac{1}{1+\rho}}\Big]^{1+\rho}$$

$$H_{\frac{1}{1+\rho}}(X|Y) = \frac{1}{\rho}\log\sum_y\Big[\sum_x P_{X,Y}(x,y)^{\frac{1}{1+\rho}}\Big]^{1+\rho}$$

$$R_\rho(D) = \max_Q\{R(Q,D) - \rho^{-1}D(Q||P)\}$$

for all $\rho > 0$.

Thank You!