Guessing Based on Compressed Side Information

Robert Graczyk*, Amos Lapidoth*, Neri Merhav†, and Christoph Pfister*

*Signal and Information Processing Laboratory ETH Zurich, 8092 Zurich, Switzerland

Email: {graczyk, lapidoth, pfister}@isi.ee.ethz.ch

†The Viterbi Faculty of Electrical and Computer Engineering Technion,

Technion City, Haifa 3200003, Israel

Email: merhav@ee.technion.ac.il

Abstract

A source sequence is to be guessed with some fidelity based on a rate-limited description of an observed sequence with which it is correlated. The trade-off between the description rate and the exponential growth rate of the least power mean of the number of guesses is characterized.

Index Terms

Compression, Guessing, Side Information.

I. Introduction

Our problem can be viewed as the guessing analogue of the Remote Sensing problem in lossy source coding [1], [2], [3]. As in that problem, the description of a source sequence is indirect: the rate-limited description is based only on a noisy version of the sequence. The problems differ, however, in their objectives: in the Remote Sensing problem the source sequence is *estimated* (with the least expected distortion), whereas in our problem it is *guessed* to within some distortion (with the least power mean of the number of required guesses). Our problem thus relates to Arıkan and Merhav's guessing-subject-to-distortion problem [4] in much the same way that the Remote Sensing problem relates to Shannon's lossy source coding problem [5].

To put our problem in context, recall that in the guessing problem pioneered by Massey [6] and Arıkan [7], a guesser seeks to recover a finite-valued chance variable $X \in \mathcal{X}$ by sequentially producing guesses of the form

"Is $X = x_1$?"

"Is $X = x_2$?"

$\vdots$

where $x_1, x_2, \ldots \in \mathcal{X}$, and each guess is answered truthfully with "Yes" or "No." The number of guesses taken until the first "Yes," i.e., until X is revealed, depends on the guesser's strategy $\mathcal{G}$ (the order in which the elements of $\mathcal{X}$ are guessed) and is denoted $G(X)$. Given the probability mass function (PMF) P_X of X and some $\rho > 0$, Arıkan showed [7] that the least achievable ρ-th moment of the number of guesses $\mathrm{E}[G(X)^\rho]$ required to recover X is closely related to its Rényi entropy:

$$\frac{1}{(1+\log|\mathcal{X}|)^\rho} 2^{\rho \, \mathrm{H}_{1/(1+\rho)}(P_X)} \leq \min_{\mathcal{G}} \mathrm{E}[G(X)^\rho] \leq 2^{\rho \, \mathrm{H}_{1/(1+\rho)}(X)}, \tag{1}$$

where $\mathrm{H}_{1/(1+\rho)}(P_X)$ denotes the order-$1/(1+\rho)$ Rényi entropy of X. When guessing a length-n random sequence $X^n \triangleq (X_1, \ldots, X_n)$ whose components are independent and identically distributed (IID) according to P_X, Inequality (1) implies that

$$\lim_{n \to \infty} \frac{1}{n} \log \Big(\min_{\mathcal{G}} \mathrm{E}[G(X^n)^\rho] \Big) = \rho \, \mathrm{H}_{1/(1+\rho)}(P_X), \tag{2}$$

so the Rényi entropy of X fully characterizes (up to the factor ρ) the exponential growth rate of the least ρ-th moment of the number of guesses required to recover X^n.

Our problem differs from Massey's and Arıkan's in the following two ways:

1) Instead of recovering X^n, the guesser need only produce a guess $\hat{X}^n \in \hat{\mathcal{X}}^n$ that is close to X^n in the sense that

$$\frac{1}{n} \sum_{i=1}^{n} d(X_i, \hat{X}_i) \leq D, \tag{3}$$

where the distortion measure $d(\cdot,\cdot) \colon \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}_{\geq 0}$ and the maximal-allowed distortion level $D > 0$ are prespecified. We assume that, for every $x^n \in \mathcal{X}^n$, (3) is satisfied by some $\hat{x}^n \in \hat{\mathcal{X}}^n$; this guarantees the existence of a guessing strategy that eventually succeeds.

2) Prior to guessing, the guesser is provided with a rate-limited description $f(Y^n) \in \{0,1\}^{nR}$ of a noisy observation $Y^n \in \mathcal{Y}^n$ of X^n. Based on $f(Y^n)$, the guesser sequentially guesses elements $\hat{X}^n$ of $\hat{\mathcal{X}}^n$ until (3) is satisfied. (The guesser's strategy $\mathcal{G}$ thus depends on $f(Y^n)$.)

We show that when $(X_1, Y_1), \ldots, (X_n, Y_n)$ are IID according to P_{XY}, the exponential growth rate of the least ρ-th moment of the number of guesses—optimized over the description function f and the guessing strategy $\mathcal{G}$—satisfies the variational characterization (13) of Theorem 1 ahead.

Along the lines of [8], this theorem can be used to assess the resilience of a password X^n against an adversary who has access to nR bits of a correlated password Y^n and is content with guessing only a fraction $1 - D$ of the symbols of X^n. (In this application, the distortion function is the Hamming distance.)

Since our guessing problem is an extension of the guessing-subject-to-distortion problem studied by Merhav and Arıkan [4], their suggested motivation (accounting for the computational complexity of a rate-distortion encoder as measured by the number of metric calculations) and proposed applications (betting games, pattern matching, and database search algorithms) also extend to our setup. Further applications include sequential decoding [7], compression [9], and task encoding [10], [11].

Numerous other variations on the Massey-Arıkan guessing problem were studied over the years. In [12], Sundaresan derived an expression for the smallest guessing moment when the source distribution is only partially known to the guesser; in [13], [14], the authors constructed and analyzed optimal decentralized guessing strategies (for multiple guessers that cannot communicate); in [15], Weinberger and Shayevitz quantified the value of a single bit of side-information provided to the guesser prior to guessing; in [16], the authors studied the guessing problem using an information-geometric approach; and in [17] and [11] the authors studied the distributed guessing problem on Gray-Wyner and Stelpian-Wolf networks.

The above distributed settings dealt, however, only with "lossless" guessing, where the guessing has to be exact. Our present setting maintains, to some degree, a distributed flavor, but allows for "lossy" guessing, i.e., with some fidelity.

II. Problem Statement and Notation

Consider n pairs $\{(X_i, Y_i)\}_{i=1}^n$ that are drawn independently, each according to a given PMF P_{XY} on the finite Cartesian product $\mathcal{X} \times \mathcal{Y}$:

$$\{(X_i, Y_i)\}_{i=1}^n \sim \text{IID}\, P_{XY}. \tag{4}$$

Define the sequences

$$X^n \triangleq \{X_i\}_{i=1}^n,\; Y^n \triangleq \{Y_i\}_{i=1}^n, \tag{5}$$

with $\{X_i\}_{i=1}^n$ being IID P_X, where P_X is the X-marginal of P_{XY}, and likewise $\{Y_i\}_{i=1}^n$ being IID P_Y. By possibly redefining $\mathcal{X}$ and $\mathcal{Y}$, we assume without loss of generality that P_X and P_Y are positive. A guesser wishes to produce a sequence $\hat{X}^n$, taking values in a finite n-fold Cartesian product set $\hat{\mathcal{X}}^n$, that is "close" to X^n in the sense that

$$\bar{\mathrm{d}}(X^n, \hat{X}^n) \leq D, \tag{6}$$

where $D > 0$ is some prespecified maximally-allowed distortion, and

$$\bar{\mathrm{d}}(x^n, \hat{x}^n) \triangleq \frac{1}{n} \sum_{i=1}^{n} \mathrm{d}(x_i, \hat{x}_i) \tag{7}$$

with

$$\mathrm{d}\colon \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}_{\geq 0} \tag{8}$$

some prespecified distortion function. We assume that $d(\cdot,\cdot)$ and D are such that for each $x^n \in \mathcal{X}^n$ there exists some $\hat{x}^n \in \hat{\mathcal{X}}^n$ for which (6) is satisfied,

$$\forall x^n \in \mathcal{X}^n\, \exists \hat{x}^n \in \hat{\mathcal{X}}^n \colon \bar{\mathrm{d}}(x^n, \hat{x}^n) \leq D. \tag{9}$$

This guarantees that such $\hat{X}^n$ can be found and in no-more-than $|\hat{\mathcal{X}}|^n$ guesses.

Courtesy of a "helper" $f_n \colon \mathcal{Y}^n \to \{0,1\}^{nR}$, the guesser is provided, prior to guessing, with an nR-bit description $f_n(Y^n)$ of Y^n. Based on $f_n(Y^n)$, the guesser produces a "guessing strategy" (also called a "guessing function")

$$\mathcal{G}_n\bigl(\,\cdot\,|f_n(Y^n)\bigr) \colon \{1, \ldots, |\hat{\mathcal{X}}^n|\} \to \hat{\mathcal{X}}^n, \tag{10}$$

with the understanding that its first guess is $\mathcal{G}_n\bigl(1\big|f_n(Y^n)\bigr)$, followed by $\mathcal{G}_n\bigl(2|f_n(Y^n)\bigr)$, etc. Thus, the guesser first asks

"Does $\mathcal{G}_n\bigl(1\big|f_n(Y^n)\bigr)$ satisfy (6)?"

If the answer is "yes," the guessing terminates and $\mathcal{G}_n\big(1\big|f_n(Y^n)\big) \in \hat{\mathcal{X}}^n$ is produced. Otherwise the guesser asks

"Does $\mathcal{G}_n\big(2\big|f_n(Y^n)\big)$ satisfy (6)?"

etc. Since guessing the same sequence twice is pointless, we assume (without loss of optimality) that, for every value of $f_n(y^n)$, the mapping $\mathcal{G}_n(\,\cdot\,|f_n(y^n))$ is injective and hence—since its domain and codomain are of equal cardinality—bijective. This and Assumption (9), allow us to define

$$G_n\big(x^n\big|f_n(y^n)\big) \triangleq \min\big\{i \geq 1\colon \bar{\mathrm{d}}(x^n, \mathcal{G}_n\big(i\big|f_n(y^n)\big) \leq D\big\} \tag{11}$$

as the number of required guesses when $X^n = x^n$ and $f_n(Y^n) = f_n(y^n)$.

Given a positive constant ρ, we seek the least exponential growth rate in n of the ρ-th moment of the number of guesses $\mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho]$:

$$\lim_{n\to\infty} \frac{1}{n} \log\Big(\min_{f_n} \min_{\mathcal{G}_n} \mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho]\Big) \tag{12}$$

(when the limit exists), where the minima in (12) are over all maps $f_n\colon \mathcal{Y}^n \to \{0,1\}^{nR}$ and all guessing strategies $\mathcal{G}_n$. Theorem 1 below asserts that the limit exists and provides a variational characterization for it.

To state the theorem, we need some additional notation. Given finite sets $\mathcal{V}$ and $\mathcal{W}$, let $\mathcal{P}(\mathcal{V})$ denote the family of PMFs on $\mathcal{V}$, and $\mathcal{P}(\mathcal{V} \mid \mathcal{W})$ the family of PMFs on $\mathcal{V}$ indexed by $\mathcal{W}$: for every $P(\cdot \mid \cdot) \in \mathcal{P}(\mathcal{V} \mid \mathcal{W})$ and every $w \in \mathcal{W}$, we have $P(\cdot \mid w) \in \mathcal{P}(\mathcal{V})$. Given PMFs $P_W \in \mathcal{P}(\mathcal{W})$ and $P_{V|W} \in \mathcal{P}(\mathcal{V} \mid \mathcal{W})$, we use $P_W\,P_{V|W}$ to denote the joint PMF $P_W(w)\,P_{V|W}(v \mid w)$ on $\mathcal{W} \times \mathcal{V}$ (in this context, $P_{V|W}(\cdot \mid \cdot)$ is the conditional PMF of V given W.)

Theorem 1. *The limit in* (12) *exists and equals*

$$\sup_{Q_Y} \inf_{Q_{U|Y}:\mathrm{I}(Q_{Y;U})\leq R} \sup_{Q_{X|YU}} \Big(\rho\,\mathrm{R}_{d,D}(Q_{X|U}) - \mathrm{D}(Q_{XYU}\|P_{XY}Q_{U|Y})\Big), \tag{13}$$

where the optimization is over $Q_Y \in \mathcal{P}(\mathcal{Y})$, $Q_{U|Y} \in \mathcal{P}(\mathcal{U} \mid \mathcal{Y})$, $Q_{X|YU} \in \mathcal{P}(\mathcal{X} \mid \mathcal{Y}\times\mathcal{U})$, and the choice of the finite set $\mathcal{U}$; where $\mathrm{I}(Q_{Y;U})$ is the mutual information between Y and U; $\mathrm{R}_{d,D}(Q_{X|U})$ is the conditional rate-distortion (R-D) function of X given U:

$$\mathrm{R}_{d,D}(Q_{X|U}) \triangleq \min_{Q_{\hat{X}|X,U}\colon\ \mathrm{E}[d(X,\hat{X})]\leq D} \mathrm{I}(Q_{X;\hat{X}|U}), \tag{14}$$

where $\mathrm{I}(Q_{X;\hat{X}|U})$ is the conditional mutual information between X and $\hat{X}$ given U; and $\mathrm{D}(\cdot\|\cdot)$ denotes relative entropy. All the expressions in (13) *are evaluated w.r.t. to $Q_{XYU} = Q_Y Q_{U|Y} Q_{X|YU}$, and those in* (14) *are w.r.t. $Q_{\hat{X}|X,U} Q_{XU}$, with Q_{XU} being the (X,U)-marginal of $Q_Y Q_{U|Y} Q_{X|YU}$.*

Remark 1. *As shown in Appendix B, restricting U to take values in a set of cardinality $|\mathcal{Y}|+1$ does not alter* (13)*. Consequently, the suprema and infimum can be replaced by maxima and minimum respectively.*

Remark 2. *In the special case where the help is direct, i.e., when Y equals X under P_{XY} so*

$$(x \neq y) \implies \big(P_{XY}(x,y) = 0\big), \tag{15}$$

Theorem 1 recovers Theorem 2 of [18].

Proof of Remark 2: This can be seen by first noting that the relative entropy in (13) is finite only when $Q_{XYU} \ll P_{XY} Q_{U|Y}$, whence $Q_{XY} \ll P_{XY}$.[1] This and (15) imply that the inner supremum in (13) is attained when X and Y are equal also under Q_{XYU}, and

$$Q_{X|YU}(x \mid y, u) = \mathbb{I}(x = y). \tag{16}$$

Using (16) and denoting expectation w.r.t. Q_{XYU} by $\mathrm{E}_{Q_{XYU}}$, we simplify $\mathrm{D}(Q_{XYU} \| P_{XY} Q_{U|Y})$ as follows:

$$\begin{aligned}
&\mathrm{D}(Q_{XYU} \| P_{XY} Q_{U|Y}) \\
&= \mathrm{D}(Q_Y Q_{U|Y} Q_{X|YU} \| P_{XY} Q_{U|Y}) && (17)\\
&= \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_Y(Y) Q_{U|Y}(U \mid Y) Q_{X|YU}(X \mid Y, U)}{P_{XY}(X,Y) Q_{U|Y}(U \mid Y)}\right)\right] && (18)\\
&= \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_Y(Y) Q_{X|YU}(X \mid Y, U)}{P_{XY}(X,Y)}\right)\right] && (19)\\
&= \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_Y(y)\, \mathbb{I}(X = Y)}{P_X(X)\, \mathbb{I}(Y = X)}\right)\right] && (20)\\
&= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} Q_Y(y)\, \mathbb{I}(x = y) \log\left(\frac{Q_Y(y)\, \mathbb{I}(x = y)}{P_X(x)\, \mathbb{I}(y = x)}\right). && (21)
\end{aligned}$$

To continue from (21), note that, by (16),

$$Q_Y(y)\, \mathbb{I}(x = y) = Q_X(x)\, \mathbb{I}(y = x), \tag{22}$$

so (21) implies that

$$\begin{aligned}
&\mathrm{D}(Q_{XYU} \| P_{XY} Q_{U|Y}) \\
&= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} Q_X(x)\, \mathbb{I}(y = x) \log\left(\frac{Q_X(x)\, \mathbb{I}(y = x)}{P_X(x)\, \mathbb{I}(y = x)}\right) && (23)\\
&= \sum_{x \in \mathcal{X}} Q_X(x) \log\left(\frac{Q_X(x)}{P_X(x)}\right) && (24)\\
&= \mathrm{D}(Q_X \| P_X). && (25)
\end{aligned}$$

Having dispensed with the inner supremum in (13), we note that, because X and Y are equal under Q_{XYU}, we can replace the outer supremum in (13) with one over Q_X, and the infimum with one over $Q_{U|X}$. From this and (25) we conclude that (13) reduces to

$$\sup_{Q_X} \inf_{Q_{U|X} : \mathrm{I}(Q_X; U) \le R} \Big(\rho\, \mathrm{R}_{d,D}(Q_{X|U}) - \mathrm{D}(Q_X \| P_X)\Big), \tag{26}$$

which recovers Theorem 2 of [18]. ■

[1] We use $Q \ll P$ to indicate that Q is absolutely continuous w.r.t. P.

Remark 3. *When the help is useless because R is zero or because X and Y are independent (under P_{XY}), Theorem 1 reduces to Corollary 1 of [4].*

Proof of Remark 3: To show this, we begin by considering the choice of U as deterministic and thus establish that (13) is upper bounded by

$$\sup_{Q_X} \Big(\rho \, \mathrm{R}_{d,D}(Q_X) - \mathrm{D}(Q_X \| P_X) \Big), \tag{27}$$

which is the expression in Corollary 1 of [4]. It remains to show that, when $R = 0$ or when X and Y are independent, this is also a lower bound.

We begin with $R = 0$. In this case, the constraint in the infimum in (13) implies that Y and U are independent under Q_{XYU}, so

$$Q_{XYU} = Q_Y \, Q_U \, Q_{X|YU}. \tag{28}$$

A lower bound results when we restrict the inner supremum to $Q_{X|YU}(x|y,u)$ that is determined by x and y, so that Q_{XYU} has the form $Q_U Q_{XY}$. With this form, the objective function in (13) reduces to

$$\Big(\rho \, \mathrm{R}_{d,D}(Q_X) - \mathrm{D}(Q_{XY} \| P_{XY}) \Big) \tag{29}$$

which depends on Q_{XYU} only via its marginal Q_{XY}. This allows us to dispense with the infimum to obtain

$$\sup_{Q_{XY}} \Big(\rho \, \mathrm{R}_{d,D}(Q_X) - \mathrm{D}(Q_{XY} \| P_{XY}) \Big), \tag{30}$$

which is attained when $Q_{Y|X}$ equals $P_{Y|X}$, whence it is equal to (27).

Having established that (27) is a lower bound on (13) when $R = 0$, we now show that it is also a lower bound on (13) when X and Y are independent. In this case we obtain the lower bound by restricting the inner supremum to be over $Q_{X|YU}(x|y,u)$ that are determined by x alone, so that Q_{XYU} has the form $Q_X Q_{UY}$. With this form (and with X and Y being independent under P_{XY}), the objective function in (13) reduces to

$$\Big(\rho \, \mathrm{R}_{d,D}(Q_X) - \mathrm{D}(Q_X Q_{YU} \| P_X P_Y Q_{U|Y}) \Big) \tag{31}$$

which simplifies to

$$\Big(\rho \, \mathrm{R}_{d,D}(Q_X) - \mathrm{D}(Q_X Q_Y \| P_X P_Y) \Big). \tag{32}$$

Again U disappears, and we are back at (30), which evaluates to the desired lower bound. ■

III. Achievability

In this section, we prove the direct part of Theorem 1, namely, that when $\{(X_i, Y_i)\}_{i=1}^n$ are IID according to P_{XY}, then for every $\epsilon > 0$ there exists a sequence of rate-R helpers $\{f_n\}$ and guessing strategies $\{\mathcal{G}_n\}$ satisfying

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log(\mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho]) \\ &\leq \sup_{Q_Y} \inf_{Q_{U|Y}: \mathrm{I}(Q_{Y;U}) \leq R} \sup_{Q_{X|YU}} \Big(\rho \, \mathrm{R}_{d,D}(Q_{X|U}) \\ &\qquad - \mathrm{D}(Q_{XYU} \| P_{XY} Q_{U|Y}) \Big) + \epsilon. \end{aligned} \tag{33}$$

Proof. Since we are only interested in the behavior of $\mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho]$ as n tends to infinity, we shall only consider large values of n.

We begin by constructing the helper f_n. To do so, we shall use the Type-Covering lemma [19, Lemma 1], [20, Lemma 9.1], [21, Lemma 2.34] that we restate here for the reader's convenience. Given finite sets $\mathcal{V}$ and $\mathcal{W}$, let $\mathcal{P}_n(\mathcal{V})$ denote the family of "types of denominator n" on $\mathcal{V}$, i.e., the PMFs $P(\cdot) \in \mathcal{P}(\mathcal{V})$ for which $nP(v)$ is an integer for all $v \in \mathcal{V}$. By a "conditional type on $\mathcal{V}$ given $\mathcal{W}$" we refer to a conditional PMF $P(\cdot \mid \cdot) \in \mathcal{P}(\mathcal{V} \mid \mathcal{W})$ for which $P(\cdot \mid w)$ is a type (of some denominator $n(w)$) for every $w \in \mathcal{W}$. Given a sequence $v^n \in \mathcal{V}^n$, the "empirical distribution of v^n" is the (unique) type $P \in \mathcal{P}_n(\mathcal{V})$ for which $P(v') = \frac{1}{n}|\{i\colon v_i = v'\}|$ for every $v' \in \mathcal{V}$. And given $P \in \mathcal{P}_n(\mathcal{V})$, we use $\mathcal{T}^{(n)}(P)$ to denote the "type class" of P, i.e., the set of all sequences $v^n \in \mathcal{V}^n$ whose empirical distribution is P.

Lemma 1 (Type-Covering lemma)**.** *Let $\mathcal{V}$ and $\mathcal{W}$ be finite sets. For every $\epsilon > 0$ there exists some $n_0(\epsilon)$ such that for all n exceeding $n_0(\epsilon)$ the following holds: For every $Q_V \in \mathcal{P}_n(\mathcal{V})$ and every conditional type $Q_{W|V}$ for which $Q_V Q_{W|V} \in \mathcal{P}_n(\mathcal{V} \times \mathcal{W})$, there exists a codebook $\mathcal{C} \subseteq \mathcal{W}^n$ satisfying*

$$|\mathcal{C}| \leq 2^{n(\mathrm{I}(Q_{V;W})+\epsilon)} \tag{34}$$

and

$$\forall v^n \in \mathcal{T}^{(n)}(Q_V)\, \exists w^n \in \mathcal{C} \colon (v^n, w^n) \in \mathcal{T}^{(n)}(Q_V Q_{W|V}). \tag{35}$$

Lemma 1 is applied as follows: For every $Q_Y \in \mathcal{P}_n(\mathcal{Y})$, we first define

$$Q^*_{U|Y}(Q_Y) \triangleq \operatorname*{arg\,min}_{Q_{U|Y}\colon \mathrm{I}(Q_{U;Y}) \leq R-\epsilon'} \max_{Q_{X|YU}} \mathrm{R}_{d,D}(Q_{X|U}), \tag{36}$$

(provided the minimum exists) where the optimization is over choice of the finite set $\mathcal{U}$, and types $Q_{U|Y}$ and $Q_{X|YU}$ for which $Q_Y Q_{U|Y} Q_{X|YU} \in \mathcal{P}_n(\mathcal{Y} \times \mathcal{U} \times \mathcal{X})$; where $\mathrm{I}(Q_{U;Y})$ and $\mathrm{R}_{d,D}(Q_{X|U})$ are computed w.r.t. $Q_Y Q_{U|Y} Q_{X|YU}$; and where ϵ' is a small positive constant (to be specified later). If the minimum in (36) does not exist, we let

$$R^*(Q_Y) \triangleq \inf_{Q_{U|Y}\colon \mathrm{I}(Q_{U;Y}) \leq R-\epsilon'} \max_{Q_{X|YU}} \mathrm{R}_{d,D}(Q_{X|U}), \tag{37}$$

where the optimization is under the same conditions as in (36), and instead define $Q^*_{U|Y}(Q_Y)$ as a conditional type satisfying

$$\max_{Q_{X|YU}} \mathrm{R}_{d,D}(Q_{X|U}) \leq R^*(Q_Y) + \epsilon'' \tag{38}$$

where the maximum is over all conditional types $Q_{X|YU}$ for which $Q_Y Q^*_{U|Y} Q_{X|YU} \in \mathcal{P}_n(\mathcal{Y} \times \mathcal{U} \times \mathcal{X})$; where $\mathrm{R}_{d,D}(Q_{X|U})$ is computed w.r.t. $Q_Y Q^*_{U|Y} Q_{X|YU}$; and where ϵ'' is a small positive constant (also to be specified later).

To construct the helper f_n, we invoke Lemma 1 (assuming that n is sufficiently large) with $Q_V \leftarrow Q_Y$, $Q_{W|V} \leftarrow Q^*_{U|Y}(Q_Y)$, and $\epsilon \leftarrow \epsilon'$ to obtain a codebook $\mathcal{C}(Q_Y) \subseteq \mathcal{U}^n$ used by f_n to produce the index of some $U^n \in \mathcal{C}(Q_Y)$ such that $(U^n, Y^n) \in \mathcal{T}^{(n)}(Q_Y Q^*_{U|Y}(Q_Y))$.

We next construct a guessing strategy $\mathcal{G}_n$. Let $U^n \in \mathcal{C}(Q_Y)$ be the codeword provided by the helper and that hence satisfies $(Y^n, U^n) \in \mathcal{T}^{(n)}(Q_Y Q^*_{U|Y}(Q_Y))$. Let Q_{XYU} denote the empirical joint distribution of (X^n, Y^n, U^n). We first argue that the guesser can be assumed cognizant of Q_{XYU}. To that end, we need the following lemma:

Lemma 2 (Interlaced-Guessing lemma [22, Lemma 5])**.** *Let V, W, and Z be finite-valued chance variables and let ρ be nonnegative. Given any guessing strategy $\mathcal{G}$ for guessing V based on W and Z, there exists a guessing strategy $\tilde{\mathcal{G}}$ based on W only such that*

$$\mathrm{E}[\tilde{G}(V \mid W)^{\rho}] \leq \mathrm{E}[G(V \mid W, Z)^{\rho}]|\mathcal{Z}|^{\rho}. \tag{39}$$

Invoking Lemma 2 with $V \leftarrow X^n$, $W \leftarrow U^n$, and $Z \leftarrow Q_{XYU}$, we see that

$$\begin{aligned} &\min_{\mathcal{G}_n} \mathrm{E}[G(X^n \mid U^n)^{\rho}] \\ &\leq \min_{\mathcal{G}_n} \mathrm{E}[G(X^n \mid U^n, Q_{XYU})^{\rho}]\big|\mathcal{P}^{(n)}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})\big|^{\rho}, \end{aligned} \tag{40}$$

where the guessing strategy on the RHS of (40) depends on both the helper's description $f_n(Y^n)$ of Y^n and the empirical joint distribution Q_{XYU} of (X^n, Y^n, U^n). Since $\big|\mathcal{P}^{(n)}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})\big|$ grows subexponentially with n,

$$\lim_{n \to \infty} \frac{1}{n} \log \big|\mathcal{P}^{(n)}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})\big|^{\rho} = 0. \tag{41}$$

Thus, by (40) and (41),

$$\begin{aligned} &\limsup_{n \to \infty} \frac{1}{n} \min_{\mathcal{G}_n} \mathrm{E}[G(X^n \mid U^n)^{\rho}] \\ &\leq \limsup_{n \to \infty} \frac{1}{n} \min_{\mathcal{G}_n} \mathrm{E}[G(X^n \mid U^n, Q_{XYU})^{\rho}]. \end{aligned} \tag{42}$$

Since the RHS of (42) cannot exceed its LHS, (42) must hold with equality, and we shall hence for the remainder of the proof assume that Q_{XYU} is known to the guesser.

Our guessing strategy $\mathcal{G}_n$ will thus depend on both the helper's description U^n of Y^n and the empirical joint distribution Q_{XYU} of (X^n, Y^n, U^n). To construct $\mathcal{G}_n$, we will use of the following corollary [18, Lemma 2] which follows from the conditional version of Lemma 1:

Corollary 1. *Let $\mathcal{V}$, $\mathcal{W}$ and $\mathcal{Z}$ be finite sets, let $\mathrm{d}(\cdot,\cdot)$ be a distortion function on $\mathcal{V} \times \mathcal{W}$, let $\bar{\mathrm{d}}(\cdot,\cdot)$ be its extension to sequences, and let D be positive. For every $\epsilon > 0$ there exists some $n_0(\epsilon)$ such that for all n exceeding $n_0(\epsilon)$ the following holds: For every $Q_{VZ} \in \mathcal{P}_n(\mathcal{V} \times \mathcal{Z})$ and every $z^n \in \mathcal{T}^{(n)}(Q_Z)$ there exists a codebook $\mathcal{C} \subseteq \mathcal{W}^n$ that satisfies*

$$|\mathcal{C}| \leq 2^{n(\mathrm{R}_{d,D}(Q_{V|Z}) + \epsilon)} \tag{43}$$

and

$$\forall v^n \in \mathcal{T}^{(n)}(Q_{V|Z}|z^n)\, \exists w^n \in \mathcal{C} \colon \bar{\mathrm{d}}(v^n, w^n) \leq D. \tag{44}$$

We invoke Corollary 1 with $Q_{VZ} \leftarrow Q_{XU}$, $z^n \leftarrow U^n$, $\mathcal{W} \leftarrow \hat{\mathcal{X}}$, and $\epsilon \leftarrow \epsilon''$, where ϵ'' is some small nonnegative constant (to be specified later) to obtain the codebook $\mathcal{C}(Q_{XYU}) \subseteq \hat{\mathcal{X}}^n$. The guessing strategy $\mathcal{G}_n$ is then chosen such that $\mathcal{G}_n|_{\{1,\ldots,|\mathcal{C}(Q_{XYU})|\}}$ is a bijection from $\{1, \ldots, |\mathcal{C}(Q_{XYU})|\}$ to $\mathcal{C}(Q_{XYU})$, i.e., such the first $|\mathcal{C}(Q_{XYU})|$ guesses are those in $\mathcal{C}(Q_{XYU})$ in some arbitrary order. Note that (44) guarantees that some $\hat{X}^n$ in $\mathcal{G}_n|_{\{1,\ldots,|\mathcal{C}(Q_{XYU})|\}}$ satisfies (6), and thus the guesser succeeds after at most $|\mathcal{C}(Q_{XYU})|$ guesses.

We now show that (33) holds for our proposed helper f_n and guessing strategy $\mathcal{G}_n$:

$$\begin{aligned}
&\mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho] \\
&\overset{(a)}{=} \mathrm{E}[G_n(X^n \mid U^n)^\rho] && (45)\\
&\overset{(b)}{=} \sum_{Q_Y} \sum_{Q_{X|YU}} \Big(\Pr[Y^n \in \mathcal{T}^{(n)}(Q_Y)] \\
&\quad \Pr[X^n \in \mathcal{T}^{(n)}(Q_{X|YU}) \mid Y^n \in \mathcal{T}^{(n)}(Q_Y)] \\
&\quad \mathrm{E}[G_n(X^n \mid U^n)^\rho \\
&\qquad \mid (X^n, Y^n, U^n) \in \mathcal{T}^{(n)}(Q_Y Q^*_{U|Y}(Q_Y) Q_{X|YU})]\Big) && (46)\\
&\overset{(c)}{\le} \sum_{Q_Y} \sum_{Q_{X|YU}} \Big(\Pr[Y^n \in \mathcal{T}^{(n)}(Q_Y)] \\
&\quad \Pr[X^n \in \mathcal{T}^{(n)}(Q_{X|YU}) \mid Y^n \in \mathcal{T}^{(n)}(Q_Y)] \\
&\quad 2^{n\rho(\mathrm{R}_{d,D}(Q_{X|U}))+\epsilon'')}\Big) && (47)\\
&\overset{(d)}{\le} \sum_{Q_Y} \sum_{Q_{X|YU}} \Big(2^{-n\,\mathrm{D}(Q_Y \| P_Y)} 2^{-n\,\mathrm{D}(Q_{X|YU} \| P_{X|Y})} \\
&\quad 2^{n\rho(\mathrm{R}_{d,D}(Q_{X|U})+\epsilon'')}\Big) && (48)\\
&\overset{(e)}{\le} \max_{Q_Y} \max_{Q_{X|YU}} \Big(2^{-n\,\mathrm{D}(Q_Y \| P_Y)} 2^{-n\,\mathrm{D}(Q_{X|YU} \| P_{X|Y})} \\
&\quad 2^{n\rho(\mathrm{R}_{d,D}(Q_{X|U})+\epsilon'')}\Big) \big|\mathcal{P}^{(n)}(\mathcal{X}\times\mathcal{Y}\times\mathcal{U})\big|^\rho && (49)\\
&\overset{(f)}{=} \max_{Q_Y} \min_{Q_{U|Y}:\mathrm{I}(Q_{U;Y}) \le R-\epsilon'} \max_{Q_{X|YU}} \Big(2^{-n\,\mathrm{D}(Q_Y \| P_Y)} \\
&\quad 2^{-n\,\mathrm{D}(Q_{X|YU} \| P_{X|Y})} 2^{n\rho(\mathrm{R}_{d,D}(Q_{X|U})+\epsilon'')} \\
&\quad \big|\mathcal{P}^{(n)}(\mathcal{X}\times\mathcal{Y}\times\mathcal{U})\big|^\rho\Big) && (50)\\
&\overset{(g)}{\le} \sup_{Q_Y} \inf_{Q_{U|Y}:\mathrm{I}(Q_{U;Y}) \le R-\epsilon'} \sup_{Q_{X|YU}} \Big(2^{-n\,\mathrm{D}(Q_Y \| P_Y)} \\
&2^{-n\,\mathrm{D}(Q_{X|YU} \| P_{X|Y})} 2^{n\rho(\mathrm{R}_{d,D}(Q_{X|U})+\epsilon'')} 2^{n\delta_n}\Big) \big|\mathcal{P}^{(n)}(\mathcal{X}\times\mathcal{Y}\times\mathcal{U})\big|^\rho && (51)\\
&\overset{(h)}{\le} \sup_{Q_Y} \inf_{Q_{U|Y}:\mathrm{I}(Q_{U;Y}) \le R} \sup_{Q_{X|YU}} \Big(2^{-n\,\mathrm{D}(Q_Y \| P_Y)} \\
&2^{-n\,\mathrm{D}(Q_{X|YU} \| P_{X|Y})} 2^{n\rho\,\mathrm{R}_{d,D}(Q_{X|U})} 2^{n\delta_n} 2^{n\epsilon}\Big) \big|\mathcal{P}^{(n)}(\mathcal{X}\times\mathcal{Y}\times\mathcal{U})\big|^\rho, && (52)
\end{aligned}$$

where (a) holds because we have assumed that the empirical distribution Q_Y of Y^n is known to the guesser who can thus recover U^n from $f_n(Y^n)$ and $\mathcal{C}(Q_Y)$; in (b) we have used the law of total expectation, averaging over the types $Q_Y \in \mathcal{P}_n(\mathcal{Y})$ and conditional types $Q_{X|YU}$ for which $Q_Y Q_{U|Y} Q_{X|YU} \in \mathcal{P}_n(\mathcal{Y}\times\mathcal{U}\times\mathcal{X})$ (recall that $Q_{U|Y} = Q^*_{U|Y}(Q_Y)$ is fixed by f_n); (c) is due to (43); (d) follows from [23, Theorem 11.1.4]; in (e) we have upper-bounded the sum by the largest term times the number of terms (the number of terms is the number of types Q_Y and $Q_{X|YU}$ that we have in turn upper-bounded by the number of types Q_{XYU}); (f) is due to (36); in (g) we have lifted the constraint on Q_Y, $Q^*_{U|Y}(Q_Y)$, and $Q_{X|YU}$ to be types at a cost of at most $2^{n\delta_n}$, where $\delta_n \downarrow 0$ as $n \to \infty$, and where the step is justified because any PMF can be approximated arbitrarily well by a type of sufficiently large denominator; and in (h) we have used the fact that all exponents are continuous functions of their respective arguments, and that ϵ' and ϵ'' were chosen sufficiently small.

Dividing the $\log$ of (52) by n, taking the $\limsup$ as n tends to infinity, and applying (41) yields (33). □

IV. Converse

In this section we prove the converse part of Theorem 1, namely, that when $\{(X_i, Y_i)\}_{i=1}^n$ are IID according to P_{XY}, then for any sequence of rate-R helpers $\{f_n\}$ and guessing strategies $\{\mathcal{G}_n\}$,

$$\begin{aligned}&\liminf_{n\to\infty} \frac{1}{n} \log(\mathrm{E}[G_n(X^n \mid f_n(Y^n))^\rho]) \\ &\geq \sup_{Q_Y} \inf_{Q_{U|Y}:\mathrm{I}(Q_Y;U)\leq R} \sup_{Q_{X|YU}} \Big(\rho\, \mathrm{R}_{d,D}(Q_{X|U}) \\ &\quad - \mathrm{D}(Q_{XYU} \| P_{XY} Q_{U|Y})\Big). \end{aligned} \tag{53}$$

Proof. Fix a sequence of helpers $\{f_n\}$ and guessing strategies $\{\mathcal{G}_n\}$. We begin by observing that for any probability law Q of (X^n, Y^n)-marginal $Q_{X^nY^n}$,

$$\begin{aligned}&\mathrm{E}_{P_{X^nY^n}}[G_n(X^n \mid f_n(Y^n))^\rho] \\ &\geq 2^{\rho\, \mathrm{E}_Q[\log(G_n(X^n|f_n(Y^n)))] - \mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n})}, \end{aligned} \tag{54}$$

where E_P denotes expectation w.r.t. the PMF P. Indeed,

$$\begin{aligned}&\mathrm{E}_{P_{X^nY^n}}[G_n(X^n \mid f_n(Y^n))^\rho] \\ &= \sum_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} P_{X^n,Y^n}(x^n,y^n)\, G_n(x^n \mid f_n(y^n))^\rho && (55)\\ &= \sum_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} Q_{X^n,Y^n}(x^n,y^n)\, G_n(x^n \mid f_n(y^n))^\rho \frac{P_{X^n,Y^n}(x^n,y^n)}{Q_{X^n,Y^n}(x^n,y^n)} && (56)\\ &= \sum_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} Q_{X^n,Y^n}(x^n,y^n)\, 2^{\log\left(G_n(x^n|f_n(y^n))^\rho \frac{P_{X^n,Y^n}(x^n,y^n)}{Q_{X^n,Y^n}(x^n,y^n)}\right)} && (57)\\ &\overset{(a)}{\geq} 2^{\sum_{x^n,y^n} Q_{X^n,Y^n} \log\left(G_n(x^n|f_n(y^n))^\rho \frac{P_{X^n,Y^n}(x^n,y^n)}{Q_{X^n,Y^n}(x^n,y^n)}\right)} && (58)\\ &= 2^{\rho\, \mathrm{E}_Q[\log(G_n(X^n|f_n(Y^n)))] - \mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n})}, && (59)\end{aligned}$$

where (a) follows from Jensen's inequality.

To describe the law Q to which we shall apply (54), let $[1:n]$ denote the set $\{1,\ldots,n\}$ and define the auxiliary variables

$$M \triangleq f_n(Y^n) \tag{60}$$

$$U_i \triangleq (X^{i-1}, Y^{i-1}, M), \quad i \in [1:n] \tag{61}$$

taking values in the sets

$$\mathcal{M} \triangleq \{0,1\}^{nR} \tag{62}$$

and

$$\mathcal{U}_i \triangleq \mathcal{X}^{i-1} \times \mathcal{Y}^{i-1} \times \mathcal{M}, \quad i \in [1:n]. \tag{63}$$

Given any $Q_Y \in \mathcal{P}(\mathcal{Y})$ and any n Markov kernels $\{Q_{X_i|Y_iU_i}\}_{i=1}^n$, define the law $Q_{X^nY^nMU^n\hat{X}^n}$ on $\mathcal{Y}^n \times \mathcal{X}^n \times \mathcal{M} \times \prod_{i=1}^n \mathcal{U}_i \times \hat{\mathcal{X}}^n$ as

$$Q_{X^nY^nMU^n\hat{X}^n} \triangleq Q_Y^{\times n} P_{M|Y^n} \prod_{i=1}^n \left(Q_{U_i|X^{i-1}Y^{i-1}M} Q_{X_i|Y_iU_i}\right) P_{\hat{X}^n|MX^n}, \tag{64a}$$

where $P_{M|Y^n}$ is specified by the helper as

$$P_{M|Y^n}(m \mid y^n) = \mathbb{I}(m = f_n(y^n)), \tag{64b}$$

$Q_{U_i|X^{i-1}Y^{i-1}M}$ is specified through the definition of U_i in (61) as

$$Q_{U_i|X^{i-1}Y^{i-1}M}(u_i \mid x^{i-1}y^{i-1}m) = \mathbb{I}(u_i = (x^{i-1}, y^{i-1}, m)), \tag{64c}$$

and $P_{\hat{X}^n|MX^n}$ is determined by the guessing strategy as

$$P_{\hat{X}^n|MX^n}(\hat{x}^n \mid m, x^n) = \mathbb{I}(\hat{x}^n = \mathcal{G}_n(G_n(x^n \mid m))). \tag{64d}$$

Thus,

$$\begin{aligned} &Q_{X^nY^nMU^n\hat{X}^n}(y^n, x^n, m, u^n, \hat{x}^n) \\ &= Q_Y^{\times n}(y^n)\, \mathbb{I}(m = f_n(y^n)) \prod_{i=1}^n \left(\mathbb{I}(u_i = (x^{i-1}, y^{i-1}, m)) Q_{X_i|Y_iU_i}(x_i \mid y_i, u_i)\right) \\ &\quad \mathbb{I}(\hat{x}^n = \mathcal{G}_n(G_n(x^n \mid m))), \end{aligned} \tag{64e}$$

where $G_n(\cdot)$ is defined in (11).

Note that (64a) implies that

$$X^{i-1} \to (M, Y^{i-1}) \to Y_i \quad \text{under } Q \tag{65}$$

because the $(X^{i-1}, Y^n, M, U^{i-1})$-marginal of Q can be written as

$$Q_{X^{i-1}Y^nMU^{i-1}} = Q_Y^{\times n}(y^n) P_{M|Y^n} \prod_{j=1}^{i-1} \left(Q_{U_j|X^{j-1}Y^{j-1}M} Q_{X_j|Y_jU_j}\right), \tag{66}$$

which impies that

$$(X^{i-1}, U^{i-1}) \to (M, Y^{i-1}) \to Y_i^n \quad \text{under } Q \tag{67}$$

because the product is a function of (m, y^{i-1}) and (x^{i-1}, u^{i-1}), and the pre-product $Q_Y^{\times n}(y^n) P_{M|Y^n}$ is a function of (m, y^{i-1}) and y_i^n.

Next define for every $i \in [1:n]$

$$D_i \triangleq \mathrm{E}[d(X_i, \hat{X}_i)], \tag{68}$$

where the expectation is w.r.t. to the PMF $Q_{X^nY^nMU^n\hat{X}^n}$. Under the latter, $\hat{x}^n = \mathcal{G}_n(G_n(x^n \mid m))$ so $\bar{\mathrm{d}}(x^n, \hat{x}^n) \leq D$ for *every* $x^n \in \mathcal{X}^n$ and, also in expectation (over $Q_{X^nY^nMU^n\hat{X}^n}$)

$$\frac{1}{n}\sum_{i=1}^{n} D_i \leq D. \tag{69}$$

Further define

$$Q^*_{\hat{X}'_i|MX^i} \triangleq \operatorname*{arg\,min}_{\substack{Q_{\hat{X}'_i|MX^i}: \\ \mathrm{E}[d(X_i,\hat{X}'_i)]\leq D_i}} \mathrm{I}(X_i; \hat{X}'_i \mid M, X^{i-1}), \tag{70}$$

where the minimum is over all conditional PMFs $Q_{\hat{X}'_i|MX^i} \in \mathcal{P}(\hat{\mathcal{X}} \mid \mathcal{M} \times \mathcal{X}^i)$, and where $\mathrm{I}(X_i; \hat{X}'_i \mid M, X^{i-1})$ and $\mathrm{E}[d(X_i, \hat{X}'_i)]$ are evaluated w.r.t. $Q_{\hat{X}'_i|MX^i}Q_{MX^i}$, with Q_{MX^i} being the (M, X^i)-marginal of $Q_{X^nY^nMU^n\hat{X}^n}$. Using $\{Q^*_{\hat{X}'_i|MX^i}\}_{i=1}^n$, we extend $Q_{X^nY^nMU^n\hat{X}^n}$ to a law Q on $\mathcal{Y}^n \times \mathcal{X}^n \times \mathcal{M} \times \prod_{i=1}^n \mathcal{U}_i \times \hat{\mathcal{X}}^n \times \hat{\mathcal{X}}^n$ as follows:

$$Q \triangleq Q_{X^nY^nMU^n\hat{X}^n} \prod_{i=1}^{n} Q^*_{\hat{X}'_i|MX^i}. \tag{71}$$

Note that the factorization in (71) implies that

$$\hat{X}'_i \to (M, X^i) \to Y^{i-1} \tag{72}$$

because it implies that—conditional on (M, X^i)—$\hat{X}'_i$ is independent of the tuple $(X^n, Y^n, M, U^n, \hat{X}^n)$ and hence also of Y^{i-1} (which is a function of this tuple). For the remainder of this section we shall assume that, unless stated otherwise, all expectations and information-theoretic quantities are evaluated w.r.t. Q. To study (54) for this Q, we begin by lower-bounding $\mathrm{E}[\log(G_n(X^n \mid M))]$ using the conditional R-D function. To this end, we note that, conditional on $M = m$, there is a one-to-one correspondence between $G_n(X^n \mid M)$ and $\hat{X}^n$ so, by the Reverse Wyner inequality of Corollary 2 in Appendix A,

$$\mathrm{E}[\log(G_n(X^n \mid M)) \mid M = m] \geq \mathrm{H}(\hat{X}^n \mid M = m) - n\delta_n \tag{73}$$

with δ_n tending to zero as n tends to infinity. Averaging over M,

$$\begin{aligned}
&\mathrm{E}[\log(G_n(X^n \mid M))] \\
&\geq \mathrm{H}(\hat{X}^n \mid M) - n\delta_n && (74)\\
&\geq \mathrm{I}(\hat{X}^n; X^n \mid M) - n\delta_n && (75)\\
&= \sum_{i=1}^{n} \Big(\mathrm{H}(X_i \mid M, X^{i-1}) - \mathrm{H}(X_i \mid M, \hat{X}^n, X^{i-1}) \Big) - n\delta_n && (76)\\
&\geq \sum_{i=1}^{n} \Big(\mathrm{H}(X_i \mid M, X^{i-1}) - \mathrm{H}(X_i \mid M, \hat{X}_i, X^{i-1}) \Big) - n\delta_n && (77)\\
&= \sum_{i=1}^{n} \mathrm{I}(X_i; \hat{X}_i \mid M, X^{i-1}) - n\delta_n && (78)\\
&\overset{(a)}{\geq} \sum_{i=1}^{n} \mathrm{I}(X_i; \hat{X}'_i \mid M, X^{i-1}) - n\delta_n && (79)\\
&= \sum_{i=1}^{n} \Big(\mathrm{H}(\hat{X}'_i \mid M, X^{i-1}) - \mathrm{H}(\hat{X}'_i \mid M, X^{i}) \Big) - n\delta_n && (80)\\
&\geq \sum_{i=1}^{n} \Big(\mathrm{H}(\hat{X}'_i \mid M, X^{i-1}, Y^{i-1}) - \mathrm{H}(\hat{X}'_i \mid M, X^{i}) \Big) - n\delta_n && (81)\\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big(\mathrm{H}(\hat{X}'_i \mid M, X^{i-1}, Y^{i-1}) - \mathrm{H}(\hat{X}'_i \mid M, X^{i}, Y^{i-1}) \Big) && (82)\\
&\quad - n\delta_n \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big(\mathrm{H}(\hat{X}'_i \mid U_i) - \mathrm{H}(\hat{X}'_i \mid U_i, X_i) \Big) - \delta_n && (83)\\
&= \sum_{i=1}^{n} \mathrm{I}(X_i; \hat{X}'_i \mid U_i) - n\delta_n, && (84)
\end{aligned}$$

where in (a) we have replaced $\hat{X}_i$ by $\hat{X}'_i$, and the inequality hence follows from (70); (b) follows from (72); and in (c) we have identified the auxiliary variable U_i defined in (61). To continue from (84), let T be equiprobable over $[1:n]$, independent of $(Y^n, M, X^n, U^n, (\hat{X}')^n)$, and define the chance variable

$$(Y, X, U, \hat{X}') \triangleq (Y_T, X_T, U_T, \hat{X}'_T) \tag{85}$$

taking values in the set $\mathcal{Y} \times \mathcal{X} \times (\cup_{i=1}^{n} \mathcal{U}_i) \times \hat{\mathcal{X}}$. Note that, since the sets $\{\mathcal{U}_i\}$ of (63) are disjoint, T is a deterministic function of U, and we can define $\iota(\cdot)$ as mapping each $u \in \cup_{i=1}^{n} \mathcal{U}_i$ to the unique $i \in [1:n]$ for which $u \in \mathcal{U}_i$. With this definition, the PMF of $(Y, X, U, \hat{X}')$ can be expressed as

$$\tilde{Q}_{YXU\hat{X}'}(y, x, u, \hat{x}') \triangleq \frac{1}{n} Q_{Y_{\iota(u)} X_{\iota(u)} U_{\iota(u)} \hat{X}'_{\iota(u)}}(y, x, u, \hat{x}'), \tag{86}$$

where $Q_{Y_i X_i U_i \hat{X}'_i}$ is the $(Y_i, X_i, U_i, \hat{X}'_i)$-marginal of Q. We next observe that, under $\tilde{Q}$, $\mathrm{E}[\mathrm{d}(X, \hat{X}')]$ is upper-bounded by D. Indeed,

$$\mathrm{E}_{\tilde{Q}}[\mathrm{d}(X,\hat{X}')] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_Q[\mathrm{d}(X_i,\hat{X}_i')] \tag{87}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} D_i \tag{88}$$

$$\leq D, \tag{89}$$

where the first inequality follows from the constraint in the optimization on the RHS of (70) and the second from (69). Also note that, since T is a deterministic function of U, the RHS of (84) can be expressed in terms of $(Y,X,U,\hat{X}')$ as

$$n\,\mathrm{I}(X;\hat{X}' \mid U) - n\delta_n, \tag{90}$$

so,

$$\mathrm{E}_Q[\log(G_n(X^n \mid M))] \geq n\,\mathrm{I}(X;\hat{X}' \mid U) - n\delta_n, \tag{91}$$

where the conditional mutual information on the RHS is w.r.t. $\tilde{Q}$. Using (89), we can lower-bound the RHS of (91) in terms of the conditional R-D function (14),

$$n\,\mathrm{I}(X;\hat{X}' \mid U) - n\delta_n \geq n\,\mathrm{R}_{d,D}(\tilde{Q}_{X|U}) - n\delta_n, \tag{92}$$

and, using (92) and (91), we obtain the desired lower bound

$$\mathrm{E}_Q[\log(G_n(X^n \mid M))] \geq n\,\mathrm{R}_{d,D}(\tilde{Q}_{X|U}) - n\delta_n. \tag{93}$$

We next return to (54) and derive a single-letter expression for $\mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n})$, where $Q_{X^nY^n}$ is the (X^n,Y^n)-marginal of Q, and

$$P_{X^nY^n} = P_{XY}^{\times n}. \tag{94}$$

We first express it as

$$\mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n}) = \mathrm{D}(Q_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM}\|P_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM}), \tag{95}$$

and then observe that $Q_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM}$ is (a factorization of) the (X^n,Y^n,M,U^n)-marginal of Q, which can be expressed as

$$Q_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM} = Q_Y^{\times n}\Big(\prod_{i=1}^{n} Q_{X_i|Y_iU_i}\Big)P_{M|Y^n}Q_{U^n|X^nY^nM}, \tag{96}$$

because, by (64a) (or (61)),

$$Q_{U^n|X^nY^nM} = \prod_{i=1}^{n} Q_{U_i|X^{i-1}Y^{i-1}M}. \tag{97}$$

From (94), (96) and (95)

$$\mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n}) = \mathrm{D}(Q_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM}\|P_{X^nY^n}P_{M|Y^n}Q_{U^n|X^nY^nM}) \tag{98}$$

$$= \mathrm{D}\left(Q_Y^{\times n}\Big(\prod_{i=1}^{n} Q_{X_i|Y_iU_i}\Big)P_{M|Y^n}Q_{U^n|X^nY^nM}\,\middle\|\,P_{XY}^{\times n}P_{M|Y^n}Q_{U^n|X^nY^nM}\right). \tag{99}$$

We now continue the derivation of a single-letter expression for $\mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n})$ by studying the RHS of (99):

$$\begin{aligned}
&\mathrm{D}(Q_{X^nY^n}\|P_{X^nY^n}) \\
&= \mathrm{D}\left(Q_Y^{\times n}\prod_{i=1}^{n} Q_{X_i|Y_iU_i}P_{M|Y^n}Q_{U^n|X^nY^nM}\middle\| P_{XY}^{\times n}P_{M|Y^n}Q_{U^n|X^nY^nM}\right) && (100)\\
&\overset{(a)}{=} \mathrm{E}_Q\left[\log\left(\frac{Q_Y^{\times n}(Y^n)\prod_{i=1}^{n} Q_{X_i|Y_iU_i}(X_i\mid Y_i,U_i)P_{M|Y^n}(M\mid Y^n)Q_{U^n|X^nY^nM}(U^n\mid X^n,Y^n,M)}{P_{XY}^{\times n}(X^n,Y^n)P_{M|Y^n}(M\mid Y^n)Q_{U^n|X^nY^nM}(U^n\mid X^n,Y^n,M)}\right)\right] && (101)\\
&= \mathrm{E}_Q\left[\log\left(\frac{Q_Y^{\times n}(Y^n)\prod_{i=1}^{n} Q_{X_i|Y_iU_i}(X_i\mid Y_i,U_i)}{P_{XY}^{\times n}(X^n,Y^n)}\right)\right] && (102)\\
&\overset{(b)}{=} \sum_{i=1}^{n}\mathrm{E}_{Q_{X_iY_iU_i}}\left[\log\left(\frac{Q_Y(Y_i)Q_{X_i|Y_iU_i}(X_i\mid Y_i,U_i)}{P_{XY}(X_i,Y_i)}\right)\right] && (103)\\
&= \sum_{i=1}^{n}\sum_{(x_i,y_i,u_i)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{U}_i} Q_{X_iY_iU_i}(x_i,y_i,u_i)\log\left(\frac{Q_Y(y_i)Q_{X_i|Y_iU_i}(x_i\mid y_i,u_i)}{P_{XY}(x_i,y_i)}\right) && (104)\\
&\overset{(c)}{=} \sum_{i=1}^{n}\sum_{(x_i,y_i,u_i)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{U}_i} Q_{X_iY_iU_i}(x_i,y_i,u_i)\log\left(\frac{Q_{Y_i}(y_i)Q_{X_i|Y_iU_i}(x_i\mid y_i,u_i)}{P_{XY}(x_i,y_i)}\right) && (105)\\
&= n\sum_{i=1}^{n}\sum_{(x_i,y_i,u_i)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{U}_i} \frac{1}{n}Q_{X_iY_iU_i}(x_i,y_i,u_i)\log\left(\frac{Q_{Y_i}(y_i)Q_{U_i|Y_i}(u_i\mid y_i)Q_{X_i|Y_iU_i}(x_i\mid y_i,u_i)\frac{1}{n}}{P_{XY}(x_i,y_i)Q_{U_i|Y_i}(u_i\mid y_i)\frac{1}{n}}\right) && (106)\\
&\overset{(d)}{=} n\sum_{i=1}^{n}\sum_{(x_i,y_i,u_i)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{U}_i} \tilde{Q}(x_i,y_i,u_i)\log\left(\frac{\tilde{Q}(x_i,y_i,u_i)}{P_{XY}(x_i,y_i)\tilde{Q}_{U|Y}(u_i\mid y_i)}\right) && (107)\\
&= n\sum_{(x,y,u)\in\mathcal{X}\times\mathcal{Y}\times(\cup_{i=1}^{n}\mathcal{U}_i)} \tilde{Q}(x,y,u)\log\left(\frac{\tilde{Q}(x,y,u)}{P_{XY}(x,y)\tilde{Q}_{U|Y}(u\mid y)}\right) && (108)\\
&= n\,\mathrm{D}(\tilde{Q}_{XYU}\|P_{XY}\tilde{Q}_{U|Y}), && (109)
\end{aligned}$$

where (a) follows from the definition of the relative entropy and the fact that $Q_Y^{\times n}\prod_{i=1}^{n} Q_{X_i|Y_iU_i}P_{M|Y^n}Q_{U^n|X^nY^nM}$ is (a factorization of) the (X^n,Y^n,M,U^n)-marginal of Q; in (b) we have used that for nonnegative x and y, $\log(xy)=\log(x)+\log(y)$, and we used $Q_{X_iY_iU_i}$ to denote the (X_i,Y_i,U_i)-marginal of Q; (c) holds because under Q, $Y^n\sim \mathrm{IID}\, Q_Y$; and in (d) we have identified $\frac{1}{n}Q_{X_iY_iU_i}$ as the (X,Y,U)-marginal of $\tilde{Q}$.

We next show that, $\mathrm{I}_{\tilde{Q}}(Y;U)$— the mutual information between Y and U under $\tilde{Q}$—is upper-bounded by R. To that end first observe that by definition of $\tilde{Q}$ (in (85) and (86)) we can express $\mathrm{I}_{\tilde{Q}}(Y;U)$ as

$$\mathrm{I}_{\tilde{Q}}(Y;U) = \frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{H}_Q(Y_i)-\mathrm{H}_Q(Y_i\mid U_i)\Big). \tag{110}$$

So continuing from the RHS of (110), with all information-theoretic quantities implicitly evaluated w.r.t. Q:

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{H}(Y_i)-\mathrm{H}(Y_i \mid U_i)\Big)=\frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{H}(Y_i)-\mathrm{H}(Y_i \mid X^{i-1},Y^{i-1},M)\Big) \tag{111}$$

$$\stackrel{(a)}{=}\frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{H}(Y_i)-\mathrm{H}(Y_i \mid Y^{i-1},M)\Big) \tag{112}$$

$$\stackrel{(b)}{=}\frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{H}(Y_i \mid Y^{i-1})-\mathrm{H}(Y_i \mid Y^{i-1},M)\Big) \tag{113}$$

$$=\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}(Y_i;M \mid Y^{i-1}) \tag{114}$$

$$=\frac{1}{n}\mathrm{I}(Y^n;M) \tag{115}$$

$$\leq\frac{1}{n}\mathrm{H}(M) \tag{116}$$

$$\stackrel{(c)}{\leq} R, \tag{117}$$

where (a) holds because, under Q, $X^{i-1}\to(Y^{i-1},M)\to Y_i$ (65); (b) holds because Y^n is IID under Q; and (c) holds because M can assume at most 2^{nR} distinct values.

We now use (54), (92), (109), and (117) to derive the converse part of Theorem 1 as stated in (53). Starting with (54), we use (92) and (109) to obtain

$$\begin{aligned}&\mathrm{E}_{P_{X^nY^n}}[G_n(X^n \mid f_n(Y^n))^{\rho}]\\&\geq 2^{n(\rho\,\mathrm{R}_{d,D}(\tilde{Q}_{X|U})-\mathrm{D}(\tilde{Q}_{XYU}\|P_{XY}\tilde{Q}_{U|Y})-\delta_n)},\end{aligned} \tag{118}$$

where the PMF $\tilde{Q}$ on the RHS of (118) is defined in (86). Taking the logarithm and dividing by n on both sides,

$$\begin{aligned}&\frac{1}{n}\log(\mathrm{E}[G_n(X^n \mid f_n(Y^n)^{\rho}])\\&\geq\rho\,\mathrm{R}_{d,D}(\tilde{Q}_{X|U})-\mathrm{D}(\tilde{Q}_{XYU}\|P_{XY}\tilde{Q}_{U|Y})-\delta_n.\end{aligned} \tag{119}$$

Since the choice of Q_Y and $\{Q_{X_i|Y_iU_i}\}_{i=1}^{n}$ in (64a) is arbitrary, so is that of $\tilde{Q}_Y$ and $\tilde{Q}_{X|YU}$ in the (X,Y,U)-marginal $\tilde{Q}_{XYU}=\tilde{Q}_Y\tilde{Q}_{U|Y}\tilde{Q}_{X|YU}$ of $\tilde{Q}$ (86). We are therefore at liberty to choose those so as to obtain the tightest bound. Things are different with regard to $\tilde{Q}_{U|Y}$, because it is influenced by the helper f_n, and we must ensure that the bound is valid for all helpters. Ostensibly, we should therefore consider the choice of $\tilde{Q}_{U|Y}$ that yields the loosest bound. However, $\tilde{Q}_{U|Y}$ cannot be arbitrary: irrespective of our choice of $\tilde{Q}_Y$, the mutual information $I_{\tilde{Q}}(U;Y)$ must be upper bounded by R (117).

These considerations allow to infer form (119) that

$$\begin{aligned}&\frac{1}{n}\log(\mathrm{E}[G_n(X^n \mid f_n(Y^n)^{\rho}])\\&\geq\sup_{\tilde{Q}_Y}\inf_{Q_{\tilde{U}|Y}:\ \mathrm{I}(\tilde{Q}_{Y;U})\leq R}\sup_{\tilde{Q}_{X|YU}}\Big(\rho\,\mathrm{R}_{d,D}(\tilde{Q}_{X|U})-\mathrm{D}(\tilde{Q}_{XYU}\|P_{XY}\tilde{Q}_{U|Y})\Big)-\delta_n,\end{aligned} \tag{120}$$

which, upon taking n to infinity, yields (53). □

Appendix A

Lemma 3. *Let X be a chance variable taking values in the finite set $\mathcal{X}$ according to some PMF P, and let f be a bijection from $\mathcal{X}$ to $[1:|\mathcal{X}|]$. Then, for $X \sim P$,*

$$\mathrm{E}[\log(f(X))] \geq \mathrm{H}(X) - \log(\ln(|\mathcal{X}|) + 3/2). \tag{121}$$

Proof. Outcomes of zero probability contribute neither to the LHS nor to the RHS of (121), and we therefore assume w.l.g. that $P(x) > 0$ for every $x \in \mathcal{X}$. We then have

$$\begin{aligned}
\mathrm{E}[\log(f(X))] &= \sum_{x\in\mathcal{X}} P(x)\log(f(x)) && (122)\\
&= \sum_{x\in\mathcal{X}} P(x)\log\left(\frac{f(x)P(x)}{P(x)}\right) && (123)\\
&= \mathrm{H}(X) + \sum_{x} P(x)\log(f(x)P(x)) && (124)\\
&= \mathrm{H}(X) - \sum_{x} P(x)\log\left(\frac{1}{f(x)P(x)}\right) && (125)\\
&\overset{(a)}{\geq} \mathrm{H}(X) - \log\left(\sum_{x}\frac{1}{f(x)}\right) && (126)\\
&\overset{(b)}{=} \mathrm{H}(X) - \log\left(\sum_{i=1}^{|\mathcal{X}|}\frac{1}{i}\right) && (127)\\
&\overset{(c)}{\geq} \mathrm{H}(X) - \log(\ln(|\mathcal{X}|) + 3/2), && (128)
\end{aligned}$$

where (a) follows from Jensen's inequality; (b) holds because f maps onto $[1:|\mathcal{X}|]$; and (c) holds because $\sum_{i=1}^{n} 1/i$ is upper-bounded by $\ln(n) + 3/2$. □

Corollary 2. *Let $\mathcal{X}$ be a finite set, and let f be a bijection from $\mathcal{X}^n$ to $[1:|\mathcal{X}|^n]$. Then, for any chance variable X^n on $\mathcal{X}^n$,*

$$\mathrm{E}[\log(f(X^n))] \geq \mathrm{H}(X^n) - n\delta_n, \tag{129}$$

where $\delta_n = \delta_n(|\mathcal{X}|)$ and for every fixed $|\mathcal{X}|$,

$$\delta_n \to 0. \tag{130}$$

Proof. The corollary follows from Lemma 3 and the fact that when $|\mathcal{X}|$ is fixed,

$$\lim_{n\to\infty} \frac{\log(\ln(|\mathcal{X}^n|) + 3/2)}{n} = 0. \tag{131}$$

□

Appendix B

We prove that restricting U to take values in a set of cardinality $|\mathcal{Y}| + 1$ does not alter (13). To that end, we first express the objective function in (13) as an expectation over U of a quantity $\Psi(Q_{Y|U=u}, Q_{X|YU=u})$ that depends explicitly on $Q_{Y|U=u}$, $Q_{X|YU=u}$ and implicitly on the given joint PMF P_{XY} and the PMF Q_Y (which is determined in the outer maximization). Specifically,

$$\rho\,\mathrm{R}_{d,D}(Q_{X|U}) - \mathrm{D}(Q_{XYU}\|P_{XY}Q_{U|Y}) = \sum_{u\in\mathcal{U}} Q_U(u)\,\Psi(Q_{Y|U=u}, Q_{X|YU=u}), \tag{132a}$$

with

$$\Psi(Q_{Y|U=u}, Q_{X|YU=u}) = \rho\,\mathrm{R}_{d,D}(Q_{X|U=u}) + H(Q_Y) - H(Q_{Y|U=u}) - D(Q_{Y|U=u}Q_{X|Y,U=u}||P_{XY}) \tag{132b}$$

where $\mathrm{R}_{d,D}(Q_{X|U=u})$ is determined by $Q_{Y|U=u}$ and $Q_{X|YU=u}$ via the relation

$$Q_{X|U=u}(x|u) = \sum_{y\in\mathcal{Y}} Q_{Y|U=u}(y|u)\,Q_{X|YU=u}(x|y,u). \tag{133}$$

Indeed, (132) follow from

$$\begin{aligned}
&\mathrm{D}(Q_{XYU}\|P_{XY}Q_{U|Y}) \\
&= \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_{XY|U}(XY\mid U)\,Q_U(U)}{P_{XY}(X,Y)\,Q_{U|Y}(U\mid Y)}\right)\right] && (134)\\
&= -H(Q_U) + H(Q_{U|Y}) + \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_{XY|U}(XY|U)}{P_{XY}(X,Y)}\right)\right] && (135)\\
&= -H(Q_Y) + H(Q_{Y|U}) + \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_{XY|U}(XY|U)}{P_{XY}(X,Y)}\right)\right] && (136)\\
&= -H(Q_Y) + H(Q_{Y|U}) + \mathrm{E}_{Q_{XYU}}\left[\log\left(\frac{Q_{Y|U}(Y\mid U)\,Q_{X|YU}(X\mid YU)}{P_{XY}(X,Y)}\right)\right] && (137)\\
&= -\sum_{u\in\mathcal{U}} Q_U(u)\left(H(Q_Y) - H(Q_{Y|U=u}) - \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} Q_{Y|U=u}\,Q_{X|YU=u}(x|y,u)\log\frac{Q_{Y|U}(y\mid u)\,Q_{X|YU=u}(x\mid y,u)}{P_{XY}(x,y)}\right). && (138)
\end{aligned}$$

The representation (132) shows that the inner maximization in (13) can be performed separately for every u. Defining

$$\Psi^*(Q_{Y|U=u}) = \max_{Q_{X|YU=u}} \Psi(Q_{Y|U=u}, Q_{X|YU=u}) \tag{139}$$

we can express (13) as

$$\sup_{Q_Y}\ \inf_{Q_{U|Y}:\mathrm{I}(Q_{Y;U})\le R} \sum_{u\in\mathcal{U}} Q_U(u)\,\Psi^*(Q_{Y|U=u}). \tag{140}$$

We next view the inner minimization above as being over all pairs $(Q_U, Q_{Y|U})$ with the objective function being

$$\sum_{u\in\mathcal{U}} Q_U(u)\,\Psi^*(Q_{Y|U=u}); \tag{141}$$

with the constraint on the Y-marginal

$$\sum_{u\in\mathcal{U}} Q_U(u)\,Q_{Y|U}(y|u) = Q_Y(y), \qquad \forall y\in\mathcal{Y}; \tag{142}$$

and the constraint on the mutual information

$$\sum_{u\in\mathcal{U}} Q_U(u)\,H(Q_{Y|U=u}) \ge H(Q_Y) - R. \tag{143}$$

Since the objective function and constraints are linear in Q_U, it follows from Carathéodory's theorem (for connected sets) that the cardinality of $\mathcal{U}$ can be restricted to $|\mathcal{Y}|+1$.

References

[1] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inf. Theory*, vol. 8, no. 5, pp. 293–304, Sep. 1962.

[2] H. Witsenhausen, "Indirect rate distortion problems," *IEEE Trans. Inf. Theory*, vol. 26, no. 5, pp. 518–521, Sep. 1980.

[3] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.

[4] E. Arıkan and N. Merhav, "Guessing subject to distortion," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1041–1056, May 1998.

[5] C. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec., Pt. 4*, vol. 44, no. 3, pp. 142–163, 1959.

[6] J. Massey, "Guessing and entropy," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 1994, p. 204.

[7] E. Arıkan, "An inequality on guessing and its application to sequential decoding," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 99–105, Jan. 1996.

[8] C. Cachin, "Entropy measures and unconditional security in cryptography," Ph.D. dissertation, ETH Zurich, 1997.

[9] R. Sundaresan, "Guessing based on length functions," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2007, pp. 716–719.

[10] A. Bracher, A. Lapidoth, and C. Pfister, "Distributed task encoding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1993–1997.

[11] ——, "Guessing with distributed encoders," *Entropy*, vol. 21, no. 3, 2019. [Online]. Available: https://www.mdpi.com/1099-4300/21/3/298

[12] R. Sundaresan, "Guessing under source uncertainty," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 269–287, Jan. 2007.

[13] N. Merhav and A. Cohen, "Universal randomized guessing with application to asynchronous decentralized brute–force attacks," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 114–129, 2020.

[14] S. Salamatian, A. Beirami, A. Cohen, and M. Médard, "Centralized vs decentralized multi-agent guesswork," in *2017 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 2258–2262.

[15] N. Weinberger and O. Shayevitz, "Guessing with a bit of help," *Entropy*, vol. 22, no. 1, 2020. [Online]. Available: https://www.mdpi.com/1099-4300/22/1/39

[16] A. Beirami, R. Calderbank, M. M. Christiansen, K. R. Duffy, and M. Médard, "A characterization of guesswork on swiftly tilting curves," *IEEE Transactions on Information Theory*, vol. 65, no. 5, pp. 2850–2871, May 2019.

[17] R. Graczyk and A. Lapidoth, "Gray-Wyner and Slepian-Wolf guessing," in *2020 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2020, pp. 2189–2193.

[18] ——, "Variations on the guessing problem," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 231–235.

[19] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, 1971.

[20] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[21] S. Moser, "Advanced topics in information theory," https://moser-isi.ethz.ch/docs/atit_script_v46.pdf, 2021, accessed: 2021-04-23.

[22] A. Bracher, E. Hof, and A. Lapidoth, "Guessing attacks on distributed-storage systems," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 6975–6998, 2019.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.