Two-Stage Guessing

Robert Graczyk and Amos Lapidoth
Signal and Information Processing Laboratory
ETH Zurich, 8092 Zurich, Switzerland
Email: {graczyk, lapidoth}@isi.ee.ethz.ch

***Abstract*—Correlated memoryless sources produce a principal and an ancillary sequence. The exponential growth of the least expected total number of guesses required to guess the principal sequence is determined when, prior to guessing it, the guesser is allowed to produce guesses (not necessarily terminating with a correct one) of the ancillary.**

I. Introduction

We study a variation on guessing with side-information: How to guess a principal sequence X^n, when we are allowed to do so in two phases. In the first we produce guesses of an ancillary sequence Y^n until it is guessed correctly or until we choose to move on to the second phase, in which we must guess X^n. The total number of guesses, namely, the sum of the guesses in the two phases, is denoted $G(X^n;Y^n)$. We study its behavior when $\{(X_i,Y_i)\}_{i=1}^n$ are IID according to some finite-support PMF P_{XY}. We prove the following variational characterization of the least achievable exponential growth of $\mathbb{E}[G(X^n;Y^n)]$:

Theorem 1. *If $\{(X_i,Y_i)\}_{i=1}^n$ are IID according to the finite-support PMF P_{XY}, then*

$$\begin{aligned}&\lim_{n\to\infty}\min_{G}\frac{\log\mathbb{E}[G(X^n;Y^n)]}{n}\\&=\sup_{Q_{XY}}\Big(\min\big(\mathrm{H}(Q_X),\max\big(\mathrm{H}(Q_Y),\mathrm{H}(Q_{X|Y})\big)\big)\\&\qquad-\mathrm{D}(Q_{XY}||P_{XY})\Big),\end{aligned}\tag{1}$$

where the minimum on the LHS is over all two-phase guessing strategies.

Remark 1. *Theorem 1 can be generalized to the ρ-th moment of $G(X^n;Y^n)$: For any $\rho\geq 0$,*

$$\begin{aligned}&\lim_{n\to\infty}\min_{G}\frac{\log\mathbb{E}[G(X^n;Y^n)^\rho]}{n}\\&=\sup_{Q_{XY}}\Big(\rho\min\big(\mathrm{H}(Q_X),\max\big(\mathrm{H}(Q_Y),\mathrm{H}(Q_{X|Y})\big)\big)\\&\qquad-\mathrm{D}(Q_{XY}||P_{XY})\Big).\end{aligned}\tag{2}$$

Example 1. *If $X_i=(Z_i,Y_i)$ with $\{(Z_i,Y_i)\}$ IID P_{ZY}, then the RHS of* (2) *equals*

$$\rho\max\big(\mathrm{H}_{1/(1+\rho)}(P_Y),\mathrm{H}_{1/(1+\rho)}(P_{Z|Y})\big),\tag{3}$$

where $\mathrm{H}_{1/(1+\rho)}(P_Y)$ (and $\mathrm{H}_{1/(1+\rho)}(P_{Z|Y})$) denote the (conditional) Rényi Entropy of order $1/(1+\rho)$

$$\mathrm{H}_{1/(1+\rho)}(P_Y)\triangleq\frac{1}{\rho}\log\left(\sum_{y\in\mathcal{Y}}P_Y(y)^{1/(1+\rho)}\right)^{1+\rho}\tag{4}$$

$$\mathrm{H}_{1/(1+\rho)}(P_{Z|Y})\triangleq\frac{1}{\rho}\log\sum_{y\in\mathcal{Y}}\left(\sum_{z\in\mathcal{Z}}P_{ZY}(z,y)^{1/(1+\rho)}\right)^{1+\rho},\tag{5}$$

with $\mathcal{Y}$ and $\mathcal{Z}$ denoting the support set of the marginals P_Y and P_Z, respectively. In fact, the exponent of (3) *is achievable by guessing Y^n until correct and then guessing X^n (see* (6) *to* (8) *ahead). For a proof of Example 1 see Appendix A.*

When Y^n is deterministic, the problem reduces to the classical Massey-Arikan guessing problem of studying the least ρ-th moment of the number of guesses required to learn the realization of a chance variable X of a PMF P_X having finite support $\mathcal{X}$. A guess is a question of the form "Is $X=x$?", and all guesses are answered truthfully. The optimal guessing order is in descending order of probabilities, and Arikan [1] showed that

$$\mathbb{E}[G^*(X)^\rho]\approx 2^{\rho\,\mathrm{H}_{1/(1+\rho)}(P_X)},\tag{6}$$

where G^* denotes an optimal guessing order, i.e., a bijection $\mathcal{X}\to\{1,2,\ldots,|\mathcal{X}|\}$ minimizing the LHS of (6). Equality in (6) holds up to a factor dominated by $\log|\mathcal{X}|$. Consequently, when guessing a sequence $X^n\sim P_X^n$, i.e., whose components $X_1,X_2,\ldots,X_n$ are IID $\sim P_X$,

$$\lim_{n\to\infty}\frac{\log\mathbb{E}[G^*(X^n)^\rho]}{n}=\rho\,\mathrm{H}_{1/(1+\rho)}(P_X).\tag{7}$$

Various extensions to the above results are known. In [1], Arikan also examined the setup where—prior to guessing X^n—side-information is revealed to the guesser in the form of a sequence Y^n that is jointly IID with X^n. He showed that

$$\lim_{n\to\infty}\frac{\log\mathbb{E}[G^*(X^n\mid Y^n)^\rho]}{n}=\rho\,\mathrm{H}_{1/(1+\rho)}(P_{X|Y}),\tag{8}$$

where $G^*(\,\cdot\mid y^n)$ is a guessing order that is optimal for the PMF $P_{X^n|Y^n=y^n}$. Generalizations to the case where only a rate-limited description of Y^n is available to the guesser were studied by Graczyk and Lapidoth [2], who sought a mapping

$$\phi_n \colon \mathcal{Y}^n \to \{0,1\}^{nR} \tag{9}$$

minimizing

$$\lim_{n\to\infty} \frac{\log \mathbb{E}[G^*(X^n \mid \phi_n(Y^n))^\rho]}{n}. \tag{10}$$

Upper and lower bounds on the least value that (10) can take are given in [2, (36) and (37)]. When $Y^n = X^n$, the optimal value is [2, Theorem 1]

$$\sup_{Q_X} \inf_{Q_{U|X}:\mathrm{I}(Q_X;U)\le R} \left(\rho\,\mathrm{H}(Q_{X|U}) - \mathrm{D}(Q_X||P_X)\right). \tag{11}$$

Weinberger and Shayevitz [3] examined one-bit descriptions

$$\phi_n \colon \mathcal{Y}^n \to \{0,1\}, \quad \forall n \in \mathbb{N}, \tag{12}$$

and derived bounds on the optimal guessing efficiency

$$\limsup_{n\to\infty} \min_{\phi_n} \frac{\mathbb{E}[G^*(X^n \mid \phi_n(Y^n))^\rho]}{\mathbb{E}[G^*(X^n)\rho]}. \tag{13}$$

Burin and Shayevitz [4] further considered a setup where the guesser is only revealed a noisy observation of $\phi_n(Y^n)$.

II. Problem Statement

Let P_{XY} be a PMF on the finite set $\mathcal{X}\times\mathcal{Y}$ and $\{(X_i,Y_i)\}_{i=1}^n$ IID according to P_{XY}. Our goal is to guess X^n in two phases. In the first we produce guesses of Y^n: We choose some subset $\emptyset \subseteq \mathcal{G} \subseteq \mathcal{Y}^n$ and define a linear order on its elements, i.e., a bijection

$$\mathrm{ord} \colon \mathcal{G} \to \{1,2,\ldots,|\mathcal{G}|\}. \tag{14}$$

We then take consecutive guesses of the form

$$\text{“Is } Y^n = y^n\text{?”}, \quad y^n \in \mathcal{G}, \tag{15}$$

until correct or until $\mathcal{G}$ is exhausted, and where y^n is guessed before $\tilde{y}^n$ whenever

$$\mathrm{ord}(y^n) < \mathrm{ord}(\tilde{y}^n). \tag{16}$$

If $\mathcal{G}$ is empty, Phase 1 is skipped and no guesses of Y^n are taken. If $\mathcal{G}$ is not empty and Y^n is in $\mathcal{G}$, then Phase 1 will terminate after $\mathrm{ord}(Y^n)$ guesses with Y^n revealed; otherwise, i.e., if Y^n is not in $\mathcal{G}$, Phase 1 will terminate after $|\mathcal{G}|$ guesses without revealing Y^n but only revealing that Y^n is not in $\mathcal{G}$. Given $\mathcal{G}$ and $\mathrm{ord}(\cdot)$, the number of guesses taken in Phase 1 is $G(Y^n)$, with $G(\cdot)$ being the mapping

$$G \colon Y^n \mapsto \begin{cases} \mathrm{ord}(Y^n) & \text{if } Y^n \in \mathcal{G} \\ |\mathcal{G}| & \text{else.} \end{cases} \tag{17}$$

We emphasize that Phase 1 need not reveal Y^n.

In the second phase we must guess X^n. To that end we choose a guessing order on $\mathcal{X}^n$ for every possible outcome of Phase 1. The chosen guessing order determines the number of guesses in Phase 2, which we denote $\tilde{G}(X^n \mid Y^n)$. The tilde in $\tilde{G}(X^n \mid Y^n)$ reminds us that, rather than Y^n, Phase 1 might only reveal that Y^n is not in $\mathcal{G}$. The functions $G(\cdot)$ and $\tilde{G}(\cdot \mid \cdot)$ (and their implicit domains) specify a two-phase guessing strategy that we denote

$$\pi = (G, \tilde{G}). \tag{18}$$

The total number of guesses required by $\pi = (G,\tilde{G})$ is denoted $G(X^n;Y^n)$, so

$$G(X^n;Y^n) \triangleq G(Y^n) + \tilde{G}(X^n \mid Y^n). \tag{19}$$

We seek the least achievable exponential growth of $\mathbb{E}[G(X^n;Y^n)]$, i.e,

$$\lim_{n\to\infty} \min_{\pi} \frac{\log \mathbb{E}\left[G(X^n;Y^n)\right]}{n}. \tag{20}$$

(We shall see that this limit exists.)

III. Analysis

In this section we prove Theorem 1, namely, that if $\{(X_i,Y_i)\}_{i=1}^n$ are IID $\sim P_{XY}$, then the limit in (20) exists and equals

$$\sup_{Q_{XY}} \Big(\min\left(\mathrm{H}(Q_X), \max\left(\mathrm{H}(Q_Y), \mathrm{H}(Q_{X|Y})\right)\right) - \mathrm{D}(Q_{XY}||P_{XY})\Big). \tag{21}$$

Proof. As in [5, Proposition 6], we first note that the cost incurred by the guesser for not knowing the (joint) empirical type of (X^n,Y^n) is at most polynomial in n. Consequently, the expectation of the total number of guesses $G^*(X^n;Y^n)$ induced by an optimal two-phase guessing strategy can be upper-bounded by

$$\mathbb{E}\left[G^*(X^n;Y^n)\right] \le \mathbb{E}\left[G^*_{\mathrm{T}}(X^n;Y^n)\right] \mathrm{poly}(n), \tag{22}$$

where $G^*_{\mathrm{T}}(\cdot\,;\cdot)$ is the total number of guesses required to guess X^n using an optimal two-phase strategy with access to the empirical type of (X^n,Y^n); and where $\mathrm{poly}(n)$ denotes a monomial in n. Since the guesser is free to ignore the type,

$$\mathbb{E}\left[G^*_{\mathrm{T}}(X^n;Y^n)\right] \le \mathbb{E}\left[G^*(X^n;Y^n)\right], \tag{23}$$

and hence

$$\lim_{n\to\infty} \frac{\log \mathbb{E}\left[G^*(X^n;Y^n)\right]}{n} = \lim_{n\to\infty} \frac{\log \mathbb{E}\left[G^*_{\mathrm{T}}(X^n;Y^n)\right]}{n}, \tag{24}$$

whenever the limit on the RHS exists.

To prove Theorem 1, we evaluate the RHS of (24). We do so by first studying our problem when, rather than IID $\sim P_{XY}$, the pair (X^n,Y^n) is drawn uniformly over a type class $\mathcal{T}^{(n)}(Q_{XY})$, where Q_{XY} is known and belongs to the family $\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})$ of joint types on $\mathcal{X}\times\mathcal{Y}$ of denominator n. This situation arises when $(X^n,Y^n) \sim P_{XY}^n$ and we condition on $(X^n,Y^n) \in \mathcal{T}^{(n)}(Q_{XY})$. We will show that in this case

$$\lim_{n\to\infty} \frac{\log \mathbb{E}\left[G^*_{\mathrm{T}}(X^n;Y^n) \mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})\right]}{n} = \min\big(\mathrm{H}(Q_X), \max\big(\mathrm{H}(Q_Y),\mathrm{H}(Q_{X|Y})\big)\big), \tag{25}$$

where the convergence is uniform w.r.t. Q_{XY}. Once (25) is established, the RHS of (24) can be evaluated by averaging over the empirical type of (X^n, Y^n) as follows:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{\log \mathbb{E}[G^*(X^n;Y^n)]}{n} \\
&\overset{(a)}{=} \lim_{n\to\infty} \frac{\log \mathbb{E}[G^*_{\mathrm{T}}(X^n;Y^n)]}{n} && (26)\\
&\overset{(b)}{=} \lim_{n\to\infty} \max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})} \Bigg(-\mathrm{D}(Q_{XY}||P_{XY}) \\
&\qquad + \frac{\log \mathbb{E}[G^*_{\mathrm{T}}(X^n;Y^n) \mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY}))]}{n}\Bigg) && (27)\\
&\overset{(c)}{=} \lim_{n\to\infty} \max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})} \Big(-\mathrm{D}(Q_{XY}||P_{XY}) \\
&\qquad + \min\big(\mathrm{H}(Q_X), \max\big(\mathrm{H}(Q_Y),\mathrm{H}(Q_{X|Y})\big)\big)\Big) && (28)\\
&\overset{(d)}{=} \sup_{Q_{XY}} \Big(\min\big(\mathrm{H}(Q_X), \max\big(\mathrm{H}(Q_Y),\mathrm{H}(Q_{X|Y})\big)\big) \\
&\qquad - \mathrm{D}(Q_{XY}||P_{XY})\Big), && (29)
\end{aligned}$$

where (a) is due to (24); (b) is justified in Appendix B; (c) follows from (25); and (d) holds because every PMF can be approximated by a type of sufficiently large denominator, and because the functions in (29) are all continuous w.r.t. Q_{XY}.

In the remainder of this section we prove (25). Achievability is straightforward: Setting $\mathcal{G}$ to be empty yields the exponent $\mathrm{H}(Q_X)$ (cf. [5, Example 1]), whereas setting $\mathcal{G}$ to be $\mathcal{T}^{(n)}(Q_Y)$ yields the exponent $\max(\mathrm{H}(Q_Y), \mathrm{H}(Q_{X|Y}))$, because with this choice of $\mathcal{G}$ Phase 1 and Phase 2 require roughly $2^{n\,\mathrm{H}(Q_Y)}$ and $2^{n\,\mathrm{H}(Q_{X|Y})}$ guesses, respectively. This strategy also results in a uniform convergence w.r.t. Q_{XY}.

Having proved that the RHS of (25) is achievable, we now show that no two-phase guessing strategy $\pi_{\mathrm{T}} = (G_{\mathrm{T}}, \tilde{G}_{\mathrm{T}})$ (cognizant of Q_{XY}) can do better. Let the exponential growth of the expected number of guesses in Phase 1 and Phase 2 under π_{T} be denoted $E_Y(\pi_{\mathrm{T}}; Q_{XY})$ and $E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY})$,

$$\begin{aligned}
&E_Y(\pi_{\mathrm{T}}; Q_{XY}) \\
&\triangleq \liminf_{n\to\infty} \frac{\log \mathbb{E}[G_{\mathrm{T}}(Y^n) \mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]}{n} && (30)\\
&E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) \\
&\triangleq \liminf_{n\to\infty} \frac{\log \mathbb{E}[\tilde{G}_{\mathrm{T}}(X^n \mid Y^n) \mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]}{n}, && (31)
\end{aligned}$$

and let $E_{XY}(\pi_{\mathrm{T}}; Q_{XY})$ denote the exponential growth of the expected total number number of guesses,

$$\begin{aligned}
&E_{XY}(\pi_{\mathrm{T}}; Q_{XY}) \\
&\triangleq \liminf_{n\to\infty} \frac{\log \mathbb{E}\left[G_{\mathrm{T}}(X^n;Y^n) \mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})\right]}{n}. && (32)
\end{aligned}$$

The exponential growth of a sum is dominated by that of the largest of the addends, so (accounting for the limit inferior)

$$E_{XY}(\pi_{\mathrm{T}}; Q_{XY}) \geq \max\big(E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}), E_Y(\pi_{\mathrm{T}}; Q_{XY})\big). \tag{33}$$

Also,

$$E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) \geq \mathrm{H}(Q_{X|Y}), \tag{34}$$

because $\mathrm{H}(Q_{X|Y})$ is the exponent in Phase 2 when Y^n is revealed in Phase 1. To conclude the proof, we will show that the RHS of (33) is lower-bounded by the RHS of (25). This is clearly the case when $E_Y(\pi_{\mathrm{T}}; Q_{XY})$ equals $\mathrm{H}(Q_Y)$ because of (34). We therefore focus on the case where $E_Y(\pi_{\mathrm{T}}; Q_{XY}) < \mathrm{H}(Q_Y)$. We will show that

$$E_Y(\pi_{\mathrm{T}}; Q_{XY}) < \mathrm{H}(Q_Y) \implies E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) = \mathrm{H}(Q_X), \tag{35}$$

and thus conclude the proof, because in this case the RHS of (33) is at least $\mathrm{H}(Q_X)$.

We establish (35) via its contrapositive,

$$E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) < \mathrm{H}(Q_X) \implies E_Y(\pi_T; Q_{XY}) = \mathrm{H}(Q_Y), \tag{36}$$

by proving that

$$E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) < \mathrm{H}(Q_X) \implies |\mathcal{G}| = \Theta(|\mathcal{T}^{(n)}(Q_Y)|), \tag{37}$$

where $|\mathcal{G}| = \Theta(|\mathcal{T}^{(n)}(Q_Y)|)$ indicates that for some $\alpha > 0$ and all sufficiently large n,

$$|\mathcal{G}| \geq \alpha |\mathcal{T}^{(n)}(Q_Y)|. \tag{38}$$

To show (37) we also argue by contraposition and therefore assume that $|\mathcal{G}| = o(|\mathcal{T}^{(n)}(Q_Y)|)$, i.e.,

$$\limsup_{n\to\infty} \frac{|\mathcal{G}|}{|\mathcal{T}^{(n)}(Q_Y)|} = 0. \tag{39}$$

We define the indicator variable

$$E \triangleq \begin{cases} 0 & \text{if } Y^n \in \mathcal{G} \\ 1 & \text{else,} \end{cases} \tag{40}$$

and observe that (39) implies

$$\lim_{n\to\infty} \mathbb{P}[E=1] = 1, \tag{41}$$

and consequently,

$$\lim_{n\to\infty} \mathrm{H}(E) = 0. \tag{42}$$

Since

$$\begin{aligned} \mathrm{H}(X^n) - \mathrm{H}(X^n \mid E) &= \mathrm{I}(X^n; E) && (43)\\ &\leq \mathrm{H}(E), && (44)\end{aligned}$$

we infer from (42) that

$$\lim_{n\to\infty} \big(\mathrm{H}(X^n) - \mathrm{H}(X^n \mid E)\big) = 0. \tag{45}$$

This and (41) implies that

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{H}(X^n) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}(X^n \mid E=1) \tag{46}$$

as can be seen by expanding $\mathrm{H}(X^n \mid E)$ and noting that

$$\frac{1}{n}\mathrm{H}(X^n \mid E=0) \leq \log|\mathcal{X}|. \tag{47}$$

To conclude the proof that (39) implies the negation of the LHS of (37) we proceed as follows:

$$\begin{aligned} &E_{X|Y}(\pi_{\mathrm{T}}; Q_{XY}) \\ &= \liminf_{n\to\infty} \frac{\log \mathbb{E}[\tilde{G}_{\mathrm{T}}(X^n \mid Y^n) \mid (X^n, Y^n) \in \mathcal{T}^{(n)}(Q_{XY})]}{n} && (48)\\ &\overset{(a)}{\geq} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\tilde{G}_{\mathrm{T}}(X^n \mid Y^n) \\ &\qquad\qquad \mid (X^n, Y^n) \in \mathcal{T}^{(n)}(Q_{XY}), E=1] && (49)\\ &\overset{(b)}{\geq} \liminf_{n\to\infty} \frac{1}{n} \mathrm{H}_{1/2}(X^n \mid E=1) && (50)\\ &\overset{(c)}{\geq} \liminf_{n\to\infty} \frac{1}{n} \mathrm{H}(X^n \mid E=1) && (51)\\ &\overset{(d)}{=} \lim_{n\to\infty} \frac{1}{n} \mathrm{H}(X^n) && (52)\\ &= \mathrm{H}(Q_X), && (53)\end{aligned}$$

where (a) follows from the law of total expectation and (41); (b) holds by the relationship between guessing and Rényi Entropy, namely (6); (c) follows from the monotonicity of the Rényi Entropy in its parameter; and (d) holds by (46). □

Appendix

A. Derivation of (3)

We prove that when X has the form (Z, Y) with Z and Y of some arbitrary joint PMF P_{ZY},

$$\sup_{Q_{XY}} \Big(\rho \min\big(\mathrm{H}(Q_X), \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{X|Y})\big)\big) - \mathrm{D}(Q_{XY}||P_{XY})\Big) \tag{54}$$

equals

$$\rho \max\big(\mathrm{H}_{1/(1+\rho)}(P_Y), \mathrm{H}_{1/(1+\rho)}(P_{Z|Y})\big). \tag{55}$$

Using the given form of X, we first provide an alternative expression for (54):

$$\begin{aligned} &\sup_{Q_{XY}} \Big(\rho \min\big(\mathrm{H}(Q_X), \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{X|Y})\big)\big) \\ &\qquad - \mathrm{D}(Q_{XY}||P_{XY})\Big) \\ &= \sup_{Q_{ZY}} \Big(\rho \min\big(\mathrm{H}(Q_{ZY}), \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big)\big) \\ &\qquad - \mathrm{D}(Q_{ZY}||P_{ZY})\Big) && (56)\\ &\overset{(a)}{=} \sup_{Q_{ZY}} \Big(\rho \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big) - \mathrm{D}(Q_{ZY}||P_{ZY})\Big), && (57)\end{aligned}$$

where (a) holds because $\mathrm{H}(Q_{ZY})$ is lower-bounded by $\max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big)$. Having established that (54) equals (57), it now suffices to show that (57) equals (55). To that end, we first argue that (57) is upper-bounded by (55):

$$\begin{aligned} &\sup_{Q_{ZY}} \Big(\rho \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big) - \mathrm{D}(Q_{ZY}||P_{ZY})\Big) \\ &\overset{(a)}{=} \sup_{Q_{ZY}} \Big(\max\big(\rho\,\mathrm{H}(Q_Y) - \mathrm{D}(Q_{ZY}||P_{ZY}), \\ &\qquad\qquad \rho\,\mathrm{H}(Q_{Z|Y}) - \mathrm{D}(Q_{ZY}||P_{ZY})\big)\Big) && (58)\\ &\overset{(b)}{\leq} \max\Big(\sup_{Q_{ZY}} \big(\rho\,\mathrm{H}(Q_Y) - \mathrm{D}(Q_{ZY}||P_{ZY})\big), \\ &\qquad\qquad \sup_{Q_{ZY}} \big(\rho\,\mathrm{H}(Q_{Z|Y}) - \mathrm{D}(Q_{ZY}||P_{ZY})\big)\Big) && (59)\\ &\overset{(c)}{=} \rho \max\big(\mathrm{H}_{1/(1+\rho)}(P_Y), \mathrm{H}_{1/(1+\rho)}(P_{Z|Y})\big), && (60)\end{aligned}$$

where (a) holds because $d\max(a,b)$ equals $\max(da, db)$ and $\max(a,b) - c$ equals $\max(a-c, b-c)$ for any real a, b, c and positive d; in (b) we independently optimize the arguments to $\max(\cdot,\cdot)$; and (c) follows from the variational characterization of the Rényi Entropy [5, Proposition 8].

To conclude the proof, we show that (55) lower-bounds (57). We do so by proving that (57) is lower-bounded by both $\rho\,\mathrm{H}_{1/(1+\rho)}(P_Y)$ and $\rho\,\mathrm{H}_{1/(1+\rho)}(P_{Z|Y})$ (the claim then follows because (55) is the maximum of the two):

$$\begin{aligned} &\sup_{Q_{ZY}} \Big(\rho \max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big) - \mathrm{D}(Q_{ZY}||P_{ZY})\Big) && (61)\\ &\overset{(a)}{\geq} \sup_{Q_{ZY}} \Big(\rho\,\mathrm{H}(Q_Y) - \mathrm{D}(Q_{ZY}||P_{ZY})\Big) && (62)\\ &= \rho\,\mathrm{H}_{1/(1+\rho)}(P_Y), && (63)\end{aligned}$$

where (a) holds because $\max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big) \geq \mathrm{H}(Q_Y)$. Likewise, $\max\big(\mathrm{H}(Q_Y), \mathrm{H}(Q_{Z|Y})\big) \geq \mathrm{H}(Q_{Z|Y})$, so (54) is also lower-bounded by $\rho\,\mathrm{H}_{1/(1+\rho)}(P_{Z|Y})$.

B. A Proof of (27)

We prove (27) by showing that for sufficiently large n,

$$\begin{aligned}&\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)]\\ &\approx \max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\Big(2^{-n\,\mathrm{D}(Q_{XY}\|P_{XY})}\\ &\qquad \mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\Big). \end{aligned}\tag{64}$$

Concretely,

$$\begin{aligned}&\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)]\\ &\overset{(a)}{=}\sum_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\\ &\qquad\qquad \mathbb{P}[(X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})] \end{aligned}\tag{65}$$

$$\begin{aligned}&\overset{(b)}{\leq}\sum_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\\ &\qquad\qquad 2^{-n\,\mathrm{D}(Q_{XY}||P_{XY})} \end{aligned}\tag{66}$$

$$\begin{aligned}&\overset{(c)}{\leq}\mathrm{poly}(n)\max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\Big(2^{-n\,\mathrm{D}(Q_{XY}||P_{XY})}\\ &\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\Big), \end{aligned}\tag{67}$$

where (a) follows from the law of total expectation; (b) is due to Sanov's Theorem [6, Theorem 11.4.1]; and (c) holds because there are at most polynomially many types of denominator n. Similarly,

$$\begin{aligned}&\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)]\\ &=\sum_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\\ &\qquad\qquad \mathbb{P}[(X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})] \end{aligned}\tag{68}$$

$$\begin{aligned}&\overset{(a)}{\geq}\max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\\ &\qquad\qquad \mathbb{P}[(X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})] \end{aligned}\tag{69}$$

$$\begin{aligned}&\overset{(b)}{\geq}\frac{1}{\mathrm{poly}(n)}\max_{Q_{XY}\in\mathcal{T}_n(\mathcal{X}\times\mathcal{Y})}\Big(2^{-n\,\mathrm{D}(Q_{XY}||P_{XY})}\\ &\mathbb{E}[G_{\mathrm{T}}^*(X^n;Y^n)\mid (X^n,Y^n)\in\mathcal{T}^{(n)}(Q_{XY})]\Big), \end{aligned}\tag{70}$$

where (a) follows from dropping all terms but one; and (b) is again due to Sanov's Theorem.

REFERENCES

[1] E. Arikan, "An inequality on guessing and its application to sequential decoding," *IEEE Trans. Inf. Theory*, vol. 42, pp. 99 – 105, Jan. 1996.

[2] R. Graczyk and A. Lapidoth, "Variations on the guessing problem," in *Int. Symp. Inf. Theory*, 06 2018, pp. 231–235.

[3] N. Weinberger and O. Shayevitz, "Guessing with a Boolean helper," in *Int. Symp. Inf. Theory*, 06 2018, pp. 271–275.

[4] A. Burin and O. Shayevitz, "Reducing guesswork via an unreliable oracle," *IEEE Trans. Inf. Theory*, vol. 64, pp. 6941 – 6953, Jul. 2018.

[5] A. Lapidoth, "Lecture notes on guessing and rényi entropy," 2017.

[6] T. Cover, *Elements of information theory (2 ed.)*. Wiley-Interscience, 2006.