Noisy Guesses

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical Engineering

Technion – Israel Institute of Technology

Haifa 3200003, Israel

ISIT 2020, June 21–26, 2020.

The Guessing Problem

Alice generates a finite–alphabet random vector,

$$\boldsymbol{X} = (X_1, \ldots, X_n) \sim P.$$

Bob submits a sequence of guesses (yes/no questions):

Is $\boldsymbol{X} = \boldsymbol{x}_1$?

Is $\boldsymbol{X} = \boldsymbol{x}_2$?

. . .

until the first hit.

Given a guessing list, $\mathcal{G} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots\}$, let $G(\boldsymbol{X}) = \min\{i : \ \boldsymbol{x}_i = \boldsymbol{X}\}$.

Ordering the guesses according to: $P(\boldsymbol{x}_1) \geq P(\boldsymbol{x}_2) \geq \cdots$

minimizes $\mathbf{E}\{f[G(\boldsymbol{X})]\}$ for every non–decreasing f.

Basic question no. 1: single–letter formula of $\min_{\mathcal{G}} \mathbf{E}\{[G(\boldsymbol{X})]^\rho\}$.

Basic question no. 2: what if P is unknown?

Motivations

♣ Relation to source coding (large deviations).

♣ Natural operational significance for the Rényi entropy.

♣ Sequential decoding (Arikan ‘96).

♣ List decoding.

♣ Security – guessing passwords.

♣ Guessing with distortion (Arikan & M, ‘98) - rate–distortion coding.

Related Work (Partial List Only)

- ♢ Massey ('94) – introduced the notion of guessing.
- ♢ Arikan ('96) – bounds on guessing moments (Rényi's entropy).
- ♢ Arikan & Merhav ('98) – guessing with a fidelity criterion; univerality.
- ♢ Malone & Sullivan ('04) – Markov sources.
- ♢ Pfitser & Sullivan ('04) – stationary sources.
- ♢ Hanawal & Sundaresan ('11) – large deviations.
- ♢ Sundaresan ('07) – guessing under source uncertainty.
- ♢ Christiansen *et al.* ('13) – guessing passwords over a channel.
- ♢ Christiansen *et al.* ('15) – a multiuser scanrio.
- ♢ Beirami *et al.* ('15) – inscrutability.
- ♢ Salamatian *et al.* ('17, '19) – multi-agent guessing.
- ♢ Merhav & Cohen ('20): universal randomized guessing.
- ♢ Merhav ('20): universal guessing individual sequences using FSM's.

Noisy Guessing

In our setting, Alice receives Bob's guesses via a noisy channel.

Formuation:

- ♣ Alice randomly draws $\boldsymbol{Y} = (Y_1, \ldots, Y_n) \sim P$ (DMS).
- ♣ Bob submits a sequence of guesses, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$.
- ♣ Each guess, $\boldsymbol{x}_i$ undergoes a DMC W, to become a noisy guess, $\boldsymbol{Y}_i$.
- ♣ A guess is successful as soon as $\boldsymbol{Y}_i = \boldsymbol{Y}$.

The number of guesses is

$$G = \min\{i : \ \boldsymbol{Y}_i = \boldsymbol{Y}\}.$$

Goal: characterize $\mathbf{E}\{G^\rho\}$ for the best strategy.

Results: 2 optimal randomized strategies, one is universal in (P, W, ρ).

Motivations

♠ Remote connection might be noisy (no coding).

♠ Alice may wish to apply a jammer for defense against attacks by Bob.

♠ Exploring properties of robustness to errors.

♠ Some of the results may be surprising...

♠ Introducing new tools: not relying on source coding for the converse.

Main Result

Define $\Gamma(Q_Y) = \inf_{Q_{X|Y}} D(Q_{Y|X} \| W | Q_X)$

and
$$E(\rho) = \sup_{Q_Y} \left\{ \rho[H(Q_Y) + \Gamma(Q_Y)] - D(Q_Y \| P) \right\}$$
$$= \ln \left(\inf_V \sum_{y \in \mathcal{Y}} \frac{P(y)}{\left[\sum_{x \in \mathcal{X}} V(x) W(y|x)\right]^\rho} \right)$$

Theorem: $\forall$ guessing strategy: $\liminf_{n \to \infty} \frac{\ln \mathbf{E}\{G^\rho\}}{n} \geq E(\rho)$.

$\exists$ guessing strategy: $\limsup_{n \to \infty} \frac{\ln \mathbf{E}\{G^\rho\}}{n} \leq E(\rho)$.

The Penalty due to the Noise

$$E(\rho) = \sup_{Q_Y} \left\{ \rho[H(Q_Y) + \Gamma(Q_Y)] - D(Q_Y \| P) \right\}$$

$$= \ln \left(\inf_{V} \sum_{y \in \mathcal{Y}} \frac{P(y)}{\left[\sum_{x \in \mathcal{X}} V(x) W(y|x)\right]^{\rho}} \right)$$

$\Gamma(Q_Y)$, in the first formula, designates the penalty due to noise.

Looking at the second formula, note that in the absence of noise,

$$E(\rho) = \ln \left(\inf_{Q} \sum_{y \in \mathcal{Y}} \frac{P(y)}{Q^{\rho}(y)} \right).$$

Here, the minimization is limited to $\mathcal{CH}\{W(\cdot|x),\ x \in \mathcal{X}\}$.

Conclusion: If $Q^* \in \mathcal{CH}\{W(\cdot|x),\ x \in \mathcal{X}\}$, there is no penalty!

Example

$P =$ binary source, $\{p = 0.25, 1 - p = 0.75\}$.

$W =$ BSC with crossover parameter, $q < \frac{1}{2}$.

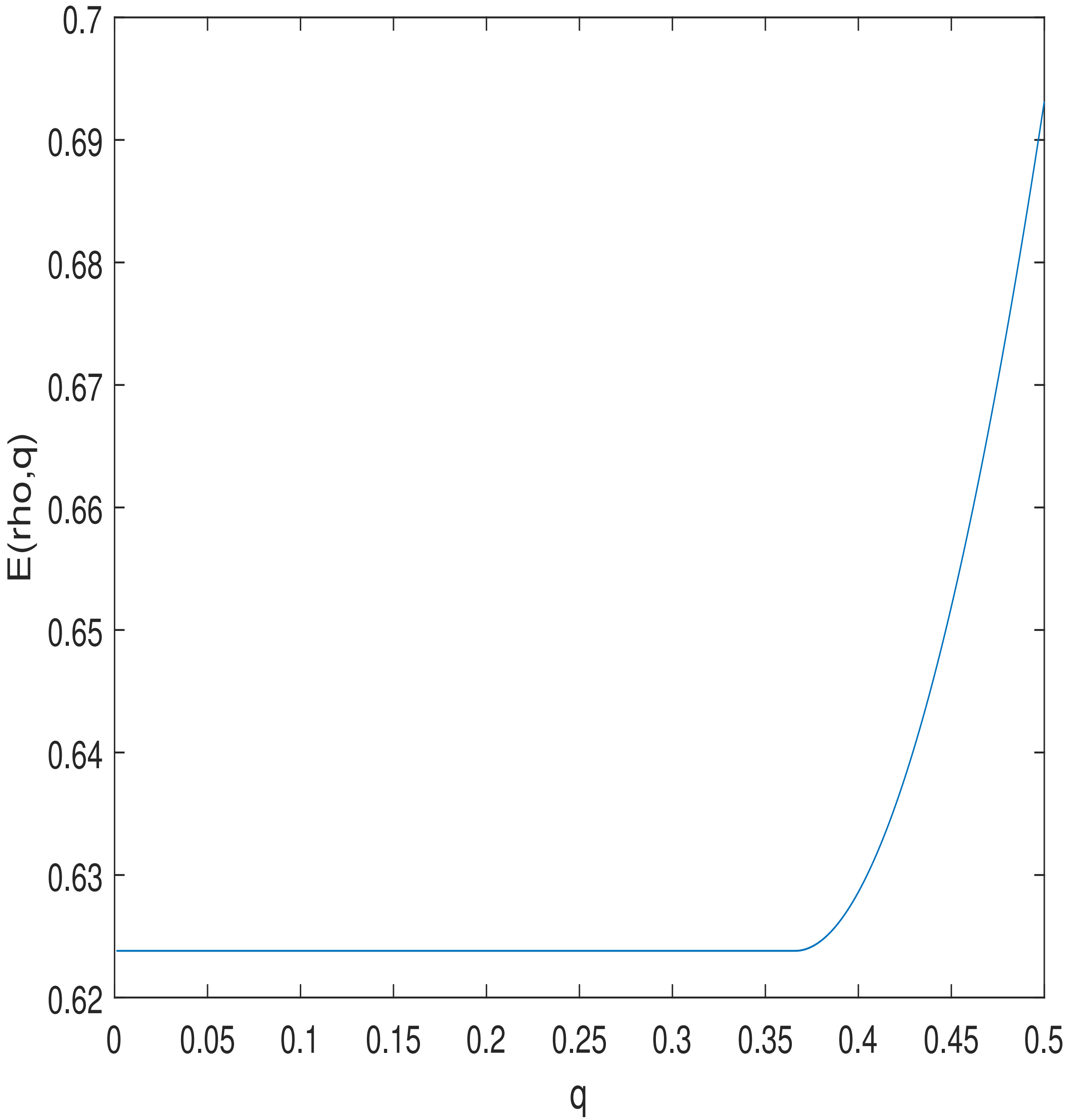
We present graphs of the guessing exponent:

♡ as a function of q for $\rho = 1$. Φ–transition at:

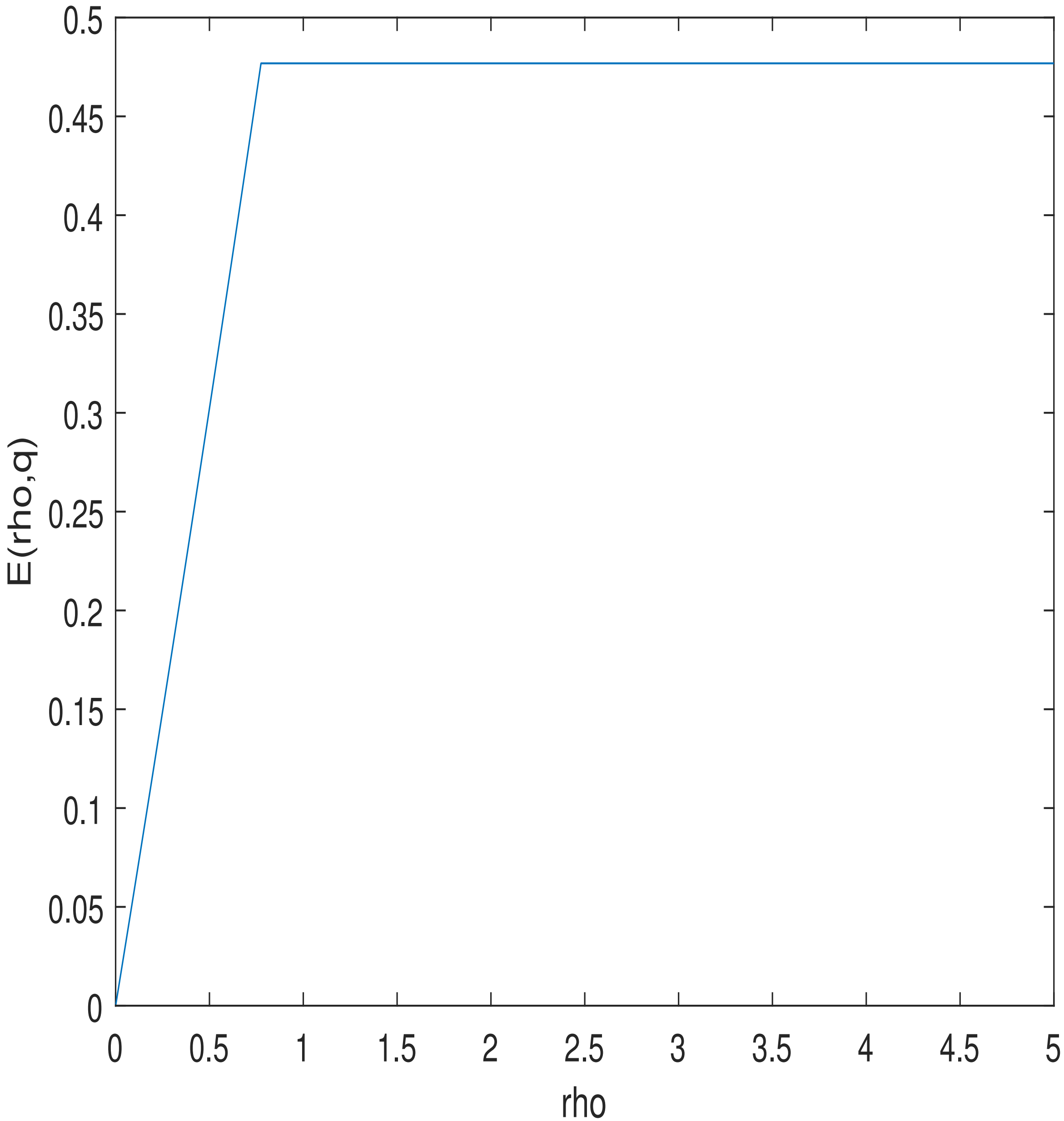
$$q = q_{\text{c}} = \frac{\sqrt{p}}{\sqrt{p} + \sqrt{1-p)}}.$$

♡ as a function of ρ for $q = 0.35$. Φ–transition at:

$$\rho = \rho_{\text{c}} = \left[\frac{\ln[(1-p)/p]}{\ln[(1-q)/q]} - 1\right]_+ .$$



0.7
0.69
0.68
0.67
0.66
0.65
0.64
0.63
0.62
E(rho,q)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
q


E(rho,q)
rho
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

Achievability

The formula

$$E(\rho) = \ln\left(\inf_V \sum_{y\in\mathcal{Y}} \frac{P(y)}{\left[\sum_{x\in\mathcal{X}} V(x)W(y|x)\right]^\rho}\right)$$

suggests a conceptually simple achievability scheme:

Draw the guesses independently at random according to

$$V^*(\boldsymbol{x}) = \prod_{i=1}^{n} V^*(x_i),$$

where V^* attains $E(\rho)$.

Disadvantage: the optimal V^* depends on P, W, and ρ.

$\exists$ universal scheme, independent of (P, W, ρ), that attains $E(\rho)$?

Achievability (Cont'd)

Consider the following random guessing distribution,

$$V(\boldsymbol{x}) = \frac{\exp\{-n\hat{H}_{\boldsymbol{x}}(X)\}}{\sum_{\boldsymbol{x}' \in \mathcal{X}^n} \exp\{-n\hat{H}_{\boldsymbol{x}'}(X)\}},$$

where $\hat{H}_{\boldsymbol{x}}(X)$ is the empirical entropy associated with $\boldsymbol{x}$.

Draw independent guesses under V, which is independent of (P, W, ρ).

♠ It is easy to show (using the method of types) that $E(\rho)$ is achieved.

♠ $V(\boldsymbol{x})$ can be implemented sequentially [Merhav & Cohen ('20)].

♠ Easy to extend to sources with memory and to availability of side info.

A Word About the Converse (Time Permits)

Different from the noiseless case – no source coding considerations.

1. Begin by conditioning on $\boldsymbol{Y} \in \mathcal{T}(Q_Y)$.
2. Use Chebychev's inequality,

$$\mathbf{E}\{G^\rho | \boldsymbol{Y} \in \mathcal{T}(Q_Y)\} \geq k^\rho \text{Pr}\{G > k | \boldsymbol{Y} \in \mathcal{T}(Q_Y)\}.$$

3. Use the relations:

$$\text{Pr}\{G > k | \boldsymbol{Y} = \boldsymbol{y}\} = \prod_{i=1}^{k} [1 - W(\boldsymbol{y}|\boldsymbol{x}_i)] = \exp\left\{\sum_{i=1}^{k} \ln[1 - W(\boldsymbol{y}|\boldsymbol{x}_i)]\right\}.$$

4. Apply the inequality, $\ln(1-w) \geq \frac{w}{1-w}$.
5. Apply Jensen's inequality to pass the expectation to the exponent.
6. Choose k properly.
7. Average over all types.

Thank You!