Guessing, Source Coding, and Channel Decoding

Hamdi Joudeh

ICT Lab, EE Department
Eindhoven University of Technology

TU/e

Workshop on Information Theory for Future Networks
Karlsruhe Institute of Technology, 04.03.2026

Guessing

- Random variable $X \sim P$ with finite support $\mathcal{X} := \{x_1, x_2, \ldots, x_K\}$
- Guess the value of X through a sequence of queries:

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

  This goes on until the answer is “Yes”
- E.g. password cracking (dictionary attack)

What would be a good guessing strategy?

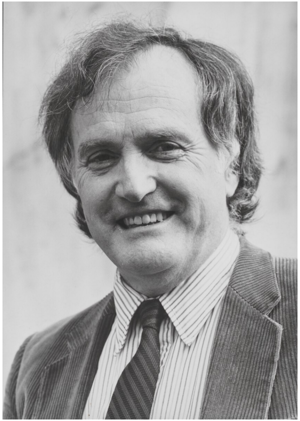
Optimal Guessing

- Guessing strategy: $G : \mathcal{X} \to \{1, 2, \ldots, K\}$
- $G(X)$ is the number of queries to correctly guess X (guesswork)
- Optimal strategy $G^\star$ queries in the order of likelihood

$$P(x) > P(x') \implies G^\star(x) < G^\star(x')$$
$$P(x) \geq P(x') \impliedby G^\star(x) < G^\star(x')$$

- Optimal in a "competitive" sense

$$\mathbb{P}\left[G^\star(X) > m\right] \leq \mathbb{P}\left[G(X) > m\right]$$



James Massey

$$\mathbb{E}[G^\star(X)]$$

Lower bound via entropy [ISIT'94]

Erdal Arıkan

$$\mathbb{E}[G^\star(X)^\rho], \ \rho > 0$$

Lower/upper bounds via Rényi entropy [Trans.IT'96]

Why Guessing Moments?

Tail probability:

$$\begin{aligned}\mathbb{P}\left[G^\star(X) > m\right] &= \mathbb{P}\left[G^\star(X)^\rho > m^\rho\right] && \rho > 0\\ &\le m^{-\rho}\mathbb{E}\left[G^\star(X)^\rho\right] && \text{Markov}\end{aligned}$$

- Chernoff bound (for $\log G^\star(X)$)
- min over ρ yields exponentially-tight bound (e.g. guessing X^n i.i.d.)

Complementary perspective:

$$\frac{1}{\rho}\log\mathbb{E}\left[G^\star(X)^\rho\right]$$

- $\rho = 1$: $\log\mathbb{E}\left[G^\star(X)\right]$
- $\rho \to 0$: $\mathbb{E}\left[\log G^\star(X)\right]$
- $\rho \to \infty$: $\max_{x\in\mathcal{X}} \log G^\star(x)$

Simple Upper Bound

$$
\begin{aligned}
G^\star(x) &= \sum_{i \in \mathcal{X}} \mathbb{1}\left[G^\star(i) \leq G^\star(x)\right] \\
&\leq \sum_{i \in \mathcal{X}} \mathbb{1}\left[P(i) \geq P(x)\right] \\
&= \sum_{i \in \mathcal{X}} \mathbb{1}\left[\frac{P(i)}{P(x)} \geq 1\right] \\
&\leq \sum_{i \in \mathcal{X}} \frac{P(i)}{P(x)} \\
&= \frac{1}{P(x)}
\end{aligned}
$$

Then

$$
\mathbb{E}\left[\log G^\star(X)\right] \leq \mathbb{E}\left[\log \frac{1}{P(X)}\right] = H(X)
$$

Aka Wyner's inequality

Slightly Less Simple Upper Bound

$$G^\star(x) \le \sum_{i\in\mathcal{X}} \mathbb{1}\left[\left(\frac{P(i)}{P(x)}\right)^\alpha \ge 1\right] \qquad \text{set } \alpha \ge 0$$
$$\le \sum_{i\in\mathcal{X}} \left(\frac{P(i)}{P(x)}\right)^\alpha$$

Then

$$\mathbb{E}\left[G^\star(X)^\rho\right] \le \sum_{x\in\mathcal{X}} P(x)\left(\sum_{i\in\mathcal{X}}\left(\frac{P(i)}{P(x)}\right)^\alpha\right)^\rho$$
$$= \sum_{x\in\mathcal{X}} P(x)^{1-\rho\alpha} \times \left(\sum_{i\in\mathcal{X}} P(i)^\alpha\right)^\rho$$
$$= \left(\sum_{x\in\mathcal{X}} P(x)^{\frac{1}{1+\rho}}\right)^{1+\rho} \qquad \text{set } \alpha = 1/(1+\rho)$$

Relation to Rényi entropy

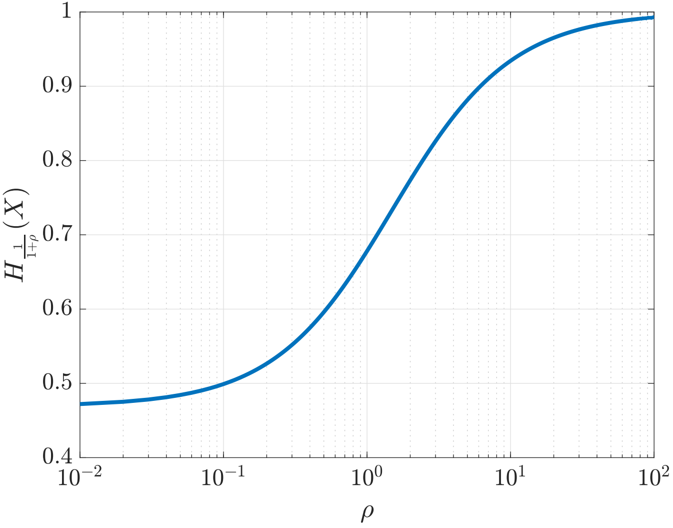
Rényi entropy:

$$H_\alpha(X) := \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} P(x)^\alpha$$

- $\alpha = 0$: $H_\alpha(X) = \log |\mathcal{X}|$
- $\alpha \to 1$: $H_\alpha(X) \to H(X)$

Back to the bound:

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



$$\frac{1}{\rho}\log \mathbb{E}\left[G^{\star}(X)^{\rho}\right] \leq H_{\frac{1}{1+\rho}}(X)$$

$$\rho = 0 \implies H(X)$$

$$\rho = 1 \implies H_{\frac{1}{2}}(X)$$

$$\rho \to \infty \implies \log|\mathcal{X}|$$

Lower Bound and Asymptotics

Arıkan's lower bound:

$$\frac{1}{\rho}\log\mathbb{E}\left[G(X)^{\rho}\right] \geq H_{\frac{1}{1+\rho}}(X) - \log(1+\ln|\mathcal{X}|)$$

Tensorization: for i.i.d. sequence $\boldsymbol{X} := X_1, \ldots, X_n$

$$H_{\frac{1}{1+\rho}}(\boldsymbol{X}) = nH_{\frac{1}{1+\rho}}(X)$$

Therefore

$$\frac{1}{n}\log\mathbb{E}\left[G(\boldsymbol{X})^{\rho}\right] \to \rho H_{\frac{1}{1+\rho}}(X)$$

New lower bound [H.J-Wu, unpublished]:

$$\frac{1}{\rho}\log\mathbb{E}\left[G(X)^{\rho}\right] \geq H_{\frac{1}{1+\rho}}(X) - \log\Big(1 + H(X_{\frac{1}{1+\rho}})\Big) - \log e$$

Source Coding

$$\boldsymbol{Z} \longrightarrow \boxed{\text{Enc}} \xrightarrow{f(\boldsymbol{Z})} \boxed{\text{Dec}} \longrightarrow \phi(f(\boldsymbol{Z}))$$

- Encode $\boldsymbol{Z} := Z_1, \ldots, Z_n$ into a message from $\{1, 2, \ldots, m\}$
- Code (i.e. compression) rate $r := \frac{1}{n} \log m$
- Fixed-length, almost lossless

$$p_e := \mathbb{P}\left[\boldsymbol{Z} \neq \phi(f(\boldsymbol{Z}))\right]$$

- Optimal code: uniquely represent m most likely source sequences
- Minimal error probability:

$$p_e^\star = \mathbb{P}\left[G^\star(\boldsymbol{Z}) > m\right]$$

Source Coding: Error Exponent

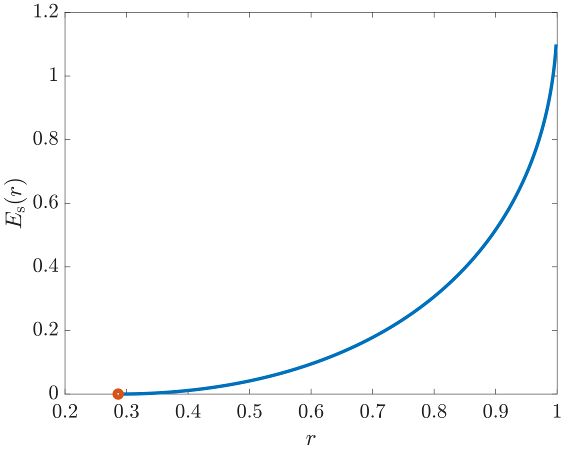
$$\begin{aligned} p_e^\star &= \mathbb{P}\left[G^\star(\boldsymbol{Z}) > m\right] \\ &\le m^{-\rho}\mathbb{E}\left[G^\star(\boldsymbol{Z})^\rho\right] \\ &\le m^{-\rho}\exp\left(\rho H_{\frac{1}{1+\rho}}(\boldsymbol{Z})\right) \end{aligned}$$

- Under i.i.d. $\boldsymbol{Z} := Z_1, Z_2, \ldots, Z_n$ and $r := \frac{1}{n}\log m$, we get

$$\frac{1}{n}\log\frac{1}{p_e^\star} \ge \max_{\rho>0} \rho\left(r - H_{\frac{1}{1+\rho}}(Z)\right)$$

- The above achievable exponent is tight [Jelinek'68]

$$E_s(r) := \lim_{n\to\infty}\frac{1}{n}\log\frac{1}{p_e^\star} = \max_{\rho>0}\rho\left(r - H_{\frac{1}{1+\rho}}(Z)\right)$$



Reliability function for Bernoulli source with $p = 0.05$ ($H(Z) = 0.2864$)

Channel Coding

$$W \longrightarrow \boxed{\text{Enc}} \xrightarrow{\boldsymbol{x}(W)} \boxed{P_{Y|X}} \xrightarrow{\boldsymbol{Y}} \boxed{\text{Dec}} \longrightarrow \varphi(\boldsymbol{Y})$$

- q-ary additive channel $(\mathcal{X}, \mathcal{Y}, \mathcal{Z} = \mathcal{A} := \{0, 1, \dots, q-1\})$

$$Y = X \oplus Z$$

- Codebook $\boldsymbol{\mathcal{C}} = \{\boldsymbol{x}(1), \dots, \boldsymbol{x}(M)\}$. Average decoding error

$$p_e(\boldsymbol{\mathcal{C}}, \varphi) = \frac{1}{M} \sum_{w \in [M]} \mathbb{P}\left[\varphi(\boldsymbol{Y}) \neq w \mid \boldsymbol{X} = \boldsymbol{x}(w)\right]$$

- ML decoding

$$\varphi^\star(\boldsymbol{y}) = \arg\max_{\hat{w} \in [M]} P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y} | \boldsymbol{x}(\hat{w}))$$

we call this a testing-based decoder

Guessing-Based Decoding

- E.g. codebook $\mathcal{C} := \{0000, 0101, 1010, 1111\}$
- Receiver observes $\boldsymbol{y} = 0100$
- Guessing:

  is 0100 in $\mathcal{C}$? No

  is 1100 in $\mathcal{C}$? No

  is 0000 in $\mathcal{C}$? Yes !! $\implies$ Decoder decides $\hat{\boldsymbol{x}} = 0000$
- In general, guessing continues $(2, 3, \ldots$ bit flips) until a codeword is identified
- Known as GRAND, introduced in this form by Duffy *et al.* [Trans.IT'19] (core idea can be traced back to at least Chase [Trans.IT'72] and others)
- Equivalent to ML decoding for BSC ($p < 0.5$)

General Procedure

- Rank all noise sequences from most to least likely $\boldsymbol{z}(1), \boldsymbol{z}(2), \boldsymbol{z}(3), \ldots$
- Given $\boldsymbol{y}$, sequentially test noise sequences:

  is $\boldsymbol{y} \ominus \boldsymbol{z}(1)$ in $\mathcal{C}$?

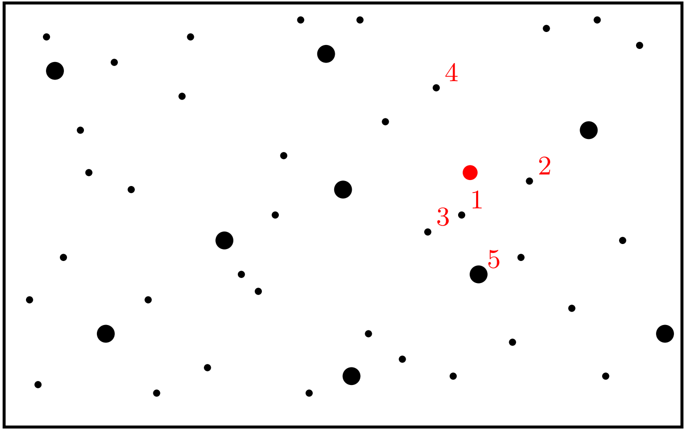
  is $\boldsymbol{y} \ominus \boldsymbol{z}(2)$ in $\mathcal{C}$?

  is $\boldsymbol{y} \ominus \boldsymbol{z}(3)$ in $\mathcal{C}$?

  $\vdots$

  until a codeword $\hat{\boldsymbol{x}} = \boldsymbol{y} \ominus \boldsymbol{z}(i)$ is encountered
- Equivalence to ML: every other codeword $\bar{\boldsymbol{x}} \in \mathcal{C}$ satisfies

$$G^\star(\boldsymbol{y} \ominus \hat{\boldsymbol{x}}) < G^\star(\boldsymbol{y} \ominus \bar{\boldsymbol{x}}) \implies P_Z(\boldsymbol{y} \ominus \hat{\boldsymbol{x}}) \geq P_Z(\boldsymbol{y} \ominus \bar{\boldsymbol{x}}) \iff P_{Y|X}(\boldsymbol{y}|\hat{\boldsymbol{x}}) \geq P_{Y|X}(\boldsymbol{y}|\bar{\boldsymbol{x}})$$


4
2
1
3
5

Random Coding Bound

Theorem [H.J, ISIT'24]

$$\bar{p}_e \leq \mathbb{E}\left[\min\left\{1, (M-1)\frac{G(\boldsymbol{Z})}{|\boldsymbol{\mathcal{A}}|}\right\}\right]$$

where $\bar{p}_e$ is the ensemble average error probability (uniform ensemble)

Proof.

- Suppose $\boldsymbol{x}(1)$ is sent and $\boldsymbol{y} = \boldsymbol{x}(1) \oplus \boldsymbol{z}$ is received. Error occurs if

$$G(\bar{\boldsymbol{z}}) < G(\boldsymbol{z}), \text{for some } \bar{\boldsymbol{z}} \text{ s.t. } \boldsymbol{y} \ominus \bar{\boldsymbol{z}} \in \boldsymbol{\mathcal{C}}$$

- Above condition equivalent to

$$G(\boldsymbol{x}(1) \oplus \boldsymbol{z} \ominus \boldsymbol{x}(\bar{w})) < G(\boldsymbol{z}), \text{for some } \bar{w} \neq 1$$

- Error probability given that message $\boldsymbol{x}(1)$ is sent

$$p_{e,1}(\mathcal{C}) = \sum_{\boldsymbol{z}} P_{\boldsymbol{Z}}(\boldsymbol{z}) \mathbb{1}\left[G(\boldsymbol{z} \oplus \boldsymbol{x}(1) \ominus \boldsymbol{x}(\bar{w})) \leq G(\boldsymbol{z}), \text{for some } \bar{w} \neq 1\right]$$

- Ensemble average conditioned on $\boldsymbol{X}(1) = \boldsymbol{x}(1)$

$$\begin{aligned}
\bar{p}_{e,1} &= \sum_{\boldsymbol{z}} P_{\boldsymbol{Z}}(\boldsymbol{z}) \mathbb{P}\left[G(\boldsymbol{z} \oplus \boldsymbol{x}(1) \ominus \boldsymbol{X}(\bar{w})) \leq G(\boldsymbol{z}), \text{for some } \bar{w} \neq 1\right] \\
&\leq \sum_{\boldsymbol{z}} P_{\boldsymbol{Z}}(\boldsymbol{z}) \min\left\{1, \sum_{\bar{w} \neq 1} \mathbb{P}\Big[G(\boldsymbol{z} \oplus \boldsymbol{x}(1) \ominus \boldsymbol{X}(\bar{w})) \leq G(\boldsymbol{z})\Big]\right\} \\
&= \sum_{\boldsymbol{z}} P_{\boldsymbol{Z}}(\boldsymbol{z}) \min\left\{1, (M-1)\mathbb{P}\Big[G(\bar{\boldsymbol{Z}}) \leq G(\boldsymbol{z})\Big]\right\}
\end{aligned}$$

where $\bar{\boldsymbol{Z}} \sim \mathrm{Unif}(\mathcal{A})$. Therefore

$$\bar{p}_e \leq \sum_{\boldsymbol{z}} P_{\boldsymbol{Z}}(\boldsymbol{z}) \min\left\{1, (M-1)\frac{G(\boldsymbol{z})}{|\mathcal{A}|}\right\}$$

Error Exponent

$$
\begin{aligned}
\bar{p}_e^\star &\le \mathbb{E}\left[\min\left\{1, (M-1)\frac{G^\star(\boldsymbol{Z})}{|\boldsymbol{\mathcal{A}}|}\right\}\right] \\
&\le \left(\frac{M}{|\boldsymbol{\mathcal{A}}|}\right)^\rho \mathbb{E}\left[G^\star(\boldsymbol{Z})^\rho\right] && \min\{1, a\} \le a^\rho, \rho \in [0,1] \\
&\le \exp\left(\rho\left[\log M - \log|\boldsymbol{\mathcal{A}}| + H_{\frac{1}{1+\rho}}(\boldsymbol{Z})\right]\right) && \mathbb{E}[G^\star(\boldsymbol{Z})^\rho] \le \exp\left(\rho H_{\frac{1}{1+\rho}}(\boldsymbol{Z})\right)
\end{aligned}
$$

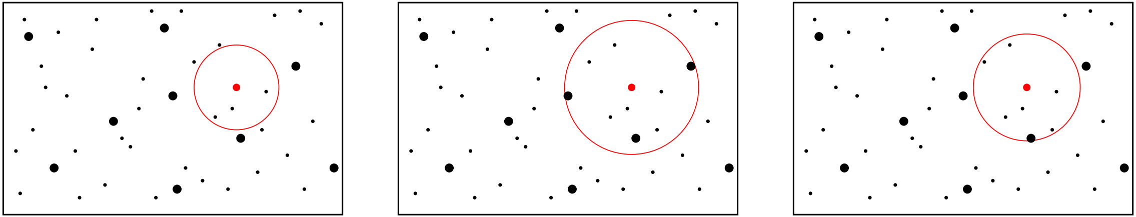
i.i.d. noise

$$
\frac{1}{n}\log\frac{1}{\bar{p}_e^\star} \ge \max_{\rho\in[0,1]} \rho\left(1 - H_{\frac{1}{1+\rho}}(Z) - R\right)
$$

- Recovers the random coding error exponent $E_{\mathrm{r}}(R)$ [Gallager, Trans.IT'65]
- $E_{\mathrm{r}}(R) > 0$ for $R < C = 1 - H(Z)$

Abandonment

- Abandon guessing after at most $m \leq |\mathcal{A}|$ queries to limit complexity
- Known as GRAND-AB [Duffy *et al.*, Trans.IT'19]
- How small can we make m?



Abandonment: Random Coding Bound

Theorem [H.J, ISIT'24]

$$\bar{p}_e(m) \leq \mathbb{E}\left[\mathbb{1}\left[G(\boldsymbol{Z}) \leq m\right] \times \min\left\{1, (M-1)\frac{G(\boldsymbol{Z})}{|\boldsymbol{\mathcal{A}}|}\right\}\right] + \mathbb{P}\left[G(\boldsymbol{Z}) > m\right]$$

Proof.

$$\bar{p}_e(m) = \mathbb{P}\left[\mathcal{E} \cup \mathcal{A}\right] = \mathbb{P}\left[\mathcal{E} \cap \mathcal{A}^c\right] + \mathbb{P}\left[\mathcal{A}\right]$$

Corollary. Using optimal guessing $G^\star$, we get $\bar{p}_e \to 0$ if

$$R < 1 - H(Z) \quad \text{and} \quad r > H(Z)$$

(observed in [Duffy *et al.*, Trans.IT'19])

Error Exponent

Corollary. Using optimal guessing $G^\star$, we get

$$\bar{p}_e^\star(R, r) \le \exp\left(-nE_{\mathrm{r}}(R)\right) + \exp\left(-nE_{\mathrm{s}}(r)\right)$$

GRAND-AB exponent:

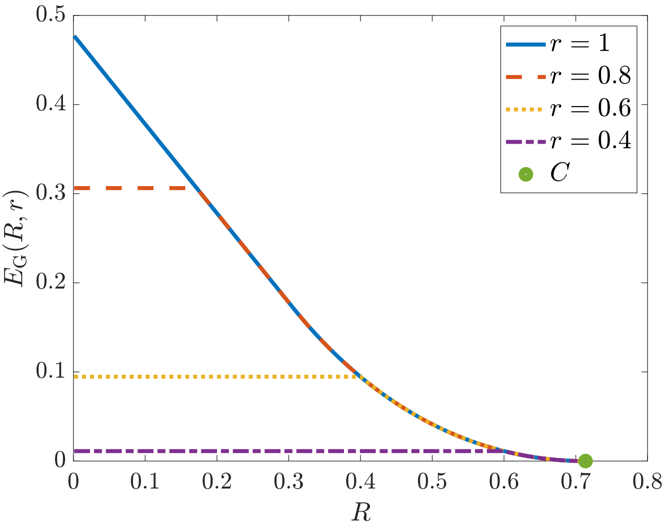
$$E_{\mathrm{G}}(R, r) = \min\left\{E_{\mathrm{r}}(R), E_{\mathrm{s}}(r)\right\}$$

High code rates: Note that $E_{\mathrm{s}}(r) \ge E_{\mathrm{r}}(1 - r)$, therefore

$$E_{\mathrm{G}}(R, r) \ge E_{\mathrm{r}}(\max\{1 - r, R\})$$

equality when $\max\{1 - r, R\} \ge R_{\mathrm{cr}}$

Sufficient to abandon after $m = 2^{n(1-R)}$ guesses to maintain reliability



$E_{\mathrm{G}}(R, r)$ of GRAND-AB for BSC with $p = 0.05$ ($C = 0.7136$)

Some extensions

- Slepian-Wolf guessing-based decoding [H.J, ISIT'24]
- General DMCs, ensemble-tightness, second-order rates [Tan-H.J, Trans.IT'25]
- Universal guessing-based decoders [Miyamoto-Yang, ITW'25]

Thank you!

Supported by the European Research Council