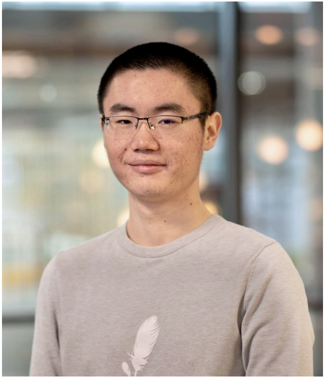
Soft Guessing Under Logarithmic Loss

Hamdi Joudeh
ICT Lab
Eindhoven University of Technology (TU/e)

Information Theory and Tapas Workshop
Madrid, Jan. 2023

Joint work with Han Wu (TU/e)



Guessing problem (Massey'94, Arikan'96)

A r.v. X is drawn from a finite set $\mathcal{X} = \{1, 2, \ldots, M\}$ according to pmf P. Assume $P(1) \geq P(2) \geq \cdots \geq P(M) > 0$.

A guesser seeks to determine X through a sequence of inquiries

$$\begin{gathered}\text{“is } X = x_1\text{?”}\\ \text{“is } X = x_2\text{?”}\\ \vdots\end{gathered}$$

until the answer is "yes".

Guessing problem (Massey'94, Arikan'96)

A r.v. X is drawn from a finite set $\mathcal{X} = \{1, 2, \ldots, M\}$ according to pmf P. Assume $P(1) \geq P(2) \geq \cdots \geq P(M) > 0$.

A guesser seeks to determine X through a sequence of inquiries

$$\text{“is } X = x_1\text{?”}$$
$$\text{“is } X = x_2\text{?”}$$
$$\vdots$$

until the answer is “yes”.

Guessing function: $G(x) \triangleq$ number of required guesses when $X = x$

Object of interest: distribution of $G(X)$

Guessing problem (Massey'94, Arikan'96)

A r.v. X is drawn from a finite set $\mathcal{X} = \{1, 2, \ldots, M\}$ according to pmf P. Assume $P(1) \geq P(2) \geq \cdots \geq P(M) > 0$.

A guesser seeks to determine X through a sequence of inquiries

$$\text{“is } X = x_1\text{?”}$$
$$\text{“is } X = x_2\text{?”}$$
$$\vdots$$

until the answer is "yes".

Guessing function: $G(x) \triangleq$ number of required guesses when $X = x$

Object of interest: distribution of $G(X)$

Motivation/Applications: security (password attacks), channel-coding (decoding effort), betting games, database search, etc.

Guessing moments

Guessing moments: The ρ-th guessing moment ($\rho > 0$) is defined as

$$\mathcal{M}_X(\rho) \triangleq \min_G \mathbb{E}\left[G(X)^\rho\right]$$

Guessing moments

Guessing moments: The ρ-th guessing moment ($\rho > 0$) is defined as

$$\mathcal{M}_X(\rho) \triangleq \min_G \mathbb{E}\left[G(X)^\rho\right]$$

The obvious guessing strategy simultaneously minimizes all moments

Theorem (Arikan'96). The ρ-th guessing moment ($\rho > 0$) satisfies

$$(1 + \log M)^{-\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P)\right) \leq \mathcal{M}_X(\rho) \leq \exp\left(\rho H_{\frac{1}{1+\rho}}(P)\right)$$

where the Rényi entropy of order $\alpha \in (0,1) \cup (1,\infty)$ is defined as

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

and remaining orders by cont. extension.

Guessing exponents

Asymptotics: Guessing a sequence $X^n \triangleq (X_1, \ldots, X_n)$ i.i.d. $\sim P$

Corollary (Arikan'96). The ρ-th guessing exponent is given by

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{M}_{X^n}(\rho) = \rho H_{\frac{1}{1+\rho}}(P)$$

for large n, we have $\mathcal{M}_{X^n}(\rho) \approx \exp\left(n\rho H_{\frac{1}{1+\rho}}(P)\right)$

Guessing exponents

Asymptotics: Guessing a sequence $X^n \triangleq (X_1, \ldots, X_n)$ i.i.d. $\sim P$

Corollary (Arikan'96). The ρ-th guessing exponent is given by

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{M}_{X^n}(\rho) = \rho H_{\frac{1}{1+\rho}}(P)$$

for large n, we have $\mathcal{M}_{X^n}(\rho) \approx \exp\left(n\rho H_{\frac{1}{1+\rho}}(P)\right)$

Why guessing moments/exponents?

- Tail probability. Chernoff bound:

$$\begin{aligned}\mathbb{P}\left[G(X^n) \geq \exp(n\gamma)\right] &\leq \inf_{\rho>0} \mathbb{E}\left[G(X^n)^\rho\right] e^{-n\rho\gamma} \\ &= \exp\left(-n \sup_{\rho>0}\left\{\rho\gamma - \frac{1}{n}\log \mathbb{E}\left[G(X^n)^\rho\right]\right\}\right)\end{aligned}$$

Lossy guessing

The goal is to guess a *reconstruction* $\hat{x} \in \hat{\mathcal{X}}$ of the r.v. X

- Loss/distortion measure: $\ell(x, \hat{x}) \geq 0$
- Lossy guessing strategy: sequence $(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N)$
- Stopping: $\ell(x, \hat{x}_u) \leq d$ for some acceptable $d \geq 0$, i.e.

$$\text{“is } \ell(x, \hat{x}_1) \leq d\text{?”}$$
$$\text{“is } \ell(x, \hat{x}_2) \leq d\text{?”}$$
$$\vdots$$

until the answer is “yes”.

Lossy guessing

The goal is to guess a *reconstruction* $\hat{x} \in \hat{\mathcal{X}}$ of the r.v. X

- Loss/distortion measure: $\ell(x, \hat{x}) \geq 0$
- Lossy guessing strategy: sequence $(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N)$
- Stopping: $\ell(x, \hat{x}_u) \leq d$ for some acceptable $d \geq 0$, i.e.

$$\text{"is } \ell(x, \hat{x}_1) \leq d\text{?"}$$
$$\text{"is } \ell(x, \hat{x}_2) \leq d\text{?"}$$
$$\vdots$$

  until the answer is "yes".

- d-admissibility: for every $x \in \mathcal{X}$, $\ell(x, \hat{x}_u) \leq d$ for some $\hat{x}_u$
- Guessing function:

$$G(x) \triangleq \text{smallest } u \in \{1, \ldots, N\} \text{ s.t. } \ell(x, \hat{x}_u) \leq d$$

- Guessing moment: $\mathcal{M}_X(d, \rho) \triangleq \min_G \mathbb{E}\left[G(X)^\rho\right]$

Guessing subject to distortion (Arikan-Merhav'98)

Asymptotics: The goal is to guess a reconstruction $\hat{x}^n \in \hat{\mathcal{X}}^n$ of an i.i.d. sequence X^n, subject to an additive distortion

$$\ell(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} \ell(x_i, \hat{x}_i)$$

Guessing subject to distortion (Arikan-Merhav'98)

Asymptotics: The goal is to guess a reconstruction $\hat{x}^n \in \hat{\mathcal{X}}^n$ of an i.i.d. sequence X^n, subject to an additive distortion

$$\ell(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} \ell(x_i, \hat{x}_i)$$

Theorem (Arikan-Merhav'98). ρ-th guessing exponent is given by

$$\lim_{n \to \infty} \frac{1}{n} \log \mathcal{M}_{X^n}(d, \rho) = \max_{Q} \{\rho R(Q, d) - D(Q \| P)\}$$

where $R(Q, d)$ is the rate-distortion function of DMS Q

Guessing subject to distortion (Arikan-Merhav'98)

Asymptotics: The goal is to guess a reconstruction $\hat{x}^n \in \hat{\mathcal{X}}^n$ of an i.i.d. sequence X^n, subject to an additive distortion

$$\ell(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} \ell(x_i, \hat{x}_i)$$

Theorem (Arikan-Merhav'98). ρ-th guessing exponent is given by

$$\lim_{n \to \infty} \frac{1}{n} \log \mathcal{M}_{X^n}(d, \rho) = \max_{Q} \{\rho R(Q, d) - D(Q \| P)\}$$

where $R(Q, d)$ is the rate-distortion function of DMS Q

- Under $d = 0$, we have $R(Q, d) = H(Q)$ and

$$\max_{Q} \{\rho H(Q) - D(Q \| P)\} = \rho H_{\frac{1}{1+\rho}}(P)$$

Soft guessing subject to logarithmic loss

Soft reconstruction and logarithmic loss

The goal is to guess a *soft reconstruction* $\hat{P}$ of the r.v. X

- Soft reconstruction: pmf $\hat{P} \in \mathcal{P}(\mathcal{X})$
- Think of $\hat{P}$ as a posterior for X (prior is P).

Soft reconstruction and logarithmic loss

The goal is to guess a *soft reconstruction* $\hat{P}$ of the r.v. X

- Soft reconstruction: pmf $\hat{P} \in \mathcal{P}(\mathcal{X})$
- Think of $\hat{P}$ as a posterior for X (prior is P).

Logarithmic loss: The loss of reconstructing x as $\hat{P}$ is

$$\ell(x, \hat{P}) \triangleq \log \frac{1}{\hat{P}(x)}$$

$\ell(x, \hat{P}) \geq 0$ with equality iff $\hat{P}$ is a hard reconstruction of x

Soft reconstruction and logarithmic loss

The goal is to guess a *soft reconstruction* $\hat{P}$ of the r.v. X

- Soft reconstruction: pmf $\hat{P} \in \mathcal{P}(\mathcal{X})$
- Think of $\hat{P}$ as a posterior for X (prior is P).

Logarithmic loss: The loss of reconstructing x as $\hat{P}$ is

$$\ell(x, \hat{P}) \triangleq \log \frac{1}{\hat{P}(x)}$$

$\ell(x, \hat{P}) \geq 0$ with equality iff $\hat{P}$ is a hard reconstruction of x

- Logarithmic loss $=$ information: $\ell(x, \hat{P}) = \imath_{\hat{P}}(x)$
- For any $d \geq 0$, x is d-covered by $\hat{P}$ whenever $\ell(x, \hat{P}) \leq d$

Soft guessing

Soft guessing strategy: sequence of pmfs $(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_N)$. For an acceptable loss level d, soft guessing goes as:

$$\text{"is } \ell(x, \hat{P}_1) \leq d\text{?"}$$
$$\text{"is } \ell(x, \hat{P}_2) \leq d\text{?"}$$
$$\vdots$$

until the answer is "yes".

Soft guessing

Soft guessing strategy: sequence of pmfs $(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_N)$. For an acceptable loss level d, soft guessing goes as:

$$\text{“is } \ell(x, \hat{P}_1) \leq d\text{?”}$$
$$\text{“is } \ell(x, \hat{P}_2) \leq d\text{?”}$$
$$\vdots$$

until the answer is “yes”.

- d-admissibility: every $x \in \mathcal{X}$ is d-covered by at least one $\hat{P}_u$
- Guessing function:

$$G(x) \triangleq \text{smallest index } u \in \{1, 2, \ldots, N\} \text{ s.t. } \ell(x, \hat{P}_u) \leq d$$

- Any *good* strategy should have $N \leq M = |\mathcal{X}|$ (why?)
- For $d = \log M$, how many guesses do we need?

Exponent under logarithmic loss

Asymptotics: For i.i.d. sequences, take $\hat{P}^n(x^n) = \prod_{i=1}^n \hat{P}(x_i)$ and

$$\ell\left(x^n, \hat{P}^n\right) = \frac{1}{n}\sum_{i=1}^n \ell\left(x_i, \hat{P}\right) = \frac{1}{n}\sum_{i=1}^n \log\frac{1}{\hat{P}(x_i)}$$

Exponent under logarithmic loss

Asymptotics: For i.i.d. sequences, take $\hat{P}^n(x^n) = \prod_{i=1}^n \hat{P}(x_i)$ and

$$\ell\left(x^n, \hat{P}^n\right) = \frac{1}{n}\sum_{i=1}^{n} \ell\left(x_i, \hat{P}\right) = \frac{1}{n}\sum_{i=1}^{n} \log\frac{1}{\hat{P}(x_i)}$$

Rate-distortion function (Courtade-Weissman'14):

$$R(Q, d) = H(Q) - d$$

Exponent under logarithmic loss

Asymptotics: For i.i.d. sequences, take $\hat{P}^n(x^n) = \prod_{i=1}^n \hat{P}(x_i)$ and

$$\ell\left(x^n, \hat{P}^n\right) = \frac{1}{n}\sum_{i=1}^n \ell\left(x_i, \hat{P}\right) = \frac{1}{n}\sum_{i=1}^n \log\frac{1}{\hat{P}(x_i)}$$

Rate-distortion function (Courtade-Weissman'14):

$$R(Q,d) = H(Q) - d$$

From (Arikan-Merhav'98) we get

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\mathcal{M}_{X^n}(d,\rho) &= \max_Q\{\rho H(Q) - \rho d - D(Q\|P)\}\\ &= \rho H_{\frac{1}{1+\rho}}(P) - \rho d\end{aligned}$$

Next: Single-shot version with no random selection (covering).

Main Result

Theorem. Define $d' \triangleq \log\lfloor \exp(d) \rfloor$. The following bounds hold

$$\mathcal{M}_X(d,\rho) \geq (1+\log M)^{-\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

and

$$\mathcal{M}_X(d,\rho) \leq 1 + 2^{\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

Main Result

Theorem. Define $d' \triangleq \log\lfloor \exp(d) \rfloor$. The following bounds hold

$$\mathcal{M}_X(d,\rho) \geq (1+\log M)^{-\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

and

$$\mathcal{M}_X(d,\rho) \leq 1 + 2^{\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

Shkel-Verdú'18: Lossy compression under log-loss $\iff$ Lossless compression + list decoding

Here:
Lossy guessing under log-loss $\iff$ lossless guessing + list decoding

Lower bound

$$\mathcal{M}_X(d,\rho) \geq (1+\log M)^{-\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

Covering under logarithmic loss

Set of realizations d-covered by $\hat{P}$:

$$\mathcal{S}_d(\hat{P}) \triangleq \left\{ x \in \mathcal{X} : \ell(x, \hat{P}) \leq d \right\}$$

Covering under logarithmic loss

Set of realizations d-covered by $\hat{P}$:

$$\mathcal{S}_d(\hat{P}) \triangleq \left\{x \in \mathcal{X} : \ell(x, \hat{P}) \leq d\right\}$$

Lemma (Shkel-Verdú'18). $|\mathcal{S}_d(\hat{P})| \leq \lfloor \exp(d) \rfloor$

Proof: Recall that $x \in \mathcal{S}_d(\hat{P}) \iff \hat{P}(x) \geq \exp(-d)$. Then

$$1 = \sum_{x \in \mathcal{X}} \hat{P}(x) \geq \sum_{x \in \mathcal{S}_d(\hat{P})} \hat{P}(x) \geq |\mathcal{S}_d(\hat{P})| \exp(-d).$$

Covering under logarithmic loss

Set of realizations d-covered by $\hat{P}$:

$$\mathcal{S}_d(\hat{P}) \triangleq \left\{ x \in \mathcal{X} : \ell(x, \hat{P}) \leq d \right\}$$

Lemma (Shkel-Verdú'18). $|\mathcal{S}_d(\hat{P})| \leq \lfloor \exp(d) \rfloor$

Proof: Recall that $x \in \mathcal{S}_d(\hat{P}) \iff \hat{P}(x) \geq \exp(-d)$. Then

$$1 = \sum_{x \in \mathcal{X}} \hat{P}(x) \geq \sum_{x \in \mathcal{S}_d(\hat{P})} \hat{P}(x) \geq |\mathcal{S}_d(\hat{P})| \exp(-d).$$

Corollary. We need at least $\left\lceil \frac{M}{\lfloor \exp(d) \rfloor} \right\rceil$ reconstructions to d-cover $\mathcal{X}$

Equivocation bound

For a d-admissible strategy given that we know $G(X)$, what is the remaining uncertainty about X?

Lemma. For $G(X)$ induced by a d-admissible strategy, we have

$$H(X|G(X)) \leq \log\lfloor \exp(d) \rfloor = d'$$

Equivocation bound

For a d-admissible strategy given that we know $G(X)$, what is the remaining uncertainty about X?

Lemma. For $G(X)$ induced by a d-admissible strategy, we have

$$H(X|G(X)) \leq \log\lfloor\exp(d)\rfloor = d'$$

Proof: From d-admissibility, we have

$$G^{-1}(u) \triangleq \{x \in \mathcal{X} : G(x) = u\} \subseteq \mathcal{S}_d(\hat{P}_u)$$

Therefore

$$\begin{aligned} H(X|G(X) = u) &\leq \log|G^{-1}(u)| \\ &\leq \log|\mathcal{S}_d(\hat{P}_u)| \\ \text{Shkel-Verdú} \quad &\leq \log\lfloor\exp(d)\rfloor \end{aligned}$$

Expected log of integer r.v.

Lemma (Arikan'96). Let $U \sim Q$ be a r.v. on $\{1, 2, \ldots, M\}$. Then

$$\mathbb{E}\left[\log U\right] \geq H(U) - \log(1 + \log M)$$

Expected log of integer r.v.

Lemma (Arikan'96). Let $U \sim Q$ be a r.v. on $\{1, 2, \ldots, M\}$. Then

$$\mathbb{E}[\log U] \geq H(U) - \log(1 + \log M)$$

Proof: Define $c \triangleq \sum_{i=1}^{M} \frac{1}{i}$ and the pmf

$$\hat{Q}(u) = \frac{1}{cu}, \ u \in \{1, 2, \ldots, M\}$$

which is well defined since $c \leq 1 + \log M$. We have

$$\begin{aligned} \mathbb{E}[\log U] &= \mathbb{E}\left[\log \frac{1}{\hat{Q}(U)}\right] - \log c \\ &= H(U) + D(Q\|\hat{Q}) - \log c \\ &\geq H(U) - \log(1 + \log M) \end{aligned}$$

Proof of lower bound

Arikan's lower bound:

Let Q be an arbitrary pmf on $\mathcal{X}$ such that $Q \ll P$. In what follows, we have $X \sim P$ and $X' \sim Q$.

Proof of lower bound

Arikan's lower bound:

Let Q be an arbitrary pmf on $\mathcal{X}$ such that $Q \ll P$. In what follows, we have $X \sim P$ and $X' \sim Q$.

$$\begin{aligned}
\mathbb{E}\left[G(X)^\rho\right] &= \sum_{x\in\mathcal{X}} P(x)G(x)^\rho \\
&= \sum_{x\in\mathcal{X}} Q(x)\exp\left(-\log\left(\frac{Q(x)}{P(x)G(x)^\rho}\right)\right) \\
\text{Jensen} \quad &\geq \exp\left(-\sum_{x\in\mathcal{X}} Q(x)\log\left(\frac{Q(x)}{P(x)G(x)^\rho}\right)\right) \\
&= \exp\left(-D(Q\|P) + \rho\mathbb{E}\left[\log G(X')\right]\right)
\end{aligned}$$

Proof of lower bound

Arikan's lower bound:

Let Q be an arbitrary pmf on $\mathcal{X}$ such that $Q \ll P$. In what follows, we have $X \sim P$ and $X' \sim Q$.

$$\begin{aligned}
\mathbb{E}\left[G(X)^\rho\right] &= \sum_{x\in\mathcal{X}} P(x)G(x)^\rho \\
&= \sum_{x\in\mathcal{X}} Q(x)\exp\left(-\log\left(\frac{Q(x)}{P(x)G(x)^\rho}\right)\right) \\
\text{Jensen} \quad &\geq \exp\left(-\sum_{x\in\mathcal{X}} Q(x)\log\left(\frac{Q(x)}{P(x)G(x)^\rho}\right)\right) \\
&= \exp\left(-D(Q\|P) + \rho\mathbb{E}\left[\log G(X')\right]\right)
\end{aligned}$$

Next we deal with the term $\mathbb{E}\left[\log G(X')\right]$

Proof of lower bound (cont.)

$G(X')$ is a r.v. defined on $\{1, 2, \ldots, M\}$. Hence:

E-log-int $\quad \mathbb{E}\left[\log G(X')\right] \geq H(G(X')) - \log(1 + \log M)$

Proof of lower bound (cont.)

$G(X')$ is a r.v. defined on $\{1, 2, \ldots, M\}$. Hence:

$$\text{E-log-int} \quad \mathbb{E}\left[\log G(X')\right] \geq H(G(X')) - \log(1 + \log M)$$

The entropy is bounded as:

$$\begin{aligned} H(G(X')) &= I(X'; G(X')) \\ &= H(X') - H(X'|G(X')) \\ \text{Equivocation} \quad &\geq H(X') - d' \\ &= H(Q) - d' \end{aligned}$$

Proof of lower bound (cont.)

$G(X')$ is a r.v. defined on $\{1, 2, \ldots, M\}$. Hence:

E-log-int $\quad \mathbb{E}\left[\log G(X')\right] \geq H(G(X')) - \log(1 + \log M)$

The entropy is bounded as:

$$\begin{aligned} H(G(X')) &= I(X'; G(X')) \\ &= H(X') - H(X'|G(X')) \\ \text{Equivocation} \quad &\geq H(X') - d' \\ &= H(Q) - d' \end{aligned}$$

Combining bounds and tightening w.r.t. Q:

$$\begin{aligned} \mathbb{E}\left[G(X)^\rho\right] &\geq (1 + \log M)^{-\rho} \exp\left(-D(Q\|P) + \rho H(Q) - \rho d'\right) \\ &\geq (1 + \log M)^{-\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right) \end{aligned}$$

Upper bound

$$\mathcal{M}_X(d, \rho) \leq 1 + 2^{\rho} \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho d'\right)$$

List guessing

Any guess d-covers at most $L = \lfloor \exp(d) \rfloor$ elements. Smallest number of guesses to d-cover $\mathcal{X}$ is $N = \lceil M/L \rceil$. Partition $\mathcal{X}$ into N lists:

$$\begin{aligned}
\mathcal{L}_1 &= \{1, 2, \ldots, L\} \\
\mathcal{L}_2 &= \{L+1, L+2, \ldots, 2L\} \\
&\vdots \\
\mathcal{L}_N &= \{(N-1)L+1, \ldots, M\}
\end{aligned}$$

List guessing

Any guess d-covers at most $L = \lfloor \exp(d) \rfloor$ elements. Smallest number of guesses to d-cover $\mathcal{X}$ is $N = \lceil M/L \rceil$. Partition $\mathcal{X}$ into N lists:

$$\begin{aligned}
\mathcal{L}_1 &= \{1, 2, \ldots, L\} \\
\mathcal{L}_2 &= \{L+1, L+2, \ldots, 2L\} \\
&\vdots \\
\mathcal{L}_N &= \{(N-1)L+1, \ldots, M\}
\end{aligned}$$

For each $u \in \{1, 2, \ldots, N\}$, define $\hat{P}_u$ as

$$\hat{P}_u(x) = \frac{\mathbb{1}\left[x \in \mathcal{L}_u\right]}{|\mathcal{L}_u|}, \quad x \in \mathcal{X}$$

List guessing

Any guess d-covers at most $L = \lfloor \exp(d) \rfloor$ elements. Smallest number of guesses to d-cover $\mathcal{X}$ is $N = \lceil M/L \rceil$. Partition $\mathcal{X}$ into N lists:

$$\begin{aligned}
\mathcal{L}_1 &= \{1, 2, \ldots, L\} \\
\mathcal{L}_2 &= \{L+1, L+2, \ldots, 2L\} \\
&\vdots \\
\mathcal{L}_N &= \{(N-1)L+1, \ldots, M\}
\end{aligned}$$

For each $u \in \{1, 2, \ldots, N\}$, define $\hat{P}_u$ as

$$\hat{P}_u(x) = \frac{\mathbb{1}\left[x \in \mathcal{L}_u\right]}{|\mathcal{L}_u|}, \quad x \in \mathcal{X}$$

- The above strategy is d-admissible
- Guessing function: $G(x) = \left\lceil \frac{x}{L} \right\rceil, \ x \in \mathcal{X}$

Proof of upper bound

Next we wish to upper bound $\mathbb{E}\left[G(X)^{\rho}\right] = \mathbb{E}\left[\left\lceil \frac{X}{L} \right\rceil^{\rho}\right]$. To this end, we start with upper bounding $\mathbb{E}\left[X^{\rho}\right]$

Proof of upper bound

Next we wish to upper bound $\mathbb{E}\left[G(X)^\rho\right] = \mathbb{E}\left[\left\lceil \frac{X}{L} \right\rceil^\rho\right]$. To this end, we start with upper bounding $\mathbb{E}\left[X^\rho\right]$

Arikan's upper bound:

Recall that $P(x') \geq P(x)$ for all $x' \leq x$, and hence

$$x = \sum_{x' \in \mathcal{X}} \mathbb{1}\left[x' \leq x\right] \leq \sum_{x' \in \mathcal{X}} \left(\frac{P(x')}{P(x)}\right)^{\frac{1}{1+\rho}}$$

Proof of upper bound

Next we wish to upper bound $\mathbb{E}\left[G(X)^\rho\right] = \mathbb{E}\left[\left\lceil \frac{X}{L} \right\rceil^\rho\right]$. To this end, we start with upper bounding $\mathbb{E}\left[X^\rho\right]$

Arikan's upper bound:

Recall that $P(x') \geq P(x)$ for all $x' \leq x$, and hence

$$x = \sum_{x' \in \mathcal{X}} \mathbb{1}\left[x' \leq x\right] \leq \sum_{x' \in \mathcal{X}} \left(\frac{P(x')}{P(x)}\right)^{\frac{1}{1+\rho}}$$

Taking the expectation of the ρ-th power, we get

$$\begin{aligned}\mathbb{E}\left[X^\rho\right] &\leq \sum_{x \in \mathcal{X}} P(x) \left(\sum_{x' \in \mathcal{X}} \left(\frac{P(x')}{P(x)}\right)^{\frac{1}{1+\rho}}\right)^\rho \\ &= \exp\left(\rho H_{\frac{1}{1+\rho}}(P)\right)\end{aligned}$$

Proof of upper bound (cont.)

The ρ-th power of the guessing function is bounded as:

$$\begin{aligned} G(x)^\rho &= \left\lceil \frac{x}{L} \right\rceil^\rho \\ &\le 1 + 2^\rho \left(\frac{x}{L}\right)^\rho \end{aligned}$$

since $\lceil z \rceil^\rho \le \max\{1, 2z\}^\rho \le 1 + 2^\rho z^\rho$ (Bunte-Lapidoth'14)

Proof of upper bound (cont.)

The ρ-th power of the guessing function is bounded as:

$$
\begin{aligned}
G(x)^\rho &= \left\lceil \frac{x}{L} \right\rceil^\rho \\
&\leq 1 + 2^\rho \left(\frac{x}{L}\right)^\rho
\end{aligned}
$$

since $\lceil z \rceil^\rho \leq \max\{1, 2z\}^\rho \leq 1 + 2^\rho z^\rho$ (Bunte-Lapidoth'14)

Taking the expectation:

$$
\begin{aligned}
\mathbb{E}\left[G(X)^\rho\right] &= \mathbb{E}\left[\left\lceil \frac{X}{L} \right\rceil^\rho\right] \\
&\leq 1 + \left(\frac{2}{L}\right)^\rho \mathbb{E}\left[X^\rho\right] \\
\text{Arikan} &\leq 1 + \left(\frac{2}{L}\right)^\rho \exp\left(\rho H_{\frac{1}{1+\rho}}(P)\right) \\
&= 1 + 2^\rho \exp\left(\rho H_{\frac{1}{1+\rho}}(P) - \rho \log\lfloor \exp(d) \rfloor\right)
\end{aligned}
$$

Some concluding remark

- Log-loss is a cheat code for lossy source coding: lossless source coding + list decoding, no random selection.
- Log-loss is a cheat code for lossy guessing.
- Result with side information Y should follow similarly (in terms of Arimoto's conditional Rényi entropy).
- Distributed encoders, compressed side information, etc.